Title

Unveiling the Path to Seasonal Flowering Responses: Exploring the Regulatory Mechanisms of Flowering Synchrony and Long-day Photoperiodic Responses

Author

Huang, Po-Kai

Publication Date

2023

Peer reviewed|Thesis/dissertation

Unveiling the Path to Seasonal Flowering Responses: Exploring the Regulatory Mechanisms of Flowering Synchrony and Long-day Photoperiodic Responses

By

PO-KAI HUANG
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Plant Biology Graduate Group

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Daniel Runcie, Chair

Johanna Schmitt

Stacey Harmer

Committee in Charge

2023

Acknowledgement

First and foremost, I am extremely grateful to my supervisor, Professor Daniel Runcie, whose invaluable guidance, unwavering support, and patience over the course of these seven years have been instrumental to my growth. I have gained immense knowledge from him in areas ranging from experimental design and data analysis to article writing. I am deeply appreciative of the assistance and guidance provided by my academic advisor, Professor Marie Jasieniuk, in preparing for the Qualifying Examination and in career development. I also extend my thanks to my dissertation committee members, Professor Johanna Schmitt and Professor Stacey Harmer, for their substantial support and patience, which was crucial during the final stages of dissertation writing.

My profound gratitude goes to my family, especially my wife and two wonderful children, whose warm support has been a constant source of strength throughout this journey, particularly amidst the challenges posed by the Covid pandemic. I am equally grateful to my family members in Taiwan, my mother and sister, for their encouragement.

I am indebted to the members of the Runcie lab for their insightful suggestions during lab meetings. I also wish to express my sincere appreciation to colleagues who generously shared their lab space and growth chamber facilities: Miki, Mimi, Beth, Lalani, Francine, Monroe lab, and Ross-Ibarra lab. A special thank you to all participants of the Tri-lab meetings; I have greatly benefited from the discussions held during these sessions.

Abstract

Annual plants possess the ability to perceive a broad spectrum of signals from their external and internal environments, enabling them to flexibly regulate flowering timing and optimize this crucial reproductive transition in seasonal conditions. However, the mechanisms underlying most of the seasonal flowering responses observed in the natural environments remain elusive, despite the wealth of molecular knowledge related to flowering-regulating genes and pathways. This dissertation presents two case studies related to seasonal flowering responses and explores the underlying molecular mechanisms. In the first chapter, I investigate vernalization-induced flowering synchrony in *Arabidopsis thaliana*. My findings unveil the contributions of both leaf-based and non-leaf-based regulatory mechanisms to flowering synchrony, emphasizing the critical role of inter-organ crosstalk. The second chapter explores the molecular mechanisms underlying the long-day photoperiodism of annual *Mimulus guttatus*. Using QTL mapping and RNAseq, I reveal distinct molecular mechanisms capable of yielding similar photoperiodic responses, highlighting the flexibility of flowering pathways. Lastly, I examine differential transcriptomic regulations in response to constant and diel temperature fluctuations using *Mimulus guttatus*. This study is a component of a broader project that focuses on how plants interact with the real-world environment, which is crucial for achieving a comprehensive understanding of the seasonal flowering responses observed in natural conditions. Overall, these studies establish a foundational understanding of seasonal flowering responses and highlight the necessity of exploring interplays among pathways and among plant organs to obtain comprehensive insights into these phenomena.

Contents

Overall Introduction

To ensure reproductive success, plants in temperate climates have evolved to align their reproductive transitions to the appropriate seasonal timing, ensuring optimal conditions for pollination and subsequent fruit and seed development. This precision is especially crucial for annual plants, which rely on a single reproductive event for their fitness. To achieve this, plants utilize a diverse array of cues from both their external and internal environments (Srikanth & Schmid, 2011; Kinoshita & Richter, 2020). The contemporary advances in our understanding of the molecular basis of flowering regulation have revealed a plethora of flowering-related genes. Additionally, multiple pathways governing the regulatory processes of flowering have been established. This complexity underpins the considerable flexibility of flowering regulatory mechanisms, allowing plants to respond effectively and promptly to their environment. However, most of this knowledge in flowering regulation is derived from experiments conducted in controlled laboratory settings, where there is relatively stringent control over the genetic background of plant materials and/or environmental conditions. Thus, there remains uncertainty regarding the practical application of these discoveries to expound upon the observed seasonal flowering responses in natural conditions, which are characterized by multi-layered fluctuations and intricate crosstalks among diverse regulatory mechanisms. The subject of this dissertation is to offer insights into the regulatory mechanisms that underlie seasonal flowering responses through two case studies with different annual plant species.

In Chapter 1, the focus lies on the molecular mechanisms driving flowering synchrony, a seasonal phenomenon that has been a subject of long-standing scientific interest (Augspurger, 1983; Primack, 1985; Rathcke & Lacey, 1985; Ims, 1990). The synchronized flowering observed across individuals within a population implies the existence of mechanisms capable of triggering flowering simultaneously despite variations in growth times and developmental stages among these plants (Miryeganeh *et al.*, 2018; Miryeganeh, 2020). However, the underlying molecular

mechanisms are largely unknown regardless of the wealth of our knowledge related to flowering. My study focuses on vernalization-induced flowering synchrony in *Arabidopsis thaliana* (hereafter, Arabidopsis). I demonstrate distinct transcriptional responses among leaves, attributed to developmental stage differences at the time of vernalization treatment. This variation across leaves suggests a potential leaf-based mechanism responsible for synchronizing flowering following vernalization induction. Furthermore, this research indicates the pivotal roles played by signals originating from organs beyond leaves in driving flowering synchrony. However, further studies are required to unravel the intricacies of the interplay between these inter- and intra-organ regulations. This chapter sheds light on the understanding of the regulatory mechanisms behind the synchronous flowering phenomena under seasonal conditions and emphasizes the importance of cross-communication between different organs and pathways.

Chapter 2 delves into the photoperiodic responses of flowering, considering daylength serves as a widely-utilized signal for plants to anticipate seasonal changes in their natural environment (Gendron & Staiger, 2023). Our specific focus centers on exploring the variation in long-day photoperiodic responses across the annual plant *Mimulus guttatus* (hereafter, Mimulus) accessions. These accessions originate from distinct habitats and harbor a wide range of genetic diversity (Sweigart & Willis, 2003; Twyford & Friedman, 2015; Puzey *et al.*, 2017). Using QTL mapping and RNAseq, I demonstrate that different Mimulus accessions can achieve the same photoperiodism phenotypes through different mechanisms. This intriguing finding highlights the plasticity of flowering pathways, which should be always taken into account for studying seasonal flowering responses within populations exhibiting substantial genetic variation. At the end of this chapter, I discuss the improvement of photoperiodism categorization and propose fine-mapping experiments to further unveil the regulatory mechanisms behind this inter-species variation in photoperiodic flowering response.

Lastly, I present findings on the transcriptomic responses of Mimulus to constant and diel fluctuating temperatures. This pilot experiment is part of a broader project aimed at understanding how plants interact with the real-world environment. This project is prompted by increasing evidence indicating that plants can exhibit markedly distinct responses under natural conditions influenced by fluctuating environmental factors (Matsuzaki *et al.*, 2015; Poorter *et al.*, 2016; Matsubara, 2018; Chiang *et al.*, 2020; Hashida *et al.*, 2022). A striking revelation by Burghardt et al. in 2016 (Burghardt *et al.*, 2016) suggests that previous studies may have exaggerated the impact of the floral repressor *FLC* on flowering time due to the constant temperature commonly used in laboratory settings. These discoveries emphasize the necessity of incorporating a wide array of stimuli to investigate the seasonal flowering responses observed in the actual environment. In this chapter, I present a time course experiment using an annual Mimulus accession, SWC. I establish pipelines that provide insights into both global and detailed differential responses to constant and diel fluctuating temperature conditions. Through our results, I unveil transcriptional variations between the two temperature profiles, shedding light on how plants respond to diel fluctuating temperatures. However, I also identify potential limitations in my analysis pipeline and experimental design. To facilitate future research in this field, I discuss the encountered challenges and propose potential improvements.

In summary, this dissertation aims to establish the foundation for the investigation of seasonal flowering responses observed in natural conditions, highlighting the importance of integrating existing discrete knowledge of flowering regulation and uncovering emerging novel patterns through a comprehensive approach. Furthermore, I propose future research directions that focus on investigating the interplay among different pathways and various plant organs.

References

Augspurger CK. **1983**. Phenology, Flowering Synchrony, and Fruit Set of Six Neotropical Shrubs. *Biotropica* **15**: 257–267.

Burghardt LT, Runcie DE, Wilczek AM, Cooper MD, Roe JL, Welch SM, Schmitt J. **2016**. Fluctuating, warm temperatures decrease the effect of a key floral repressor on flowering time in Arabidopsis thaliana. *New Phytologist* **210**: 564–576.

Chiang C, Bånkestad D, Hoch G. **2020**. Reaching Natural Growth: The Significance of Light and Temperature Fluctuations in Plant Performance in Indoor Growth Facilities. *Plants* **9**: 1312.

Gendron JM, Staiger D. **2023**. New Horizons in Plant Photoperiodism. *Annual Review of Plant Biology* **74**: null.

Hashida Y, Tezuka A, Nomura Y, Kamitani M, Kashima M, Kurita Y, Nagano AJ. **2022**. Fillable and unfillable gaps in plant transcriptome under field and controlled environments. *Plant, Cell & Environment* **45**: 2410–2427.

Ims RA. **1990**. The ecology and evolution of reproductive synchrony. *Trends in Ecology & Evolution* **5**: 135–140.

Kinoshita A, Richter R. **2020**. Genetic and molecular basis of floral induction in Arabidopsis thaliana. *Journal of Experimental Botany* **71**: 2490–2504.

Matsubara S. **2018**. Growing plants in fluctuating environments: why bother? *Journal of Experimental Botany* **69**: 4651–4654.

Matsuzaki J, Kawahara Y, Izawa T. **2015**. Punctual Transcriptional Regulation by the Rice Circadian Clock under Fluctuating Field Conditions. *The Plant Cell* **27**: 633–648.

Miryeganeh M. **2020**. Synchronization of senescence and desynchronization of flowering in Arabidopsis thaliana. *AoB PLANTS* **12**.

Miryeganeh M, Yamaguchi M, Kudoh H. **2018**. Synchronisation of Arabidopsis flowering time and whole-plant senescence in seasonal environments. *Scientific Reports* **8**: 10282.

Poorter H, Fiorani F, Pieruschka R, Wojciechowski T, van der Putten WH, Kleyer M, Schurr U, Postma J. **2016**. Pampered inside, pestered outside? Differences and similarities between plants growing in controlled conditions and in the field. *New Phytologist* **212**: 838–855.

Primack RB. **1985**. Patterns of Flowering Phenology in Communities, Populations, Individuals, and Single Flowers. In: White J, ed. Handbook of Vegetation Science. The Population Structure of Vegetation. Dordrecht: Springer Netherlands, 571–593.

Puzey JR, Willis JH, Kelly JK. **2017**. Population structure and local selection yield high genomic variation in Mimulus guttatus. *Molecular Ecology* **26**: 519–535.

Rathcke B, Lacey EP. **1985**. Phenological Patterns of Terrestrial Plants. *Annual Review of Ecology and Systematics* **16**: 179–214.

Srikanth A, Schmid M. **2011**. Regulation of flowering time: all roads lead to Rome. *Cellular and Molecular Life Sciences* **68**: 2013–2037.

Sweigart AL, Willis JH. **2003**. Patterns of Nucleotide Diversity in Two Species of Mimulus Are Affected by Mating System and Asymmetric Introgression. *Evolution* **57**: 2490–2506.

Twyford AD, Friedman J. **2015**. Adaptive divergence in the monkey flower Mimulus guttatus is maintained by a chromosomal inversion. *Evolution* **69**: 1476–1486.

Chapter 1 Exploring the Molecular Regulation of Vernalization-induced Flowering Synchrony in Arabidopsis

1.1 Abstract

Flowering synchrony has been a long-standing observation with potential advantages in plant reproduction, yet its underlying molecular mechanisms remain largely unknown. To shed light on this phenomenon, we conducted sequential planting experiments with varying germination times to study vernalization-induced flowering synchrony in the Arabidopsis Col-FRI accession. A simplified chamber environment was able to capture the synchronous flowering patterns previously observed in a fluctuating field environment. To unravel the molecular regulation of flowering synchrony, we experimentally tested whether flowering synchrony could be related to the differential activity of flower-promoting *FT* signals across leaf ranks. *FT* was efficiently expressed only in leaves developing within or after vernalization, and this uncharacterized expression pattern suggests a leaf-based mechanism involved in flowering synchrony. Further analyses including *FLC* and five other flowering regulatory genes identified from our RNAseq data showed differences in the dynamics of this leaf-based regulation among target genes. However, by manipulating daylength after vernalization, we discovered that signals originating from sources other than leaves played pivotal roles in synchronizing flowering time, especially in germination cohorts with prolonged growth before vernalization. To summarize our findings, we propose three conceptual models of vernalization-induced flowering synchrony and suggestions for future research in this field.

1.2 Introduction

Flowering synchrony can offer several advantages to plants, including enhancing the attractiveness of floral displays to pollinators, satiating predators, and promoting the chance of outcrossing (Rathcke & Lacey, 1985; Ims, 1990). To synchronize flowering across individuals in a population, plants utilize various external signals, including temperature, photoperiod, and precipitation. Vernalization, the response to prolonged exposure to low temperatures during winter, is used by many temperate plant species to promote synchronized flowering in the spring. This mechanism is also used in horticulture and agriculture to synchronize flowering (Suzuki & Metzger, 2001; Tyler *et al.*, 2014; Rubin & Friedman, 2018; Somssich, 2020). Flowering synchrony is particularly noticeable in natural populations with considerable variation in germination time, leading to distinct differences in age, size, and physical characteristics among flowering individuals. The simultaneous flowering of individuals within such populations indicates the presence of specific mechanisms driving synchronization. Nonetheless, despite its wide recognition, the underlying molecular mechanisms driving vernalization-induced flowering synchrony remain poorly elucidated. Due to its ease of handling and well-understood flowering pathways, *Arabidopsis thaliana* (hereafter, Arabidopsis) provides opportunities to deepen our understanding of vernalization-induced flowering synchrony and gain insight into the molecular pathways involved.

Extensive research on the regulation of flowering time in Arabidopsis has revealed multiple interacting pathways modulating responses to various factors, including photoperiod, ambient temperature, vernalization, gibberellin, and aging (Srikanth & Schmid, 2011). These pathways, although partially distinct, converge on a set of integrator genes including *FLOWERING LOCUS T (FT), TWIN SISTER OF FT (TSF),* and *SUPPRESSOR OF OVEREXPRESSION OF CO 1 (SOC1)*, which activate meristem identity genes such as *LEAFY (LFY)* and *APETALLA 1 (AP1)*. The role of FT in signal integration is particularly pivotal, as it

acts as the main component of florigen and can be transported over long distances throughout the plant (Putterill & Varkonyi-Gasic, 2016). FT protein synthesized in the leaves is transported via the phloem vascular system and eventually reaches the shoot apical meristem (SAM) (Lin *et al.*, 2007; Corbesier *et al.*, 2007; Tamaki *et al.*, 2007; Jaeger & Wigge, 2007; Notaguchi *et al.*, 2008). This spatial movement facilitates leaf-to-apex communication, which triggers the onset of flowering and influences the rate of flowering under inductive photoperiods.

Given these features, many mechanistic flowering time models use *FT* as the link between molecular pathways and physiological growth, based on the assumption that total *FT* levels are positively correlated with the size, number, or biomass of leaves (Salazar *et al.*, 2009; Satake, 2010; Jaeger *et al.*, 2013; Kinmonth-Schultz *et al.*, 2019). Such "leaf-counting models" have successfully integrated qualitative gene relationships and quantitative expression dynamics, and have proven to be powerful tools in breaking down intricate mechanisms and uncovering hidden details that may not have been revealed through experiments alone (Salazar *et al.*, 2009; Antoniou-Kourounioti *et al.*, 2021). As vernalization-induced flowering synchronization is a quantitative characteristic associated with the regulation of flowering time, these models are particularly suitable for shedding light on this process.

However, a largely untested assumption of leaf-counting models is that total *FT* production is proportional to leaf number (or other relevant leaf properties), such that larger plants with more leaves will always have higher *FT* levels and consequently flower earlier than smaller plants with fewer leaves. Recent findings by Kinmouth-Shultz et al. (Kinmonth-Schultz *et al.*, 2019) demonstrated that this assumption is an oversimplification, as *FT* expression actually varies with the developmental stage of the leaves. Additionally, based on our understanding of the regulation of *FT* by the vernalization pathway, we hypothesize that *FT* production may differ considerably among leaves in ways contrary to typical modeling assumptions.

In the vernalization pathway, *FT* is regulated by several MADS-box proteins. Prolonged cold treatment leads to epigenetic reprogramming of the chromatin surrounding those key

MADS-box genes, either repressing or activating their expression (Alexandre & Hennig, 2008; Kim & Sung, 2013). The details of this process are best understood for the floral repressor *FLOWERING LOCUS C* (*FLC)* (Dennis & Peacock, 2007; Xu & Chong, 2018; Kim, 2020). In Arabidopsis accessions with a strong vernalization requirement, *FLC* expression is activated by the FRIGIDA (FRI) complex and remains high before vernalization (Johanson *et al.*, 2000; Choi *et al.*, 2011). High levels of *FLC* lead to delayed flowering by repressing *FT*, *SOC1*, and *FLOWERING LOCUS D* (Michaels *et al.*, 2005; Helliwell *et al.*, 2006; Sheldon *et al.*, 2006; Searle *et al.*, 2006). After vernalization, *FLC* is repressed by the Polycomb repressive complex 2 (PRC2) (De Lucia *et al.*, 2008) in a process that involves the cold-induced induction of the plant-homeodomain (PHD) finger protein *VERNALIZATION INSENSITIVE3* (*VIN3*) and histone modifications (primarily trimethylation of histone 3 on lysine 27, H3K27me3) (Sung & Amasino, 2004; Wood *et al.*, 2006). Other MADS-box proteins, such as FLOWERING LOCUS M (FLM)/MADS AFFECTING FLOWERING1 (MAF1) and MAF2-MAF5, also play negative roles in flowering and are downregulated after vernalization (Kim & Sung, 2013), whereas AGAMOUS-LIKE 19 (AGL19) is a floral activator that is induced after exposure to low temperatures (Schönrock *et al.*, 2006; Kang *et al.*, 2015).

In one of the few molecular studies investigating the effects of the leaf developmental stage during vernalization treatment in Arabidopsis, Finnegan and Dennis (Finnegan & Dennis, 2007) observed that *FLC* was incompletely repressed in leaves that had matured prior to vernalization relative to leaves that developed after vernalization. In particular, they observed a depletion of H3K27me3 epigenetic marks, which are responsible for maintaining the repression of *FLC* after vernalization, in leaves that developed before vernalization. These findings suggest that only leaves developing after (or perhaps during) the vernalization treatment “remember” the cold stimulus. We term this phenomenon of vernalization memory differing among leaves with different developmental histories "leaf-specific vernalization response”. If the leaf-specific vernalization responses of *FLC* result in leaf-specific regulation of downstream *FT*, this will have

important implications for how we understand and model post-vernalization flowering dynamics such as flowering synchronization.

In addition to the leaf-based external signal sensing discussed earlier, recent studies have started to unveil the important roles of aging pathways in regulating flowering in plants with prolonged vegetative growth. The aging pathway is comprised of two microRNA-related modules coordinating plant development in a timely manner: the miR156-SPL and miR172-AP2 modules, which act as count-down and count-up timers for flowering, respectively (Huijser & Schmid, 2011; Teotia & Tang, 2015; Hyun *et al.*, 2017). During the juvenile stage, miR156 is highly expressed but miR172 is low. As plants progress into the adult and reproductive stages, miR156 expression decreases, while miR172 expression increases. These expression features of miR156 and miR172 enable plants to gradually gain flowering competence during growth since miR156 represses the expression of the *SPL* family, positive regulators of flowering (Huijser & Schmid, 2011; He *et al.*, 2018), while miR172 represses the *AP2* family, negative regulators of flowering (Chen, 2004). The aging pathway can influence the integrator genes in both the leaf and SAM, so their crosstalk is expected. For example, a recent study demonstrated that *SPL15* in the SAM functions in parallel with *FT* in leaves to induce flowering, but with different effects at different developmental stages, thus underscoring the dynamic roles of different plant organs during growth (Hyun *et al.*, 2019). However, how plants balance the signals from these two locations is still poorly understood.

In this study, we aimed to link the molecular pathways of vernalization and flowering induction to the ecologically important trait of synchronous flowering, with a specific focus on the potential roles of leaf-specific vernalization responses on flowering synchrony. We performed a series of analyses to measure flowering synchrony across Arabidopsis germination cohorts in both field and controlled chamber environments. Subsequently, we tested the leaf-specific vernalization response hypothesis by studying the expression patterns of the *FT*, *FLC*, and other vernalization-pathway genes across leaf ranks in different germination cohorts to test if

leaves of different developmental stages exhibited different vernalization responses. Finally, we evaluated the importance of variation in *FT* production relative to other flowering induction pathways in governing the synchronization of flowering induced by vernalization. Our study provides new insights into the molecular basis of vernalization-induced flowering synchrony.

1.3 Results

1.3.1. Flowering synchrony is observed in Col-FRI under natural conditions

Previous overwinter experiments demonstrated Arabidopsis accessions with vernalization requirements exhibited flowering synchrony in both natural fields and greenhouses, while those without vernalization requirements displayed flowering de-synchrony under the same conditions (Miryeganeh *et al.*, 2018; Miryeganeh, 2020). Here, we investigated whether the vernalization-responsive genotype Col-FRI (Col with functional FRI introgressed from SF-2 (Lee & Amasino, 1995)) also exhibited flowering synchrony after vernalization. Although not included in Miryeganeh et al.'s studies, Col-FRI is widely used in vernalization research due to its Col background, which is a commonly used wild-type accession that can flower promptly without vernalization. We re-analyzed published data on spring flowering times of Col-FRI planted in seven sequential cohorts from September to November 2007 in Cologne, Germany (Wilczek *et al.*, 2009). Additionally, we included Col data from the same experiment as a control. To quantify flowering synchrony, we employed the synchronization index (SI) described by Miryeganeh et al. (Miryeganeh *et al.*, 2018; Miryeganeh, 2020). The SI is defined as the ratio of the variance of germination time to the variance of the flowering time on the log2 scale. A positive SI value indicates flowering synchronization, whereas a negative value indicates flowering de-synchronization.

As shown in Fig. 1.1A, Col exhibited a strong de-synchronization of flowering times in relation to the variation in germination times, with an SI of -3.61. This finding is consistent with the results using Col in the previous article (Miryeganeh, 2020). In contrast, Col-FRI displayed an SI of 1.34, indicating that the variation in flowering times was approximately 40% of the variation in germination time and demonstrated flowering synchrony for this vernalization-responsive genotype (Fig. 1.1B). However, we noticed that the synchronization of flowering was still incomplete, even in Col-FRI. In plants exhibiting complete synchronization,

we would expect flowering to occur at the same time on average, regardless of their germination time. However, we detected differences in the flowering dates between germination cohorts, as measured by photothermal time accumulated from Jan 1st ($F(6,44) = 14.649$ and $p = 4.153e-09$). Specifically, we found that the 1st to 4th cohorts of Col-FRI showed significantly earlier flowering dates compared to the 6th and 7th cohorts. These differences among cohorts implied the presence of factors counteracting the mechanisms promoting synchronization, preventing complete synchronization.

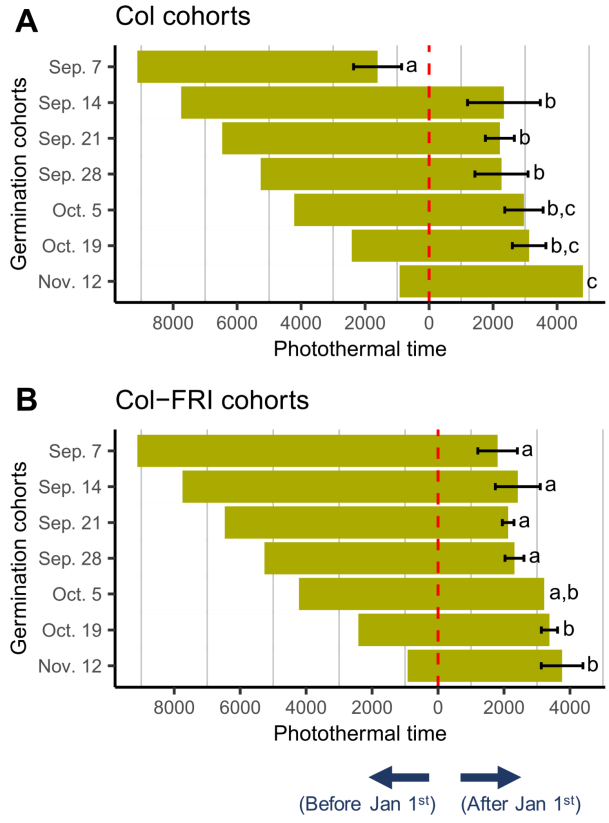

Fig. 1.1: The distribution of flowering time across germination cohorts of Col (A) and Col-FRI (B) in the field data published by Wilczek et al, 2009. The plants were grown sequentially from September to November 2007 in Cologne, Germany. Different cohorts were labeled by their germination date on the y-axis. The flowering and germination times are presented on the photothermal time scale, with photothermal time calculated from Jan 1st (red dashed line). Error bars represent standard deviations. Samples without error bars do not have biological replicates in the dataset. Means with different letters are significantly different based on the Tukey test at the 5% significance level.

1.3.2. Controlled chamber experiments capture the trends of flowering synchrony observed in Col-FRI under natural conditions

To investigate the specific impacts of vernalization and photoperiod on vernalization-induced flowering synchrony, separate from the complex array of environmental factors encountered in field conditions, we conducted a series of sequential planting experiments in growth chambers using Col-FRI (see Fig. 1.2A). We grew plants in short days (SD, 8/16 h day/night cycle) for 36, 22, 7, or 1 day before an 8-week vernalization treatment (G36, G22, G7, G1, respectively, with “G” representing the “Germination cohort” and the number representing the number of days). After vernalization, plants were moved to the long day (LD, 16/8 h day/night cycle) condition until flowering. Non-vernalization plants were included as a control group (NV).

The results of the day-to-flowering after vernalization are depicted in Fig. 1.2B, with an SI of 5.99, indicating substantial synchronization, and the flowering time variation was approximately 2% of the germination time variation. Consistent with the major trend of flowering times observed in the natural field for Col-FRI, pairwise comparisons among germination cohorts revealed significant differences in flowering times, demonstrating that synchronization was incomplete. Overall, the results from our controlled chamber environment mirrored the major patterns observed in the natural fields, suggesting that we can investigate this phenomenon in a simplified experimental setting. Furthermore, the absence of gradual changes in photoperiods in chamber experiments ruled out the possibility that the synchronicity observed in the field experiment was dependent on increasing spring daylength.

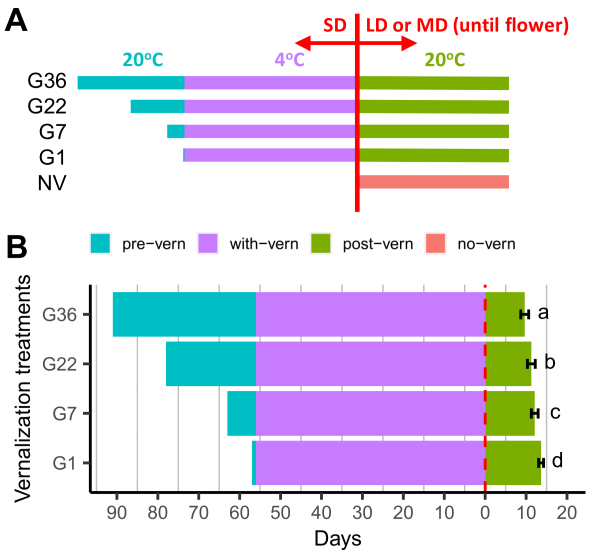

Fig. 1.2: Col-FRI demonstrated flowering synchrony in chamber experiments. (A) Schematic figure of the sequential planting experiments in growth chambers using Col-FRI. (B) The distribution of flowering time across vernalization treatments (G1, G7, G22, and G36). The flowering and germination times are calculated from the end of the vernalization (red dashed line). Error bars represent standard deviations. Means with different letters are significantly different based on the Tukey test at the 5% significance level. Cyan, purple, green, and salmon represent the growth before vernalization (pre-vern), within vernalization (with-vern), after vernalization (post-vern), and without vernalization (no-vern), respectively.

1.3.3. *FT* expression patterns indicate leaf-specific vernalization responses across leaf ranks

Several quantitative models of flowering time regulation were constructed based on the assumption that the number of leaves (or total leaf area) producing *FT* is a key determinant of flowering time (Salazar *et al.*, 2009; Satake, 2010; Jaeger *et al.*, 2013; Kinmonth-Schultz *et al.*, 2019). The premise underlying this assumption is that all leaves (and leaf areas) produce *FT* at an equal rate. However, the observations of Finnegan and Dennis (Finnegan & Dennis, 2007) that *FLC* was not fully repressed in older leaves suggested that *FT* might not be highly expressed in leaves that developed before vernalization, and therefore not contribute equally to the total *FT* pool in a plant. To test whether leaves varied in *FT* expression after vernalization, we investigated *FT* expression across leaf ranks in Col-FRI plants exposed to the G7 and G36

treatments, as described in the preceding section. We collected every fifth leaf, starting from the youngest fully extended rosette leaf and progressing basipetally until leaves began to senesce, and measured the *FT* expression levels in each leaf using qPCR.

The G36 treatment resulted in plants with approximately 25, 50, and 70 macroscopically visible rosette leaves before vernalization, after vernalization, and at flowering, respectively. The expression of *FT* across leaves followed a sigmoidal pattern, with low expression levels in leaves with ranks less than 45 and high expressions in later-produced leaves (Fig. 1.3). In contrast, plants in the G7 treatment had only cotyledons (zero true leaves), 10, and 20 macroscopically visible rosette leaves before vernalization, after vernalization, and at flowering, respectively. In these plants, *FT* expression was low in the fifth leaf but increased sharply by the 20th leaf (see Fig. 1.3). The low expression of *FT* in lower leaf ranks of the G7 treatment was expected because those were juvenile leaves known to have lower *FT* expression (Debernardi *et al.*, 2022). Interestingly, significant differences in *FT* levels were observed even among the leaves developed after vernalization. For instance, there were differences in the maximum *FT* levels between the G7 and G36 treatments, and the 20th true leaves of the G7 treatment exhibited higher *FT* levels compared to the 70th true leaves of the G36 treatment. Moreover, even within the G7 treatment, the expression of *FT* was higher in the 20th true leaves compared to the 15th true leaves. Nevertheless, our results show a remarkable pattern that only leaves developing within or after vernalization can effectively express *FT*, suggesting *FT* also exhibited leaf-specific vernalization responses. This pattern suggests a possible role for leaf-specific vernalization responses in flowering synchrony, with only ~20 leaves in both treatments

contributing the majority of the *FT* pool, irrespective of the total number of leaves.

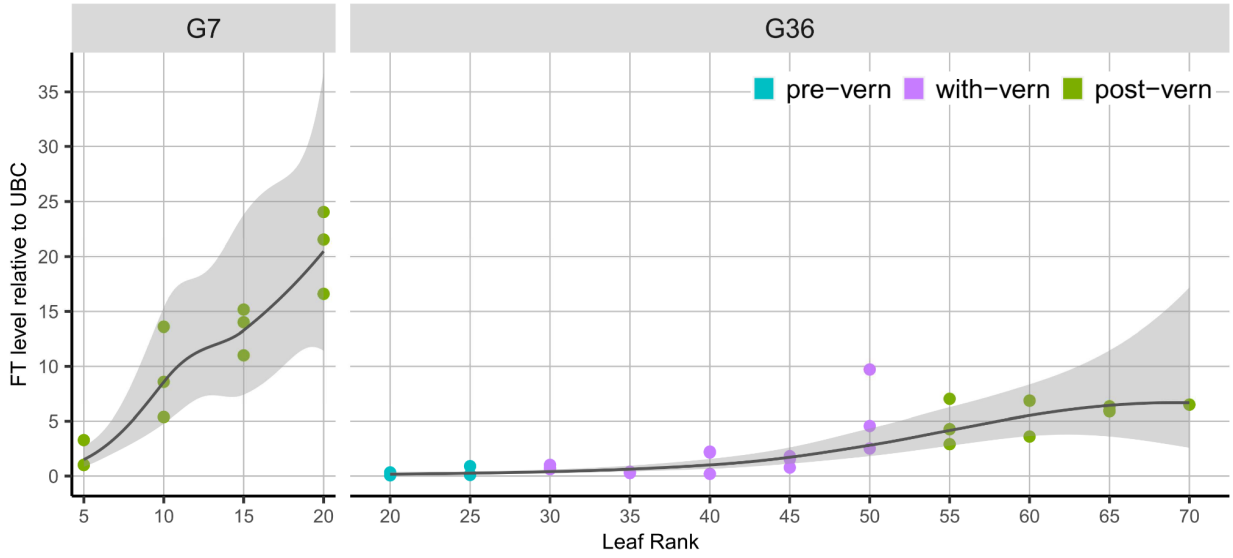

Fig. 1.3: *FT* expressions across leaf ranks indicate leaf-specific vernalization responses. The left panel represents the expression curve of the G7 treatment, and the right panel represents the expression curve of the G36 treatment. Each dot represents a biological replicate of leaves collected from different plants. Dots are colored based on the best inference of when each leaf developed, with respect to the vernalization treatment: before (pre-vern), during (with-vern), or after (post-vern) the plants were vernalized, respectively. The grey area around the curve indicates the 95% confidence interval.

1.3.4. The differential expression patterns between *FT* and *FLC* indicate a complex regulatory mechanism of leaf-specific vernalization responses

Although the above experiments provided evidence for systematic differences in *FT* expression as a function of the developmental environment of each individual leaf, we could not directly conclude these differences were caused by vernalization because vernalization timing was confounded with daylength across treatments: leaves that developed after vernalization also developed under longer daylength than those developed before vernalization. For example, in the G7 treatment, the 15th leaf developed after vernalization in LD, while the corresponding leaf in the G36 treatment developed before vernalization in SD. Also, as mentioned above, the *FT* expression levels varied across leaf ranks in the G7 treatment due to differences in leaf

development stages (i.e., juvenile vs. adult leaves) despite all developing after vernalization. To address these potential confounding effects, we ran a follow-up experiment to isolate the effects of vernalization from daylength and expression variations across leaf ranks. We kept the same germination-time-variation design and grew plants in warm SDs for 36, 22, 8, or 1 day before the vernalization treatment (G36, G22, G8, G1, respectively). However, instead of transferring directly to LD after vernalization, we moved the plants back to warm SD conditions immediately after vernalization and only exposed them to inductive LD on the day of sample harvest. Previous studies have shown that a single LD after SD growing is sufficient to induce maximum transcriptional responses of *FT* (Yanovsky & Kay, 2002; Krzymuski *et al.*, 2015). Thus, all plant leaves developed exclusively under SD conditions, with the only difference among treatments being the timing of the vernalization treatment during development. We collected samples at three chronological ages, measured as days at 20°C after planting: 22 d, 36 d, and 43 d or 50 d (WSD22, WSD36, WSD43/50, respectively, with "WSD" representing the "Warm Short Day" and the number representing the number of days). At each time point, we harvested three sets of leaves: early-emerging leaves (ranks 3-7), mid-development leaves (ranks 14-17), and late-emerging leaves (ranks 30+), if those leaves were present on the plants. We divided similarly staged leaves into three groups: leaves that developed before (pre-vern), during (with-vern), or after (post-vern) vernalization. We also included non-vernalized plants as a negative control, which were sampled one week earlier at the last harvest because they were significantly larger in size and were beginning to flower (see Fig. 1.4 and Table 1.1 for the summary of the experimental design). We quantified *FT* and *FLC* expression by qPCR.

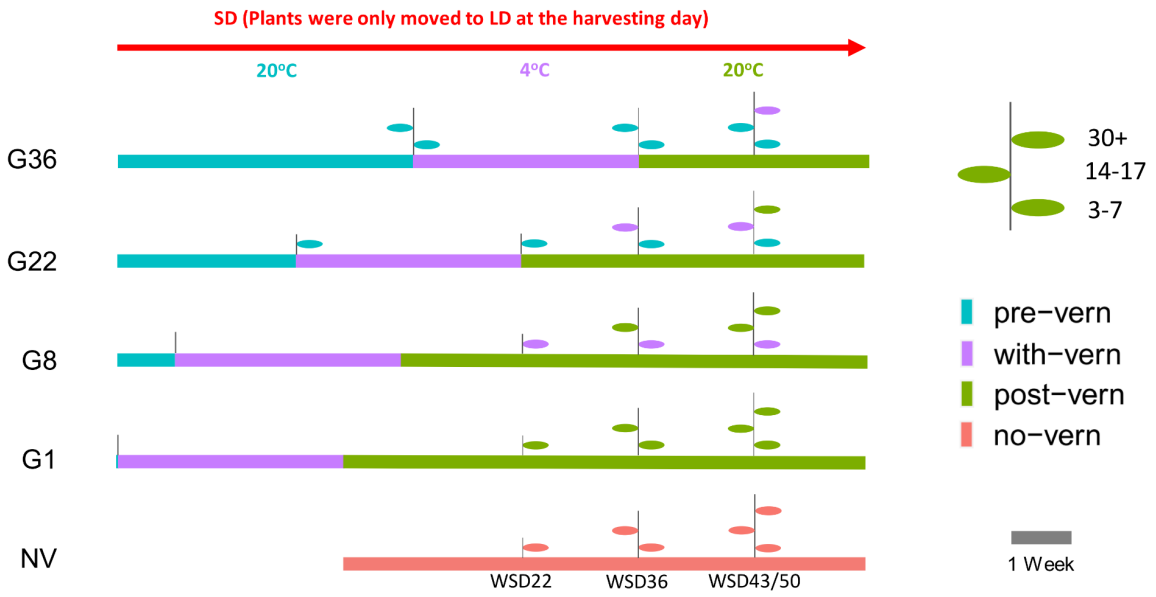

Fig. 1.4: Schematic figure of the sequential planting experiments in growth chambers using Col-FRI. We controlled the confounding daylength effects by growing plants in SD before, within, and after vernalization. The plants were exposed to LD only on the day of sampling. Cyan, purple, green, and salmon represent the growth before vernalization (pre-vern), within vernalization (with-vern), after vernalization (post-vern), and without vernalization (no-vern), respectively. Leaves are colored based on when they are presumed to have developed in the shoot apical meristem.

Table 1.1: Summary of sequential planting experiments depicted in Fig. 1.4. Rows represent vernalization treatments, and columns represent sampling time points. The leaf developmental stage at the time when a plant experienced vernalization is indicated in parentheses and color-coded corresponding to Fig. 1.4.

	WSD22	WSD36	WSD43/50
NV	3-7(NV)	14-17(NV), 3-7(NV)	30+(NV), 14-17(NV), 3-7(NV)
G1	3-7(post-vern)	14-17(post-vern), 3-7(post-vern)	30+(post-vern), 14-17(post-vern), 3-7(post-vern)
G8	3-7(with-vern)	14-17(post-vern), 3-7(with-vern)	30+(post-vern), 14-17(post-vern), 3-7(with-vern)
G22	3-7(pre-vern)	14-17(with-vern), 3-7(pre-vern)	30+(post-vern), 14-17(with-vern), 3-7(pre-vern)
G36	-	14-17(pre-vern), 3-7(pre-vern)	30+(with-vern), 14-17(pre-vern), 3-7(pre-vern)

Fig. 1.5 shows the expression levels of *FLC* (Fig. 1.5A) and *FT* (Fig 1.5B) across different vernalization treatments (G1, G8, G22, G36, and NV) for three sets of leaf ranks (3-7, 14-17, and 30+) at three different sampling time points (WSD22, WSD36, and WSD43/50). To account for any ontogenetic changes in gene expression or vernalization sensitivity related to leaf rank, we compared *FLC* or *FT* expression only among leaves of the same leaf rank collected at the same sampling time point. Within such comparably-staged leaves, expression differences can clearly be attributed to the leaf developmental stage at the time when a plant experienced vernalization (pre-vern, with-vern, post-vern, or no-vern). ANOVA for testing differences in *FLC* or *FT* expression between vernalization treatments revealed significant

differences ($p < 0.05$) within each set of comparable leaves (i.e. each panel in Fig. 1.5) (Table S1.1 for the ANOVA results), indicating the impact of vernalization on the expression of these two genes.

We observed consistent patterns across leaf ranks and sample harvesting time points for each gene. *FLC* expression was highest in leaves on un-vernalized plants (no-vern leaves), followed by the leaves developing before, within, and after vernalization, consistent with the observations of Finnegan and Dennis. On the other hand, *FT* expression was similar between the leaves developing either after or during vernalization (post-vern and with-vern leaves), but was much higher in both these sets of leaves than in the leaves developing before vernalization or in comparable leaves from non-vernalized plants (pre-vern and non-vern leaves). Surprisingly, although both genes exhibited leaf-specific vernalization responses, *FT* expression patterns were not directly opposite to *FLC* expression patterns: *FLC* expression markedly decreased even in leaves developing before vernalization relative to leaves from non-vernalization plants (pre-vern vs. non-vern leaves), albeit this decrease was less pronounced compared to the decrease observed in leaves that developed after or during vernalization (post-vern/with-vern vs. no-vern leaves). In contrast, *FT* expression remained barely detectable in leaves that developed before vernalization, similar to leaves from plants that were not vernalized (pre-vern vs. non-vern leaves), but expression levels were much higher in all leaves that developed during or after vernalization (with-vern and post-vern leaves). Therefore, the leaf developmental stage-dependent vernalization responses observed by Finnegan and Dennis for *FLC*'s expression appeared to be even more pronounced in *FT*.

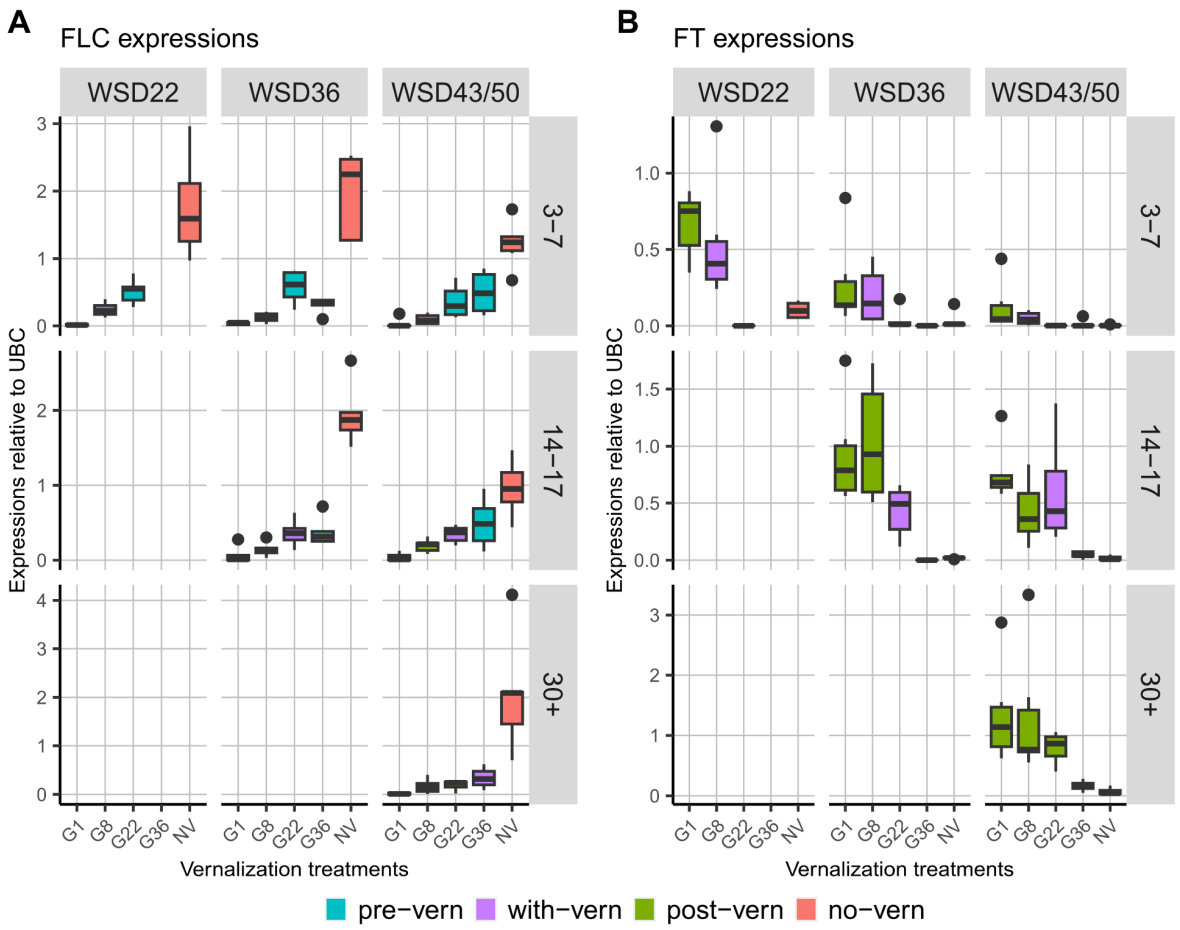

Fig. 1.5: Expressions of *FLC* (A) and *FT* (B) show leaf-specific vernalization responses. Each plot grid's columns represent sampling time points, while the rows represent leaf ranks. Within each panel, gene expression values are grouped based on the vernalization treatments (G1, G8, G22, G36, and NV) and color-coded according to the leaf's vernalization status (pre-vern, with-vern, post-vern, and no-vern), as referenced in Fig. 1.4 and Table 1.1.

To quantify how a leaf's vernalization memory is affected by its developmental stage when the plant was vernalized, we compared vernalization responses in leaves that developed before and after the plant underwent vernalization. In each case, we defined the “vernalization effect” as the difference in expression between leaves of plants that had been vernalized (pre-vern or post-vern leaves) to corresponding leaves of plants that had not been vernalized (no-vern leaves). We measured the ratio of these two effects to quantify the effect of the leaf developmental stage on vernalization memory:

The effect of the leaf developmental stage on vernalization memory

$$= |\text{Pre-vern} - \text{No-vern}| \,/\, |\text{Post-vern} - \text{No-vern}| \qquad \text{(Equation 1)}$$

A smaller value of this ratio indicates a reduced vernalization response in leaves developing before vernalization relative to leaves developing after vernalization. For example, a ratio close to zero indicates that expression levels in the leaves developing before vernalization are similar to those in the leaves on non-vernalized plants. This also suggests a larger effect of the leaf developmental stage on vernalization memory because the vernalization effects are limited to the leaves developing after vernalization. As shown in Fig. 1.6, the vernalization effect ratio was much smaller for *FT* than *FLC*, indicating a substantial difference in leaf-specific vernalization responses between these two genes.

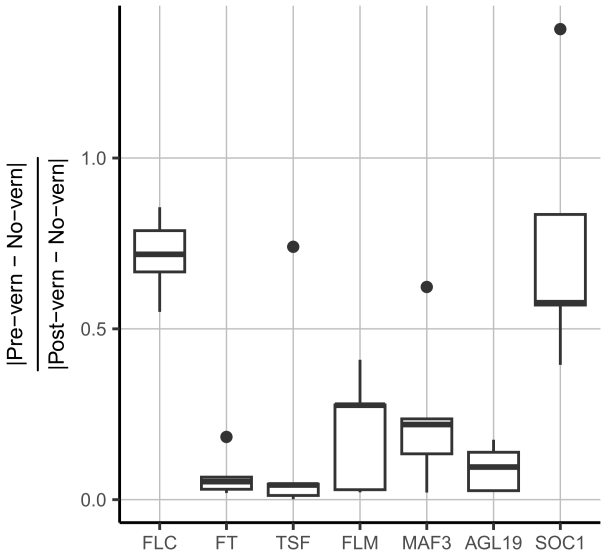

Fig1.6: The ratio of vernalization effects of leaves developed before vernalization (pre-vern) to the vernalization effects of leaves developed after vernalization (post-vern). Each vernalization effect is measured relative to corresponding leaves from plants without vernalization treatment (no-vern). A smaller ratio indicates a relatively smaller vernalization response in pre-vern leaves, and therefore a more dramatic effect of the leaf developmental stage on vernalization memory.

1.3.5. Differential gene expression associated with vernalization memory in leaves is largely restricted to seven key members of the vernalization pathway

Since the leaf developmental-stage effect on vernalization memory was stronger for the downstream integrator *FT* than the upstream regulator *FLC*, we asked whether other *FT* regulators are involved in shaping the pattern of *FT*'s leaf-specific vernalization responses. To achieve this, we used RNA-seq to identify genes exhibiting vernalization memory in Col-FRI. Specifically, we compared the gene expression levels between non-vernalized leaves and vernalized leaves that underwent six weeks of vernalization before returning to standard growth conditions for 14 and 18 days (n=3 per time point). To ensure accurate comparisons, we grew and sampled non-vernalized plants (n=3 per time point) simultaneously with vernalized plants of the same size and stage that had completed the cold treatment. We evaluated 19,319 genes for differential expression between vernalized and non-vernalized leaves and identified 19 genes with a significant response to vernalization at a false discovery rate (FDR) threshold of 0.05, and seven at an FDR threshold of 0.001 (Table S1.3). This latter list of genes was composed of well-known components of the vernalization response pathway: *FLM*, *SOC1*, *AGL19*, *TSF*, *MAF3*, *FT*, and *FLC* (ordered by increasing p-value). Consequently, we chose to focus our study on these seven genes to further characterize the transcription dynamics of the leaf-specific vernalization responses.

1.3.6. Vernalization-pathway genes exhibit diverse patterns of leaf-specific vernalization responses

We conducted detailed analyses of these seven genes using the same RNA samples and analysis procedures described in Section 1.3.4. For brevity, we only present the expression patterns of these genes in the middle leaf (ranks 14-17) at the final time point (WSD43/50) in

Fig. 1.7. Results for the other two time points and two sets of leaves are available in Supplementary Fig. S1.1-S1.5.

We observed strong ($p < 0.05$) evidence of developmental-stage-dependent vernalization responses in all genes in most leaf sets (Table S1.1). However, the vernalization responses varied among genes, resulting in different expression patterns. *FLM* and *MAF3*, which are two MADS-box repressors of flowering time, exhibited similar expression patterns. Their expression levels were high in leaves without vernalization, as well as in leaves that developed before and during vernalization (no-vern, pre-vern, and with-vern leaves). In contrast, their expression was relatively lower in leaves developing after vernalization (post-vern leaves). Thus, vernalization memory was restricted to leaves that fully developed after vernalization for these two genes. On the other hand, the MADS-box gene *AGL19*, which accelerates flowering time, displayed different leaf-specific vernalization memory patterns. *AGL19* showed higher expression levels in leaves developing after and within vernalization (post-vern and with-vern leaves), compared to the leaves developing before vernalization or leaves on non-vernalized plants (pre-vern and no-vern leaves). Among the downstream integrator genes, *TSF* exhibited expression patterns similar to *FT*, albeit with generally lower expression levels, consistent with previous observations (Michaels *et al.*, 2005; Jang *et al.*, 2009). In contrast, *SOC1* expression demonstrated an approximately inverse pattern relative to *FLC*, with the lowest levels observed in leaves without vernalization, followed by the leaves developed before, within, and highest after vernalization.

As in Section 1.3.4, we quantified the effect of the leaf developmental stage on vernalization memory by comparing the vernalization effects of leaves developing before vernalization to the vernalization effects of leaves developing after vernalization (Equation 1) across all the combinations of leaf ranks and sampling time points (Fig. 1.6). The ratio of these two vernalization effects was generally low for the genes of interest, indicating that the

expressions of these genes in the leaves developing before vernalization were similar to those observed in the leaves without vernalization. However, SOC1 stood out as an exception, showing a higher ratio similar to FLC. Taken together, our findings illustrate the diverse patterns of leaf-specific vernalization responses for the seven target genes.

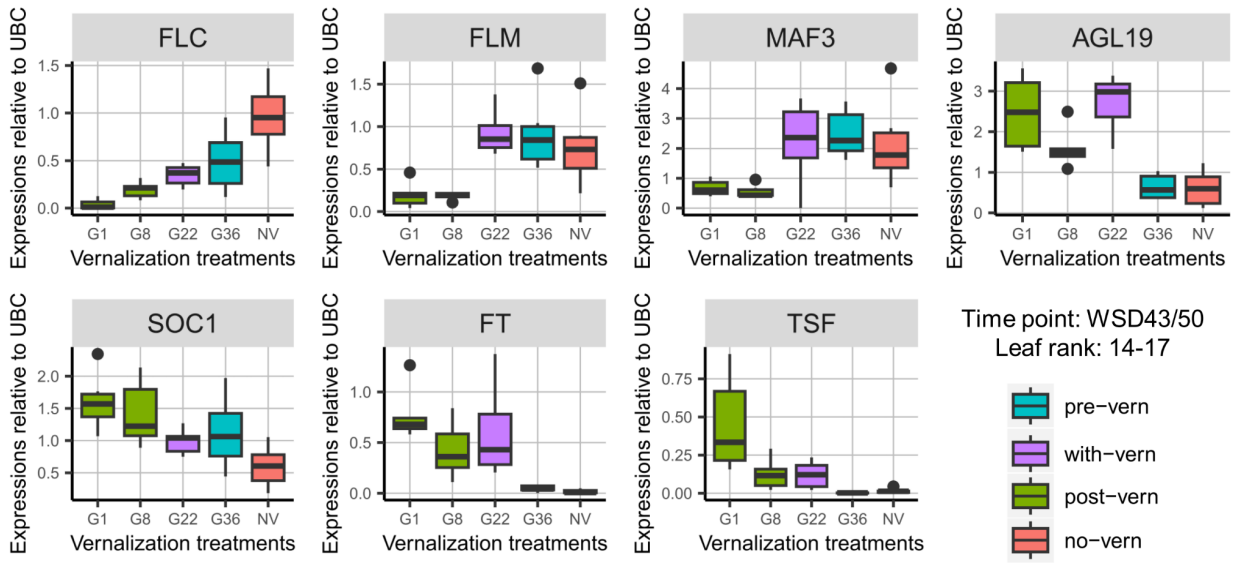

Fig. 1.7: The expression dynamics of seven candidate genes in the middle leaf (ranks 14-17) at the final time point (WSD43/50). Each panel represents a single gene. Within each panel, gene expressions are grouped based on the vernalization treatments (G1, G8, G22, G36, and NV) and color-coded according to the leaf's vernalization status (no-vern, post-vern, pre-vern, and with-vern), as referenced in Fig. 1.4 and Table 1.1.

1.3.7. *AGL19* alone cannot explain the expression pattern differences between *FLC* and *FT*

Among candidate regulators of *FT*, *AGL19*'s expression paralleled *FT*'s leaf-specific vernalization response most closely. Specifically, for *FT*, *TSF*, and *AGL19*, expression levels in the leaves developing within and after vernalization were similar, and dramatically different from those in the leaves developing before vernalization or without vernalization (i.e., with-vern similar to post-vern leaves). In contrast, for *FLM* and *MAF3*, the leaves developing within vernalization had expressions similar to those developing before and without vernalization (i.e.,

with-vern leaves similar to pre-vern and no-vern leaves). *FLC* expression patterns were different as well, as discussed above. *AGL19* is a member of the MADS-box protein family and is an activator of *FT*. Therefore, we hypothesized that the leaf-specific responses of *FT* may be dependent on *AGL19* in addition to its well-known dependence on *FLC*.

To test this hypothesis, we generated an *agl19* mutant in the Col-FRI background. We crossed Col-FRI with the T-DNA insertion mutant *agl19-1* in the Col background to create the *agl19-1/Col-FRI* genotype. This new genotype has high *FLC* and very low (but non-zero) *AGL19* levels (Fig. 1.8C-D). We predicted that *agl19-1*/Col-FRI plants would exhibit extremely late flowering compared to Col-FRI, particularly in the absence of vernalization. We also predicted that this genotype would flower later than Col-FRI after vernalization. However, we did not observe the expected difference in flowering time between these two genotypes (Fig. 1.8A-B) or any significant alteration in *FT* levels between the *agl19-1* mutants and the corresponding wild types (Fig. 1.8E-F). One possible explanation for the lack of the expected phenotype could be the low but still existing expression of the *AGL19* gene. Nevertheless, considering that the *agl19-1* mutants exhibited only approximately 1/15 of the *AGL19* expression levels compared to their corresponding wild types, which was much lower than the fluctuations observed across leaf ranks in the mentioned experiment (all less than a four-fold difference between vernalized and non-vernalized leaves, for example, expression of *AGL19* in Fig. 1.7), our results suggest that *AGL19* does not play a critical role in *FT* expression or the regulation of flowering time in our conditions. Taken together, *AGL19* alone is unlikely to reconcile the discrepancy between the *FT* and *FLC* patterns observed previously.

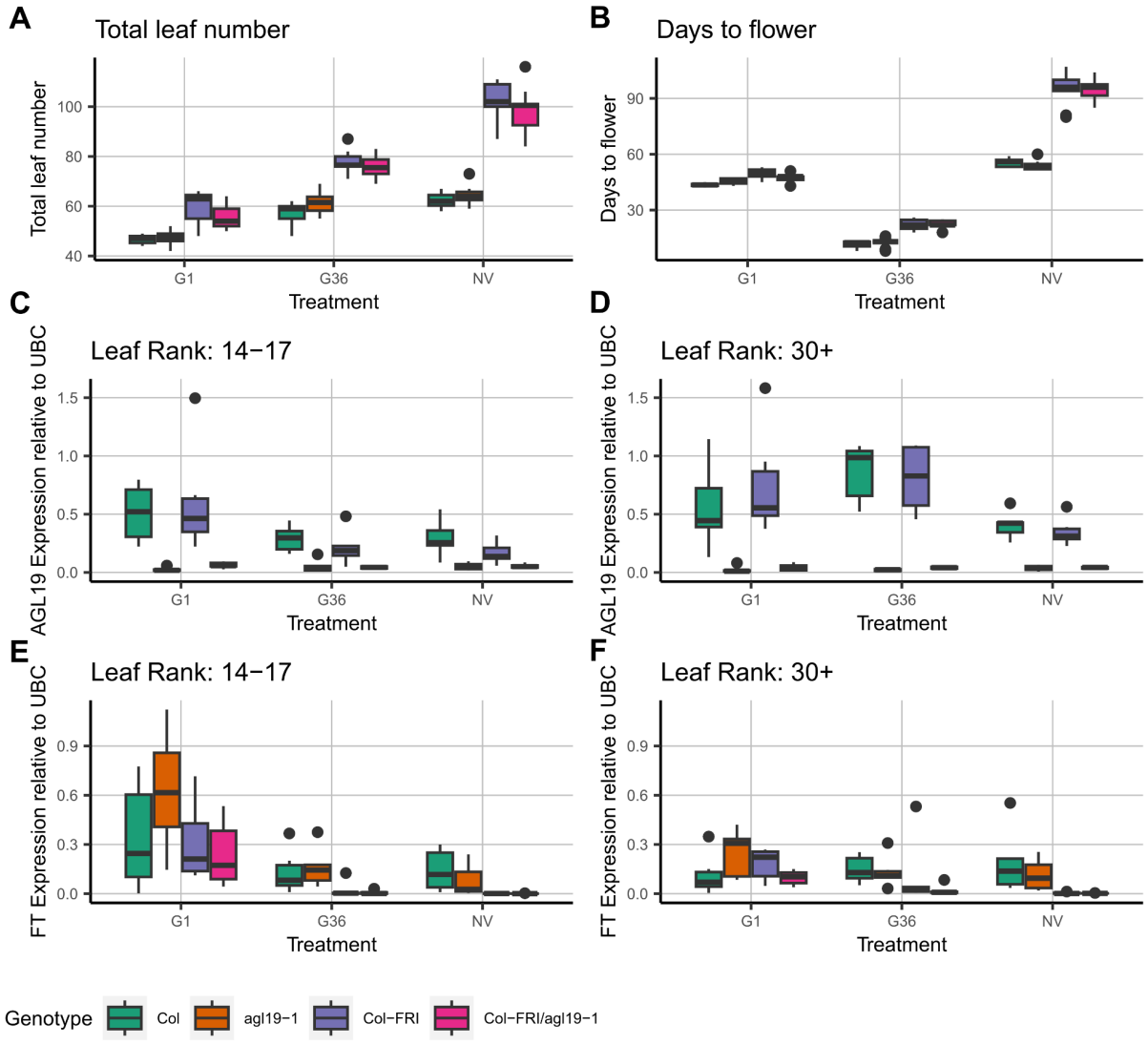

Fig. 1.8: The *agl19* mutants did not show obvious differences in flowering time (A-B) and *FT* expressions (E-F) compared to the corresponding wild types. We tested the *agl19* mutants with three vernalization treatments, G1, G36, and NV. Total leaf number (A) and days to flower (B) were recorded to estimate flowering time. Expressions of *AGL19* (C-D) and *FT* (E-F) were analyzed by qPCR in two leaf ranks (14-17 and 30+). The order of genotypes for each treatment in each panel is as follows: Col (green), *agl19-1* (orange), Col-FRI (dark purple), and Col-FRI/*agl19-1* (pink).

1.3.8. Other tissues or mechanisms may be involved in the vernalization-induced flowering synchrony

While our focus was on the expression dynamics in leaves, a recent article revealed that parallel regulatory networks expressed primarily in the shoot apical meristem are important in flowering regulation, particularly in plants with prolonged vegetative growth (Hyun *et al.*, 2019). To assess the relative importance of leaf-based vernalization responses, we reduced the total *FT* expression of leaves by shortening the post-vernalization photoperiod. Specifically, we repeated the flowering time analysis described in Section 1.3.3 but moved the plants to a 12/12 h medium photoperiod (MD) after vernalization instead of the usual LD condition (Fig. 1.2A). Previous studies have shown that *FT* has similar transcript patterns in MD and LD, with both exhibiting a peak at dusk. However, in MD, *FT* exhibits a much smaller peak intensity compared to LD conditions (Song *et al.*, 2012; Krzymuski *et al.*, 2015). Thus, we expected to observe a general delay in flowering across all treatments under MD conditions compared to those under LD if the total *FT* expression produced by all leaves was a major determinant of flowering time regulation.

We first examined whether *FT* responses to the MD conditions varied across leaf ranks of the G36 treatment. We found a sigmoidal shape similar to that in LD, but the extent of the induction under MD was much lower, with only 1/5 of the expression compared to LD (Fig. 1.9A). If the total *FT* production integrated across leaves was the primary determinant of flowering time, around five times as many *FT*-producing leaves would need to develop to induce flowering compared to the LD condition, requiring considerably more developmental time. However, we observed only a slight delay in flowering time in each germination cohort, from 1.3 days (~10% delay in time post-vernalization) for the earliest germinating G36 cohort, to 4.3 days (~33% delay in time) for the latest germinating G1 cohort. Differences in total leaf number

showed similar trends. These results suggest that the leaf-based pathways are likely not the dominant determinant of flowering synchronization under these conditions.

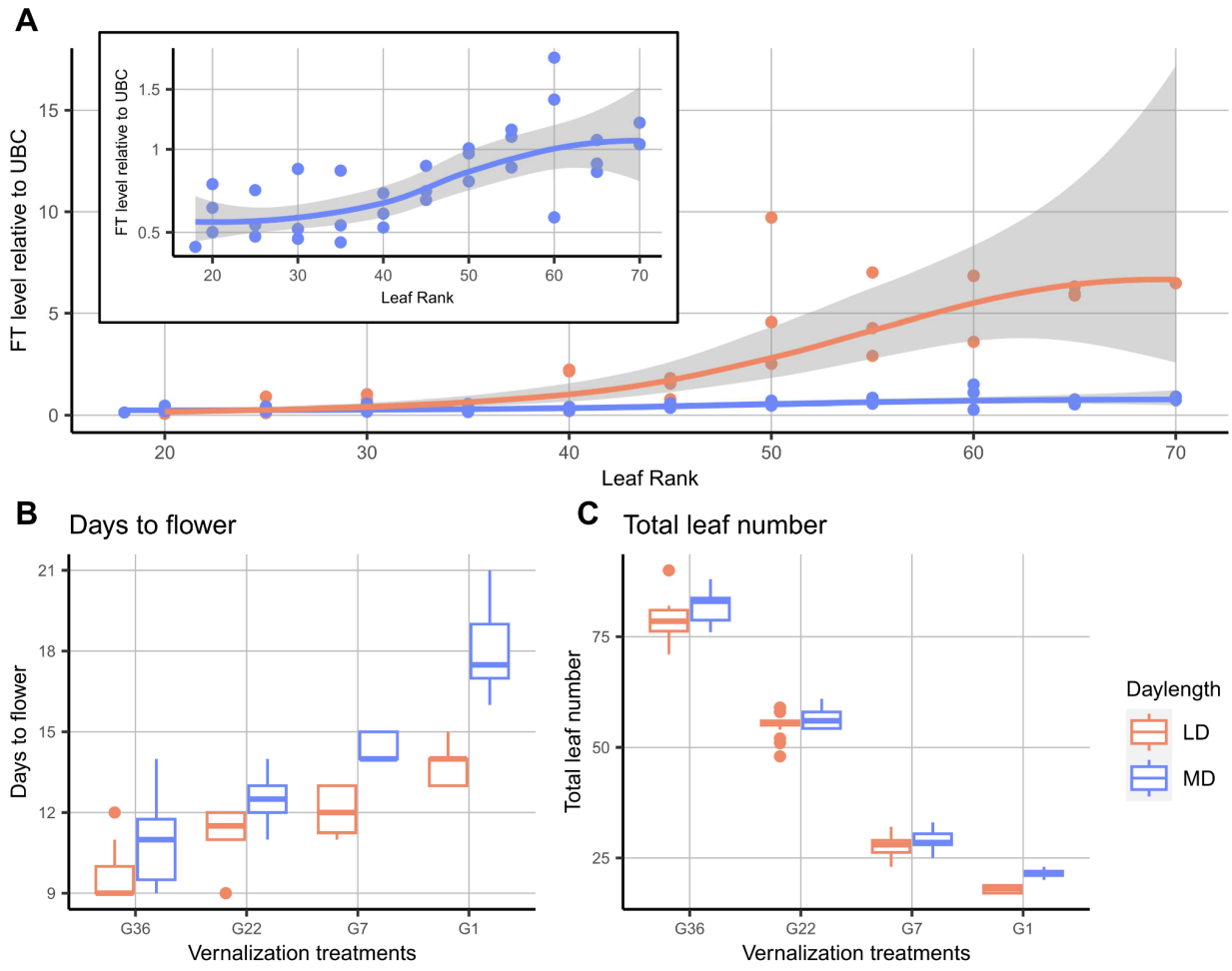

Fig. 1.9: *FT* levels responded to daylength differences (A) but the flowering time (B and C) did not show as large a change relative to *FT* expression. (A) *FT* expression across leaf ranks in LD (light orange curve) and MD (violet curve). The zoom-in of the *FT* curve in MD is shown in the subplot in the upper left. The grey area around the curve indicates the 95% confidence interval. (B) and (C) show the comparisons of days to flower and total leaf number between plants moved to LD (light orange) or MD (violet) after vernalization treatments.

1.4 Discussion

Flowering synchrony is a widely recognized phenomenon in natural conditions (Primack, 1985; Rathcke & Lacey, 1985; Ims, 1990). It refers to a reduced variation in the day of flowering among individuals within a population in comparison to the levels of variation observed in other developmental stage transitions, such as germination time. However, few studies have investigated the underlying molecular mechanisms. Studies on flowering at the molecular level typically compare flowering times among different genotypes or treatments without incorporating perturbations in germination time, and thus without differences in developmental stages at which environmental cues are received. For example, most studies of the vernalization pathway compared genotypes under conditions that vary only in the duration or intensity of vernalization. However, to study how vernalization contributes to flowering synchrony, it is crucial to vary germination time, so that plants would be asynchronous if not for active synchronizing mechanisms (Miryeganeh *et al.*, 2018; Miryeganeh, 2020). Consequently, despite the wealth of molecular knowledge on flowering, the challenge of bridging the gap between our understanding and the observed phenomenon of flowering synchrony remains unresolved.

To address this question, we conducted a series of experiments focusing on flowering synchrony across Arabidopsis germination cohorts in both field and controlled chamber environments. We employed Miryeganeh et al.'s synchrony index (SI), which takes into account both the variation in germination time and flowering time, providing a quantitative measure of synchrony. Traditionally, flowering synchrony is considered a trait primarily associated with flowering regulation. Previous studies commonly measured flowering synchrony in two ways: (1) quantifying the variability in flowering time within a population using statistical measures such as variance, standard deviation, or coefficient of variation, or (2) assessing the overlap in flowering time among different populations or individuals under investigation (Augspurger, 1983). However, we believe that Miryeganeh et al.'s SI provides a more robust approach to

understanding the underlying mechanisms of flowering synchrony. This choice was motivated by the recognition of the cascading effect of germination timing on flowering (Donohue *et al.*, 2010, 2015; Chiang *et al.*, 2013; Postma & Ågren, 2016) and the significant influence of internal flowering signals, such as the aging pathway, which are highly related to the timing of germination.

To the best of our knowledge, our article represents the first attempt to establish a connection between the understanding of molecular flowering pathways and the phenomenon of flowering synchrony. We provide insights into molecular mechanisms underlying the synchronization of flowering after vernalization in *Arabidopsis thaliana*. While the benefits of synchrony are most widely appreciated for outcrossing species and Arabidopsis is typically considered a selfing species, Arabidopsis does have an estimated outcrossing rate of 1-3% and as high as 20% in some populations (Abbott & Gomes, 1989; Shimizu & Purugganan, 2005; Platt *et al.*, 2010; Bomblies *et al.*, 2010). In addition, flowering synchrony could still provide fitness benefits through predator satiation in Arabidopsis, although further studies are required to test this hypothesis. Finally, since most flowering pathways are conserved across the Brassicaceae family, especially the vernalization pathway (Leijten *et al.*, 2018), the insights gained from this study can be extended to other outcrossing species, such as *Brassica napus*, *Brassica rapa*, and *Brassica oleracea*. Therefore, our study lays the groundwork for future research on flowering synchrony.

1.4.1. Transcriptional responses to vernalization are leaf-specific and indicate complexity in vernalization-dependent flowering regulation

We conducted an extended analysis of the leaf-specific vernalization responses first observed for *FLC* by Finnegan and Dennis (Finnegan & Dennis, 2007), testing this observation across six additional genes (*FT*, *FLM*, *MAF3*, *AGL19*, *TSF*, and *SOC1*) and sampling leaves at

different time points and leaf ranks. All seven genes showed some degree of leaf-specific vernalization responses, but the dynamics differed among genes. To facilitate further discussion of the potential factors shaping the patterns of leaf-specific vernalization responses of these genes, we roughly divided them into two groups.

First, extensive studies have been conducted on the molecular memory derived from epigenetic regulation of the MADS-box genes: *FLC*, *FLM*, *MAD3*, and *AGL19*. The VIN3 protein family plays a coordinated role in repressing *FLC*, *FLM*, and *MAF3* through the deposition of repressive epigenetic marks on these genes (Sung *et al.*, 2006; Sheldon *et al.*, 2009; Kim & Sung, 2013, 2017). VIN3 is also involved in the activation of *AGL19* following vernalization (Schönrock *et al.*, 2006). Surprisingly, we observed a striking contrast in vernalization memory between *FLC* and the other MADS-box genes (Fig. 1.7, pre-vern vs no-vern leaves). While *FLC* responded to vernalization in all leaves (although with a lower dynamic range in leaves that developed before vernalization), the other MADS-box genes only responded to vernalization in newly developed leaves (i.e., post- and with-vern leaves). The cause of the different vernalization memory patterns among these genes is unclear. One potential explanation is variation in target specificities of the chromatin remodeling factors, such as members of the VIN3 protein family, along with the differences in their timing of function during the vernalization process. The distinct patterns of PRC2-mediated H3K27me3 accumulation across the loci of these MADS-box genes during vernalization may also contribute to the observed differences (Kim & Sung, 2013). Additionally, aside from memory formation, another possibility is related to the maintenance of memory and/or de-vernalization. Finnegan and Dennis (Finnegan & Dennis, 2007) demonstrated that the leaf-specific memory pattern of *FLC* was associated with a failure to transition into the repression maintenance phase due to the lack of DNA replication in those leaves. However, the underlying mechanism of memory maintenance and/or de-vernalization remains largely unknown for other MADS-domain genes.

The formation of leaf-specific memory dynamics for the integrator genes, *FT*, *TSF*, and *SOC1*, can be influenced by two distinct processes: (1) leaf-specific memory of upstream genes (such as *FLC*, *FLM*, etc), resulting in variations in the activities of these integrator genes and (2) the leaf-specific epigenetic regulation of the integrator genes themselves. While these integrator genes are regulated by MADS-domain genes (Gu *et al.*, 2013; Kang *et al.*, 2015), the interaction effects of these MADS-box genes on their downstream targets remain unknown. As a case study, we investigated the involvement of *AGL19* in shaping the leaf-specific memory patterns of *FT*, as the leaf-specific expression patterns of *FT* and *AGL19* were the most similar. However, the *agl19* mutation in a genetic background with high *FLC* expression did not decrease *FT* levels or delay flowering time. Since expression patterns of *FLC* and *FT* were not highly correlated, these results together suggest that *FLC* and *AGL19* cannot fully explain the expression dynamics of *FT*. In addition to integrating the signals from the upstream MADS-domain genes, *FT* and *SOC1* are directly regulated through chromatin-mediated repression and activation (López-González *et al.*, 2014; Bratzel & Turck, 2015; Liu *et al.*, 2023). Therefore, similar molecular memory mechanisms for *FLC* or other MADS-box genes may also play roles in leaf-specific memory for these integrator genes.

1.4.2. Conceptual models for the regulation of vernalization-based flowering synchrony

We constructed three conceptual models (Fig. 1.10) to explore the potential role of leaf-specific vernalization responses in synchronizing flowering in Arabidopsis. First, Model 1 is inspired by the idea that flowering is triggered by the total *FT* produced across all leaves of a plant, so a specific number of leaves is required to trigger flowering as a function of photoperiod (since *FT* production per leaf increases in longer photoperiods). The principles underlying Model 1 have also been incorporated into previous quantitative models (Salazar *et al.*, 2009; Satake, 2010; Jaeger *et al.*, 2013; Kinmonth-Schultz *et al.*, 2019). As depicted in Fig. 1.10, Model 1, applied to cohorts of plants that germinated at different times pre-winter, this conceptual model

would predict largely desynchronized flowering in the spring. Early-germinating cohorts reaching the leaf number threshold would flower immediately after the conclusion of winter, but later germinating cohorts would require the development of additional leaves to meet the total *FT* threshold. Under this model, we would only see synchronization among the earliest germinating cohorts: those that produced a large number of leaves before winter but could not flower at that point because of a lack of vernalization.

However, it is evident from our presented experiments that the assumption made by Model 1, which suggests an equal *FT* production rate in all leaves, does not hold true. We observed that approximately equal numbers of leaves produced *FT* in early and late germinating plants at the time of flowering (around 20 leaves in our case, Fig. 1.3) and that only leaves developing during or after vernalization produced high levels of *FT*. As a result, regardless of the germination timing, if a fixed total *FT* production was required for flowering, flowering would be triggered at approximately the same time across germination cohorts since earlier developing leaves contribute little to the total *FT* pool. Under this model, only very late-developing plants - those that do not start developing leaves until after winter - would exhibit asynchrony in flowering (Fig. 1.10, Model 2).

While promising, our experimental results also revealed several factors that are not consistent with Model 2. First, we observed variation in *FT* levels among leaves developed after vernalization (Fig. 1.3, Section 3 in Results), which contradicts the assumption made in Model 2 of uniform *FT* production rate across all these leaves. As mentioned earlier, some of this variation can be attributed to the developmental stage of leaves. For example, the 5th true leaf in the G7 treatment, which is considered a juvenile leaf (He *et al.*, 2018), exhibited low *FT* expression due to the high miR156 and low miR172 levels overriding external signals. However, the reasons behind the remaining variation in *FT* levels remain unclear, indicating the presence of additional factors regulating leaf-specific vernalization responses that were not included in the

model. Second, Model 2 predicts largely synchronized flowering across germination cohorts, except perhaps for very late germinating cohorts that do not produce leaves during vernalization. Additionally, the synchronized flowering is expected to be theoretically complete, meaning that all the plants exhibiting flowering synchrony should flower at the same or almost the same time point, except for the variation in germination timing. However, both the natural field observations and controlled chamber experiments consistently revealed incomplete flowering synchrony, as demonstrated by the significant differences among cohorts of Arabidopsis plants (Fig. 1.1 and 1.2). This phenotypic deviation from the prediction could be attributed to the missing factor related to leaf-specific vernalization responses mentioned above, or it could indicate the involvement of other sources apart from the leaves.

To examine the relative importance of the leaf-based regulations in vernalization-induced flowering synchrony, we attenuated the influence of leaves on flowering by reducing total *FT* expression through a shortened post-vernalization photoperiod from LD to MD. As shown in Fig. 1.9, we observed only a 1-day difference in flowering time between the MD and LD conditions for the G36 and G22 cohorts, suggesting that in early-germinating cohorts, *FT* expression, and thus leaf-derived signals may no longer play a decisive role in determining flowering time. This observation is likely due to the increased dominance of the aging pathway operating in the SAM. Recent research has demonstrated that *SPL15* expressed in the SAM can induce flowering independently of *FT* produced in the leaves, and its effects are particularly important in non-inductive photoperiods and older plants (Hyun *et al.*, 2019). Specifically, Hyuan et al found that the *ft/tsf* double mutant flowered significantly later than the wild type after growing in SD for two weeks followed by vernalization, but the flowering times of the *ft/tsf* double mutant and wild type were approximately the same after growing in SD for six weeks followed by vernalization. These findings are consistent with our flowering time patterns for the late and early germination cohorts, respectively.

Taken together, these findings suggest that the importance of the leaf and SAM in flowering may shift as plants grow older. In younger plants, leaves likely play a more dominant role in determining flowering time, while in older plants, the SAM may be more decisive. This hypothesis is reminiscent of the need for researchers to test the phenotypes of Arabidopsis in short days to delay plant growth when investigating the GA or aging pathways. In long days, plants flower too rapidly due to the dominance of *FT* signals, making it difficult to discern the effects of other pathways. The shift of dominance between leaf and SAM (or external to internal cues) may be due to increased activity of SAM-based pathways, or reduced sensitivity to *FT* in older plants. However, the precise mechanism by which plants integrate these two sources of signals to regulate flowering is not yet fully understood. Furthermore, it remains unclear at what point during plant development the shift in the relative importance of the leaf and SAM occurs. Thus, further research is needed to answer these questions, and our MD and LD comparison may provide a promising strategy to reveal this shift in flowering determinations.

Based on these arguments, we propose Model 3 as a model for vernalization-induced flowering synchrony (Fig. 1.10, Model 3). Building upon Model 2, Model 3 retains the concept of leaf-specific vernalization responses as the leaf-based signal, but its effect is limited to late-germinating individuals. In addition, Model 3 incorporates the contribution of flowering cues from the aging pathway, represented by the blue line in Fig. 1.10, Model 3. According to this model, early-germinating individuals with strong aging pathway activation surpassing its threshold can initiate flowering immediately after vernalization, irrespective of the leaf count. This is exemplified by the left two individual in Fig.1.10, Model 3, which possesses only two leaves developed within or after vernalization, fewer than the four-leaf threshold. In contrast, later germinating (younger) plants have reduced aging pathway activity and so require a high level of *FT* to trigger flowering. Such high levels of *FT* require the development of additional leaves after the end of vernalization to meet the flowering threshold. In such cases, plants can

initiate flowering either by accumulating four leaves (the leftmost plant in Fig. 1.10, Model 3) or by accumulating a combination of leaves and aging signals from the SAM to reach the flowering threshold (the second plant from the left in Fig 1.10, Model 3). While flowering dynamics predicted by Models 2 and 3 are similar under long photoperiods (solid lines), the effect of reducing photoperiods to intermediate lengths is predicted to differ and Model 3 seems more consistent with our observations (dash lines).

By considering both leaf-based vernalization signals and the SAM-based aging pathway, Model 3 provides a more comprehensive hypothesis for understanding the mechanisms underlying vernalization-induced flowering synchrony. We believe that this framework lays a cornerstone for future research and model development in this field. However, it is important to acknowledge that our findings also reveal gaps in our understanding of these mechanisms, emphasizing the necessity of a comprehensive approach that considers multiple regulatory factors to accurately capture the intricacies of this phenomenon.

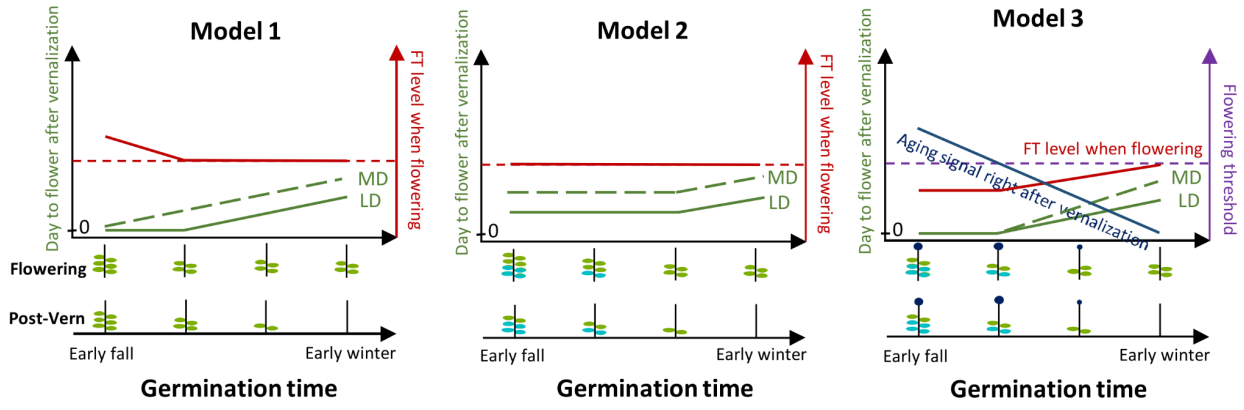

Fig. 1.10: Three conceptual models proposed to explain flowering synchrony in Arabidopsis. Each model illustrates four different plant sizes before exposure to vernalization, representing germination timing from early fall (early-germination individuals) to early winter (late-germination individuals). The anticipated sizes at the end of winter ("Post-Vern") and at the time of flowering ("Flowering") are depicted pictorially according to the assumption of each model, and the total *FT* expression integrated across leaves at the time of flowering is represented by the red solid curves based on the number of leaves capable of

producing *FT* (represented by green leaves). Cyan leaves indicate those that cannot efficiently contribute to *FT* production. The predicted days to flower after vernalization are depicted by green curves, with the solid curve representing the LD condition and the dashed curves representing the MD condition. The minimum threshold of *FT*-induced flowering requirement is four leaves in this case.

Model 1 is based on the "leaf-counting model" and assumes that all leaves can produce *FT* at the same rate. Under this model, plants must have at least four leaves to initiate flowering. Plants with fewer leaves at the end of winter (“Post-Vern”) require additional time to grow in the spring, while those with at least four leaves at the end of winter flower immediately at this time. This results in an apparent broken-stick pattern with synchronous flowering for very early cohorts, but germination-dependent flowering time for the remaining cohorts. Note that total *FT* levels at flowering for early germinating cohorts are above the threshold - they have more leaves (six) than needed but failed to flower earlier because they reached four leaves before vernalization was complete.

Model 2 incorporates the leaf-specific vernalization responses observed in this study, where only leaves developing within or after vernalization can effectively contribute to the *FT* pool. We depict a scenario where two leaves develop during the winter (with-vern) in all but the latest germinating cohort. In this scenario, the three plants on the left need to grow two more leaves in the spring to meet the flowering requirements (four *FT*-producing (green) leaves) and thus flower synchronously. In contrast, late-germination individuals require more time to grow in the spring to reach the four-leaf requirement and exhibit delayed flowering time compared to the majority. Since all the plants flower with four leaves, the total *FT* levels are expected to be the same at flowering.

Model 3 builds upon Model 2 but introduces an additional flowering signal source from the SAM represented by blue circles on the top of plants. The SAM's flowering signal increases as plants grow, as shown by the blue line in Model 3. A plant will flower if the sum of the *FT* signal (red curve) and the SAM signal (blue line) exceeds the flowering threshold (purple dashed line). In this scenario, despite having only two *FT*-producing leaves, the left two plants are ready to flower at the end of winter, as their combined signal with the SAM flowering signal is sufficiently high to reach the flowering threshold. For the plants germinating later in the fall (right two individuals), the plant with extremely late germination (rightmost) reaches the flowering threshold mainly by accumulating enough leaves and *FT*, as the SAM flowering signal is weak in this case. In contrast, the second individual from the right combines signals from both leaves and SAM, requiring only three leaves to meet the threshold.

The differences between Models 2 and 3 are more prominent when considering the flowering time in MD. In Model 2, the flowering time is always delayed in MD compared to LD due to relatively low *FT* production across all the leaves in MD. This delay in flowering is expected to be constant across all germination cohorts, as the leaf-based *FT* signal is the only flowering cue. In contrast, for the two individuals on the left side of the figure in Model 3, the flowering time is expected to be the same in both

MD and LD because the major determinant for flowering is the SAM-based aging signal, and the leaf-based signal becomes less important. Thus, in Model 3, the flowering delay in MD will only be observed for the plants germinating later, as the leaf-based signal still plays a role in flowering induction.

1.5 Conclusion

This study suggests valuable insights into the molecular mechanisms underlying vernalization-induced flowering synchrony using the Arabidopsis Col-FRI accession. Our finding of leaf-specific vernalization responses among seven target genes suggests a potential leaf-based mechanism contributing to flowering synchrony. Additionally, we found signals originating from tissues other than leaves, likely the aging pathway signals in the SAM, as pivotal regulators of flowering synchrony, particularly in early germination cohorts with extended vegetative growth. These findings emphasize the importance of considering the crosstalks within and among organs when studying intricate plant responses to their environment. Overall, our research contributes to a better understanding of the complex interplay between vernalization and flowering synchrony, paving the way for further investigations in this field.

1.6 Materials and methods

Plant materials and growth conditions

The Col-FRI genotype was derived by introgressing the FRI locus from the San Felui 2 (SF-2) genotype into the Col background, rendering it vernalization-dependent for promoting flowering (Lee & Amasino, 1995). Seeds of Col-FRI were cold-stratified in either water or a 0.15% agar solution for 3 to 4 days in the dark at 4°C. After stratification, seeds were planted in 5 cm pots filled with Sunshine Mix #1 potting soil. Pots were randomly assigned to different vernalization treatments (G1, G7 or G8, G22, G36, and NV) in Conviron E7/2 growth chambers. Trays of pots (four per chamber cabinet) were rotated three times per week, and plants in warm temperature conditions were watered twice a week with fertilizer water supplied by the UC Davis Controlled Environment Facility. During the warm temperature growth phase, conditions were set at 20°C, with either 8 hours (short day, SD), 12 hours (medium day, MD), or 16 hours (long day, LD) of light at 250 umol m^{-2} s^{-1} PAR provided by fluorescent tubes without additional far-red supply. For the vernalization treatments, plants were subjected to 4°C with an 8-hour light period at 50 umol m^{-2} s^{-1} PAR, provided by fluorescent tubes without extra far-red supplementation. To maintain humidity during vernalization, trays were covered with clear plastic lids, eliminating the need for additional watering.

For in-chamber phenotyping and FT expression testing, plants were grown under SD warm conditions before vernalization and randomly assigned to receive either G1, G7, G22, or G36 treatment (Fig. 1.2A). After 8-week vernalization, we moved the plants to either warm LD or MD chambers until flowering and recorded the flowering date and the number of rosette and cauline leaves. At ZT12 for MD and ZT16 for LD, when FT expression levels were expected to be at their maximum, we harvested every 5th fully expanded leaf basipetally until leaves were senescing (around the 5th and 20th true leaf for the 7d and 35d treatment, respectively).

To analyze the expression levels of the 7 candidate genes identified from RNAseq, we grew plants in SD warm chambers and randomly assigned them to receive one of four

treatments: G1, G8, G22, or G36 (Fig. 1.4). After four weeks of vernalization, the plants were moved back to SD warm conditions until the date of tissue collection. At this point, the plants were transferred to LD warm chambers and harvested a single leaf at each defined leaf rank at ZT16. We selected three ranges of leaves on each plant by leaf rank (3-7, 14-17, 30+) such that within each range, the leaves were visible (and thus largely post-mitotic) before vernalization (pre-vern) in a set of the treatments, the leaves emerged during vernalization (with-vern) in other treatments, and the leaves emerged after vernalization (post-vern) in the remaining vernalization treatments (see Fig. 1.4 and Table 1.1). We also selected corresponding leaves by leaf rank in non-vernalized plants (non-vern). We harvested plants in each treatment at three chronological ages: 22d, 36d, and 43 or 50d under the warm SD condition (WSD22, WSD36, WSD43/50, respectively). Non-vernalized plants were sampled one week early at the last harvest because they were beginning to bolt. On the WSD22 and WSD36 harvests, only the first or first two leaf sets were visible and could be sampled.

For the RNAseq experiment, we subjected six plants to six weeks of vernalization, starting 8 days after sowing. Six additional plants were grown without vernalization as a control. The non-vernalized plants were planted 17 days before the vernalized plant were removed from the cold treatment to match their size and developmental stage. Following vernalization, plants were grown for an additional 14-18 days in the warm SD before sampling a single leaf (rank ~10). This treatment ensured the observed transcriptional differences were related to the vernalization memory. Three plants from each treatment were sampled on each day.

For AGL19 experiments, the T-DNA insertion lines *agl19-1* (SALK_N578786, in the Col background) were obtained from the SALK Institute (http://signal.salk.edu/). We generated the *agl19-1*/Col-FRI mutant in the Col-FRI background by crossing the homozygous *agl19-1* mutants with Col-FRI and self-pollination of the F1 generation. The homozygous F2s for FRI and *agl19* alleles were determined by PCR amplification using the primers (Table S1.2). For

phenotyping and the qPCR, we subjected the plants to either G1 or G36 treatments with eight weeks of vernalization and returned them to the warm LD condition until flowering.

All the leaf samples were cut at the base of the blade, immediately submerged in liquid N2, and stored at -80°C before conducting RNA extraction for gene expression analysis using qPCR or RNAseq.

qPCR analysis

Total RNA was extracted with the ZR Plant RNA Miniprep Kit (Zymo Research, Irvine, CA, USA) with DNase digestion. RNA concentrations were standardized to ~40 ng/ul, and then 10ul was used for cDNA synthesis with the High-Capacity cDNA Reverse Transcription Kit (ThermoFisher Scientific, Grand Island, NY, USA) with random primers. cDNA was diluted 5-20x and analyzed by qPCR on a BioRad CFX96 instrument (BioRad, Hercules CA USA) using the SsoAdvanced Universal SYBR Green Supermix (BioRad). Each reaction was run at a volume of 20ul. Assay primers are listed in Table S1.2.

RNAseq analysis

Illumina HiSeq compatible RNAseq libraries were prepared from cell lysates using a custom protocol based on Kumar et al 2012 with some modifications (Kumar *et al.*, 2012). The six libraries analyzed in this study were barcoded, pooled with 18 other libraries, and sequenced in a single lane on the HiSeq 2000 at the UC Davis Genome Center DNA Technologies & Expression Analysis Core, resulting in 9 million 50bp PE reads. Reads were aligned to the *Arabidopsis thaliana* genome (TAIR10) (Berardini *et al.*, 2015) using TopHat v2.1.0 with the `--mate-inner-dist` parameter set to 20 (Kim *et al.*, 2013). Reads uniquely mapped to any exon of any isoform of each gene were counted with the featureCounts (Liao *et al.*, 2014), and the differential expression was assessed with limma-voom (Law *et al.*, 2014).

Statistical analysis

All statistical analyses for phenotyping and gene expressions were conducted using the R language (R version 4.2.2). For ANOVA, we first fitted a linear model using the `lm` function in the `stats` package and then processed the `lm` results using the `anova` function in the same package. For pairwise comparisons, we used the `emmean` and `contrast` functions in the `emmeans` package to process the `lm` results. Data visualization was performed using the `ggplot2` package.

For determining flowering synchrony, we adopted the synchronization index (SI) described in Miryeganeh et al.'s article (Miryeganeh *et al.*, 2018; Miryeganeh, 2020). In brief, the SI was developed to compare the variation in germination time to the variation in flowering time, calculated as

$$SI = \log2 [(\text{variance of germination timing})/(\text{variance of flowering timing})]$$

A positive SI indicates synchronization, while a negative SI indicates desynchronization.

1.7 References

Abbott RJ, Gomes MF. **1989**. Population genetic structure and outcrossing rate of Arabidopsis thaliana (L.) Heynh. *Heredity* **62**: 411–418.
Alexandre CM, Hennig L. **2008**. FLC or not FLC: the other side of vernalization. *Journal of Experimental Botany* **59**: 1127–1135.
Antoniou-Kourounioti RL, Zhao Y, Dean C, Howard M. **2021**. Feeling Every Bit of Winter – Distributed Temperature Sensitivity in Vernalization. *Frontiers in Plant Science* **12**.
Augspurger CK. **1983**. Phenology, Flowering Synchrony, and Fruit Set of Six Neotropical Shrubs. *Biotropica* **15**: 257–267.
Berardini TZ, Reiser L, Li D, Mezheritsky Y, Muller R, Strait E, Huala E. **2015**. The arabidopsis information resource: Making and mining the “gold standard” annotated reference plant genome. *genesis* **53**: 474–485.
Bomblies K, Yant L, Laitinen RA, Kim S-T, Hollister JD, Warthmann N, Fitz J, Weigel D. **2010**. Local-Scale Patterns of Genetic Variability, Outcrossing, and Spatial Structure in Natural Stands of Arabidopsis thaliana. *PLOS Genetics* **6**: e1000890.
Bratzel F, Turck F. **2015**. Molecular memories in the regulation of seasonal flowering: from competence to cessation. *Genome Biology* **16**: 192.
Chen X. **2004**. A MicroRNA as a Translational Repressor of APETALA2 in Arabidopsis Flower Development. *Science* **303**: 2022–2025.
Chiang GCK, Barua D, Dittmar E, Kramer EM, Casas RR de, Donohue K. **2013**. Pleiotropy in the Wild: The Dormancy Gene Dog1 Exerts Cascading Control on Life Cycles. *Evolution* **67**: 883–893.
Choi K, Kim J, Hwang H-J, Kim S, Park C, Kim SY, Lee I. **2011**. The FRIGIDA Complex Activates Transcription of FLC, a Strong Flowering Repressor in Arabidopsis, by Recruiting Chromatin Modification Factors. *The Plant Cell* **23**: 289–303.
Corbesier L, Vincent C, Jang S, Fornara F, Fan Q, Searle I, Giakountis A, Farrona S, Gissot L, Turnbull C, *et al.*2007. FT Protein Movement Contributes to Long-Distance Signaling in Floral Induction of Arabidopsis. *Science* **316**: 1030–1033.
De Lucia F, Crevillen P, Jones AME, Greb T, Dean C. **2008**. A PHD-Polycomb Repressive Complex 2 triggers the epigenetic silencing of FLC during vernalization. *Proceedings of the National Academy of Sciences* **105**: 16831–16836.
Debernardi JM, Woods DP, Li K, Li C, Dubcovsky J. **2022**. MiR172-APETALA2-like genes integrate vernalization and plant age to control flowering time in wheat. *PLOS Genetics* **18**: e1010157.
Dennis ES, Peacock WJ. **2007**. Epigenetic regulation of flowering. *Current Opinion in Plant Biology* **10**: 520–527.
Donohue K, Burghardt LT, Runcie D, Bradford KJ, Schmitt J. **2015**. Applying developmental threshold models to evolutionary ecology. *Trends in Ecology & Evolution* **30**: 66–77.
Donohue K, Rubio de Casas R, Burghardt L, Kovach K, Willis CG. **2010**. Germination, Postgermination Adaptation, and Species Ecological Ranges. *Annual Review of Ecology, Evolution, and Systematics* **41**: 293–319.
Finnegan EJ, Dennis ES. **2007**. Vernalization-Induced Trimethylation of Histone H3 Lysine 27 at FLC Is Not Maintained in Mitotically Quiescent Cells. *Current Biology* **17**: 1978–1983.
Gu X, Le C, Wang Y, Li Z, Jiang D, Wang Y, He Y. **2013**. *Arabidopsis* FLC clade members form flowering-repressor complexes coordinating responses to endogenous and environmental cues. *Nature Communications* **4**: 1947.
He J, Xu M, Willmann MR, McCormick K, Hu T, Yang L, Starker CG, Voytas DF, Meyers BC, Poethig RS. **2018**. Threshold-dependent repression of SPL gene expression by miR156/miR157 controls vegetative phase change in Arabidopsis thaliana. *PLOS Genetics* **14**: e1007337.

Helliwell CA, Wood CC, Robertson M, James Peacock W, Dennis ES. **2006**. The Arabidopsis FLC protein interacts directly in vivo with SOC1 and FT chromatin and is part of a high-molecular-weight protein complex. *The Plant Journal* **46**: 183–192.
Huijser P, Schmid M. **2011**. The control of developmental phase transitions in plants. *Development* **138**: 4117–4129.
Hyun Y, Richter R, Coupland G. **2017**. Competence to Flower: Age-Controlled Sensitivity to Environmental Cues. *Plant Physiology* **173**: 36–46.
Hyun Y, Vincent C, Tilmes V, Bergonzi S, Kiefer C, Richter R, Martinez-Gallegos R, Severing E, Coupland G. **2019**. A regulatory circuit conferring varied flowering response to cold in annual and perennial plants. *Science* **363**: 409–412.
Ims RA. **1990**. The ecology and evolution of reproductive synchrony. *Trends in Ecology & Evolution* **5**: 135–140.
Jaeger KE, Pullen N, Lamzin S, Morris RJ, Wigge PA. **2013**. Interlocking Feedback Loops Govern the Dynamic Behavior of the Floral Transition in Arabidopsis. *The Plant Cell* **25**: 820–833.
Jaeger KE, Wigge PA. **2007**. FT Protein Acts as a Long-Range Signal in Arabidopsis. *Current Biology* **17**: 1050–1054.
Jang S, Torti S, Coupland G. **2009**. Genetic and spatial interactions between FT, TSF and SVP during the early stages of floral induction in Arabidopsis. *The Plant Journal* **60**: 614–625.
Johanson U, West J, Lister C, Michaels S, Amasino R, Dean C. **2000**. Molecular Analysis of FRIGIDA, a Major Determinant of Natural Variation in Arabidopsis Flowering Time. *Science* **290**: 344–347.
Kang M-J, Jin H-S, Noh Y-S, Noh B. **2015**. Repression of flowering under a noninductive photoperiod by the HDA9-AGL19-FT module in Arabidopsis. *New Phytologist* **206**: 281–294.
Kim D-H. **2020**. Current understanding of flowering pathways in plants: focusing on the vernalization pathway in Arabidopsis and several vegetable crop plants. *Horticulture, Environment, and Biotechnology* **61**: 209–227.
Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. **2013**. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology* **14**: R36.
Kim D-H, Sung S. **2013**. Coordination of the Vernalization Response through a VIN3 and FLC Gene Family Regulatory Network in Arabidopsis. *The Plant Cell* **25**: 454–469.
Kim D-H, Sung S. **2017**. The Binding Specificity of the PHD-Finger Domain of VIN3 Moderates Vernalization Response. *Plant Physiology* **173**: 1258–1268.
Kinmonth-Schultz HA, MacEwen MJS, Seaton DD, Millar AJ, Imaizumi T, Kim S-H. **2019**. An explanatory model of temperature influence on flowering through whole-plant accumulation of FLOWERING LOCUS T in Arabidopsis thaliana. *in silico Plants* **1**.
Krzymuski M, Andrés F, Cagnola JI, Jang S, Yanovsky MJ, Coupland G, Casal JJ. **2015**. The dynamics of FLOWERING LOCUS T expression encodes long-day information. *The Plant Journal* **83**: 952–961.
Kumar R, Ichihashi Y, Kimura S, Chitwood DH, Headland LR, Peng J, Maloof JN, Sinha NR. **2012**. A High-Throughput Method for Illumina RNA-Seq Library Preparation. *Frontiers in Plant Science* **3**.
Law CW, Chen Y, Shi W, Smyth GK. **2014**. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biology* **15**: R29.
Lee I, Amasino RM. **1995**. Effect of Vernalization, Photoperiod, and Light Quality on the Flowering Phenotype of Arabidopsis Plants Containing the FRIGIDA Gene. *Plant Physiology* **108**: 157–162.
Leijten W, Koes R, Roobeek I, Frugis G. **2018**. Translating Flowering Time from Arabidopsis thaliana to Brassicaceae and Asteraceae Crop Species. *Plants* **7**: 111.
Liao Y, Smyth GK, Shi W. **2014**. featureCounts: an efficient general purpose program for

assigning sequence reads to genomic features. *Bioinformatics* **30**: 923–930.
Lin M-K, Belanger H, Lee Y-J, Varkonyi-Gasic E, Taoka K-I, Miura E, Xoconostle-Cázares B, Gendler K, Jorgensen RA, Phinney B, *et al.*2007. FLOWERING LOCUS T Protein May Act as the Long-Distance Florigenic Signal in the Cucurbits. *The Plant Cell* **19**: 1488–1506.
Liu S, He M, Lin X, Kong F. **2023**. Epigenetic regulation of photoperiodic flowering in plants. *The Plant Genome* **n/a**: e20320.
López-González L, Mouriz A, Narro-Diego L, Bustos R, Martínez-Zapater JM, Jarillo JA, Piñeiro M. **2014**. Chromatin-Dependent Repression of the Arabidopsis Floral Integrator Genes Involves Plant Specific PHD-Containing Proteins[C][W]. *The Plant Cell* **26**: 3922–3938.
Michaels SD, Himelblau E, Kim SY, Schomburg FM, Amasino RM. **2005**. Integration of Flowering Signals in Winter-Annual Arabidopsis. *Plant Physiology* **137**: 149–156.
Miryeganeh M. **2020**. Synchronization of senescence and desynchronization of flowering in Arabidopsis thaliana. *AoB PLANTS* **12**.
Miryeganeh M, Yamaguchi M, Kudoh H. **2018**. Synchronisation of Arabidopsis flowering time and whole-plant senescence in seasonal environments. *Scientific Reports* **8**: 10282.
Notaguchi M, Abe M, Kimura T, Daimon Y, Kobayashi T, Yamaguchi A, Tomita Y, Dohi K, Mori M, Araki T. **2008**. Long-Distance, Graft-Transmissible Action of Arabidopsis FLOWERING LOCUS T Protein to Promote Flowering. *Plant and Cell Physiology* **49**: 1645–1658.
Platt A, Horton M, Huang YS, Li Y, Anastasio AE, Mulyati NW, Ågren J, Bossdorf O, Byers D, Donohue K, *et al.*2010. The Scale of Population Structure in Arabidopsis thaliana. *PLOS Genetics* **6**: e1000843.
Postma FM, Ågren J. **2016**. Early life stages contribute strongly to local adaptation in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences* **113**: 7590–7595.
Primack RB. **1985**. Patterns of Flowering Phenology in Communities, Populations, Individuals, and Single Flowers. In: White J, ed. Handbook of Vegetation Science. The Population Structure of Vegetation. Dordrecht: Springer Netherlands, 571–593.
Putterill J, Varkonyi-Gasic E. **2016**. FT and florigen long-distance flowering control in plants. *Current Opinion in Plant Biology* **33**: 77–82.
Rathcke B, Lacey EP. **1985**. Phenological Patterns of Terrestrial Plants. *Annual Review of Ecology and Systematics* **16**: 179–214.
Rubin MJ, Friedman J. **2018**. The role of cold cues at different life stages on germination and flowering phenology. *American Journal of Botany* **105**: 749–759.
Salazar JD, Saithong T, Brown PE, Foreman J, Locke JCW, Halliday KJ, Carré IA, Rand DA, Millar AJ. **2009**. Prediction of Photoperiodic Regulators from Quantitative Gene Circuit Models. *Cell* **139**: 1170–1179.
Satake A. **2010**. Diversity of plant life cycles is generated by dynamic epigenetic regulation in response to vernalization. *Journal of Theoretical Biology* **266**: 595–605.
Schönrock N, Bouveret R, Leroy O, Borghi L, Köhler C, Gruissem W, Hennig L. **2006**. Polycomb-group proteins repressthe floral activator AGL19 in the FLC-independent vernalization pathway. *Genes & Development* **20**: 1667–1678.
Searle I, He Y, Turck F, Vincent C, Fornara F, Kröber S, Amasino RA, Coupland G. **2006**. The transcription factor FLC confers a flowering response to vernalization by repressing meristem competence and systemic signaling in Arabidopsis. *Genes & Development* **20**: 898–912.
Sheldon CC, Jean Finnegan E, Dennis ES, James Peacock W. **2006**. Quantitative effects of vernalization on FLC and SOC1 expression. *The Plant Journal* **45**: 871–883.
Sheldon CC, Jean Finnegan E, James Peacock W, Dennis ES. **2009**. Mechanisms of gene repression by vernalization in Arabidopsis. *The Plant Journal* **59**: 488–498.
Shimizu KK, Purugganan MD. **2005**. Evolutionary and Ecological Genomics of Arabidopsis. *Plant Physiology* **138**: 578–584.
Somssich M. **2020**. *A Short History of Vernalization*. Zenodo.

Song YH, Lee I, Lee SY, Imaizumi T, Hong JC. **2012**. CONSTANS and ASYMMETRIC LEAVES 1 complex is involved in the induction of FLOWERING LOCUS T in photoperiodic flowering in Arabidopsis. *The Plant Journal* **69**: 332–342.
Srikanth A, Schmid M. **2011**. Regulation of flowering time: all roads lead to Rome. *Cellular and Molecular Life Sciences* **68**: 2013–2037.
Sung S, Amasino RM. **2004**. Vernalization in Arabidopsis thaliana is mediated by the PHD finger protein VIN3. *Nature* **427**: 159–164.
Sung S, Schmitz RJ, Amasino RM. **2006**. A PHD finger protein involved in both the vernalization and photoperiod pathways in Arabidopsis. *Genes & Development* **20**: 3244–3248.
Suzuki A, Metzger JD. **2001**. Vernalization in a Greenhouse Promotes and Synchronizes Flowering of Osteospermum ecklonis Norl. *HortScience* **36**: 658–660.
Tamaki S, Matsuo S, Wong HL, Yokoi S, Shimamoto K. **2007**. Hd3a Protein Is a Mobile Flowering Signal in Rice. *Science* **316**: 1033–1036.
Teotia S, Tang G. **2015**. To Bloom or Not to Bloom: Role of MicroRNAs in Plant Flowering. *Molecular Plant* **8**: 359–377.
Tyler L, Fangel JU, Fagerström AD, Steinwand MA, Raab TK, Willats WG, Vogel JP. **2014**. Selection and phenotypic characterization of a core collection of Brachypodium distachyon inbred lines. *BMC Plant Biology* **14**: 25.
Wilczek AM, Roe JL, Knapp MC, Cooper MD, Lopez-Gallego C, Martin LJ, Muir CD, Sim S, Walker A, Anderson J, *et al.*2009. Effects of Genetic Perturbation on Seasonal Life History Plasticity. *Science* **323**: 930–934.
Wood CC, Robertson M, Tanner G, Peacock WJ, Dennis ES, Helliwell CA. **2006**. The Arabidopsis thaliana vernalization response requires a polycomb-like protein complex that also includes VERNALIZATION INSENSITIVE 3. *Proceedings of the National Academy of Sciences* **103**: 14631–14636.
Xu S, Chong K. **2018**. Remembering winter through vernalisation. *Nature Plants* **4**: 997–1009.
Yanovsky MJ, Kay SA. **2002**. Molecular basis of seasonal time measurement in Arabidopsis. *Nature* **419**: 308–312.

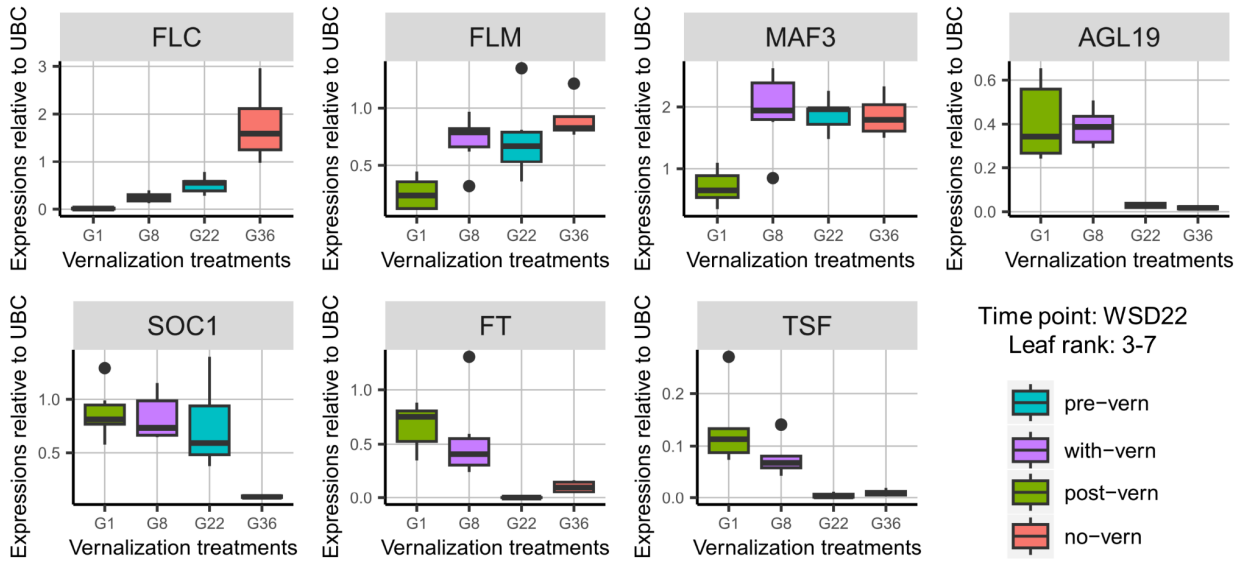

Fig. S1.1: The expression dynamics of seven candidate genes in the leaf rank 3-7 at WSD22. Each panel represents a single gene. Within each panel, gene expressions are grouped based on the vernalization treatments (G1, G8, G22, G36, and NV) and color-coded according to the leaf's vernalization status (no-vern, post-vern, pre-vern, and with-vern), as referenced in Fig. 1.4 and Table 1.1.

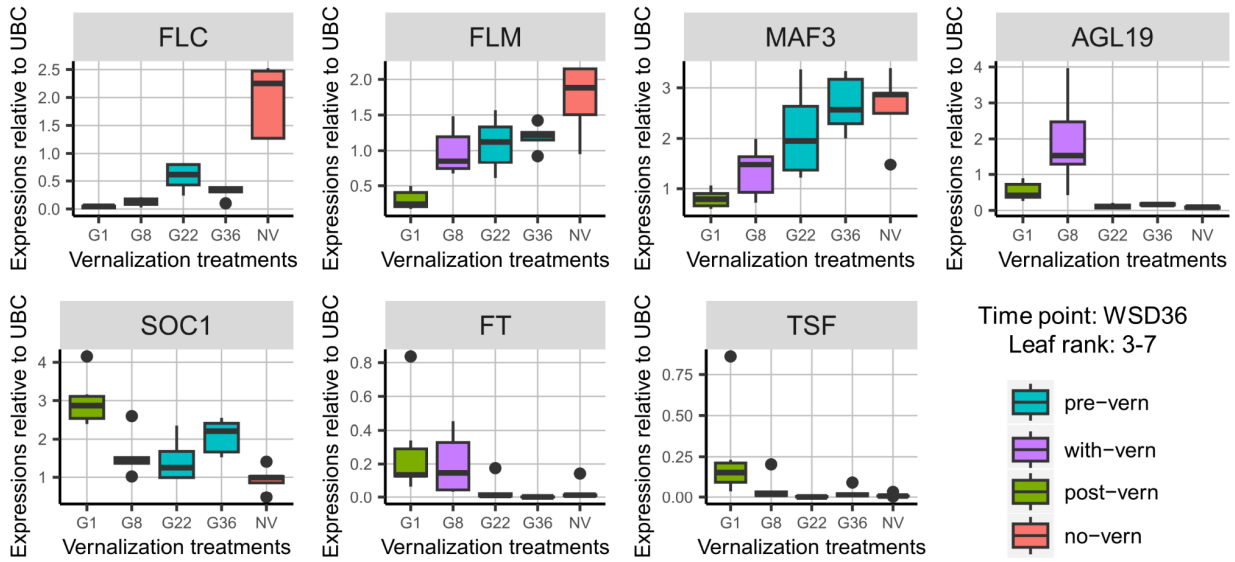

Fig. S1.2: The expression dynamics of seven candidate genes in the leaf rank 3-7 at WSD36. Each panel represents a single gene. Within each panel, gene expressions are grouped based on the vernalization treatments (G1, G8, G22, G36, and NV) and color-coded according to the leaf's vernalization status (no-vern, post-vern, pre-vern, and with-vern), as referenced in Fig. 1.4 and Table 1.1.

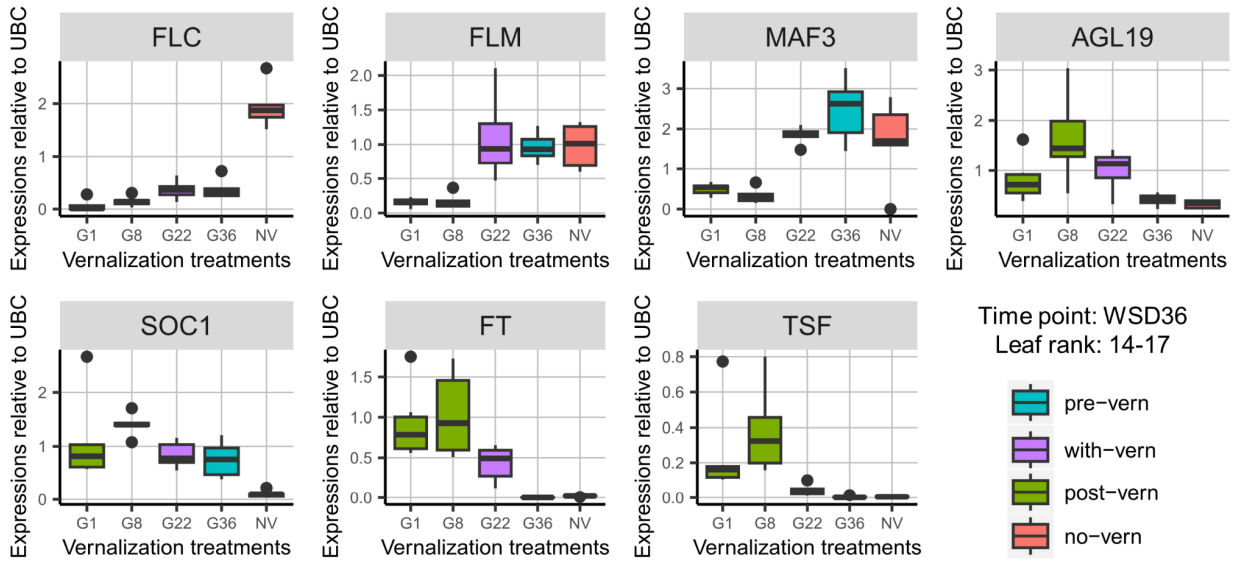

Fig. S1.3: The expression dynamics of seven candidate genes in the leaf rank 14-17 at WSD36. Each panel represents a single gene. Within each panel, gene expressions are grouped based on the vernalization treatments (G1, G8, G22, G36, and NV) and color-coded according to the leaf's vernalization status (no-vern, post-vern, pre-vern, and with-vern), as referenced in Fig. 1.4 and Table 1.1.

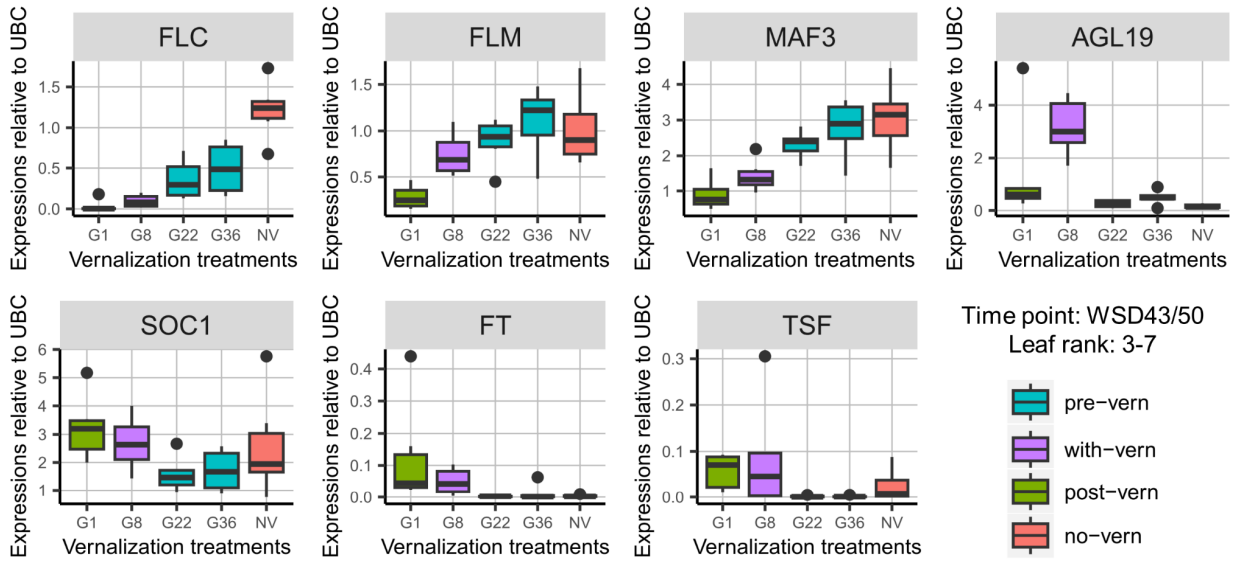

Fig. S1.4: The expression dynamics of seven candidate genes in the leaf rank 3-7 at WSD43/50. Each panel represents a single gene. Within each panel, gene expressions are grouped based on the vernalization treatments (G1, G8, G22, G36, and NV) and color-coded according to the leaf's vernalization status (no-vern, post-vern, pre-vern, and with-vern), as referenced in Fig. 1.4 and Table 1.1.

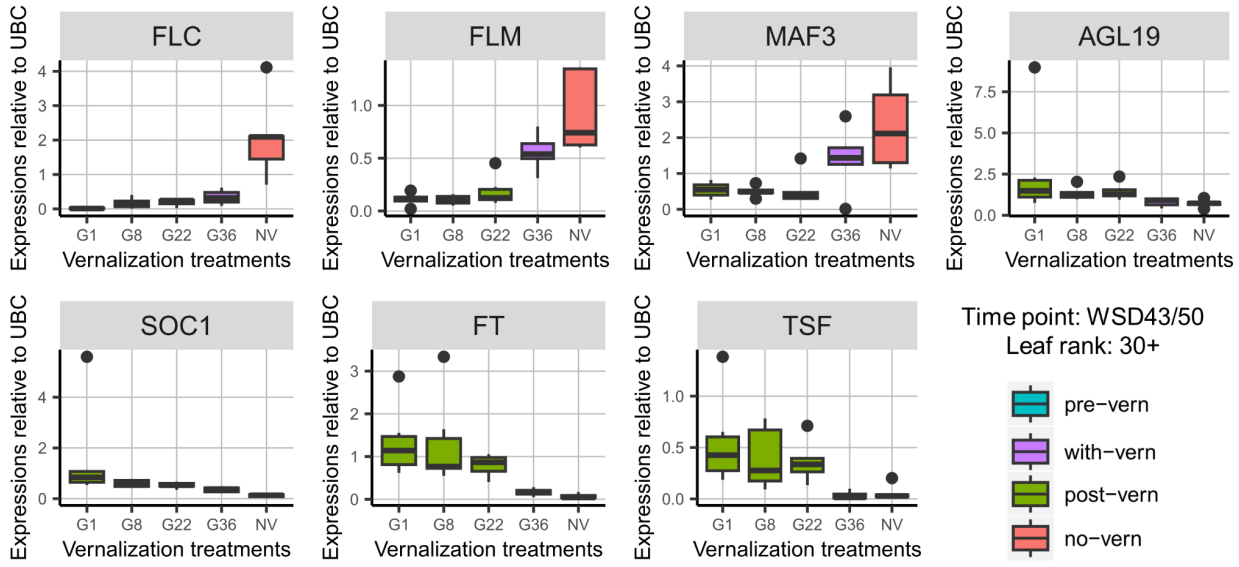

Fig. S1.5: The expression dynamics of seven candidate genes in the leaf rank 30+ at WSD43/50. Each panel represents a single gene. Within each panel, gene expressions are grouped based on the vernalization treatments (G1, G8, G22, G36, and NV) and color-coded according to the leaf's vernalization status (no-vern, post-vern, pre-vern, and with-vern), as referenced in Fig. 1.4 and Table 1.1.

Table S1.1: ANOVA tests for the expression differences within each combination of vernalization treatment and leaf rank for seven target genes. The significance of the results indicates the presence of developmental-stage-dependent vernalization responses in at least one pair of gene expression comparisons among G1, G8, G22, G36, and NV. P-values greater than 0.05 are highlighted in red.

Gene	Vernalization treatment	Leaf rank	p-value
FLC	WSD22	3-7	5.19E-06
FLC	WSD36	3-7	4.35E-10
FLC	WSD36	14-17	1.57E-12
FLC	WSD43/50	3-7	3.05E-08
FLC	WSD43/50	14-17	2.67E-06
FLC	WSD43/50	30+	6.11E-06
FT	WSD22	3-7	0.000306
FT	WSD36	3-7	0.02983
FT	WSD36	14-17	1.23E-05
FT	WSD43/50	3-7	0.04968
FT	WSD43/50	14-17	7.65E-05
FT	WSD43/50	30+	0.003628
FLM	WSD22	3-7	2.41E-03
FLM	WSD36	3-7	4.82E-06

FLM	WSD36	14-17	1.22E-05
FLM	WSD43/50	3-7	3.41E-04
FLM	WSD43/50	14-17	1.13E-04
FLM	WSD43/50	30+	2.53E-07
MAF3	WSD22	3-7	1.76E-04
MAF3	WSD36	3-7	4.33E-05
MAF3	WSD36	14-17	2.47E-06
MAF3	WSD43/50	3-7	1.38E-05
MAF3	WSD43/50	14-17	2.29E-03
MAF3	WSD43/50	30+	3.47E-04
AGL19	WSD22	3-7	2.35E-06
AGL19	WSD36	3-7	5.05E-05
AGL19	WSD36	14-17	7.06E-04
AGL19	WSD43/50	3-7	1.08E-04
AGL19	WSD43/50	14-17	1.45E-06
AGL19	WSD43/50	30+	1.97E-01
TSF	WSD22	3-7	2.08E-04

TSF	WSD36	3-7	3.59E-02
TSF	WSD36	14-17	1.29E-03
TSF	WSD43/50	3-7	8.43E-02
TSF	WSD43/50	14-17	1.98E-04
TSF	WSD43/50	30+	1.05E-02
SOC1	WSD22	3-7	1.22E-03
SOC1	WSD36	3-7	1.23E-05
SOC1	WSD36	14-17	6.86E-04
SOC1	WSD43/50	3-7	8.47E-02
SOC1	WSD43/50	14-17	4.22E-03
SOC1	WSD43/50	30+	9.57E-02

Table S1.2: Primers used in this article.

Primer Name	Sequence	Purpose
AGL19_F	ATGGTGAGGGGCAAAACGGAG	qPCR
AGL19_R	CCAGATGTTTCGTCTCTCGC	qPCR
FLC_393F	AGCCAAGAAGACCGAACTCA	qPCR
FLC_550R	TTTGTCCAGCAGGTGACATC	qPCR
FLM_B_F	CATGCTGATGAACTTAGAGCCTTAGATC	qPCR
FLM_B_R	CAGCAACGTATTCTTTCCCAT	qPCR
FT_372F	CTGGAACAACCTTTGGCAAT	qPCR
FT_590R	AGCCACTCTCCCTCTGACAA	qPCR
MAF3_F	TCGGAATTATCTTCCACACAAGGAG	qPCR
MAF3_R	GCCAGAATCTGGTTCTCTTCTATCAGC	qPCR
SOC1_F	AGCTGCAGAAAACGAGAAGC	qPCR
SOC1_R	TGAAGAACAAGGTAACCCAATG	qPCR
TSF_F	GAGTCCAAGCAACCCTCACCAA	qPCR
TSF_R	CACAATACGATGAATTCCCGAG	qPCR
UBC_F	CTGCGACTCAGGGAATCTTCTAA	qPCR
UBC_R	TTGTGCCATTGAATTGAACCC	qPCR

ListerFRI_F	GCGAGACAAGTTTCGCTTCT	Col-FRI genotyping
ListerFRI_R	GCAAAGGTGGTTCCTTTTGT	Col-FRI genotyping
AGL19_LP	TCTGATCTACACACATGCGATG	*agl19-1* genotyping
AGL19_RP	TGTGATGCTGAAGTTGCTTTG	*agl19-1* genotyping
LBb1-3	ATTTTGCCGATTTCGGAAC	*agl19-1* genotyping

Table S1.3: Target genes in RNAseq analysis with adjusted p-values smaller than 0.001.

Gene ID	Adjusted p-value	Gene name
AT1G77080	1.67E-07	FLM
AT2G45660	1.06E-05	SOC1
AT4G22950	1.06E-05	AGL19
AT4G20370	8.59E-06	TSF
AT5G65060	8.59E-06	MAF3
AT1G65480	1.90E-05	FT
AT5G10140	8.59E-06	FLC

Chapter 2 Understanding the molecular mechanisms of obligate photoperiodism in *Mimulus guttatus*

2.1 Abstract

Daylength is an environmental factor that plays a vital role in regulating the flowering responses of plants. While our understanding of the genetic components and regulatory networks involved in the photoperiod pathway has advanced, our knowledge is still limited to facultative long-day or short-day plants. To gain insights into the molecular mechanisms underlying the stringent photoperiodic responses observed in plants with obligate-type photoperiodism, we conducted a series of experiments using the annual *Mimulus guttatus*, a well-known obligate long-day plant. We generated several lines exhibiting facultative long-day photoperiodism phenotypes. We employed QTL mapping to identify the genetic basis of photoperiodism and flowering time of two mapping populations under the short-day condition. Interestingly, we identified specific QTLs related to photoperiodism but not flowering time in short days. Furthermore, the QTL results suggested that different accessions of *Mimulus guttatus* achieved similar photoperiodism phenotypes through distinct genetic mechanisms. Our RNAseq analyses provided further support for this hypothesis. Overall, this study offers new insights into the molecular basis of obligate long-day photoperiodism, shedding light on the intricate regulation of plant flowering in response to daylength cues.

2.2 Introduction

To maximize reproductive success, plants integrate multiple environmental cues to precisely align their flowering time with optimal conditions. Daylength, also known as photoperiod, is an environmental factor that is highly correlated with seasonal changes and is used as a signal for regulating flowering time in many plant species. Based on their flowering responses to daylength, plants can be categorized into three main photoperiodism groups: day-neutral plants, long-day plants (LDPs), and short-day plants (SDPs). Day-neutral plants are capable of flowering at the same time regardless of the photoperiod, while LDPs demonstrate accelerated flowering when the daylength exceeds a specific threshold. Conversely, SDPs exhibit delayed flowering when the daylength extends beyond a maximum threshold. The flowering responses of LDPs and SDPs can be further classified as obligate or facultative, depending on whether a specific daylength is essential for initiating flowering. For example, facultative LDPs can eventually flower in any daylength, but longer photoperiods can expedite the process. In contrast, obligate LDPs have a minimum daylength requirement, known as the critical photoperiod, and they will only flower if this requirement is met. Despite comparable numbers of species in each photoperiodism group (Thomas & Vince-Prue, 1997), our understanding of the underlying molecular mechanisms of photoperiodic responses is predominantly limited to facultative LDPs and SDPs. Consequently, we know little about the mechanisms that control the stringent daylength requirements observed in plants with obligate photoperiodism.

Photoperiodic flowering regulation is best understood in *Arabidopsis thaliana* (hereafter referred to as Arabidopsis) and rice, which are facultative LDP and SDP, respectively. In both species, a canonical photoperiod-measuring regulation known as the CO-FT module plays a key role. In Arabidopsis, this module consists of the genes *CONSTANS* (*CO*) and *FLOWERING LOCUS T* (*FT*), while in rice, their homologs *Heading date1* (*Hd1*) and *Heading date3a* (*Hd3a*),

respectively, are involved (Song *et al.*, 2015; Hill & Li, 2016; Gendron & Staiger, 2023). The FT protein functions as the main component of florigen, which is transported from leaves to the shoot apical meristem to trigger flowering (Lin *et al.*, 2007; Corbesier *et al.*, 2007; Tamaki *et al.*, 2007; Jaeger & Wigge, 2007; Notaguchi *et al.*, 2008). Under inductive photoperiods (long days for Arabidopsis and short days for rice), *FT* expression is upregulated in leaves (Kardailsky *et al.*, 1999; Kobayashi *et al.*, 1999; Kojima *et al.*, 2002). The regulation of *FT* expression is controlled by the transcription factor CO, which exhibits a circadian oscillation and acts as a bridge between the endogenous circadian rhythm and external photoperiodic sensitivity (Suárez-López *et al.*, 2001; Yanovsky & Kay, 2002; Kojima *et al.*, 2002). Interestingly, CO exerts distinct effects on *FT* expression in Arabidopsis and rice. In Arabidopsis, CO functions as an activator, promoting *FT* expression under inductive long-day conditions (Kobayashi *et al.*, 1999; Samach *et al.*, 2000; Yanovsky & Kay, 2002). However, in rice, CO serves as a bifunctional regulator of *FT*: the rice CO homolog, Hd1, induces the *FT* homolog, *Hd3a*, in inductive short-day conditions, while Hd1 acts as an repressor of *Hd3a* transcription in non-inductive long-day conditions (Yano *et al.*, 2000; Hayama *et al.*, 2003).

In addition to the canonical FT-CO module, plants recruit a wide range of genes to respond to varying daylengths and precisely regulate flowering based on local conditions. The interplay between the circadian clock, metabolic oscillation, and photoreceptors also influences photoperiodic flowering (Ballerini & Kramer, 2011; Song *et al.*, 2015; Gendron & Staiger, 2023). Interestingly, different plant species can display distinct levels of dependence on specific mechanisms, and some species even possess their own unique species-specific or lineage-specific pathways. For instance, in rice, alongside the FT-CO module, a monocot-specific pathway involving *Grain number and heading date7* (*Ghd7*) and *Early heading date1* (*Ehd1*) operates in parallel to regulate the expression of *Hd3a* (Doi *et al.*, 2004; Itoh *et al.*, 2010; Osugi *et al.*, 2011). This intrinsic complexity within the photoperiodic regulation network

equips plants with the flexibility and robust regulatory capacity required to fine-tune their flowering time in accordance with different environmental conditions and habitats. Moreover, this complexity forms the foundation for the diversity observed in plant photoperiodic responses.

The diversity in photoperiodic responses manifests through the variations in two distinct floral traits: flowering time and photoperiodism. While variation in flowering time is common (e.g., numerous documented Arabidopsis mutants exhibit significant differences in flowering time compared to their wild-type counterparts), transitions among photoperiodism groups can also occur, and numerous examples have been documented across diverse species (Thomas & Vince-Prue, 1997). Interestingly, mutations in only a few genes can switch plants from one form of photoperiodism to another. For instance, the *ga1-3* mutant in Arabidopsis, which is deficient in gibberellin synthesis, displayed obligate LDP characteristics and failed to flower in short days even after an extended growing period (> 117 days) (Wilson *et al.*, 1992). This gibberellin-dependent short-day flowering regulation is mediated by the floral pathway integrators, *SUPPRESSOR OF OVEREXPRESSION OF CO 1* (*SOC1*) and *AGAMOUS-LIKE 24* (*AGL24*). Consistently, the *soc1-2/agl24-1* double mutant also displayed similar obligate LDP characteristics and a flowering defect in short days (Liu *et al.*, 2008). In rice, flowering under non-inductive long-day conditions relies on *OsMADS50* and *Rice FT-like 1* (*RFT1*), an *FT* homolog other than *Hd3a*. *OsMADS50* acts as an upstream activator of *RFT1* in long days (Ryu *et al.*, 2009; Komiya *et al.*, 2009). Loss-of-function mutations in *RFT1* or *OsMADS50* resulted in the obligate SDP responses, demonstrated by failure to flower under long-day conditions (Komiya *et al.*, 2009; Ogiso-Tanaka *et al.*, 2013).

However, the insights gained from the aforementioned transitions among photoperiodism groups do not directly elucidate the underlying mechanism governing flowering regulation in obligate LDPs or obligate SDPs. The inherent flexibility and vast diversity of plant photoperiodic

responses pose challenges in extrapolating our current knowledge to infer potential mechanisms in new species or across different photoperiodism groups. Furthermore, the scarcity of molecular-level studies conducted on plants exhibiting obligate photoperiodisms hinders comparative analyses between photoperiodism groups. In this study, we focus on obligate LDPs and propose two models based on our existing understanding of photoperiodic regulation (Fig. 2.1). The observed existence of simple genetic bases facilitating transitions among photoperiodism groups suggests the presence of certain genes that may act as toggle switches in determining photoperiodic responses. Expanding on this concept, we propose Model 1, assuming the presence of switch-like genes responsible for determining whether a plant can initiate flowering under specific daylength conditions. Once the plant is capable of flowering, a separate set of genes acting as dials regulate the flowering time. The switch-like genes might interact epistatically with the dial-like genes, or they might be located upstream to activate the expression of the dial-like genes. Hence, flowering regulation under this model is controlled by two distinct sets of genes: switch-like genes and dial-like genes. In contrast, Model 2 hypothesizes the absence of switch-like genes. Instead, plants growing under non-inductive photoperiods prolong the flowering process indefinitely by adjusting the dial-like genes, resulting in the absence of observable flowering within their lifespan.

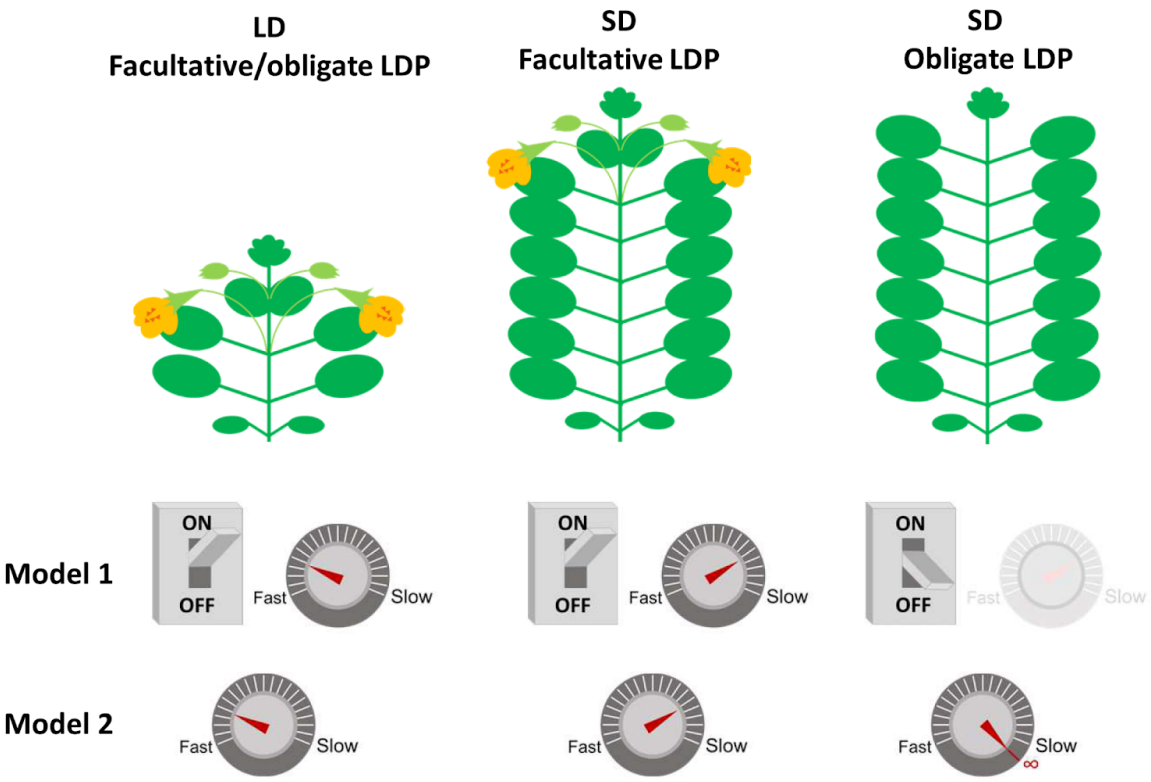

Fig. 2.1: Switch-vs-dial Model. The primary distinction between Model 1 and Model 2 lies in the presence of switch-like genes that specifically determine whether a plant will flower under certain daylength conditions. In Model 1, when exposed to conditions that can induce flowering, such as long days for both obligate and facultative LDPs or short days for facultative LDPs, the switch-like genes are in the "ON" state, and the timing of flowering is regulated by dial-like genes. However, if the daylength does not exceed the critical photoperiod of obligate LDPs, the switch-like genes remain in the "OFF" state, resulting in a non-flowering phenotype. In this situation, the state of the dial-like genes either has no effect on flowering, or they are not activated due to the absence of inductive signals from the switch-like genes. On the other hand, Model 2 represents the dial-only model, where switch-like genes are absent. In this case, plants unable to flower in short days indefinitely delay their flowering date through the action of dial-like genes, leading to a perpetual postponement without reaching the flowering stage.

To test these hypotheses, we conducted a series of experiments using annual populations of *Mimulus guttatus* (hereafter referred to as Mimulus). Mimulus is generally considered to be an obligate LDP with significant phenotypic (Friedman *et al.*, 2015; Kooyers *et al.*, 2015) as well as genetic diversity (Sweigart & Willis, 2003; Twyford & Friedman, 2015; Puzey *et al.*, 2017). Its distribution spans a wide geographic range throughout western North

America, encompassing diverse habitats ranging from sea level to elevations exceeding 3,000 m. These habitats differ in various environmental factors, including variations in precipitation, seasonal temperature, soil compositions, and biotic factors (Wu *et al.*, 2008). Such factors collectively shape the timing of the growing season, and different Mimulus populations are able to precisely respond to certain specific external cues and properly align their life cycle with the optimal growth period. An example highlighting this local adaptation is the association observed between the critical photoperiod required for flowering and the onset of the growing season (Kooyers *et al.*, 2015). Intriguingly, according to the detailed records from previous studies, several annual Mimulus populations exhibit remarkably short critical photoperiods, some as brief as 8 hours (Friedman & Willis, 2013; Kooyers *et al.*, 2015), suggesting naturally-occurred non-obligate LD accessions may be present within Mimulus.

In this study, we aimed to identify the mechanisms underlying the regulation of critical photoperiod and short-day flowering time in Mimulus, which serves as an example of a predominantly obligate LD species. We first screened and characterized non-obligate LD Mimulus plants. We then used QTL mapping to explore the genetic basis underlying the differences in short-day flowering time and photoperiodism, thereby testing our switch-vs-dial model. Finally, we conducted RNAseq to investigate the gene expression patterns associated with the distinct flowering responses observed among accessions exhibiting different photoperiodisms. Our results provide new insights into the molecular basis of obligate LD photoperiodism.

2.3 Results

2.3.1 A screen of natural accessions identifies several facultative LD lines in Mimulus

To investigate the genetic basis of photoperiodism, it is essential to study genotypes that exhibit variation in this trait. In a study by Kooyers *et al.* (Kooyers *et al.*, 2015), the critical photoperiods of 52 Mimulus accessions representing much of the Mimulus geographic range were evaluated. Among these accessions, three (BEL, MAC, and TER) were observed to achieve over 50% flowering even when exposed to daylength as short as 8 hours. The ability to flower under an 8-hour daylength is typically considered short-day flowering, especially since winter daylengths across the range of these three accessions never fall below 9.5 hours. However, due to it being a predominantly outcrossing species, Mimulus accessions exhibit significant genetic and phenotypic diversity and maintain high levels of heterozygosity throughout their genomes (Sweigart & Willis, 2003; Twyford & Friedman, 2015; Puzey *et al.*, 2017). Taking this into account, we conducted a screening within BEL, MAC, and TER accessions to identify maternal lines capable of successful short-day flowering. A minimum of eight maternal lines per accession were assessed, with at least five individuals from each line, under 9:15 hour day/night cycles (short days, SD) over a duration of 16 weeks. We successfully generated several lines from BEL, MAC, and TER accessions that exhibited reliable SD flowering phenotypes (i.e., non-obligate LD lines) after at least four generations of selection and inbreeding. To serve as negative controls in the subsequent experiments, we also generated typical obligate LD lines from these three accessions as well as four other accessions with longer critical photoperiods (Fig. S2.1).

In this study, we focused on two flowering phenotypes: photoperiodism and flowering time. The photoperiodism group assigned to each line was based on its ability to flower in SD, which allowed us to differentiate between obligate LDPs and facultative LDPs. Additionally, we

examined whether the flowering times differ under SD and the long-day (16:8 day/night cycles, LD) conditions to determine if a plant exhibits day-neutral characteristics. To compare flowering times under different daylengths, we used the node of the first flower bud as a measure of flowering time instead of the node of the first fully open flower. Although the latter trait has been commonly used in previous flowering-related studies, we observed that many Mimulus plants produced aborted flower buds under shorter daylengths (Fig. 2.2). The presence of aborted buds indicated that the initiation of flowering occurred but further development was halted, potentially due to inadequate metabolic resources or a lack of signaling for subsequent developmental stages. This phenomenon also indicated a decoupling between flowering induction and development in these situations. In addition, the number of aborted flower buds differed among lines, suggesting variation in the regulation of flower bud development.

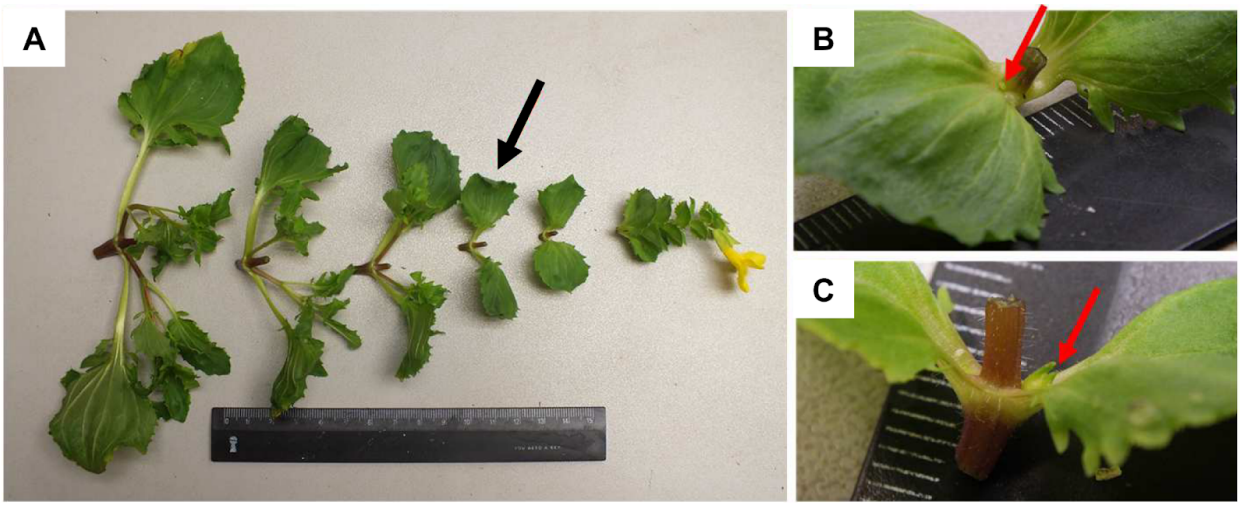

Fig. 2.2: Aborted flower buds observed in SD. (A) The black arrow indicates the node where aborted flower buds were observed. Note that the first flower was produced at a node distant from the node of the first flower bud. Zoomed-in views (B) and (C) highlight the nodes with aborted flower buds (indicated by red arrows) halted at different developmental stages.

After characterizing the flowering phenotypes of the Mimulus lines we generated, we found that all the non-obligate LD lines derived from BEL, MAC, and TER exhibited facultative LD photoperiodism. As shown in Fig. 2.3, these lines were able to initiate flowering with an

average of fewer than 5 nodes under the LD condition. However, under the SD condition, a minimum of 10 nodes on average was required for flowering induction, highlighting each line's sensitivity to photoperiod and the delaying effect of SD on flowering. In contrast, the obligate lines, serving as the negative control, did not flower in SD, but their flowering in LD was as rapid as that of the facultative lines (Fig. S2.2). In addition to the observed reproductive stage variation, we also noticed a wide range of vegetative variation across different lines in SD (Fig. S2.3), which is consistent with the Mimulus phenotype diversity described in previous studies.

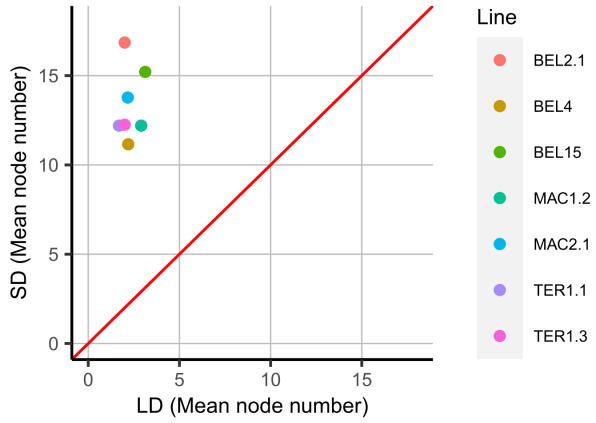

Fig. 2.3: A comparison of the node number of the first flower bud under SD and LD conditions indicates that all the non-obligate LD lines generated in this study exhibit facultative LD photoperiodism. Flowering times are substantially delayed in SD compared to LD. Each dot represents the average node number of the first flower bud for each facultative LD line. The red 1:1 diagonal line depicts the exact identical flowering time under the two photoperiod conditions.

2.3.2 F2 mapping populations display a large degree of variation in flowering phenotypes

To investigate the genetic basis of obligate LD photoperiodism and test the switch-vs-dial hypothesis, we performed QTL mapping. We generated two F2 mapping populations by crossing a facultative LD line and an obligate LD line derived from each of the two accessions: BEL4 x BEL4.2 and MAC1.2 x MAC7, where the first line in each pair represents the facultative LD line, and the second line represents the obligate LD line. We grew a total of 250 F2 plants for

each mapping population in SD for a duration of 22 weeks. For each F2 individual, we evaluated both the photoperiodism and flowering time, which correspond to the “switch” and “dial” components, respectively, in the switch-vs-dial model (see the next section). The photoperiodism of each F2 individual was categorized as either facultative LDP or obligate LDP based on whether flowering was eventually induced within the 22-week experimental period, without considering the time taken to flower. Regarding flowering time, we considered not only the node of the first flower bud, as described in the preceding screening section, but also the day-to-flower since we focused exclusively on the SD condition.

We observed wide phenotypic variation in the flowering-related traits within each F2 mapping population (Fig. 2.4 and Fig. S2.4). For example, morphologically, facultative LD individuals usually exhibited a prominent main stem in SD, with the first flower bud emerging on it. In contrast, obligate LD individuals usually displayed a "bushy" phenotype characterized by numerous side branches but no conspicuous main stem. We also observed intermediate phenotypes with varying degrees of main stem prominence; however, these individuals ultimately failed to develop flower buds and were classified as obligate LDPs. Additionally, a few individuals developed several aborted flower buds in earlier nodes, suggesting that they underwent the reproductive phase change, but the buds in later nodes reverted back to the vegetative phase, ultimately resulting in a failure to flower in SD. Due to the ambiguity of characterization, these plants were excluded from the data analyses below.

Fig. 2.4: Morphological variation within the F2 mapping population of BEL in SD. The leftmost individual exhibits a typical facultative LD phenotype with a prominent main stem. In contrast, the individual with a typical obligate LD phenotype appears "bushy" in SD (rightmost). The population also displays a wide spectrum of obligate LD individuals showing intermediate phenotypes with varying degrees of main stem extension but without flowering induction (middle two).

In addition to the observed morphological variation, we also examined the distribution of the node of the first flower bud and day-to-flower for the F2 individuals that eventually flowered (i.e., facultative LD individuals) (Fig. 2.5). The distributions of these two flowering time-related traits exhibited right-skewed patterns, indicating that the majority of individuals from these populations flowered within a relatively narrow time window in SD, but there were some individuals that displayed a much later flowering time. These findings suggest that photoperiodism is not simply a binary trait - the various morphological and developmental characteristics of photoperiodism groups are not always aligned among individuals.

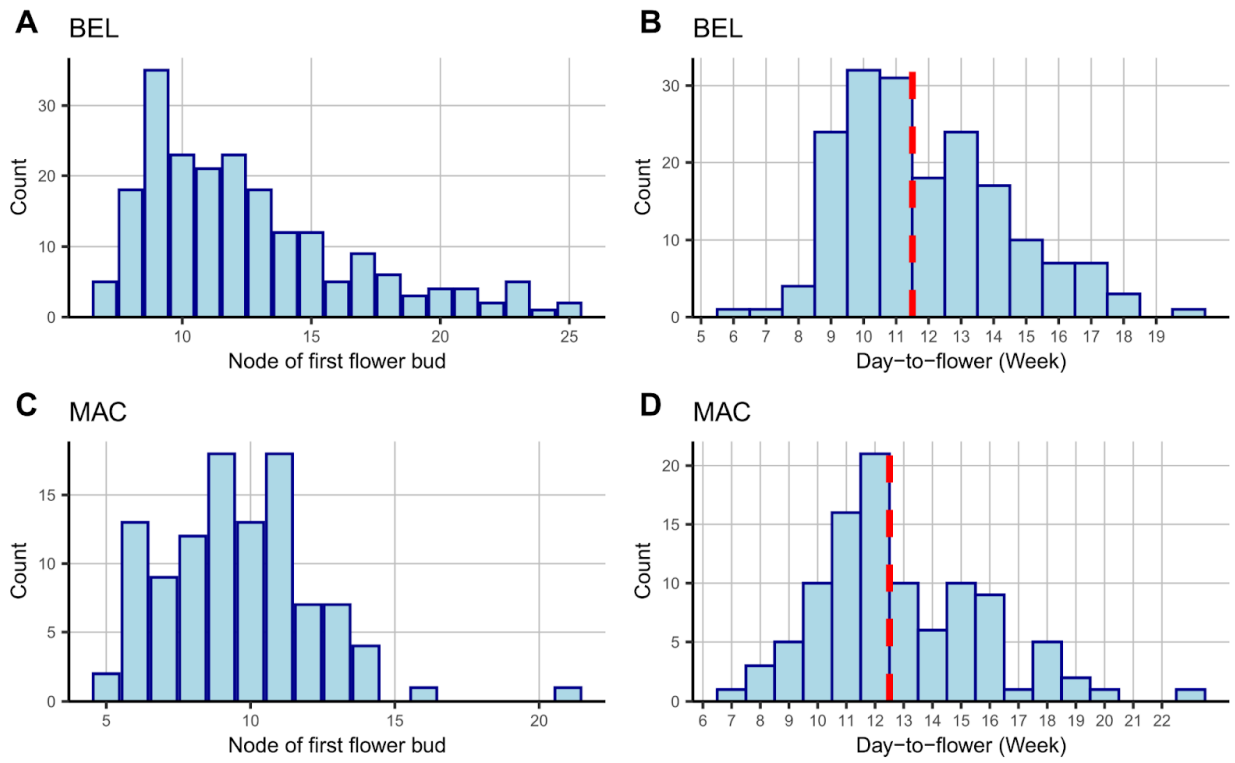

Fig. 2.5: Distribution of flowering time of the facultative LD individuals from the BEL (A and B) and MAC (C and D) F2 mapping populations, displaying right-skewed patterns. Two proxies were used to estimate flowering time: the node of the first flower bud (A and C) and day-to-flower (B and D). Obligate LD individuals are not included in this figure as they did not induce flowering during the experimental period. The red dashed line in (B) and (D) indicates the threshold for selecting individuals that are ready to flower in SD for QTL mapping. The threshold is approximately set at the peak of the day-to-flower distribution.

2.3.3 QTL mapping detects photoperiodism-specific QTLs, providing evidence for the existence of switch-like genes

To investigate our switch-vs-dial hypothesis, we conducted QTL mapping within the two F2 mapping populations for three flowering phenotypes in SD: day-to-flower, the node of the first flower bud, and photoperiodism (facultative vs obligate LD photoperiodism). Day-to-flower and the node of the first flower bud served as indicators of the plant's speed of flowering in SD, representing the "dial" component in our models. For these two flowering time-related traits, obligate LD individuals were not included in the QTL mapping, as they did not have exact

flowering times. On the other hand, we classified the F2 individuals into binary photoperiodism groups based on their ability to flower in SD for the third trait, which represented the "switch" component in our models. If all three traits share the same QTLs, it would support the dial-only model, suggesting that the non-flowering phenotype of obligate LD lines in SD is the result of extremely delayed flowering. Conversely, the presence of photoperiodism-specific QTLs would support the switch model, indicating the involvement of specific genes in determining the ability to flower in SD.

We constructed linkage maps for BEL and MAC by whole-genome resequencing. We identified a total of 134,831 high-quality SNPs for the BEL mapping population and 123,011 high-quality SNPs for the MAC mapping population. Following the removal of redundant markers in strong linkage disequilibrium, the linkage maps for BEL and MAC were generated using 5,319 and 4,961 SNPs, respectively (Fig. 2.6). The genetic maps spanned 5,319 cM for the BEL mapping population and 4,961 cM for the MAC mapping population, distributed across 14 chromosomes. The average interval length between markers was 0.4 cM for both mapping populations.

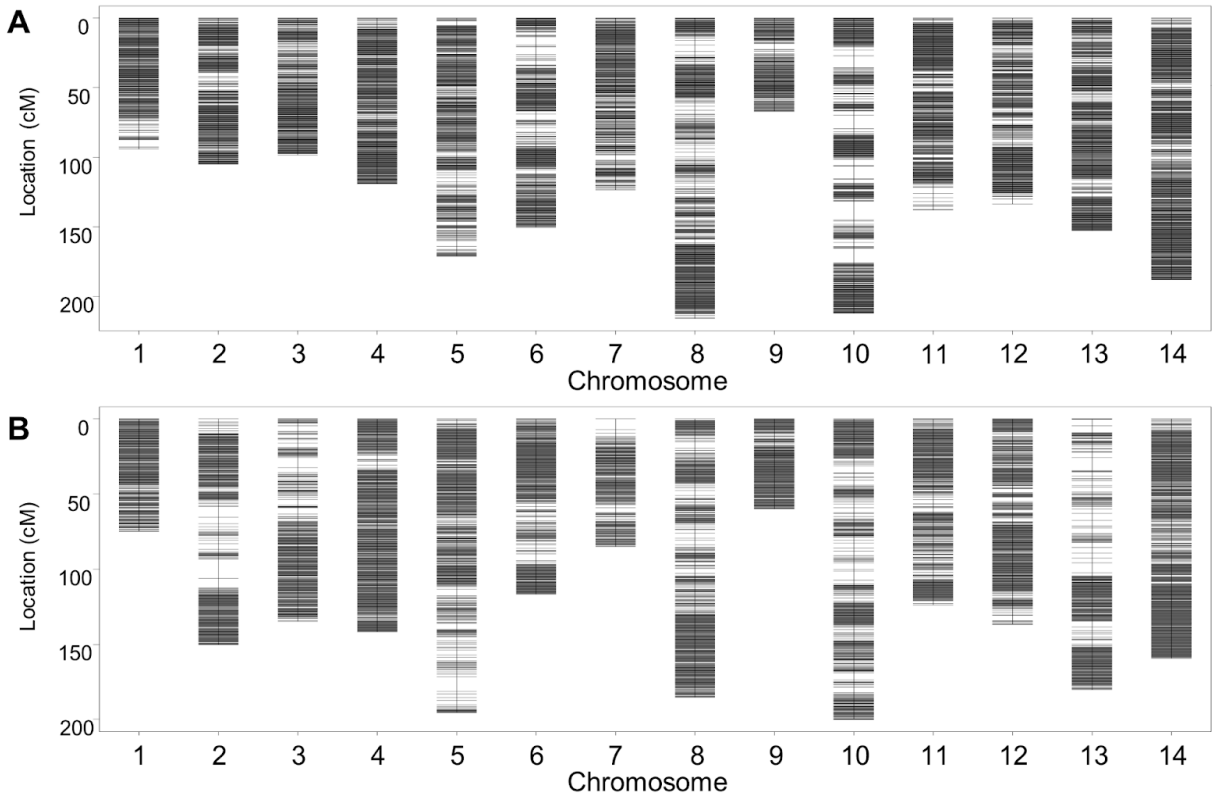

Fig. 2.6: Linkage maps of BEL (A) and MAC (B) based on SNP markers obtained from whole-genome resequencing. The black lines represent the positions of the markers on 14 chromosomes.

Using the BEL mapping population, we identified two QTLs for both day-to-flower and the node of the first flower bud (Fig. 2.7A), with a shared QTL located on chromosome 10 (refer to Table 1 for the effects of these QTLs). Interestingly, a unique QTL was associated with each flowering time-related trait: a QTL on chromosome 12 for day-to-flower and a QTL on chromosome 7 for the node of the first flower bud. Discrepancies in the results between these two flowering time-related traits might stem from inherent developmental differences underlying these proxies for flowering time, or from potential limitations in statistical power due to our sample size. Further discussions of this observation can be found in the discussion section.

We did not detect any genome-wide significant QTLs for photoperiodism when comparing facultative LD individuals to obligate LD individuals within the BEL F2 population. Considering the substantial variation in the F2 mapping population, as described in the previous

section, we suspected that the binary photoperiodism groups might not adequately capture the diversity of photoperiodic responses. With the right-skewed patterns observed in day-to-flower and the node of the first flower bud (Fig. 2.5), it is plausible that the facultative LD group designated in our study actually encompassed multiple subgroups, each with different flowering-inducing mechanisms in SD. Grouping them together would potentially obscure the genetic associations between QTLs and phenotypic variations. Hence, we narrowed our focus to the facultative LD individuals that exhibited early flowering in SD, characterized by day-to-flower no larger than 11 weeks. This threshold was chosen based on the peak of the day-to-flower distribution shown in Fig. 2.5B.

This approach led to the identification of three QTLs (Fig. 2.7B), among which two QTLs on chromosomes 7 and 10 overlapped with those associated with flowering time. Given that this analysis examined QTLs associated with differences between obligate LD individuals and early-flowering facultative LD individuals, it's challenging to untangle the confounding factors of photoperiodism group differences and the early flowering phenotype. Therefore, regarding the QTLs identified on chromosomes 7 and 10, it remained uncertain whether they contributed to both flowering time and photoperiodism group determination, or if they merely manifested specific allele frequencies within the early flowering individuals. Nonetheless, we did detect a QTL on chromosome 14 that was not observed in the flowering time QTL mapping, indicating the existence of the photoperiodism-specific QTL in the context of early flowering facultative LD individuals.

To further investigate the photoperiod-specific QTL, we carried out a parallel analysis comparing obligate LD individuals to late-flowering facultative LD individuals (day-to-flowering > 11 weeks). Despite not discovering any genome-wide significant QTLs in this scenario, we found that the previously mentioned QTL on chromosome 14 still demonstrated a significant association ($p = 0.016$) with this comparison based on the individual marker analysis.

Furthermore, this particular QTL displayed a significant effect when investigated in the context of the comparison between all facultative LD individuals and obligate LD individuals (p = 0.001). In contrast, this QTL showed no significant association with either day-to-flowering or the node of the first flower bud across all the facultative LD individuals (Fig. S2.5). Collectively, our results suggest the existence of a QTL specifically influencing photoperiodism but not flowering time in the BEL population, thereby providing support to the presence of switch-like genes.

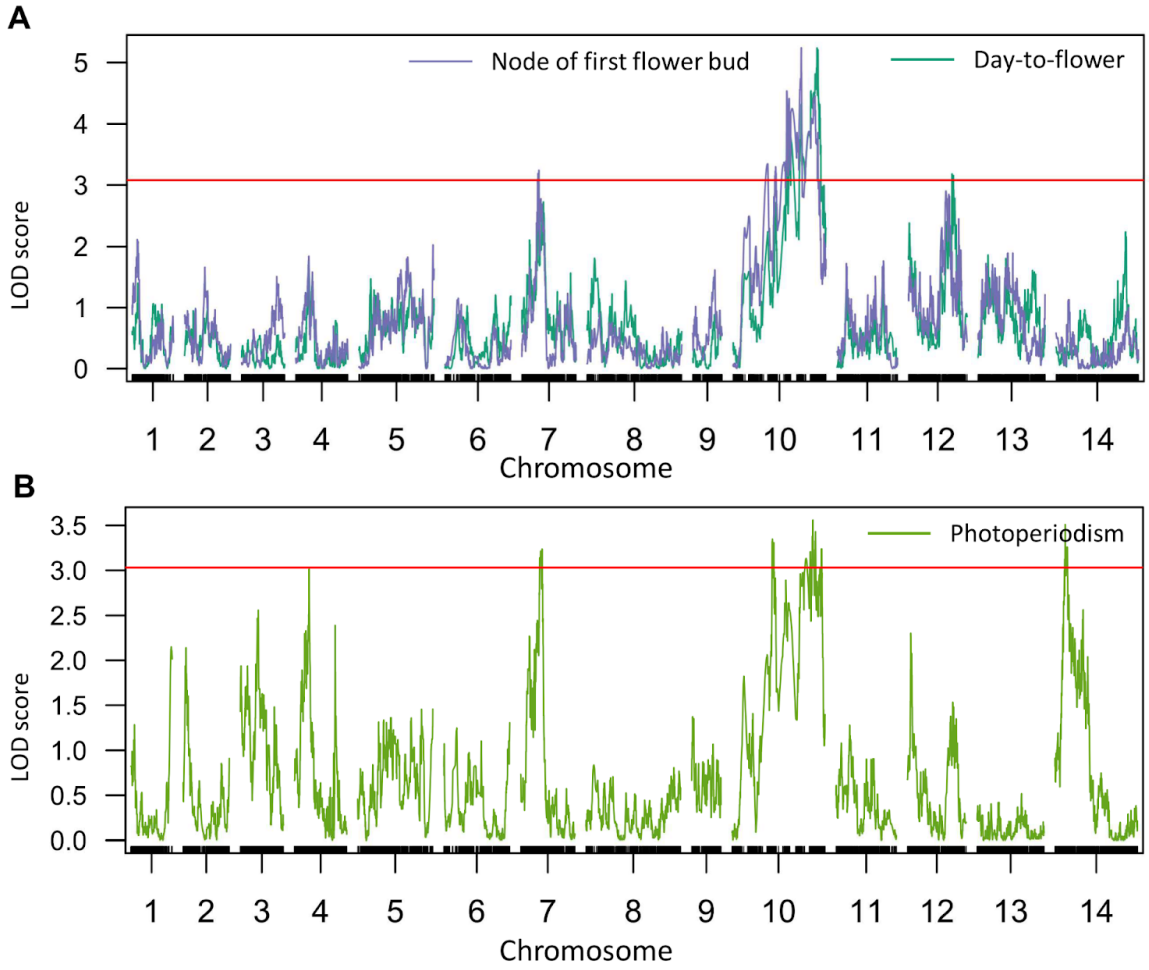

Fig. 2.7: QTL analysis of the BEL mapping population. Panel (A) displays the genome-wide LOD score plots for day-to-flower (dark green curve) and the node of the first flower bud (purple curve). No genome-wide significant LOD peak was observed when comparing obligate and facultative LD individuals directly. However, a specific LOD peak (B) associated with photoperiodism was identified on chromosome 14 when comparing the obligate LD individuals with those flowered early in SD (day-to-flower < 11

weeks). The horizontal red line represents the genome-wide significance level (α = 0.1) with 1000 permutations.

For the MAC mapping population, we observed a single QTL on chromosome 4 that accounted for 13.0% of the variance in day-to-flower (Fig. 2.8A). However, no genome-wide significant QTLs were found for the node of the first flower bud. Similar to the BEL mapping population, no genome-wide significant QTLs were identified when comparing facultative LD individuals to obligate LD individuals within the MAC F2 population. Consequently, we adopted the aforementioned approach, focusing on the comparison between obligate LD individuals and early-flowering facultative LD individuals. In this case, the criterion for early flowering was a day-to-flower not exceeding 12 weeks, corresponding to the peak of the distribution shown in Fig. 2.5D. Our investigation unveiled two QTLs on chromosomes 3 and 5, neither of which overlapped with the QTL linked to flowering time (Fig. 2.8B). While no genome-wide significant QTLs emerged in the comparison between obligate LD individuals and late-flowering facultative LD individuals (day-to-flower > 12 weeks), the QTL on chromosome 5 exhibited a significant association (p = 0.032) with this comparison based on the individual marker analysis. Intriguingly, regarding the QTL on chromosome 3, we did not detect evidence linked to this comparison (p = 0.581), possibly due to sample size constraints or its exclusive presence within the early-flowering facultative LD subgroup. Both of these particular QTLs also demonstrated significant effects when examining the comparison between all facultative LD individuals and obligate LD individuals using individual marker analysis (p = 0.043 and 0.005 for the QTL on chromosomes 3 and 5, respectively). In contrast, the analysis of the QTL effect for these two QTLs indicated no significant associations with day-to-flower or the node of the first flower across all the facultative LD individuals (Fig. S2.6 and S2.7), implying their specificity to photoperiodism.

Collectively, the findings in both the BEL and MAC mapping populations suggest the presence of photoperiodism-specific QTLs, supporting the idea that the short-day flowering phenotype of facultative LD individuals is regulated by both switch-like and dial-like genes. Interestingly, we noted different photoperiodism QTLs between the BEL and MAC populations. One possibility for this phenomenon is that all the identified photoperiodism QTLs contribute to this trait, but different loci segregate or possess polymorphisms between the BEL and MAC populations. Alternatively, it's possible that different accessions achieve the same photoperiodism phenotype through diverse genetic mechanisms.

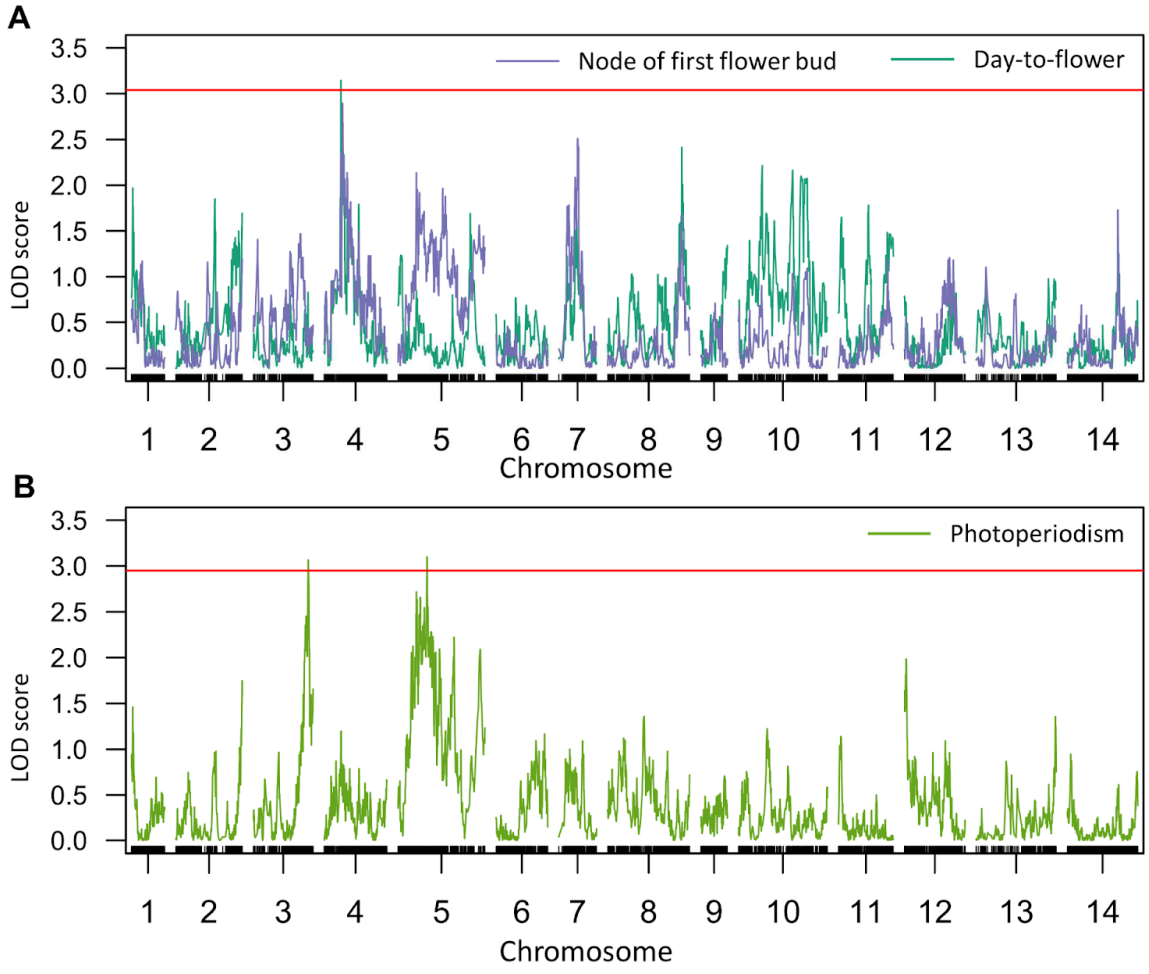

Fig. 2.8: QTL analysis of the MAC mapping population. Panel (A) displays the genome-wide LOD score plots for day-to-flower (dark green curve) and the node of the first flower bud (purple curve). Similar to the

results of BEL, no significant genome-wide LOD peak was observed when comparing obligate and facultative LD individuals directly. However, two specific LOD peaks (B) associated with photoperiodism were identified on chromosomes 3 and 5 when comparing the obligate LD individuals with those flowered early in SD (day-to-flower < 12 weeks). The horizontal red line represents the genome-wide significance level (α = 0.1) with 1000 permutations.

Table 1: QTL positions, peak positions, 1.5-LOD score intervals in both cM and Mb, percentage of the variance explained by each QTL, additive effects (a) with corresponding standard errors (SE), dominance effects (d) with corresponding SE, LOD scores, and significance levels are presented for each trait in each accession. The photoperiodism QTLs result from the comparison between obligate LD individuals and those flowered early in SD.

Accession	Trait	Chromosome	Peak position (cM)	1.5 LOD interval (cM)	1.5 LOD interval (Mb)	Percentage of variance explained	a(SE)	d(SE)	LOD	p-value
BEL	Day-to-flower	10	191.5	122.00-199.09	1.92-2.65; 5.95-17.13	14.029	-12.91 (2.52)	-5.53 (3.61)	5.24	<0.0001
BEL	Day-to-flower	12	99.4	0.00-111.93	0.03-24.34	8.583	0.67 (2.69)	-14.53 (3.57)	3.18	0.08
BEL	Node of first flower bud	7	38.7	31.06-51.85	2.21-4.71	7.41	-1.42 (0.48)	-1.46 (0.65)	3.24	0.083
BEL	Node of first flower bud	10	155.7	121.92-193.85	1.92-2.65; 5.95-15.99	12.64	-1.97 (0.45)	-1.79 (0.65)	5.24	0.004
BEL	Photoperiodism	7	48.7	16.06-51.85	1.21-4.71	12.79	1.45 (0.51)	1.01 (0.59)	3.24	0.081
BEL	Photoperiodism	10	184.4	79.53-206.67	1.92-2.65; 3.72-18.69	8.63	1.37 (0.41)	0.43 (0.55)	3.56	0.046
BEL	Photoperiodism	14	23.1	17.00-74.78	1.02-8.42	7.98	1.19 (0.39)	0.50 (0.54)	3.51	0.048
MAC	Day-to-flower	4	37.9	35.78-78.65	2.11-5.47	13	11.77 (3.31)	-12.31 (4.40)	3.14	0.08
MAC	Photoperiodism	3	123.9	112.28-134.67	16.55-17.98	9.34	-1.02 (0.32)	0.62 (0.43)	3.07	0.09
MAC	Photoperiodism	5	65.1	28.72-187.39	1.55-24.05	9.46	-1.29 (0.39)	-0.32 (0.46)	3.1	0.09

2.3.4 No common gene expression responses to daylength are apparent between facultative LD and obligate LD lines across different Mimulus accessions

Although distinct QTLs are responsible for photoperiodism variation in the BEL and MAC mapping populations, it is possible that the effects of these QTLs converge on shared

downstream pathways, resulting in similar photoperiodic responses. If this is the case, we would expect common gene expression differences between facultative and obligate LD lines across accessions. To test this hypothesis, we conducted RNAseq on the leaf samples of seven obligate LD lines and seven facultative LD lines (Fig. S2.8) at three different time points: dawn, noon, and dusk. Initially, all plants were grown in SD until they reached a minimum of four pairs of true leaves. This strategy was adopted to postpone flowering time in LDs, as plants in this condition would otherwise commit to flowering very early. Subsequently, half of the plants were transferred to the LD condition, while the other half continued to grow in SD. We harvested the 6th or 7th fully expanded true leaf under both conditions to ensure a comparable developmental stage among the samples. For each combination of daylength and time point, three biological replicates from different plants were collected for each line.

To analyze transcriptomic patterns, we conducted separate analyses for each time point. First, to identify genes that exhibited different responses to daylengths between facultative and obligate LD photoperiodism, we performed a 2-way ANOVA on each gene to test for a significant interaction between photoperiodism and daylength. Second, since facultative and obligate LD lines primarily differed phenotypically in short days, we screened for genes differentially expressed between these two groups specifically in SD (i.e. regardless of any interaction with daylength).

This combined analysis did not reveal any genes that reached genome-wide significance. One potential reason is the presence of high variation across different Mimulus accessions (as mentioned earlier in the first section of the results), and the limited sample size of only seven lines per photoperiodism group may result in insufficient statistical power to identify candidate genes. Another plausible explanation is the absence of a universally shared pathway among accessions within the same photoperiodism group, despite exhibiting the same

ultimate photoperiodism phenotype. Under this hypothesis, different accessions may recruit distinct sets of genes to achieve their photoperiodism phenotypes.

2.3.5 A very limited number of genes related to the transcriptomic differences in photoperiodism are shared among the BEL, MAC, and TER accessions

To investigate whether different accessions utilized distinct transcriptional regulations to achieve similar photoperiodism phenotypes, we subsetted our RNAseq data and focused on gene expression variation within the BEL, MAC, and TER accessions separately, each of which contained both facultative and obligate lines. In the BEL accession, we examined three facultative LD lines (BEL2.1, BEL4, and BEL15) and a single obligate LD line (BEL6.1). Similarly, the MAC and TER accessions consisted of two facultative LD lines each (MAC1.2 and MAC2.1, and TER1.1 and TER1.3, respectively) along with a single obligate LD line (MAC7 and TER1.10). Given the lack of replicates for obligate LD photoperiodism within each accession, we applied filtering criteria to identify candidate genes. Similar to our previous approach, we performed separate analyses for the data collected at different time points (dawn, noon, and dusk) and employed two strategies to identify preliminary candidate genes: (1) conducting a 2-way ANOVA analysis to detect genes displaying line-by-daylength interactions and (2) examining differential expression specifically in SD across all lines. Subsequently, we filtered the genes based on their expression patterns, with a particular focus on identifying genes that exhibited similar expression patterns among all facultative lines and distinct expression patterns in the obligate line compared to all facultative lines. Table 2 and Table 3 provide the numbers of the filtered candidate genes obtained from the 2-way ANOVA analysis and the differential expression analysis in SD, respectively.

Table 2: Numbers of filtered candidate genes identified by 2-way ANOVA detecting line-by-daylength interactions within each accession at three distinct time points.

	BEL accession	MAC accession	TER accession
Dawn	128	430	10
Noon	9	145	0
Dusk	2	1111	172

Table 3: Numbers of filtered candidate genes identified by differential expression analysis in SD within each accession at three distinct time points.

	BEL accession	MAC accession	TER accession
Dawn	771	1676	922
Noon	1040	1991	999
Dusk	560	1735	1170

Next, we examined the presence of shared candidate genes among the BEL, MAC, and TER accessions. While no genes were found to be shared among all three accessions based on the 2-way ANOVA analysis, the differential expression analysis in SD revealed 51 candidate genes that were shared. These 51 genes were further filtered using Gene Ontology (GO) terms, leading to the identification of a single gene, Migut.I00979, which is annotated with a GO term associated with flowering regulation. Migut.I00979 encodes a basic helix-loop-helix protein, and its homolog in rice was shown to be involved in the photoperiodic pathway of flowering (see the discussion section for more information on this gene). The expression levels of Migut.I00979 were found to be generally higher in obligate LD lines compared to facultative LD lines under the SD condition across all three accessions, and this pattern is specifically observed at dusk (Fig. 2.9). Collectively, our analysis revealed a very limited number of shared genes among the

BEL, MAC, and TER accessions, even considering the higher false positive rates due to the lack of replicates for obligate LD photoperiodism. Thus, our results suggest that those accessions may achieve similar photoperiodism phenotypes through distinct underlying molecular mechanisms.

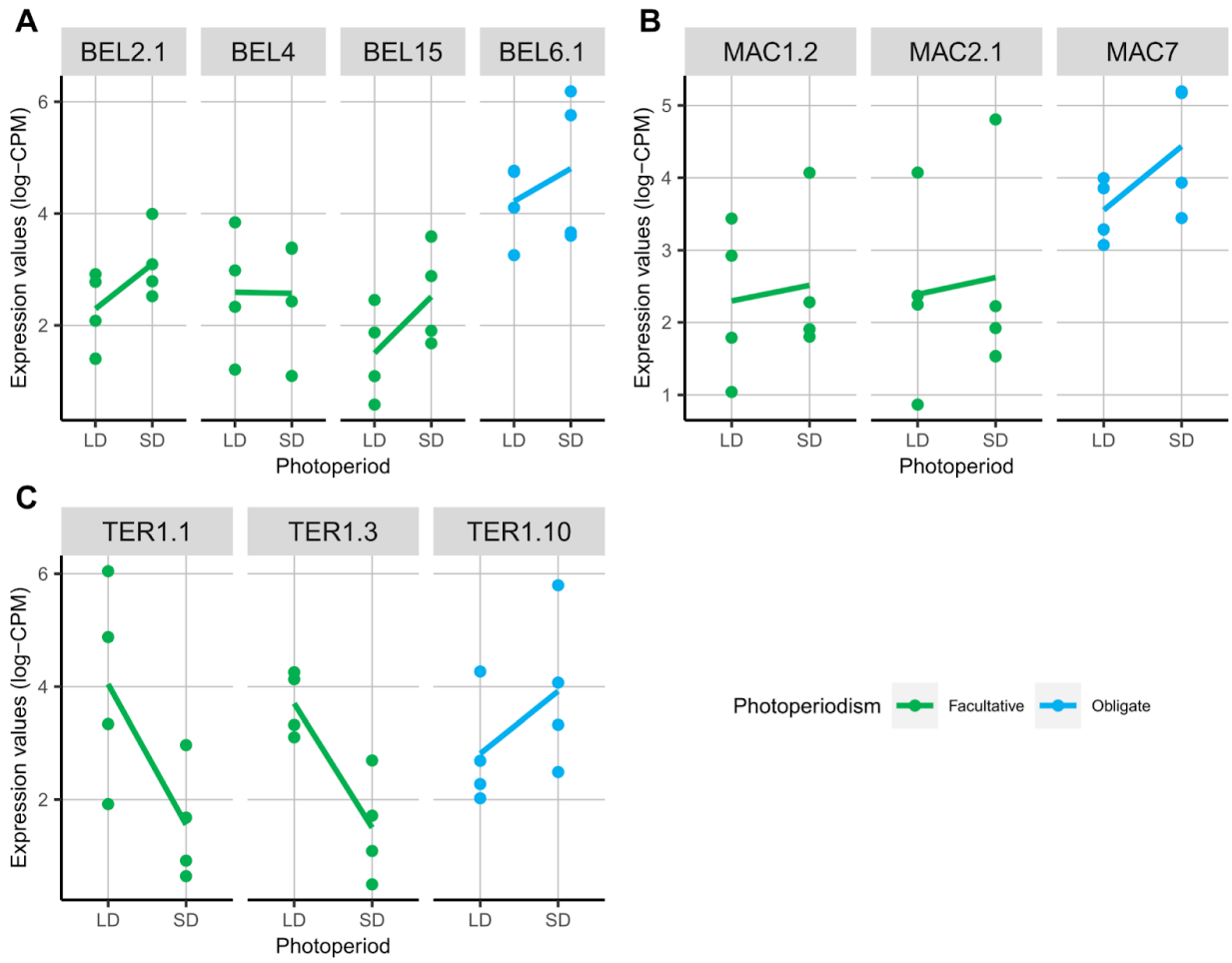

Fig. 2.9: Expression patterns of Migut.I00979 in BEL (A), MAC (B), and TER (C) accessions at dusk. In SD, this gene exhibits generally higher expression levels in obligate LD lines compared to the corresponding facultative LD lines. Each sub-panel represents the expression levels in a single line in either LD or SD. Facultative LD lines are represented by green, while obligate LD lines are represented by blue.

2.3.6 Integration of QTL and RNAseq results identifies multiple candidate genes related to flowering regulation in BEL and MAC accessions

To integrate the information from QTL mapping and RNAseq analyses in the BEL and MAC accessions, we examined whether any candidate genes identified in the RNAseq data were located within the 1.5-LOD intervals of the QTLs specifically related to photoperiodism (i.e., the QTL on chromosome 14 for BEL and the QTLs on chromosomes 3 and 5 for MAC). Subsequently, we applied a filtering process to these candidate genes using GO terms associated with flowering regulation.

In the BEL accession, none of the genes identified from the RNAseq results using the 2-way ANOVA met the criteria of being located within the photoperiodism QTL and being associated with the GO terms of interest. However, two flowering-related genes, Migut.N00999 and Migut.N00245, exhibited differential expressions between two photoperiodism groups in BEL under the SD condition and were located within the photoperiodism QTL interval (Table 4).

Table 4: Flowering-related candidate genes identified by differential expression analysis in SD and located within the photoperiodism-specific QTL on chromosome 14 of the BEL accession.

GeneID	Locus	Arabidopsis homolog PLAZA	Description of Arabidopsis homolog
Migut.N00999	scaffold_14 : 7174776-7190149 : negative	AT1G33410	SUPPRESSOR OF AUXIN RESISTANCE1 (SAR1)
Migut.N00245	scaffold_14 : 1141380-1142384 : negative	AT3G12145	FLOR1, Leucine-rich repeat family protein

On the other hand, in the MAC accession, we identified several candidate genes related to flowering regulation that were located within the photoperiodism-specific QTLs (chromosomes 3 and 5). Four genes, Migut.C01074, Migut.E00408, Migut.E01232, and Migut.E01656, exhibited line-by-photoperiod interactions (Table 5). Additionally, two genes, Migut.E01551 and Migut.E01667, showed differential expressions between obligate and facultative LD

photoperiodisms in SD (Table 6). Further discussion regarding the candidate genes listed above is provided in the discussion section.

Table 5: Flowering-related candidate genes identified by 2-way ANOVA detecting line-by-daylength interactions and located within the photoperiodism-specific QTLs on chromosomes 3 and 5 of the MAC accession.

GeneID	Locus	Arabidopsis homolog PLAZA	Description of Arabidopsis homolog
Migut.C01074	scaffold_3 : 17286028-17289219 : negative	AT1G77250	RING/FYVE/PHD-type zinc finger family protein
Migut.E01232	scaffold_5 : 15410864-15416574 : negative	AT2G46340	SUPPRESSOR OF PHYA-105 1 (SPA1)
Migut.E01656	scaffold_5 : 20303677-20307235 : positive	AT2G35720	ORIENTATION UNDER VERY LOW FLUENCES OF LIGHT 1 (OWL1)
Migut.E00408	scaffold_5 : 2213992-2214852 : negative	AT2G01290	ribose-5-phosphate isomerase 2 (RPI2)

Table 6: Flowering-related candidate genes identified by differential expression analysis in SD and located within the photoperiodism-specific QTLs on chromosomes 3 and 5 of the MAC accession.

GeneID	Locus	Arabidopsis homolog PLAZA	Description of Arabidopsis homolog
Migut.E01667	scaffold_5 : 20788017-20792032 : negative	AT1G12110	NITRATE TRANSPORTER 1.1 (NRT1.1)
Migut.E01551	scaffold_5 : 18092042-18093704 : negative	AT2G25930	EARLY FLOWERING 3 (ELF3)

2.4 Discussion

Photoperiod is a crucial external signal that plays a pivotal role in the adaptation of plants to their local environment. Consequently, understanding how plants respond to photoperiods has received considerable attention in evolutionary and ecological research. In agricultural contexts, the artificial adaptation of crops to broad geographic ranges is a key step during crop domestication. The extensive manipulation of photoperiodic responses in crops has resulted in a diverse array of varieties that can be successfully cultivated at latitudes beyond the range occupied by their wild progenitors (Nakamichi, 2015; Hill & Li, 2016; Brambilla *et al.*, 2017). While traditional breeding relies on natural genetic variation, gaining a deeper understanding of the molecular mechanisms underlying a wide range of photoperiodic responses is essential for efficient crop improvement. However, most of our current understanding of photoperiodic responses is limited to plants with facultative photoperiodism. In this study, we aimed to further our understanding of photoperiodism by studying the molecular mechanism underlying obligate photoperiodism using Mimulus as a case study. Our research focused on addressing two key questions: (1) What are the mechanisms that lead to the stringent photoperiodic responses observed in obligate LDPs? and (2) Do different accessions exhibiting similar photoperiodic responses share the same underlying mechanisms?

To address our first aim, we proposed and tested the "switch-vs-dial" hypothesis (Fig. 2.1). In Model 2, dial-like genes contribute to both the regulation of flowering time and determining whether a plant can flower under certain photoperiod conditions. In contrast, Model 1 suggests that switch-like genes specifically dictate whether a plant can flower or not, while dial-like genes modulate the timing of flowering. Through our QTL mapping using the BEL and MAC populations, we identified several photoperiodism-specific QTLs, and those QTLs did not exhibit significant effects on the flowering time phenotypes. These findings provide evidence for the existence of switch-like genes in Mimulus. To gain deeper insights into the interplay between

dial-like and switch-like genes in regulating short-day flowering phenotypes, further experiments are required. For instance, exploring whether switch-like genes directly activate dial-like gene expression or exert epistatic effects on dial-like genes would shed light on the underlying molecular mechanisms of obligate-type photoperiodism.

Interestingly, our investigation uncovered distinct photoperiodism-specific QTLs within the BEL and MAC populations (Fig. 2.7 and Fig. 2.8). One possible explanation for this discovery is that similar photoperiodic responses across different accessions could be controlled by different genetic mechanisms. Our RNAseq analysis supports this hypothesis because of a lack of consistent transcriptional responses to photoperiod between facultative LD lines and obligate LD lines in Mimulus. Finally, the integration of QTL and RNAseq results allowed us to identify several candidate genes related to flowering regulation, which could be the focus of future studies. Thus, in the first part of the discussion, we provide some information about those candidate genes based on the functions of their homologs in other species.

We also recognized some potential challenges and caveats associated with the intrinsic characteristics of our materials and the considerations in experimental design. Therefore, we discuss these points in detail in the remaining part of our discussion. We believe that these insights and experiences can be valuable for future research on organisms like Mimulus, which receives relatively less attention compared to well-established model organisms.

2.4.1 Candidate genes related to photoperiodism in this study

In our RNAseq analysis (Section 2.3.5 of the results), we explored the presence of shared candidate genes among the BEL, MAC, and TER accessions that were related to the differences in photoperiodism. We identified a single candidate gene, Migut.I00979, which exhibited higher expression in the obligate LD lines specifically at dusk under the SD condition. Migut.I00979 encodes a protein belonging to the basic helix-loop-helix DNA-binding superfamily,

with its predicted ortholog in Arabidopsis being AT3G21330. The function of AT3G21330 remains largely unknown due to the lack of study. However, its rice homolog, *OsLF* (Os05g0541400), is implicated as a flowering repressor under both long-day and short-day conditions. *OsLF* is proposed to interact with components of light signaling and the circadian clock, ultimately influencing flowering-regulating genes including *OsGI* and *Hd1* (Zhao *et al.*, 2011). The higher expression of Migut.I00979 in the obligate LD lines in SD is consistent with its putative repressor function. Considering its potential role in transcriptional regulation, we expected to identify other downstream flowering-related genes in our RNAseq analyses. However, we did not detect any other candidate genes shared by these three accessions. Therefore, further molecular experiments are necessary to confirm the role of Migut.I00979 in the photoperiodism of Mimulus.

In BEL, we identified two flowering-related genes located within the photoperiodism-specific QTL interval on chromosome 14: Migut.N00999 and Migut.N00245. The homolog of Migut.N00999 in Arabidopsis is *Nucleoporin 160* (*Nup160*), which is a nucleoprotein that forms nuclear pore complexes involved in nucleocytoplasmic transport. In a recent article (Li *et al.*, 2020), *Nup160* was demonstrated to mediate flowering regulation through the diurnal abundance of the CO protein by ensuring the proper localization of an E3-ubiquitin ligase called HIGH EXPRESSION OF OSMOTICALLY RESPONSIVE GENES1 (HOS1). Mutations in *Nup160* result in an additional morning peak of *FT* expression and accelerated flowering. On the other hand, the homolog of Migut.N00245 in Arabidopsis is *FLOR1*, which encodes a leucine-rich protein. *FLOR1* has been reported to interact with the floral organ identity gene *AGAMOUS* (Gamboa *et al.*, 2001). Mutations in *FLOR1* result in a slight delay in flowering under long-day conditions (Torti *et al.*, 2012).

In MAC, we identified several genes located within the QTL interval that are potentially involved in flowering regulation through different pathways. Migut.E01551 and Migut.E01232

encode homologs of Arabidopsis *EARLY FLOWERING 3* (*ELF3*) and *SUPPRESSOR OF PHYA-105 1* (*SPA1*), respectively, suggesting their role in circadian rhythm and photoperiodic regulation of flowering (Zagotta *et al.*, 1996; Hicks *et al.*, 2001; Laubinger *et al.*, 2006). We also found genes involved in flowering regulation through carbohydrate metabolism and nitrate transportation, such as Migut.E00408 (homologous to Arabidopsis *ribose-5-phosphate isomerase 2*, *RPI2*) (Xiong *et al.*, 2009) and Migut.E01667 (homologous to Arabidopsis *NITRATE TRANSPORTER 1.1*, *NRT1.1*) (Teng *et al.*, 2019). Migut.E01656 shares homology with *ORIENTATION UNDER VERY LOW FLUENCES OF LIGHT 1* (*OWL1*) in Arabidopsis, which is known to sense light signals under very low light conditions (Kneissl *et al.*, 2009). While the function of the homolog of Migut.C01074 in Arabidopsis remains unclear, its homolog in rice, *Early heading date 3* (*Ehd3*), has been associated with long-day flowering through the Ghd7-Ehd1 pathway. Interestingly, the *ehd3* mutant failed to activate *RFT1* in long days, leading to a phenotype of obligate SD photoperiodism (Matsubara *et al.*, 2011).

2.4.2 Challenges and caveats in this study

(A) QTL mapping

To detect the QTL associated with photoperiodism, we categorized the F2 individuals in the mapping populations as either facultative LDP or obligate LDP based on their ability to flower in SD. However, we observed a wide range of diversity in morphology (Fig. 2.4 and Fig. S2.4) and flowering time among the F2 individuals (Fig. 2.5). In addition, when comparing these binary groups directly, we did not detect any genome-wide significant QTLs. In contrast, we found QTLs related to photoperiodism by comparing the obligate LD individuals with early-flowering facultative LD individuals and then tested their QTL effects based on individual marker analyses to determine if they were specifically linked to photoperiodism.

These observations raise doubts about the effectiveness of binary categorization in accurately capturing the variations present in our F2 populations for QTL mapping. It is plausible that photoperiodic responses are influenced by multiple switch genes. Thus, if individuals flowering early in SD tend to possess a certain combination of alleles of these genes compared to those flowering late, grouping them together under a single photoperiodism category may hinder the detection of individual switch-like genes involved. Similarly, inadequate categorization could occur in individuals assigned to obligate LDP if they are unable to flower in SD due to different combinations of causal genes, thereby compromising the accuracy of QTL mapping. The challenge of properly categorizing similar flowering phenotypes is not uncommon, as these phenotypes are typically modulated by multiple genes or pathways and exhibit high flexibility. For example, in the case of rice, it has been reported that lines showing similar photoperiod sensitivity can actually harbor different combinations of underlying flowering gene alleles (Zong *et al.*, 2021). In addition to the challenge of proper categorization, the relatively small sample size in our experiment (192 F2s sequenced for each accession) further hinders the robust detection of QTL signals, especially when we subsetted the dataset based on their flowering time in some analyses. To address these issues, a more systematic recording across different phenotypes is required for defining possible sub-categories within each photoperiodism. Additionally, larger sample sizes are crucial to ensure sufficient replicates within each sub-category, enabling more robust QTL mapping analyses.

Besides, we noticed that the QTLs associated with the two flowering time-related phenotypes were not exactly identical. In BEL, we identified QTLs on chromosomes 10 and 12 for day-to-flower, while QTLs on chromosomes 7 and 10 were found for the node of the first flower bud. On the other hand, in MAC, we only detected a QTL for day-to-flower, which was on chromosome 4, while no genome-wide significant QTLs were found for the other trait. One possibility is that these two flowering time-related phenotypes actually share the same

underlying QTLs, but due to the small sample size used in this study, we might not have detected all of the QTLs with sufficient statistical power. Consistent with this hypothesis, we observed significant effects of the day-to-flower QTL on the node of the first flower bud in both BEL and MAC ($p = 0.013$ and $p = 0.005$, respectively) when testing each individual QTL effect separately. Additionally, in BEL, the QTL for the node of the first flower bud on chromosome 7 also exhibited a significant effect ($p = 0.001$) on the day-to-flower phenotype. Another possibility is that these two flowering time-related phenotypes indeed have their own specific QTLs. As suggested by Thomas and Vince-Prue (Thomas & Vince-Prue, 1997), different flowering time proxies may hold distinct biological meanings for different species under varied experimental conditions. Further fine-mapping is required to test these two hypotheses.

(B) RNAseq

To identify candidate genes associated with photoperiodism differences within accessions, we subsetted our original RNAseq dataset and compared the transcriptomes of facultative LD lines with a single obligate LD line in the BEL, MAC, and TER accessions. While we had three biological replicates for each line at every single time point, it is important to note the limitation of this analysis due to the lack of replication of obligate LD photoperiodism. This limitation makes it challenging to distinguish between genes truly related to photoperiodism and those influenced merely by the uniqueness of the single obligate LD line. To mitigate this issue, we applied additional criteria and specifically focused on genes that showed similar expression patterns among facultative lines but differed from the obligate line. However, it is expected that this approach may include more false positives compared to the analysis with true replicates of obligate LD photoperiodism. Interestingly, even with this lenient approach, we only detected a single gene, Migut.I00979, which was shared by all three accessions. This finding supports the possibility that different accessions achieve similar photoperiodic phenotypes through distinct

molecular mechanisms, and it is consistent with our QTL results and the original RNAseq analysis that directly compared seven obligate and facultative LD lines across accessions.

Furthermore, sample selection is another important factor to consider in future studies. In our current investigation, we primarily focused on the transcriptome regulations in leaves, as it has been well-established that plants perceive photoperiod signals through their leaves. However, emerging evidence suggests the involvement of other pathways and tissues in flowering regulation, particularly in plants with prolonged growth. For instance, the aging pathway in the shoot apical meristem, specifically the *SPL15* gene, has been shown to act in parallel with *FT* in inducing flowering in older plants (Hyun *et al.*, 2019). Therefore, incorporating additional pathways and tissues into future studies may provide a more comprehensive understanding of the regulatory mechanisms underlying photoperiodism.

(C) Integration of the QTL and RNAseq results

To integrate the results from QTL and RNAseq, we analyzed the candidate genes obtained from the RNAseq that were located in the QTL interval. We identified several candidate genes and discussed their possible functions in flowering regulation above.

Although the candidate genes identified in our study provide a valuable starting point for future experiments, we acknowledge that the assumption underlying our analysis—that the causal genes within the QTLs would exhibit transcriptional differences—may not hold true in all cases. In addition, it is important to recognize that the GO annotation in Mimulus is still incomplete. Most of the available evidence for Mimulus GO annotations is based on the prediction using homologous genes in other plant species, primarily Arabidopsis and rice. However, due to the intrinsic diversity and complexity of the flowering pathway, certain lineage- or species-specific genes may not be well annotated. Furthermore, recent studies have highlighted crosstalk between the classical flowering pathway and other metabolic pathways,

such as nutrition or sugar metabolism (Gendron & Staiger, 2023). This poses challenges in precisely defining "flowering-related genes" for GO filtering. Therefore, further fine mapping is necessary to identify the genetic basis of photoperiodism, and functional studies combining mutagenesis will be crucial in elucidating the detailed mechanisms involved.

2.5 Materials and Methods

Plant materials and growth conditions

The annual *Mimulus guttatus* seeds used in this study were collected by Benjamin K. Blackman's lab, and the accessions were described in Kooyers et al (Kooyers *et al.*, 2015). To initiate germination, seeds were cold stratified on a soil mixture consisting of 50% Sunshine Mix #1 potting soil and 50% vermiculite in the dark at 4°C for 7 days. Following cold stratification, seeds were transferred to Conviron E7/2 growth chambers until the second true leaves emerged (approximately 10 to 14 days). Seedlings were then transplanted into plastic rose pots measuring 5 cm square by 8 cm deep. A fresh mixture of Sunshine Mix and vermiculite was used as the growing medium. Pots were randomized within chambers and rotated three times per week. Seedlings were watered every 2 days with fertilizer water provided by the UC Davis Controlled Environment Facility. Growth chambers were maintained at a temperature of 22°C and photosynthetically active radiation of approximately 250 umol $m^{-2}s^{-1}$, with a red-to-far-red light ratio of about 1.5.

For the phenotypic screening, the plants were germinated and grown under a short day (SD) condition, with a day/night cycle of 9:15 hours. This daylength was chosen because it corresponds to the minimum daylength observed in the natural habitat of the accessions used in this study, which is not shorter than 9.5 hours. By using this SD condition, we aimed to ensure that the individuals capable of flowering under these conditions would also be capable of flowering in any daylength in their natural habitat throughout the year, assuming other environmental factors such as drought or abiotic stress are not considered. The experiment was conducted for a duration of 16 weeks to align with the typical life cycle of annual Mimulus accessions. These plants typically germinate in the spring and undergo senescence during

summer droughts. To establish stable flowering responses, the candidate lines were selfed for an additional 2 generations, reaching at least the F4 generation.

To evaluate flowering traits, including photoperiodism and flowering time, the plants were grown under either SD or long day (LD, 16:8 hours day/night cycle) conditions. The lines that could only flower in LD were classified as obligate LD lines. Among the lines capable of flowering in SD, we distinguished between day-neutral lines and facultative LD lines by comparing their flowering time in SD and LD. Day-neutral lines were expected to flower at the same time regardless of daylength, while flowering in facultative LD lines was anticipated to be accelerated in LD. To estimate flowering time, we recorded day-to-flower, as well as the node of the first flower bud and the node of the first fully-opened flower. The experiment concluded at the end of the 16th week, as described previously.

For RNAseq analysis, we took into consideration the high sensitivity of the transcriptome to plant age and the timing of harvest. To ensure accurate sampling, we established two criteria: (1) We harvested the 6th or 7th true leaves (excluding leaves on reproductive nodes) to ensure the comparisons were made across adult leaves. (2) We collected samples at three time points: Dawn (ZT0 for both SD and LD), Noon (ZT4.5 for SD and ZT8 for LD), and Dusk (ZT9 for SD and ZT16 for LD). To meet these criteria, all plants designated for RNAseq analysis were germinated and grown under the SD condition until the emergence of the 4th true leaves. This approach was adopted to delay flowering in the LD treatment, as plants germinating directly under LD conditions typically flowered before reaching the 6th node. Subsequently, the seedlings were randomly assigned to either SD or LD chambers, and when available, the 6th or 7th fully extended true leaves were collected for RNA extraction. For each line, we harvested three leaves from three different biological replicate plants.

For QTL mapping, the BEL and MAC F2 mapping populations were grown under SD conditions, as described above. The experiment duration was extended by 6 weeks (a total of 22 weeks) to identify plants that did not flower more accurately. For the plants that flowered by the 22nd week, we recorded the days to flowering, the node of the first flower bud, and the node of the first flower.

All leaf samples harvested during our experiments were promptly submerged in liquid nitrogen to rapidly halt any biological processes. The samples were then stored at -80°C to maintain their integrity until RNA or DNA extraction was performed.

Generating mapping populations and conducting QTL mapping analysis

We constructed two F2 mapping populations by crossing lines with distinct photoperiodism from the same accession for QTL mapping of the obligate photoperiodism trait. We selected BEL4 (facultative LD line) and BEL4.2 (obligate LD line) for the BEL accession and MAC1.2 (facultative LD line) and MAC7 (obligate LD line) for the MAC accession. The F1 generations were generated by crossing one individual from each line, followed by selfing of a single F1 individual to produce recombinant F2 mapping progenies.

To obtain genomic data for our mapping populations, we employed a whole-genome resequencing approach. Total genomic DNA was isolated from 192 F2 samples and both parents in each mapping population using a high-throughput column-based method with slight modifications, as described by Anderson et al. (Anderson *et al.*, 2018). DNA integrity was assessed using agarose electrophoresis, and the concentration was determined using Quant-iT™ PicoGreen™ dsDNA Assay Kits (ThermoFisher, Cat. # P7589). High-quality DNA samples were sent to the DNA Technologies & Expression Analysis Core at UC Davis for library

preparation using the SeqWell DNA Library kit. Pooled libraries were sequenced on an Illumina HiSeq X sequencer, generating 150-bp paired-end reads.

To obtain SNP markers for subsequent linkage map construction and QTL mapping, we conducted variant calling and marker selection. The sequencing results were first subjected to quality analysis using FastQC (Andrews, 2010) and trimmed using HTStream (Angell & S4HTS, 2017). Subsequently, the reads were aligned to the *Mimulus guttatus* reference genome (*Mimulus guttatus* v2.0 from Phytozome v13 (Goodstein *et al.*, 2012)) using the BWA-MEM alignment tool (Li & Durbin, 2009). Next, we called variants for the parental lines using GATK4 based on its best practices workflows (Poplin *et al.*, 2018) and filtered the resulting SNPs using bcftools (Li, 2011) with setting `-g ^miss` and `-e 'N_PASS(GT="AA")>1'`. After obtaining high-confidence SNP markers, we performed variant calling for the F2 individuals at these marker loci using bcftools mpileup. The VCF files generated from each individual were then merged, and an additional round of marker filtering was carried out to remove markers with unusual allele frequencies. This filtering was carried out by vcftools with setting MAF=0.35 and QUAL=50. Due to potential issues such as low coverage and noise, we developed a custom R script to re-score genotypes for fixed windows of 1 Kb. Subsequently, the TIGER script (Rowan *et al.*, 2015) was utilized for genotype imputation. The procedures are summarized in Fig. S2.9.

Linkage maps construction and subsequent QTL mapping were performed using R/qtl (Broman *et al.*, 2003). Linkage groups and the order of SNP markers were determined based on the 14 chromosomes defined in the reference genome. The linkage maps were then assessed by plots of the pairwise recombination fractions and LOD scores. We identified a clear inversion on chromosome 10. To identify possible breakpoints of the inversion fragments, we developed a custom R script that employed a sliding window strategy and calculated the correlation of allele frequencies between marker blocks. The inversion fragments were manually rearranged and the LOD scores of pairwise recombination fractions were analyzed. For QTL mapping, single-QTL

genome scans were conducted using the Haley-Knott regression method implemented with the `scanone` function. To establish genome-wide significance thresholds, 1000 permutations were conducted for each phenotype individually, setting the significance level at $\alpha = 0.1$. The QTL intervals were estimated by calculating 1.5-LOD support intervals using the `lodint` function. Furthermore, multiple-QTL models were fitted using the `fitqtl` function to determine the additive effect, dominance deviation, and the percentage of phenotypic variance explained by each QTL among the F2 individuals. Non-additive interactions among QTLs were also examined using the `addint` function, but no significant interactions were detected.

RNAseq library preparations and data analyses

Samples were treated with liquid nitrogen and ground with two or three 2.8 mm ceramic beads using the SPEX™ SamplePrep 2010 Geno/Grinder at 1100 rpm for 2 minutes. Subsequently, mRNA was extracted from the samples using the Dynabeads™ mRNA DIRECT Kit (Invitrogen, Cat. #: 61012). The concentration of mRNA was assessed using the QuantiFluor RNA System (Promega, Cat. #: E3310) to ensure sufficient input mRNA for library preparation. RNAseq libraries were generated following the BrAD-seq protocol (Townsley *et al.*, 2015) with some modifications, and then cleaned and size-selected using homemade SPRI beads. Gel electrophoresis was performed to examine the libraries, and they were pooled together at the same concentration. The pooled libraries were then sequenced on a single lane of an Illumina HiSeq X sequencer, generating 150-bp paired-end reads.

The sequencing reads were analyzed using FastQC (Andrews, 2010) for quality assessment, and then mapped to the Mimulus reference genome using the STAR aligner (Dobin *et al.*, 2013). The resulting count tables generated by STAR were merged to create a count table, which served as the input for the limma-voom package (Law *et al.*, 2014). Data obtained from different time points were analyzed separately. The model fitting was performed following

the guidelines provided in the limma-voom tutorial (Law *et al.*, 2018). Candidate genes were identified based on a false discovery rate (FDR) threshold of 10% (i.e., with a Benjamini-Hochberg adjusted p-value below 0.1 (Benjamini & Hochberg, 1995)).

For filtering gene lists by GO terms, GO annotations for each gene were obtained from Dicots PLAZA 5.0 (Van Bel *et al.*, 2022). The predicted homologs of Mimulus genes in Arabidopsis or rice were obtained by querying the Dicots PLAZA 5.0 website.

2.6 References

Anderson CB, Franzmayr BK, Hong SW, Larking AC, van Stijn TC, Tan R, Moraga R, Faville MJ, Griffiths AG. **2018**. Protocol: a versatile, inexpensive, high-throughput plant genomic DNA extraction method suitable for genotyping-by-sequencing. *Plant Methods* **14**: 75.
Andrews S. **2010**. FASTQC. A quality control tool for high throughput sequence data.
Angell J, S4HTS. **2017**. HTStream: A high throughput sequence read toolset using a streaming approach facilitated by Linux pipes.
Ballerini E, Kramer E. **2011**. In the Light of Evolution: A Reevaluation of Conservation in the CO–FT Regulon and Its Role in Photoperiodic Regulation of Flowering Time. *Frontiers in Plant Science* **2**.
Benjamini Y, Hochberg Y. **1995**. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**: 289–300.
Brambilla V, Gomez-Ariza J, Cerise M, Fornara F. **2017**. The Importance of Being on Time: Regulatory Networks Controlling Photoperiodic Flowering in Cereals. *Frontiers in Plant Science* **8**.
Broman KW, Wu H, Sen Ś, Churchill GA. **2003**. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19**: 889–890.
Corbesier L, Vincent C, Jang S, Fornara F, Fan Q, Searle I, Giakountis A, Farrona S, Gissot L, Turnbull C, *et al.*2007. FT Protein Movement Contributes to Long-Distance Signaling in Floral Induction of Arabidopsis. *Science* **316**: 1030–1033.
Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. **2013**. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15–21.
Doi K, Izawa T, Fuse T, Yamanouchi U, Kubo T, Shimatani Z, Yano M, Yoshimura A. **2004**. Ehd1, a B-type response regulator in rice, confers short-day promotion of flowering and controls FT-like gene expression independently of Hd1. *Genes & Development* **18**: 926–936.
Friedman J, Twyford AD, Willis JH, Blackman BK. **2015**. The extent and genetic basis of phenotypic divergence in life history traits in Mimulus guttatus. *Molecular Ecology* **24**: 111–122.
Friedman J, Willis JH. **2013**. Major QTLs for critical photoperiod and vernalization underlie extensive variation in flowering in the Mimulus guttatus species complex. *New Phytologist* **199**: 571–583.
Gamboa A, Paéz-Valencia J, Acevedo GF, Vázquez-Moreno L, Alvarez-Buylla RE. **2001**. Floral Transcription Factor AGAMOUS Interacts in Vitro with a Leucine-Rich Repeat and an Acid Phosphatase Protein Complex. *Biochemical and Biophysical Research Communications* **288**: 1018–1026.
Gendron JM, Staiger D. **2023**. New Horizons in Plant Photoperiodism. *Annual Review of Plant Biology* **74**: null.
Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, *et al.*2012. Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research* **40**: D1178–D1186.
Hayama R, Yokoi S, Tamaki S, Yano M, Shimamoto K. **2003**. Adaptation of photoperiodic control pathways produces short-day flowering in rice. *Nature* **422**: 719–722.
Hicks KA, Albertson TM, Wagner DR. **2001**. EARLY FLOWERING3 Encodes a Novel Protein That Regulates Circadian Clock Function and Flowering in Arabidopsis. *The Plant Cell* **13**: 1281–1292.
Hill CB, Li C. **2016**. Genetic Architecture of Flowering Phenology in Cereals and Opportunities for Crop Improvement. *Frontiers in Plant Science* **7**.
Hyun Y, Vincent C, Tilmes V, Bergonzi S, Kiefer C, Richter R, Martinez-Gallegos R, Severing E, Coupland G. **2019**. A regulatory circuit conferring varied flowering response to

cold in annual and perennial plants. *Science* **363**: 409–412.
Itoh H, Nonoue Y, Yano M, Izawa T. **2010**. A pair of floral regulators sets critical day length for Hd3a florigen expression in rice. *Nature Genetics* **42**: 635–638.
Jaeger KE, Wigge PA. **2007**. FT Protein Acts as a Long-Range Signal in Arabidopsis. *Current Biology* **17**: 1050–1054.
Kardailsky I, Shukla VK, Ahn JH, Dagenais N, Christensen SK, Nguyen JT, Chory J, Harrison MJ, Weigel D. **1999**. Activation Tagging of the Floral Inducer FT. *Science* **286**: 1962–1965.
Kneissl J, Wachtler V, Chua N-H, Bolle C. **2009**. OWL1: An Arabidopsis J-Domain Protein Involved in Perception of Very Low Light Fluences. *The Plant Cell* **21**: 3212–3225.
Kobayashi Y, Kaya H, Goto K, Iwabuchi M, Araki T. **1999**. A Pair of Related Genes with Antagonistic Roles in Mediating Flowering Signals. *Science* **286**: 1960–1962.
Kojima S, Takahashi Y, Kobayashi Y, Monna L, Sasaki T, Araki T, Yano M. **2002**. Hd3a, a Rice Ortholog of the Arabidopsis FT Gene, Promotes Transition to Flowering Downstream of Hd1 under Short-Day Conditions. *Plant and Cell Physiology* **43**: 1096–1105.
Komiya R, Yokoi S, Shimamoto K. **2009**. A gene network for long-day flowering activates RFT1 encoding a mobile flowering signal in rice. *Development* **136**: 3443–3450.
Kooyers NJ, Greenlee AB, Colicchio JM, Oh M, Blackman BK. **2015**. Replicate altitudinal clines reveal that evolutionary flexibility underlies adaptation to drought stress in annual Mimulus guttatus. *New Phytologist* **206**: 152–165.
Laubinger S, Marchal V, Gentilhomme J, Wenkel S, Adrian J, Jang S, Kulajta C, Braun H, Coupland G, Hoecker U. **2006**. Arabidopsis SPA proteins regulate photoperiodic flowering and interact with the floral inducer CONSTANS to regulate its stability. *Development* **133**: 3213–3222.
Law CW, Alhamdoosh M, Su S, Dong X, Tian L, Smyth GK, Ritchie ME. **2018**. RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. *F1000Research* **5**.
Law CW, Chen Y, Shi W, Smyth GK. **2014**. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biology* **15**: R29.
Li H. **2011**. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**: 2987–2993.
Li H, Durbin R. **2009**. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**: 1754–1760.
Li C, Liu L, Teo ZWN, Shen L, Yu H. **2020**. Nucleoporin 160 Regulates Flowering through Anchoring HOS1 for Destabilizing CO in Arabidopsis. *Plant Communications* **1**: 100033.
Lin M-K, Belanger H, Lee Y-J, Varkonyi-Gasic E, Taoka K-I, Miura E, Xoconostle-Cázares B, Gendler K, Jorgensen RA, Phinney B, *et al.*2007. FLOWERING LOCUS T Protein May Act as the Long-Distance Florigenic Signal in the Cucurbits. *The Plant Cell* **19**: 1488–1506.
Liu C, Chen H, Er HL, Soo HM, Kumar PP, Han J-H, Liou YC, Yu H. **2008**. Direct interaction of AGL24 and SOC1 integrates flowering signals in Arabidopsis. *Development* **135**: 1481–1491.
Matsubara K, Yamanouchi U, Nonoue Y, Sugimoto K, Wang Z-X, Minobe Y, Yano M. **2011**. Ehd3, encoding a plant homeodomain finger-containing protein, is a critical promoter of rice flowering. *The Plant Journal* **66**: 603–612.
Nakamichi N. **2015**. Adaptation to the Local Environment by Modifications of the Photoperiod Response in Crops. *Plant and Cell Physiology* **56**: 594–604.
Notaguchi M, Abe M, Kimura T, Daimon Y, Kobayashi T, Yamaguchi A, Tomita Y, Dohi K, Mori M, Araki T. **2008**. Long-Distance, Graft-Transmissible Action of Arabidopsis FLOWERING LOCUS T Protein to Promote Flowering. *Plant and Cell Physiology* **49**: 1645–1658.
Ogiso-Tanaka E, Matsubara K, Yamamoto S, Nonoue Y, Wu J, Fujisawa H, Ishikubo H, Tanaka T, Ando T, Matsumoto T, *et al.*2013. Natural Variation of the RICE FLOWERING LOCUS T 1 Contributes to Flowering Time Divergence in Rice. *PLOS ONE* **8**: e75959.
Osugi A, Itoh H, Ikeda-Kawakatsu K, Takano M, Izawa T. **2011**. Molecular Dissection of the

Roles of Phytochrome in Photoperiodic Flowering in Rice. *Plant Physiology* **157**: 1128–1137.
Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, Auwera GAV der, Kling DE, Gauthier LD, Levy-Moonshine A, Roazen D, *et al.*2018. Scaling accurate genetic variant discovery to tens of thousands of samples. : 201178.
Puzey JR, Willis JH, Kelly JK. **2017**. Population structure and local selection yield high genomic variation in Mimulus guttatus. *Molecular Ecology* **26**: 519–535.
Rowan BA, Patel V, Weigel D, Schneeberger K. **2015**. Rapid and inexpensive whole-genome genotyping-by-sequencing for crossover localization and fine-scale genetic mapping. *G3 (Bethesda, Md.)* **5**: 385–398.
Ryu C-H, Lee S, Cho L-H, Kim SL, Lee Y-S, Choi SC, Jeong HJ, Yi J, Park SJ, Han C-D, *et al.*2009. OsMADS50 and OsMADS56 function antagonistically in regulating long day (LD)-dependent flowering in rice. *Plant, Cell & Environment* **32**: 1412–1427.
Samach A, Onouchi H, Gold SE, Ditta GS, Schwarz-Sommer Z, Yanofsky MF, Coupland G. **2000**. Distinct Roles of CONSTANS Target Genes in Reproductive Development of Arabidopsis. *Science* **288**: 1613–1616.
Song YH, Shim JS, Kinmonth-Schultz HA, Imaizumi T. **2015**. Photoperiodic Flowering: Time Measurement Mechanisms in Leaves. *Annual Review of Plant Biology* **66**: 441–464.
Suárez-López P, Wheatley K, Robson F, Onouchi H, Valverde F, Coupland G. **2001**. CONSTANS mediates between the circadian clock and the control of flowering in Arabidopsis. *Nature* **410**: 1116–1120.
Sweigart AL, Willis JH. **2003**. Patterns of Nucleotide Diversity in Two Species of Mimulus Are Affected by Mating System and Asymmetric Introgression. *Evolution* **57**: 2490–2506.
Tamaki S, Matsuo S, Wong HL, Yokoi S, Shimamoto K. **2007**. Hd3a Protein Is a Mobile Flowering Signal in Rice. *Science* **316**: 1033–1036.
Teng Y, Liang Y, Wang M, Mai H, Ke L. **2019**. Nitrate Transporter 1.1 is involved in regulating flowering time via transcriptional regulation of FLOWERING LOCUS C in Arabidopsis thaliana. *Plant Science* **284**: 30–36.
Thomas B, Vince-Prue D. **1997**. *Photoperiodism in plants*. San Diego: Academic Press.
Torti S, Fornara F, Vincent C, Andrés F, Nordström K, Göbel U, Knoll D, Schoof H, Coupland G. **2012**. Analysis of the Arabidopsis Shoot Meristem Transcriptome during Floral Transition Identifies Distinct Regulatory Patterns and a Leucine-Rich Repeat Protein That Promotes Flowering[C][W][OA]. *The Plant Cell* **24**: 444–462.
Townsley BT, Covington MF, Ichihashi Y, Zumstein K, Sinha NR. **2015**. BrAD-seq: Breath Adapter Directional sequencing: a streamlined, ultra-simple and fast library preparation protocol for strand specific mRNA library construction. *Frontiers in Plant Science* **6**.
Twyford AD, Friedman J. **2015**. Adaptive divergence in the monkey flower Mimulus guttatus is maintained by a chromosomal inversion. *Evolution* **69**: 1476–1486.
Van Bel M, Silvestri F, Weitz EM, Kreft L, Botzki A, Coppens F, Vandepoele K. **2022**. PLAZA 5.0: extending the scope and power of comparative and functional genomics in plants. *Nucleic Acids Research* **50**: D1468–D1474.
Wilson RN, Heckman JW, Somerville CR. **1992**. Gibberellin Is Required for Flowering in Arabidopsis thaliana under Short Days. *Plant Physiology* **100**: 403–408.
Wu CA, Lowry DB, Cooley AM, Wright KM, Lee YW, Willis JH. **2008**. *Mimulus* is an emerging model system for the integration of ecological and genomic studies. *Heredity* **100**: 220–230.
Xiong Y, DeFraia C, Williams D, Zhang X, Mou Z. **2009**. Deficiency in a cytosolic ribose-5-phosphate isomerase causes chloroplast dysfunction, late flowering and premature cell death in Arabidopsis. *Physiologia Plantarum* **137**: 249–263.
Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K, Umehara Y, Nagamura Y, *et al.*2000. Hd1, a Major Photoperiod Sensitivity Quantitative Trait Locus in Rice, Is Closely Related to the Arabidopsis Flowering Time Gene CONSTANS. *The Plant Cell* **12**: 2473–2483.

Yanovsky MJ, Kay SA. **2002**. Molecular basis of seasonal time measurement in Arabidopsis. *Nature* **419**: 308–312.
Zagotta MT, Hicks KA, Jacobs CI, Young JC, Hangarter RP, Meeks-Wagner DR. **1996**. The Arabidopsis ELF3 gene regulates vegetative photomorphogenesis and the photoperiodic induction of flowering. *The Plant Journal* **10**: 691–702.
Zhao X-L, Shi Z-Y, Peng L-T, Shen G-Z, Zhang J-L. **2011**. An atypical HLH protein OsLF in rice regulates flowering time and interacts with OsPIL13 and OsPIL15. *New Biotechnology* **28**: 788–797.
Zong W, Ren D, Huang M, Sun K, Feng J, Zhao J, Xiao D, Xie W, Liu S, Zhang H, *et al.*2021. Strong photoperiod sensitivity is controlled by cooperation and competition among Hd1, Ghd7 and DTH8 in rice heading. *New Phytologist* **229**: 1635–1649.

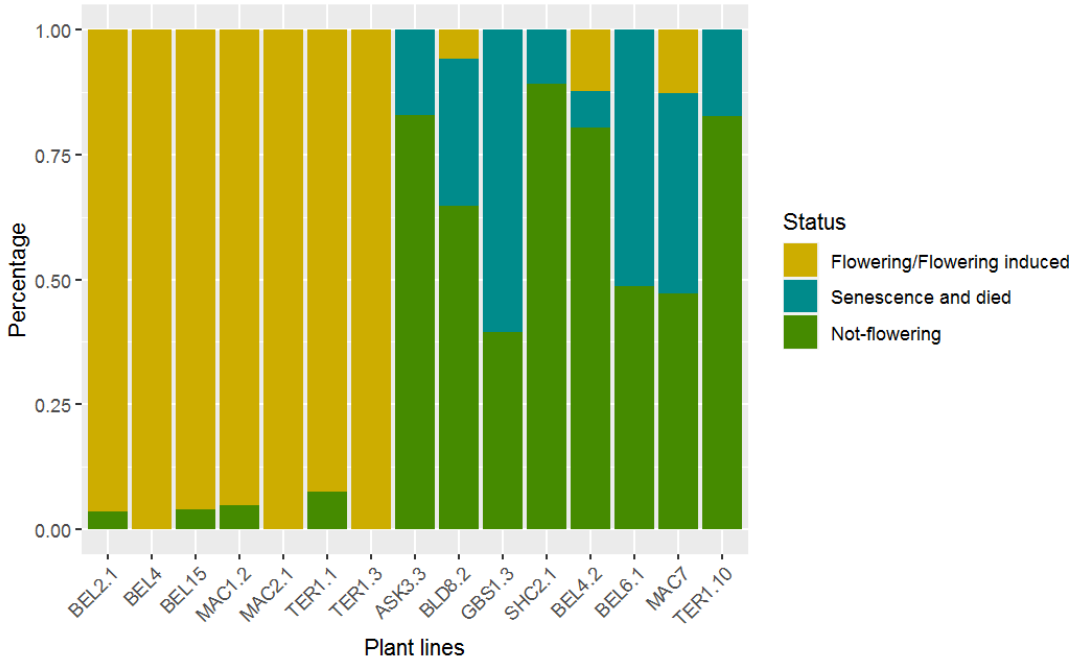

Fig. S2.1: The percentage of plants that either flowered or failed to flower in SD among the generated obligate and facultative LD lines. The yellow bars represent the proportion of plants that flowered or induced flowering in SD. The cyan bars represent the plants that senesced and died before the end of the 16-week experiment. The green bars represent the plants that did not flower but still maintained vegetative growth in SD. All lines underwent at least 4 generations of inbreeding to stabilize their flowering responses.

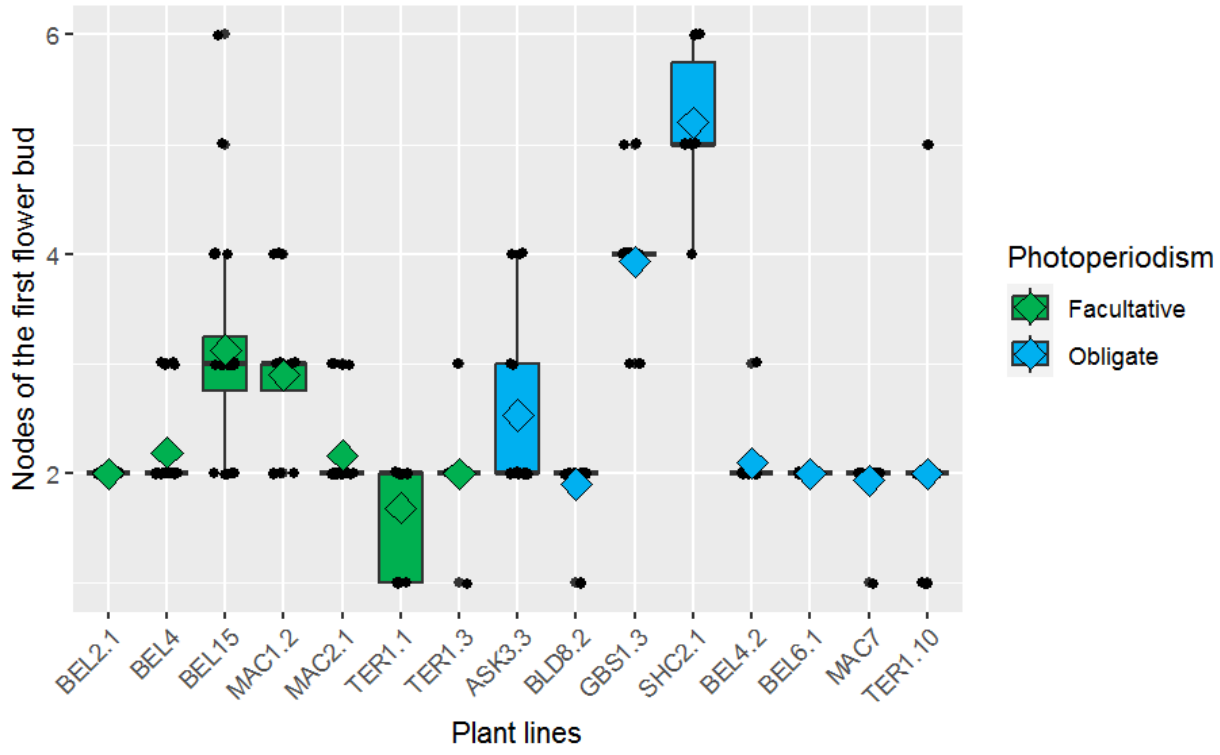

Fig. S2.2: Nodes of the first flowering bud for both facultative LD lines (green) and obligate LD lines (blue) under the LD condition. The diamonds in each box plot represent the mean of the node of the first flower bud.

Fig. S2.3: A wide variation in 7-week-old Mimulus seedlings in SD. The upper row displays obligate LD lines (from left to right): ASK3.3, BLD8.2, GBS1.3, SHC2.1, and TER1.10. The lower row exhibits facultative LD lines (from left to right): BEL2.1, BEL15, MAC1.2, MAC2.1, and TER1.1. Red bars located beside the shoot apical meristem (SAM) of each plant in the panels indicate a length of 1 cm.

Fig. S2.4: Morphological variation within the F2 mapping population of MAC. (A) An individual exhibiting a typical facultative LD phenotype. (B) Obligate LD individuals display a diverse range of phenotypes.

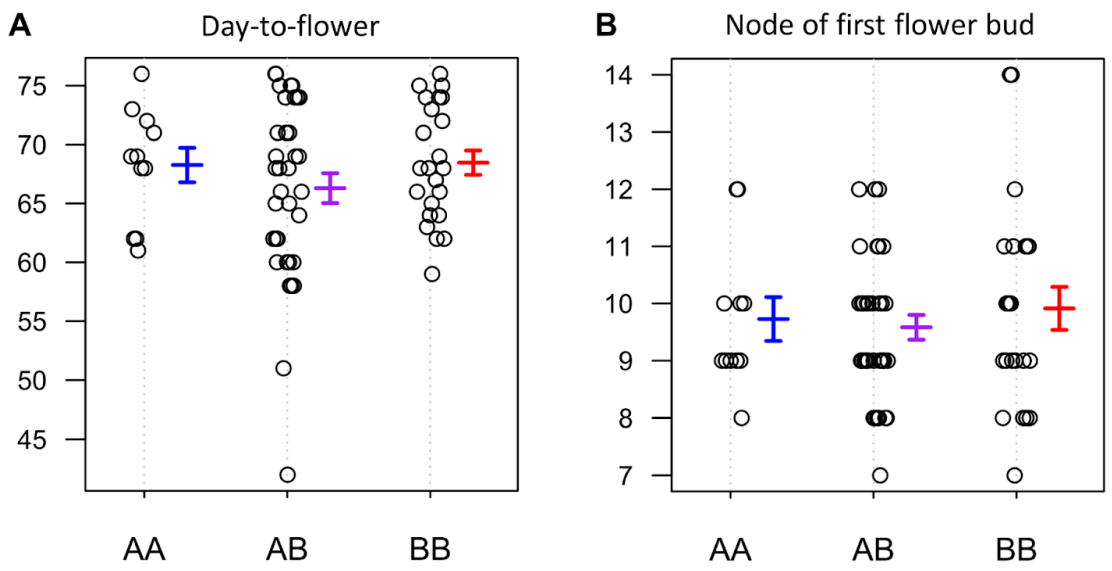

Fig. S2.5: Phenotypic effects of the photoperiodism-specific QTL on chromosome 14 in BEL on (A) day-to-flower and (B) the node of the first flower bud. “A” represents the allele from BEL4.2, an obligate LD line, while “B” represents the allele from BEL4, a facultative LD line. Error bars indicate standard errors.

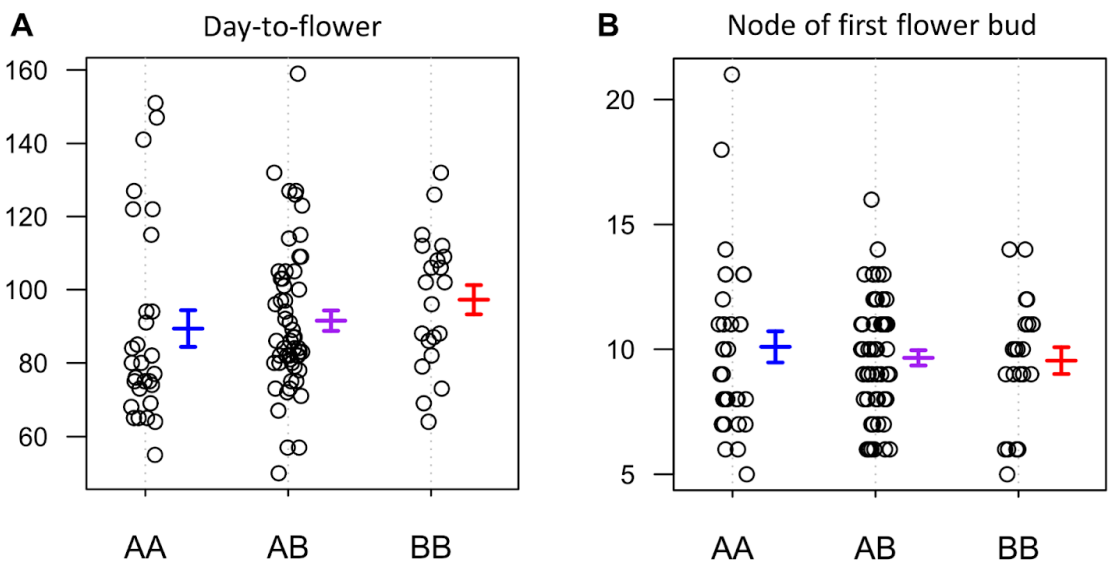

Fig. S2.6: Phenotypic effects of the photoperiodism-specific QTL on chromosome 3 in MAC on (A) day-to-flower and (B) the node of the first flower bud. “A” represents the allele from MAC1.2, a facultative LD line, while “B” represents the allele from MAC7, an obligate LD line. Error bars indicate standard errors.

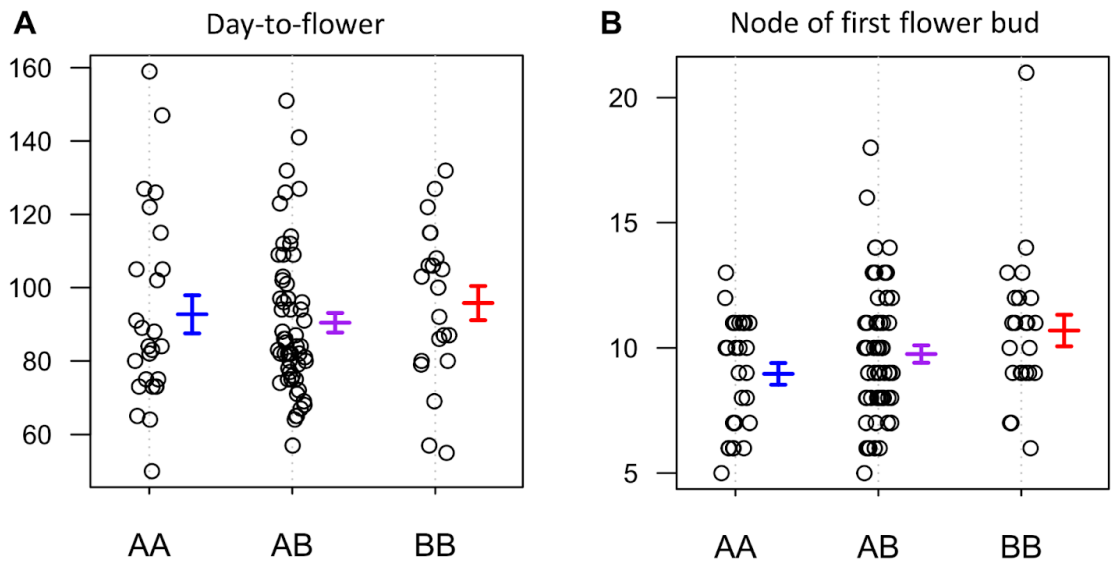

Fig. S2.7: Phenotypic effects of the photoperiodism-specific QTL on chromosome 5 in MAC on (A) day-to-flower and (B) the node of the first flower bud. “A” represents the allele from MAC1.2, a facultative LD line, while “B” represents the allele from MAC7, an obligate LD line. Error bars indicate standard errors.

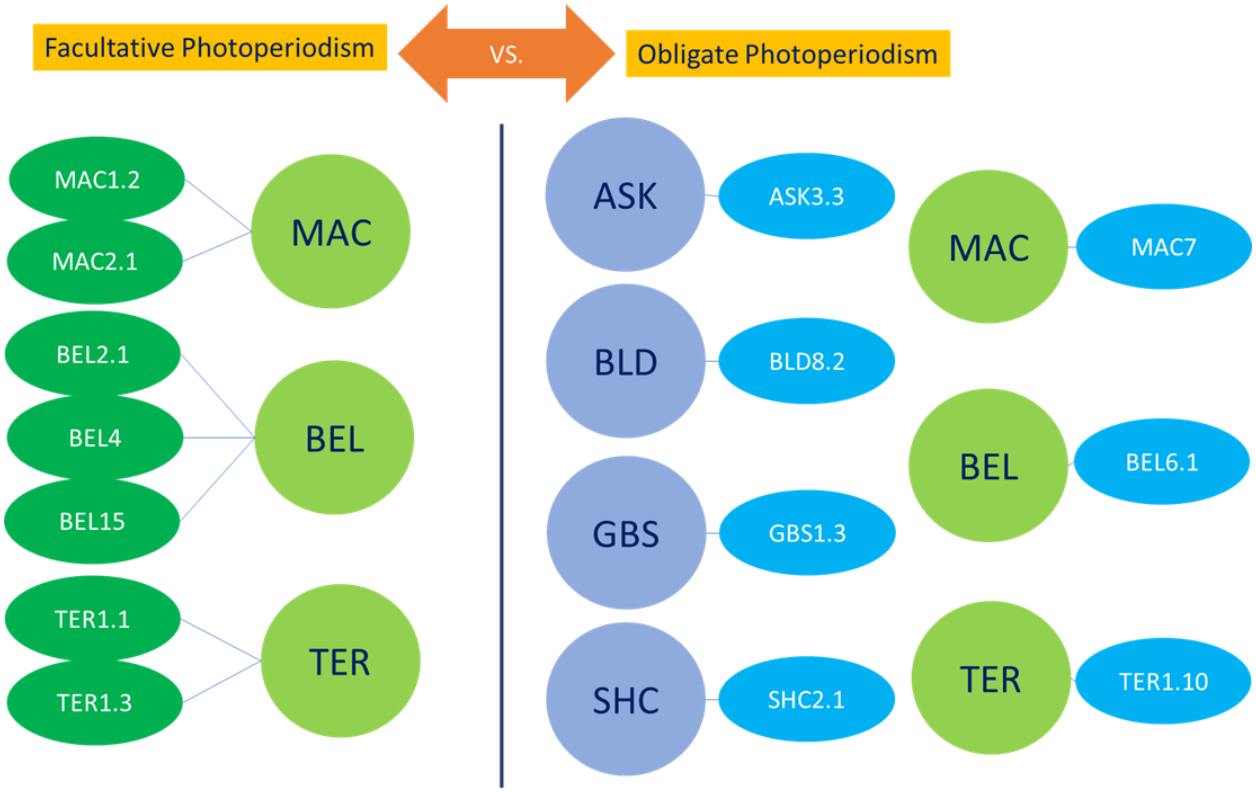

Fig. S2.8: Experimental design of RNAseq comparing photoperiodism across Mimulus accessions. The study included 7 lines from each facultative LD photoperiodism (dark green ellipses) and obligate LD photoperiodism (blue ellipses). The circles connected to the ellipses indicate the accessions from which each line was derived.

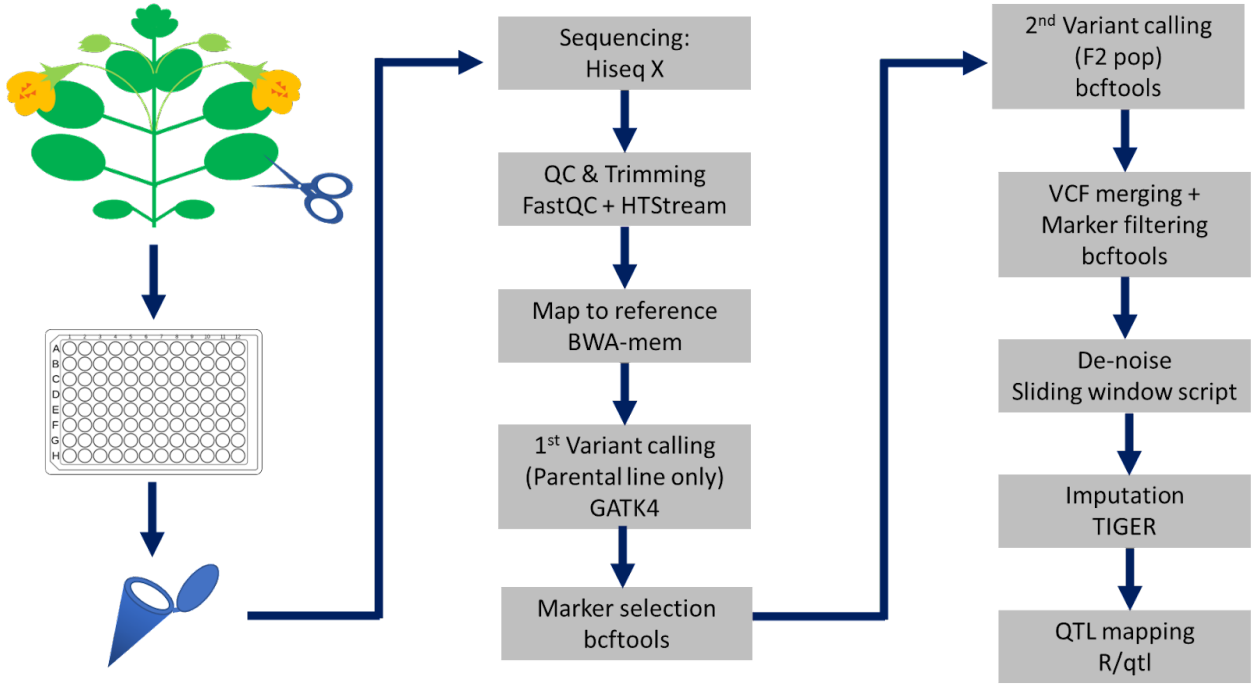

Fig. S2.9: Schematic workflow illustrating the process of SNP marker acquisition prior to QTL mapping.

Chapter 3 Transcriptomic Responses to Temperature Fluctuations of Mimulus

3.1 Abstract

Controlled environment conditions are extensively used in plant biology research and have provided valuable insights. However, extrapolating findings from controlled environments to natural field conditions is challenging due to the substantially greater fluctuations of environmental factors in the latter conditions. To bridge this knowledge gap, we incorporated diel fluctuating temperature into the controlled growth system and investigated the transcriptomic differences between constant and diel fluctuating temperature conditions in the Mimulus SWC accession. As part of a broader project, we developed pipelines for analyzing our time-course transcriptomic data. We summarized global expression variations using the molecular timetable method and identified specific genes displaying differential expression patterns between the two temperature profiles using a B-spline fitting strategy. Different expression patterns of circadian core clock genes were observed in the diel fluctuating temperature condition. Our study sheds light on the influence of diel fluctuations of temperature on gene expression and facilitates further research in understanding the impact of fluctuating environmental factors on plant transcriptomes. Moreover, we offer insights into potential caveats in this project that require subsequent refinement.

3.2 Introduction

Deciphering the molecular mechanisms that underlie plant responses to dynamic external environments has long been a fundamental objective in plant research. However, the intrinsic complexity of these processes poses challenges in studying all possible environmental factors simultaneously. To overcome this, a common strategy is to dissect environmental factors in controlled environments, where plants are cultivated under consistent conditions, with the exception of the specific factor being investigated. Although this approach has successfully contributed to our understanding of plant responses to environmental stimuli, increasing evidence shows that this strategy has certain limitations, particularly in extrapolating findings to real-world field conditions (Matsuzaki *et al.*, 2015; Poorter *et al.*, 2016; Matsubara, 2018; Chiang *et al.*, 2020; Hashida *et al.*, 2022). Notably, several recent articles shed light on the impact of natural temperature profiles on plant development and suggest that temperature can be one of the missing factors connecting laboratory results and field observations (Burghardt *et al.*, 2016; Annunziata *et al.*, 2018; Song *et al.*, 2018; Hashida *et al.*, 2022). Unlike growth chambers or greenhouses, which typically maintain constant or relatively stable temperatures, natural environments exhibit diel fluctuations in temperature, with peak values often occurring during the afternoon and lowest values at dawn. The diel temperature fluctuations exert wide-ranging influences on plants, including primary metabolism (Annunziata *et al.*, 2018), flowering regulation (Burghardt *et al.*, 2016; Song *et al.*, 2018), and the circadian clock (Annunziata *et al.*, 2018; Hashida *et al.*, 2022). These findings underscore the importance of considering diel fluctuating temperature in plant studies. Therefore, to achieve a comprehensive understanding of the effects of diel temperature fluctuations, further investigations encompassing different species and/or temperature profiles are necessary.

Time-course transcriptome analysis can be used to explore the differences in plant responses to constant and fluctuating conditions. Such analysis provides valuable information along the time axis, offering insights into the dynamic regulatory networks of plants. However,

the intricate nature of time-course transcriptomic data, along with the abundance of information it provides, presents challenges in terms of concise interpretation. To obtain a straightforward summary of time-course transcriptome analysis, we utilized the molecular timetable method, initially described by Ueda et al (Ueda *et al.*, 2004). This method comprises two main modules: circadian time group (CT group) binning and internal time prediction (Ueda *et al.*, 2004; Kerwin *et al.*, 2011; Fukuda *et al.*, 2018; Wang *et al.*, 2021). Time points in this analysis are usually represented on the scale of Zeitgeber time (ZT), which is a temporal unit determined by the period of external cues (day-night cycle in our case and ZT0 corresponds to the onset of light). CT group binning involves the identification and grouping of cycling genes based on the timing of their peak expression levels. For example, a gene exhibiting maximum expression at ZT10 will be allocated into the CT10 group. To obtain summarized expression patterns for each CT group, the expression levels of each gene are standardized across all the samples. Subsequently, the average expression of all genes within the same CT group is calculated at each sampling time point. With a designated CT group, the global patterns can be compared between temperature profiles, either by visualizing the average expression curves or by utilizing rhythmicity prediction programs to acquire estimated rhythmic parameters.

The second component of the molecular timetable method involves internal time prediction. The primary objective of internal time prediction is to leverage the oscillating patterns within a pre-constructed training set to predict the internal time of a sample collected at a single time point. The establishment of the training set is based on the CT group binning method described earlier. Within this approach, genes belonging to a predetermined CT group can serve as “time indicators” for a specific time point. For instance, genes in CT10 act as time indicators for ZT10 due to their peak expression at that time. Utilizing this information, we are able to predict the internal time for a sample by identifying which group of time indicator genes displays the highest average expression at the time of collection. In practice, the average expression for each CT group is computed for a sample, and subsequently, a curve is fitted to

these average expressions across the CT groups. The internal time is expected to coincide with the peak of the fitted curve. The relationship between CT groups and the predicted internal time can be illustrated using a three-dimensional wave plot. For detailed derivation of this relationship, please refer to the previous articles (Fukuda *et al.*, 2018; Talamanca & Naef, 2020). Similar to CT group binning, with a defined CT group, the disparities among the predicted internal times of the samples collected under different temperature profiles can provide global summaries of the impact of temperature on the rhythmic regulation of gene expression.

This study is part of a broader project aimed at investigating the influence of diel fluctuating temperature on plant transcriptomes. Building upon unpublished data by Johanna Schmitt and Daniel Runcie, who examined transcriptomic differences in *Arabidopsis thaliana*, our objective is to expand this project by focusing on *Mimulus guttatus* (hereafter referred to as Mimulus), another species used in our laboratory. To replicate the natural diel temperature fluctuations experienced by Mimulus in its native habitat, we constructed an hourly-based temperature profile using publicly available weather data. Our primary goal is to establish a systematic analysis workflow that incorporates the molecular timetable method and spline-fitting modeling to analyze time course data. Through the application of these approaches, we gained deeper insights into the molecular dynamics and regulatory mechanisms underlying Mimulus' response to diel temperature fluctuations. However, we also recognized issues in the analysis pipeline we developed and suggested potential improvements for future research.

3.3 Materials and methods

Plant materials and growth conditions

In this study, we utilized the annual *Mimulus guttatus* SWC accession collected by Benjamin K. Blackman's laboratory and described in the article by Kooyers et al (Kooyers *et al.*, 2015). The seeds underwent cold stratification on a 1:1 mixture of Sunshine Mix #1 potting soil and vermiculite in the dark at 4°C for 7 days prior to germination. Germination took place under controlled conditions of 22°C and 14:10 hours light/dark cycles (constant condition) in Conviron E7/2 growth chambers. After germination, the young seedlings were transplanted into plastic rose pots measuring 5 cm square by 8 cm deep, filled with a fresh mixture of Sunshine Mix #1 potting soil and vermiculite. The pots were randomized within the growth chambers and rotated three times per week and were watered every 2 days using fertilizer water provided by the UC Davis Controlled Environment Facility. Upon the development of the 2nd true leaves, half of the plants were transferred to a diel fluctuating temperature condition, also with 14:10 hours light/dark cycles. Details of the temperature profile are provided below. On the 12th day after the transfer, fully expanded 2nd or 3rd true leaves were collected from the plants in both temperature profiles every 3 hours, starting at dusk on the first day, spanning a 24-hour duration, for RNAseq library preparation. We collected four biological replicates from different plants for each combination of time point and temperature profile.

Temperature treatments

For the constant temperature profile, we maintained the temperature at 22°C, as this is commonly used for conducting Mimulus experiments. The daylength was set as a 14:10 hours day/night cycle, based on the critical photoperiod (the minimum daylength required for flowering) for the SWC accession.

To recreate a diel fluctuating temperature profile that mimics the temperature conditions during the growing season of Mimulus, we obtained and analyzed hourly-based ground

temperature data from the KRDM weather station on the Weather Underground database (https://www.wunderground.com/). This weather station was chosen because of its proximity to Iron Mountain in Oregon, USA, which is the habitat of the widely utilized Mimulus IM accession series. We applied and fitted the temperature model described by Cesaraccio et al (Cesaraccio *et al.*, 2001). Among the available temperature data, we selected the dataset from 6/6, 2016, as it closely matched the desired temperature profile we intended to evaluate (14:10 hours day/night cycle, average temperature 22°C) and displayed a smooth diel fluctuating temperature pattern. Based on this dataset, we generated an hourly temperature profile (Fig. S3.1) with a maximum temperature of 32°C (at ZT10 and 11) and a minimum temperature of 12°C (at ZT1).

RNAseq library preparation and data analysis

Samples were frozen using liquid nitrogen and then homogenized with two or three 2.8 mm ceramic beads using the SPEX™ SamplePrep 2010 Geno/Grinder at 1100 rpm for 2 minutes. mRNA extraction was conducted using the Dynabeads™ mRNA DIRECT Kit (Invitrogen, Cat. #: 61012), and the concentration of mRNA was assessed using the QuantiFluor RNA System (Promega, Cat. #: E3310) to ensure sufficient input mRNA for library preparation. RNAseq libraries were prepared following the BrAD-seq protocol (Townsley *et al.*, 2015) with some modifications. The libraries were then cleaned and size-selected using homemade SPRI beads. After performing gel electrophoresis to assess qualities, the libraries were pooled together at the same concentration and sequenced on a single lane of an Illumina HiSeq X sequencer, producing 150-bp paired-end reads.

The sequencing reads were subjected to quality assessment using FastQC (Andrews, 2010) and trimmed with Trimmomatic (Bolger *et al.*, 2014). Mapping to the Mimulus reference genome was performed using the STAR aligner (Dobin *et al.*, 2013), and the count tables generated were merged to create the input for the limma-voom package (Law *et al.*, 2014).

RNAseq data normalization and model fitting were performed following the guidelines presented in the limma-voom tutorial (Law *et al.*, 2018). For the time course data, B-splines were applied using the `pbs` function in the `pbs` package, with degrees of freedom set to 5. Candidate genes were identified based on a false discovery rate (FDR) threshold of 5% (i.e., with a Benjamini-Hochberg adjusted p-value below 0.05 (Benjamini & Hochberg, 1995)).

Molecular timetable analysis

The first step of the molecular timetable method involved circadian time group (CT group) binning. To identify cycling genes and determine their CT group, we utilized the JTK_CYCLE algorithm version 3.0 (Hughes *et al.*, 2010) on our normalized RNAseq data (on the counts per million, CPM, scale) following the user guide provided on the JTK_CYCLE website (https://openwetware.org/wiki/HughesLab:JTK_Cycle). We applied a false discovery rate (FDR) threshold of 5% to identify target cycling genes. Since we collected a total of nine time points in a 24-hour period, JTK_CYCLE automatically binned the cycling genes into 16 CT groups for each temperature profile based on their peak time. The cycle genes within a CT group can be defined by either the dataset of constant or diel fluctuating temperature conditions.

To compare expression patterns between the two temperature profiles within the same CT group, we analyzed the datasets from the different temperature profiles separately and visualized them together with curve plots (see Fig. 3.2) using `ggplot2` package in R. Specifically, we standardized the expressions of each gene across samples within the same temperature profile to ensure all cycling genes had the same average expression level and standard deviation, thus eliminating the effect of intrinsic expression level variation among genes. Next, we calculated the average expressions of all genes within the same CT group for each sample. Since four samples were collected at each ZT time point, each sample represented a single replicate of that specific ZT time point. The expressions were then standardized again across samples within the same CT group to make them comparable among

different CT groups. In addition to visualizing CT curves, we also used JTK_CYCLE to estimate the phase and amplitude of each CT curve for the purpose of comparison.

To predict the internal time for each sample based on the information obtained from the CT group binning, we applied JTK_CYCLE on each sample's CT group expression data to detect the CT group with the highest expression, which was defined as the internal time of that sample.

GO enrichment analysis

The Mimulus GO annotations for each gene were obtained from Dicots PLAZA 5.0 (Van Bel *et al.*, 2022). GO enrichment analysis was performed using GOseq (Young *et al.*, 2010) with read count bias correction. For function predictions, we retrieved the homologs of Mimulus genes in Arabidopsis or rice by querying the Dicots PLAZA 5.0 website (Van Bel *et al.*, 2022).

Statistical analysis

All statistical analyses for phenotyping and gene expressions were conducted using the R language (R version 4.2.2). We used the `pheatmap` package to construct heatmaps and other data visualization was performed using the `ggplot2` package. For pairwise comparisons, we used the `emmean` and `contrast` functions in the `emmeans` package to process the `lm` results.

3.4 Results

3.4.1 Identification of cycling genes under constant and diel fluctuating temperature conditions

To investigate the transcriptomic differences in Mimulus under different temperature conditions, we conducted a time-course experiment using the SWC accession. The plants were exposed to either constant or diel fluctuating temperatures with an average of 22°C, following a 14:10-hour light/dark cycle (Fig. S3.1). We collected leaf samples at 3-hour intervals over a 24-hour period, starting at dusk on the first day (ZT-10) and ending at dusk on the second day (ZT14). Four biological replicates were obtained, each harvested from different plants. RNAseq was used to obtain the transcriptome data. After data preprocessing and filtering, we obtained 17,912 genes with a sufficiently large number of counts for downstream analyses.

We applied the JTK_CYCLE algorithm (Hughes *et al.*, 2010) to our RNAseq data to identify cycling transcripts, resulting in the detection of 1,557 and 2,945 genes exhibiting rhythmic expression patterns under constant temperature and diel fluctuating temperature conditions, respectively (Fig. 3.1A and B). Based on the timing of their peak expression as predicted by JTK_CYCLE, we binned these genes into 16 CT groups for each temperature condition. The CT groups ranged from CT-10 to CT12.5, with each group representing a 1.5-hour window, which is the resolution of our analysis. As shown in Fig. 3.1C and D, the distributions of the bin sizes differed between the two temperature profiles. Under the diel fluctuating temperature condition, we observed two peaks at CT-1 and CT11 with a high frequency of cycling genes. In contrast, the constant temperature condition exhibited a single peak at CT11.

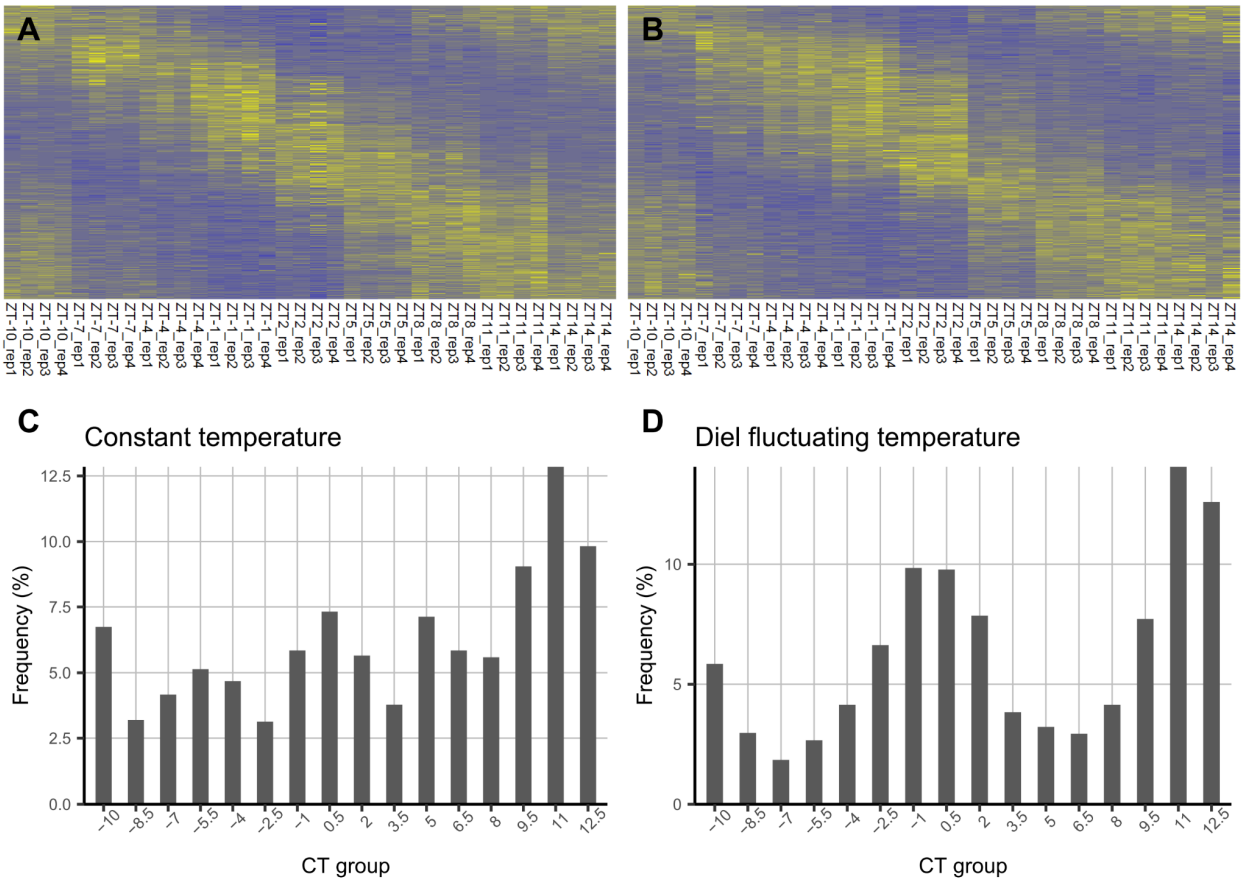

Fig. 3.1: Identification of cycling genes under constant (A and C) and diel fluctuating (B and D) temperature conditions. Heat maps (A and B) illustrate the standardized expression levels of cycling genes under constant (A) and diel fluctuating (B) temperature conditions, comprising 1,557 and 2,945 genes, respectively. The adjacent columns represent four biological replicates for each time point, ordered from ZT-10 to ZT14, left to right. Genes are sorted in rows based on their peak expression times. Yellow and blue represent high and low expression levels, respectively. The bar plots depict the differences in the frequency distributions of the cycling genes across 16 CT groups under constant (C) and diel fluctuating (D) temperature conditions.

3.4.2 Diel temperature fluctuations have influence on the global expression patterns of cycling genes

To investigate the global pattern variation in cycling genes under two different temperature conditions, we analyzed expression curves of CT groups: The standardized expression levels of all genes within a CT group were averaged along the ZT time axis

separately for the datasets of constant and diel fluctuating temperature conditions. Subsequently, the resulting average expression curves (CT curves) were subjected to analysis using JTK_CYCLE to obtain the phase and amplitude information. Meanwhile, CT curves were visualized and compared directly.

First, we examined the expression patterns using the CT groups defined under the constant temperature condition (Fig. 3.2). Using JTK_CYCLE, all the CT curves derived from samples collected under the constant temperature were shown to have the peak times at the corresponding ZT time points (Table 3.1). However, visualizing these CT curves (Fig. 3.2, salmon curves) revealed certain deviations. In general, most CT curves displayed prominent peaks around their corresponding ZT time points. For example, the CT-1 curve exhibited the highest expression levels at ZT-1. Nonetheless, certain CT groups, such as CT3.5 and CT5, demonstrated deviations by peaking at approximately ZT2 and ZT3, respectively. The disparities between the JTK_CYCLE calculations and the visualized CT curves could be attributed to the limited resolution of our dataset. Since we sampled at intervals of 3 hours within a single 24-hour period, only 16 bins were generated using JTK_CYCLE. This restricted number of bins might not adequately capture the rhythmic patterns of the cycling genes, subsequently leading to imprecise predictions or categorizations of the peak time of the cycling genes. We discuss these deviations and suggest potential improvements in the discussion section. Nevertheless, the comparisons of CT curves between two temperatures remained valid, as we only examined the timing of expression peaks within the same CT groups. We found that the majority of the expression peaks observed using the fluctuating temperature dataset aligned with those observed in the constant temperature dataset, except for CT-7, CT-5.5, CT8, and CT9.5, which JTK_CYCLE indicated delayed maximum expressions under the fluctuating temperature condition. The visualized results approximately demonstrated these delayed patterns (Fig. 3.2, cyan curves). However, we also noticed that certain variations in rhythmic characteristics, such as waveform, were challenging to detect using JTK_CYCLE but were evident in the visualized

curves. For example, the CT8 curves may also display waveform differences between two temperature profiles, not just phase shift.

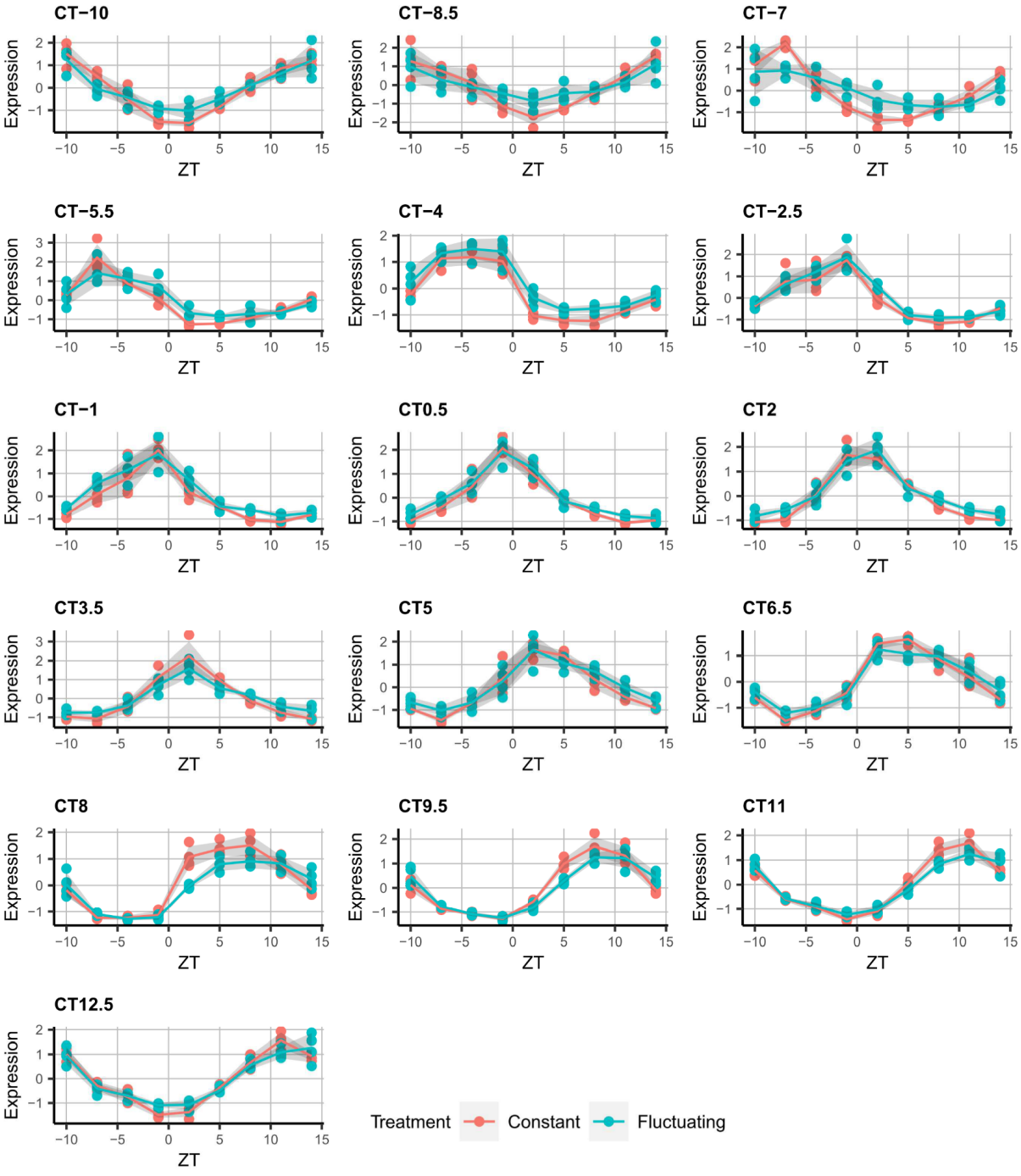

Fig. 3.2: Phase shifts observed in some CT groups determined under the constant temperature conditions. JTK_CYCLE identified delayed maximum expressions for CT-7, CT-5.5, CT8, and CT9.5 under the fluctuating temperature condition (see Table 3.1). Expression levels of cycling genes within

each CT group were standardized, averaged separately for the constant temperature dataset (salmon curves) and the diel fluctuating temperature dataset (cyan curves), and arranged from ZT-10 to ZT14. Dots indicate mean expression values in a single sample, while shaded ribbon plots represent standard deviations.

Table 3.1: Predicted phase of each CT curve using JTK_CYCLE. CT groups were defined under the constant temperature condition. Therefore, all predicted peak times of CT curves using the constant temperature dataset are consistent with CT group binning. CT groups highlighted in red indicate those showing delayed peak times under the diel fluctuating temperature condition.

CT groups	Constant	Diel Fluctuating
CT-10	-10	-10
CT-8.5	-8.5	-8.5
CT-7	-7	-5.5
CT-5.5	-5.5	-4
CT-4	-4	-4
CT-2.5	-2.5	-2.5
CT-1	-1	-1
CT0.5	0.5	0.5
CT2	2	2
CT3.5	3.5	3.5
CT5	5	5

CT6.5	6.5	6.5
CT8	8	9.5
CT9.5	9.5	11
CT11	11	11
CT12.5	12.5	12.5

Next, we conducted the same analysis, but using the CT groups defined under the diel fluctuating temperature condition (Table 3.2 and Fig. 3.3). Similar to the previous analysis, JTK_CYCLE indicated that all CT curves derived from the dataset of diel fluctuating temperature exhibited peak times aligned with the expected ZT time points. The disparities between JTK_CYCLE calculations and the visualized CT curves were also observed in this case, with several CT groups displaying notable deviations from their expected peak times in Fig. 3.3 (cyan curves). Additionally, based on JTK_CYCLE calculations, phase shifts were identified by comparing the CT curves between constant and diel fluctuating temperature conditions (Table 3.2). Intriguingly, more CT groups exhibited phase shifts between two temperature profiles in this case compared to the previous one. Specifically, we observed delayed maximum expression in the diel fluctuating temperature condition for CT-10, -2.5, -1, 0.5, 9.5, 11, and 12.5, while accelerated maximum expression was noted for CT-7, -5.5, and 3.5.

The more pronounced disparity between CT curves of different temperature profiles using the CT group defined under the diel fluctuating temperature condition could be attributed to the higher number of cycling genes identified in this dataset (2,945 genes vs 1,557 predicted cycling genes in the dataset of constant temperature). However, it is challenging to differentiate between the possibility that the higher count of cycling genes resulted from better RNAseq

experimental performance in the diel fluctuating temperature dataset, or whether the diel temperature introduced additional signals to entrain more cycling genes. Interestingly, a consistent pattern emerged when investigating CT8 and CT9.5 for CT groups defined under the constant temperature condition (Table 3.1) with CT9.5 and CT11 for CT groups defined under the diel fluctuating temperature condition (Table 3.2). In these cases, the predicted peak times of the CT curves derived under the diel fluctuating temperature condition occurred at ZT9.5 and ZT11, exhibiting delayed peak expressions. Given this time window corresponded to the period with the highest temperature (above 30°C) in the diel fluctuating temperature profile, it's plausible that these shared shifts in CT curves were influenced by the temperature. Taken together, our findings of CT curve comparisons suggest that the global expression patterns of cycling genes are influenced by diel temperature fluctuations.

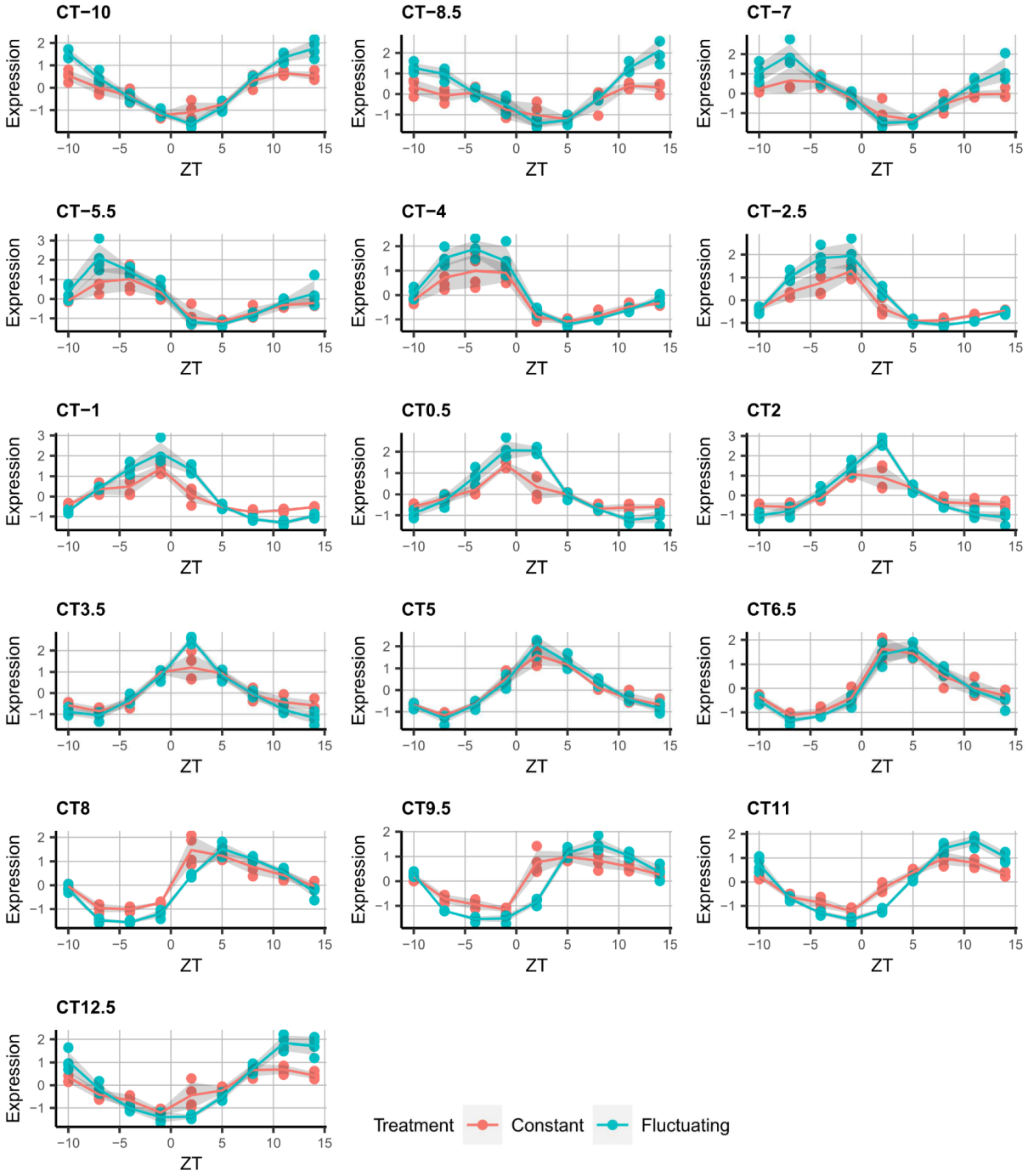

Fig. 3.3: Phase shifts observed in some CT groups determined under the diel fluctuating temperature conditions. JTK_CYCLE identified delayed maximum expressions for CT-10, CT-2.5, CT-1, CT0.5, CT9.5, CT11, and CT12.5 under the constant temperature condition (see Table 3.2). Expression levels of cycling genes within each CT group were standardized, averaged separately for the constant temperature dataset (salmon curves) and the diel fluctuating temperature dataset (cyan curves), and arranged from

ZT-10 to ZT14. Dots indicate mean expression values in a single sample, while shaded ribbon plots represent standard deviations.

Table 3.2: Predicted phase of each CT curve using JTK_CYCLE. CT groups were defined under the diel fluctuating temperature condition. Therefore, all predicted peak times of CT curves using the diel fluctuating temperature dataset are consistent with CT group binning. CT groups highlighted in red and blue indicate those showing delayed and advanced peak times under the diel fluctuating temperature condition, respectively.

CT groups	Constant	Diel Fluctuating
CT-10	-11.5	-10
CT-8.5	-8.5	-8.5
CT-7	-5.5	-7
CT-5.5	-4	-5.5
CT-4	-4	-4
CT-2.5	-4	-2.5
CT-1	-2.5	-1
CT0.5	-1	0.5
CT2	2	2
CT3.5	5	3.5
CT5	5	5
CT6.5	6.5	6.5

CT8	8	8
CT9.5	8	9.5
CT11	9.5	11
CT12.5	11	12.5

In addition to the CT curve analyses, we conducted additional assays to investigate the differences between two temperature profiles. Instead of examining the rhythmic parameters after obtaining the CT curves by averaging across the cycling genes, we estimated the rhythmic parameters for each gene individually. In other words, we treated each individual cycling gene within a CT group as a replicate for estimating that specific CT group.

Using the CT groups defined under the constant temperature condition, we observed six CT groups exhibited later average phases under the diel fluctuating temperature condition compared to the constant temperature condition (Fig. S3.3A). Conversely, when using the CT groups defined under the diel fluctuating temperature condition (Fig. S3.3C), earlier average phases under the constant condition for seven CT groups were detected. These results indicate a delayed maximum expression in the diel fluctuating temperature condition, which is similar to the trends observed in CT curve analyses. However, we did not observe any advanced peaks under the diel fluctuating temperature condition, as we detected in analyses shown in Fig. 3.3 and Table 3.2. Furthermore, while there were some overlapping results, several differences within CT groups detected by this strategy differed from the CT curve analyses. These deviations could be attributed to the influence other than peak time that contribute to shaping the final CT curve during the averaging process. For instance, a gene with a wide peak or a unique waveform could exert effects on time points other than the maximum peak. Some of

those information may be lost when focusing on the estimated rhythmic parameters for each individual gene.

However, investigating rhythmic parameters for each gene in a CT group separately provided a better understanding of the amplitude distributions. The ratio of the amplitude under different temperature conditions can be examined on a logarithmic scale. Deviations from zero in the ratio distribution indicate differences in amplitudes under the two temperature profiles. We observed consistent patterns using CT groups defined by either constant or fluctuating temperature conditions, with the amplitudes of cycling genes tending to be greater in CT-2.5, -1, and 0.5 under the diel fluctuating temperature condition. These CT groups had their expected peak time around dawn, corresponding to the time window with the lowest temperature in the diel fluctuating temperature condition. Thus, this amplitude difference could be related to the temperature-sensing or temperature-responding genes. Taken together, our results suggest that diel temperature fluctuations can have global effects on the amplitude and phase of cycling genes. However, further investigations and strategies are required to obtain closer insights into other rhythmic parameters, such as waveform.

3.4.3 Internal time prediction of the molecular timetable method is not suitable for summarizing the global differences in our transcriptome data

To further characterize the global differences in cycling genes under different temperature conditions, we used the molecular timetable method to predict internal times based on the CT group information established above. Specifically, we estimated the peak expressions across CT groups within a single sample and assessed if the predicted internal time aligns with the actual sample collecting time. This approach also allows us to compare the predicted internal times between samples collected under constant and diel fluctuating temperatures.

As in the preceding section, internal time predictions can be conducted based on the CT groups defined under either constant or diel fluctuating temperature conditions. When using the

CT groups defined under the constant temperature condition (Fig. 3.4A), the predicted internal times of the samples collected under the constant temperature approximately reflected their actual sample collecting ZT times. However, the predictions were not perfect, and a general acceleration of predicted internal times was observed (i.e., all the salmon boxes are positioned above the 1:1 line in Fig. 3.4A). The only exception was the prediction at ZT11, where the 95% confidence interval encompassed the actual ZT time, so we did not have evidence showing whether the predicted internal time differed from the actual time. Using the same CT groups, the internal time predictions for samples collected under the diel fluctuating temperature also exhibited some extent of deviation between the predicted internal time and the actual sample collecting time. Similar observations were found using the CT groups defined under the diel fluctuating temperature condition (Fig. 3.4B). Surprisingly, we did not detect any internal time alternation corresponding to the phase shift observed in the CT curve analyses above. Several possible reasons can contribute to this inability to precisely predict the internal time, including imprecise CT group binning, sparse sampling time points within a single day of data collection, and variation in peak amplitude across CT groups (please refer to the discussion section).

Interestingly, we consistently detected significant differences ($p < 0.05$) in predicted internal time at ZT2 between the samples collected in different temperature profiles (Fig. 3.4A and B). However, due to the concerns of imperfect internal time prediction, we cannot rule out the possibility of artifacts, and thus it is challenging to explain the biological meaning of this observation. Based on our current findings, it appears that the internal time prediction of the molecular timetable method may not be suitable for summarizing the global differences in our transcriptome data, and further investigations and improvements in this method are required.

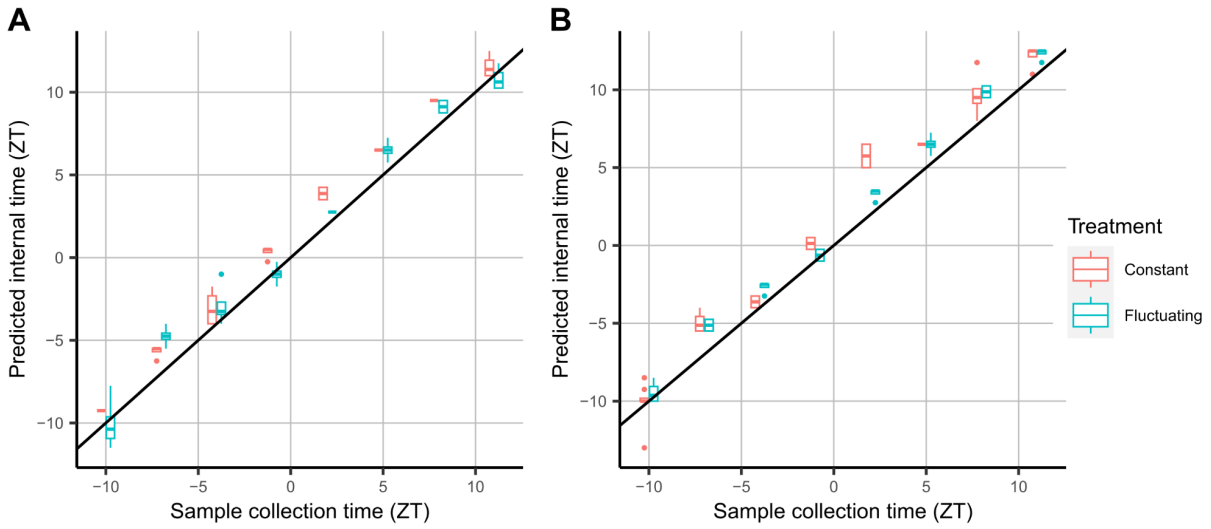

Fig. 3.4: Internal time prediction did not provide precise estimation in our case. CT groups were defined under the constant (A) and diel fluctuating (B) temperature conditions. The samples collected under the constant temperature condition are shown in salmon, while those collected under the diel fluctuating temperature condition are shown in cyan. Samples from the same conditions are grouped based on the actual sampling time. The black line represents the 1:1 line, indicating perfect alignment between predicted and actual sampling times.

3.4.4 Genes with distinct expression patterns between two temperature profiles are enriched with GO terms associated with temperature responses

Given the observations in the section of CT group comparisons, which demonstrated the global variation in cycling genes related to temperature profile differences, our subsequent goal was to identify specific genes exhibiting different expression patterns. To exploit the information thoroughly from the time-course experimental design, we fitted B-spline curves to our transcriptomic data, allowing us to assess the pattern variations between the two temperature treatments. We found 2,066 genes that displayed a significant response to diel fluctuating temperature at a false discovery rate (FDR) threshold of 0.05.

To investigate the biological relevance of those genes, we performed a Gene Ontology (GO) enrichment analysis. This analysis identified six significantly enriched GO terms, with an

FDR threshold of 0.1 (Table 3.3). Notably, three of these enriched terms were specifically related to temperature responses, including "response to heat," "response to temperature stimulus," and "cellular response to heat." These temperature-related terms were nested within the broader category of "response to abiotic stimulus," which was another enriched GO term. Additionally, among the target genes associated with the "protein folding" GO term, several encoded heat shock protein homologs in Arabidopsis known to respond to heat stress, such as Migut.A00100, Migut.C00037, Migut.C00039, Migut.F00194, etc. As for the remaining enriched GO term related to flowering time regulation, GO:0048577, we offer a detailed discussion in the discussion section due to certain potential issues with this enrichment. Taken together, these findings suggest that the target genes identified in our RNAseq experiment are enriched with GO terms associated with temperature responses, which is consistent with the temperature treatment employed in our experiment.

Table 3.3: List of significantly enriched GO terms and the adjusted p-value.

Category	Term	FDR-adjusted p-value
GO:0009408	response to heat	2.11E-08
GO:0009266	response to temperature stimulus	3.01E-04
GO:0006457	protein folding	2.57E-03
GO:0009628	response to abiotic stimulus	3.06E-02
GO:0034605	cellular response to heat	7.84E-02
GO:0048577	negative regulation of short-day photoperiodism, flowering	8.87E-02

3.4.5 Diel fluctuating temperature modulates the expression patterns of circadian clock-related genes in Mimulus

In our analysis of the 2,066 target genes showing significant responses to diel fluctuating temperature, we identified 33 genes annotated with the GO term "circadian rhythm" (GO:0007623). Although not enriched, circadian clock genes have been known to be influenced

by external signals such as temperature. Hence, we specifically examined the expression patterns of core clock genes within these 33 genes, including homologs of Arabidopsis genes *CCA1*, *LHY*, *PRR7*, *TOC1*, *LUX*, *RVE6*, *RVE8*, and *FKF1*.

As depicted in Fig. 3.5 and Table 3.4, diel fluctuating temperature exerted different effects on the expression of the target core circadian genes. We detected phase shifts in Migut.N01518 (*CCA1*/*LHY*), Migut.D01764 (*TOC1*), Migut.D00884 (*LUX*), Migut.F00285 (*RVE6*), and Migut.E00487 (*FKF1*) based on the JTK_CYCLE estimations, all of which exhibited a delay in the expression peak under the diel fluctuating temperature condition. However, it's important to note that certain genes displayed noticeable variations in the waveform, and for these changes, JTK_CYCLE can not offer effective estimations.

For the remaining target genes, Migut.L01650 (*PRR7*) and Migut.J00406 (*RVE8*), JTK_CYCLE predicted the same phase but different amplitudes between the two temperature profiles. The prediction of Migut.J00406 (*RVE8*) was unexpected, given that the expression curves of this gene appeared to have different peak times between two temperature profiles upon visual inspection. One possible explanation for this discrepancy could be the limited resolution mentioned earlier, which resulted in the truncated peak pattern between ZT-1 and ZT2 for the expression curve of diel fluctuating temperature. This truncated pattern might pose challenges in directly interpreting the data. For *PRR7*, although there was a difference in amplitude estimated by JTK_CYCLE, the result should be interpreted with caution because JTK_CYCLE only provides a single estimation, making it infeasible for conducting statistical tests.

Furthermore, we investigated the Mimulus homolog of *PIF4* (Migut.M00105), which was also among the 2,066 target genes responding to diel fluctuating temperature. In Arabidopsis, *PIF4* is a pivotal gene known to integrate thermal signals and circadian rhythms and regulate various developmental processes. Interestingly, we observed distinct expression patterns of *PIF4* between our two temperature profiles. Under the diel fluctuating temperature condition,

PIF4 exhibited a sharp peak, whereas under constant temperature conditions, it displayed a broad hump-shaped expression pattern. These findings collectively demonstrate the influence of diel fluctuating temperature on the expression patterns of genes associated with circadian rhythm in Mimulus. Further investigations are required to unravel the specific roles of individual genes in temperature response mechanisms within Mimulus.

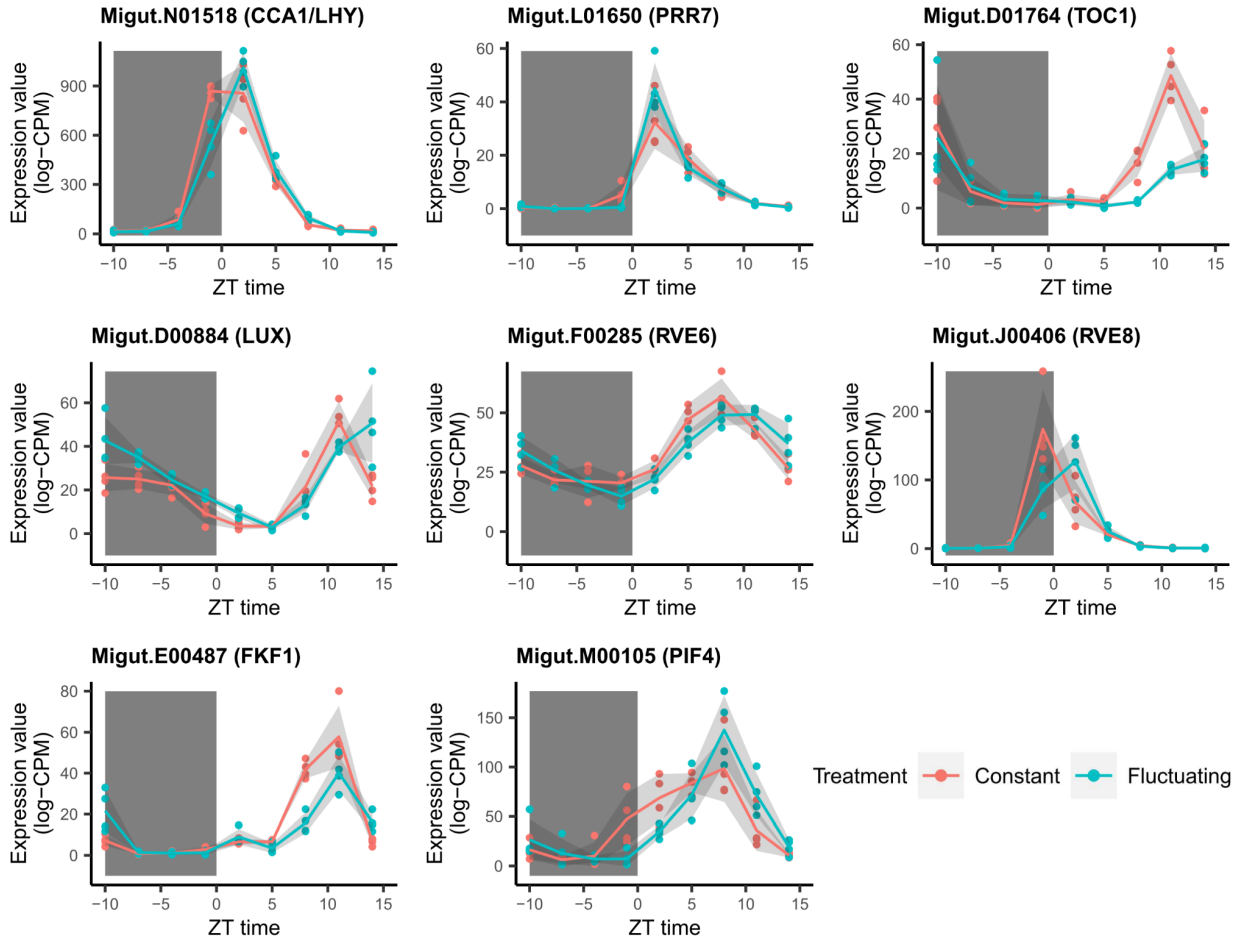

Fig. 3.5: Expression patterns of circadian clock genes under constant (salmon lines/dots) and diel fluctuating temperature (cyan lines/dots) conditions. Each dot represents the expression value of a biological replicate at a specific time point, and the shaded ribbon plots depict the standard deviation.

Table 3.4: Predicted phases and amplitudes of circadian clock genes using JTK_CYCLE. C.Phase and F.Phase represent the phase under the constant and diel fluctuating temperature conditions, respectively. C.Amp and F.Amp represent the amplitude under the constant and diel fluctuating temperature conditions, respectively.

Mimulus gene	Arabidopsis homologs	C.Phase	F.Phase	C.Amp	F.Amp
Migut.N01518	*CCA1/LHY*	2	3.5	247.5	267.9
Migut.L01650	*PRR7*	6.5	6.5	7.4	6.6
Migut.D01764	*TOC1*	12.5	14	12.3	6.5
Migut.D00884	*LUX*	-10	-8.5	14.0	17.2
Migut.F00285	*RVE6*	9.5	11	15.0	15.9
Migut.J00406	*RVE8*	2	2	15.9	20.4
Migut.E00487	*FKF1*	9.5	11	7.9	9.8
Migut.M00105	*PIF4*	6.5	8	40.4	41.3

3.5 Discussion

Laboratory investigations of plant physiological mechanisms commonly involve the utilization of controlled environments, such as growth chambers or greenhouses. While this approach contributes to our understanding of the components and architectures of molecular mechanisms across various topics in plant science, it is crucial to recognize that plants in their natural environments are subjected to far greater fluctuations and variability. Natural environments encompass a range of factors that exhibit regular fluctuations, such as diel temperature and light changes, as well as seasonal variations in photoperiod, along with unpredictable elements like weather factors and biotic interactions. The disparity between laboratory and field responses has been receiving increasing attention, prompting the exploration of strategies aimed at bridging this gap and enhancing the practical applicability of knowledge (Matsuzaki *et al.*, 2015; Poorter *et al.*, 2016; Matsubara, 2018; Chiang *et al.*, 2020; Hashida *et al.*, 2022). One approach directly examines plant responses in the field and attempts to explain the observed differences using existing knowledge or incorporating modeling for more comprehensive prediction and explanation (Nagano *et al.*, 2012; Matsuzaki *et al.*, 2015; Iwayama *et al.*, 2017; Dantas *et al.*, 2021). Alternatively, some studies “reintroduce” natural fluctuating factors gradually into controlled environments to mimic natural field conditions (Annunziata *et al.*, 2017, 2018; Song *et al.*, 2018; Chiang *et al.*, 2020; Hashida *et al.*, 2022). This approach provides researchers with greater control over experimental conditions and enables the dissection of confounding factors influencing specific responses of interest. In this article, we adopt the second strategy and focus on the effect of diel fluctuating temperature. Specifically, we investigated the transcriptomic dynamics of *Mimulus guttatus* SWC accession under both constant and diel fluctuating temperature conditions. The objective of this pilot study is to develop pipelines for examining time-course RNAseq data, providing both global and detailed insights into the differences between the two temperature profiles. The ultimate goal of

the broader project is to establish links between these patterns and whole-plant characteristics, such as flowering time and growth.

To gain the global summaries of the transcriptomic differences responding to constant and diel fluctuating temperatures, we utilized the molecular timetable method (Ueda *et al.*, 2004), which comprises two components: CT group binning and internal time prediction. This approach enabled us to detect general variations in expression patterns when comparing the two temperature profiles in our experiments. Nevertheless, we also encountered some caveats and limitations, which we discuss below.

In the CT group binning analysis, we identified shifts in expression peaks using CT groups defined under both constant and diel fluctuating temperature conditions. However, the extent of CT curve differences between the two temperature profiles varied, with more pronounced distinctions observed when CT groups were defined under the diel fluctuating temperature condition. Interestingly, a shared trend emerged, specifically when focusing on the time window where CT curves from the fluctuating temperature condition exhibited predicted peak times at ZT9.5 and 11, resulting in delayed maximum expressions (Table 3.1 and 3.2). Notably, these time points coincided with the period of highest temperature in the diel fluctuating temperature condition, where temperatures remained greater than 30°C. Consequently, it's plausible that these phase shifts are influenced by factors like heat-shock responses or temperature fluctuations between day and night. Another possible regulatory hub contributing to these phase shifts is the Evening Complex (EC) in the core clock, consisting of LUX, ELF3, and ELF4, as the EC has been shown to play a vital role in entraining ambient temperature cues (Box *et al.*, 2015; Ezer *et al.*, 2017) and can influence various developmental regulations through *PIF4* and *PIF5* (Nusinow *et al.*, 2011). It is possible that other core clock genes are also involved, as discussed below. However, additional experiments are necessary to fully understand the underlying mechanisms of these CT group expression shifts. For example, it is

important to distinguish if the shifts are the result of delayed/accelerated expression or if they are related to the stabilities of the mRNA transcripts.

Although we detected variation in global expression patterns by comparing CT curves of different temperature profiles, we noticed the disparities between the JTK_CYCLE estimations and the visualized CT curves, and the peak times of certain CT groups did not precisely coincide with the expected ZT time points. These observations suggest that the binning of cycling genes by JTK_CYCLE may be imprecise for our experiment. It was reported that phase estimation using JTK_CYCLE can be inaccurate when the input data is sparse (Hughes *et al.*, 2017). Even if the estimation was accurate in our case, it's worth noting that we generated only 16 CT groups from our time-course data, which encompassed 9 time points across a single 24-hour cycle. Such resolution may not be sufficient for accurately describing the patterns of cycling genes. In addition, JTK_CYCLE only produces a single estimate for the phase and amplitude of each CT curve, thereby making it infeasible to assess differences through statistical tests. To address this issue, we can improve the resolution by increasing the frequency of sampling time points. Moreover, in future experiments, extending the duration of the study from 24 hours to a minimum of 48 hours will be beneficial. Simulations have indicated that experiments with fewer than 48 hours (equivalent to 2 cycles in the time series) can render the resulting data sensitive to outliers and elevate the risk of encountering false negatives. Another improvement can be achieved by using more sophisticated algorithms that align better with our experimental design or that are capable of better estimating different rhythmic parameters, such as waveforms. For example, eJTK demonstrates better performance for experiments involving replicates compared to repeated cycles (Hutchison *et al.*, 2018), making it a potentially more fitting choice for our experimental design. In general, this issue of imprecise CT group binning did not substantially affect the CT curve comparisons, as we only examined the pattern differences within the same CT group. However, it can pose challenges when conducting cross-analysis comparisons, like comparing results between analyses using CT

groups defined under different temperature profiles (i.e., Fig. 3.2 vs Fig. 3.3), because the CT groups with the same name may actually have very different peak times. Additionally, this could potentially raise issues with downstream predictions of internal time.

In the internal time prediction section, we found this method was not suitable for summarizing the global expression differences for our RNAseq data. As internal time prediction is based on the results of CT group binning, the primary issue lies in the imprecise CT group binning mentioned in the previous paragraph, which may cause cascading effects and amplify the bias. Furthermore, we observed that the maximum expression levels of CT curves varied among CT groups even after standardization. Therefore, a local peak within a CT group may not necessarily represent the highest expression compared to other CT groups for a sample (see Fig. S3.2 for an example). Finally, if a shift in CT curves is not significant enough, the low resolution of our molecular timetable construction may not effectively detect it. These issues make it difficult to interpret any discoveries from internal time predictions and confirm their accuracy. As the field is rapidly evolving and newer approaches continue to emerge, future analyses could incorporate updated algorithms other than the molecular timetable method for internal time predictions, as outlined in the recent review by Talamanca and Naef (Talamanca & Naef, 2020). Taken together, for our RNAseq data, only the CT curve comparison can provide a relatively reliable global summary of transcriptomic differences.

To gain a deeper insight into the variation in the transcriptome, we identified genes that responded to diel fluctuating temperature by applying B-spline fitting to our time-course data. The subsequent GO enrichment analysis revealed a significant enrichment of GO terms associated with temperature responses (GO:0009408, GO:0009266, and GO:0034605), which was consistent with our investigation of the impact of diel fluctuating temperature. Interestingly, we also identified an enriched gene ontology term: GO:0048577, "negative regulation of short-day photoperiodism, flowering". Among all the 22,019 genes with GO annotations in Mimulus, only eight genes were associated with GO:0048577, and six of them were among our

identified candidate genes: Migut.D00803, Migut.D01811, Migut.H00566, Migut.J00012, Migut.J00795, and Migut.L01278. Migut.L01278 and Migut.D01811 encode homologs of the CYCLING DOF FACTOR (CDF) family, which were demonstrated to function as repressors of *Constans* (*CO*), a positive circadian regulator of flowering time in Arabidopsis (Imaizumi *et al.*, 2005; Sawa *et al.*, 2007; Fornara *et al.*, 2009). In line with this, Migut.H00566 encodes a homolog of *CONSTANS-like 2* in Arabidopsis, which is a close homolog of *CO* (Simon *et al.*, 2015). While *CONSTANS-like 2* in Arabidopsis is primarily involved in circadian rhythm regulation rather than flowering regulation (Ledger *et al.*, 2001), its homolog in rice (Os09g0240200, *OsCO3*) acts as a repressor of short-day flowering (Kim *et al.*, 2008). The three remaining genes, Migut.D00803, Migut.J00012, and Migut.J00795, encode proteins containing K-homology (KH) RNA-binding domains, and their homologs affect flowering time in rice (Cai *et al.*, 2014) or Arabidopsis (Ripoll *et al.*, 2009) under both long-day and short-day conditions. While the detection of these six genes aligned with the previous finding that flowering genes can respond differently to diel fluctuating temperatures compared to constant temperature conditions (Song *et al.*, 2018), it is crucial to interpret the results of the enrichment in GO:0048577 with caution. First, the limited number (only eight in total) of genes associated with this specific GO term may be due to incomplete GO annotation of Mimulus, which can introduce bias and affect the statistical analysis based on this incomplete GO term. Moreover, GO:0048577 focuses on flowering regulation under short-day conditions, which can be misleading as it encompasses evidence from both long-day plants (such as Arabidopsis) and short-day plants (such as rice). It is important to note that short days act as inductive cues for short-day plants but repressive cues for long-day plants. This essential difference can potentially result in different biological implications of this GO term across different species. Nonetheless, there is room for refinement in the GO enrichment analysis. In this study, we assessed enrichment by considering all the target genes collectively. However, to more effectively

incorporate temporal information, binning these target genes into distinct CT groups and subsequently conducting GO enrichment analysis might unveil more specific patterns.

Finally, we examined the influence of diel fluctuating temperature on core circadian clock genes in Mimulus. Previous studies investigating various species and temperature profiles have indicated that diel fluctuating temperatures can alter the expression of circadian clock genes (Annunziata *et al.*, 2018; Hashida *et al.*, 2022). This observation is not surprising, as circadian clocks have the ability to utilize external temperature cues to entrain their rhythms with the environmental day-night cycles (Gil & Park, 2019). Consequently, some studies suggested that while constant conditions are necessary for dissecting individual components of the circadian clock, understanding the adaptive roles of the circadian clock requires investigations under conditions that more closely resemble the natural environment (Rubin *et al.*, 2017; Panter *et al.*, 2019; Paajanen *et al.*, 2021; Laosuntisuk *et al.*, 2023).

Among the 33 candidate genes associated with the circadian rhythm, we specifically focused on seven genes that are known to play key roles in the core clock of Arabidopsis. The peak expression times of these seven genes were generally consistent with their Arabidopsis homologs, supporting the notion that the fundamental architecture of the core clock is conserved across these diverse plant species (Filichkin *et al.*, 2011; Nakamichi, 2020; Laosuntisuk *et al.*, 2023).

Interestingly, the majority of these genes exhibited a delayed peak in our diel fluctuating temperature condition compared to the constant temperature condition, except for the homologs of *PRR7* and *RVE8*. These phase shifts may be related to the peak shifts in the CT groups described earlier, given the core clock genes can influence a wide array of downstream genes and cause global changes. It is important to note that although certain target core clock genes did not belong to the CT groups exhibiting the phase shifts, we cannot rule out their role in shaping the CT groups. It is possible that there are phase shifts in those clock genes, but the CT curves do not reflect this change due to their averaging nature. Surprisingly, we did not

observe a significant difference in the expression of *ELF3* or *ELF4,* known to play a crucial role in integrating ambient temperature signals (Thines & Harmon, 2010; Box *et al.*, 2015; Ezer *et al.*, 2017; Jung *et al.*, 2020). This could be related to the thermosensory mechanism of ELF3 (Jung *et al.*, 2020), which involves regulation at the protein level and may not be accurately reflected in mRNA expression levels. Interestingly, we did detect a prominent pattern change in *PIF4*, a gene regulated by *ELF3* and involved in integrating multiple environmental signals, including temperature cues (Choi & Oh, 2016). Under the constant temperature condition, *PIF4* exhibited a broad expression peak throughout the day. However, under the diel fluctuating temperature condition, we detected a sharper peak at ZT8, which corresponds to the hottest time point in our diel fluctuating temperature profile. This finding is consistent with the role of *PIF4* in thermoresponsive growth reported previously.

As part of a larger research project focused on investigating the impact of diel fluctuating temperature on plant transcriptomes, we developed pipelines for analyzing time-course transcriptome data. However, some improvements are suggested to better test rhythmic parameter differences and to optimize the application of the molecular timetable method (Hughes *et al.*, 2017; Parsons *et al.*, 2020). Our analyses provide both global and detailed insights into the expression patterns under constant and diel fluctuating temperature conditions. Interestingly, the differences observed between constant and diel fluctuating temperature conditions in Mimulus are not as pronounced as those found in our unpublished Arabidopsis dataset. Since different diel fluctuating temperature programs were employed for these two species, it is difficult to attribute these differences solely to species or temperature profile variations. Nonetheless, this discovery highlights the dynamic nature of plant transcriptomic responses to temperature and underscores the challenges of using non-constant environments in experiments. Thus, to gain a better understanding of the effects of diel temperature fluctuations, further experiments across different species and temperature profiles are

necessary, and subsequent meta-analyses may provide an improved resolution to address this question.

3.6 References

Andrews S. **2010**. FASTQC. A quality control tool for high throughput sequence data.
Annunziata MG, Apelt F, Carillo P, Krause U, Feil R, Koehl K, Lunn JE, Stitt M. **2018**. Response of Arabidopsis primary metabolism and circadian clock to low night temperature in a natural light environment. *Journal of Experimental Botany* **69**: 4881–4895.
Annunziata MG, Apelt F, Carillo P, Krause U, Feil R, Mengin V, Lauxmann MA, Köhl K, Nikoloski Z, Stitt M, *et al.*2017. Getting back to nature: a reality check for experiments in controlled environments. *Journal of Experimental Botany* **68**: 4463–4477.
Benjamini Y, Hochberg Y. **1995**. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**: 289–300.
Bolger AM, Lohse M, Usadel B. **2014**. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.
Box MS, Huang BE, Domijan M, Jaeger KE, Khattak AK, Yoo SJ, Sedivy EL, Jones DM, Hearn TJ, Webb AAR, *et al.*2015. ELF3 Controls Thermoresponsive Growth in Arabidopsis. *Current Biology* **25**: 194–199.
Burghardt LT, Runcie DE, Wilczek AM, Cooper MD, Roe JL, Welch SM, Schmitt J. **2016**. Fluctuating, warm temperatures decrease the effect of a key floral repressor on flowering time in Arabidopsis thaliana. *New Phytologist* **210**: 564–576.
Cai Y, Vega-Sánchez ME, Park CH, Bellizzi M, Guo Z, Wang G-L. **2014**. RBS1, an RNA Binding Protein, Interacts with SPIN1 and Is Involved in Flowering Time Control in Rice. *PLOS ONE* **9**: e87258.
Cesaraccio C, Spano D, Duce P, Snyder RL. **2001**. An improved model for determining degree-day values from daily temperature data. *International Journal of Biometeorology* **45**: 161–169.
Chiang C, Bånkestad D, Hoch G. **2020**. Reaching Natural Growth: The Significance of Light and Temperature Fluctuations in Plant Performance in Indoor Growth Facilities. *Plants* **9**: 1312.
Choi H, Oh and E. **2016**. PIF4 Integrates Multiple Environmental and Hormonal Signals for Plant Growth Regulation in <italic>Arabidopsis</italic>. *Moleucles and Cells* **39**: 587–593.
Dantas LLB, Dourado MM, de Lima NO, Cavaçana N, Nishiyama Jr. MY, Souza GM, Carneiro MS, Caldana C, Hotta CT. **2021**. Field microenvironments regulate crop diel transcript and metabolite rhythms. *New Phytologist* **232**: 1738–1749.
Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. **2013**. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15–21.
Ezer D, Jung J-H, Lan H, Biswas S, Gregoire L, Box MS, Charoensawan V, Cortijo S, Lai X, Stöckle D, *et al.*2017. The evening complex coordinates environmental and endogenous signals in Arabidopsis. *Nature Plants* **3**: 1–12.
Filichkin SA, Breton G, Priest HD, Dharmawardhana P, Jaiswal P, Fox SE, Michael TP, Chory J, Kay SA, Mockler TC. **2011**. Global Profiling of Rice and Poplar Transcriptomes Highlights Key Conserved Circadian-Controlled Pathways and cis-Regulatory Modules. *PLOS ONE* **6**: e16907.
Fornara F, Panigrahi KCS, Gissot L, Sauerbrunn N, Rühl M, Jarillo JA, Coupland G. **2009**. Arabidopsis DOF Transcription Factors Act Redundantly to Reduce CONSTANS Expression and Are Essential for a Photoperiodic Flowering Response. *Developmental Cell* **17**: 75–86.
Fukuda H, Tanigaki Y, Moriyuki S. **2018**. Detection and Utilization of Biological Rhythms in Plant Factories. In: Kozai T, ed. Smart Plant Factory: The Next Generation Indoor Vertical Farms. Singapore: Springer Singapore, 367–384.
Gil K-E, Park C-M. **2019**. Thermal adaptation and plasticity of the plant circadian clock. *New Phytologist* **221**: 1215–1229.
Hashida Y, Tezuka A, Nomura Y, Kamitani M, Kashima M, Kurita Y, Nagano AJ. **2022**.

Fillable and unfillable gaps in plant transcriptome under field and controlled environments. *Plant, Cell & Environment* **45**: 2410–2427.
Hughes ME, Abruzzi KC, Allada R, Anafi R, Arpat AB, Asher G, Baldi P, de Bekker C, Bell-Pedersen D, Blau J, *et al.*2017. Guidelines for Genome-Scale Analysis of Biological Rhythms. *Journal of Biological Rhythms* **32**: 380–393.
Hughes ME, Hogenesch JB, Kornacker K. **2010**. JTK_CYCLE: An Efficient Nonparametric Algorithm for Detecting Rhythmic Components in Genome-Scale Data Sets. *Journal of Biological Rhythms* **25**: 372–380.
Hutchison AL, Allada R, Dinner AR. **2018**. Bootstrapping and Empirical Bayes Methods Improve Rhythm Detection in Sparsely Sampled Data. *Journal of biological rhythms* **33**: 339–349.
Imaizumi T, Schultz TF, Harmon FG, Ho LA, Kay SA. **2005**. FKF1 F-Box Protein Mediates Cyclic Degradation of a Repressor of CONSTANS in Arabidopsis. *Science* **309**: 293–297.
Iwayama K, Aisaka Y, Kutsuna N, Nagano AJ. **2017**. FIT: statistical modeling tool for transcriptome dynamics under fluctuating field conditions. *Bioinformatics* **33**: 1672–1680.
Jung J-H, Barbosa AD, Hutin S, Kumita JR, Gao M, Derwort D, Silva CS, Lai X, Pierre E, Geng F, *et al.*2020. A prion-like domain in ELF3 functions as a thermosensor in Arabidopsis. *Nature* **585**: 256–260.
Kerwin RE, Jimenez-Gomez JM, Fulop D, Harmer SL, Maloof JN, Kliebenstein DJ. **2011**. Network Quantitative Trait Loci Mapping of Circadian Clock Outputs Identifies Metabolic Pathway-to-Clock Linkages in Arabidopsis. *The Plant Cell* **23**: 471–485.
Kim S-K, Yun C-H, Lee JH, Jang YH, Park H-Y, Kim J-K. **2008**. OsCO3, a CONSTANS-LIKE gene, controls flowering by negatively regulating the expression of FT-like genes under SD conditions in rice. *Planta* **228**: 355–365.
Kooyers NJ, Greenlee AB, Colicchio JM, Oh M, Blackman BK. **2015**. Replicate altitudinal clines reveal that evolutionary flexibility underlies adaptation to drought stress in annual Mimulus guttatus. *New Phytologist* **206**: 152–165.
Laosuntisuk K, Elorriaga E, Doherty CJ. **2023**. The Game of Timing: Circadian Rhythms Intersect with Changing Environments. *Annual Review of Plant Biology* **74**: 511–538.
Law CW, Alhamdoosh M, Su S, Dong X, Tian L, Smyth GK, Ritchie ME. **2018**. RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. *F1000Research* **5**.
Law CW, Chen Y, Shi W, Smyth GK. **2014**. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biology* **15**: R29.
Ledger S, Strayer C, Ashton F, Kay SA, Putterill J. **2001**. Analysis of the function of two circadian-regulated CONSTANS-LIKE genes. *The Plant Journal* **26**: 15–22.
Matsubara S. **2018**. Growing plants in fluctuating environments: why bother? *Journal of Experimental Botany* **69**: 4651–4654.
Matsuzaki J, Kawahara Y, Izawa T. **2015**. Punctual Transcriptional Regulation by the Rice Circadian Clock under Fluctuating Field Conditions. *The Plant Cell* **27**: 633–648.
Nagano AJ, Sato Y, Mihara M, Antonio BA, Motoyama R, Itoh H, Nagamura Y, Izawa T. **2012**. Deciphering and Prediction of Transcriptome Dynamics under Fluctuating Field Conditions. *Cell* **151**: 1358–1369.
Nakamichi N. **2020**. The Transcriptional Network in the Arabidopsis Circadian Clock System. *Genes* **11**: 1284.
Nusinow DA, Helfer A, Hamilton EE, King JJ, Imaizumi T, Schultz TF, Farré EM, Kay SA. **2011**. The ELF4–ELF3–LUX complex links the circadian clock to diurnal control of hypocotyl growth. *Nature* **475**: 398–402.
Paajanen P, Lane de Barros Dantas L, Dodd AN. **2021**. Layers of crosstalk between circadian regulation and environmental signalling in plants. *Current Biology* **31**: R399–R413.
Panter PE, Muranaka T, Cuitun-Coronado D, Graham CA, Yochikawa A, Kudoh H, Dodd AN. **2019**. Circadian Regulation of the Plant Transcriptome Under Natural Conditions. *Frontiers*

in Genetics **10**.
Parsons R, Parsons R, Garner N, Oster H, Rawashdeh O. **2020**. CircaCompare: a method to estimate and statistically support differences in mesor, amplitude and phase, between circadian rhythms. *Bioinformatics* **36**: 1208–1212.
Poorter H, Fiorani F, Pieruschka R, Wojciechowski T, van der Putten WH, Kleyer M, Schurr U, Postma J. **2016**. Pampered inside, pestered outside? Differences and similarities between plants growing in controlled conditions and in the field. *New Phytologist* **212**: 838–855.
Ripoll JJ, Rodríguez-Cazorla E, González-Reig S, Andújar A, Alonso-Cantabrana H, Perez-Amador MA, Carbonell J, Martínez-Laborda A, Vera A. **2009**. Antagonistic interactions between Arabidopsis K-homology domain genes uncover PEPPER as a positive regulator of the central floral repressor FLOWERING LOCUS C. *Developmental Biology* **333**: 251–262.
Rubin MJ, Brock MT, Davis AM, German ZM, Knapp M, Welch SM, Harmer SL, Maloof JN, Davis SJ, Weinig C. **2017**. Circadian rhythms vary over the growing season and correlate with fitness components. *Molecular Ecology* **26**: 5528–5540.
Sawa M, Nusinow DA, Kay SA, Imaizumi T. **2007**. FKF1 and GIGANTEA Complex Formation Is Required for Day-Length Measurement in Arabidopsis. *Science* **318**: 261–265.
Simon S, Rühl M, de Montaigu A, Wötzel S, Coupland G. **2015**. Evolution of CONSTANS Regulation and Function after Gene Duplication Produced a Photoperiodic Flowering Switch in the Brassicaceae. *Molecular Biology and Evolution* **32**: 2284–2301.
Song YH, Kubota A, Kwon MS, Covington MF, Lee N, Taagen ER, Laboy Cintrón D, Hwang DY, Akiyama R, Hodge SK, *et al.*2018. Molecular basis of flowering under natural long-day conditions in Arabidopsis. *Nature Plants* **4**: 824–835.
Talamanca L, Naef F. **2020**. How to tell time: advances in decoding circadian phase from omics snapshots.
Thines B, Harmon FG. **2010**. Ambient temperature response establishes ELF3 as a required component of the core Arabidopsis circadian clock. *Proceedings of the National Academy of Sciences* **107**: 3257–3262.
Townsley BT, Covington MF, Ichihashi Y, Zumstein K, Sinha NR. **2015**. BrAD-seq: Breath Adapter Directional sequencing: a streamlined, ultra-simple and fast library preparation protocol for strand specific mRNA library construction. *Frontiers in Plant Science* **6**.
Ueda HR, Chen W, Minami Y, Honma S, Honma K, Iino M, Hashimoto S. **2004**. Molecular-timetable methods for detection of body time and rhythm disorders from single-time-point genome-wide expression profiles. *Proceedings of the National Academy of Sciences* **101**: 11227–11232.
Van Bel M, Silvestri F, Weitz EM, Kreft L, Botzki A, Coppens F, Vandepoele K. **2022**. PLAZA 5.0: extending the scope and power of comparative and functional genomics in plants. *Nucleic Acids Research* **50**: D1468–D1474.
Wang X, Xu Y, Zhou M, Wang W. **2021**. Assessing Global Circadian Rhythm Through Single-Time-Point Transcriptomic Analysis. In: MUKHTAR S, ed. Methods in Molecular Biology. Modeling Transcriptional Regulation: Methods and Protocols. New York, NY: Springer US, 215–225.
Young MD, Wakefield MJ, Smyth GK, Oshlack A. **2010**. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biology* **11**: R14.

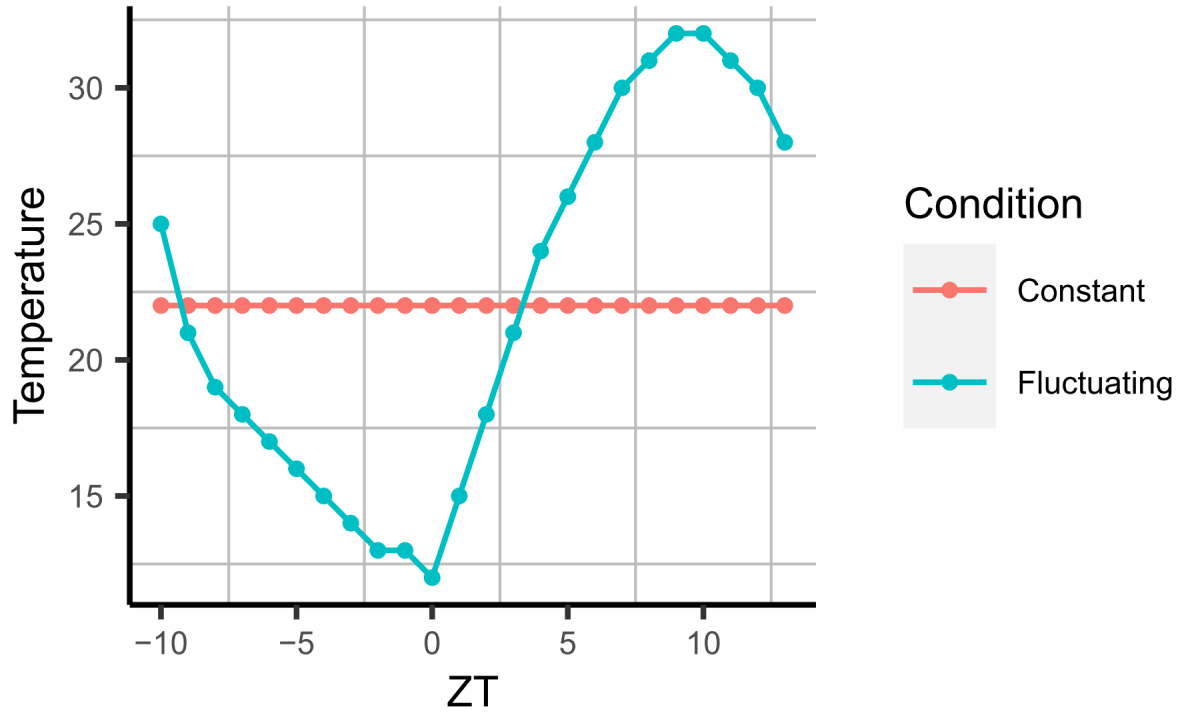

Fig. S3.1: Temperature profile of the constant (salmon) and fluctuating (cyan) temperature conditions.

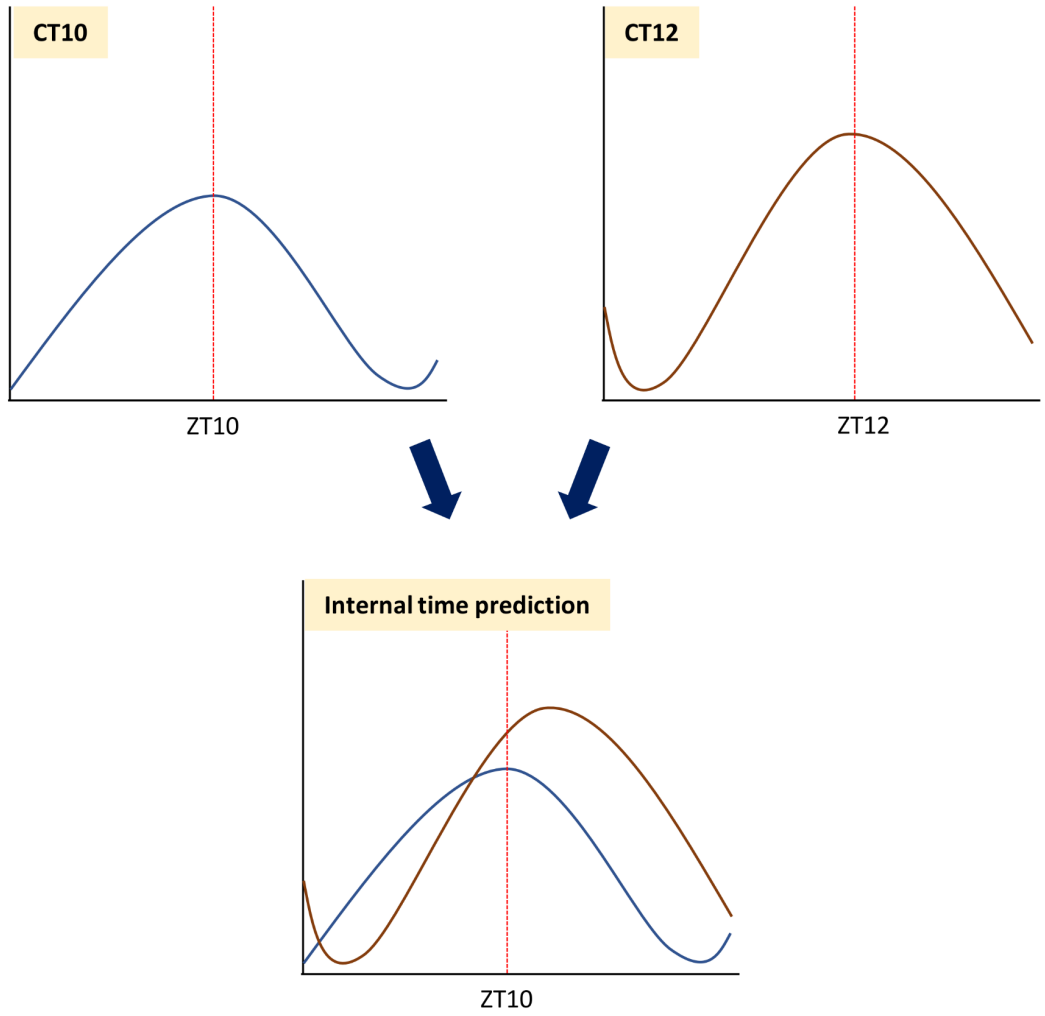

Fig. S3.2: Possible issue of internal time predictions caused by the amplitude differences between CT groups. As depicted in the upper panels, the CT groups exhibit their expression peaks at the corresponding ZT time points. For instance, CT10 exhibits the maximum expression at ZT10. Note the amplitude of CT12 is higher than CT10 in this case. For the samples collected at ZT10, the internal times are predicted as the highest expression among CT groups (as shown in the lower panel). However, due to the amplitude differences between CT10 and CT12, the internal time of those samples may be incorrectly predicted as ZT12 instead of ZT10 since the expression of CT12 is higher than CT10 at ZT10.

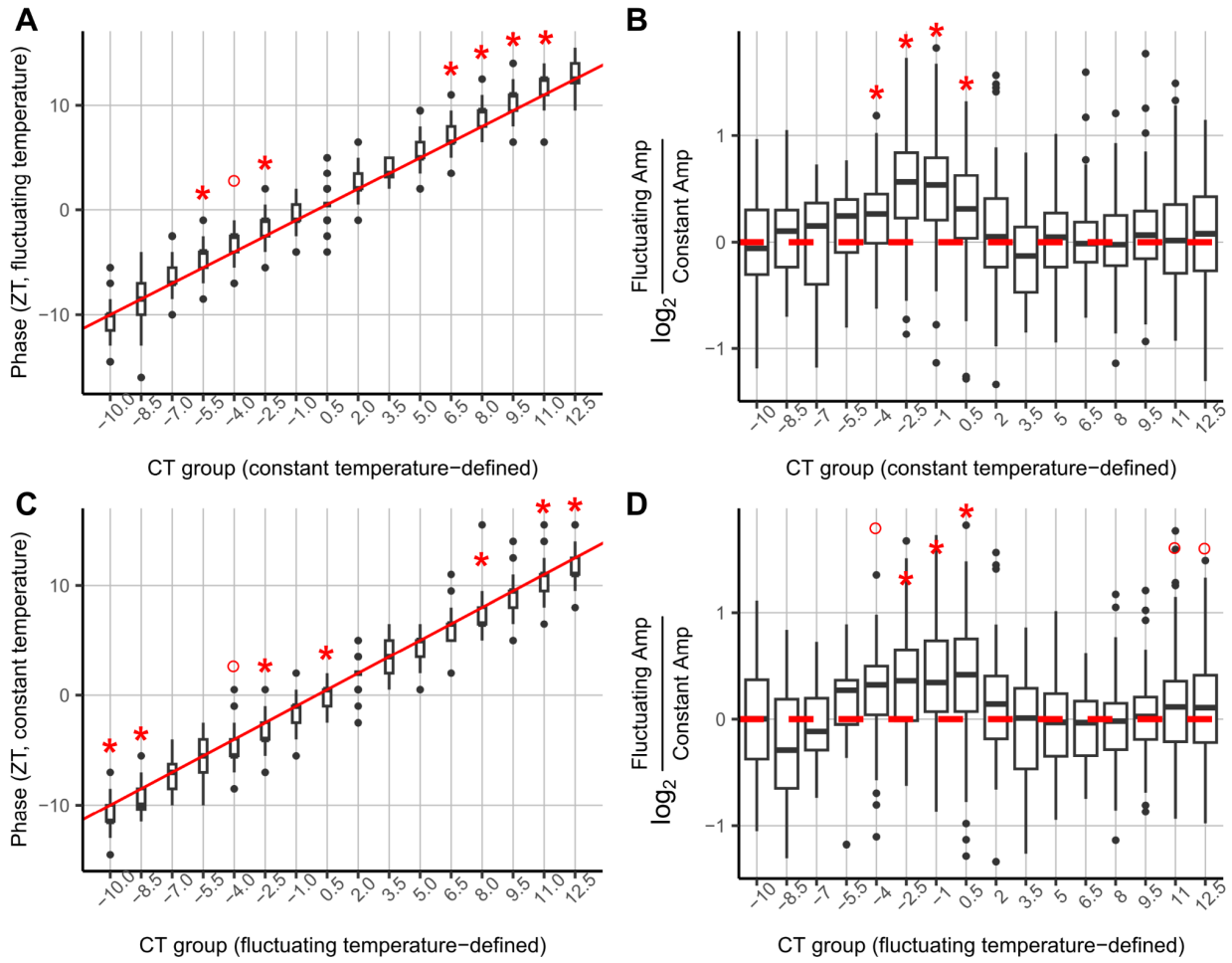

Fig. S3.3: Investigation of the phase (A and C) and amplitude (B and D) of individual cycling genes within the same CT group reveals the global influence of diel fluctuating temperature. Panels A and B display results using CT groups derived under the constant temperature condition, while Panels C and D show results using CT groups derived under the diel fluctuating temperature condition. In the phase analyses (A and C), the red diagonal lines represent an exact match of the expected CT group peak times and the peak times derived from the tested dataset (diel fluctuating temperature dataset for A and constant temperature dataset for C). Delayed peak times compared to the corresponding CT groups are indicated by the boxes above the diagonal line. In the amplitude analyses, the ratio of the amplitude observed under the diel fluctuating temperature condition to that observed under the constant temperature condition is depicted on a log2 scale. Positive deviations from zero (red dashed lines) indicate larger amplitudes under the diel fluctuating temperature condition. Stars represent significant differences in phase or amplitude from the expected value (peak times of the corresponding CT groups for phase analyses, and 0 for amplitude analyses) using one-sample t-tests with Bonferroni corrections ($p < 0.05$), while circles represent observations with $0.05 < p < 0.1$.