# Connecting the dots for flowering time genes in wheat

## Adam Gauley

A thesis submitted to the University of East Anglia for the degree of Doctor of Philosophy

John Innes Centre
October 2020

# Abstract

There is an urgent need to increase food security. The world's population is growing, the climate is changing, and yet the annual gains in crop yields are plateauing. To meet the demands of the future, we must take new approaches to improve crop productivity. Plants integrate seasonal progression in daylength and temperature to determine the optimal time to flower and set seed. However, in wheat, we understand very little about this process. The overall aim of this thesis is to understand how the leaf and developing inflorescence of bread wheat detects and responds to the changing seasons, and to investigate crosstalk between these tissues.

Using lines containing variant alleles for the key photoperiod gene, *Photoperiod-1* (*Ppd-1*), I analysed the molecular processes controlling flowering in the field. I find discrete photoperiod changes cause a step-wise increase in the transcription of *FLOWERING LOCUS T1* (*FT1*) as the major floral activator. This seasonal induction is partially regulated by *Ppd-1*, which dynamically responds to changes in daylength to control the rate of inflorescence development in a 'checkpoint' dependent manner. Photoperiod insensitive alleles of *Ppd-1* override this step-wise increase in *FT1* expression, resulting in accelerated inflorescence development. Within the developing inflorescence, these leaf-derived signals have a powerful influence over gene expression, with *Ppd-1* allelism altering gene expression patterns, amplitude and genome biases. *Ppd-1* mediated inflorescence development involves many genes, with large clusters of gene expression focused to each key developmental stage. Investigating the genes involved in these transitions has revealed four previously uncharacterised genes that help regulate inflorescence development. In addition, temperature can influence the rate of these stage transitions, likely through leaf- and inflorescence-based pathways.

This research has expanded our understanding of how wheat regulates flowering, providing a strong foundation to increase yield by fine-tuning photoperiod-depended control over spikelet and floret development.

# Acknowledgements

# Table of contents

# Table of figures

# Table of tables

# List of abbreviations

| | |
|---|---|
| aa | Amino acid |
| ARF | AUXIN RESPONSE FACTOR |
| BLAST | Basic Local Alignment Search Tool |
| bp | Base pair |
| C | Cluster |
| CRISPR | Clustered Regularly Interspaced Short Palindromic Repeats |
| Cas9 | CRISPR associated protein 9 |
| cv | Cultivar |
| DNA | Deoxyribonucleic acid |
| DR | Double ride |
| F | Field |
| FAC | Floral activation complex |
| FP | Floret primordium |
| FRC | Florigen repression complex |
| GA | Green anther |
| GFP | Green fluorescent protein |
| GH | Glasshouse |
| GO | Gene ontology |
| GP | Glume primordium |
| H | Hours |
| IM | Inflorescence meristem |
| KASP | Kompetitive allele-specific PCR |
| kb | Kilobase |
| LP | Lemma primordium |
| LP | Long days |
| MB | Moving bench |
| NC | Non-clustered |
| NIL | Near-isogeneic lines |
| Ppd-1 | Photoperiod-1 |
| PRR | Pseudo-response regulator |
| qPCR | Quantitative Polymerase Chain Reaction |
| RNA | Ribonucleic acid |
| RNA-seq | RNA-sequencing |
| SAM | Shoot apical meristem |
| SD | Short days |
| SEM | Standard error mean |
| TGW | Thousand grain weight |
| TILLING | Targeting Induced Local Lesions In Genomes |
| TPM | Transcripts per million |
| TS | Terminal Spikelet |
| VG | Vegetative |
| WA | White anther |
| WT | Wild-type |
| ZT | Zeitgeber Time |
| °C | Degrees Celsius |
| TSD | Time since dawn |

# Chapter 1    General introduction

## 1.2 **Wheat as a crop, and its importance for humans**

In 1798 Thomas Malthus proposed that population growth is exponential, whereas the growth of food supply is linear. It is therefore inevitable that population growth will outpace food production resulting in widespread famine until the population is returned to sustainable levels. It is the duty of scientists to prevent this prediction from coming true, to provide for the growing population and avoid a Malthusian catastrophe (Malthus, 1798). The world's population is predicted to reach 9.7 billion by 2050, which combined with a worldwide increase in living standard, is expected to see wheat demand rise by 60% (United Nations, 2015). Over half of the world's calories come from three main crops: rice, wheat and maize. Developing the productivity of these will help feed this growing population (IDRC, 2017).

Wheat is the most widely grown crop in the world (FAO, 2015). On average, we eat 50 wheat plants per day, providing one-fifth of all calories consumed by humans. It is an extremely nutritious and efficient food source, containing 327 calories and 14 grams of protein per 100 grams, providing more protein globally than all meats combined (Wheat Initiative, 2017). In the late 20th century, significant wheat yield gains were achieved through the discovery of dwarfism traits and improved agricultural techniques, such as the use of nitrogen-based fertilisers, in what is known as the "green revolution" (Borlaug *et al.*, 1969). However, the yield increases led by these advances are stagnating. For humanity to meet the nutritional needs of the future, a second green revolution must occur, and genetic approaches will spearhead this (Pingali, 2012).

## 1.3 **Modern wheat is an allopolyploid**

Hexaploid bread wheat is an allopolyploid plant generated by the hybridisation of three genomes that occurred through at least two events. Each of these events caused the addition of a genome in the resulting hybrid (Figure 1.1). The first is thought to have occurred between *Triticum urartu* (A genome) and *Aegilops speltoides* (B genome), circa 100,000 to 150,000 years ago to produce tetraploid wheat. A second hybridisation event 10,000 years ago between tetraploid wheat and *Triticum tauschii* (D genome) produced hexaploid wheat, *Triticum aestivum* (Figure 1.1) (Salamini *et al.*, 2002; Petersen *et al.*, 2006). Due to wheat's three homoeoalleles of each gene, there is much genetic redundancy. This, however, presents advantages in addition to the challenges such complexity creates. For example, it is possible to obtain variations in gene expression by knocking out specific numbers of alleles. In the past 5 years, there has been significant

development of genomic tools for wheat, including a fully-sequenced reference genome, speed breeding and development of a sequenced TILLING (Targeting Induced Local Lesions In Genomes) mutant resource (Adamski *et al.*, 2020). In addition, Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/ CRISPR associated protein 9 (Cas9) gene editing is becoming faster and more efficient (Debernardi *et al.*, 2020). These techniques combined with a well annotated genome are revolutionising the speed and quality of science in this field.
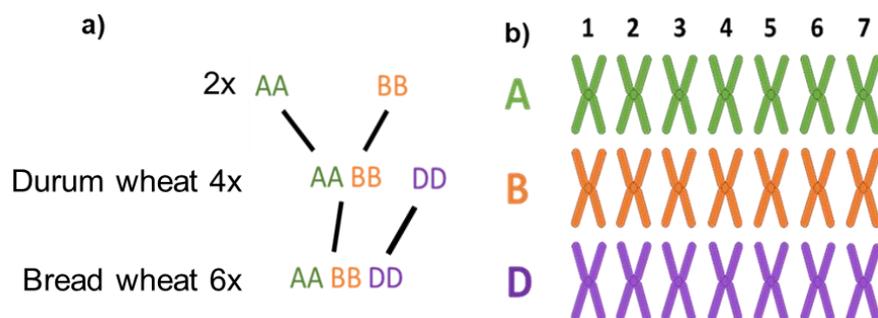


**Figure 1.1 - Genetic evolution of hexaploid wheat**

**a)** The first hybridisation event occurred between *T.urartu* (A genome) and *A.speltoides* (B genome). A second hybridisation event between tetraploid wheat and *T.tauschii* (D genome) produced hexaploid wheat *T.aestivum*. **b)** The allopolyploid makeup of the *T.aestivum* genome consisting of 21 pairs of wheat chromosomes, 7 from each genome.

## 1.4 **Floral induction is controlled by photoperiod**

The detection of photoperiod length is vital for timing key developmental processes, such as flowering, within a seasonal context. In addition to flowering, the signals induced by photoperiod play a key role in the production of grain in wheat, barley, rice and maize. Thus, enhanced regulation of flowering provides an attractive target for increasing yields. This thesis will focus on bread wheat (*Triticum aestivum*) which is considered a long-day plant, i.e. it requires long daylength photoperiods to flower and aim to understand the connection between flowering signals produced in the leaf with the developmental events that occur within the inflorescence. Compared to other cereals, such as rice and maize, little is known about the genetic mechanisms of flowering and inflorescence development in wheat.

In the primary plant model system, *Arabidopsis thaliana*, research has shown that floral meristem development is regulated by a complex gene network. When *Arabidopsis*

senses environmental changes, such as rising temperatures and an extending photoperiod, a pathway is initiated that transmits a signal from the leaf to the meristem in the form of FLOWERING LOCUS T (FT) protein (Andrés & Coupland, 2012). In *Arabidopsis*, the external coincidence model explains how light signals are integrated with the clock to regulate flowering. Light stabilises CONSTANS (CO) protein, allowing it to induce the expression of *FT.* The circadian clock oscillates the expression of *CO*, resulting in a diurnal expression with the peak in expression occurring in light only during long-days. Flowering is prevented during short days due to low expression of *CO* during light hours*,* in addition to degradation of CO protein in darkness (Suárez-López *et al.*, 2001; Yanovsky & Kay, 2002; Andrés & Coupland, 2012). FT then translocates from the leaf to the meristem. Transport of a florigen from the leaf to the shoot apical meristem (SAM) was first theorised through grafting experiments (Chailakhyan, 1937). More recently, the florigen was characterised as FT (Corbesier *et al.*, 2007; Tamaki *et al.*, 2007; Jaeger *et al.*, 2007; Mathieu *et al.*, 2007). In the SAM, FT triggers the expression of central floral integrator genes, including *SUPPRESSOR OF OVEREXPESSION OF CONSTANS1* (*SOC1*), and *LEAFY* (*LFY*) (Kardailsky *et al.*, 1999). With the advance of genomics, it has become possible to use the work carried out in model plants such as *Arabidopsis* to translate outcomes into staple crops, such as wheat. In wheat, the pathway varies from the model organism; however, there are some key similarities. Like in *Arabidopsis*, it is predicted that FT1 is the mobile signal that connects the leaf with the inflorescence meristem, which is under the control of *Photoperiod-1* (*Ppd-1*). While wheat does have a *CONSTANS* (*CO*) homologue, its importance for the flowering pathway is thought to be greatly reduced (Peng *et al.*, 2015).

## 1.4.1 Flowering Locus T

FLOWERING LOCUS T (FT) is a phosphatidylethanolamine binding protein (PEBP) that is well conserved throughout plant species. Due to its importance, *FT* expression is tightly regulated. In many angiosperms including rice and *Arabidopsis*, FT is responsible for delivering the signal to flower to the floral meristem, inducing a cascade of developmental events including spikelet development, floret initiation and eventual spike emergence (Turner *et al.*, 2005; Jaeger *et al.*, 2007; Corbesier *et al.*, 2007).

In wheat, the PEBP family of floral promoting genes is particularly varied due to gene duplication. The genes of this family have roles in flowering time, spike development and seed dormancy. There are three main clades, *FT*-like genes, *TERMINAL FLOWER1*-like (*TFL1*-like) and *MOTHER OF FT AND TFL1* (*MFT*-like) (Kobayashi *et al.*, 1999). It is thought that the *MFT*-like genes are the ancestors of the other two subfamilies, and it is the duplication and diversification of *MFT* from basal plant species such as mosses that gave rise to *FT*-like and *TFL1*-like genes (Hedman *et al.*, 2009; Wang *et al.*, 2015). *MFT*-

*like* genes have documented roles in flowering time, spikelet number as well as germination. The other two subfamilies have antagonistic roles, relative to each other (Kobayashi *et al.*, 1999). *FT*-like genes induce flowering and inflorescence development, whereas the *TFL1*-like genes repress flowering through anti-florigen activity (Liu *et al.*,2019; Wickland and Hanzawa., 2015).

Throughout plant evolution, the manipulation of the *FT and TFL* related genes have been tailored to regulate plant development. In wheat, *FT* has undergone further gene duplication, and as a result, there are at least 5 *FT-like* genes in wheat, each with a copy on the A, B, and D genomes (Lv *et al.*, 2014; Dixon *et al.*, 2018). *FT1* is the primary florigen; however, several of the gene copies have diverged functionality. The actions of many of these copies are still being elucidated, but some have been characterised in cereals. In *Brachypodium distachyon*, barley and tetraploid wheat, *FT2* is expressed in the developing inflorescence, with *ft2* null lines flowering slightly later and producing more spikelets (Shaw *et al.*, 2019). In hexaploid wheat, *ft-b1* mutants show severely delayed inflorescence development, but promote seed germination, responding differently under varied temperatures (Dixon *et al.*, 2018). The other *FT* genes in wheat are as of yet uncharacterised; however, down-regulation of *FT1* by RNAi results in down-regulation of *FT2* and *FT5* in wheat and *FT2* and *FT4* in *Brachypodium* (Lv *et al.*, 2014). When *FT1* is overexpressed in wheat, *FT2* and *FT3* were up-regulated. Considering the difference in the functionality of *FT1* and *FT2*, the other *FT* genes present interesting areas for future study.

The flowering pathway in wheat is poorly understood relative to other crops; although, we can use other cereals as models to understand what the photoperiod pathway in wheat may look like. For example, the complexity of *FT* family expression, and its regulation, have been examined in rice (*Oryza sativa*) (Brambilla *et al.*, 2017). As it is currently understood, the rice flowering pathway revolves around two rice *FT* homologues named *HEADING DATE 3a (Hd3a)* and *RICE FLOWERING LOCUS T 1 (RFT1)* (Hayama *et al.*, 2003; Komiya *et al.*, 2008). Rice is a short-day plant with two florigens, and when it experiences a day length that falls under a critical threshold, *Hd3a* is induced, and under long days *RFT1* is expressed. Like in *Arabidopsis*, these florigens then translocate to the shoot apical meristem (Komiya *et al.*, 2009). These two proteins are essential to the flowering process, as demonstrated by knocking-out expression of both genes causing plants to remain eternally vegetative (Komiya *et al.*,2008; Tamaki *et al.,*2015; Tamaki *et al.*,2007). In the meristem, Hd3a and RFT1 create a complex with transcription factors FLOWERING LOCUS D 1 (FD1) and a 14-3-3 protein to form a floral activation complex. This complex then dimerises, which allows it to target DNA sequences (Zhao *et al.*,2015; Taoka *et al.*, 2011). Hd3a and RFT1 act to promote flowering, but also function to repress themselves in the leaves in a negative feedback loop by interacting with *Hd3a BINDING*

*REPRESSOR FACTOR1* (*HBF1*) and *HBF2*, which reduce *Hd3a* and *RFT1* expression to delay flowering. This balancing of activating and repressing complexes using the same core proteins illustrates how the flowering pathway has different layers of regulation and control, which is still being unravelled (Brambilla *et al.*, 2017).

## 1.4.2 Photoperiod-1

*Ppd-1* is a member of the pseudo-response regulator (PRR) family, which encode proteins that contain a CCT (CONSTANS, CONSTANS-like, TIMING OF CAB1 (TOC1)) and a pseudo-receiver domain (Turner *et al.*, 2005; Beales *et al.*, 2007). *Ppd-1* is the most important known regulator of flowering time in wheat. Gain-of-function insensitive alleles promote increased expression of *FT1*, creating a constitutive long-day response under controlled conditions, whereas *ppd-1* null lines delay flowering (Turner *et al.*, 2005; Beales *et al.*, 2007; Díaz *et al.*, 2012; Shaw, Turner and Laurie, 2012; Shaw *et al.*, 2013). In *Arabidopsis* there is no clear homolog of *Ppd-1*; however, there are several closely related *PRR* genes including *PRR9,PRR7* and *PRR5* that when mutated together convey very late flowering and photoperiod insensitivity (Nakamichi *et al.*, 2005).

In hexaploid wheat, there are three *Ppd-1* homoeologues, *Ppd-A1*, *Ppd-B1*, *Ppd-D1*. Allelic variants in the *Ppd-1* gene can greatly alter flowering time. Lines that carry *Ppd-1* photoperiod insensitive alleles in one or multiple genomes constitutively promote the expression of *FT1* and promote flowering, irrespective of photoperiod. There are two widely used insensitive lines, the *Ppd-D1a* near-isogeneic lines (NIL)*,* where photoperiod insensitivity is caused by a 2 kb deletion in the promoter of *Ppd-D1* (Beales *et al.*, 2007). Additionally, *Ppd-B1a* insensitive lines have insensitivity conferred through an increase in copy number (Díaz *et al.*, 2012; Bentley *et al.*, 2013). Genetic variation in these genes has contributed significantly to wheat breading throughout the 20th century (Worland, 1996). A greater understanding of mechanisms underlying photoperiod insensitivity could help adapt wheat to more varied growth conditions and to further optimise yields (Hunt *et al.*, 2015). For example, in Western Australia, fast-developing wheat lines that carry photoperiod insensitive alleles of *Ppd-1* are favoured. Photoperiod insensitivity and the corresponding early maturity allows the plant to escape the increased drought and heat conditions in the later stages of the season (Flohr *et al.*, 2015). However, photoperiod insensitive lines that develop faster have a decreased yield, relative to sensitive lines. This yield penalty is due to accelerated development through early inflorescence stages resulting in fewer spikelets and florets forming (Coventry *et al.*, 1993; Kirkegaard *et al.*, 2014; Boden *et al.*, 2015; Ochagavía *et al.*, 2018; Prieto *et al.*, 2018).

### 1.4.3 Pathways converge to dynamically regulate flowering

While *Ppd-1* is a key regulator of *FT1* expression in wheat*,* in *Arabidopsis* several genes play significant roles including *CO*. Under long days *CO* is regulated by light signals in addition to the circadian clock to induce the expression of *FT* (Suárez-López *et al.*, 2001; Yanovsky & Kay, 2002; Valverde *et al.*, 2004; Andrés & Coupland, 2012). In Arabidopsis the FT protein then translocates to the meristem where it forms a floral activating complex (FAC) with FD and 14-3-3 proteins (Jaeger *et al.*, 2007; Corbesier *et al.*, 2007; Mathieu *et al.*, 2007; Tamaki *et al.*, 2007; Zhao *et al.*,2015; Taoka *et al.*, 2011). It is likely that *FT1* in wheat fulfils a similar function translocating to the meristem where it interacts with FLOWERING LOCUS D-LIKE (FDL) and 14-3-3 proteins to form a complex that induces the expression of key meristem identity genes (Li & Dubcovsky, 2008; Li *et al.*, 2015). Combining these studies and those in *Arabidopsis* we would expect *co* null lines in wheat to have delayed flowering. However, a recent study has shown that knockouts in *CO1* and *CO2* in the A and B genomes of tetraploid wheat show accelerated flowering time in both *Ppd-1* sensitive and insensitive lines under long and short days (Shaw *et al.*, 2020). It is theorised that whilst *Ppd-1* and *CO* detect and respond to photoperiod independently, they function together to fine-tune the responses to photoperiod. This characterisation of *CO* in wheat has created two lines of thought for the regulation of photoperiod-mediated flowering (Shaw *et al.*, 2020). Either there are two separate photoperiod sensing mechanisms, through *Ppd-1* and *CO*. Alternatively, both of these pathways are under the control of a single mediator of flowering, possibly through *EARLY FLOWERING 3* (*ELF3*), that acts downstream of the photoreceptor PHYTOCHROME C (PHYC) during the night. Regardless, it is clear from this recent study that *CO* is not an inducer of flowering in wheat, and there is significant interplay between *Ppd-1* and *CO* (Shaw *et al.*, 2020)*.*

The most important environmental controls of flowering are photoperiod and vernalization; however, after these requirements have been satisfied variations in flowering time are influenced by factors known as *earliness per se* (*Eps*) effects. *Eps* genes has been shown to mediate these responses, and are generally considered to have small effects fine-tuning developmental patterns (Ochagavía *et al.*, 2019). In wheat, *ELF3* has been identified as the *Eps-A$^m$1* locus (Alvarez *et al.*, 2016). *ELF3* deletions exhibit advanced flowering (Zikhali *et al.*, 2016). *ELF3* is a member of the evening complex of the circadian clock, which has been shown in *Arabidopsis* to provide temperature responsiveness. A recent study has found that a polyglutamine (polyQ) repeat is embedded within a predicted prion domain (PrD), and that the length of this domain correlates directly with temperature responsiveness (Jung *et al.*, 2020). The characterisation of the role of *ELF3* and *CO* in wheat are very recent discoveries, demonstrating that major components of the flowering pathway are only now being discovered, displaying the importance of research

in this field. Understanding and improving both the photoperiod and temperature pathways will be essential to adapting wheat to an increasingly warming climate (Jacott & Boden, 2020).

The vernalization pathway has a central role in temperature-mediated control of flowering. *VERNALIZATION1* (*VRN1*) and *VERNALIZATION2* (*VRN2*) help mediate this process, functioning as activators and repressors of flowering, respectively. Variation of the vernalization pathway in wheat creates plants with winter and spring growth habits. These growth types are characterised by natural variation within *VRN1* and *VRN2* (Yan *et al.*, 2003, 2004; Danyluk *et al.*, 2003; Trevaskis *et al.*, 2003, 2006)*.* During prolonged periods of cold together with short daylengths, conditions prevalent in winter, the expression of *VRN2* drops, decreasing VRN2-mediated repression in *VRN1.* This consequently induces the expression of both *VRN1* and *FT1,* triggering the flowering response (Yan *et al.*, 2006; Hemming *et al.*, 2008; Li & Dubcovsky, 2008). A spring growth type is conferred in wheat and barley cultivars that possess *VRN1* alleles with either insertions or deletions in first intron or mutations in the *cis*-regulatory regions. These variations of *VRN1* result in expression without the need for a cold treatments (Yan *et al.*, 2003; Danyluk *et al.*, 2003; Trevaskis *et al.*, 2003). Promotion of *VRN1* expression and a consequential spring habit can also be conferred through deletion of *VRN2* (Yan *et al.*, 2004, 2006; Trevaskis *et al.*, 2006)*.* This mechanism has been well described; however, the role of temperature in flowering beyond the vernalization pathway is relatively understudied. Efforts to unravel the role of ambient temperature on flowering provides interesting possibilities for increasing yields (Jacott & Boden, 2020).
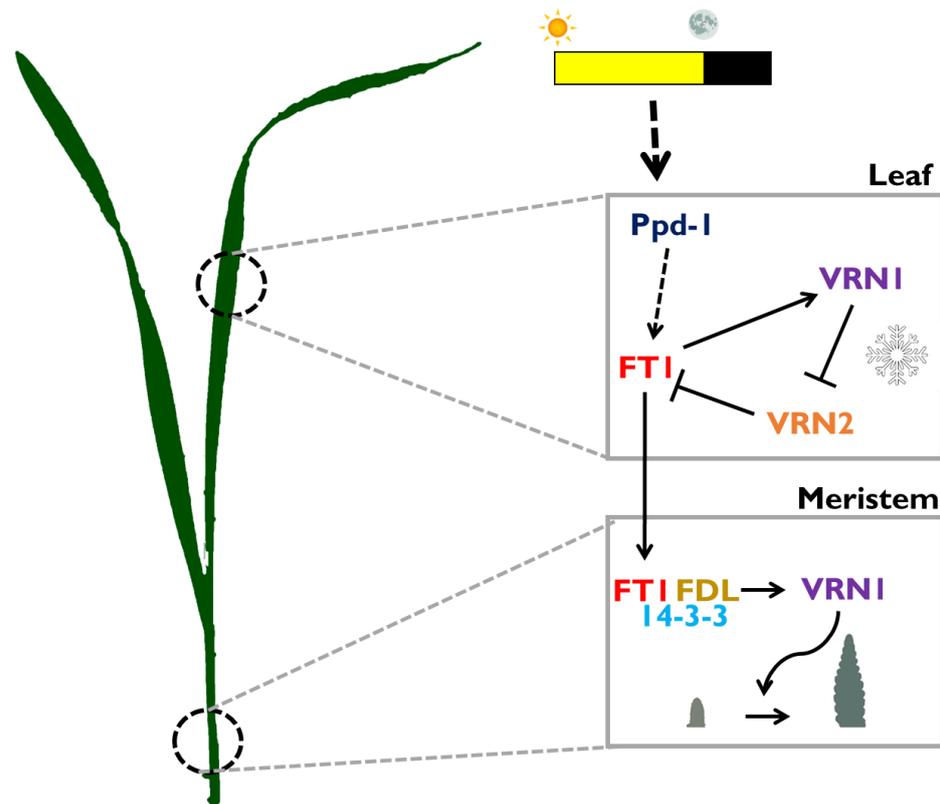
**Figure 1.2 - The flowering pathway in wheat**

A simplified model of the flowering pathway in wheat. Under long days, *Ppd-1* induces the expression of *FT1*. *VRN2* represses the expression of *FT1*, however after a period of prolonged cold *VRN1* repressed *VRN2* subsequently releasing the repression on *FT1*, consequently promoting the expression of *VRN1*. FT1 then translocates from the leaf to the meristem where it interacts with activator complex genes such as FDL and 14-3-3 to promote meristem identity genes such as VRN1, that mediate the floral transition (Chen *et al*., 2014).

## 1.4.4 Flowering in the field vs. controlled conditions.

Flowering occurs under inductive seasonal conditions; these conditions reflect gradual changes in both photoperiod and temperature that are difficult to replicate accurately in controlled conditions. To capture how wheat detects and responds to seasonal changes, this study will examine field grown plants and compare them to plants grown under glasshouse condition. A recent study highlights the differences in the expression of the *Arabidopsis* flowering time pathway between the field and laboratory growth chambers (Song *et al*., 2018). Differences in the expression profiles of *FT* were observed between the two conditions resulting from variations in R/FR light, in addition to fluctuations in the daily temperature. Field grown plants were more effective at inducing morning expression of *FT*. In the lab-grown conditions, there is commonly a small peak during the morning 4 hours after dawn (ZT 4) and a much greater peak in the evening at ZT 16 h; however, in

19

the field, this pattern is reversed (Song *et al.*, 2013; Song *et al.*, 2018). In the morning, *PHYA* and *ELF3* are primarily involved in the regulation of *FT,* with *phya* and *elf3* mutants showing altered morning expression patterns of *FT* (Song *et al.*, 2018).

A similar trend has been detected when examining the vernalization response between the field and controlled conditions in Brassicacea (O'Neill *et al.*, 2019). Typically, experiments investigating vernalization are carried out within a range of 5 °C–10 °C. This led to the conclusion that since temperatures are only maintained within this range during the winter months, it is during this period that the floral transition is repressed. However, a recent study in oilseed rape, observed that expression of *FLOWERING LOCUS C* (*FLC*)*,* which represses flowering in the Brassicacea*,* decreases during autumn when the air temperature is between 10 °C–15 °C. This results in the floral transition occurring before winter, with the plants overwintering as inflorescence meristems, before continuing development in spring (O'Neill *et al.*, 2019).

These examples demonstrate the importance of investigating flowering time genes under seasonal conditions in the field. Controlled conditions often do not adequately reflect the dynamic nature of environmental factors in the field.

## 1.5 **Inflorescence development**

One of the most important points in wheat development is the decision to progress from a vegetative to a floral state. As discussed, the timing of this transition is essential – too early and the developing floral meristem could succumb to frost damage, while flowering too late could mean grain-filling is reduced by drought and/or heat damage. The initial transition of the vegetative meristem to an inflorescence meristem marks the beginning of reproductive development, which includes the development of spikelets and florets. Much of the research into flowering and floral meristem development focuses on this key transition. However, beyond the transition, the plant must progress through several key early development stages at which important yield components such as spikelets and florets are formed (Kirby & Appleyard, 1984).

### 1.5.1 **The key structures of a developing inflorescence**

As the inflorescence meristem transitions from a vegetative stage to emergence, it undergoes significant morphological changes (Figure 1.3). Two studies are commonly used to define key wheat developmental stages, Kirby & Appleyard (1984) and Zadoks *et al.* (1974); the stages described in these studies are referred to extensively in this thesis. The first stage of wheat inflorescence development is the vegetative stage (VG) when the meristem characteristically has a dome shape and is surrounded by immature leaves. The

plant remains vegetative until signalling by *FT1* triggers the meristem to transition from making leaf primordia to spikelet primordia (Kirby & Appleyard, 1984; Li *et al.*, 2015).

Immediately after reception of the florigen signal, the meristem progresses to the next stage of development, known as the double ridge stage (DR). At DR, the primordia develop to a point where the spikelet primordia (top ridge) differentiate from the leaf primordia (bottom ridge), giving the characteristic double ridge patterning. Following DR is the glume primordium (GP) stage, at which point the glume primordia are formed. The GP stage is characterised by an enlargement of the spikelet primordium, such that it outgrows the leaf primordium to form an immature spikelet. This is the stage where wheat and barley inflorescence development begin to diverge. In wheat, the spikelet has several florets enclosed by glumes, whereas barley only has one floret. This difference between the species results in the GP stage in wheat having much larger spikelet primordia compared to its barley counterpart. This is also when the developing inflorescence gains its characteristic oval shape, with development of the inflorescence progressing in the central region first, before expanding outwards to the apex and base. The following stage of development is where we see rapid floral development occur. During the lemma primordium (LP) stage, floret primordium development is initiated. The first structures of these florets are the lemma primordia before the axillary spikelet meristems differentiate to form additional floral structures. The next developmental stage is the floret primordium stage (FP), when development of floret meristems at the axil of each lemma is visible. This stage can be defined by a tilting of the apical spikelet. The final stage of early inflorescence development is known as the terminal spikelet stage (TS), at which point all spikelets have been initiated and spikelet number is fixed – beyond this point, rapid floral maturation and elongation of the inflorescence occurs. This stage is defined by the formation of the terminal spikelet, which has clear floral organs.

Beyond early inflorescence development, there are several more defined developmental stages. These include: the white anther stage (WA), when floret initiation is complete, and glumes enclose the floret and the lemma enclose the white anthers; the green anther stage (GA), when the glumes almost cover the entire florets and the anthers are green. Beyond GA, the spike begins to boot, which is closely followed by flowering, defined by half the spike emerging from its leaf sheath or at anthesis.

Whilst the flowering and inflorescence pathways vary between *Arabidopsis*, tomato, maize, rice, barley and wheat, there are key characteristics and important gene families conserved throughout their evolution and divergence. For example, in barley (*Hordeum vulgare)*, the inflorescence is very similar to wheat in that it is unbranched and progresses through three main phases: the inflorescence meristem, spikelet meristem and the floral meristem (Kirby & Appleyard, 1984). The most striking difference between wheat and

barley is that the wheat spike is determinate, forming the terminal spikelet at the apex of the meristem. Conversely, the spikelet meristems are indeterminate and produce many florets. In barley, the opposite is true with the spike being indeterminate and not producing a terminal spikelet. Whereas, the barley spikelets are determinate, producing one floret each, but forming three spikelets at each rachis node (Youssef and Hansson., 2019). In rice, the panicle inflorescence forms branch meristems that produce primary and secondary branches. The primary and secondary branches produce lateral spikelet meristems. The spikelet meristem then gives rise to a single floral meristem, generating one grain-producing floret (Bommert, 2005; Wang *et al.*, 2008). These steps can be divided into key stages, the first is the inflorescence meristem specification, followed by the initiation of the branch meristem and sequentially the spikelet meristem that develops into the floral meristem (Yoshida & Nagato, 2011a). These stages correspond to very similar stages in wheat and barley, differing mainly in the lack of a branched meristem phase and likely share genetic components (Derbyshire & Byrne, 2013). The similarities of inflorescence development among the cereals indicate that we can learn much about the relatively unknown genetics underpinning the wheat inflorescence development through the existing studies in cereals and model organisms. However, the key differences, such as determinacy and lack of branching, suggest there will be unique gene pathways underpinning wheat inflorescence development.
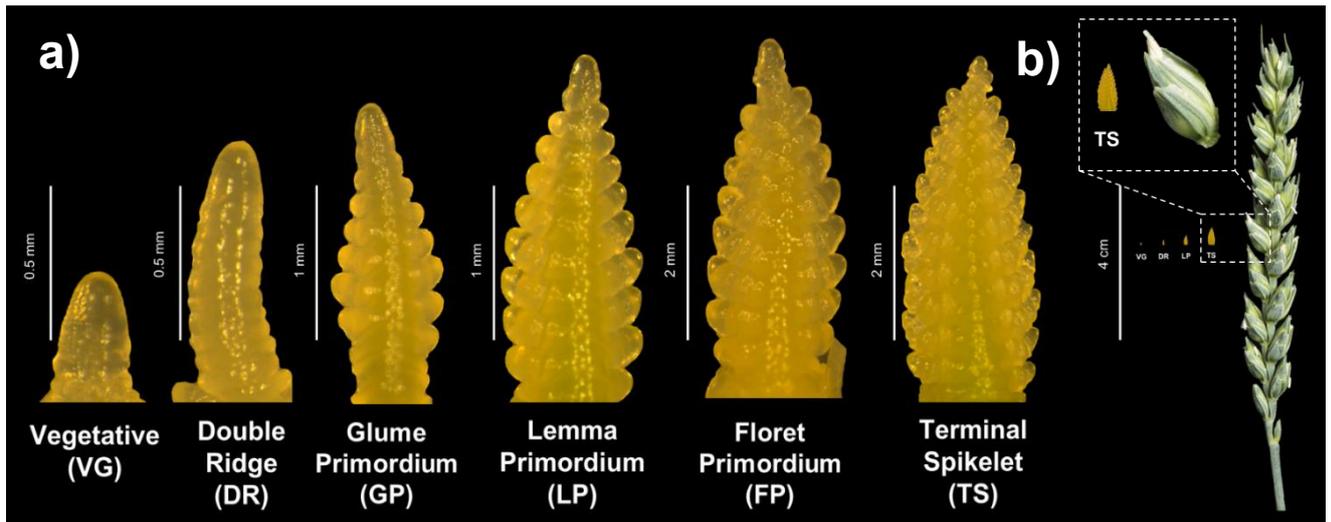
**Figure 1.3 - Wheat inflorescence development**

The early stages of the wheat inflorescence pathway, **a)** the vegetative (VG), double ridge (DR), glume primordium (GP), lemma primordium (LP), floret primordium (FP) and terminal spikelet (TS) stages. **b)** The inflorescence stages compared to a fully developed wheat spike (*cv.* Paragon). Scale is represented by a white bar.

## 1.5.2 The floral transition

In *Arabidopsis*, complex regulatory mechanisms in the leaf orchestrate the induction of *FT* and its subsequent translocation to the SAM (Jaeger *et al.*, 2007; Corbesier *et al.*, 2007; Mathieu *et al.*, 2007; Tamaki *et al.*, 2007). In the SAM, regulation of meristem identity genes by FT is also tightly controlled. FT forms a complex with the bZIP transcription factor, FLOWERING LOCUS D (FD). The FT/FD complex then interacts with 14-3-3 proteins to form the florigen activation complex (FAC). The FAC localises to the cell nucleus, with the 14-3-3 proteins acting as a scaffold for FT/FD interactions (Jaeger *et al.*, 2007; Corbesier *et al.*, 2007; Taoka *et al.*, 2011). The FAC then activates key meristem identity genes such as *APETLALA 1* (*AP1*) to trigger the transition of the meristem from a vegetative to floral state (Wigge *et al.*, 2005; Abe *et al.*, 2019).

Transcription factors are important regulators of gene expression. One of the largest transcription factor families in wheat are the MADS-box genes (Riechmann *et al.*, 2000; Schilling *et al.*, 2018). This family can be divided into two distinct groups known as type I and type II genes. The roles of Type I genes remain to be revealed. The type II MADS-box genes on the other hand have well-documented roles in many developmental processes (Schilling *et al.*, 2018). These type II MADS-box genes are also known as MIKC-type MADS-box genes due to their domain structure, consisting of the MADS, I, K and C-terminal domains (Kaufmann *et al.*, 2005). Many of the MICK-type MADS-box genes have

key roles during floral meristem development in a variety of species. In wheat, these genes are known to have had a major role in its domestication, in particular through the regulation of *VRN1*, itself a MADS-box gene (Danyluk *et al.*, 2003; Trevaskis *et al.*, 2003, 2006; Yan *et al.*, 2003, 2004).

The role of MADS-box genes in floral development is best characterised in *Arabidopsis*. *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1* (*SOC1*) is a MADS-box gene, and the first gene to be induced in the shoot apical meristem by the FAC (Lee *et al.*, 2000). The regulation of *SOC1* in *Arabidopsis* is complex. It is accepted that *SOC1* is up-regulated through the action of SQUAMOSA PROMOTER BINDING PROTEIN-LIKE3 (SPL3), SPL4 and SPL5 proteins which themselves are regulated by the FAC (Moon *et al.*, 2003; Wang *et al.*, 2009; Lee & Lee, 2010). There is also evidence for direct activation of *SOC1* by the FAC (Lee & Lee, 2010). Subsequently, *SOC1* can induce the expression of *LFY* that induces the expression of *AP1* (*VRN1*) and *FUL*, in addition to the *FAC* directly binding to the promoters of these genes (Moon *et al.*, 2005; Abe *et al.*, 2005; Wigge *et al.*, 2005; Teper-Bamnolker & Samach, 2005; Yoo *et al.*, 2005; Jung *et al.*, 2012; Collani *et al.*, 2019). This regulation of *SOC1* results in a cascade of transcriptional changes which eventually culminates in flowering.

The genes underpinning the vegetative to floral transition in crops are still be elucidated, although several studies in rice have been carried out to this end. In the rice SAM, the two florigens *Hd3a* and *RFT1* also act to regulate the floral transition (Tamaki *et al.*, 2007; Komiya *et al.*, 2009; Brambilla *et al.*, 2017). *Hd3a and RFT1* both interact with OsFD1 and 14-3-3 to create a FAC (Taoka *et al.*, 2011; Zhao *et al.*, 2015). The importance of MADS-box genes in the flowering and inflorescence developmental is also evident in rice with key similarities and differences to *Arabidopsis* (Komiya *et al.*, 2009). In rice, *OsMADS14, OsMADS15* and *OSMADS18,* all of which belong to the *APETALA1* (*AP1*)/ *FRUITFUL* (*FUL*)-like genes, in addition to *PAP2,* a *SEPELLATA* (*SEP*)-like gene, are activated by the FAC (K. Kobayashi *et al.*, 2012; Litt & Irish., 2003). Mutants in the *PAP2* gene do not undergo the floral transition, whereas all three of the *AP1*/*FUL*-like genes must be knocked down to impede the progression. These provide a promising target for altering spikelet development as the pathway is directly downstream of the rice photoperiod signals. However, in the leaves of the *pap2* mutant, there is a decline of the *Hd3a* and *RFT1* expression levels, which could indicate the *AP1*/*FUL*-like genes have unique roles in both the leaves in the meristems, similar to the role of wheat *VRN1* (Kobayashi *et al.*, 2012). In addition to its role in specifying the inflorescence meristem, *PAP2* has a role in promoting spikelet meristem identity (Kobayashi *et al.*, 2012). The multiple roles of *PAP2* highlight its function during rice inflorescence development and by analysing its role in wheat through its orthologue (*TaAGLG1*) could provide a greater understanding of inflorescence development and create opportunities for increasing yield.

A recent study explored the role of these genes in tetraploid wheat (Li *et al.*, 2019). The orthologs of *MADS14, MADS15* and *MADS18* are *VRN1, FUL2* and *FUL3* respectively. As expected from the characterised role of *VRN1* in the floral transition, *vrn1* null lines flower late, more so than individual *ful2* or *ful3* mutants. These genes have overlapping roles, with the spikelet meristem of the triple mutant remaining vegetative, whilst overexpression of any of these genes is sufficient to accelerate flowering. Studies in wheat have pointed to positive feedback loops regulating *VRN1* expression in the SAM to accelerate the floral transition. With the FAC, *VRN1* and *FT1* acting within a positive regulatory loop (Deng *et al.*, 2015; Li *et al.*, 2015).

## 1.5.3 Floral organ specification

Beyond the vegetative to floral transition, the SAM develops creating floral structures, a process critical to yield in crops. In *Arabidopsis*, floral organ identity, involving the MADS-box genes has been broken down into a stage-specific model (Murai *et al*, 2013). This is known as the ABCDE model for floral development (Coen & Meyerowitz, 1991; Murai, 2013). Each gene class functions in a specific layer of floral development named 'whorls' (Coen & Meyerowitz, 1991; Rijpkema *et al.*, 2010; Kitagawa *et al.*, 2012; Murai, 2013). A and E class protein complexes specify sepals in the first whorl, in the second whorl the A, B and E class specify the petals and in the third, B, C and E class proteins specify the stamens. The C and E class transcription factors function in the fourth whorl to specify carpels (Smaczniak *et al.*, 2012). Studies in *Arabidopsis* have assigned MADS-box genes to these families. *AP1* belongs to the A family (Jofuku *et al.*, 1994), *AP3* and *PISTILLATA* (*PI*) represent the B family (Jack *et al.*, 1992; Goto & Meyerowitz, 1994). The *AGAMOUS* (*AG*) genes embody C (Yanofsky *et al.*, 1990), D are *SHATTERPROOF1* (*SHP1*), *SHP2* and *SEEDSTICK* (*STK*) (Pinyopich *et al.*, 2003; Favaro *et al.*, 2003). Class E are *SEPALLATA1* (*SEP1*), *SEP2, SEP3* and *SEP4* and act to develop sepals, petals, stamens and carpels (Pelaz *et al.*, 2000; Ditta *et al.*, 2004). Whilst little is known about the role of the MADS-box genes in wheat, a recent study provides a genome-wide analysis of MIKC-type (type II) MADS-box genes (Schilling *et al.*, 2020). The study highlights the many duplications of these MADS-box genes, which are likely to be a product of wheat's adaptation to a variety of locations. 202 MADS-box genes were identified in wheat within the A, B and D genomes compared to just 54 in rice. A list of the wheat orthologue of all the known rice genes is shown in Table 1.1 (Schilling *et al.*, 2020). This highlights this gene family as an important candidate for detailed analysis as it is likely gene function has diverged from what we know largely from *Arabidopsis*, as well as other crops.

Like at the floral transition, there is currently a greater understanding of floral specification in rice relative to other cereals. Examining studies in rice to understand key processes such as spikelet meristem development is particularly useful due to the likely conserved

pathways amongst the cereals (Liu *et al.*, 2019). For example, *TAWAWA1* (*TAW1*) is an ALOG domain-containing rice gene that has a characterised role in inflorescence development. In a dominant gain of function mutant named *tawawa1-D,* the duration of inflorescence development is extended such that determination of spikelet number is delayed (Yoshida *et al.*, 2013). The mutant produces more spikelets and elongated branches. The role for *TAW1* in controlling spikelet meristem determinacy is supported by the effect it has on reducing expression of *OsMADS1* or *FRIZZY PANICLE* (*FZP*) as markers for spikelet meristem identity, and the accelerated development of panicles within *taw1* RNA-interference lines. MADS-box genes belonging to the *SHORT VEGETATIVE PHASE* (*SVP*) suppressor of flowering subfamily were shown to be up-regulated in *taw1* mutants. Other MADS genes such as *OsMADS7* (*SEP3*)*, OsMADS8* (*SEP3*)*, OsMADS16* (*AP3*)*, OsMADS4* (*PI*)*, OsMADS3* (*AG*)*, and OsMADS58* (*AG*) show down-regulation in the mutant line (Yoshida *et al.*, 2013). This work builds a pathway where downstream of the floral signal *TAW1* regulates inflorescence development through the action of SVP MADS-box genes.

In maize, few MADS-box transcription factors have been characterised to have roles in floral meristem development; however, the type II MADS-box *FUL-like* genes have been identified. By comparing plants in vegetative and floral states, two maize (*Zea mays*) *MADS-box* (*ZMM*) *ZMM4* and *ZMM15* were identified as two MADS-box genes belonging to the *FUL1* family that are up-regulated during the floral transition and are potentially involved in inflorescence development (Danilevskaya *et al.*, 2008). A second study showed that overexpression of *ZmCCT10* (CO, CONSTANS, *CO-LIKE* and *TIMING OF CAB1 10*) leads to a delay in switch from vegetative to reproductive floral meristem development, and vegetative reversion of the tassels to a branched leafy structure, which correlated with down-regulation of the *ZMM* meristem identity genes (Stephenson *et al.*, 2019).

Taken together, these studies in different cereals show that reduced expression of floral meristem identity MADS genes allows for more elaborately branched inflorescences to form.

**Table 1.1 - List of wheat genes with their rice orthologue.**

Wheat genes alongside their rice orthologues as annotated by (Schilling *et al.*, 2020)

| Wheat Gene name | Rice ortholog | Wheat Gene name | Rice ortholog |
|---|---|---|---|
| *TaAG-1* | *OsMADS58* | *TaFLC-1* | *OsMADS37* |
| *TaAG-2* | *OsMADS3* | *TaFLC-2* | *OsMADS51-2* |
| *TaAGL12-1* | *OsMADS26* | *TaFLC-3* | *OsMADS51-3* |
| *TaAGL12-2* | *OsMADS33* | *TaFLC-4-1* | *OsMADS51-4* |
| *TaAGL17-1* | *OsMADS61* | *TaMADS32-1* | *OsMADS32* |
| *TaAGL17-2* | *OsMADS25-1* | *TaMIKC-1* | *OsMADS68* |
| *TaAGL17-3* | *OsMADS57* | *TaMIKC-2* | *OsMADS62* |
| *TaAGL17-4* | *OsMADS25-2* | *TaPI-1* | *OsMADS4* |
| *TaAGL17-5-2* | *OsMADS25-3* | *TaPI-2* | *OsMADS2* |
| *TaAGL17-6-1* | *OsMADS59-1* | *TaSEP1-1* | *OsMADS1* |
| *TaAGL17-7* | *OsMADS59-2* | *TaSEP1-2* | *OsMADS1-1* |
| *TaAGL17-8* | *OsMADS59-3* | *TaSEP1-3* | *OsMADS1-2* |
| *TaAGL17-9* | *OsMADS59-4* | *TaSEP1-4* | *OsMADS5-1* |
| *TaAGL6-1* | *OsMADS6/OsMADS17* | *TaSEP1-5* | *OsMADS5* |
| *TaAP1-1* | *OsMADS14* | *TaSEP1-6* | *OsMADS34* |
| *TaAP1-2* | *OsMADS18* | *TaSEP3-1* | *OsMADS7* |
| *TaAP1-3* | *OsMADS15* | *TaSEP3-2* | *OsMADS8* |
| *TaAP3-1* | *OsMADS16* | *TaSOC1-1* | *OsMADS56-1* |
| *TaAP3-2* | *OsMADS16-2* | *TaSOC1-3* | *OsMADS50* |
| *TaBS-1* | *OsMADS29* | *TaSOC1-4* | *OsMADS50-2* |
| *TaBS-2* | *OsMADS31* | *TaSOC1-5* | *OsMADS56* |
| *TaBS-3* | *OsMADS30-1* | *TaSTK-1* | *OsMADS13* |
| *TaBS-4* | *OsMADS30-2* | *TaSTK-2* | *OsMADS21* |
| *TaBS-5-1* | *OsMADS30-3* | *TaSVP-1* | *OsMADS22* |
| *TaBS-6-1* | *OsMADS30-4* | *TaSVP-2* | *OsMADS55* |
| *TaBS-8* | *OsMADS30-6* | *TaSVP-3* | *OsMADS47* |
| *TaBS-9* | *OsMADS30-5* | | |

## 1.5.4 Transcriptome analysis of cereal inflorescence development

Efforts to unravel the genetic pathways underpinning inflorescence development in cereals have focused on barley, rice and maize. For example, the role of *Ppd-1* on inflorescence development, has been investigated in barley (Digel *et al.*, 2015). This study in barley is significant as it is one of the first to examine how altering genes primarily expressed in the leaf impact floral meristem development in cereals. It shows that an

extended inflorescence development correlates with increased yield, with the early phase of barley inflorescence development being particularly important. The number of floret primordia produced correspond to the final number of seeds, highlighting the importance of these early phases with useful marker genes identified. For example, MADS-box genes *Barley MADS1* (*BM1*) and *BM10* expression decreases throughout inflorescence development. Conversely, the expression of *KNOTTED1*, *SPL4* and *SOC1* increase throughout inflorescence development. It has also been shown that the expression of *SOC1*, *VRN-H1* and *BM3* correlates positively with the number of floret primordia formed throughout development (Digel *et al.*, 2015). The investigation suggests that much of gene expression changes associated with the early stages of inflorescence development including floral transition and floret primordia formation, were regulated largely independently of photoperiod. The group hypothesised that the role of *Ppd-1* in accelerating inflorescence development under long days comes instead through its up-regulation of *FT1*. The induction of *FT1* alters source-sink relationships such as nutrient transport, cell cycle regulation and carbohydrate metabolism at the floral transition stage, causing increased floret fertility. This supports the key role of photoperiod perception and flowering signals must play in inflorescence development whilst highlighting the need for a much greater understanding of the intricacies of inflorescence development.

A more recent study has investigated the transcriptome underpinning wheat inflorescence development (Feng *et al.*, 2017). Four stages of early wheat inflorescence development were examined, and based on comparisons to the characterisation of Kirby & Appleyard,. (1984), these stages correlate to the double ridge (DR), floret primordia stage (FP), late terminal spikelet (TS) and green anther stages (GA). Notably, this investigation found several hundred genes expressed specifically at each stage, highlighting the large transcription changes that occur as the inflorescence develops beyond the handful of genes currently characterised in wheat. The study also emphasizes the conserved gene expression patterns of several rice MADS-box genes in wheat, including *OsMADS22, OsMADS47* and *OsMADS55* of the SVP family, known to be involved in inflorescence branching and in repressing spikelet meristem identity (Liu *et al.*, 2013). These genes are expressed at the DR stage and lowly expressed at the subsequent stages.

Together these studies highlight how little we know about the genes underpinning wheat inflorescence development. Studies to understand this process focus largely on the MADS-box genes, despite hundreds of uncharacterised genes being differentially regulated between developmental stages.

## 1.5.5 Influence of key leaf genes on inflorescence architecture

Two key studies have highlighted the interaction between genes expressed in the leaf with *FT1* which go on to alter inflorescence architecture. As discussed, spikelet number, and therefore yield, has been altered by extending or delaying development time from vegetative stage to terminal spikelet. However, it is also possible to alter spikelet number by promoting the formation of 'paired spikelets', whereby two spikelets form at single node where there is commonly only a single spikelet (Boden *et al.*, 2015; Sharman, 1944; Yen *et al.*, 1992; Figure 1.4). Interestingly, *Ppd-1* and *FT1* influence the formation of these structures in wheat (Boden *et al.*, 2015). Photoperiod sensitive wheat NILs produce a high number of paired spikelets under short day conditions (12 h light/12 h dark) with 20% of rachis nodes producing paired spikelets, whereas NILs with a photoperiod insensitive allele do not. Under long days (16h light/8 h dark), reduced paired spikelet formation was reported in both lines, indicating that long days through the action of *Ppd-1* suppress the formation of paired spikelets. Similarly, *ppd-d1* null lines have an increased rate of paired spikelet formation under short days. These paired spikelet phenotypes are intricately linked with *FT1* expression levels, with reduced expression of *FT1* correlating with increased occurrences of paired spikelets. Furthermore, *ft-B1* null lines show increased rates of paired spikelet formation, proving that *FT1* is the link between *Ppd-1* mediated paired spikelet formation. In lines with reduced expression of *FT1*, there was a reduction in transcript level of key inflorescence development genes including *LFY, VRN1, AGLG1, AGL10* and *AGL29*, determined to be the likely cause of the altered inflorescence architecture (Boden *et al.*, 2015).

A second study in hexaploid wheat, identified a role for *TEOSINTE BRANCHED1* (*TB1*) in regulating wheat inflorescence architecture (Dixon, *et al.*, 2018; Figure 1.4). In maize, *TB1* is a major domestication gene where it is involved in concentrating resources to the main stem of the plant and suppression of axillary branches (apical dominance) (Doebley *et al.,* 1997). A dominant allele of *Tb1* confers the lack of branching in modern maize, compared to its highly branched progenitor, teosinte, which carries a *tb1* allele (Doebley *et al.,* 1997). *TB1* orthologs also have reported roles regulating inflorescence architecture and tiller number in rice and barley (Takeda *et al.*, 2003; Ramsay *et al.*, 2011). In wheat, *TB1* also promotes formation of paired spikelets in a dosage dependent manner. Increased expression of *TB1* correlates with a decrease in the expression of key meristem identity genes, similar to that seen in the *ppd-d1* null lines (Boden *et al.*, 2015). It is postulated that TB1 binds to FT1 in the floral meristem, reducing its availability for the FAC, thereby restricting its promotion of meristem identity genes (Dixon *et al.*, 2018).

In addition to the formation of paired spikelets, the inflorescence can be altered by promotion of spike branching. This occurrence produces a dramatic phenotype, whereby

spikelets are replaced by lateral branch-like structures that are essentially small secondary spikes (Poursarebani *et al.,* 2015). This phenomenon has been reported in wheat producing 'Miracle-Wheat', in addition to in barley 'Compositum-Barley'. The gene underpinning 'Miracle-wheat' was first identified in barley as *compositum 2* (*com2*). This gene is orthologous to the *branched head*$^t$ (*bh*$^t$) locus in tetraploid wheat, in addition to *BRANCHED SILKLESS 1* (*BD1*) in maize and *FRIZZY PANICLE/BRANCED FLORERLESS 1* (*FZP/BFL1*) in rice, all of which encode AP2/ERF transcription factors (Chuck *et al.*, 2002; Komatsu *et al.*, 2003; Zhu *et al.*, 2003). The function of this gene is highly conserved, with mutant alleles that disrupt the AP2 domain suppressing the ability of this protein to repress spikelet branching, which facilitates outgrowth of lateral branches. In barley, *COM2* expression begins during early spikelet development at the triple mound stage, when spikelet primordia start to differentiate. The orthologue in bread wheat has been identified as wheat *FRIZZY PANICLE* (*WFZP*) (Dobrovolskaya *et al.*, 2015; Figure 4).

These studies demonstrate that the developing inflorescence is capable of producing more elaborate branching patterns; however, these alternate arrangements of spikelets are normally suppressed genetically. Understanding the mechanisms underpinning these processes presents promising possibilities for releasing the repression and increasing yields (Gauley and Boden, 2019).
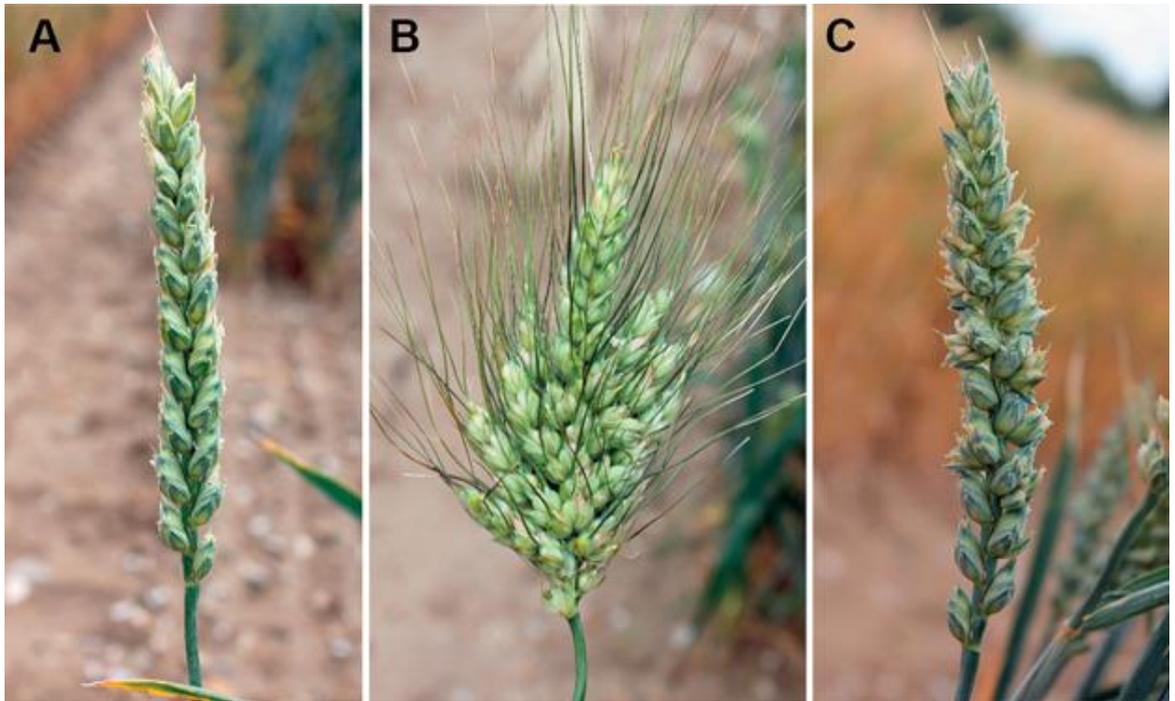
**Figure 1.4 - Inflorescence architecture phenotypes of wheat**

**(A–C)** Wheat inflorescences of **A**) a wild-type plant (*cv*. Cadenza), **B**) a 'Miracle wheat' plant and **C**) a Cadenza mutant line (CAD1290) that forms paired spikelets. Plants were grown under field conditions, Norwich, United Kingdom. Figure and legend from (Gauley and Boden, 2019).

## 1.6 **Summary**

At the start of my *PhD,* there were significant gaps in our understanding of the flowering pathway, particularly in wheat. The regulation of *FT1* in the leaf is well understood in *Arabidopsis* and rice; however, in wheat and barley the induction of *FT1* is controlled by *Ppd-1* (Suárez-López *et al.,* 2001; Yanovsky and Kay, 2002; Andrés and Coupland, 2012; Komiya *et al.*,2008; Tamaki *et al.,*2015; Tamaki *et al.*,2007; Shaw *et* al., 2012; Shaw *et al.*, 2013) We know little about how *Ppd-1* is regulated in the leaf, and how it interacts genetically with *FT1.* Photoperiod insensitive alleles of *Ppd-1* have been used extensively to adapt wheat to different growing seasons, but how *Ppd-1* integrates those seasonal cues, particularly under field conditions, is unknown (Beales *et al.*, 2007). The role of the MADS-box genes has been investigated in many plants, not only their role in the floral transition but also through floral organ specification as the inflorescence develops (Schilling *et al.*, 2018). We can use the knowledge gained in plants such as *Arabidopsis* and rice to inform us of the likely function of these genes in wheat. However, gene duplication, in addition to unique floral architecture, suggests that the regulatory mechanisms will differ in wheat. Wheat spikelet number is tightly maintained; despite this,

efforts to target individual genes demonstrates there is great potential to increase yields through targeted mutagenesis. This thesis will use transcriptomics in addition to targeted mutagenesis to connect the dots for flowering time genes in wheat.

## 1.7 **Thesis aims**

The overall aim of this thesis is to understand how the leaf and developing inflorescence detect and respond to the changing seasons, and to investigate the crosstalk between these two important tissues. I hypothesise that there is a significant regulation between the leaf and the developing inflorescence that is modulated by *Ppd-1.* By using *Ppd-1* allelism, I aim to reveal significant interaction between the leaf and the developing inflorescence, in addition to uncovering a complex regulatory landscape underpinning wheat inflorescence development. Specifically, this thesis looks at:

- The expression patterns of major flowering genes in the leaf and how they respond to changing seasonal conditions (**Chapter 2**).
- A comprehensive examination of the genetics underpinning wheat inflorescence development (**Chapter 3**).
- The influence of ambient temperature on inflorescence development (**Chapter 4**).
- A novel method for rapid identification of genes involved in inflorescence development (**Chapter 5**).

# Chapter 2    The impact of photoperiod

All results in this chapter apart from Ppd-1 localisation and *FT1* meristem expression have been published in the following manuscript (Appendix 1):

**Gauley, A., & Boden, S. A. 2020.** Step-wise increases in *FT1* expression regulate seasonal progression of flowering in wheat (*Triticum aestivum*). *New Phytologist.* https://doi.org/10.1111/NPH.16910

## 2.2 **Chapter summary**

In this chapter, I investigate the molecular signals coordinating flowering and inflorescence development under natural, field-based conditions. Using photoperiod insensitive and null *Ppd-1* lines, I find discrete photoperiod changes underpin the initiation of flowering and the control of yield-based traits. This occurs as days become longer from winter into spring and involves a step-wise increase in transcription of *FT1. Ppd-1* contributes to this seasonal induction, dynamically integrating environmental signals to mediate flowering. I find photoperiod insensitive *Ppd-D1a* lines over-ride this step-wise induction, resulting in accelerated inflorescence development and earlier flowering. In addition, the completion of spikelet formation is promoted by *FLOWERING LOCUS T2* (*FT2*)*,* which is activated by Ppd-1*.*

## 2.3 **Introduction**

Plants detect and respond to environmental signals such as temperature and daylength to optimally time key developmental processes such as flowering and setting seed. Cereals such as wheat and barley detect the cold temperatures of winter through vernalization in concert with the extending daylengths of spring so seed production can occur under the most optimal conditions (Worland *et al.*, 1998; Fjellheim *et al.*, 2014). Alleles that alter the plants' ability to respond to seasonal changes have been used by breeders to expand the geographical range of wheat and barley (Chouard, 1960; Limin & Fowler, 2006). These alleles typically accelerate flowering by limiting the need for either long days or temperatures, creating plants that are better suited to marginal growth environments with shorter growing seasons (Danyluk *et al.*, 2003; Trevaskis *et al.*, 2003; Yan *et al.*, 2003, 2004; Turner *et al.*, 2005; Beales *et al.*, 2007).

### 2.3.1 The role of *Ppd-1*

The responsiveness of wheat and barley to daylength (photoperiod), is largely determined by allelic diversity in *Photoperiod-1* (*Ppd-1*) (Laurie *et al.*, 1995; Turner *et al.*, 2005;

Beales *et al.*, 2007). *Ppd-1* is expressed in the leaf where it induces the expression of *FLOWERING LOCUS T1 (FT1)* a highly conserved gene among angiosperms (Turner *et al.*, 2005; Yan *et al.*, 2006; Beales *et al.*, 2007; Wilhelm *et al.,* 2009; Kitagawa *et al.*, 2012; Boden *et al.*, 2015; Bratzel & Turck, 2015). *FT1* is expressed in the leaf, but it is likely as in *Arabidopsis* and rice that the encoded protein translocates from the leaf to the shoot apical meristem where it triggers the vegetative to floral transition. It is believed that this transition is mediated through the formation of a complex with FLOWERING LOCUS D-like and 14-3-3 proteins, to activate the expression of key meristem identity genes, the expression of these genes allows the subsequent progression of the inflorescence meristem (IM) (Corbesier *et al.*, 2007; Komiya *et al.*, 2008; Li & Dubcovsky, 2008). Insensitive varieties of *Ppd-1* in hexaploid wheat can induce the expression of *FT1* irrespective of photoperiod, largely because of copy number variations in the B genome or through the deletion of the *cis*-regulatory element in the D genome, termed *Ppd-B1a* and *Ppd-D1a* alleles respectively (Beales *et al.*, 2007; Bentley *et al.,* 2011; Díaz *et al.*, 2012; Kitagawa *et al.*, 2012; Shaw *et al.*, 2012; Shaw *et al.*, 2013).

## 2.3.2 The *Ppd-1* NILs used in this study

The cultivar used in this study is 'Paragon' which contains copies of *Ppd-1* on each sub-genome. The *Ppd-A1* gene contains a 39-bp deletion at the predicted transcriptional start site resulting in decreased expression, whereas the *Ppd-D1* copy contains a 4.8-kb *mariner*-type transposon in intron 1 containing a spice site resulting in transcripts with frameshift mutations and premature stop codons (Beales *et al.*, 2007; Shaw *et al.*, 2012). The resulting protein is likely to be non-functional. The *Ppd-B1* homoeologue has no mutations and is predicted to be functional (Beales *et al.*, 2007; Wilhelm *et al.*, 2009; Shaw *et al.*, 2013).

This study uses a *Ppd-D1a* photoperiod insensitive NIL with a 2 kb deletion in its promoter region introgressed from Sonora64 into Paragon. The *Ppd-D1a* allele was introgressed into Paragon to the $BC_4F_2$ stage (Bentley *et al.*, 2011). Subsequently, these lines were developed through self-pollination to $BC_4F_3$ and $BC_4F_4$ (Shaw *et al.*, 2012). This study uses *Ppd-D1a* insensitive lines, as they conferred the strongest flowering phenotype; however, *Ppd-B1a* and *Ppd-A1a* lines were also generated resulting in faster flowering (Bentley *et al.*, 2011).

A triple *ppd-1* null line was also used in this study (Shaw *et al.*, 2013). Both the *Ppd-A1* and *Ppd-D1* null alleles were introgressed to 'Paragon' from 'Norstar' (Beales *et al.*, 2007; Shaw *et al.*, 2013). The *Ppd-A1* copy contains a 303 bp deletion that removes proportions of exon 5 and 6 in addition to a part of intron 5. This deletion causes a frameshift mutation predicted to cause a truncated protein lacking the CCT domain. The Norstar *Ppd-D1a*

allele also has a 5 bp deletion in exon 7 causing a frameshift mutation resulting in a lack of CCT domain, predicted to cause a non-functional protein (Beales *et al.*, 2007). A *Ppd-B1* null line was developed from a gamma irradiated 'Paragon' population whereby *Ppd-B1* was entirely deleted (Shaw *et al.*, 2013). Paragon lines with *Ppd-1* null copies on each genome were then generated through crossing. At the BC$_3$ stage plants heterozygous for the *ppd-D1* were crossed to plants heterozygous for *ppd-A1* deletions, progeny which were heterozygous for both the *ppd-D1* and *ppd-A1* deletions were then crossed to BC$_2$F$_2$ paragon plants homozygous for the *ppd-B1* deletion. Offspring that were homozygous for deletions in all three genome copies were then used in this study (Shaw *et al.*, 2013).

The *Ppd-1* alleles used in this study have come from several backgrounds but all have been introgressed and backcrossed into 'Paragon' to generate NILs which vary at the Ppd-1 loci. Therefore, 'Paragon' is used in this thesis as a control and will be referred to as wild-type. Future studies may want to examine how these NILs compare to the ancestral cultivars; however, for the scope and resources of this thesis only Paragon is used to facilitate comparisons between the insensitive and null lines.

The genotypes used are near isogenic lines backcrossed 4 times and therefore, if they followed a mendelian pattern of inheritance will be 93.75% Paragon. However, I cannot rule out that genes introgressed with *Ppd-1,* particularly those which could have been within the marker selection region on 2D, are responsible for some of the effects observed in this thesis. This is, however, unlikely due to other *Ppd-1* insensitive lines showing similar phenotypes of accelerated flowering and spike characteristics (Bentley *et al.*, 2013).

### 2.3.3 The impact of *Ppd-1* allelism on flowering

Photoperiod insensitivity does not accelerate flowering equally, with the ranking generally following the *Ppd-D1a* > *Ppd-B1a* > *Ppd-A1a* order of strength (González *et al.*, 2005; Díaz *et al.*, 2012; Shaw *et al.*, 2012; Bentley *et al.*, 2013; Pérez-Gianmarco *et al.*, 2019). This accelerated flowering and associated spike architecture phenotypes are most evident under short day conditions and is greatly reduced by growing under long days. Interestingly there is little benefit to combining photoperiod insensitivity alleles under long day growing conditions but under short-day conditions the photoperiod insensitivity is additive (Bentley *et al.*, 2013; Pérez-Gianmarco *et al.*, 2019). There are many different *Ppd-1a* alleles across cultivars and knowledge regarding how these alleles affect the progression of specific growth phases under different environmental conditions has been vitally important for breeders for tailoring new cultivars for photoperiod conditions (Whitechurch & Slafer, 2002; Bentley *et al.*, 2013; Pérez-Gianmarco *et al.*, 2019).

*Ppd-1* photoperiod insensitive alleles have been shown to accelerate flowering under field, glasshouse and controlled conditions. However, the majority of our understanding of the effects of *Ppd-1* on *FT1* expression come from controlled conditions that use extreme short (8-9 h) or long (15-16 h) daylengths. These artificial conditions do not adequately reflect the dynamic changes in photoperiod and temperature that occur in the field to mediate the vegetative to floral transition (Beales *et al.*, 2007; Wilhelm *et al.*, 2009; Díaz *et al.*, 2012; Kitagawa *et al.*, 2012; Shaw *et al.*, 2012; Shaw *et al.*, 2013; Boden *et al.*, 2015).

## 2.3.4 The circadian clock in wheat

In addition to photoperiod and temperature control, these pathways are intertwined with the circadian clock. In plants, the circadian clock is a network of genes that act together in a rhythmic pattern of gene loops, to create consistent patterns of gene expression throughout 24 hours (Block & Page, 1978; Ford *et al.*, 2016). This inbuilt clock rhythmically controls many biological processes (Harmer, 2009; Ford *et al.*, 2016). The clock can sense and respond to changing photoperiod and temperature to induce and regulate flowering pathways. In wheat, the clock is only beginning to be unravelled, however, the pathways appear to have significant conservation from model plants such as *Arabidopsis* (Calixto *et al.*, 2015). The key genes in the *Arabidopsis* circadian clock are the *PSEUDO RESPONSE REGULATORS* (*PRRs*), *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*) and *LATE ELONGATED HYPOCOTYL* (*LHY*), as well as *LUX ARRHYTHMO* (*LUX*) / *PHYTOCLOCK1* (*PCL1*), *EARLY FLOWERING 3* (*ELF3*), and *ELF4* (Schaffer *et al.*, 1998; Nakamichi *et al.*, 2005). These genes are fixed into a regulatory loop wherein the morning, *LHY* and *CCA1* are under sequential repression by *TIMING OF CAB EXPRESSION 1* (*TOC1*), *PRR9* and *PRR7*. In the evening, the *PRRs* and *TOC1* are repressed by the "evening complex", consisting of *EARLY FLOWERING 3* and 4 (*AtELF3*, *AtELF4)* and *LUXARRHYTHMIO* (*LUX*). The cycle is then reset by the action of *LHY* and *CCA1* that repress the evening complex in the morning (Pokhilko *et al.*, 2012). Whilst this pathway does appear to be largely conserved from *Arabidopsis* to crops such as barley, there appears to be a much greater impact of changing photoperiod on the clock, and responds differently to certain temperature cues (Griffiths *et al.*, 2009; Gawroński *et al.*, 2014; Deng *et al.*, 2015; Ford *et al.*, 2016; Zikhali *et al.*, 2016). For example, the circadian clock in barley will entrain the period of the clock to photoperiod conditions and requires a lights-on and lights-off signal to begin robust rhythmic gene expression (Deng *et al.*, 2015). Ppd-1 in barley and wheat is a PRR protein, and unlike in *Arabidopsis*, it is crucial in mediating the flowering response in long days (Beales *et al.*, 2007; Shaw *et al.*, 2012; Shaw *et al.*, 2013). We are only beginning to understand the key role the circadian clock plays in wheat flowering. In this chapter, I analyse how *Ppd-1* allelism influences and is influenced by the circadian clock diel gene expression

throughout the growing season. This will be key to fully understand flowering and subsequent inflorescence development.

## 2.3.5 Inflorescence development

Most of our understanding of *Ppd-1* regulation and its impact on the molecular processes underpinning flowering has come from studies performed in the leaves. We know little about the impact of insensitive and null alleles of *Ppd-1* on the genes that underpin meristem identity in the inflorescence. Recent studies have demonstrated the effects of *Ppd-1* allelism on yield-related components of inflorescence development, including floret fertility and spikelet architecture (González *et al.*, 2005; Boden *et al.*, 2015; González-Navarro *et al.*, 2015; Prieto *et al.*, 2018). In wheat, early inflorescence development is characterized by the formation of spikelets and the initiation of florets, from the floral transition until the formation of the terminal spikelet (Kirby & Appleyard, 1984). After the inflorescence receives the *FT1* mediated signal to flower, it progresses to the double ridge stage, at this stage leaf and spikelet primordia appear at each node. As the meristem develops, the leaf primordia are suppressed and the spikelet primordia enlarge through the glume and lemma primordium stages, whilst florets form. At the terminal spikelet stage, spikelets initiation is halted, defining the determinate inflorescence characteristic of wheat. How quickly the IM reaches the TS stage and progresses through the preceding stages reflects how many spikelets and florets will be produced. Understanding how these critical early stages of inflorescence are linked with leaf signals, will help inform new strategies for increasing yield (Gauley and Boden, 2019).

To advance our understanding of the seasonal regulation of flowering, I investigated the regulation of the photoperiod-dependent flowering pathway and early inflorescence development under the natural photoperiod transitions of a standard growing season. I used near-isogenic lines of *Ppd-1*, which included photoperiod insensitive, sensitive and null alleles, to understand how modified photoperiod responsiveness alters flowering and inflorescence development under field-based conditions. I also draw comparisons between the field and glasshouse conditions to understand how these different environments alter seasonal changes in gene expression. This work demonstrates and dissects the complexity of the flowering process under natural field-based conditions, presenting new genetic targets for the improvement of yield potential in wheat.

## 2.4 **Results**

### 2.4.1 Impact of *Ppd-1* on yield components

Previous studies have illustrated the effect of *Ppd-1* allelism on flowering and yield components (Shaw *et al.*, 2012; Shaw *et al.*, 2013; Ochagavía *et al.*, 2018). However, in this study, I compare the yield components in both the field and controlled conditions (Figure 2.1). Three genotypes were analysed in this study, a NIL carrying a photoperiod insensitive copy of *Ppd-1* on the D genome only (*Ppd-D1a*), a NIL carrying null alleles of *Ppd-1* on all three genomes (*ppd-1*) and wild-type Paragon (WT) (Beales *et al.*, 2007; Bentley *et al.*, 2011; Shaw *et al.*, 2012, 2013). The field was located in Norwich, Norfolk, UK and the controlled conditions consisted of an unheated and unlit glasshouse (referred to as natural conditions) located at the John Innes Centre, Norwich, Norfolk, UK. Seeds were planted in October, allowed to overwinter, and develop under a seasonal context. Under both conditions, the insensitive *Ppd-D1a* NILs have fewer spikelets relative to wild-type (Figure 2.1), whereas the slower flowering *ppd-1* null line has more spikelets (Figure 2.1). All genotypes develop more spikelets under glasshouse conditions compared to the field (Figure 2.1). In the field, spike length is longer on average in the *ppd-1* line relative to wild-type (Figure 2.1a), and *Ppd-D1a* lines in both the field and the glasshouse, correlating with spikelet number (Figure 2.1). Interestingly, thousand grain weight is higher in the wild-type compared to the insensitive line and the *ppd-1* NIL under controlled conditions (Figure 2.1a), but not in the field (Figure 2.1b). This trend correlates with grain area, the area is smaller in both the *Ppd-D1a* and *ppd-1* lines, relative to wild-type. However, the grain width of the *ppd-1* and *Ppd-D1a* lines are equally affected, both being significantly lower relative to wild-type (Figure 2.1). The grain length is much shorter in the *ppd-1* null line relative wild-type compared to the *Ppd-D1a* line. This data suggests the yield decreases shown by the insensitive lines are primarily a combination of reduced spikelet number and grain width whilst the yield decrease associated with *ppd-1* null lines are associated with a reduction in total grain size.
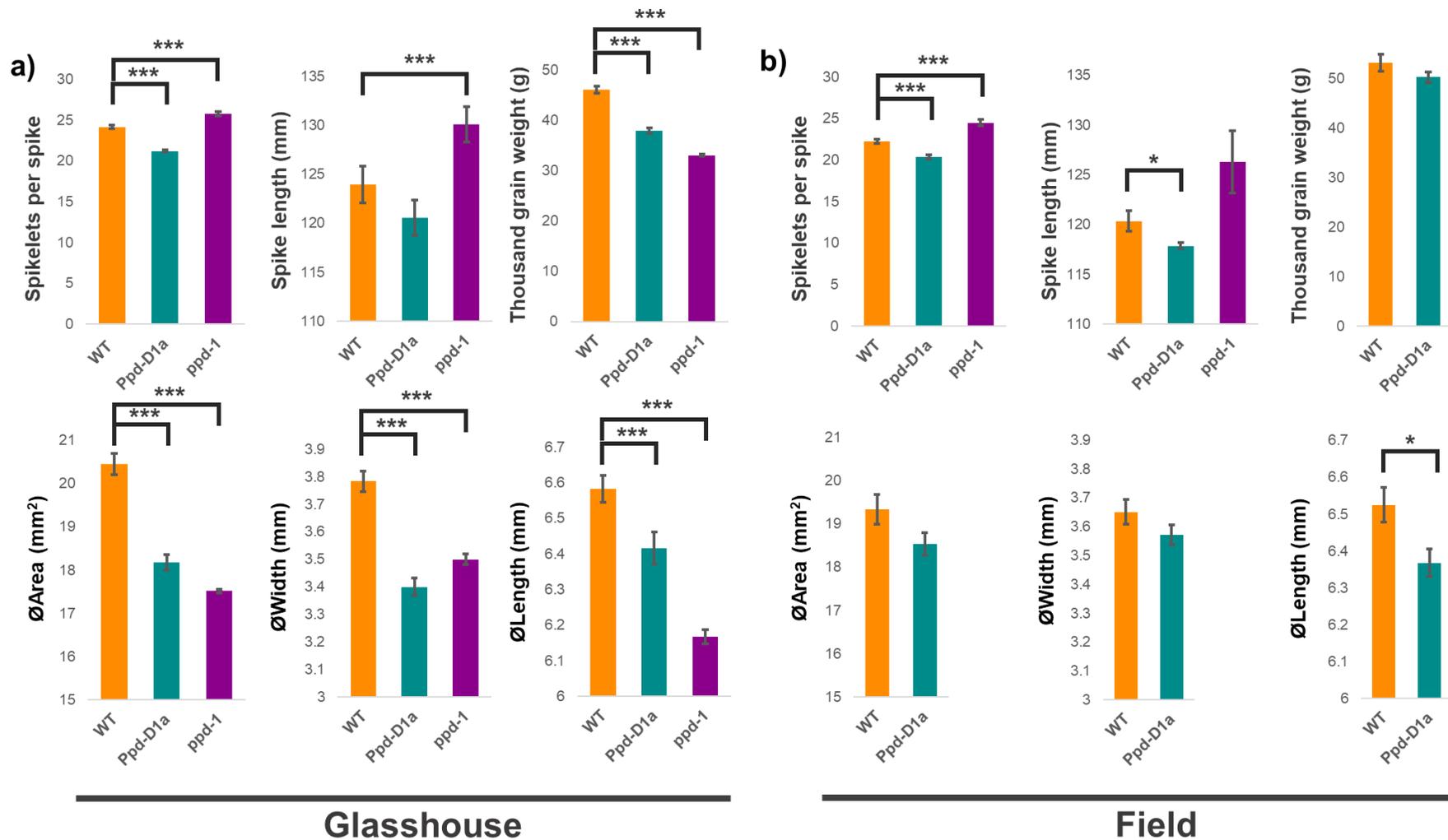
**Figure 2.1 - Seed and spike phenotypes of *Ppd-1* NILs**

The seed and spike characteristics of the Wildtype (Orange), *ppd-1* (Magenta), *Ppd-D1a* (Cyan) NILs. Spikelets per spike, spike length (mm), thousand grain weight (g), area (mm2), width (mm) and length (mm) are shown for each genotype under; **a)** glasshouse and **b)** field conditions. Error bars are ± SEM of at least 5 biological replicates for seed characteristics and 10 biological reps for the spike phenotypes. A one-way ANOVA was used to compare measurements across all three genotypes *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$. A two-way ANOVA compares measurements between the Wild-type and *Ppd-D1a* genotypes and the conditions glasshouse and field (Table 2.1).

**Table 2.1 – A two-way ANOVA analysis for seed and spike phenotypes of wildtype and *Ppd-D1a* NIL in the glasshouse and field.**

| Spikelet Number | SS | df | MS | F | P-value | F crit |
|---|---|---|---|---|---|---|
| Genotype | 176.4 | 2 | 88.2 | 30.84715026 | 1.16201E-09 | 3.168245967 |
| Condition | 15 | 1 | 15 | 5.24611399 | 0.025926884 | 4.01954096 |
| Genotype * Condition | 3.6 | 2 | 1.8 | 0.629533679 | 0.536705004 | 3.168245967 |
| Spike Length | SS | df | MS | F | P-value | F crit |
| Genotype | 1535.033333 | 2 | 767.5166667 | 6.018252574 | 0.004369798 | 3.168245967 |
| Condition | 40.01666667 | 1 | 40.01666667 | 0.313778733 | 0.577687186 | 4.01954096 |
| Genotype * Condition | 26.43333333 | 2 | 13.21666667 | 0.103634542 | 0.901733589 | 3.168245967 |
| TGW(g) | SS | df | MS | F | P-value | F crit |
| Genotype | 152.042364 | 1 | 152.042364 | 24.89373244 | 0.000133707 | 4.493998478 |
| Condition | 470.1909681 | 1 | 470.1909681 | 76.98386056 | 1.64025E-07 | 4.493998478 |
| Genotype * Condition | 33.11739878 | 1 | 33.11739878 | 5.422276018 | 0.033316432 | 4.493998478 |
| ØArea | SS | df | MS | F | P-value | F crit |
| Genotype | 11.44431238 | 1 | 11.44431238 | 33.15129304 | 2.93951E-05 | 4.493998478 |
| Condition | 0.622331086 | 1 | 0.622331086 | 1.802736547 | 0.198121711 | 4.493998478 |
| Genotype * Condition | 2.499961242 | 1 | 2.499961242 | 7.24175861 | 0.016063624 | 4.493998478 |
| ØWidth | SS | df | MS | F | P-value | F crit |
| Genotype | 0.263244543 | 1 | 0.263244543 | 38.52819846 | 1.25591E-05 | 4.493998478 |
| Condition | 0.002257664 | 1 | 0.002257664 | 0.330429326 | 0.573401196 | 4.493998478 |
| Genotype * Condition | 0.113321243 | 1 | 0.113321243 | 16.5855797 | 0.000886164 | 4.493998478 |
| ØLength | SS | df | MS | F | P-value | F crit |
| Genotype | 0.142446776 | 1 | 0.142446776 | 16.18741956 | 0.000982579 | 4.493998478 |
| Condition | 0.009619945 | 1 | 0.009619945 | 1.093194928 | 0.311304107 | 4.493998478 |
| Genotype * Condition | 0.000628556 | 1 | 0.000628556 | 0.071428072 | 0.792683072 | 4.493998478 |

## 2.4.2 Seasonal responses of key flowering genes in the leaf

### 2.4.2.1   Regulation of *Photoperiod-1*

To investigate the seasonal regulation of the photoperiod-dependent flowering pathway, I first measured the expression of *Ppd-D1* and *Ppd-B1* in leaf tissue using quantitative PCR (qPCR). This study focuses on these *Ppd-1* homoeologues, because *Ppd-B1* is the primary functional copy of *Ppd-1* in wild-type (WT) wheat (*cv.* Paragon). In wild-type plants *Ppd-D1* is non-functional due to a 4.8 kb mariner transposable element insertion within the first exon, however, this study uses a NIL (near isogenic line) containing a photoperiod insensitive copy of *Ppd-1* in the D genome (*Ppd-D1a)* (Beales *et al.*, 2007). *Ppd-D1* expression was examined to observe the impact of insensitivity on *Ppd-1* gene expression. In wheat, *Ppd-1* is known to be under the influence of photoperiod, inducing the expression of *FT1* under long-days. To capture this seasonal regulation, expression was examined under field conditions as the daylengths naturally increased by 1-hour increments, from short days in winter (9 h light/15 h dark) to long days in summer (13 h light/11 h dark). In addition to its likely regulation by photoperiod, *Ppd-1* is influenced by the circadian clock, thus, it is important to examine its diel rhythm. To reflect this, expression of *Ppd-1* was examined over a 24-hour period. To investigate if there were differences between the field and controlled environments, expression studies were carried out in the glasshouse in addition to the field (Figure 2.2).

The expression profiles of *Ppd-D1* and *Ppd-B1* are very similar in the field, with the diel rhythm of both being tightly maintained across all photoperiods (Figure 2.2a-b). Both exhibit peaks in expression during the day (Time since dawn (TSD) 3-6 h) and dusk, followed by down-regulation during the night (TSD 16-24 h). The rhythm of expression for both *Ppd-D1* and *Ppd-B1* is maintained across all photoperiods, adjusting to the changing daylengths. The amplitude of *Ppd-D1* expression was stable across all of the photoperiods tested, and *Ppd-B1* maintained normalised peaks of expression in a range between 0.04 and 0.09, which was slightly higher at 10 and 13 h photoperiods (Figure 2.2a,b).

Previous studies have investigated the effects of *Ppd-1* allelism under controlled conditions (Beales *et al.*, 2007; Shaw *et al.*, 2012; Boden *et al.*, 2015). To investigate if there are differences between field and glasshouse conditions, I examined expression under both conditions (Figure 2.2). In the glasshouse, *Ppd-D1* expression profiles show similar trends to that observed in the field, there is low expression at 0 TSD but a sharp increase between the 3 and 6 TSD timepoints (Figure 2.2a,c). Expression then sharply decreased at dusk, with a shoulder of expression just at or during dusk. During the night, I see a similar pattern of gene expression with low expression after dusk, remaining low

until dawn. In the field, gene expression is generally lower compared to the glasshouse, and at the 9 h/15 h photoperiod in the field, there is not the same degree of down-regulated expression after dusk relative to the glasshouse.

These results suggest that neither the altered diel rhythms nor quantitative changes in transcript levels for *Ppd-1* are the cause of photoperiod responsiveness for seasonal regulation of flowering in wheat.
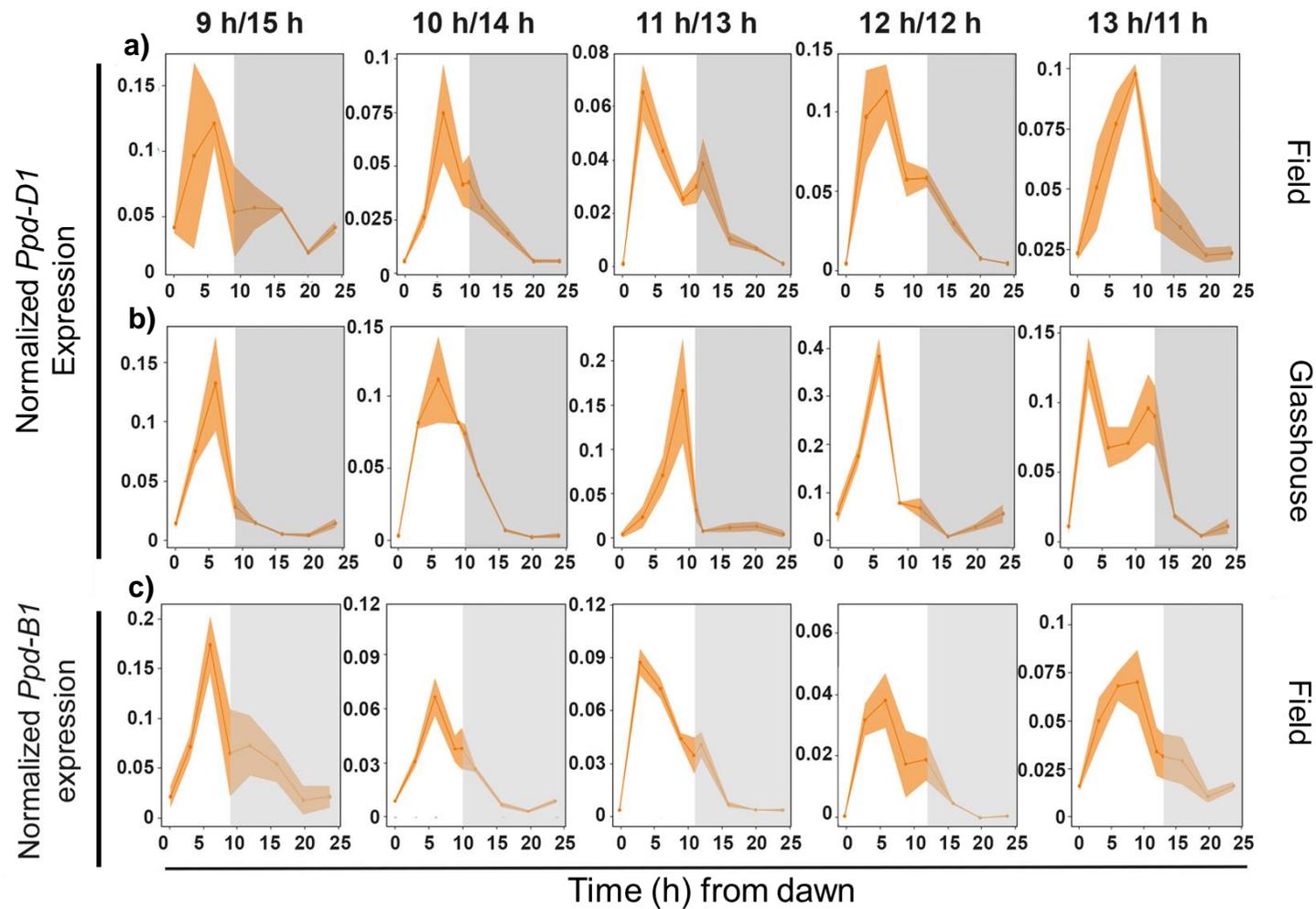
**Figure 2.2 - Seasonal regulation of *Ppd-1* under field- and glasshouse-based conditions**

**a-b)** Diel expression profiles of *Ppd-D1* in wild-type plants (orange) under field conditions. **b)** Expression of *Ppd-D1* under glasshouse conditions. **c)** Expression of *Ppd-B1* under field conditions. All expression profiles are shown over a 24-hour period at hourly incremental increases in daylength. Y axis is a variable scale. Graphs are presented as ribbon plots which show expression, each point is the normalised mean transcript levels (solid line with data points) ± SEM (shaded region) of three biological replicates. Figure adapted from (Gauley & Boden, 2020).

### 2.4.2.2  Genetic regulation of *Photoperiod-1*

A single insensitive allele on the D genome is sufficient to produce an accelerated flowering time; this photoperiod insensitivity is caused by a large deletion (2.09 kb) in the promoter region of *Ppd-D1* (Beales *et al.*, 2007; Shaw *et al.*, 2013). To gain an insight into how this insensitivity manifests in gene regulation, I compared the gene expression in these *Ppd-D1a* lines in the field and glasshouse, relative to wild-type plants (Figure 2.3). To facilitate comparison, *Ppd-1* expression in the *Ppd-D1a* lines is overlayed with wild-type data (Figure 2.3).

In the field, from the 10 h/14 h photoperiod through to the 13 h/11h photoperiods, I observe the insensitive allele altering the diel rhythm of *Ppd-D1* expression (Figure 2.3a). There is significantly higher expression during the night (TSD 20 hr, 24 hr) from the 10-13 h photoperiods, relative to wild-type, particularly at the 11 and 12 h daylengths ($p$ value < 0.05; Appendix 2). The 9 h photoperiod does not show the same trend. The insensitive allele did not significantly affect the amplitude of *Ppd-D1* expression during the day, with *Ppd-D1* peaking between 3 and 6 h after dawn in both genotypes. The presence of the *Ppd-D1a* insensitive allele affected *Ppd-B1* activity. I detected a general trend of higher *Ppd-B1* expression in the *Ppd-D1a* line relative to wild-type, particularly at the 10 h and 11 h photoperiods, no significant difference was detected at the 12 and 13 h photoperiods ($p$ value > 0.05; Appendix 2) (Figure 2.3c).

In the glasshouse, I see a similar trend, although the expression patterns are more robust, with the glasshouse plants showing significantly higher expression at all hours of the night relative to the field (Figure 2.3b). This is consistent with previous reports that examined *Ppd-1* insensitivity under controlled conditions (Beales *et al.*, 2007; Shaw, Turner and Laurie, 2012; Boden *et al.*, 2015). These data suggest that controlled conditions exaggerate the impact of photoperiod-insensitive alleles on *Ppd-D1a* expression, relative to field grown plants. From these results, I conclude that the insensitive allele misregulates *Ppd-D1* expression during the late hours of the night, particularly under day-neutral conditions, and there is an interaction, genetically between *Ppd-1* homoeologous.

**Figure 2.3 - Seasonal regulation of *Ppd-1* under field- and glasshouse-based conditions in wild-type and *Ppd-D1a* NILs**

**(a, b)** Diel expression profiles of *Ppd-D1* in wild-type photoperiod sensitive (wild-type, orange)*, *photoperiod insensitive (*Ppd-D1a*; cyan) NILs under field conditions. **(b)** Expression of *Ppd-D1* in wild-type and *Ppd-D1a* NILs under glasshouse conditions **(c)** Expression of *Ppd-B1* in wild-type*, Ppd-D1a* and *ppd-1* NILs under field conditions. All expression profiles are shown over a 24-hour period at hourly incremental increases in daylength. Y axis is a variable scale. Graphs are presented as ribbon plots which show expression, each point is the normalised mean transcript levels (solid line with data points) ± SEM (shaded region) of three biological replicates. Figure adapted from (Gauley & Boden, 2020).

### 2.4.2.3 A *ppd-1* null line has a dramatic effect on expression

To understand the role of *Ppd-1* in the leaf, I used NILs that contain loss-of-function alleles for all three homoalleles (*ppd-1* NILs). The lines used contain a deletion of *Ppd-B1* and *Ppd-A1*. The non-functional copy in the D genome comes from a premature stop codon that causes a non-functional protein (Beales *et al.*, 2007; Shaw *et al.*, 2013). There is no expression of *Ppd-B1* in these *ppd-1* null lines due to the deletion (Figure 2.4d). The *Ppd-D1* expression, however, shows a dramatic up-regulation of expression in *ppd-1* relative to wild-type or the *Ppd-D1a* lines at all daylengths, which is significant during the day for the 10-13 h photoperiods (Figure2.4a,c). Notably, they maintain a similar expression pattern relative to wild-type with a significant peak in expression in the morning and a dramatic down-regulation at night. In the glasshouse, I observe the same trend (Figure 2.4b). This lends evidence to a feedback loop regulating *Ppd-1* in wheat, whereby *Ppd-1* regulates its own expression, which is disrupted when there is no functional protein.

**Figure 2.4 - Seasonal regulation of *Ppd-1* under field- and glasshouse-based conditions in wild-type, *Ppd-D1a* and *ppd-1* NILs**

**a-b)** Diel expression profiles of *Ppd-D1* in wild-type photoperiod sensitive (orange), photoperiod insensitive *Ppd-D1a* (cyan) and null *ppd-1* (magenta) NILs under field conditions. **b)** Expression of *Ppd-D1* in wild-type, *Ppd-D1a* and *ppd-1* under natural photoperiod glasshouse conditions. All expression profiles are shown over a 24-hour period at hourly incremental increases in daylength. Each point is the normalised mean transcript levels ± SEM of three biological replicates. Y axis is a variable scale. **(c-d)** Summary of *Ppd-1* expression in the leaf. Data summarising the peak of **c)** *Ppd-D1* and **d)** *Ppd-B1* expression for each daylength, plotted on a logarithmic scale. All graphs are presented as ribbon plots showing expression, each point is the normalised mean transcript levels (solid line with data points) ± SEM (shaded region) of three biological replicates. Figure adapted from (Gauley & Boden, 2020).

### 2.4.2.4 Feedback loop regulating *Ppd-1* expression

Higher expression of *Ppd-D1* in the *ppd-1* NIL led us to hypothesize that *Ppd-1* is involved in a self-regulating feedback loop. Clues to *Ppd-1* regulation can be seen from its predicted protein domains and the projected binding sites in its promoter. *Ppd-1* has two primary functional domains, a Pseudo-response regulator domain (PRR) and a CCT (CONSTANS, CO-like, and TOC1) domain (Beales *et al.*, 2007; Figure 2.5). PRR domains are primarily responsible for the regulation of gene expression but also contain conserved residues that are phosphorylated to regulate the function of the domain (Fujiwara *et al.*, 2008). CCT domains are often found in proteins associated with light signal transduction, are involved in nuclear localisation and have roles in protein-protein interactions (Strayer *et al.*, 2000).

Upstream of *Ppd-1* in its promoter region, I identified several binding sites of interest (Figure 2.5). A CIRCADIAN CLOCK-ASSOCIATED 1 (CCA1) binding site was identified, which is notable as *Ppd-1* is a circadian regulated gene and *CCA1* is a core component of the circadian clock. A binding site of PHYTOCHROME INTERACTING FACTOR 3 (PIF3), which is a transcription factor that functions down-stream of the phytochrome photoreceptors was also identified (Chen *et al.*, 2014). There is also an ARR binding site, which is a potential binding site for *Ppd-1*'s PRR domain (Figure 2.6).

Together, this information gives an exciting insight to *Ppd-1*s regulation. The PIF3 and CCA1 binding sites are within the 2 kb deletion of insensitive alleles, and so they could be attributed to the altered *Ppd-1* regulation in these lines. The ectopic expression and apparent lack of circadian regulation of *Ppd-1* I see in the insensitive lines could be due to the lack of CCA1 function on *Ppd-1*. The earlier flowering phenotype, with apparent disregard to the changing photoperiod, could be due to a lack of phytochrome influence through PIF3. The existence of the ARR domain outside this deleted region, could be an avenue for how *Ppd-1* is negatively regulating itself by binding to the ARR site.
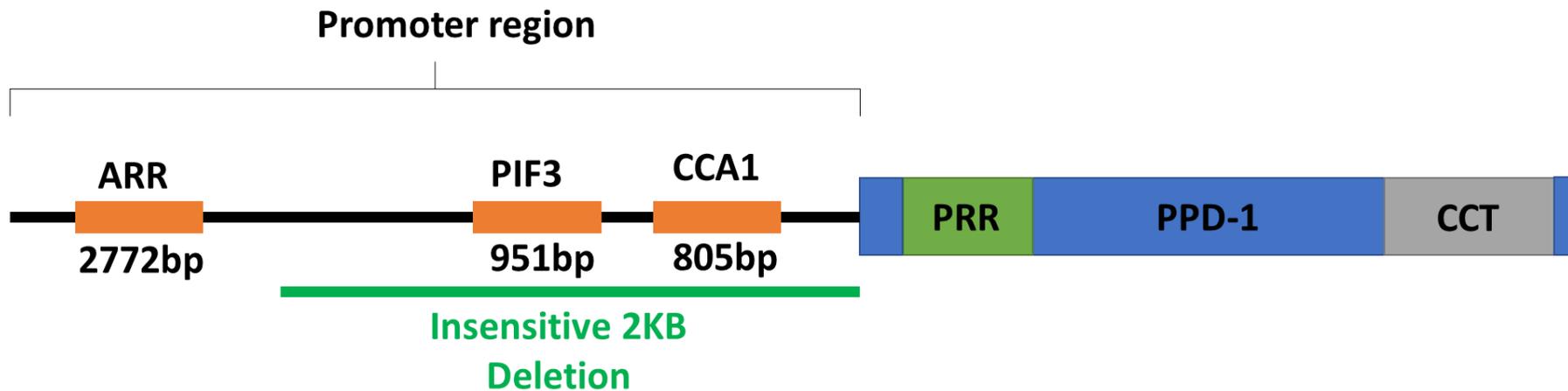
**Figure 2.5 - Predicted protein domains and promoter binding sites of *Ppd-1***

Based on BLAST sequence two *Ppd-1* protein domains are predicted, a CCT (grey) and a PRR domain (green). Key upstream promoter binding sites CCA1, PIF3 and ARR shown in orange, with the distance upstream from the gene start codon noted below the site. The region deleted in *Ppd-D1a* insensitive lines is denoted in green.

### 2.4.2.5  The floral transition occurs under short days

The triggering of flowering in wheat is characterized by the vegetative to floral transition. To understand how flowering is initiated in response to seasonal changes, I carried out a detailed phenotypic analysis of the developing inflorescence in the wild-type, *Ppd-D1a* and *ppd-1* NILs under field conditions. Phenotypically, I observe the characteristic differences in spikelet number observed in these lines becoming clear from the lemma primordia (LP) stage, with the differences being maintained through to terminal spikelet (TS) when spikelet number is fixed (Figure 2.1; Figure 2.6a). In wild-type, I observe the floral transition occurring between the 10 and 11 h photoperiods, with a significant delay at the DR stage until the 12.5 h photoperiod (Figure 2.6b,c). There is a comparatively rapid progression beyond DR to TS stage. The IMs remain at the LP stage until daylengths reach 13.5 h and the TS stage at 14 h. The insensitive line follows a similar trend but transitions earlier between the 9 and 10 h stages and experiences a greatly reduced delay at the DR stage. The *ppd-1* null line transitions at a similar stage relative to wild-type, but progression past the DR stage is delayed. A comparison of development between wild-type plants grown in the glasshouse to those grown in the field shows that the floral transition occurs when the day is approximately one hour shorter in the GH and with subsequent transitions occurring earlier relative to field grown plants (Figure 2.6c). This data shows the vegetative to floral transition occurs during short-day neutral daylengths and not during long days, and the inflorescence has a delay in development at the DR stage.
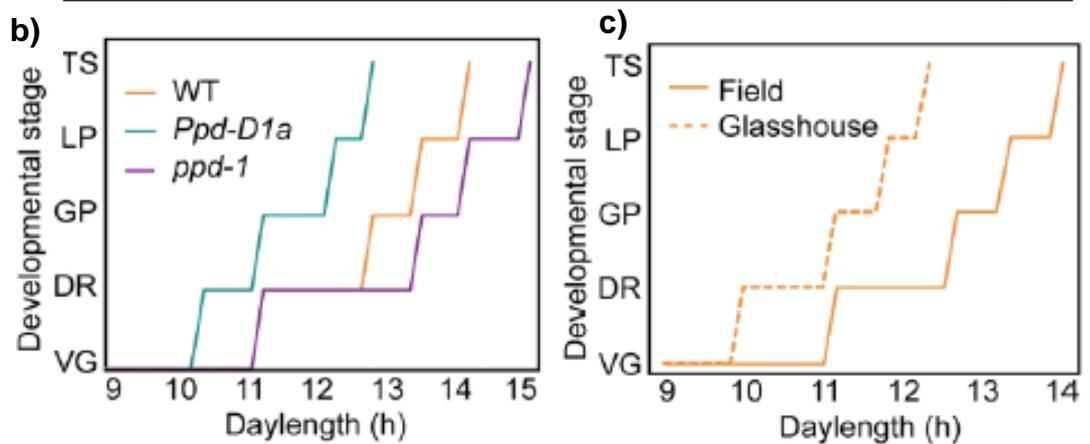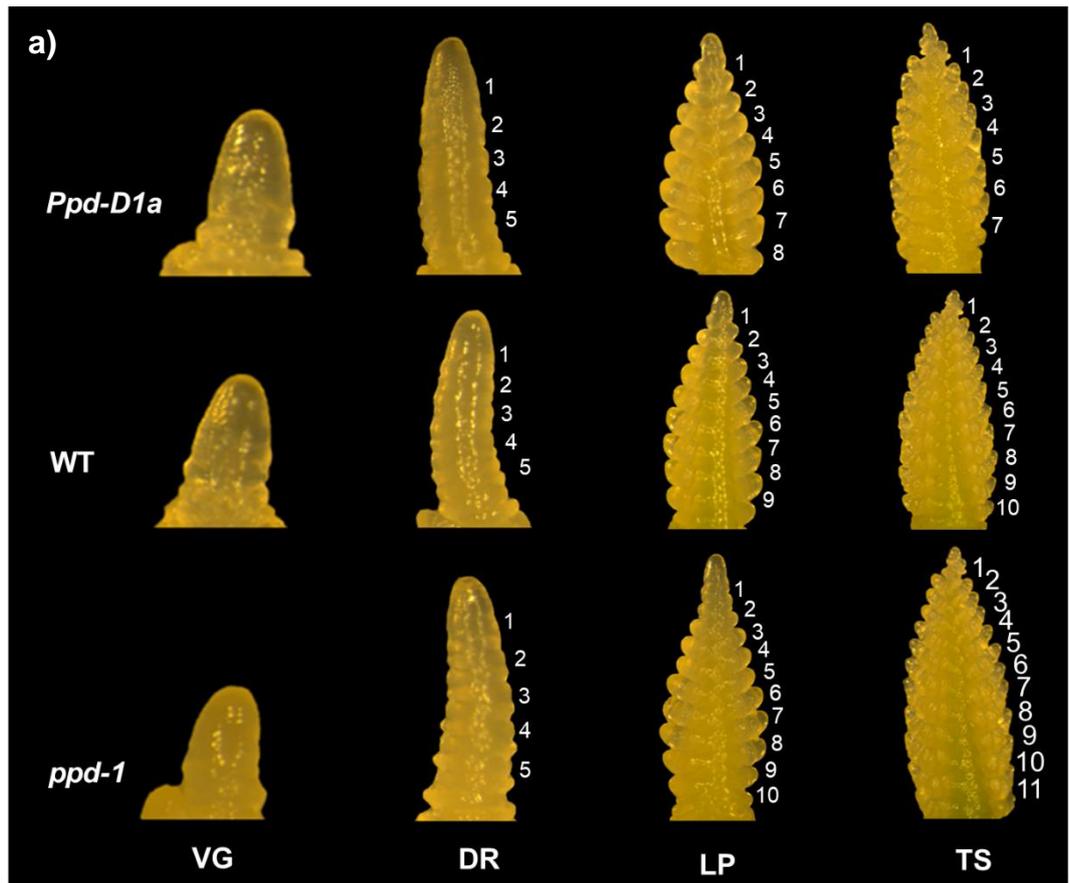
**Figure 2.6 - Seasonal progression of inflorescence meristem development in *Ppd-1* NILs**

**a)** Representative images of inflorescence meristems from each genotype at the four developmental stages. **b)** Progression of inflorescence meristem development throughout the season in wild-type (orange)*, Ppd-D1a* photoperiod-insensitive (cyan) and null *ppd-1* (magenta) NILs under field conditions, at the vegetative (VG), double ridge (DR), glume primordia (GP), lemma primordia (LP) and terminal spikelet (TS) stages. **c)** Comparative progression of wild-type inflorescence meristems grown under field (solid) and glasshouse (dashed) conditions. **b-c)** Data are the average of 4-5 replicates per developmental stage. Figure adapted from (Gauley & Boden, 2020).

### 2.4.2.6 A step-wise induction of *FT1* expression

The seasonal regulation of *Ppd-1* alone does not explain the differences between flowering phenotypes of the *Ppd-1* NILs relative to wild-type. Therefore, I hypothesized that changes in *FT1* expression may provide insights. *FT1* is the primary florigen in wheat. It has been described previously that *FT1* expression is regulated by *Ppd-1,* however, a detailed analysis under field conditions has been lacking (Beales *et al.*, 2007; Boden *et al.*, 2015; Bratzel & Turck, 2015; Kitagawa *et al.*, 2012; Turner *et al.*, 2005; Wilhelm *et al.*, 2009; L. Yan *et al.*, 2006).

The expression of *FT1* was examined in leaf tissue from both the field and glasshouse in wild-type Paragon in addition to *Ppd-D1a* and *ppd-1* NILs using qPCR. In wild-type, *FT1* is induced between the 10 and 11 h photoperiods, where I observe the characteristic patterns of *FT1* expression with a peak during the morning between 3 – 6 TSD, and another at dusk (Figure 2.7a). This first induction correlates directly with when the plant undergoes the floral transition in the field (Figure 2.6b). However, I see a second much greater induction of *FT1* between the 12 and 13 h photoperiods, when *FT1* expression reaches levels 10-fold higher than those observed at the 11 h photoperiod (Figure 2.7a). This induction correlates with when the plants surpass the DR stage to the GP stage, levels then remain at this amplitude whilst growth and development of the meristem proceeds rapidly beyond this point, eventually reaching the terminal spikelet stage when daylengths are 14.75 h (Figure 2.6b). In the glasshouse the induction of *FT1* occurs between the 10 and 11 h photoperiods, the same photoperiod transition observed in the field (Figure 2.7b). However, in the glasshouse *FT1* is induced to a level comparable with the 12 to 13 h transition in the field immediately (Figure 2.7a-b). This high induction of *FT1* correlates with a much-accelerated progression through the DR stage and the inflorescence meristem progressing through key development stages when the day is approximately one hour shorter (Figure 2.6c).

Taken together, these results demonstrate that the development of wild-type IMs to the double ridge stage in the field correlates with an initial rise in *FT1* expression. The IMs then stall at the double ridge stage until the second increase in *FT1* activity promotes the transition of the IM to the lemma primordium and terminal spikelet stages.
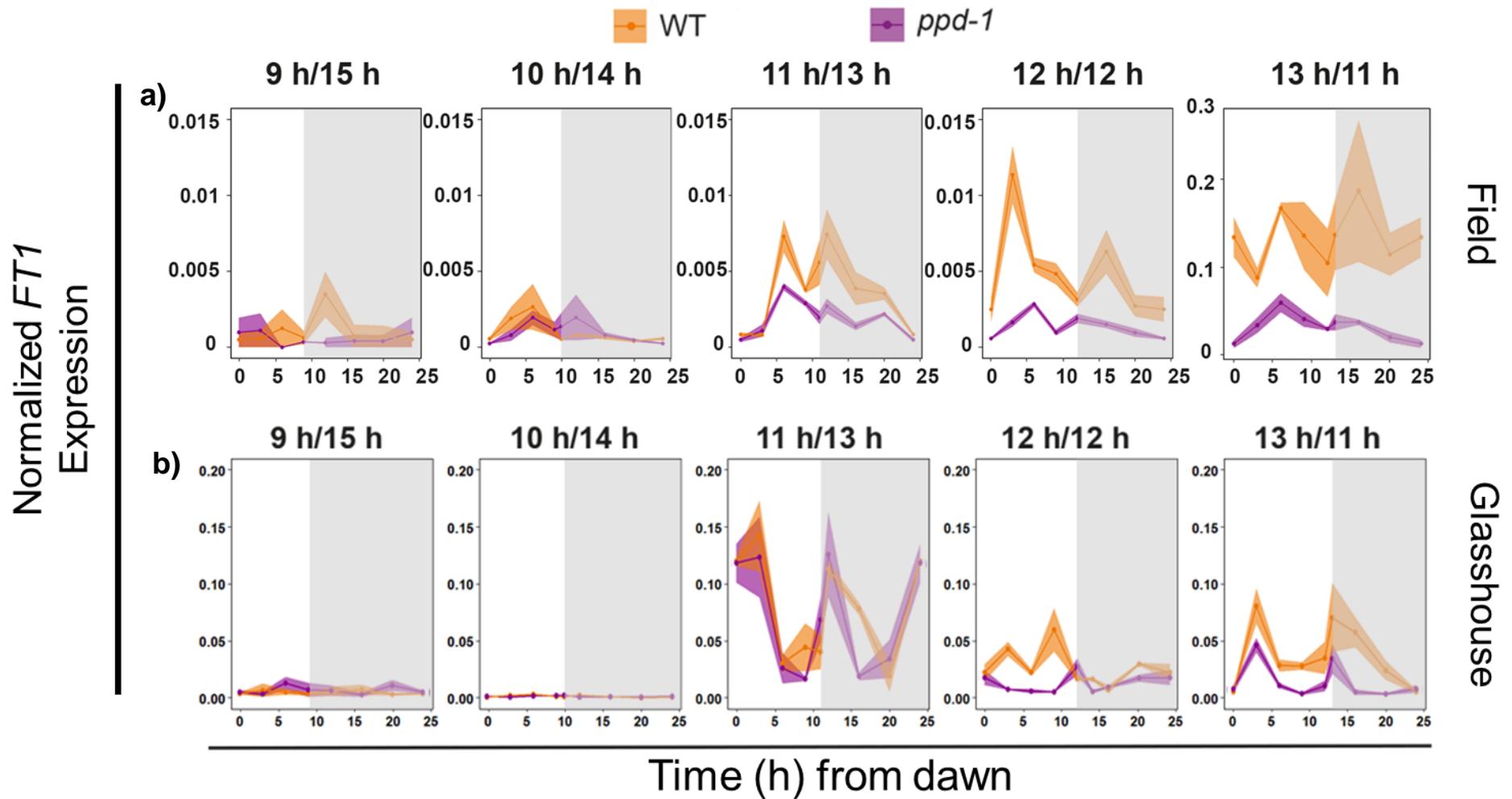
**Figure 2.7 - Seasonal regulation of *FT1* under field- and glasshouse-based conditions in wild-type and *ppd-1* NIL**

Diel expression profile of *FT1* in wild-type (orange) and null *ppd-1* (magenta) NILs under **a)** field and **b)** glasshouse conditions. The grey shading highlights night-time data points. Data are presented as ribbon plots which show the normalised mean transcript levels (solid line with data points) ± SEM (shaded region) of three biological replicates. Figure adapted from (Gauley & Boden, 2020).

### 2.4.2.7  Effects of *Ppd-D1a* and *ppd-1* on *FT1* expression

The development of the inflorescence in *Ppd-D1a* insensitive NILs corresponds closely with the induction of *FT1.* Relative to wild-type, in the field, I have observed an earlier induction of *FT1*, occurring at the 9 to 10 h transition (Figure 2.8a), compared with 10 to 11 in wild-type. At this first induction of *FT1* expression in the *Ppd-D1a* line, it reaches amplitudes similar to wild-type at the 13 h/ 11 h photoperiod. This induction correlates with the IM in the *Ppd-D1a* NIL transition from VG to DR 1-hour daylength before wild-type in the field, (Figure 2.6b). A second induction of *FT1* expression is also evident in the *Ppd-D1a* line, however, it vastly surpassed wild-type levels reaching levels 4-fold higher (Figure 2.8a). Strikingly, in the *Ppd-D1a* line at the 11 h photoperiod, the diel pattern observed in the wild-type plant is not maintained. Interestingly in the NILs that experienced this increase of *FT1* amplitude, the IM progressed much faster through the glume primordium and lemma primordium stages, reaching the terminal spikelet stage when the daylengths were 12.75 h. The same pause at DR seen in wild-type was not observed in the *Ppd-D1a* NIL.

In the *ppd-1* null line, *FT1* expression shares a similar pattern to wild-type, particularly at the 12 h and 13 h daylengths, the expression is lower throughout the photoperiod. Again, this corresponds to inflorescence development transitioning to DR at the 11 h photoperiod and proceeding to the lemma primordium and terminal spikelet stages when daylengths were 14 and 15 h respectively. The accelerated development of the *Ppd-D1a* NIL and the slower development of the *ppd-1* line correlates with spikelet number (Figure 2.6a). Differences first become clear at the lemma primordium stage with variations in final spikelet number apparent at the terminal spikelet stage, with the wild-type plant producing 20 spikelets, compared to the *Ppd-D1a* (16 spikelets) and the *ppd-1* (22 spikelets) lines*.*

To observe trends in *FT1* expression as the daylength increases, the morning peak in expression was plotted on a logarithmic scale (Figure 2.8c). There is a step-wise induction of *FT1* expression for all three genotypes tested, suggesting *FT1* mediated control over inflorescence development beyond the floral transition. The effect of the *Ppd-1* NILs on days to emergence was tested under field and glasshouse conditions (Figure 8d). For each genotype, the glasshouse-grown plants flowered earlier than the field-grown plants. Emergence was accelerated in the *Ppd-D1a* NIL compared to wild-type under both conditions and delayed in the *ppd-1* NIL. In the field, *Ppd-D1a* emerged 13 days earlier than wild-type, while the *ppd-1* NILs flowered 11 days later, in the glasshouse photoperiod-insensitive lines flowered 20 days earlier than wild-type, while the *ppd-1* NILs flowered 11 days later.

Taken together, this step-wise induction of *FT1* points to an intimate connection between the leaf and the developing inflorescence, mediating key stage transitions.
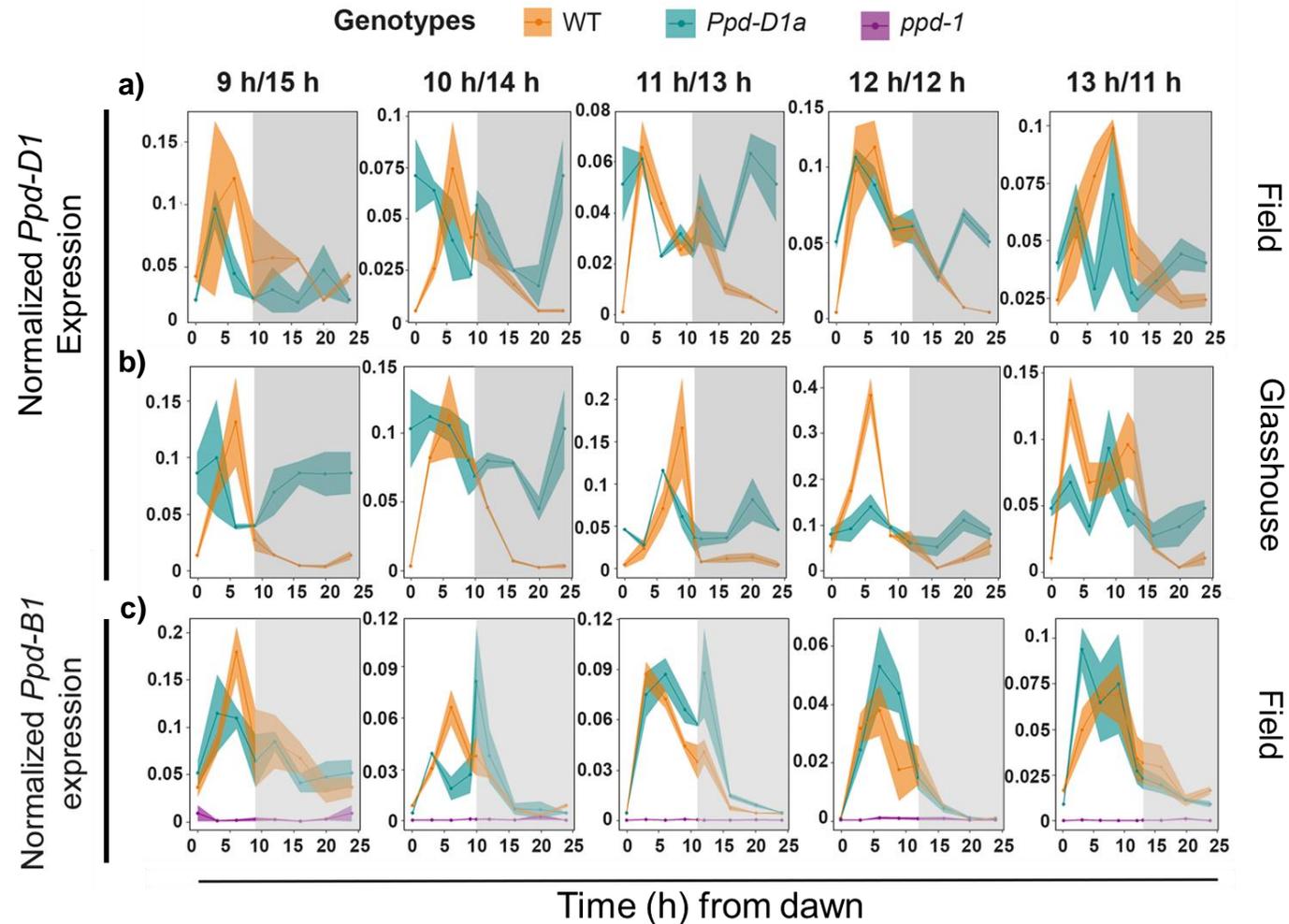
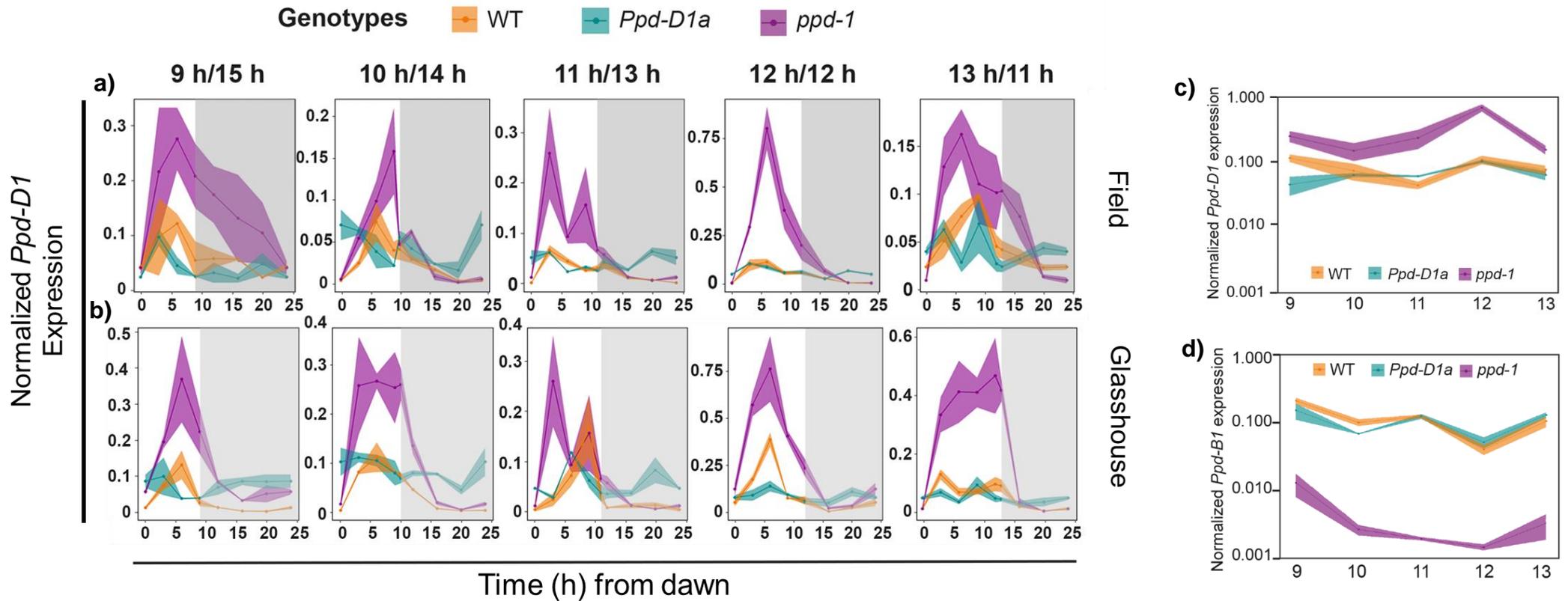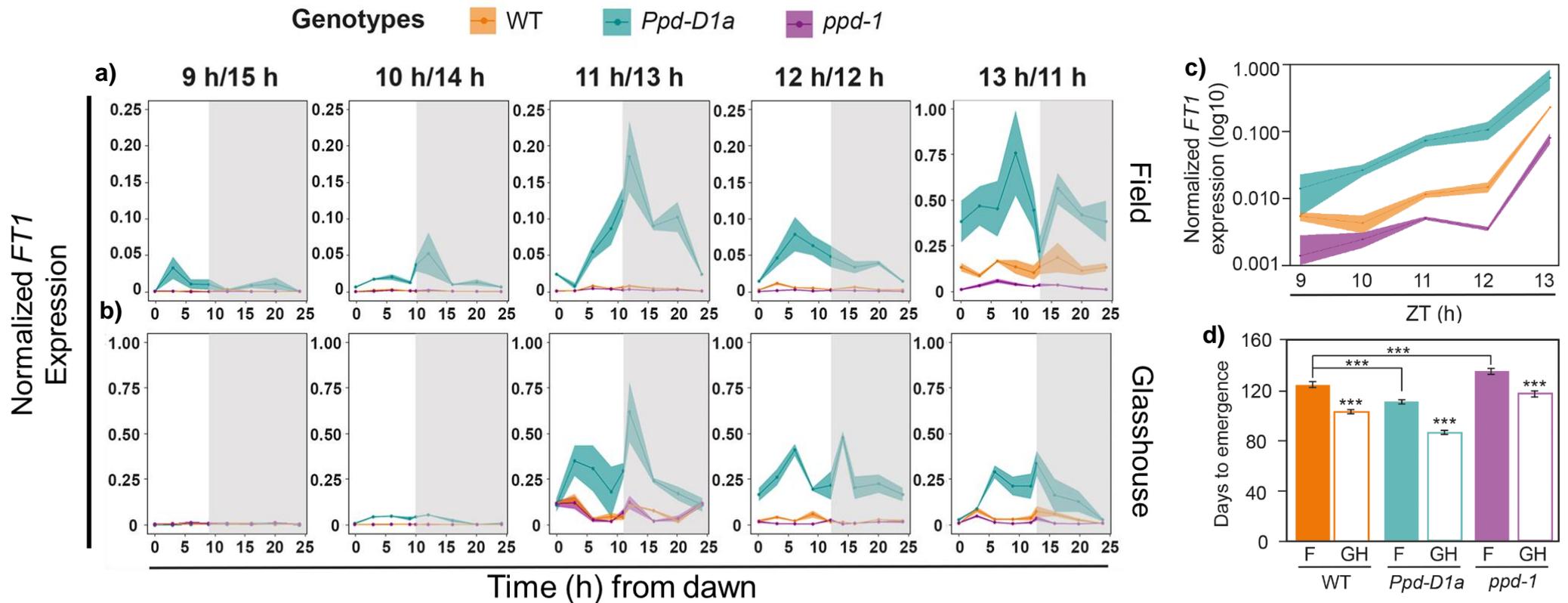**Figure 2.8- Seasonal regulation of *FT1* under field- and glasshouse-based conditions in wild-type, *Ppd-D1a* and *ppd-1* NILs**

Diel expression profile of *FT1* in wild-type (orange)*, Ppd-D1a* photoperiod-insensitive (cyan) and null *ppd-1* (magenta) NILs under **a)** field and **b)** glasshouse conditions. The grey shading highlights night-time data points. **c)** Data summarising the peak of *FT1* expression for each daylength, plotted on a logarithmic scale. **d)** Field and glasshouse flowering time phenotypes for the three *Ppd-1* NILs, normalised as days to emergence from the 9 h daylength. **(a-c)** Data are presented as ribbon plots that show the normalised mean transcript levels (solid line with data points) ± SEM (shaded region) of three biological replicates. The grey shading highlights night-time data points. **d)** Data is mean ± SEM, with field replicates being 5 independent plots and glasshouse replicates being 10-15 plants. *** $P < 0.001$. Figure adapted from (Gauley & Boden, 2020).

### 2.4.2.8 Regulation of the core clock genes *CCA1, TOC1* and *ELF3*

To determine if seasonal variation in photoperiod insensitive alleles of *Ppd-1* influence the circadian clock, I analysed the expression of several core clock components under field conditions (Figure 2.9). The genes examined included the core clock genes *CCA1* and *TOC1*, in addition to *ELF3* which is a member of the evening loop that suppresses flowering in wheat and barley (Wang & Tobin, 1998; Alabadí *et al.*, 2001; Dixon *et al.*, 2011; Nusinow *et al.*, 2011; Faure *et al.*, 2012; Zakhrabekova *et al.*, 2012; Alvarez *et al.*, 2016; Zikhali *et al.,* 2016). For *CCA1*, I found expression peaked at 3 h after dawn in each photoperiod before dropping during the day and beginning to increase again between TSD 16 h and dawn. As daylengths increased, the peak in expression shifted to be maintained at TSD 3 h (Figure 2.9a). Intriguingly, the high expression seen at TSD 3 h indicates a few hours of daylength are required for field-grown plants to generate the characteristic peak of *CCA1* transcripts around a subjective dawn, as observed in plants grown under controlled conditions (Wang *et al.*, 1997; Wang & Tobin, 1998; Campoli *et al.*, 2012; Deng *et al.*, 2015). The shift of *CCA1* expression according to photoperiod was mirrored by *TOC1*, for which expression peaked at dusk for all photoperiods, and was lowest at 3 h after sunrise (Figure 2.9b). The diel pattern of *CCA1* and *TOC1* expression did not change in the presence of the *Ppd-D1a* allele, suggesting the *Ppd-1* insensitive alleles do not accelerate flowering by modifying the expression of core circadian clock genes (Figure 2.9a-b). *ELF3* expression was arrhythmic in wild-type across all photoperiods, with no consistent diel pattern observed across the different daylengths, which is different from the expression reported under controlled conditions (Figure 2.9c) (Dixon *et al.*, 2011; Faure *et al.*, 2012; Alvarez *et al.*, 2016). Insensitive and triple null alleles of *Ppd-1* did not appear to affect *ELF3* expression, which is consistent with the suggested model of *Ppd-1* acting downstream of *ELF3* to regulate flowering in grasses (Faure *et al.*, 2012).
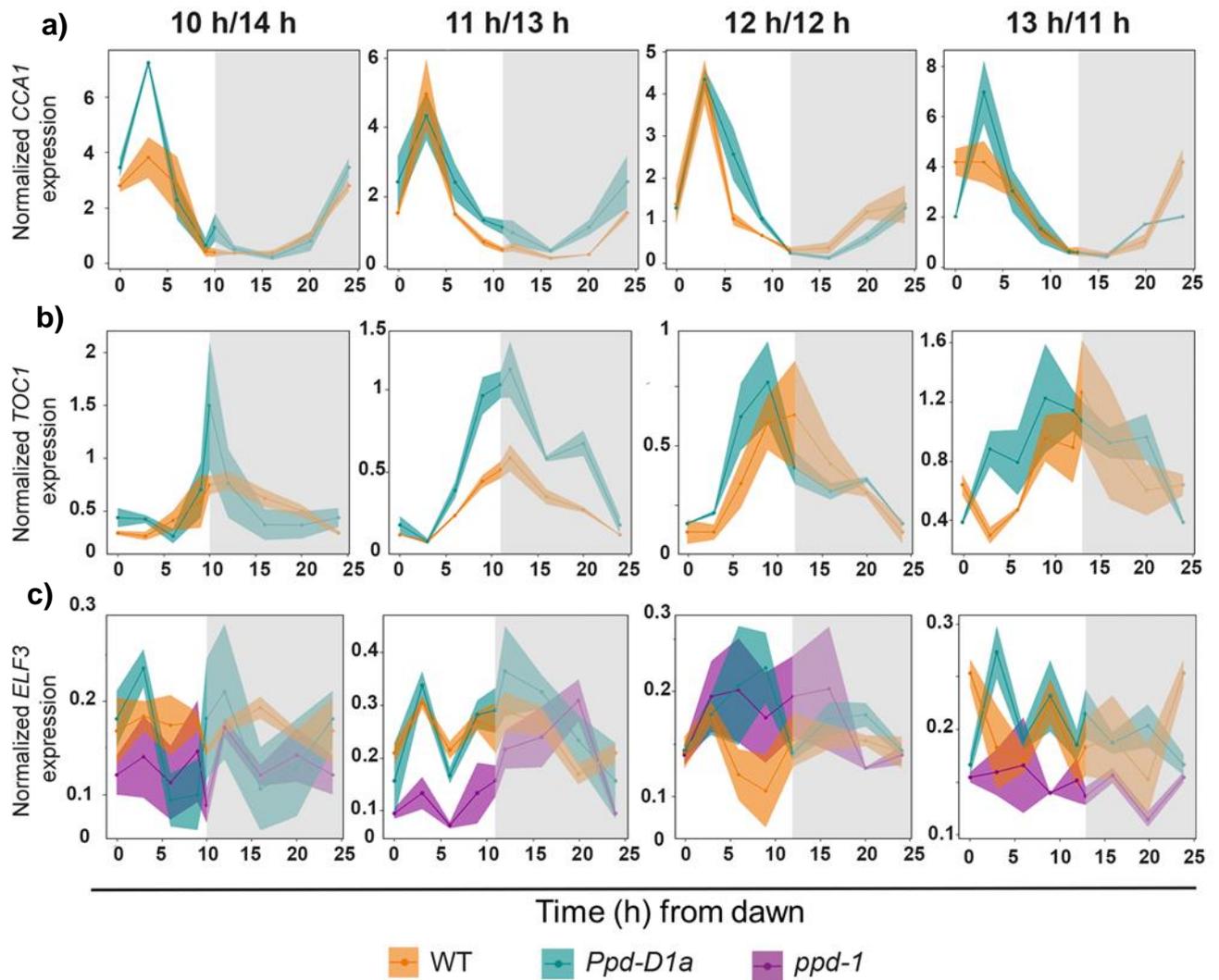
**Figure 2.9- Seasonal expression patterns of core circadian clock genes**

Diel expression profiles of **a)** *CCA1.* **b)** *TOC1* **c)** *ELF3* in wild-type (Orange), photoperiod insensitive *Ppd-D1a* (cyan) and *ppd-1* (Magenta) NILs under field conditions. Each gene was analysed at four different photoperiods (10-13 h). Data are presented as ribbon plots that show the normalised mean transcript levels (solid line with data points) ± SEM (shaded region) of three biological replicates. The grey shading highlights night-time data points.

### 2.4.2.9  The localisation of *Ppd-B1* protein

Despite its importance as a regulator of flowering time, very little is known about the mechanism of *Ppd-1* action, particularly on a protein level. It is predicted that *Ppd-1* acts as a transcription factor, binding to the promoter of *FT1*; however, there is little evidence of this direct interaction. To start to investigate the molecular action of Ppd-1, transgenic plants containing *Ppd-B1* with a green fluorescent protein (GFP) tag under an overexpressing ubiquitin promoter were developed in hexaploid wheat (*cv.* Fielder). Based on preliminary phenotypes of the $T_0$ transgenic lines, plants flowered much faster and produced a spike with dramatically fewer spikelets, relative to wild-type (Figure 2.10a). It is unknown if *Ppd-1* functions through protein-protein interaction in the cytoplasm, through its CCT domain, or direct transcription regulation, likely through its PRR domain (Figure 2.5). Clues to its action can be found in its localisation – to investigate the localisation, confocal imaging was used on the leaf tissue of the *UBI::PPD-B1::GFP* lines. In these lines, GFP tagged Ppd-B1 localises to the nucleus of cells in the leaf, which is consistent with a role in transcription (Figure 2.10b). Notably, it is not localized to the nucleus in every cell, suggesting a potential dynamic regulation of *Ppd-1* through protein localisation.
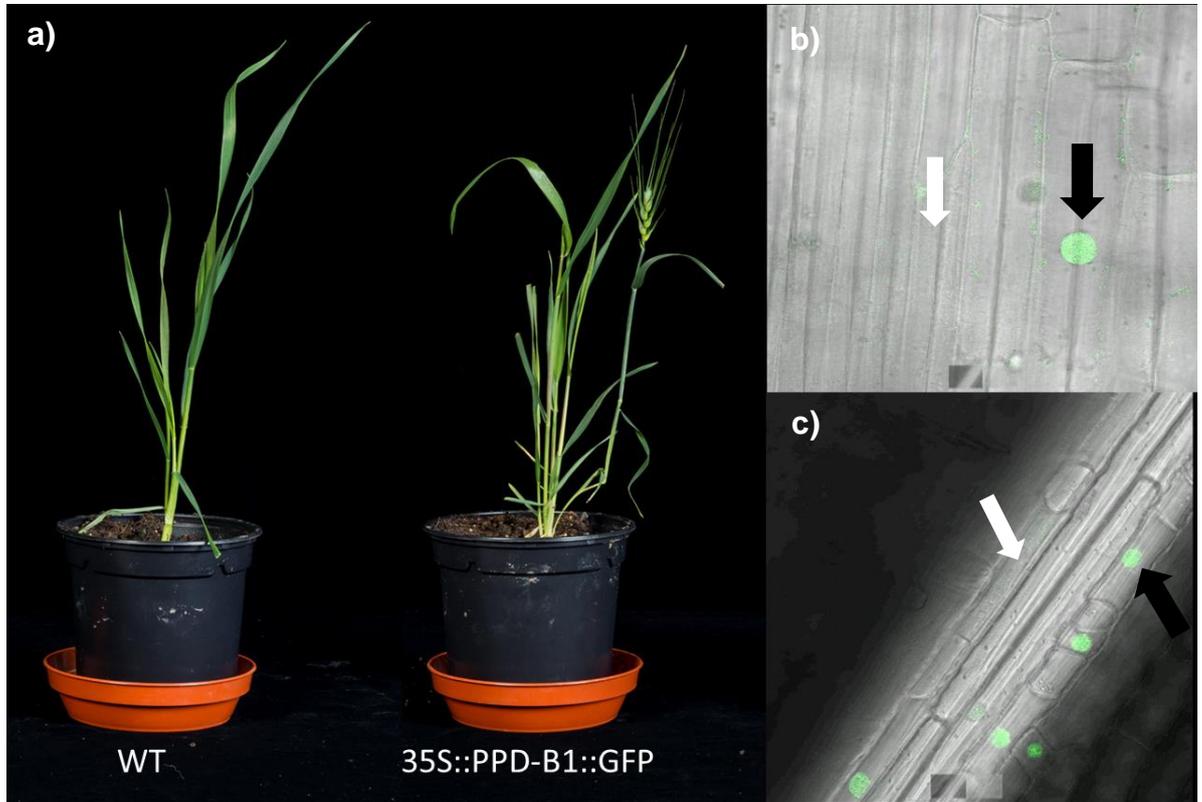
**Figure 2.10 - Effect and localisation of overexpressed GFP tagged *Ppd-B1***

**(a)** The flowering phenotype of overexpressed *Ppd-B1* (UBI::PPD-B1::GFP)*. **(b, c)** Localisation of GFP tagged *Ppd-B1* to leaf cell nucleus*. A black arrow denotes a cell with localisation to the nucleus. A white arrow denotes a cell without any nucleus localisation visible.

## 2.4.3 Checkpoints underpinning inflorescence development

### 2.4.3.1  A checkpoint at the lemma primordium stage

To investigate if the first induction of *FT1* in the leaf is sufficient to induce IM development to flowering, I performed a moving bench experiment. I hypothesized that the first induction of *FT1* promoted the transition from a vegetative to floral state, but the second induction of *FT1* is required for continuation through to the later stages of inflorescence development. To test this, wild-type and *Ppd-D1a* plants were grown under glasshouse conditions until the photoperiod naturally reached 10 h light/14 h dark, at which point the plants were maintained at a 10 h/14 h photoperiod on moving benches (MB). The MB plants reached the lemma primordium stage at the same rate to glasshouse grown plants, which at lemma primordia stage had surpassed the 12 h photoperiod. However, the MB plants stalled at the lemma primordium stage, whilst the plants grown under glasshouse conditions developed normally. This stalling delayed transition to the terminal spikelet stage by 30 days compared to plants maintained under natural photoperiods, and produced more spikelets, respectively (24 $\pm$ 0.3 for glasshouse vs 29 $\pm$ 0.5 spikelets for MB) (Figure 2.11d,e). The delay in inflorescence development was also observed in *ppd-1* NILs, which produced more spikelets and transitioned to the terminal spikelet stage significantly later than plants maintained under natural photoperiods (25.4 $\pm$ 0.2 vs 28.3 $\pm$ 0.3 spikelets). In addition, the wild-type plants maintained at 10 h/14 h photoperiod produced inflorescences with elongated internodes, similar to that exhibited in the *ppd-1* NILs (Figure 2.11c; and Shaw *et al*, 2013). The elongated internodes suggest that *Ppd-1* functions at much later developmental stages and photoperiods than previously thought. Our findings show that by growing wild-type Paragon plants under these 10 h conditions, the elongated internode phenotype observed in *ppd-1* null plants is replicated. The insensitive *Ppd-1Da* lines, however, have a regular internode length. This suggests *Ppd-1* controls later stages of inflorescence developing in a photoperiod-dependent way. Under 10 h/ 14 h MB conditions, *Ppd-D1a* lines did not exhibit the same delay at the lemma primordium stage as seen in wild-type under these conditions; however, there was a slight delay resulting in more spikelets forming (MB 25.8 $\pm$ 0.4 vs GH 21.5 $\pm$ 0.2). The wild-type and *ppd-1* NIL flower later in the MB relative to the GH. However, the *Ppd-D1a* line flowered at the same time under both conditions (Figure 2.11d). These data indicate that the photoperiod insensitivity conferred by these lines is sufficient to overcome this checkpoint regulation.
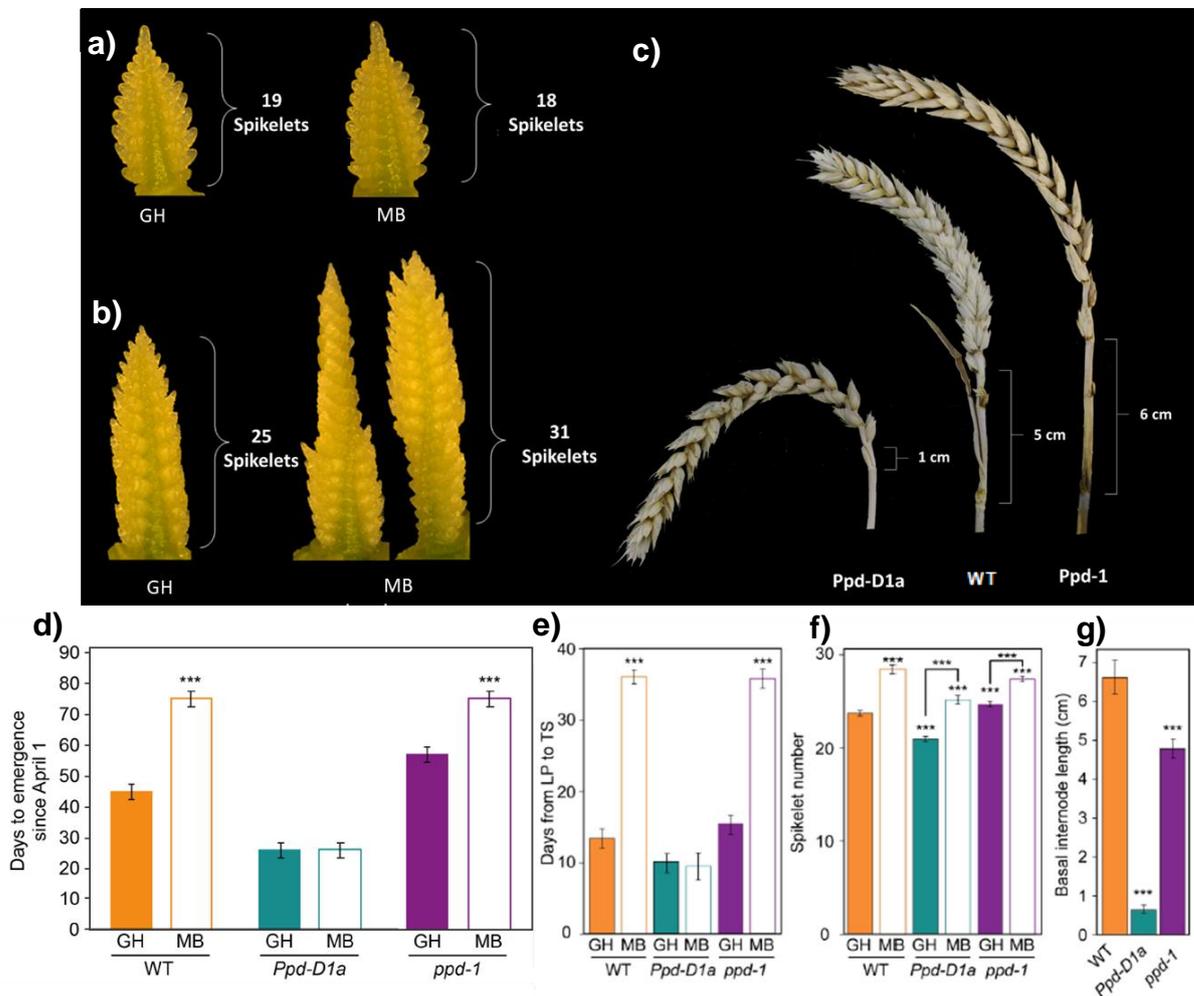
**Figure 2.11 - Effect of maintained short days on inflorescence development**

**a)** Inflorescences of wild-type plants were grown under natural glasshouse and 10 h moving bench (MB) conditions develop at the same rate until the LP stage. **b)** IM development of plants grown under MB conditions is significantly delayed between the LP and TS stages, resulting in more spikelets. **c)** Wild-type and *ppd-1* NIL plants show elongated basal rachis internodes, but not *Ppd-D1a* NILs under MB conditions. **d)** Flowering time measurements of wild-type (orange), *Ppd-D1a* NIL (cyan) and *ppd-1* null (Magenta) NILs grown under natural photoperiod glasshouses (GH, solid) or the moving bench conditions (MB, white). **e)** IM development of wild-type, *Ppd-D1a* and *ppd-1* NILs grown under MB conditions is significantly delayed between the LP and TS stages. **f)** Spikelet numbers for all three genotypes grown under natural GH or MB conditions. **g)** Wild-type and *ppd-1* NIL plants show elongated basal rachis internodes, but not *Ppd-D1a* NILs. Data are the average ± SEM of five biological replicates. *** P < 0.001. Figure adapted from (Gauley & Boden, 2020).

### 2.4.3.2 *FT2* mediates the LP to TS transition

To investigate the genes involved in this process, I first theorised that the *FT* family of genes would be involved in the photoperiod control of inflorescence development. Previous analysis in *Brachypodium distachyon* and tetraploid wheat indicated the importance of *FT2* in flowering time and spikelet number (Shaw *et al*, 2019). I compared the expression of *FT2* at the LP stage in GH and MB plants, since it was at LP that differences between the conditions manifested. The expression of *FT2* is significantly lower in the MB bench (Figure 2.12d).

To test the role of *FT2* genetically, I analysed two independent lines containing missense mutations in *FT2* of the B genome (*FT-B2*) that are predicted to be deleterious for protein function (PROVEAN scores of -5.1 and -6.8) (Choi & Chan, 2015; Figure 2.12a). Progression of inflorescence development to the terminal spikelet stage was delayed in the *ft-b2* lines, with mutant plants producing significantly more spikelets per inflorescence than the wild-type NILs (Figure 2.12c). The *ft-b2* mutants also flowered later than wild-type (Figure 2.12f). These results are consistent with *FT2* having an important role during the stages of inflorescence development when spikelet numbers are determined, coinciding with the second induction of *FT1* expression.
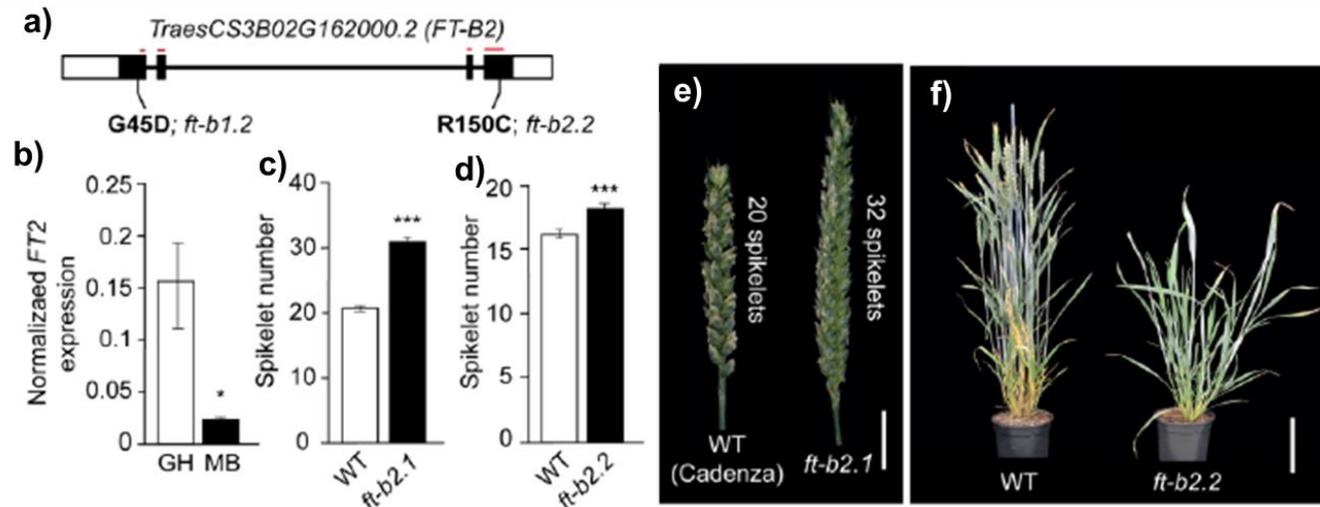
**Figure 2.12 - *FT2* influences inflorescence development and spikelet number**

**a)** A *FT-B2* schematic showing sites of the two missense mutations, *ft-b1.2* and *ft-b2.2 :* exons (black boxes), introns (black lines), untranslated regions (white boxes) and the PEBP domain (red line). **b)** *FT2* expression is lower in developing IMs of wild-type plants shifted to a 10 h photoperiod using a moving bench (MB), relative to unshifted plants maintained under natural photoperiods. **c-e)** Increased number of spikelets on inflorescences of (**c**) *ft-b2.1* mutants and **d)** *ft-b2.2* mutants, relative to wild-type, **e)** images of representative inflorescences. **f)** *ft-b2.2* mutant lines flower later that wild-type (WT) under long daylengths. **b)** Data are the mean ± SEM of three biological replicates, and five replicates for **c-d).** * $P < 0.05$; *** $P < 0.001$. Figure adapted from (Gauley & Boden, 2020).

### 2.4.3.3  Induction of *FT1* in the meristem

Through observation of plants grown at the 10 h MB, I observed abortion between the WA and GA stages in the wild-type, indicating there may be another checkpoint of photoperiod mediated inflorescence development. I hypothesized that this seemingly photoperiod-mediated abortion would associate with *FT1* expression. As a result, I examined the expression of *FT1* during later stages of inflorescence development. *FT1* was not expressed at the early stages (VG-TS), but during later stages (WA-GA) there is considerable rise in expression at GA (Figure 2.13a). To identify which tissue *FT1* is expressed and to provide clues of its function, *FT1* expression was tested in the glume, floret and the rachis (Figure 2.13b). To examine if this induction, like in the leaf, is dependent on *Ppd-1, FT1* expression was tested in the *Ppd-D1a* and *ppd-1* NILs. In the wild-type, there is significantly higher expression in the rachis and no expression in the glume and floret. In the *ppd-1* NIL I see a similar expression pattern to that seen in the wild-type. In the insensitive *Ppd-D1a* NIL, there is a significantly higher expression of *FT1* in the rachis, indicating photoperiod responsiveness of the induction of *FT1* at the green anther stage. This expression indicates *FT1* has a role in this late stage of spike development.
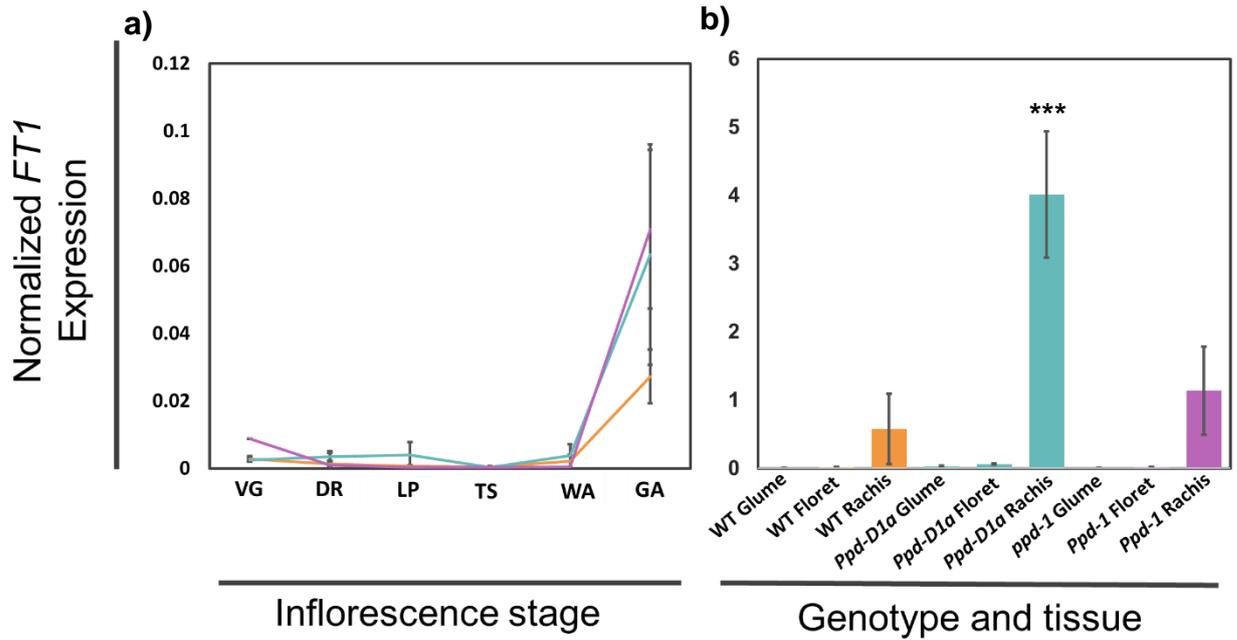
**Figure 2.13 - Expression of *FT1* at the green anther stage**

**a)** Expression of *FT1* plotted by developmental stage from plants grown under field conditions. Data includes expression profiles for wild-type (orange), *Ppd-D1a* photoperiod-insensitive (cyan) and null *ppd-1* (magenta) NILs. The defined stages are vegetative (VG), double ridge (DR), lemma primordium (LP), terminal spikelet (TS), white anther (WA), green anther (GA). **b)** The expression of *FT1* in wild-type, *Ppd-D1a* and *ppd-1* NILs in the glume, floret, and rachis at the GA stage. Error bars are ± SEM of three-four biological replicates. *** $P < 0.001$.

### 2.4.3.4 Impact of leaf gene expression on key meristem identity genes

A previous study highlighted the impact of *Ppd-D1* loss-of-function alleles on spikelet architecture (Boden *et al.*, 2015). It was found that the expression of meristem identity genes are reduced, relative to wild-type. Because of the *FT1* expression patterns observed in this study (Figure 2.7) and the altered inflorescence development in the *Ppd-D1a* and *ppd-1* NILs (Figure 2.6), I hypothesized that the expression of key meristem identity genes would occur earlier and to higher levels in the *Ppd-D1a* line and lower and delayed in the *ppd-1* NILs, relative to wild-type. To investigate this, I tested the expression of several genes in inflorescence meristems extracted from field grown plants, these include the MADS-box transcription factors *VERNALIZATION1* (*VRN1*), *SEPALLATA 1-6* (*SEP1-6*), *APETALA1-2* (*AP1-2*), *APETALA1-3* (*AP1-3*), *SUPPRESSOR OF CONSTANS1* (*SOC1*) and *AGAMOUS-like 6* (*AGL6*). I also examined genes that are known to influence spikelet number and floret fertility in wheat and barley, namely *FLOWERING LOCUS T2* (*FT2*) and *HOMEOBOX- PROTEIN HOX2* (*HOX2*) (Figure 2.14; Boden *et al.*, 2015; Sakuma *et al.*, 2019; Schilling *et al.*, 2020; Shaw *et al.*, 2019; Zhao *et al.*, 2006). The expression of *FT2*, *SEP1-6*, *AP1-3*, *AP1-2*, and *SOC1* were detected in the *Ppd-D1a* NIL and were induced earlier in the growing season relative to wild-type, when plotted by daylengths (Figure 2.14a). The expression of these genes was detected between the 10-11 h photoperiods in *Ppd-D1a* but not in wild-type until 12-13 h photoperiods. This closely correlated with the daylengths where *FT1* was induced in the leaf and the inflorescence transitioned from vegetative to double ridge. The expression of *AGL6* and *HOX2* occurred later in development with expression being detected in the insensitive lines earlier, relative to wild-type. The expression of these genes was the opposite in the *ppd-1* null line, with expression detected later than wild-type, at daylengths corresponding to the delay in *FT1* induction and inflorescence development observed in these plants.

The maximum amplitude of *FT2*, *VRN1*, *SEP1-6*, *AP1-3*, *AP1-2*, *SOC1* and *AGL6* expression was the same in both the *Ppd-D1a* and wild-type lines. This demonstrates that the peak of expression in these meristem identity genes is not affected by the *Ppd-D1a* allele (Figure 2.14). However, the transcripts were significantly lower in the *ppd-1* null lines, indicating that *Ppd-1* is required for robust expression of these genes. At the green anther stage, *HOX2* transcripts have a peak in expression which is greater in the *Ppd-D1a* NILs relative to wild-type and *ppd-1* NILs. Because of this shift in expression pattern in the insensitive and null lines relative to wild-type, I hypothesized the pattern was changing based on developmental stage rather than daylength. To examine this, I normalised the gene expression by developmental stage, namely the vegetative, double ridge, lemma primordium, terminal spikelet stages (Figure 2.14b). The normalization resulted in

transcript levels being very similar between the wild-type and *Ppd-D1a* NILs, highlighting that the insensitive allele affects meristem expression through accelerating expression without increasing the absolute peak of transcript levels (Figure 2.14b). The exception to this is *FT2,* which exhibits much higher expression in the *Ppd-D1a* NILs at the LP stage, relative to wild-type. In wild-type, the transcripts did not increase significantly until the TS stage (Figure 2.14b). Expression in the *ppd-1* line, however, was much lower and tended to peak later in development relative to wild-type (Figure 2.14b). These results, carried out in the field, show insensitive alleles cause accelerated but comparable expression of key meristem identity genes in the inflorescence meristem as it develops, relative to wild-type. In addition, it is clear *Ppd-1* is required for timely and robust expression of genes that promote floret and spikelet development and that expression of all these genes are intimately linked with the induction of *FT1* in the leaf.
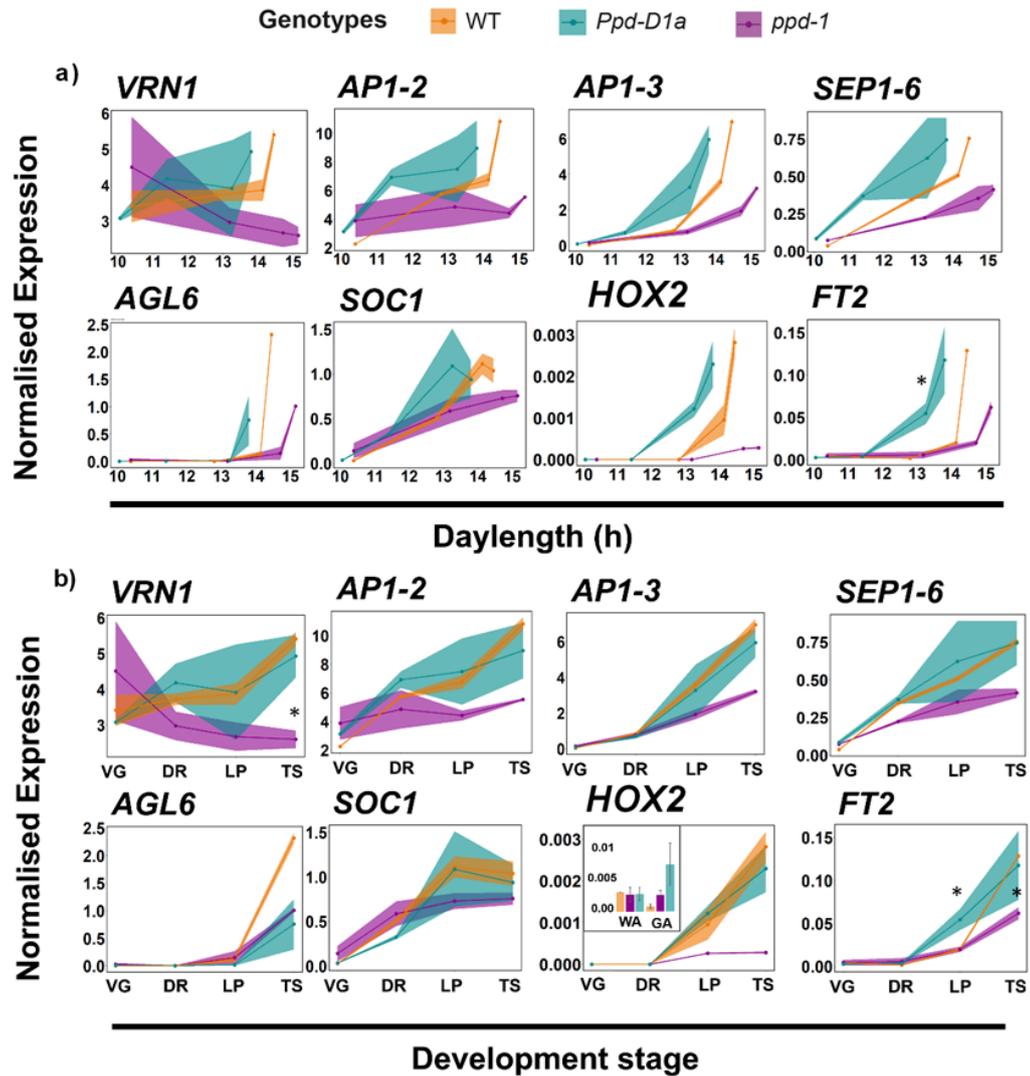
**Figure 2.14 - Seasonal and stage-specific expression analysis of meristem identity genes**

Expression of key meristem identity genes plotted by daylength **(a)** and developmental stage **(b)** from plants grown under field conditions. Data includes expression profiles for wild-type (orange), *Ppd-D1a* photoperiod-insensitive (cyan) and null *ppd-1* (magenta) NILs. The defined stages are vegetative (VG), double ridge (DR), lemma primordium (LP), terminal spikelet (TS), white anther (WA) and green anther (GA).WA and GA expression data for *HOX2* is shown in an insert box. Data are presented as ribbon plots which show the normalised mean transcript levels (solid line with data points) ± SEM (shaded region) of three-four biological replicates. * $P < 0.05$. Figure adapted from (Gauley & Boden, 2020).

2.5 **Discussion**

Allelic variation for *Ppd-1* is known to effect both flowering time and key yield traits, such as spikelet number and floret fertility. These effects have contributed greatly to altering the geographical range of wheat cultivation, maximizing yield in marginal growth environments (Figure 2.1) (Beales *et al.*, 2007; Boden *et al.*, 2015; González-Navarro *et al.*, 2015; González *et al.*, 2005; Pérez-Gianmarco *et al.*, 2019; Prieto *et al.*, 2018; Wilhelm *et al.*, 2009; Worland *et al.*, 1998). Studies to understand *Ppd-1* dependent regulation of flowering have been performed under laboratory conditions, usually with extreme short or long days, these conditions do not adequately represent the daylengths under which spikelet number and floret initiation are decided (Beales *et al.*, 2007; Boden *et al.*, 2015; Campoli *et al.*, 2012; Ejaz & von Korff, 2017; Kitagawa *et al.*, 2012; Shaw *et al.*, 2012, 2013; Turner *et al.*, 2005; Wilhelm *et al.*, 2009). In this chapter, I have investigated flowering and inflorescence development under field-based conditions, providing new understanding for how these processes are regulated during natural seasonal changes.

## *2.5.1 A step-wise induction of FT1*

In angiosperms, the induction of *FLOWERING LOCUS T-like* gene in the leaf is an essential part of the vegetative to floral transition, and this role is conserved in wheat, with FT1 acting as the trigger for flowering (Yan *et al.*, 2006; Bratzel & Turck, 2015; Finnegan *et al.*, 2018; Dixon *et al.*, 2018). In this chapter, I observe a previously unidentified step-wise induction of *FT1* expression in the leaf that occurs under field conditions. There is an initial rise in transcript levels under 11 h daylengths with a second rise occurring as the photoperiod extends to 13 h (Figure 2.8c). The initial induction of *FT1* at 11 h daylengths demonstrates that long days are not required for the floral transition to occur in wheat. The development of the inflorescence meristem from vegetative to double ridge stage, correlates with this initial induction at 11 h, with development to the terminal spikelet occurring in early spring – a process that is promoted by *FT1* (González-Navarro *et al.*, 2015; Dixon *et al.*, 2018).

The rise in *FT1* expression observed at 13 h daylengths is consistent with the increase in expression observed under controlled conditions when plants are shifted to extreme long days (22 h/2 h). It is also consistent with studies in *Arabidopsis* grown under natural summer photoperiods where *FT* expression levels were higher under 16 h compared to 14 h photoperiods (Dixon *et al.*, 2018; Song *et al.*, 2018). Another outcome of this study is that the diel pattern of *FT1* gene expression is different in the field and glasshouse compared to that reported in wheat and barley under controlled environment conditions. Here, I observe peaks in expression at both the morning and the afternoon. Whereas in

controlled environments with 16 h daylengths there is a dominant peak in the evening at lights off (Boden *et al.*, 2015; Campoli *et al.*, 2012; Chen *et al.*, 2014; Turner *et al.*, 2005). These results are consistent with a recent study in *Arabidopsis* that compared plants grown in natural photoperiods with those grown under controlled laboratory conditions. This report shows both light quality and temperature are responsible for the peak of *FT* expression during the morning, which was not represented under controlled conditions (Song *et al.*, 2018). These data demonstrated that by carrying out field-based experiments you can uncover a level of complexity underpinning the seasonal regulation of flowering time that would not be readily detectable using laboratory-based analyses.

Genetic variation for *Ppd-1* has a large effect on flowering time through varying *FT1* activity. The photoperiod insensitive alleles cause higher *FT1* expression under controlled laboratory conditions, that consequently results in earlier flowering (Beales *et al.*, 2007; Boden *et al.*, 2015; Shaw *et al.*, 2012, 2013; Wilhelm *et al.*, 2009). Here I show that *Ppd-D1a* photoperiod insensitive lines cause *FT1* to be expressed both earlier in the season and to a higher level relative to wild-type. The initial increase in expression is comparable to the transcript levels, which are detected in wild-type late in the growing season (Figure 2.8)

The impact of *Ppd-D1a* on *FT1* expression is highly significant when the daylengths are 10-12 h long, which correlates to photoperiods at which the early stages of inflorescence development occur in the field and spikelet number is decided (Figure 2.6; Figure 2.8) (González-Navarro *et al.*, 2015). The dramatic effect that the insensitive line has on *FT1* expression during short days is consistent with the acceleration of flowering time that is observed in these NILs under a constant 9 h photoperiod. Under short days, these insensitive lines flower 60 to 190 days earlier relative to wild-type but only accelerate flowering by 6-10 days under normal field conditions (Figure 2.8e) (Beales *et al.*, 2007; Bentley *et al.*, 2013; Cane *et al.*, 2013; Díaz *et al.*, 2012; Worland *et al.*, 1998).

## 2.5.2 Regulation of *Ppd-1* expression

It is interesting that the mis-regulation of *Ppd-D1* expression in the photoperiod insensitive line was less dramatic in the field compared to glasshouse and laboratory conditions (Figure 2.4). In the field, the constant high evening expression was not observed at the 9 h photoperiod and as the photoperiod extended was limited to the late hours of the night (Beales *et al.*, 2007; Boden *et al.*, 2015). Considering that field and glasshouse plants were both grown under natural photoperiods, the altered *Ppd-1* expression is likely to be attributed to differences in temperature and/or light quality, this is consistent with *Ppd-1* being regulated by phytochromes, the photoreceptors which detect both light quality and temperature in plants (Chen *et al.*, 2014; Jung *et al.*, 2016; Rockwell *et al.*, 2006). In the

*ppd-1* null lines, *FT1* expression still responded to photoperiod and displayed the same diel pattern as wild-type plants, however *FT1* transcript levels were significantly lower relative to wild-type (Figure 2.8). Interestingly, these data do not agree with analysis in barley under controlled environment conditions, where barley plants with loss of function *Ppd-1* alleles do not express *FT* under long day photoperiods (Turner *et al.*, 2005; Campoli *et al.*, 2012). Taken together, I conclude that photoperiod insensitive alleles promote the expression of *FT1* both to higher levels and earlier in the season, relative to wild-type. In addition, *Ppd-1* is not the only factor regulating the seasonal regulation of *FT1.*

In this study, I observe transcriptional interaction between *Ppd-1* homoeologues (Figure 2.3). This observation is consistent with previous studies that have shown *Ppd-B1* expression is higher in *Ppd-D1a* mutants that contain a splice site mutation (Boden *et al.*, 2015). This same effect has not been reported in barley; however, the insensitive alleles used in barley contain missense mutation that are likely to have a much weaker effect compared to the null alleles used in this study (Turner *et al.*, 2005; Campoli *et al.*, 2012; Ejaz & von Korff, 2017). In this study, strong evidence of a feedback loop regulating *Ppd-1* expression is presented (Figure 2.4). Feedback loops are prevalent in biology, particularly in the circadian clock that is based upon them. In the clock, the negative feedback loops are highly conserved and generate the circadian rhythms that mediate many processes at a molecular level (Weitz *et al.*, 2013).

In Arabidopsis, the central oscillator of the circadian clock is formed by the genes *CCA1*, *LHY* and *TOC1*, which function though a loop of negative transcriptional feedback (Schaffer *et al.*, 1998; Nakamichi *et al.*, 2005; Pokhilko *et al.*, 2012). The PRR proteins have been shown to be part of this negative feedback loop, binding to the promoters of *CCA1* and *LHY* during the day to regulate expression (Alabadí *et al.*, 2001). This chapter presents data providing evidence for a feedback regulating the expression of *Ppd-1*, itself a PRR protein. Higher levels of *Ppd-D1* transcripts, when there is non-functional *Ppd-1* protein present suggests the feedback loop is negative. In addition, the identification of potential binding sites of ARR, PIF and TOC1 proteins suggest regulation both directly and as part of the wider clock (Figure 2.5). Investigating whether this negative feedback is direct or indirect is an interesting area for future study.

The mechanism of how Ppd-1 controls the transcription of *FT1* is unknown. Here, I see evidence for *Ppd-1* as a transcription factor, with *Ppd-1* protein localizing to the nucleus (Figure 2.10). The localisation did not occur in every cell, suggesting a dynamic localisation in response to stimulus, however future studies are required to confirm this. Other PRR proteins have been shown to be involved in the flowering pathway*.* In Arabidopsis, PRR proteins help stabilize the transcription factor CO under long days

(Hayama *et al.*, 2017). This direct interaction mediates CO binding to the *FT* promoter. Localisation of PRR proteins to the nucleus has been observed in *Arabidopsis* protoplasts (Hayama *et al.*, 2017). Further investigation of Ppd-1 protein dynamics, in response to different light conditions will be an interesting area for future study.

For flowering to occur a signal must be transmitted from the leaf to the shoot apical meristem, in wheat this signal is predicted to be FT1 protein, from which inflorescence development in promoted through the expression of key meristem identity genes (Corbesier *et al.*, 2007; Li & Dubcovsky, 2008; Tamaki *et al.*, 2007). In both *Arabidopsis* and rice, it has been shown that over expression of *FT*/*Hd3a* (*Heading date 3a*) results in hyper-activation of meristem identity genes (Igor Kardailsky *et al.*, 1999; Yoo *et al.*, 2005; Taoka *et al.*, 2011; Kaneko-Suzuki *et al.*, 2018). I hypothesized that meristem identity genes would be higher expressed in *Ppd-D1a* NILs, relative to wild-type, because *FT1* expression is higher in photoperiod insensitive lines this combined with a recorded reduction in expression of meristem identity genes in *ft-b1* mutants (Beales *et al.*, 2007; Boden *et al.*, 2015; Shaw *et al.*, 2013).

A surprising result was that the whilst the insensitive line accelerated the expression of meristem genes to earlier in the growing season, relative to wild-type, the amplitude of transcripts was identical (Figure 2.14). In *ppd-1* null lines, however, the induction of the meristem identity gene expression was delayed and the transcript level lower relative to wild-type, corresponding to the reduction in *FT1* activity in these lines. The step-wise increase in *FT1* expression in wild-type plants closely aligns with the seasonal up-regulation of meristem identity genes, the first induction closely correlates with the floral transition from a vegetative state to the double ridge stage, and the second, stronger induction correlates with the promotion onwards from the double ridge stage (Figure 2.14).

The earlier expression of the meristem identity genes in the *Ppd-D1a* lines, relative to wild-type, correlated with the accelerated arrival of the inflorescence at the terminal spikelet stage, occurring earlier in the season. These data indicate that the impact on inflorescence meristem development and spikelet number by *Ppd-1* allelism is not determined by absolute levels of meristem identity genes, but rather by the timing of expression peaks. The behaviour of *FT1* in the *Ppd-D1a* lines reported here is consistent with the ability of these lines to accelerate past the lemma primordium stage to terminal spikelet when the daylengths are maintained at 10 h, relative to wild-type and *ppd-1* (Figure 2.6)*.* As well as reducing spikelet number, the photoperiod insensitive lines reduce floret fertility; this decrease in floret fertility may be explained by the expression of *HOX2* at later stages in the *Ppd-D1a* lines, relative to wild-type (Prieto *et al.*, 2018; Sakuma *et al.*, 2019). Together the examination of seasonal *FT1* expression presented in this study, alongside the phenotypic analysis of inflorescence development, points to an intimate

connection between floral signals generated in the leaves with the development of the inflorescence meristem, illustrating how the wheat meristem dynamically responds to environmental signals.

### 2.5.3 *FT-like* genes in the inflorescence

This chapter has identified *FT2* as a key regulator of spikelet development in wheat. This importance is shown by the higher number of spikelets in the loss-of-function mutants in addition to the significant increase in expression when the plant reaches the lemma primordia stage (Figure 2.12). I propose *FT2* is involved in determining the number of spikelets on a developing inflorescence by promoting the transition past the lemma primordium stage to the terminal spikelet stage, accelerating or delaying this transition decreases or increases the final number of spikelets. This conclusion is supported by evidence from tetraploid wheat, whereby it was shown *FT2* is expressed strongly in the developing inflorescence, and loss of function both delays flowering and increases spikelet number (Shaw *et al.*, 2019).

Evidence for the role of *FT2* in spikelet number is reflected in its expression in the *Ppd-1* NILs. The expression of *FT2* is inversely proportional to final spikelet number and inflorescence development rate (Figure 2.12, 2.14). The accelerated increase in *FT2* expression in the photoperiod insensitive line, could explain the increased ability of the inflorescence meristem to develop rapidly through the early developmental stages to TS, even when the plants were maintained under 10 h short days. The expression of *FT2* in wild-type plants is photoperiod dependent with expression being much lower in plants maintained at 10 h relative to plants grown under natural photoperiods (Figure 2.12). There is evidence for *FT* paralogues regulating reproductive development in rice, barley, beet, poplar, onion, *Arabidopsis* and *Brachypodium,* all these plants have shown *FT-like* genes in addition to *FT/FT1/Hd3a* influencing reproductive development phases, which include vernalization responsiveness, floral transition, spikelet development and bulb formation (Yamaguchi *et al.*, 2005; Hsu *et al.*, 2006; Kikuchi *et al.*, 2009; Pin *et al.*, 2010; Lee *et al.*, 2013; Mulki *et al.*, 2018; Woods *et al.*, 2019).

Notably, the increase of *FT2* expression from the lemma primordium to terminal spikelet stages and its relationship with spikelet number is the opposite to the expression profiles of genes including *SEP1-6, AP1-3* and *AP1-2* (Figure 2.14). The expression of these genes increases earlier in the inflorescence meristem relative to *FT2*, and are associated with spike architecture phenotypes, including multi-row and paired spikelet formation (Boden *et al.*, 2015; Dobrovolskaya *et al.*, 2015). Together with the data described here, these results suggest a system wherein genes that are expressed earlier in development regulate inflorescence architecture, whereas genes expressed later in development

between the lemma primordium and terminal spikelet stages contribute to the determination of spikelet number. This new understanding of inflorescence development could provide new strategies for identifying genes to increase spikelet number and consequently yield (Wang *et al.*, 2017; Kuzay *et al.*, 2019). To this end, genes such as *WHEAT ORTHOLOGUE OF APO1* (*WAPO1*) and *TERMINAL FLOWER 1* (*TFL1*) have higher expression during the glume and floret primordium stages and are associated with increased spikelet number (Wang *et al.*, 2017; Kuzay *et al.*, 2019).

Interestingly I have recorded another induction of *FT1* in the inflorescence at the green anther stage, at this point it is expressed very highly, specifically in the rachis. This stage is crucial for floret development and so may indicate a role for photoperiod at this late stage of development (Kirby & Appleyard, 1984). The plants reach GA during long days (15 h/ 9 h), with expression being drastically increased by a photoperiod insensitive allele of *Ppd-1* (Figure 2.13b). Together these data point to another photoperiod dependent checkpoint late in inflorescence development, characterized by induction of *FT1* in the meristem itself, that to our knowledge has not been identified in any plant before this study.

In summary this chapter has provided a unique insight into the dynamic and complex regulation of the molecular processes that control inflorescence development and flowering time in wheat. The analysis presented here exhibits the *Ppd-1*-dependent regulation of *FT1*, showing that a step-wise increase in *FT1* expression mediates early inflorescence development. This control is overridden in *Ppd-D1a* insensitive lines that transition to terminal spikelet without delay. I demonstrate the important role of *FT2* in regulating spikelet number, highlighting the varied role of *FT-like* genes in floral development. The knowledge gained here reveals the importance of complementing laboratory-based analysis with field-based experiments, especially when considering flowering or other season responding processes. These conclusions are supported by recent work form *Arabidopsis,* rice and oil seed tape that have demonstrated seasonal regulation of vernalization and flowering time (Duncan *et al.*, 2015; Gómez-Ariza *et al.*, 2015; Hepworth *et al.*, 2018; Song *et al.*, 2018; O'Neill *et al.*, 2019). The data presented here highlights intimate crosstalk between the leaf and Inflorescence, and chapter 3 will explore this connection in greater detail.

## 2.6 **Methods**

### 2.6.1 Plant material and growth conditions

This study used genotypes of spring hexaploid wheat (*Triticum aestivum*) *cv.* Paragon; photoperiod sensitive wild-type Paragon, *Ppd-D1a* photoperiod insensitive lines (Shaw *et*

*al.*, 2012; Bentley *et al.*, 2011) and *ppd-1* lines carrying null *ppd-1* alleles on the A, B and D genomes (Shaw *et al.*, 2013). In addition, two missense *ft-b2* mutants *ft-b2.1* (*Cadenza0122*) and *ft-b2.2* (*Cadenza1655*) were obtained from the hexaploid wheat TILLING population (Krasileva *et al.*, 2017), these mutant NILs were compared to both *cv.* Cadenza and sibling lines containing the wild type allele. The TILLING point mutations were verified using segregation analysis on two independent mutations. Transgenic UBI::PPD-B1::GFP lines were developed with BRACT (John Innes Centre, Norwich, UK). The construct contains the *Ppd-B1* coding region as determined through the Ensembl Plants website (https://plants.ensembl.org/index.html).

Wild-type Paragon and the *Ppd-D1a* and *ppd-1* NILs were grown at field sites of the John Innes Centre based at Church Farm, John Innes Centre, Bawburgh, Norfolk, UK (52°62'25.7"N, 1°21'83.2"E) in 1 m$^2$ plots. These field grown plants experienced natural increases in photoperiod and temperature (Figure 2.15a). The Paragon lines were also grown in a glasshouse under natural photoperiod and temperature conditions. Seed were sown between week two and three of October, this allows germination to occur over winter with subsequent flowering occurring in spring. Phenotype data were collected over the 2017 and 2018 growing seasons and gene expression data collected over the 2017 and 2019 growing seasons. The *ft-b2.1* TILLING lines and wild-type Cadenza were grown under 16 h light/ 8 h dark and 16 °C in a controlled glasshouse, whereas the *ft-b2.2* lines with relevant controls were grown under (22 h/ 2 h) conditions. The UBI::PPD-1::GFP lines were grown in a controlled environment room under 16 h light/ 8 h dark and 20 °C day temperature and 15 °C night. Light intensity was set at 300 µmol/m2/s (using Plantastar 400-W HQI bulbs (Osram) and Maxim 60-W tungsten bulbs).

To conduct the moving bench (MB) experiment, wild-type, *Ppd-D1a* and *ppd-1* lines were germinated and grown under natural temperature and photoperiod conditions in a uncontrolled glasshouse. When the daylength reached 10 h, plants were shifted to a moving bench that maintained daylength at 10 h / 14 h night (Figure 2.15b). This experiment was repeated over two growing seasons 2018 and 2019.

## 2.6.2 Inflorescence architecture measurements

For fully developed inflorescences, all spikelets (viable and inviable) were counted. Spikelet number was counted at each stage of inflorescence development. Spikelet meristems are first visible at the double ridge stage (DR), and at this stage the spikelet meristem ridge (the upper ridge) was counted as a spikelet. As the inflorescence develops the spikelet meristem becomes more pronounced and is clearly defined at the lemma primordium (LP) and terminal spikelet (TS) stages where it is counted as a spikelet. Spikelet number for early inflorescences are the average ± SEM of at least 3 replicates.

For fully developed inflorescences, the final spikelet numbers are the average ± SEM for at least 10 replicates. For the *ft-b2* TILLING mutants, spikelet numbers are the average ± SEM of at least 7 replicates.
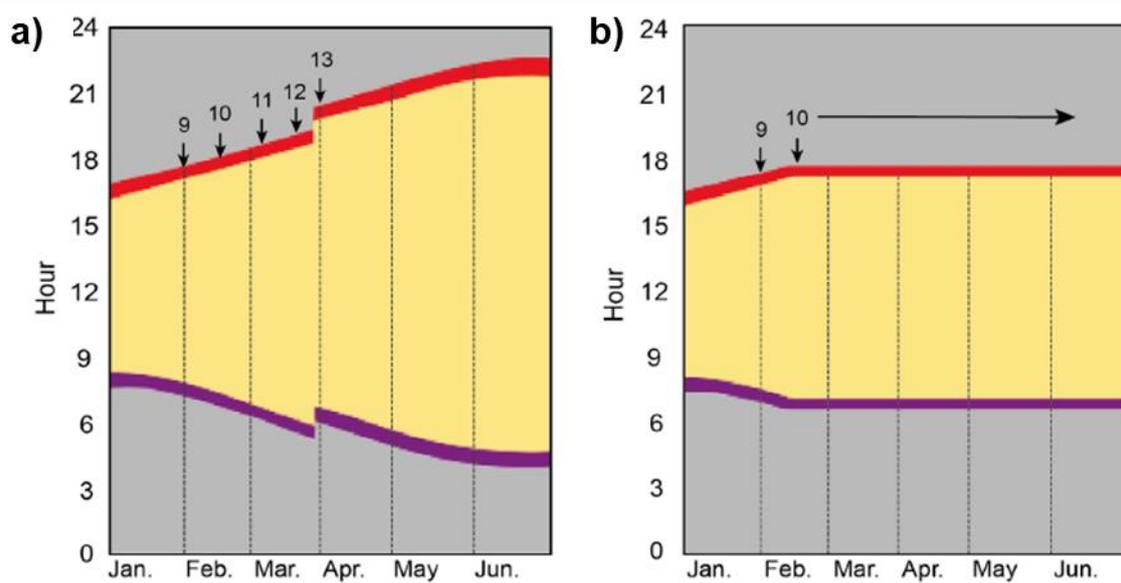


**Figure 2.15 - Experimental design of the leaf gene expression and photoperiod shift experiment**

Schematic diagrams of seasonal changes in daylength for plants grown **a)** under natural photoperiods, and **b)** plants that were grown under natural photoperiods until daylengths were 10 h, and subsequently shifted to constant 10 h conditions. Dawn (purple), daylight (yellow), dusk (red) and night (grey) are represented. Arrows denote when a photoperiod checkpoint was reached from 9 h to 13 h. Graphs show day hours against months of the year.

## 2.6.3 Grain phenotyping

The measurements for grain width, area, length and a thousand grain weight (TGW) estimate were carried out on more than 50 grains per genotype using a MARVIN grain analyser (GTA Sensorik GmbH, Germany). Data are the average ± SEM of at least 5 biological replicates.

## 2.6.4 Heading date measurements

The recorded heading dates for all field and natural glasshouse grown plants are the calendar date of flowering. In the field, the date was determined per plot, with the day

marked when 50% of the inflorescence had emerged from 50% of plants in a given plot. The data is the average ± SEM of 5 plots. For the glasshouse experiments the heading date was recorded on an individual plant basis. The heading date for these plants is the ± SEM of 20 plants in the glasshouse and 7 replicates in the moving bench experiment.

## 2.6.5 Sampling of *Ppd-1* lines for leaf gene expression analysis

Wild-type Paragon, and *Ppd-D1a* and *ppd-1* NILs leaves were sampled every 3 hours during the day and every 4 hours during the night, and at dusk. The sample time-points are described in terms of time since dawn (TSD), with dawn regarded as 0 h. For both the field and natural glasshouse experiments, the most recently emerged leaf was sampled at 1 h increments in daylength increases, the first sample daylength was the 9 h photoperiod and last was 13 h (Figure 2.15a). For each time point, three biological replicates were collected. Samples were placed in a 2 ml collection tubes and immediately frozen in liquid nitrogen.

## 2.6.6 Sampling of *Ppd-1* lines for inflorescence meristem phenotyping and gene expression analysis

As the plants developed in the field and glasshouse, representative inflorescences were examined using a Leica MZ16 Stereo binocular dissecting microscope and imaged using a Leica DFC420 colour camera. Inflorescences were imaged at regular intervals, including when the daylength naturally increased by 1 h to correlate with leaf gene expression studies. Plants from the field were transplanted into pots and transported to the lab for processing. At least three inflorescences per genotype were phenotyped and imaged at each sample point. For expression analysis, inflorescences were collected from wild-type, *Ppd-D1a* and *ppd-1* NILs when they reached one of 4 key stages as determined by (Kirby & Appleyard, 1984); vegetative (VG), double ridge (DR), lemma primordium (LP) and terminal spikelet (TS). Three replicates were collected per stage, with each sample a pool of 5 – 15 inflorescences per replicate dependent on stage, which were immediately frozen using liquid nitrogen. For glasshouse plants additional sampling was performed at the white anther (WA) and green anther (GA) stages. At the GA stage glumes, florets and rachis were dissected using the same method used for inflorescences. When measuring the rate of inflorescence development measurements consisted of the 4 – 5 biological replicates.

## 2.6.7 RNA extraction and expression analysis

To extract RNA from the leaf a Spectrum Plant Total RNA Kit (Sigma-Aldrich, St. Louis, United States) was used according to manufacturers instructions. To extract RNA from

developing inflorescences a RNeasy Plant Mini Kit (Qiagen, Hilden, Germany) was used according to manufacturer's instructions. RNA was treated with RQ1 DNase I (Promega, Madison, United States) and subsequently reverse-transcribed with SuperScript III reverse transcriptase (Life Technologies, Carlsbad, United States), as per manufacturer's instructions. Quantitative RT-PCR was carried out on a LightCycler® 480 (Roche, Basel, Switzerland) using GoTaq® qPCR and RT-qPCR Systems (Promega) as per manufacturer's instructions. The oligonucleotides used for RT-qPCR analysis are listed in (Table 2.1). Candidate gene expression from leaf and inflorescence was normalised using TraesCS6D02G145100 (Traes_6DS_ BE8B5E56D.1; Borrill *et al.*, 2016). All RT-qPCR data points are the average of three biological replicates, with two technical replicates performed for each reaction.

## 2.6.8 DNA extractions and sequence analysis

Leaf tissue from seedlings were sampled and genomic DNA extracted as described previously (Paterson *et al.*, 1993). The adapted protocol is as follows: Using a 96 well 3 mm bead dispenser (Qiagen), add one 3 mm Tungsten Carbine bead (Qiagen) to each well of a 96-well 1.2 mL storage plate (ABgene, Epsom, UK); add a leaf tissue sample approximately of 100mg, whilst maintaining box and samples on dry ice; grind tissue using a Geno/Grinder® - Automated Tissue Homogenizer and Cell Lyser (Spex, Metuchen, USA) for 45 seconds at 1250 rpm; remove from ice box, and add 200µl of warm cotton lysis buffer to each tube. Add new caps to avoid contamination; invert several times and place box in 65 °C for 1 h. Invert at 20 minute intervals to ensure mixing. After 1 h place in freezer to cool for 10 minutes; Add 150 µl chloroform and spin for 10 minutes at full speed; remove 115 µl of supernatant and add to new 96-well 1.2 mL storage plate; add 115 µl isopropanol to wells, seal and spin for 45 min; flick off isopropanol, add 200 µl 70% ethanol, spin for 20 minutes at 5000 rpm; leave plate to dry and add 50 µl dH2O to suspend pellet.

All gene sequences were obtained by BLAST search from the Ensembl Plants website (https://plants.ensembl.org/index.html). The primers used to amplify and sequence the *ft-b2* mutant alleles are listed in (Table 2.1). DNA fragments were amplified using Phusion DNA polymerase (New England Biolabs, Ipswich, USA), and resulting amplicon sequencing was carried out with Mix2Seq Kit (Eurofins, Luxembourg, Luxembourg).

## 2.6.9 Kompetitive Allele-Specific PCR analysis

To analyse TILLING mutant lines, Kompetitive allele-specific PCR (KASP) analysis was performed on extracted gDNA. Oligonucleotides to verify mutant were designed using Polymarker (Ramirez-Gonzalez *et al.*, 2015), these oligos contain wither FAM or HEX

compatible tails (FAM tail, 5'-GAAGGTGACCAAGTTCATGCT-3' HEX tail, 5'-GAAGGTCGGAGTCAACGG -ATT-3') (Table 2.1). The KASP assay was carried out as described previously (Ramirez-Gonzalez *et al.*, 2015). In summary:

KASP assays were performed using a standard recipe and protocol:

**PCR recipe:** Template 10-20 ng, Primer mix 0.07 µL, KASP mix 2.43 µL, ddH2O to volume of 5 µL.

**KASP protocol:**

Hotstart at 95 °C for 15 min, followed by 10 touchdown cycles (95 °C for 20 s; touchdown 65 °C, 21 °C per cycle, 25 s), followed by 40 cycles of amplification (95 °C for 10 s; 57 °C for 60 s).

## 2.6.10    Confocal microscopy

Confocal analysis was performed on overexpressed GFP tagged Ppd-B1 protein (UBI::PPD-B1:GFP) using a Zeiss LSM 780 confocal microscope with a 40/1.2 water immersion objective lens. A 488 nm line from an argon ion laser was used to excite GFP fluorescence a GAsP spectral array 499-579 nm was then used to capture the fluorescence.

## 2.6.11    Promoter region analysis

The 2 kb promoter sequence of *Ppd-D1* was downloaded from Ensembl Plants. Promoter transcription factor binding site analysis was performed using PlantPAN 2.0 (Chow *et al.*, 2016).

## 2.6.12    Statistical analysis

Statistical differences between treatments and sample points were tested by two-tailed Students *t* test. Even distribution of data was determined using a Barlett's test. Data in figures are mean ± standard error of the mean (SEM). *P* values for gene expression comparisons in this Chapter can be found in Appendix 2.

**Table 2.2 - Oligonucleotides used in chapter 2**

| Gene | Direction | Type | Primer | Source |
|------|-----------|------|--------|--------|
| *SEP1-6* | Forward | Q-PCR | CCTCTACCAGTTCTCCTCCTCC | Boden *et al* |
| | Reverse | Q-PCR | CATATACTCCAGATAGTTGTT | Boden *et al* |
| *VRN1* | Forward | Q-PCR | GGAAACTGAAGGCGAAGGTTGA | Oliver *et al* |
| | Reverse | Q-PCR | TGGTTCTTCCTGGATCTGATATG | Oliver *et al* |
| *AP1-3* | Forward | Q-PCR | TCTATGAGTACGCCACCGACT | Boden *et al* |
| | Reverse | Q-PCR | CACCAATTTCCCTCACTTTCA | Boden *et al* |
| *AP1-2* | Forward | Q-PCR | AGCTCACCGTCACCTACACC | Boden *et al* |
| | Reverse | Q-PCR | TTGTTTGCTTGTGCTGGAGA | Boden *et al* |
| *AGL6* | Forward | Q-PCR | CCAGACAGCGAAAGACACAA | Gauley & |
| | Reverse | Q-PCR | CTTGTGCTTGAGTTGCCTGT | Boden |
| *FT1* | Forward | Q-PCR | GTCGTTCGGGCAGGAG | Shaw *et al* |
| | Reverse | Q-PCR | TGGAAGAGTACGAGCACGA | Shaw *et al* |
| *Ppd-D1* | Forward | Q-PCR | AAGACAAGGCTGATGAAATGAG | Shaw *et al* |
| | Reverse | Q-PCR | GAAGGATTGACCACATTGGA | Shaw *et al* |
| *Ppd-B1* | Forward | Q-PCR | AAGACAAGGTTGATGACGTGA | Shaw *et al* |
| | Reverse | Q-PCR | GAGGGATTGATCACGTTGG | Shaw *et al* |
| *SOC1* | Forward | Q-PCR | CAGCAAGTCAAAGCTGATGC | Pearce *et al* |
| | Reverse | Q-PCR | AACGCGGAGACTCTTCTCAA | Pearce *et al* |
| *HOX2* | Forward | Q-PCR | AGCTTATGGAGGAGGAGTTCG | Gauley & |
| | Reverse | Q-PCR | CTCCCAGCCTCTCCTTCAG | Boden |
| *TOC−B1* | Forward | Q-PCR | TTGAGACGCCTGTGCAG | Shaw *et al* |
| *TOC−B1* | Reverse | Q-PCR | AAAATGGGTAATGATATACAGGAGG | Shaw *et al* |
| *CCA1* | Forward | Q-PCR | CCTGGAATTGGAGATGGAGA | Pearce *et al* |
| *CCA1* | Reverse | Q-PCR | TGAGCATGGCTTCTGATTTG | Pearce *et al* |
| *ELF3* | Forward | Q-PCR | AGCGATTTCCAGCTGCCTTC | Shaw *et al* |
| *ELF3* | Reverse | Q-PCR | TGCGAAGAGGCCAGTCAGTC | Shaw *et al* |
| *FT2* | KASP | Cadenza122 | gaaggtgaccaagttcatgctGGACGGCCTCAGCTCGCAGC | |
| *FT2* | KASP | Cadenza122 | gaaggtcggagtcaacggattGGACGGCCTCAGCTCGCAGT | |
| *FT2* | KASP | Cadenza122 | CGGCCGCCTTCATTAACATA | |
| *FT2* | Forward | Cadenza1655 Sequencing | GAGGTGGTGTGCTACGAGG | Gauley & |
| *FT2* | Reverse | Cadenza1655 Sequencing | CCCATCTGATTCCCGTACGA | Boden |

# Chapter 3     Inflorescence transcriptomics

## 3.2 Chapter summary

In this chapter, I summarise a novel approach to understand the genes controlling inflorescence development in bread wheat. I aim to explore the role of *Ppd-1* after it induces the expression of *FT1* and the impact this induction in the leaf has on the developing inflorescence as described in chapter 2. My results show *Ppd-1* has a dramatic effect on the transcriptome likely through direct and indirect influence on the expression of genes in the developing meristem. This role is genome dependent, likely because of epigenetic effects on the availability of genome copies for regulation.

## 3.3 Introduction

The key yield traits of spikelet number and floret initiation are determined between the early developmental stages of vegetative and terminal spikelet (Kirby & Appleyard, 1984). Inflorescence development and flowering time are intimately linked – plants that flower earlier generally proceed through early floral meristem stages faster than late flowering lines (Shaw *et al.*, 2012, 2013). Given the flowering pathway initiates in the leaf and culminates in the emergence of the spike from the sheath, it is important consider both the leaf and the developing spike when investigating flowering-time. In the leaf, a key component of the flowering pathway that influences inflorescence development is *Photoperiod-1* (*Ppd-1*). *Ppd-1* promotes the expression of *FLOWERING LOCUS T1* (*FT1*), which encodes a protein that subsequently translocates from the leaf to the meristem (Turner *et al.*, 2005; Beales *et al.*, 2007; Díaz *et al.*, 2012). In the meristem, FT1 then interacts with FDL and 14-3-3 proteins to form a complex that induces the expression of key meristem identity genes, such as *VERNALISATION-1* (*VRN1*) (Li & Dubcovsky, 2008; Li, Lin and Dubcovsky, 2015). Throughout the history of wheat breeding, genetic variation for *Ppd-1* has been used to manipulate flowering time and the rate of inflorescence development (Worland, 1996; Worland *et al.*, 1998). A deletion in the promoter region confers a photoperiod insensitive characteristic (*Ppd-1a*) that confers accelerated flowering, as shown extensively under field conditions (Bentley *et al.*, 2011, 2013; Shaw *et al.,* 2012). Conversely, *ppd-1* null lines dramatically delay flowering (Gauley and Boden, 2020; Shaw *et al.*, 2013). Regarding the wheat inflorescence, very little is known about the genes underpinning spikelet and floret development. However, by examining the literature from other cereals and model organisms, we can form an approach to expand our understanding in wheat. For example, MIKC-type MADS-box genes contribute significantly to developmental processes in plants (Schilling *et al.*, 2018). They have key roles during floral meristem development in a variety of species. In wheat,

these genes have contributed to breeding, in particular through the regulation of *VERNALIZATION1* (*VRN1*), which is itself a MADS-box transcription factor (Deng *et al.*, 2015)*.*

## 3.3.1 MADS-box genes and floral development

Across plants, the MADS-box genes have an essential role in floral development. In *Arabidopsis*, floral organ identity is determined by spatially defined patterns of expression for different classes of MADS-box genes (Coen & Meyerowitz, 1991). The regulation of floral development is explained by the ABCDE model, in which the expression of these gene classes corresponds to 4 distinct layers of flower development, named 'whorls' (refer to 1.5.3) (Coen & Meyerowitz, 1991; Rijpkema *et al.*, 2010; Kitagawa *et al.,* 2012; Murai, 2013). Investigating the role of these genes in wheat will help us understand how it regulates floral development. Whilst the ABCDE model of flowering is generally conserved throughout plant species, there are several genes whose expression pattern and function have diversified. In rice, the function of class B genes are generally conserved, specifying the stamen and petals like in *Arabidopsis* – an  example is the rice orthologue to *AP3,* known as *SUPERWOMAN1* (*SPW1*) (Yoshida & Nagato, 2011; Sugiyama *et al.*, 2019). Null *spw1* lines exhibit transformation of stamen into carpels, similar to *ap3* mutants in *Arabidopsis* that transform stamens and petals into carpels and sepals (Nagasawa *et al.*, 2003). A similar effect is seen in mutants of the *YABBY* gene *DROOPING LEAD* (*DL*) in rice, an ortholog of *CRABS CLAW* (*CRC*) from *Arabidopsis* that regulates carpel formation. Rice *dl* mutants fail to develop carpels, which instead differentiate into stamens. In crops such as rice, there have been duplication events for these transcription factors, and as such the regulatory mechanisms have become more complex (Schilling *et al.*, 2018, 2020). For example, *MADS2* and *MADS4* are orthologues of *PI.* The function of these paralogues has diverged, with *MADS2* expressed in the lodicule and the stamen primordia, whereas *MADS4* is only expressed in the stamen primordia. In the lodicule, *MADS2* plays a major role, whereas *MADS4* does not. In wheat, there has been more gene duplications, suggesting there will be greater complexity of regulatory networks to uncover (Schilling *et al.*, 2020).  Whilst little is known about the role of the MADS-box genes in wheat, a recent study reported a genome-wide analysis of MIKC-type (type II) MADS-box genes (Schilling *et al.*, 2020). The study highlights duplications of MADS-box genes, which are likely to be a product of adaptation to a variety of geographical locations and environments. This work highlights that the MADS-box gene family as an important candidate for detailed analysis, as it is likely gene function has diverged from what we know from *Arabidopsis*, as well as other crops.

## 3.3.2 Investigations into the genes underpinning inflorescence development in cereals

Throughout this project, studies have been published that examine the transcriptional events of inflorescence development. The first compares wild-type *Ppd-H1* barley lines with later flowering *ppd-h1* lines using RNA sequencing in the developing inflorescence (Digel *et al.*, 2015). They examined the transcriptome in 4 key early inflorescence stages of barley, ranging from vegetative to the early reproductive meristems, capturing a similar range to the vegetative and terminal spikelet stages in wheat. The authors concluded that the effect of *Ppd-1* is dependent on the induction of *FT1* and *FT*2, and that *Ppd-1* does not play a direct role in inflorescence development. A following study performed transcriptomic analysis in wheat at the double ridge stage, stamen primordia stage, terminal spikelet stage and green anther stage (also known as the Waddington(W) W2.0, W3.0, W4.0 and the W7.5 stages). The authors investigated gene expression between these stages, and point to genes such as the *CLV1* homologs that appear to have a conserved function (Feng *et al.*, 2017). Similar studies using RNA-sequencing to examine early inflorescence development have been carried out in barley and rice (Harrop *et al.*, 2016; Li *et al.*, 2018). These studies have highlighted the large number of genes expressed at these stages using clustering analysis to pull out patterns of expression. They highlight the stage specific expression patterns of key genes including MADS-box and auxin genes. These studies were carried out under controlled conditions, however a defining study that used microarray analysis on early rice inflorescence stages to characterise the roles of *PAP2, OsMADS14, OsMADS15* and *OsMADS18* was carried out on plants grown in the field (Kobayashi *et al.*, 2012).

What separates the results presented in this chapter from similar studies in wheat is the use of a combination of *Ppd-1Da* insensitive and *ppd-1* null NILs to determine the role of *Ppd-1* on meristem gene expression patterns during early inflorescence development. In addition, I consider the impact on key gene families such as auxin response and signalling genes as well as MADS-box genes, all in the context of a polyploid organism. Crucially, my study was performed using meristems collected from the field, so that the investigation of the genes underpinning inflorescence development are responding to the seasonal cues experienced by plants grown in their native environment.

My study used hexaploid bread wheat, specifically the cultivar Paragon (*cv.* Paragon) that has been used as a reference genotype, especially in the UK (AABBDD). The polyploidisation of wheat occurred through hybridisation of diploid genomes (refer to 1.3). Polyploidy itself is common across plants and is thought to confer adaptive plasticity, it is thought that between 30-80% of plants are polyploid (Masterson, 1994; Wood *et al.*, 2009). It has been hypothesised that polyploidy promotes adaptive evolutionary change

providing a relatively faster way to evolve new beneficial alleles (Otto & Whitton, 2000). However, despite the implications of this polyploidy, the effects it has on the transcription of key gene families is poorly understood, especially during development. This study investigates large scale as well as gene family-specific implications of polyploidy.

### 3.3.3 RNA-Sequencing, a powerful tool for genetic discovery

RNA-sequencing (RNA-seq) uses next-generation sequencing to observe the quantity of RNA in a biological sample. It provides a rapid method for identifying genes that may contribute to a given biological process, especially for organisms such as wheat where few genes have been functionally characterised to the standard of reference plants.

The expression is quantified through counting the number of reads mapped to each locus against the reference transcript annotation. The data presented in this chapter is shown as transcripts per million (TPM). TPM is used to normalise the data and examine expression levels among genes and between genotypes (or treatments, in other cases), and it is defined as the number of RNA molecules out of 1,000,000 molecules that came from the gene in question. TPM is very similar to other RNA-seq analysing methods such as RPKM and FPKM. However, it is better suited to comparisons between genes (Li & Dewey, 2011; Wagner *et al.*, 2012). This is because of the order of operations: first, the read counts are divided by the gene length measured in kilobases, producing the reads per kilobase (RPK). The RPK is then divided by 1,000,000 giving you the "per million" scaling factor. Each RPK value in question is then divided by the scaling factor, producing the TPM of each gene. The key difference is that gene length is normalised prior to further analysis. The data presented here allows us to examine the expression of every gene at each major stage of early inflorescence development, namely the vegetative (VG), double ridge (DR), lemma primordium (LP) and terminal spikelet (TS) stages.

To identify expression patterns within the RNA-seq data and determine co-expressed genes within the dataset, I have used the python package *clust* (Abu-Jamous & Kelly, 2018). Before analysis, gene expression data is normalised to the input data to facilitate fair comparisons. To do this, *clust* uses a combination of quantile normalisation to ensure the conditions are distributed equally. All TPM values that are less than 1.0 are set to 1.0, which is followed by calculating the log2 of each TPM to remove the effect of large differences in expression. Further normalisation is carried out using Z-score which is when you subtract the mean of the data and then divide by its standard deviation. As such, if the TPM for your gene of interest is exactly equal to the mean, the normalised value will be 0, if it is below the mean TPM it will be a negative number, and if it is above the mean it will be a positive number. These combinations of normalisation methods allow us to compare the patterns of gene expression independently to the absolute levels of gene expression.

At the inception of this study I hypothesised there would key gene families up- and down-regulated between key stages of inflorescence development. By varying the expression of *Ppd-1* in the leaf, I expected there would be an effect on meristem expression, but that this would be focused on the vegetative to double ridge transition. Furthermore, I hypothesised there would be much more mis-regulation in the *Ppd-D1a* insensitive line compared to the wild-type and *ppd-1* null lines. This hypothesis was based on the assumption that the impact of *Ppd-1* would only be through regulation of *FT1* in the leaf. In this chapter I use *Ppd-1* allelism to determine the stages of inflorescence development that experience the greatest transcriptional changes, identifying clusters of gene expression that mark key developmental stages. I Investigate if gene families are regulated on a family level or gene by gene basis, in addition to visualising the impact of polyploidization on gene expression in the floral meristem.

## 3.4 **Results**

### 3.4.1 *Ppd-1*-dependent regulation of inflorescence gene expression

To examine how *Ppd-D1a* insensitive and *ppd-1* null lines influence gene expression in the developing inflorescence, meristem samples were collected from the field 6 hours after dawn. Four key developmental stages were examined: the vegetative (VG), double ridge (DR), Lemma primordia (LP) and terminal spikelet (TS) stages. These developmental stages were selected as they mark inflorescence development from a vegetative meristem through to spikelet number fixation at the terminal spikelet stage. These genotypes were grown under field conditions and were sampled on different calendar days, consequently the expression data presented in this chapter may be influenced by environmental factors. These results aim to examine the effects of *Ppd-1* allelic variation in field grown plants, this includes the consequences of earlier or delayed plant development and the different environmental factors associated with it.

#### 3.4.1.1 Expression of *Ppd-1* in the floral meristem

*Ppd-1* is a key regulator of flowering; however, the mechanism underlying its function is poorly understood. Before this study, the only characterised role of *Ppd-1* in wheat involved regulation of *FT1* in the leaf, and an influence on selected genes within the developing inflorescence (Shaw *et* al., 2012; Boden *et al.*, 2015). However, here I observe significant expression of *Ppd-1* in the developing inflorescence (Figure 3.1). The B genome copy of *Ppd-1* (*Ppd-B1*) is the primary functional copy in Paragon, and it is expressed highly in the meristem at all stages (Figure 3.1b; Beales *et al.*, 2007).

TreasCSU02G196100.3 is the functional spice variant, with variant
TreasCSU02G196100.1 expressed but lacking a 373 bp region in the first exon. *Ppd-B1*
expression appears to be affected by the insensitive allele of *Ppd-D1*, as *Ppd-B1*
transcripts are much higher at the TS stage in the photoperiod-insensitive NILs, relative to
wild-type. Together, these results indicate that *Ppd-B1* has a role regulating transcription
in the meristem in addition to its reported role in the leaf, and that there is interplay
between the *Ppd-1* homoeologues.

In wild-type plants, the *Ppd-A1* allele is also functional with 3 transcript variants producing
valid gene copies. In the meristem, over all expression of *Ppd-A1* is low with the
TraesCS2A02G081900.3 variant expressed the highest indicating that, like in the leaf,
*Ppd-A1* plays a minor role relative to *Ppd-B1* (Figure 3.1a; Shaw, Turner and Laurie,
2012). The *Ppd-D1* allele is not functional in wild-type plants due to a large deletion
resulting in a non-functional truncated protein. In the leaf, *Ppd-D1* is expressed to the
same amplitude as *Ppd-B1* and displays a very similar diel expression pattern (Figure
2.3). In the developing inflorescence, it is expressed very lowly (<1 TPM) in wild-type,
pointing to a different system of regulation of *Ppd-1* in the floral meristem, relative to the
leaf (Figure 3.1a).

In the photoperiod-insensitive *Ppd-D1a* line, which contains a functional, photoperiod
insensitive allele of *Ppd-1*, there is a small but significant increase in *Ppd-D1* expression
at the lemma primordium and terminal spikelet stages, relative to wild-type ($p$ value <0.5;
Figure 3.1a; Beales *et al.*, 2007). Strikingly, the most dramatic effect is on the *Ppd-B1*
allele, with significant up-regulation at the lemma primordia and terminal spikelet stages,
as expression levels reach 7.1 TPM and 8.63 TPM in the developing inflorescence
compared to 4.1 TPM and 5.25 TPM in wild-type, respectively (Figure 3.1b). The
expression of *Ppd-A1* is similar in wild-type and photoperiod-insensitive NILs (Figure
3.1a). The *ppd-1* null line contains a deletion of *Ppd-B1* and non-function copies of *Ppd-A1* and *D1* (Beales *et al.*, 2007; Shaw *et al.*, 2013). In this line, we observe up-regulation
of *Ppd-D1*, in a similar manner to that observed in the leaf, indicating a feedback loop may
regulate *Ppd-1* in both tissues (Figure 3.1a; Figure 2.4). The expression of the *Ppd-A1*
remains low, whereas there is no expression of the deleted *Ppd-B1* allele.

This analysis points to a role for *Ppd-1* in the developing inflorescence. This chapter will
examine the implications of this expression; however, future studies may examine how the
mechanisms regulating *Ppd-1* differ in the meristem compared to the leaf.
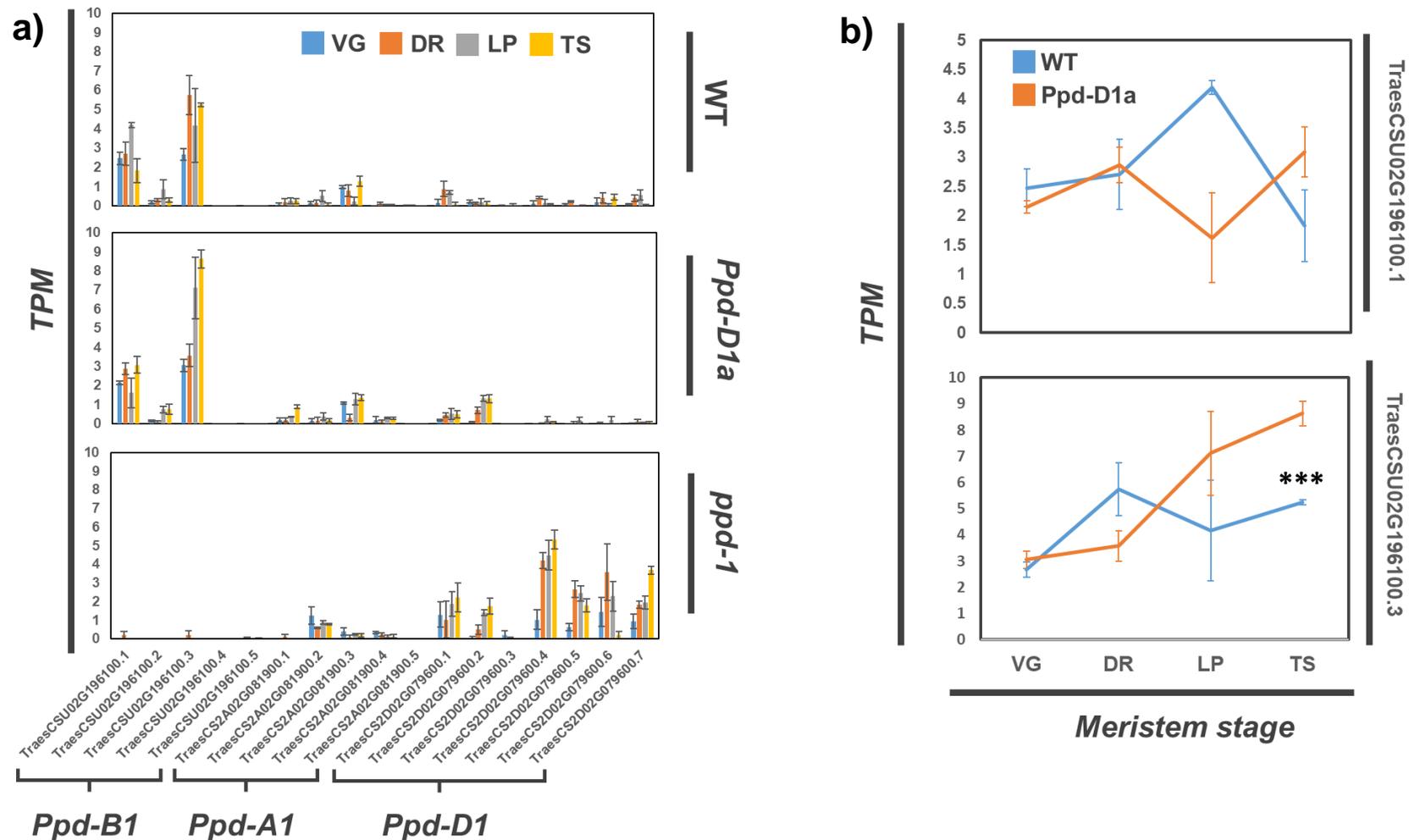
**Figure 3.1 - The expression of all *Ppd-1* genes and spice variants in the developing inflorescence**

**a)** The expression of each of the *Ppd-1* alleles and splice variants in wheat. The B genome copy is: TraesCSU02G196100 with 4 splice variants, the .3 transcript being functional. The A genome copy gene ID is TraesCS2A02G081900 with 4 splice variants. The D genome copy is TraesCS2D02G079600 with 7 splice variants. Data shown is for vegetative (VG)= blue, double ridge (DR) = orange, lemma primordium (LP)= grey and terminal spikelet (TS)= yellow. **b)** Shows the expression of *Ppd-B1* (TreasCSU02G196100.3*)* in the developing inflorescence stages VG, DR, LP, TS. Blue = wild-type, orange = *Ppd-D1a, Ppd-B1* showed no expression in the *ppd-1* NIL. Error bars ± SEM of three biological replicates. *** $P < 0.001$.

### 3.4.1.2 There are significant differences in gene expression between stages and genotypes

To observe the impact of *Ppd-1* allelism on the developing floral meristem, I initially examined the number of genes that were significantly (q value < 0.05 and a difference of > 0.5 TPM) up- and down-regulated between stages and genotypes (Figure 3.2). In wild-type, there are a large number of genes expressed at the VG, DR, LP and TS stages; 62880, 64197, 65376 and 64597, respectively. A focal point for differential expression is the VG to DR transition, where 3786 and 5098 genes are up- and down-regulated, respectively (Figure 3.2). This is the point at which we hypothesised the greatest activation of gene expression in the meristem would occur. However, many more genes are down-regulated than up-regulated, suggesting more genes are involved in repressing the transition opposed to promoting it. As the meristem continues to develop, many genes continue to be differentially regulated. At the DR to LP stage there are twice as many genes up-regulated compared to down-regulated, whereas from the LP to TS stage there are fewer genes differentially expressed, with a similar number of genes up- and down-regulated. At each transition between stages, the *Ppd-D1a* line shows a similar trend to wild-type, however, 1873 fewer genes are down-regulated at the VG to DR transition. As the stages progress, overall, more genes are down-regulated in the insensitive line with a similar number of genes up-regulated between each transition.

Regarding the differences between wild-type and the *Ppd-D1a* NIL at each key stage, many genes are differentially expressed (Figure 3.2). However, comparatively less genes were detected than those between the stage transitions. For example, at the DR stage, there are only 78 genes up-regulated and 68 down-regulated between the genotypes, compared to thousands at the transition from VG to DR. At the other stages, many more genes are differentially expressed relative to the DR stage. For example, both the LP and TS stages have approximately 650 genes down-regulated in the *Ppd-D1a* line, compared to approximately 400 up-regulated. While at the VG stage, 1134 genes are down-regulated in the *Ppd-D1a* line, relative to WT, compared to 484 up-regulated. This indicates that an equally important impact of the photoperiod-insensitive allele is to down-regulate genes that may be involved in the repression of meristem development, in addition to up-regulating genes that promote floral development (Figure 2.14). Overall, the *ppd-1* null line influences the expression of significantly more genes during the early stages of inflorescence development, relative to wild-type or *Ppd-D1a* NILs (Figure 3.2). There are two-fold more genes differentially regulated at every stage for both up- and down-regulated genes, relative to wild-type. For example, between the DR and LP stage, there are 4765 up-regulated and 4311 down-regulated compared to 1938 and 1018 respectively in wild-type, the same trend applies for the LP to TS transition. The greatest point of differential regulation occurs at the VG to DR stage transition, with 7076 genes

up-regulated and 7215 genes down-regulated in the *ppd-1* NIL, compared to wild-type that has 3738 genes up-regulated and 5098 genes down-regulated between the same transition. Notably, the *ppd-1* NIL, has a very similar number of genes that are up- and down- regulated between each stage, which is greater than the changes observed in wild-type, indicating a more general disruption of gene expression.

These results are surprising if you consider the proposed role of *Ppd-1* as an initiator of flowering in the leaf (Beales *et al.*, 2007; Shaw *et al.*, 2012; Boden *et al.*, 2015). They suggest an additional role for *Ppd-1* in the meristem both as an activator and repressor of gene expression at these key developmental stages. These data indicate that in addition to *Ppd-1* having an essential role in the leaf, it has another essential role as a transcriptional regulator in the meristem.
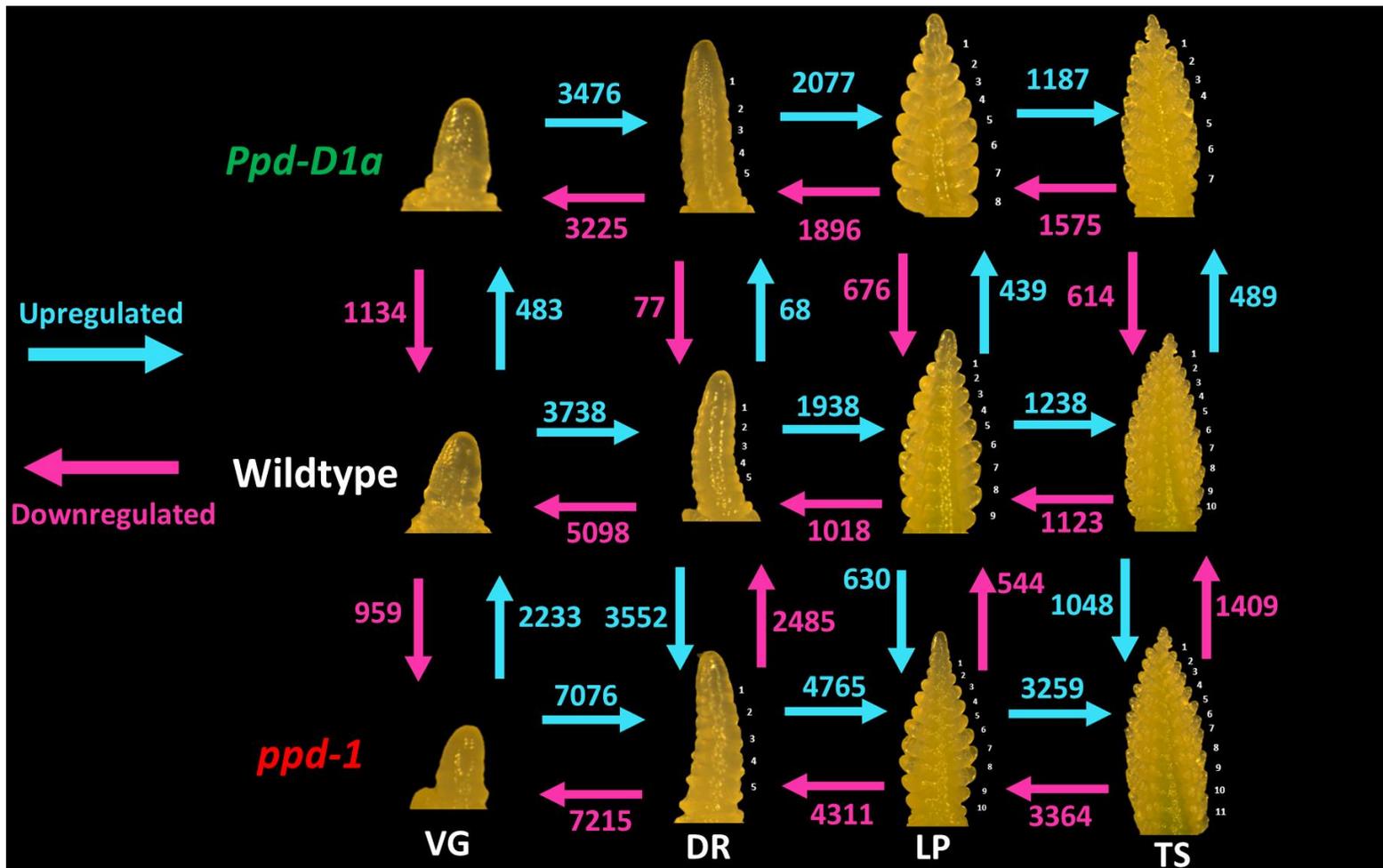
**Figure 3.2 - Summary of the transcriptional changes that occur across early inflorescence development mediated by _Ppd-1_**

The number of genes differentially regulated between stages and genotypes. Stages sampled are the vegetative (VG), double ridge (DR), lemma primordium (LP) and terminal spikelet (TS) stages of inflorescence development in the wild-type, _Ppd-D1a_ and _ppd-1_ lines. Up-regulated genes are denoted by a cyan arrow, down-regulated genes by the magenta arrows. A gene was considered differentially expressed with a q value < 0.05.

### 3.4.1.3 Clustered gene expression profiles representative of key stages

In addition to examining gene expression levels, it is important to consider their profiles across the stages being examined. Knowing where a gene peaks and is down-regulated gives an insight into the stage where the gene performs its role. Using *clust*, I clustered the expression profiles of the top 30,000 expressed genes in the developing floral meristem (Abu-Jamous & Kelly, 2018). I hypothesised that this analysis would provide stage-specific profiles for each genotype, and that genes would shift their expression to be earlier in the *Ppd-D1a* line and later in the *ppd-1* null line.

In wild-type plants, I identified individual clusters that represent the key stages of meristem development (Figure 3.3). The cluster representing the VG stage is characterised by profiles that show significant down-regulation from the VG to DR transition, remaining low until TS (Figure 3.3a). Within this cluster, genes are enriched for gene ontology (GO) terms related to transferase activity, catalytic activity and serine/threonine kinase activity (Table 3.1). The cluster representative of the DR stage in wild-type shows a large peak in expression at the DR stage, with a gradual down-regulation as the stages transition to TS (Figure 3.3b). In this cluster, the genes are enriched for GO terms related to binding, including heterocyclic compound binding and organic cyclic compound binding (Table 3.1). I also identified a LP representative cluster (Figure 3.3c). This cluster experiences flatlined expression from the VG to DR stage with a dramatic peak at the LP stage and subsequent down-regulation at the TS stage to levels comparable with the VG and DR stages. In this cluster, genes are enriched for GO terms related to protein heterodimerization activity, carboxylic ester hydrolase activity and glutathione binding (Table 3.1). For the TS stage, the representative cluster chosen shows a gradual decrease from the VG to LP stage and a subsequent dramatic up-regulation from the LP to TS transition (Figure 3.3d). Within this cluster there are genes enriched in GO terms indicating binding, heterocyclic compound binding and organic cyclic compound binding were highlighted (Table 3.1), these terms are similar to those identified at the DR stage.

In the *Ppd-D1a* and *ppd-1* lines, the expression of genes detected within each of the wild-type clusters are highly disrupted. Examining the VG-specific cluster of wild-type in the *Ppd-D1a* NIL shows that the same genes do not, on average, experience the same degree of down-regulation during the VG to DR transition, relative to wild-type (Figure 3.3a). However, from the LP to TS stages, the patterns of expression in *Ppd-D1a* remain similar to those detected in wild-type. This indicates that mis-regulation of genes caused by the insensitive allele is focused to the VG to DR transition. Conversely, in the *ppd-1* line there is a comparable down-regulation of expression between the VG to DR stages, relative to wild-type. However, unlike in wild-type and *Ppd-D1a* NILs, there is significant

up-regulation of these genes in the LP to TS transition, with average expression at the TS stage reaching comparable levels to the VG stage. The genes represented in the wild-type DR cluster show maintained expression profiles in the *Ppd-D1a* line from the VG to LP stage (Figure 3.3b). Whereas at the LP to TS transition, these genes are up-regulated in the photoperiod-insensitive NILs despite being down-regulated in wild-type. In the *ppd-1* line, the expression profiles are generally similar to wild-type, but the peak of transcripts is dampened from VG to DR stage. These results indicate that *Ppd-1* is required to maintain the developmental progression of gene expression as well as the amplitude of transcript levels. Within the LP cluster, there are substantial changes in gene profiles from the wild-type to the NILs (Figure 3.3c). In the *Ppd-D1a* line, the peak of expression for these genes shifts from the LP stage to DR, with the average expression remaining high at the LP and TS stages. In the *ppd-1* null line, the peak at LP is maintained, but instead of the decrease in expression from LP to TS that occurs in wild-type, transcript levels are maintained at a higher level in the absence of *Ppd-1*. The TS cluster was the largest representative cluster selected, consisting of 5176 genes (Figure 3.3d). Even with this number of genes examined, *Ppd-D1a* allelism still has a dramatic effect on gene expression. In the *Ppd-D1a* line, we see a similar profile relative to wild-type; however, there is a general dampening effect on gene expression. The effect is particularly strong at the TS stage, with up-regulation still occurring from the LP stage, but to a much lesser degree that that detected in the wild-type. In the *ppd-1* null line, the genes display a similar profile to wild-type at the VG, DR and LP stages, with a greater average down-regulation at the DR stage. Conversely, during the LP to TS transition, I detect down-regulation between in the *ppd-1* null line compared to considerable up-regulation in the wild-type.

Taken together, these data show that genetic variation for *Ppd-1* disrupts the pattern and amplitude of gene expression profiles, relative to wild-type. The disruption is generally focused to the DR and LP stages, suggesting that *Ppd-1* has its greatest impact on meristem gene expression during early developmental stages when spikelet development occurs.
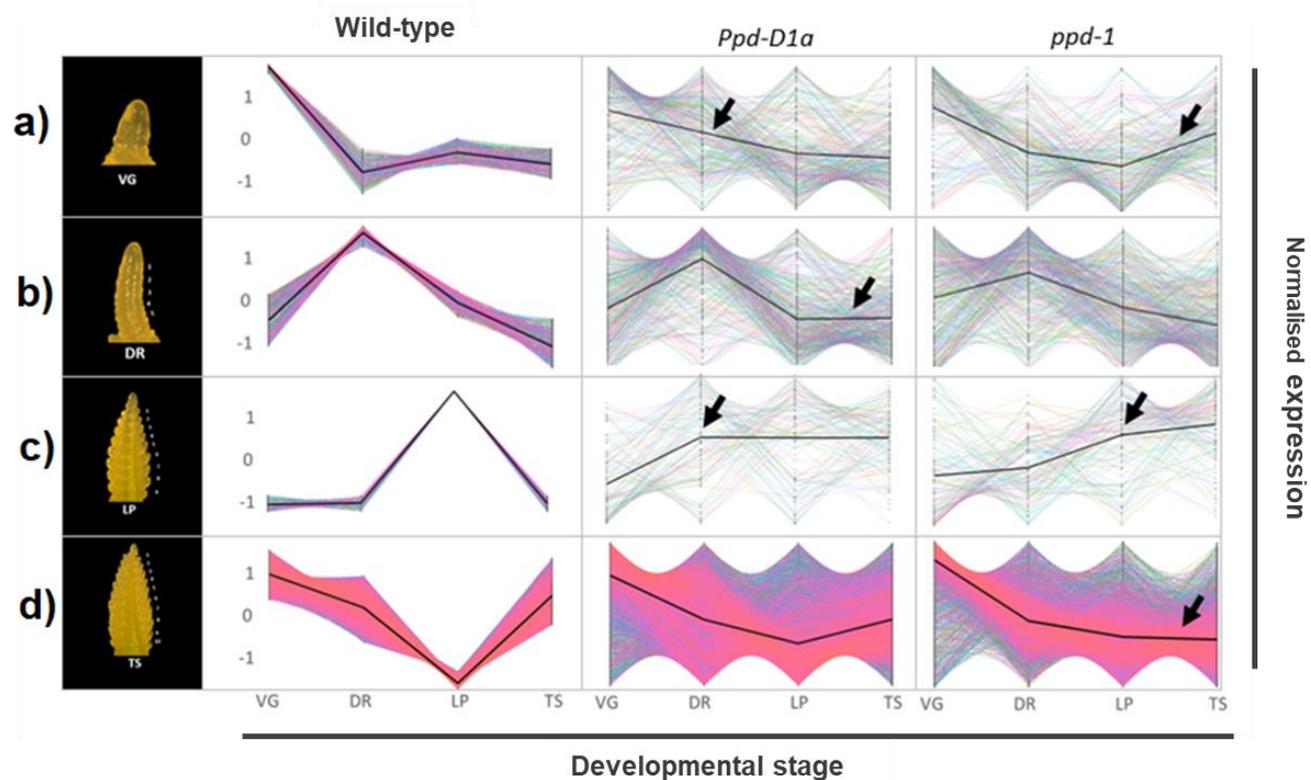
**Figure 3.3 - Clustering analysis of the inflorescence transcriptome at key developmental stages**

Normalised gene expression cluster profiles selected as representative of the key meristem development stages, **a)** vegetative (VG), **b)** double ridge (DR), **c)** lemma primordium (LP) and **d)** terminal spikelet (TS) in the wild-type plants. Corresponding expression profiles of these genes are shown for *Ppd-D1a* and *ppd-1* NILs. The coloured gene expression profiles consist of normalised gene expression, a line of best fit is shown in black. Arrows show notable points of differentiation in the NILs relative to the wild-type.

## Table 3.1 - Summary of GO terms enriched for stage-specific clusters

Table of enriched go terms for the stage-specific genes identified for each of the VG, DR, LP and TS stages. Molecular GO term is shown, with GOID alongside the number of genes represented.

| VG | Number of genes | DR | Number of genes |
|---|---|---|---|
| transferase activity (GO:0016740) | 54 | binding (GO:0005488) | 138 |
| catalytic activity, acting on a protein (GO:0140096) | 42 | heterocyclic compound binding (GO:1901363) | 93 |
| protein serine/threonine kinase activity (GO:0004674) | 16 | organic cyclic compound binding (GO:0097159) | 93 |
| structural molecule activity (GO:0005198) | 10 | catalytic activity (GO:0003824) | 76 |
| structural constituent of ribosome (GO:0003735) | 9 | nucleic acid binding (GO:0003676) | 63 |
| ubiquitin-protein transferase activity (GO:0004842) | 9 | DNA binding (GO:0003677) | 43 |
| ubiquitin-like protein transferase activity (GO:0019787) | 9 | protein binding (GO:0005515) | 34 |
| lipid binding (GO:0008289) | 6 | transferase activity (GO:0016740) | 27 |
| calmodulin binding (GO:0005516) | 4 | RNA binding (GO:0003723) | 21 |
| calcium-dependent protein kinase activity (GO:0010857) | 3 | transferase activity, transferring phosphorus-containing groups (GO:0016772) | 13 |
| calmodulin-dependent protein kinase activity (GO:0004683) | 3 | enzyme regulator activity (GO:0030234) | 12 |
| calcium-dependent protein serine/threonine kinase activity (GO:0009931) | 3 | molecular function regulator (GO:0098772) | 12 |
| chaperone binding (GO:0051087) | 3 | oxidoreductase activity (GO:0016491) | 11 |
| copper ion binding (GO:0005507) | 3 | protein kinase activity (GO:0004672) | 9 |
| organic anion transmembrane transporter activity (GO:0008514) | 3 | cysteine-type peptidase activity (GO:0008234) | 8 |
| rRNA binding (GO:0019843) | 3 | enzyme binding (GO:0019899) | 8 |
| uridylate kinase activity (GO:0009041) | 2 | enzyme activator activity (GO:0008047) | 7 |
| ATPase binding (GO:0051117) | 2 | transporter activity (GO:0005215) | 6 |
| acetylglucosaminyltransferase activity (GO:0008375) | 2 | GTPase activator activity (GO:0005096) | 5 |
| nucleoside monophosphate kinase activity (GO:0050145) | 2 | GTPase regulator activity (GO:0030695) | 5 |

| LP | Number of genes | TS | Number of genes |
|---|---|---|---|
| protein heterodimerization activity (GO:0046982) | 4 | molecular_function (GO:0003674) | 209 |
| carboxylic ester hydrolase activity (GO:0052689) | 3 | binding (GO:0005488) | 158 |
| glutathione binding (GO:0043295) | 2 | heterocyclic compound binding (GO:1901363) | 104 |
| oligopeptide binding (GO:1900750) | 2 | organic cyclic compound binding (GO:0097159) | 104 |
| modified amino acid binding (GO:0072341) | 2 | ion binding (GO:0043167) | 78 |
| ubiquitin conjugating enzyme activity (GO:0061631) | 2 | nucleic acid binding (GO:0003676) | 65 |
| ubiquitin-like protein conjugating enzyme activity (GO:0061650) | 2 | small molecule binding (GO:0036094) | 45 |
| sulfur compound binding (GO:1901681) | 2 | hydrolase activity (GO:0016787) | 44 |
| thiol-dependent ubiquitin-specific protease activity (GO:0004843) | 2 | cation binding (GO:0043169) | 40 |
| thiol-dependent ubiquitinyl hydrolase activity (GO:0036459) | 2 | metal ion binding (GO:0046872) | 39 |
| hydro-lyase activity (GO:0016836) | 2 | DNA binding (GO:0003677) | 37 |
| ubiquitinyl hydrolase activity (GO:0101005) | 2 | protein binding (GO:0005515) | 30 |
| omega peptidase activity (GO:0008242) | 2 | RNA binding (GO:0003723) | 27 |
| peptide binding (GO:0042277) | 2 | zinc ion binding (GO:0008270) | 16 |
| ubiquitin-like protein-specific protease activity (GO:0019783) | 2 | nucleoside-triphosphatase activity (GO:0017111) | 12 |
| amide binding (GO:0033218) | 2 | mRNA binding (GO:0003729) | 11 |
| glutathione transferase activity (GO:0004364) | 2 | phosphoric ester hydrolase activity (GO:0042578) | 10 |
| ribokinase activity (GO:0004747) | 1 | phosphatase activity (GO:0016791) | 9 |
| phytanoyl-CoA dioxygenase activity (GO:0048244) | 1 | cytoskeletal protein binding (GO:0008092) | 8 |
| superoxide dismutase copper chaperone activity (GO:0016532) | 1 | nucleotidyltransferase activity (GO:0016779) | 7 |

### 3.4.1.4  Unpicking the LP cluster

To understand the impact of *Ppd-1* allelism on stage specific clusters, I hypothesised that the disrupted expression patterns in the *Ppd-D1a* and *ppd-1* lines would consist of several different clustered patterns, reflecting the effects of the different genotypes. Therefore, I unpicked the constituting patterns underpinning the LP cluster (Figure 3.3c). The LP cluster was selected because my analysis suggested the LP stage is a focal point for *Ppd-1* influence on inflorescence gene expression. Several well-defined clusters make up the tangled *Ppd-D1a* LP gene expression patterns (Figure 3.4a-d). Genes that are expressed highly at the VG to DR transition and remain high constitute the majority of represented genes (Figure 2.4a-b). A second important set includes genes that peak at the LP stage, but also show higher expression at the DR stage (Figure 3.4c). Of particular note are genes that show a shift of the complete expression pattern (Figure 3.4d), which peak earlier at the DR stage and subsequent down-regulated transitioning to LP. Similarly, several different clusters constituting the major patterns were identified for the *ppd-1* line (Figure 3.4e-h). The first identified cluster is similar to the wild-type cluster, with a peak at the LP stage (Figure 3.4e); however, higher comparative expression is detected at the VG stage, with reduced down-regulation from the LP to TS stage. The second *ppd-1* subcluster shows a gradual increase in expression as the inflorescence progresses past the DR stage, with little or no down-regulation at the LP to TS transition (Figure 3.4f). The third subcluster shows high expression after the VG stage that remains high throughout this phase of development (Figure 3.4g). The fourth subcluster shows a dramatic down-regulation from the VG to DR stage transition that remains low as the stages progress (Figure 3.4h). Based on this subcluster analysis, I conclude that the same genes are affected by *Ppd-1* allelism in different ways. There is a trend that a photoperiod-insensitive allele of *Ppd-1* mis-regulate expression during the VG to DR stage transition, while *ppd-1* null alleles tend to mis-regulate gene activity during the LP to TS transition.
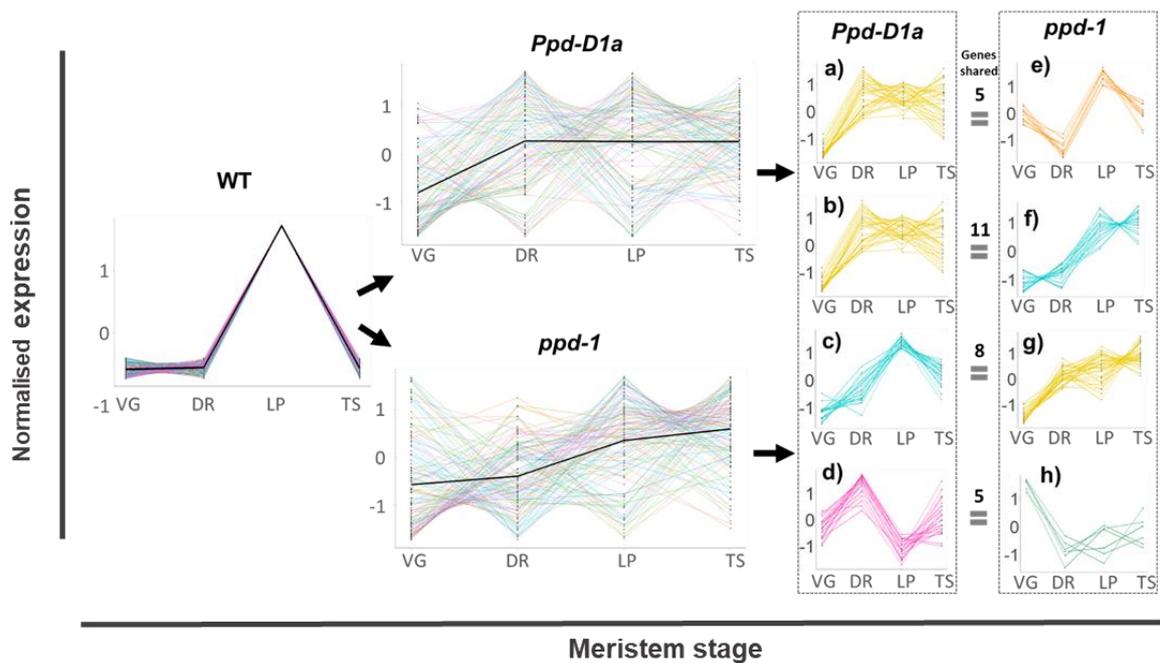
**Figure 3.4 - Subclustering analysis for the LP representative cluster.**

The breakdown of the gene expression profiles selected as representative of the lemma primoridium (LP) meristem development stage, in the wild-type, *Ppd-D1a* and *ppd-1* lines over the vegetative (VG), double ridge (DR), lemma primordium (LP) and terminal spikelet (TS) stages. The coloured gene expression profiles consist of normalised gene expression, a line of best fit is shown in black and highlights the trends of gene expression. **a-d)** Represent constituting subclusters of *Ppd-D1a.* **e-h)** Represent constituting subclusters of *ppd-1*.

### 3.4.1.5  Identification of LP cluster genes affected by *Ppd-D1a*

In the previous analysis (Figure 3.4), a subcluster of genes that show an earlier gene expression profile relative to wild-type were identified (Figure 3.4d). I hypothesised that the expression of genes represented in this subcluster are under the influence of *Ppd-1*. To investigate the genes further, I plotted their non-normalised expression patterns (Figure 3.5; Table 3.2). All genes experience a shift in their peak of expression from the LP to DR stage. However, some genes show significant down-regulation, relative to wild-type. This can be observed in TraesCS3A02G068100.1 (tRNA N6-adenosine threonylcarbamoyltransferase), TraesCS5A02G385200.1 (Sigma non-opioid intracellular receptor 1), TraesCS5A02G525900.1 (COP9 signalosome complex subunit 1) and TraesCS7A02G213500.1 (TMV-MP30 binding protein 2C). Of particular interest are the genes extracted from this analysis that have a striking expression in the *ppd-1* line. For example, TraesCS5D02G207500.1 (U-box domain-containing protein) and TraesCSU02G092100.1 (Histone H3), both display increasing expression as the inflorescence develops in the *ppd-1* null line, relative to the stage specific profiles in wild-type and *Ppd-D1a* NILs. None of the above-mentioned genes have described roles during inflorescence development, and yet, here we observe expression patterns changing on a stage- and genotype-specific basis. These genes may prove to be promising candidates for future study.
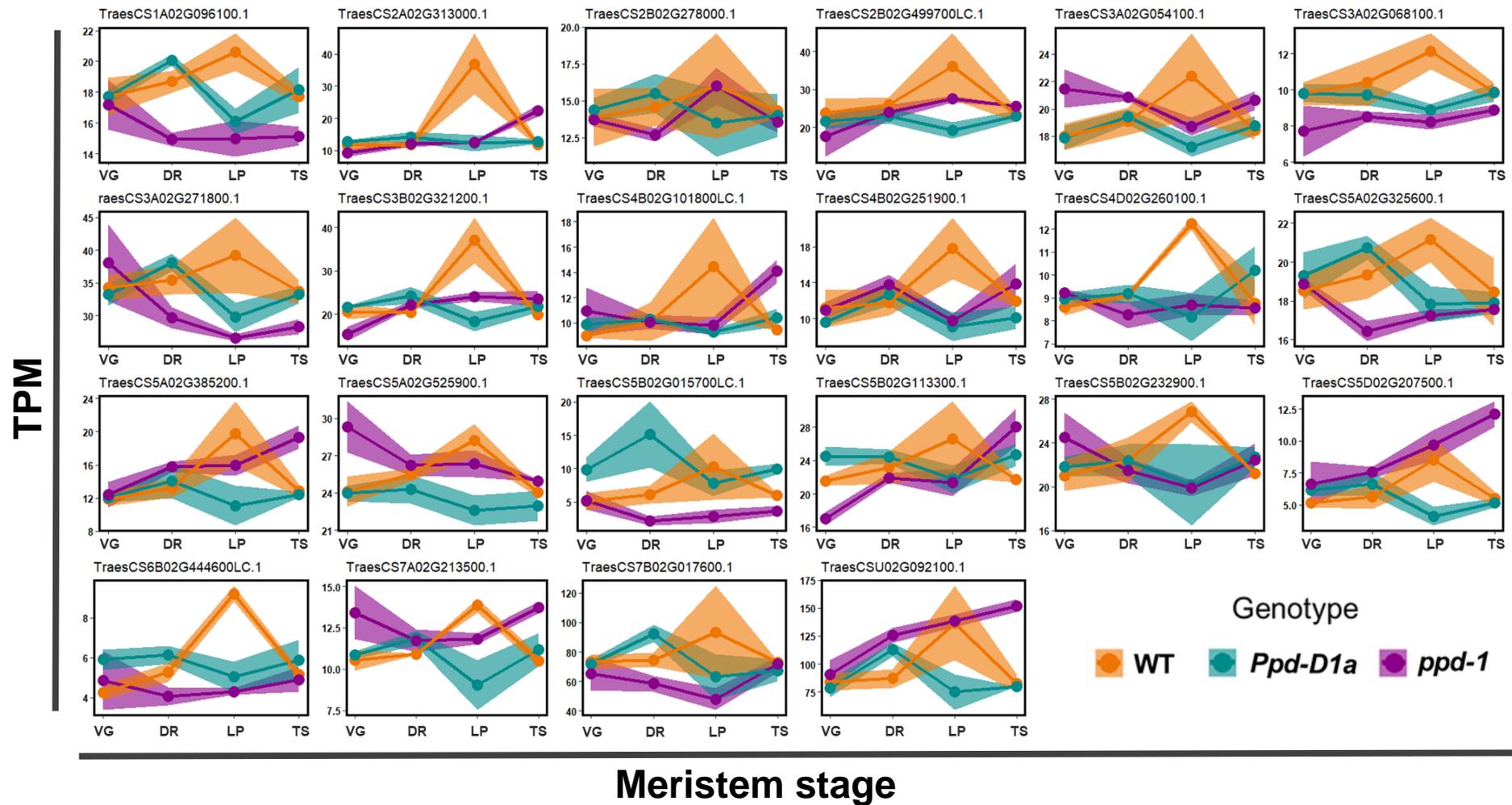
**Figure 3.5 - The TPM expression profiles of a *Ppd-D1a* subcluster**

Expression of genes within the *Ppd-D1a* subcluster Figure 3.4d in wild-type (orange), *Ppd-D1a* photoperiod-insensitive (cyan) and null *ppd-1* (magenta) NILs under field conditions. Expression is plotted against the meristem stages: vegetative (VG), double ridge (DR), lemma primordium (LP) and terminal spikelet (TS). Graphs are presented as ribbon plots which show the transcripts per million (TPM) (solid line with data points) ± SEM (shaded region) of three biological

**Table 3.2 - Functional annotation of *Ppd-D1a* subcluster genes**

The functional annotation for each of the Gene.ID represented in Figure 3.5, from the Figure 3.4d subcluster.

|  | Gene.ID | Functional Annotation |
|---|---|---|
| 1 | TraesCS1A02G096100.1 | Mediator of RNA polymerase II transcription subunit 28 |
| 2 | TraesCS2A02G313000.1 | RmlC-like cupins superfamily protein |
| 3 | TraesCS2B02G278000.1 | helicase with zinc finger protein |
| 4 | TraesCS2B02G499700LC.1 | COP9 signalosome complex subunit 8 |
| 5 | TraesCS3A02G054100.1 | Katanin p60 ATPase-containing subunit, putative |
| 6 | TraesCS3A02G068100.1 | tRNA N6-adenosine threonylcarbamoyltransferase |
| 7 | TraesCS3A02G271800.1 | Ubiquitin-conjugating enzyme E2 |
| 8 | TraesCS3B02G321200.1 | Interferon-induced GTP-binding protein Mx2 |
| 9 | TraesCS4B02G101800LC.1 | Cellulose synthase |
| 10 | TraesCS4B02G251900.1 | DUF538 family protein (Protein of unknown function, DUF538) |
| 11 | TraesCS4D02G260100.1 | ADP-ribosylation factor, putative |
| 12 | TraesCS5A02G325600.1 | Protein XAP5 CIRCADIAN TIMEKEEPER |
| 13 | TraesCS5A02G385200.1 | Sigma non-opioid intracellular receptor 1 |
| 14 | TraesCS5A02G525900.1 | COP9 signalosome complex subunit 1 |
| 15 | TraesCS5B02G015700LC.1 | Xanthine/uracil permease family protein |
| 16 | TraesCS5B02G113300.1 | electron transporter, putative (Protein of unknown function) |
| 17 | TraesCS5B02G232900.1 | Protein disulfide isomerase |
| 18 | TraesCS5D02G207500.1 | U-box domain-containing protein |
| 19 | TraesCS6B02G444600LC.1 | Ribosome maturation protein SBDS |
| 20 | TraesCS7A02G213500.1 | TMV-MP30 binding protein 2C |
| 21 | TraesCS7B02G017600.1 | Histone H3 |
| 22 | TraesCSU02G092100.1 | Histone H3 |

### 3.4.1.6 Global expression patterns for WT, *Ppd-D1a* and *ppd-1* and their interactions

My previous analysis demonstrates considerable disruption of gene expression profiles between the wild-type and the two *Ppd-1* NILs at key meristem stages (Figure 3.1-5). Based on these findings, I hypothesised that genetic variation for *Ppd-1* influences the global inflorescence transcriptome and that certain developmental stages may emerge as focal points for regulation. To investigate this, clustering analysis was carried out on the top 30,000 genes expressed in the developing inflorescence for wild-type, *Ppd-D1a* and *ppd-1* NILs. The clusters of each genotype were then aligned and graphed into an alluvial diagram to show how genes are represented between the clusters (Figure 3.6).

An immediate outcome of this analysis is that many genes fall into the non-clustered category (NC). Of the 30,000 genes examined in wild-type, 14535 genes were clustered with the remainder of these genes falling into the NC category. In the *Ppd-D1a* line, 19459 genes were clustered, and in the *ppd-1* line, 17474 genes were clustered. This difference is likely because either the expression profile has become flat between the stages and therefore don't fall into a cluster, or that the genes in wild-type that were not in the top 30000 most expressed have increased in expression in the *Ppd-D1a* and *ppd-1* lines. In general, these results show that genetic variation for *Ppd-1* influences the expression of multiple gene families during inflorescence development, beyond the set of flowering-related transcription factors that have previously been shown to alter transcript abundance in photoperiod-insensitive lines (Boden *et al.*, 2015).

Based on the identified clusters, I explored the genes that fall within the most-populated groups. In the wild-type, the largest clusters are C0 and C8, which account for 58.17% of clustered gene expression. Both clusters crescendo at the LP stage showing a mirror image of gene expression patterns. The C0 cluster consists of 5176 genes that show a gradual down-regulation from the VG stage until LP, and a subsequent uptrend in expression to the TS stage. The C8 cluster consists of 3278 genes that show gradual up-regulation after the VG stage, and a subsequent downturn in expression after the LP stage. Both these profiles suggest the LP stage is an important gene expression checkpoint for inflorescence development, with the positive and negative regulation of a total of 8454 genes converging at this stage. The second and third largest clusters are C1 (1219 genes) and C11 (760 genes) respectively; interestingly, these both represent apparent antagonistic patterns of expression with up- or down-regulation after the VG stage, with the C1 cluster showing down-regulation at the LP stage that remains low, and the C11 cluster showing up-regulation at the DR stage that then remains high. This trend of equal amounts of genes up- and down-regulated is consistent across all clusters, with a total of 7611 genes represented in down-regulated clusters and 6789 genes in up-

regulated clusters. This data demonstrates that inflorescence development is an intricately controlled process, mediated by a balancing of up and down-regulated genes.

The largest cluster in wild-type, C0 shows little maintenance in the *Ppd-D1a* line (5.89%). Instead, the genes represented in the wild-type C0 lines are largely present in the C1 (20.48%) and C16 (26.37%) clusters of the *Ppd-D1a* NIL. The C1 cluster of *Ppd-D1a* shows a similar expression profile to the C0 cluster of wild-type; however, the uptrend in expression between the LP and TS stage is absent. The C16 cluster of *Ppd-D1a* also shows significant down-regulation after the VG stage; however, this downturn in expression occurs earlier at the DR stage as opposed to the LP stage in both the C0 and C1 clusters of wild-type. The exact same trend is replicated in the wild-type C8 cluster to *Ppd-D1a* C10 (17.27%) and C11 (18.97%) clusters only mirror imaged. This trend illustrates both a shifting of expression profiles to earlier stages (C16 and C11) and a loss of regulation at the LP to TS transition (C1 and C10). In multiple instances, gene expression profiles shift to earlier between the wild-type and photoperiod-insensitive line, while maintaining an overall consistent pattern of expression. For example, from the wild-type C0 cluster, 276 genes are represented in the C17 cluster of *Ppd-D1a*, in which the down-regulation of expression is shifted from the LP stage to the DR stage.

In the *ppd-1* null line, the two largest clusters are C2 and C11 consisting of 4903 genes (28.06%) and 4458 genes (25.51%) of total genes clustered, respectively. Comparing the *ppd-1* clustered expression to wild-type, many genes are shared between the C0 wild-type cluster and the C2 cluster of *ppd-1* (42.75%). In wild-type, there is a gradual drop in gene expression between the VG and LP stages, with a significant uptrend in transcripts between the LP and TS stages. However, the *ppd-1* expression in C2 drops immediately from the VG to DR stage, and low expression is maintained until TS. These results indicate that *Ppd-1* is required for high expression of these genes after VG. Similarly, the genes represented in the wild-type C8 cluster are largely present in the *ppd-1* C11 cluster (41.06%). The C11 cluster shows an up-regulation after the VG stage and subsequent plateau of gene expression from the DR stage onwards, indicating that any repressive activity of *Ppd-1* at the TS stage is lost for these genes. The third largest cluster in the *ppd-1* line is C16, consisting of 1590 genes. The largest wild-type cluster represented in C16 is C0, sharing 503 genes – these genes are down-regulated at the DR stage, as opposed to the LP stage, but expression tends to increase as inflorescence development proceeds.

Together, these data paint a picture for a highly influential role of *Ppd-1* over gene expression in the developing inflorescence in the morning either directly, or through regulation of genes such as *FT1* and *FT2*. Trends from wild-type inflorescence development highlight the LP stage as a focal point for gene regulation. The insensitive

*Ppd-D1a* line both mis-regulates gene expression at the LP to TS stage and shifts gene expression profiles to earlier in development. The *ppd-1* null line shows a dramatic loss of gene regulation after the VG stage, affecting both up and down-regulated genes.
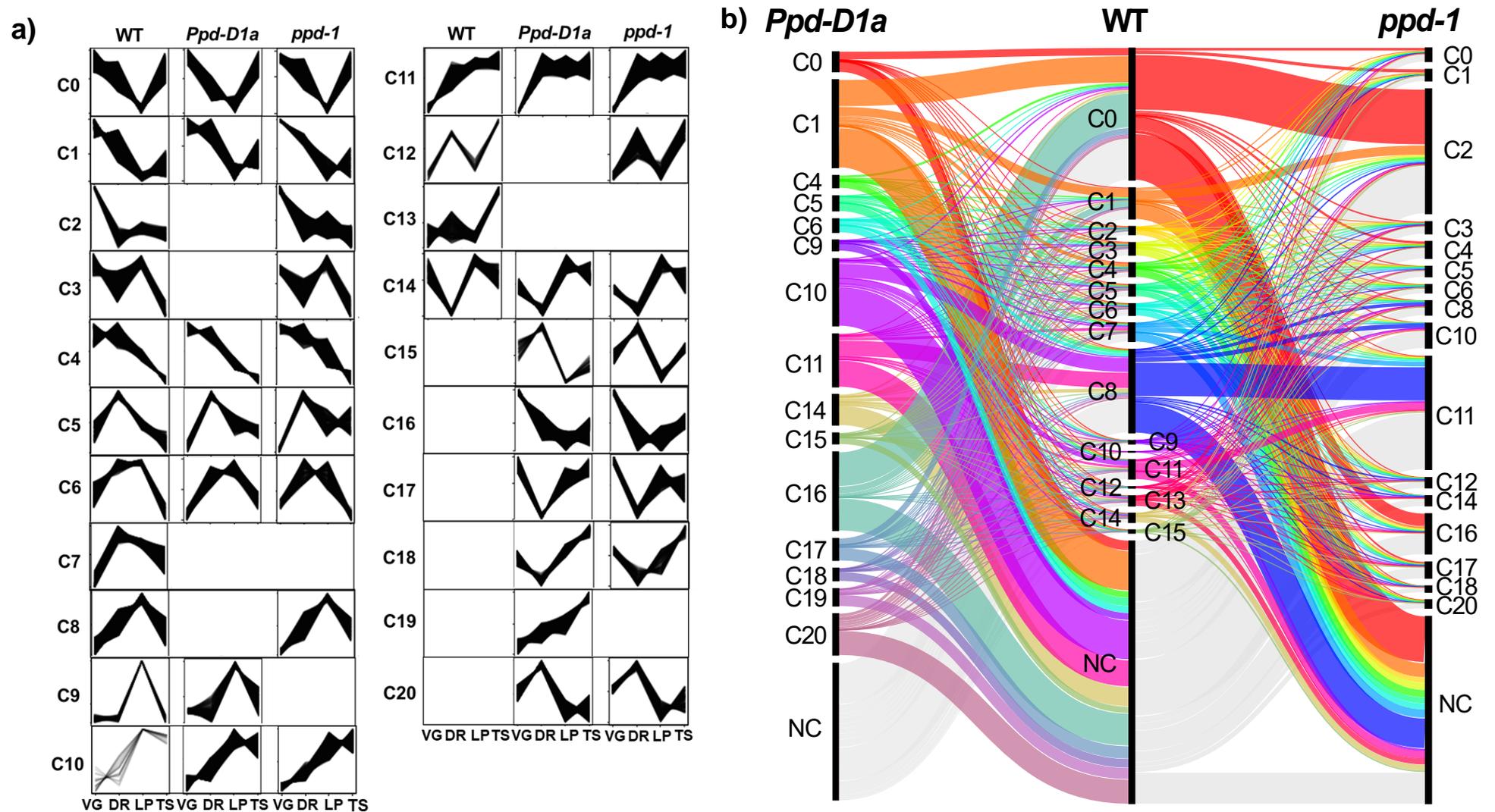
**Figure 3.6 - Clustering analysis of the top 30,000 expressed genes in the inflorescence**

**a)** Shows the clustered gene expression profiles of the top 30,000 genes represented in the WT, *Ppd-D1a* and *ppd-1* NILs over the meristem stages: vegetative (VG), double ridge (DR), lemma primordium (LP) and terminal spikelet (TS). Clusters are denoted by the numbering C0 to C20. Minimum cluster size of 22 gene. Where a cluster profile is not represented in a genotype it is represented by a blank box. **b)** Shows the genes shared between the clusters of *Ppd-D1a* and *ppd-1* relative to WT plotted using an alluvial diagram. The thickness of the connection represents number of genes. NC= No cluster shared between genotypes.

## 3.4.2 *Ppd-1* mediated regulation of auxin-related genes

### 3.4.2.1 Expression patterns of auxin-related genes in the developing inflorescence

Having determined that *Ppd-1* influences the global inflorescence transcriptome (Figure 3.6), I hypothesised that individual gene families that coordinate important aspects of spikelet and floret development may respond to *Ppd-1* allelism. Auxin and auxin related genes have a well-documented role in plant development, including the formation of lateral organs such as spikelets in maize (Gallavotti *et al.*, 2008; Zhao, 2010; Xing *et al.*, 2011; Li *et al.*, 2018). In addition, treatment with the auxin analogue 2,4-D profoundly alters inflorescence architecture in wheat, promoting the formation of inflorescence branches (Sharman, 1978). However, the role and regulation of auxin and auxin regulated genes in wheat is poorly understood. To determine the transcriptome landscape of auxin-related processes in wheat, I performed cluster analysis using all genes with a functional annotation related to auxin (1111 genes), in each of the three genotypes (Figure 3.7). The clusters were aligned and compared using an alluvial diagram. Notably, many genes fall into the no cluster (NC) category, indicating that either the expression of these genes flatlined during these developmental stages, their expression was below the 1 TPM threshold, or they did not fit into a cluster with the minimum size of 11 genes. In the comparison between wild-type and *Ppd-D1a* clusters, 120 genes fell into the NC category. Of these, the TPM of 44 genes were below 1 TPM in wild-type, the expression profile of 22 genes flatlined, and 54 did not fall into a cluster with 11 or more genes. Between the *ppd-1* line and wild-type, 114 genes fell into NC, of which 54 were expressed too lowly in wild-type, 20 displayed a flat expression profile, and 50 did not fall into a cluster with 11 or more genes.

In contrast to the global expression analysis (Figure 3.6), there is much greater conservation of gene expression profiles among the genotypes for the auxin-related gene subset (Figure 3.7). For example, the C5 cluster is largely maintained between the wild-type and *ppd-1* line. Likewise, the C2 cluster between the wild-type and the *Ppd-D1a* line is conserved. However, there is still significant disruption caused by the *Ppd-1* NILs on clustered gene expression, though to a lesser extent than the global analysis (Figure 3.6).

In wild-type, 250 genes were clustered into 9 defined groups, with the largest clusters (C2 and C5) representing 18.4% and 16.4% of total clustered genes. These clusters represent genes that are significantly up- or down-regulated following the VG to DR transition and remain at that level. This likely reflects many genes that are involved in activating or repressing gene expression at every stage post vegetative to floral transition. In addition, I identified examples of genes whose expression coincides with different developmental stages. For example, the DR stage is a focal point for high expression in the C0 (11.2%)

and C4 (5.2%) clusters, and a point for suppressed expression of the C9 (4.4%) cluster. Similarly, at the LP stage, the C7 (8%) and C9 (4.4%) clusters peak and the C3 (7.2%) and C4 (6.4%) clusters are repressed. At the TS stage, the C8 (6%) cluster peaks and no clusters with strongly suppressed expression were detected. These data show that the DR and LP are both focal points for altering the regulation of auxin-related genes. Overall, the DR stage showed a prevalence of genes peaking (16.4%) as opposed to dipping (4.4%). This result indicates that auxin-related genes have a positive role in activating the processes at DR, including spikelet meristem formation, consistent with studies from rice and maize (Gallavotti *et al.*, 2008; Yang *et al.*, 2017). Conversely, the LP stage showed a slight preference for down-regulation in expression (13.6% of genes), compared to 12.4% transcripts that peak at this stage, indicating a balance of auxin regulated responses at this stage. Based on this analysis, the DR and LP stages are two major focal points for regulation of auxin-related genes. These stages are critical for the formation of floral structures, as spikelet meristems initiated at the DR stage and floral structures such as the lemma and palea differentiate at the LP stage (Kirby & Appleyard, 1984). The profiles represented by the auxin-related genes suggest a significant involvement of this gene family in wheat inflorescence development.

In the *Ppd-D1a* NIL, 262 auxin-related genes were determined to cluster. Compared to wild-type, similar expression profiles are represented; however, there are major differences, as the largest clusters represented are the C1 (17.9%) and C7 (12.21%) clusters, compared to C2 and C5 in the wild-type. C1 is highly expressed at the VG and DR stages, but subsequently experiences a significant drop in expression at the LP stage, remaining at the same level until the TS stage. Conversely, C7 has the exact opposite expression pattern, showing low transcript levels at the VG and DR stages and high at the LP and TS stages. Unlike wild-type, there is no C5 cluster, with the genes of this group from wild-type being split between several other clusters and the NC category. The C5 cluster in wild-type is characterised by significant up-regulation from the VG to DR stage, before expression plateaus for LP and TS. The movement of many of these genes from C5 in wild-type to NC in *Ppd-D1a* is most likely due to consistently high expression at all stages, including VG. The C0 cluster is represented in both the wild-type and *Ppd-D1a* NILs, showing conservation of gene profiles between genotypes.

In the *ppd-1* null line, 275 auxin-related genes were clustered, with the largest clusters represented being the C2 (24%) and C5 (25.9%) clusters. These represent half of the clustered auxin-related genes. They follow expression patterns that are either significantly down-regulated after VG, such as those in C2 and those with expression significantly higher expression after VG, including those in C5. In wild-type, both of these major clusters are compiled of many different clusters. This relationship points to regulation of these genes by *Ppd-1* whereby knocking out its function removes either its positive or

negative regulation. This trend is very similar to that seen in the global clustering analysis, indicating that the major effect of *Ppd-1* null alleles on a global scale is present in the auxin-specific analysis.

Together, these data point to *Ppd-1*-mediated stage-specific regulation of auxin-related genes in the developing inflorescence. The DR and LP stages are focal points for regulation, likely correlating with the formation of important yield traits such as spikelet initiation/development.
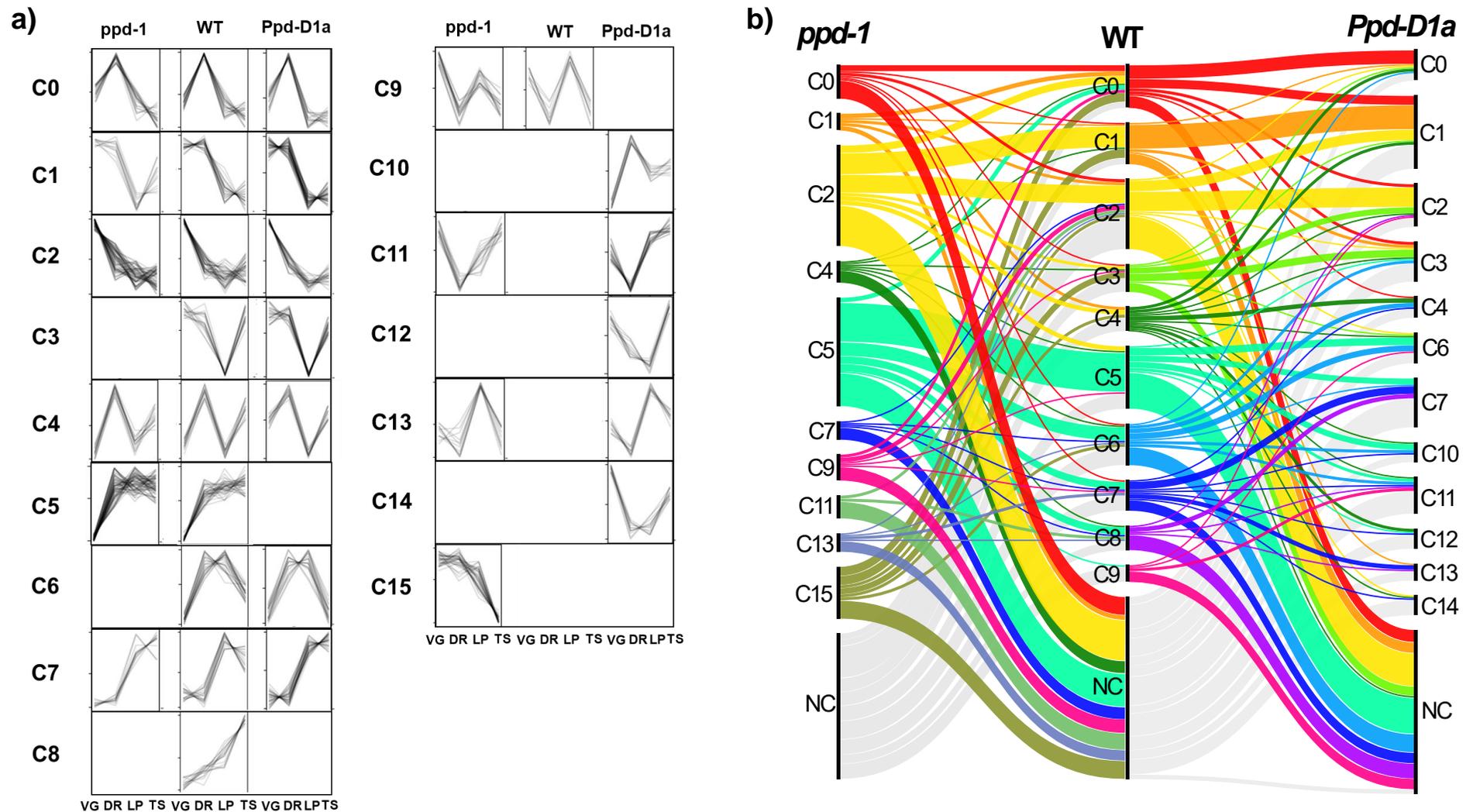
**Figure 3.7 - Clustering analysis of all auxin-related genes expressed in the wheat inflorescence**

a) Shows the clustered gene expression profiles of the 1111 auxin -related genes, identified through functional annotation. Genes are represented in the wild-type, *Ppd-D1a* and *ppd-1* NILs over the meristem stages: vegetative (VG), double ridge (DR), lemma primordium (LP) and terminal spikelet (TS). Clusters are denoted by the numbering C0 to C15. Minimum cluster size of 11 gene. Where a cluster profile is not represented in a genotype it is represented by a blank box. b) Shows the genes shared between the clusters of *Ppd-D1a* and *ppd-1* relative to WT plotted using an alluvial diagram. The thickness of the connection represents number of genes. NC= No cluster shared between genotypes.

### 3.4.2.2  Phylogenetic analysis of the *AUXIN RESPONSE FACTORS* in wheat

Auxin signalling plays a cardinal role in many aspects of plant development. Auxin-response genes are responsible for providing precise control over auxin regulated processes (Abel & Theologis, 1996; Guilfoyle *et al.*, 1998). These genes are regulated by conserved promoter elements including AuxRE (auxin response element, TGTCTC) (Xing *et al.*, 2011). These auxin-response genes are subsequently under the regulation of *AUXIN RESPONSE FACTORS* (*ARF*), characterised by an N-terminal DNA-binding domain (DBD), with either an activation (AD) or repression (RD) domain and a C terminal carboxyl-terminal dimerization domain (CTD). This CTD domain mediates both the homo- and hetero-dimerization between ARF proteins in addition to Aux/IAA proteins, to facilitate their action (Guilfoyle & Hagen, 2007; Kim *et* al., 1997; Piya *et al.*, 2014; Li *et al.*, 2016).

In rice and *Arabidopsis* there are 23 and 25 *ARF* genes, respectively (Wang *et al.*, 2007; Rademacher *et al.*, 2011). Of all plants, maize has the most identified *ARF* genes, with phylogenetic analysis uncovering 37 genes (Xing *et al.*, 2011; Liu *et al.*, 2011). To investigate these genes in wheat, I used BLAST analysis to detect the ortholog of each maize gene. A subsequent phylogenetic analysis was performed on the D genome copy of each wheat gene identified (Figure 8a). One genome copy was used for simplicity, as there is strong homology between genome copies.

In maize (*Zea mays*), the auxin response factors fall into six classes: Class I (*AtARF3/4*-like), II (*AtARF10/16/17*-like), III (*AtARF1/2*-like), IV(*AtARF5*-like), V(*AtARF6/8*-like) and VI(*AtARF7/19*-like) (Figure 8b) (Xing *et al.*, 2011). With many of the genes falling into related sister pairs *ZmARF2* and *ZmARF17*, *ZmARF3* and *ZmARF30*, *ZmARF4* and *ZmARF29*, *ZmARF5* and *ZmARF31*, *ZmARF6* and *ZmARF14*, *ZmARF8* and *ZmARF15*, *ZmARF10* and *ZmARF25*, *ZmARF12* and *ZmARF24*, *ZmARF13* and *ZmARF28*, *ZmARF19* and *ZmARF21*, or triplets, *ZmARF11*, *ZmARF23* and *ZmARF26*, *ZmARF1*, *ZmARF20* and *ZmARF27* and quadruplets in the case of *ZmARF9*, *ZmARF16*, *ZmARF18* and *ZmARF22* (Xing *et al.*, 2011). The genes in the same pairs, triplets or quadruplets have the same number of exons and introns, indicating they are the result of duplicated genomic regions in maize.

I have annotated the wheat *ARF* genes by the nomenclature of orthologues characterised in maize (Figure 3.8a; Xing *et al.*, 2011). Several genes have at least one orthologue in wheat. These include *ZmARF4* and *ZmARF29* here named *TaARF4*. *ZmARF3* and *ZmARF30* here named *TaARF3*. *ZmARF9* and *ZmARF34* here named *TaARF9*. *ZmARF18* and *ZmARF22* named here as *TaARF18*. *ZmARF11* and *ZmARF26* named here as *TaARF11*. *ZmARF6* and *ZmARF38* named here as *TaARF6*. *ZmARF2* and *ZmARF17* named here as *TaARF2*. *ZmARF19* and *ZmARF21* named here as *TaARF19*.

Several genes don't have orthologues in wheat, these include *ZmARF36*, *ZmARF14*, *ZmARF13*, *ZmARF8*, *ZmARF28*, *ZmARF32*, *ZmARF33*, *ZmARF10* and *ZmARF31*.

In maize, 13 ARFs are expressed in the developing inflorescence (Galli *et al.*, 2015; Figure 8a). These genes exclusively fall into classes V, IV and VI, indicating that these classes are particularly important during inflorescence development. As such, we hypothesised these sub-families would also have a role in wheat. The sub-localisation of expression for these genes has been identified in the developing floral meristem of maize. *ARF1* and *ARF35* showed broad expression throughout the inflorescence; *ARF4*, *ARF18*, *ARF20*, *ARF22*, *ARF29*, and *ARF34* showed strong expression at the peripheral zone of the inflorescence meristem; and *ARF3*, *ARF27*, and *ARF30* showed narrow expression in developing primordia. Expression patterns of the different *ARFs* varied in developing axillary meristems; the majority were predominantly restricted to the core (*ARF1, ARF4, ARF9*, *ARF16*, *ARF20*, *ARF22*, *ARF29*, *ARF34*, *ARF35*), and others such as *ARF3* and *ARF30* appeared to be localised to the base and boundaries of the meristems, whereas *ARF18* and *ARF22* localized to the glume primordia and bracts (Gallivoti *et al.*, 2015). These genes fall into the *AtARF5*-like, *AtARF6/8*-like and *AtARF7/19*-like categories.

Interestingly, several wheat homologs of the maize genes are expressed highly during inflorescence development (>10 TPM) (Figure 3.8a), such as *TaARF35* (class VI), *TaARF4* (class IV) and *TaARF3*, *TaARF9* and *TaARF18* of class V. However, in wheat, unlike maize, genes from other *ARF* classes are expressed highly (Gallivotti *et al.*, 2015). *TaARF25* from class III (*ARF1/2*-like) shows the highest expression of all *ARF* genes examined (>50 TPM), and *TaARF11*, *TaARF23* and *TaARF24* from class I are expressed strongly. Several genes in class II are also expressed robustly. This analysis indicates that whilst the classes of genes appear to be conserved through species, their expression patterns and resulting role may vary.

**Figure 3.8- A phylogenetic analysis of the *ARF* genes in wheat**

**a)** A phylogenetic analysis of the *ARF* genes in maize and wheat. Colour of branch denotes class of gene; class I = light blue, class II = green, class III = purple, class IV = dark blue class V = red, class VI = yellow. Colour next to gene ID denotes expression of the gene; green= expressed in maize, dark blue= expressed over 50 TPM in wheat IM, yellow= expressed over 10 TPM in wheat IM, light blue = expressed over 1 TPM in wheat IM. Amino acids were compared, obtained through BLAST. Unrooted tree generated using MEGAX, using the maximum likelihood method and Dayhoff matrix based model. **b)** Figure taken from (Xing *et al.*, 2011), phylogenetic analysis of the *ARF* genes in maize, *Arabidopsis* and rice.

110

### 3.4.2.3   A detailed examination of the *AUXIN RESPONSE FACTORS*

My analysis identified 20 *ARF* genes in wheat homologous to maize ARF genes, of which 19 are represented on all three genomes, with one gene (*ARF5*) having two (Figure 8a). To our knowledge, none of these genes have been functionally characterised in wheat inflorescence development.

I identified four class I *ARF* (*AtARF3/4-like)* genes in wheat. *ARF11*, *ARF23*, *ARF24* and *ARF12*. *ARF23* homoeologues show a consistent expression pattern, with a peak at DR, and significantly lower expression in the *ppd-1* null line relative to wild-type (Figure 3.9). *ARF11* is expressed highly at every stage of inflorescence development (>10 TPM), with higher expression observed in the *Ppd-D1a* line at LP and lower expression in the *ppd-1* line at TS. *ARF12* is expressed in all three genomes, with *ARF-A12* expressed 5-fold higher in wild-type, relative to the B and D copies. It is expressed at every stage with a decrease in expression at LP. The effect of *Ppd-1* allelism on *ARF-12* expression varies on a stage-specific bases, and in the *ppd-1* null line there is an increase in expression at the LP stage, whereas in the *Ppd-D1a* line, there is a dip in expression earlier at DR.

I identified five class II *(AtARF10, 16, 17-like)* genes in wheat (Figure 3.8). One gene (*ARF5*) is not expressed, while four genes (*ARF6*, *ARF15*, *ARF2*, *ARF19*) are expressed lowly (<10 TPM) (Figure 3.9). *ARF19 and ARF2* show similar expression profiles, both have all three genomes expressed lowly with a decreasing expression pattern as the stages progress from 3 TPM to 1 TPM. *ARF15* and *ARF6* are both expressed very lowly at every stage, with only the A and D copies expressed (<2 TPM).

In wheat, we see only one class III (*AtARF1/2*-like) gene expressed, namely *ARF25* (Figure 3.9). *ARF25* shows the highest expression of any *ARF* gene examined in the wheat inflorescence. The expression is high throughout all stages examined, with a peak at the DR stage. The transcripts were higher in *ppd-1* lines, relative to the *Ppd-D1a* insensitive and wild-type genotypes, at all stages. All *ARF25* homeologs are expressed highly, with *ARF-25B* expressed lowest (Figure 3.9).

In wheat, there is only one gene belonging to class IV (*AtARF5-like*) genes, *ARF4*. Here, I detected high expression throughout meristem development, particularly for the *ARF-A4* copy (Figure 3.9). *ARF-A4* transcripts peaked at the DR stage (30 TPM), relative to the other three stages (20 TPM). The expression trends to be lower in the *ppd-1* line compared to wild-type, particularly at the DR stage.

I identified four wheat class V (*AtARF6/8-like*) genes, namely *ARF3*, *ARF9*, *ARF16* and *ARF18* (Figure 3.8a). *ARF3* exhibits high expression throughout inflorescence development, with all three homoeologues expressed (Figure 3.9). *ARF-B3* is expressed

three-fold lower than *ARF-A3* or *ARF-D3*, increasing as the meristem develops with highest expression at the LP stage. I did not detect any significant effect of *Ppd-1* alleles on *ARF3* expression. *ARF9* shows high expression throughout inflorescence development, with stronger expression as development progresses, with *ARF-B9* the highest expressed homoeolog. *ARF16* shows only very slight expression at all stages with a peak in expression at the LP stage. Expression was two-fold higher in the *Ppd-D1a* line at the LP stage. Expression in the *ppd-1* null line is like wild-type. *ARF18* is expressed at every stage, but predominantly from the B genome copy and is unaffected by *Ppd-1*.

There are four class VI (*ARF7/19*) in wheat: *ARF20, ARF35, ARF1* and *ARF27. ARF20* is expressed lowly at all stages of inflorescence development with a peak at DR, *ARF-D20* is the highest expressed homoeologue and *ARF-A20* the lowest. *ARF1* is not expressed at any stage in any genotype. *ARF24* is expressed very highly at every stage, being the second-highest expressed *ARF*. All homoeologues are expressed robustly, with *ARF-D24* being the most abundant. There is a peak in expression at DR, at which point the null and *Ppd-D1a* line are expressed lower in all three genomes compared to wild-type. Notably, *ARF-D27* is expressed higher in the *Ppd-D1a* line and lower in the *ppd-1* null line, relative to wild-type at the TS stage. *ARF35* is expressed highly at all stages throughout inflorescence development, with *ARF35* showing different stage-dependent expression patterns of each genome copy. *ARF-B35* is the highest expressed homoeologue throughout development. In wild-type, *ARF-B35* peaks at the DR stage showing lower expression in the *Ppd-D1a* line and the null line relative to wild-type. Whereas, at the VG stage expression is higher in the *Ppd-D1a* and *ppd-1* null line.

Together, the expression of these genes suggests significant involvement of auxin-related processes throughout inflorescence development, with examples of genes being positively, negatively regulated, or unaffected, by *Ppd-1* allelism. Investigating the function and spatial expression of these genes represents interesting avenues for future research.
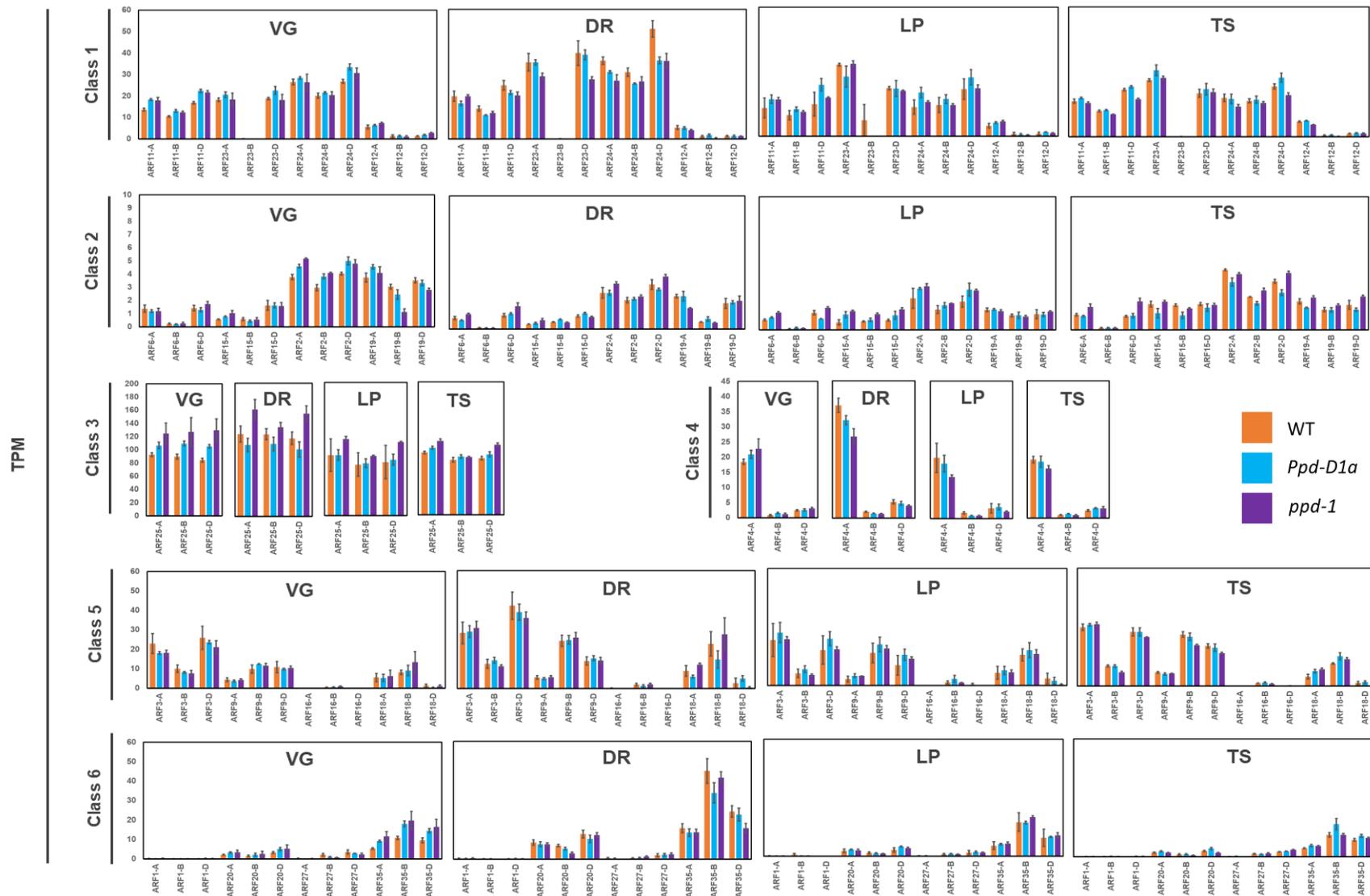
**Figure 3.9 - The TPM expression of the *ARF* family of genes**

Expression of all *ARF* genes identified in this study in wild-type (orange), *Ppd-D1a* photoperiod-insensitive (cyan) and null *ppd-1* (purple) NILs under field conditions. Expression of each gene is shown for the stages: vegetative (VG), double ridge (DR), lemma primordium (LP) and terminal spikelet (TS). Graphs are bar graphs that show the TPM ± SEM error bars of three biological replicates.

### 3.4.3 The impact of *Ppd-1* allelism on three key gene family's expression clusters.

In this study, I used cluster analysis on a global transcriptome level to investigate how expression patterns shifted between *Ppd-1* genotypes (Figure 3.6). However, to examine the genes that fall into these clusters we have performed a subsequent clustering analysis on key gene families and compared the top two represented clusters for each genotype (Figure 3.10a). Cyclins were selected due to their reported role in the cell cycle and cell division, which are processes likely to be involved with the significant morphological changes in the meristem as it transitions from the VG to the TS stage (Engler *et al.*, 2009; Gaamouche *et al.*, 2010). Within the gene family, the largest expression profiles change in a genotype-dependent manner. We see a general shift in expression to an earlier stage in the *Ppd-D1a* line, with Cluster 1 down-regulated at the LP stage before spikelet differentiation terminates (Figure 3.10). Down-regulation of Cluster 2 genes that in wild-type, occurs at LP happens earlier in the *Ppd-D1a* lines, with down-regulation at the DR stage (Figure 3.10). Interestingly, this precise regulation is lost in the *ppd-1* line, falling into the patterns associated with C5 and C11 in the whole genome clustering analysis (Figure 3.6).

We also looked at the histone family as histones and chromatin regulators have reported roles in the regulation of chromatin states affecting gene expression (Parvathaneni *et al.*, 2020; Figure 3.10b). These clusters appear to follow a similar trend to the major differences we see in the whole genome comparison (Figure 3.6). In the photoperiod-insensitive line, we see a lack of up- or down-regulation at the LP to TS transition, pointing to *Ppd-1* playing a role at this point. The genes in this cluster are also mis-regulated in the *ppd-1* line after the VG stage (Figure 3.10b).

The MADS-box family that have a well-reported role in floral development show a remarkably conserved patterning of gene expression. However, in the *ppd-1* line, there appears to be a slight delay in expression. Cluster 2 of MADS-box genes from wild-type, in which genes are up-regulated between the VG and DR, shifts to being induced between DR and LP in the *ppd-1* line.

Together, these patterns point to a system whereby processes such as cyclin mediated cell division and histone-based regulation are mis-regulated by genetic diversity for *Ppd-1*. Whereas the MADS-box gene expression profiles are much more robustly maintained in the developing inflorescence.
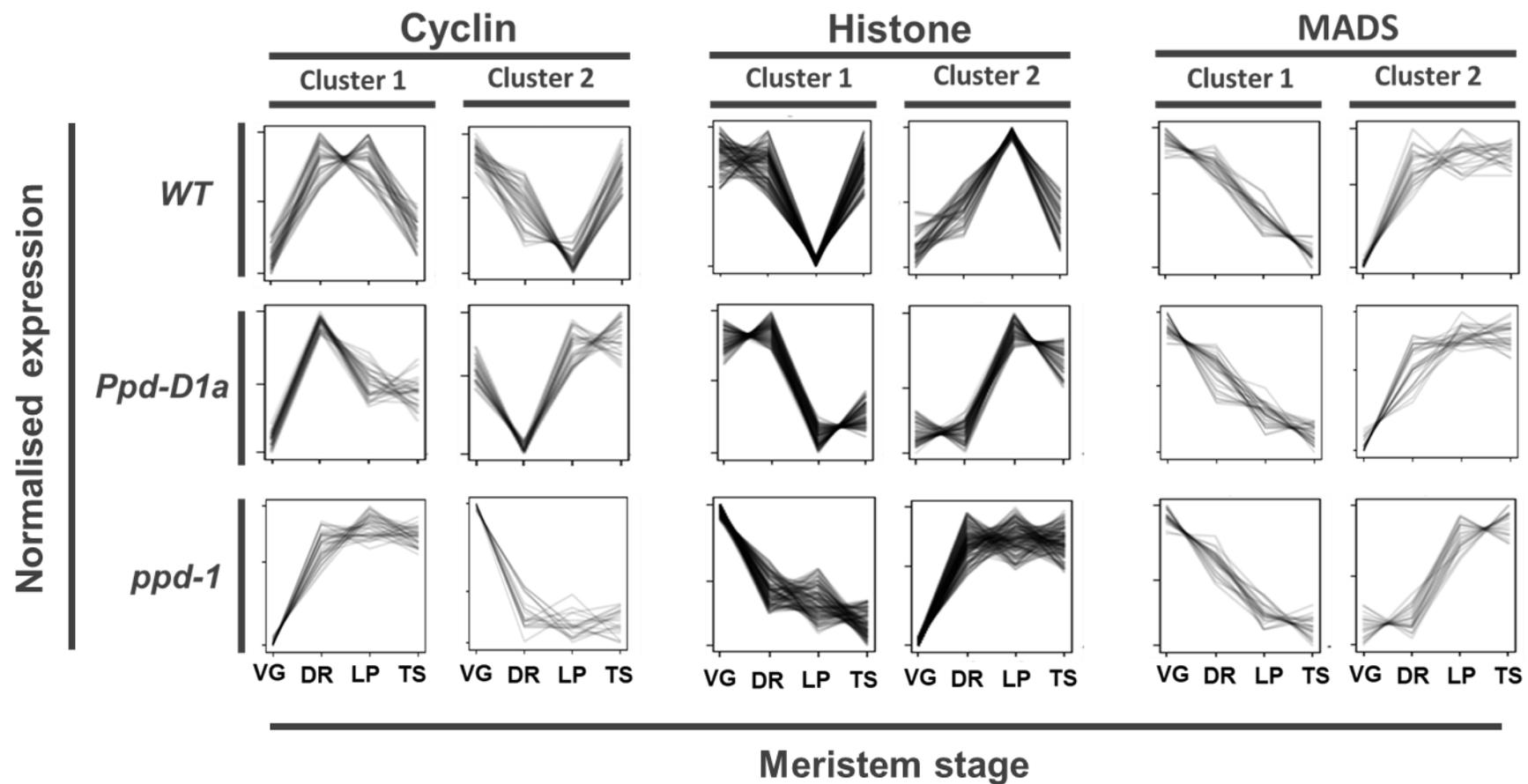
**Figure 3.10 - The top two gene clusters for major gene families**

The top two clustered profiles for the Cyclin, Histone and MADS-box associated genes. The top profiles for each gene family are shown for the wild-type, *Ppd-D1a* and *ppd-1 lines*. Gene families were identified through functional annotation (Ensembl plants biomart). Clustered expression profiles are plotted against the vegetative (VG), double ridge (DR), lemma primordium (LP) and terminal spikelet (TS) stages.

### 3.4.4 The role of MADS-box genes in the inflorescence

#### 3.4.4.1 ABCDE flowering pathway genes

The flowering pathway is well studied in model organisms such as *Arabidopsis*, and is embodied by the ABCDE pathway (Coen & Meyerowitz, 1991; Rijpkema *et al.*, 2010; Kitagawa *et al.*, 2012; Murai, 2013). However, very few studies have documented this pathway in wheat. Here, I aim to use the RNA-seq analysis to determine conservation of gene function and characterise unexpected expression patterns. Using a list of 201 wheat MIKC-type MADS-box genes (MADS-box genes) recently identified through phylogenetic analysis, I have performed clustering analysis on the expression profiles as represented by their normalised expression patterns at the VG, DR, LP and TS stages (Figure 3.11; Schilling *et al.*, 2020). The VG to TS stages encompass the induction of many floral organs in wheat, from the glume in the glume primordium stage through to the florets at the terminal spikelet stage (Kirby & Appleyard, 1984). As such, I have been able to examine the expression of the MADS-box genes classically involved in the ABCDE pathway and determine their putative function within the ABCDE model based on reports from *Arabidopsis* and rice.
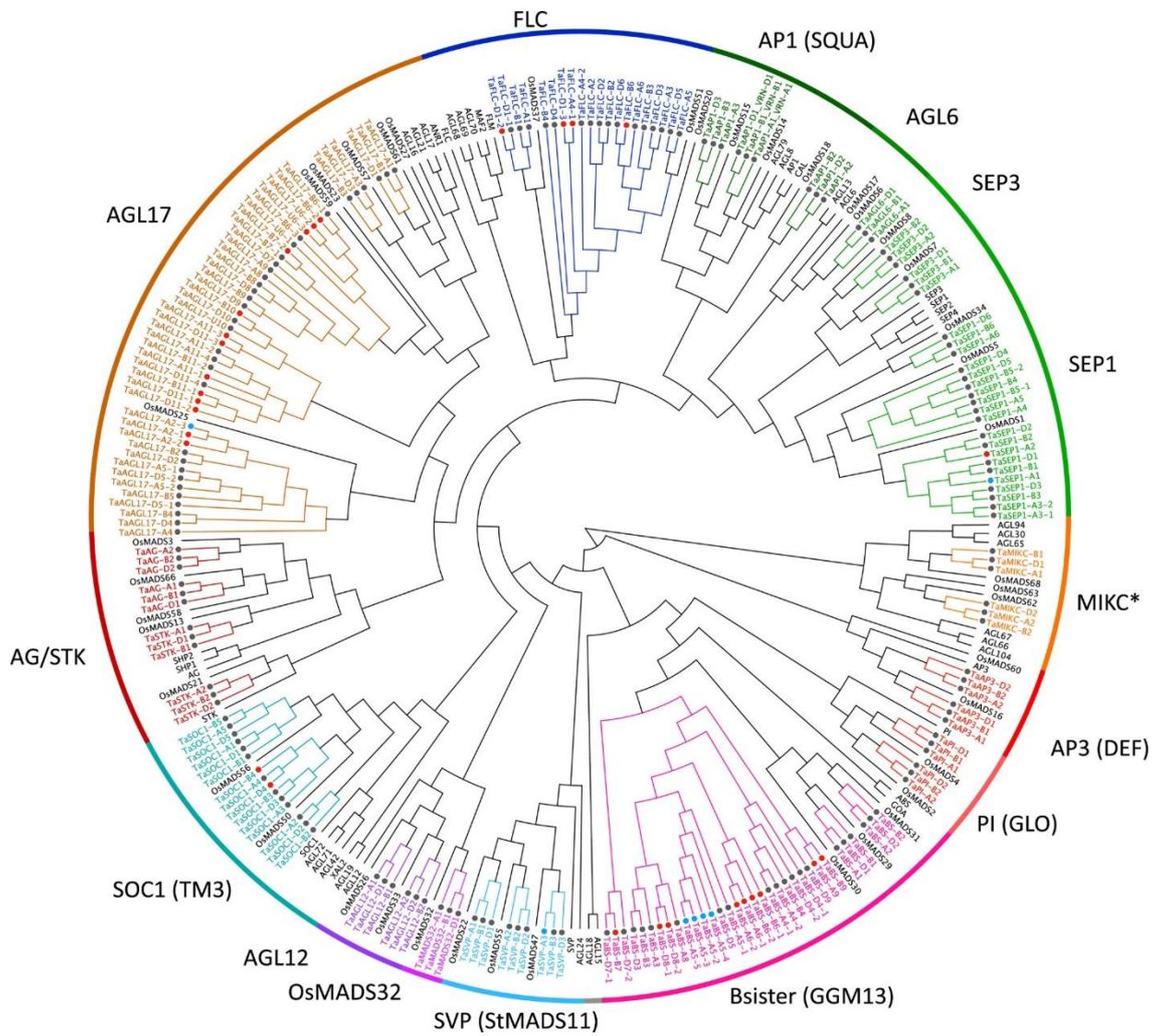
**Figure 3.11 - A phylogenetic analysis of the MADS-box genes**

Figure and partial legend taken from (Schilling *et al.*, 2020). A phylogenetic tree of MADS-domain proteins from bread wheat, cultivated rice and *Arabidopsis*. Wheat genes are coloured subclade-specific, whereas rice and *Arabidopsis* genes are in black.

### 3.4.4.2  Predicting ABCDE family function based on MADS-box clustering

To investigate the function of these genes in the developing inflorescence, I preformed clustering analysis on all MADS-box genes expressed in the developing wheat inflorescence. Based on clustered expression profiles, we can infer the roles of these genes in inflorescence development (Figure 3.12). *SOC1-D2*, *SVP-B1*, *SVP-D1* and *SVP-D2* were down-regulated from the VG stage onwards, indicating their involvement at these early stages of inflorescence development. The genes up-regulated directly after the VG stage, are *AP1-B1, AP1-B2, AP1-B3, MADS32-A1, MADS32-B1, SEP1-A6, SEP1-B6, SEP1-D5* and *SOC1-A1*. These genes are up-regulated post VG stage and are expressed highly at the DR, LP and TS stages, suggesting a role in promoting spikelet and floret meristem development. *STK-A1* and *STK-D1* show a peak in expression at the DR stage, and gradual reduction in expression as the meristem develops, potentially indicating a role in spikelet meristem development. *AG-A2*, *AP1-D1*, *PI-B1* and *PI-B2* show high expression at every stage bar the LP stage, where they are dramatically down-regulated, potentially indicating a role a role at the LP stage as florets begin to form. The *AG-A1*, *AG-B1*, *AG-D1*, *PI-A1*, *SEP1-D1*, *SEP3-A2* and *SEP3-B1* genes show a significant upturn in expression from the LP to TS stage, indicating a role in the promotion of floral organs, as it is the TS stage where the floret primordia begin to differentiate. *AGL12-A2*, *AGL12-B2*, *TaAGL12-D2*, *AGL17-A1*, *BS-A6-1*, *PI-D1*, *STK-B1*, *STK-B1*, *SVP-A1*, *SVP-A2*, *SVP-B3* and *SVP-D3* show a gradual down-regulation as stages progress, indicating a role during early development. *SEP1-A5* shows an uptrend in expression from the DR to LP stage where it remains highly expressed, suggesting a role in floral organ development. *BS-AB*, *BS-B6-2* and *BS-DB*, show high expression profile until a dramatic down-regulation at the TS stage, pointing towards a role for these genes during glume or early floret initiation.
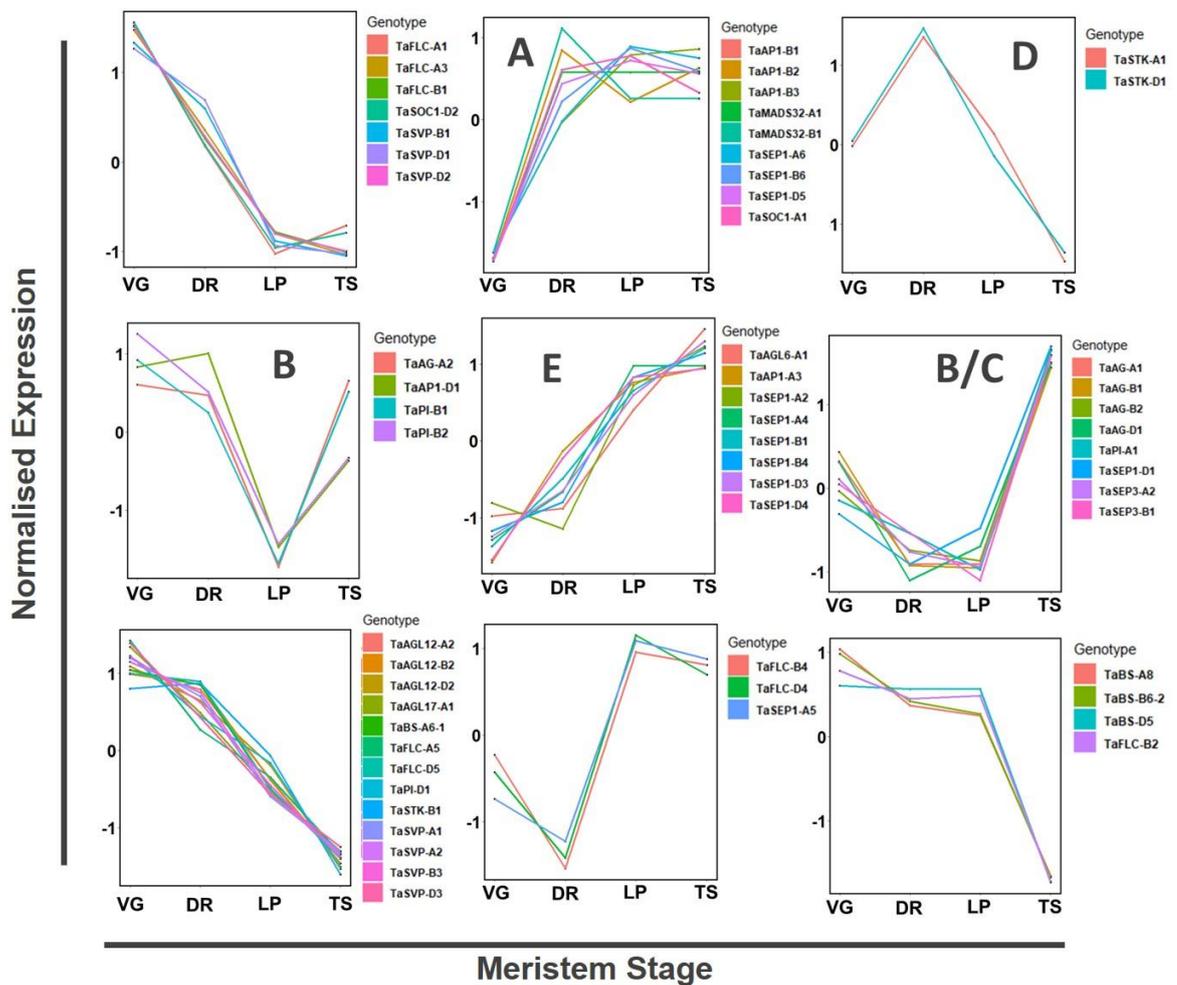
**Figure 3.12 - Clustered gene expression of MADS-Box genes**

Eight clusters of expression from a clustering analysis of all the MADS-box genes currently identified in wheat for wild-type plants. Expression pattern is shown at the vegetative (VG), double ridge (DR), lemma primordium (LP) and terminal spikelet (TS) stages. The key next to each sub-graph shows the name and representative colour of the gene. Patterns of clustered were associated with genes from the ABCDE flowering pathway, denoted within each graph.

### 3.4.4.3 Expression of MADS-box genes in wheat

The MADS-box genes have a diverse range of functions within floral meristem development, as shown by the broad range of expression patterns (Figure 3.12), but the expression profiles of the MADS-box genes are not always consistent amongst gene copies. To explore this trend further, I plotted the TPM of all the MADS-box genes expressed in the wheat inflorescence (Figure 3.13).

 *AP1* and *AP2* represent the A family of the ABCDE genes (Jofuku *et al.*, 1994). In this data set, the *AP1* genes show varied expression patterns, with high expression throughout development in the *AP1-1* (VRN1) genes and increasing expression in the *AP1-2* and *AP1-3* genes. These genes will be covered in detail in Chapter 4. The *AP2* genes are the only ABCDE family genes that are not MADS-box genes (Kim *et al.*, 2006).

*AP3* and *PI* genes represent the B family (Jack *et al.*, 1992; Goto & Meyerowitz, 1994). The *AP3* genes exhibit varied expression profiles. *AP3-A1* and *D1* both show a conserved pattern, with the *AP3-A1* gene showing an initial peak at the DR stage (7.29 TPM), down-regulation at the LP stage and a subsequent peak in expression at the TS stage (12.87 TPM). *AP3-D3,* the only other *AP3* gene expressed, shows a gradual increase in expression with a peak at the TS stage (23.07 TPM). The *PI* genes show varied expression patterns. *PI-A1*, *B1* and *D2* share a similar profile, being expressed lowly at the VG, DR and LP stages but with a peak at the TS stage, *PI-A1* reaching 22.30 TPM. *PI-B2* and *D1* display a much more disordered expression pattern, with low expression throughout inflorescence development.

The class C family are represented by the *AG* genes (Yanofsky *et al.*, 1990). These genes share a consistent pattern, low expression until the LP stage and a strong upturn in expression towards TS. The highest expressed *AG* gene is *AG-A1*, reaching 3.7 TPM at TS. Notably, unlike the other *AG* genes, *AG-A2* is expressed at the DR stage (0.55 TPM). Possibly indicating a role for this paralogue earlier than its counterparts.

The class D genes consist of the *AGL* genes and *STK* (Pinyopich *et al.*, 2003; Favaro *et al.*, 2003). Four *AGL* genes are expressed in this data set, two *AGL12* genes (*AGL12-A2* and *B2*) share a similar expression pattern, with a peak in expression at the DR stage (1.2 TPM). *AGL6-A1* shows a robust expression pattern, with no expression of this gene until the TS stage, at which point it is expressed highly (21.79 TPM). There are two *STK* genes expressed in wheat, *STK-A1* and *B1*. In wild-type, *STK-B1* is the highest expressed and shows consistent expression throughout inflorescence development with a peak at the DR stage (2.2 TPM).

The *SEP* genes constitute the class E genes (Ditta *et al.*, 2004). The *SEP* family of genes is the largest family expressed in the developing inflorescence, with 15 *SEP1* genes and 3 *SEP3* genes being expressed. There is generally a pattern of increasing expression as the meristem develops, although it is interesting that the wheat paralogs show differences to each other. For example, *SEP1-B1* is expressed only lowly at the LP stage (1.9 TPM) with a significant five-fold increase in transcripts at the TS stage (10.3 TPM). Whereas *SEP1-D2* shows expression of (8.85 TPM) at LP and a 2.5-fold increase (20.2 TPM) at the TS stage. *SEP1-D2* is also expressed significantly at the VG (2.55 TPM) and DR (3 TPM) stages, whereas the *SEP1-1* genes are not expressed at these stages (<0.5 TPM).

Beyond the ABCDE pathway there are several other MADS-box genes with interesting expression profiles. For example, the monocot specific gene *MADS32* is expressed highly at all stages, with a gradual increase in expression as the inflorescence develops (Wang *et al.*, 2015). The highest expressed copy, *MADS32-D1*, peaks in expression at the TS stage (231.8 TPM). *SHORT VEGETATIVE PHASE* (*SVP*) genes, as their names suggest in *Arabidopsis* are expressed during vegetative development in the *Arabidopsis* SAM and act with *FLC* to negatively regulate *FT* and *SOC1,* suppressing the floral transition (Willmann & Poethig, 2011; Liu *et al.*, 2018). The eight *SVP* genes represented show a very robust and consistent expression pattern, *SVP-A1*, *-B1* and *-D1* all show equally high expression at the VG stage (57 TPM) and are subsequently down-regulated as development proceeds. The *SOC1* genes are core floral regulators, with expression stimulated as a result of *FT* arriving to the SAM (Nilsson *et al.*, 1998; Samach *et al.*, 2000). *SOC1-A3*, *B3* and *D3* all have consistent expression with an increase in expression until the DR stage where it flatlines until down-regulation at the TS stage. *SOC1-A3* is the highest expressed of these genes reaching a peak of expression of 27 (TPM) at the DR stage. *SOC1-A1*, *D1*, *B1* and *B5* show a similar pattern but peak later at the LP stage, with *SOC1-B5* showing highest expression (10.37 TPM). *SOC1-D2* has a unique expression pattern, with high expression at the VG stage (6.96 TPM) with no significant expression thereafter. The *Bsister* genes play an important role in ovule and seed development (Schilling *et al.*, 2015). These show a trend of constant expression throughout the inflorescence, with the highest expressed gene being *BS-A6-1* with a peak in expression at the DR stage (5.89 TPM).

The varied expression profiles of MADS-box genes and their paralogues indicate that whilst their core function may be conserved in wheat relative to *Arabidopsis* and rice, sub-functionalisation may be occurring.

### 3.4.4.4 *Ppd-1*-mediated expression of MADS-box genes in wheat

My analysis examined the expression of all the MIKC-type MADS-box genes currently identified in wheat (Figure 3.13; Schilling *et al.*, 2020). This chapter has identified a

considerable impact of *Ppd-1* allelism on inflorescence gene expression. As discussed, the MADS-box genes have a key role in SAM development. I therefore investigated the influence the *Ppd-D1a* insensitive and *ppd-1* null alleles on MADS-box gene expression. Consistent with previous analysis of selected MADS-box genes, the expression of all members of this family remains consistent between genotypes when normalised to meristem stage (Figure 2.14; Gauley & Boden, 2020; Boden *et al.*, 2015; Dixon *et al.*, 2018). The exception to this trend is that several genes are downregulated in the *ppd-1* line, which shows decreased expression of *SEP1-6, AP1-1* (*VRN1*), *STK* and the *SVP* genes (Figure 2.7; Figure 3.13). *STK-B1* shows a particularly striking profile at the DR stage with the wild-type showing moderate expression (2.2 TPM) compared to higher expression in the *Ppd-D1a* NIL (3.7 TPM) and dramatically lower expression in the *ppd-1* NIL (0.88 TPM). Conversely, the *AG, AP3* and *Bsister* genes trend lower in the insensitive line compared to both the null and wild-type lines.

The expression of these important genes show that their profiles are generally conserved based on stage. This trend is not universal, with clear examples of *Ppd-1* having an impact on gene regulation.
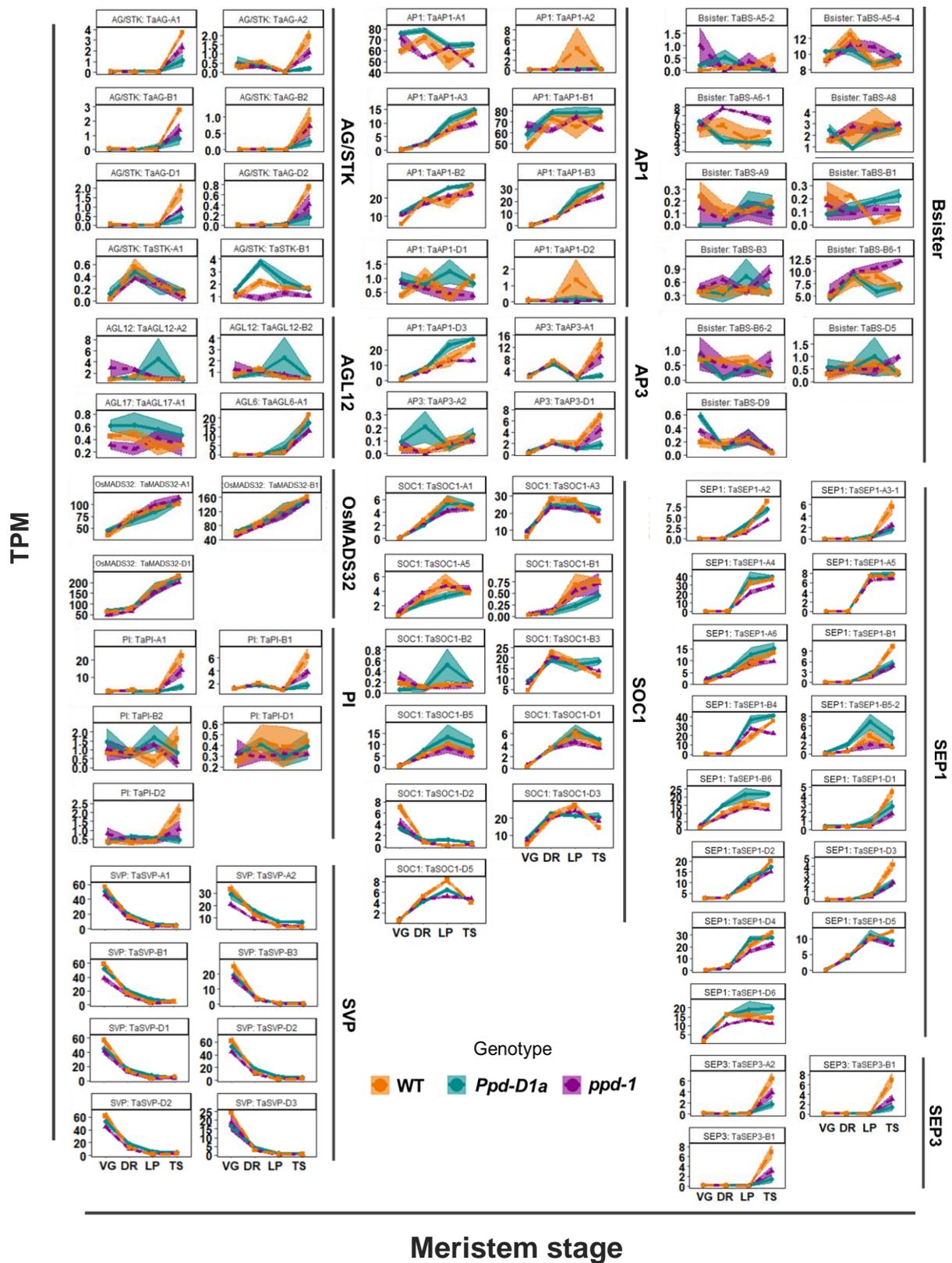
**Figure 3.13 - The TPM expression of the MADS-box gene in wheat**

Shows the TPM expression profiles of all the MADS-box genes currently identified in wheat. In wild-type (orange), *Ppd-D1a* photoperiod-insensitive (cyan) and null *ppd-1* (magenta) NILs under field conditions. Expression is plotted against the meristem stages: vegetative (VG), double ridge (DR), lemma primordium (LP) and terminal spikelet (TS). Graphs are presented as ribbon plots which show the TPM (solid line with data points) ± SEM (shaded region) of three biological replicates.

### 3.4.5 Genome biases in the inflorescence

#### 3.4.5.1 Impact of *Ppd-1* alleles on total genome bias

Polyploidy is a fundamental characteristic of modern wheat, with each genome hybridisation event providing a paradigm shift in gene regulation. However, as well as additional gene copies being added during genome hybridisation, there is evidence for *biased fractionation*. *Biased fractionation* is a process by which the function of a sub-genome gene copy is lost during hybridisation as the gene doubling (tetraploid) or tripling (hexaploid), is unstable (Emery *et al.*, 2018). This study examines 39,474 genes with expression in more than one genome of hexaploid wheat. I observe 24,946 of these genes carrying copies on the A, B and D genomes, 4,975 in B and D, 4,859 in A and D and 4,694 in the A and B. These numbers suggest a pattern of *biased fractionation*, whereby the A genome as the oldest genome is more likely to lose a gene copy as hybridisation events occurred, with the D genome as the most recent addition showing the greatest number of retained copies (Salamini *et al.*, 2002; Petersen *et al.*, 2006).

Due to the hexaploid nature of wheat, it is important to examine the effects of gene regulation in the polyploid context. I hypothesise that individual genomes play a more active role in the regulation of biological processes, and that this may be influenced by *Ppd-1* allelism. To investigate this, I interrogated the expression of all gene triads in wheat (21,627) at the VG, DR, LP and TS stages (Figure 3.14). Overall, there is a shift away from the B genome to towards the A and D genomes at all points, with a preference for the D genome. The average expression contribution of the A, B and D genomes across all four stages is 33.446% (± SEM 0.100) 30.5104% (± SEM 0.069) and 36.0436% (± SEM 0.070), respectively (Table 3.3), with the relationship being highly significant (Kruskal-Wallis *p* value < 0.001). This expression bias indicates gene transcription in the meristem is dominated by the D genome to a greater extent than that shown in other wheat tissues (Ramírez-González *et al.*, 2018; Jordan *et al.*, 2020). Through examination of the average TPM expression at the key stages for each of the genomes, it is clear that VG expression is higher in the *ppd-1* line, indicating a role for *Ppd-1* in negatively regulating gene expression at this stage (Figure 3.15). It is also higher in the D and A genome for the *Ppd-D1a* NIL, to levels almost identical to the DR wild-type expression, furthering the idea that the insensitive line shifts gene expression patterns earlier but not to higher levels, as occurs in the leaf (Figure 2.14). An important take home message is that *Ppd-1* affects the expression of a considerable number of genes in the meristem, detectable on a global genome-scale. It is also apparent that an insensitive copy of *Ppd-1* on the D genome can influence transcription over all three genomes, pointing to intergenomic interaction.
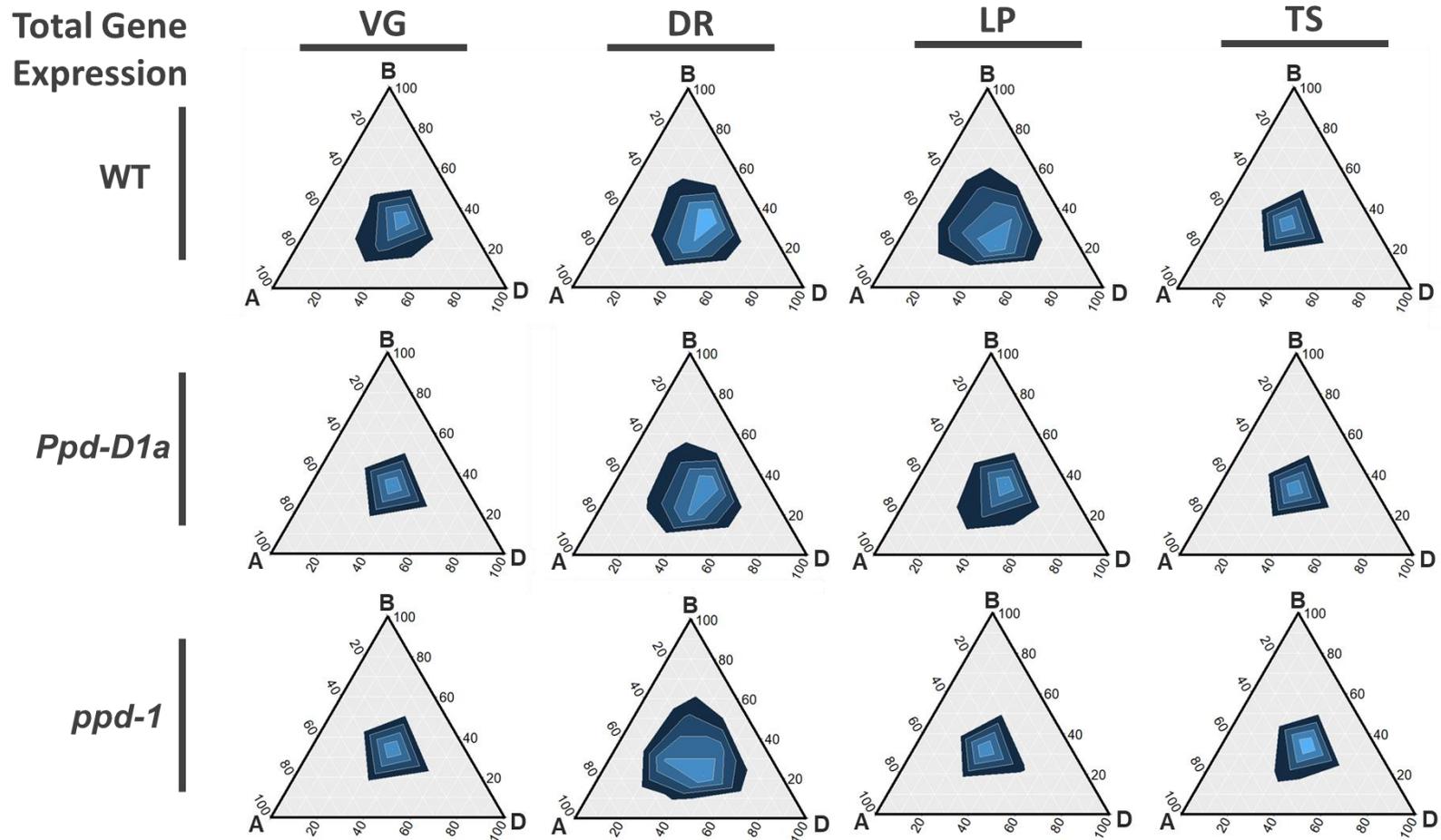
**Figure 3.14 - The percentage contribution of genome expression for every gene triad in wheat**

Ternary plot showing relative expression abundance of 21,627 gene triads. Each corner represents the A, B or D genome, with the scale on each side representing the percentage contribution by that genome. Density of data points have been visualised using Kde2d two-dimensional kernel density estimation with increasingly lighter shades of blue representing higher density. Graphs show genome percentage distribution at the vegetative (VG), double ridge (DR), lemma primordium (LP) and terminal spikelet (TS) stages in the Wild-type, *Ppd-D1a* and *ppd-1* lines.
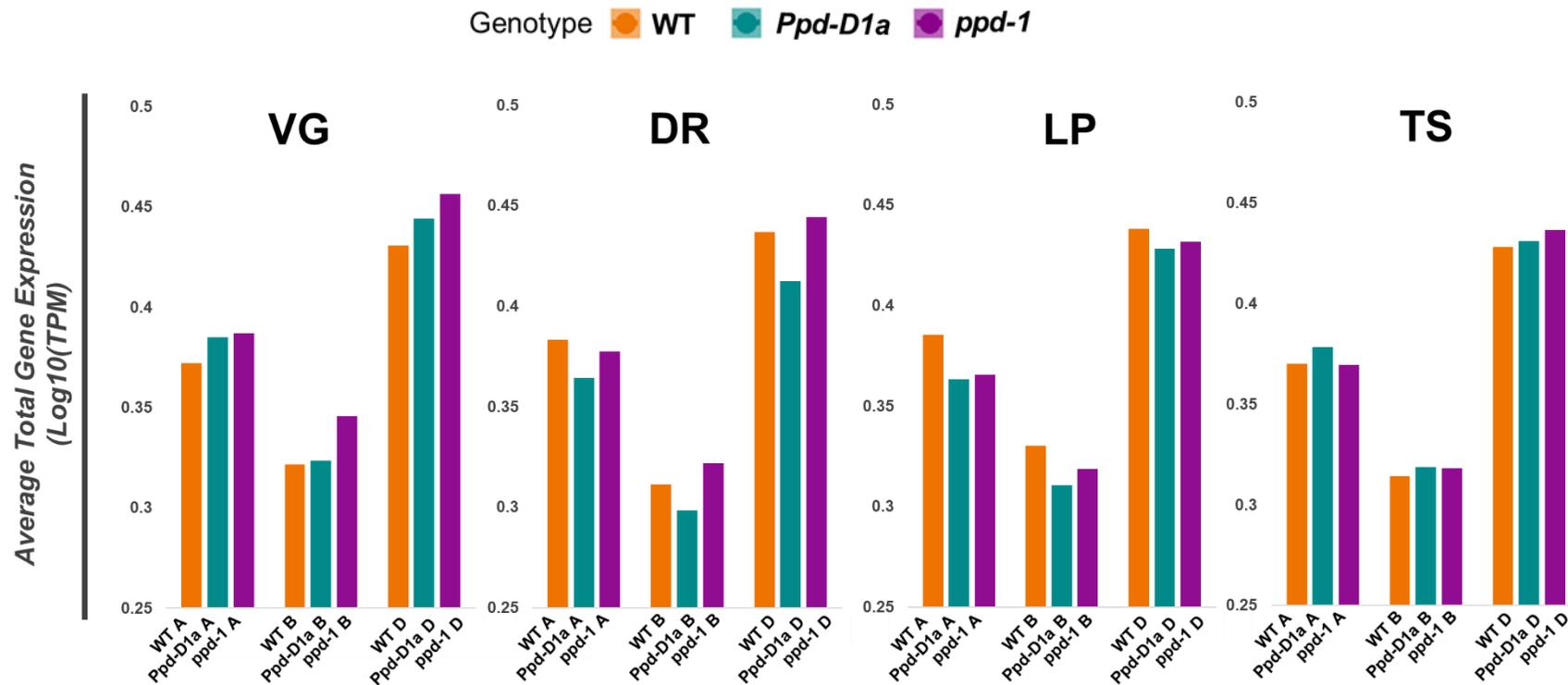
**Figure 3.15 - Average total gene expression**

The average Log 10(TPM) expression of all gene triads expressed in the vegetative (VG), double ridge (DR), lemma primordium (LP) and terminal spikelet (TS) stages. 21,627 gene triads total. Each graph shows the expression of the A, B and D genomes. Wild-type (orange), *Ppd-D1a* photoperiod-insensitive (cyan) and null *ppd-1* (magenta) NILs under field conditions. Data was logged to normalise gene expression.

### 3.4.5.2   Impact of *Ppd-1* alleles on MADS-box genome bias

The MADS-box genes play an essential role in the regulation of inflorescence development in wheat; however, the majority of our understanding comes from diploid plants (Coen & Meyerowitz, 1991; Koornneef & Meinke, 2010; Schilling *et al.*, 2018). This leaves a gap in our understanding of these genes in wheat as a polyploid, relative to model species.

When investigating key gene families such as the MADS-box genes, considering the contribution of each genome to total expression and how they change activity throughout development could provide insights into their function. In this study, I examined 20 MADS-box gene triads. During the vegetative stage of wild-type, there is a tendency towards expression from the B (36%) and D (39%) genomes, compared to the A genome (24%) (Figure 3.16, Table 3.3). This trend continues throughout the development; however, the bias fluctuates from a high of 40.5% at the DR stage to a low of 32.8% at the TS stage. This trend is maintained in the photoperiod insensitive *Ppd-D1a* lines. In the *ppd-1* line there are changes in percentage contributions of genomes, particularly at the LP stage with the B genome (38%) contributing more expression than the A (31.5%) or D genome (30.37%). Interestingly, there are stage specific differences of percentage genome contribution. The A genome contribution increases as the inflorescences develop, from 24.7% at the VG stage to 34.5% at the TS stage. The D genome, interestingly, has the opposite effect, contributing 39% of expression at the VG stage and gradually decreasing to 32.8% at the TS stage, whilst the B genome remains relatively consistent in wild-type.

Perhaps the most interesting finding is that *Ppd-1* allelism doesn't affect genome expression equally in MADS-box gene regulation. For example, at the LP stage the average expression in the null lines, increases on the B genome, but decreases on the D genome. Together, these data point to genome specific regulation of MADS-box genes that varies based on stage and genotype.
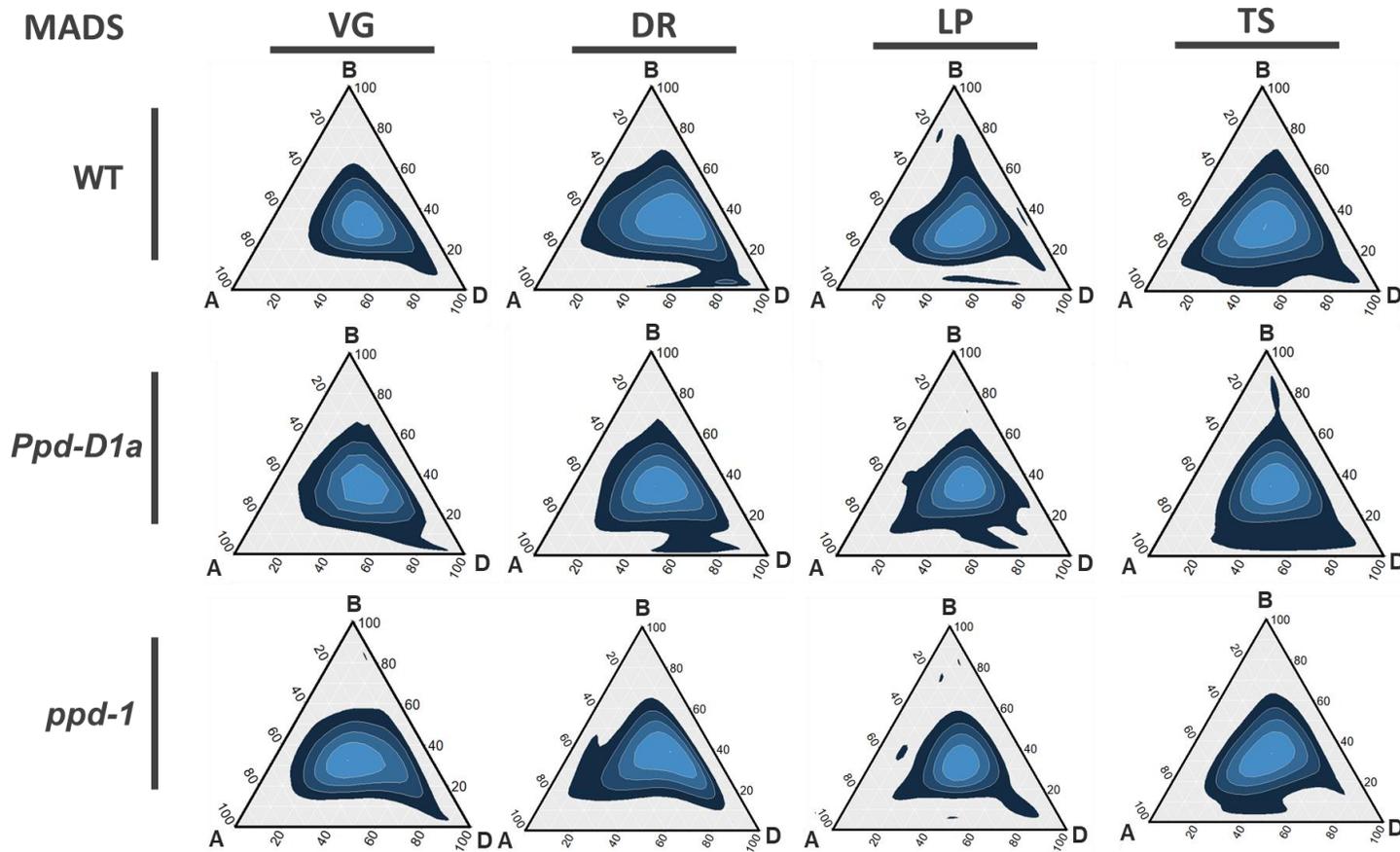
**Figure 3.16 - The percentage contribution of genome expression for every MADS-box gene in wheat**

Ternary plot showing relative expression abundance of 20 gene triads (60 genes total). Each corner represents the A, B or D genome, with the scale on each side representing the percentage contribution by that genome. Density of data points have been visualised using Kde2d two-dimensional kernel density estimation with increasingly lighter shades of blue representing higher density. Graphs show genome percentage distribution at the vegetative (VG), double ridge (DR), lemma primordium (LP) and terminal spikelet (TS) stage in the Wild-type, *Ppd-D1a* and *ppd-1 lines*.
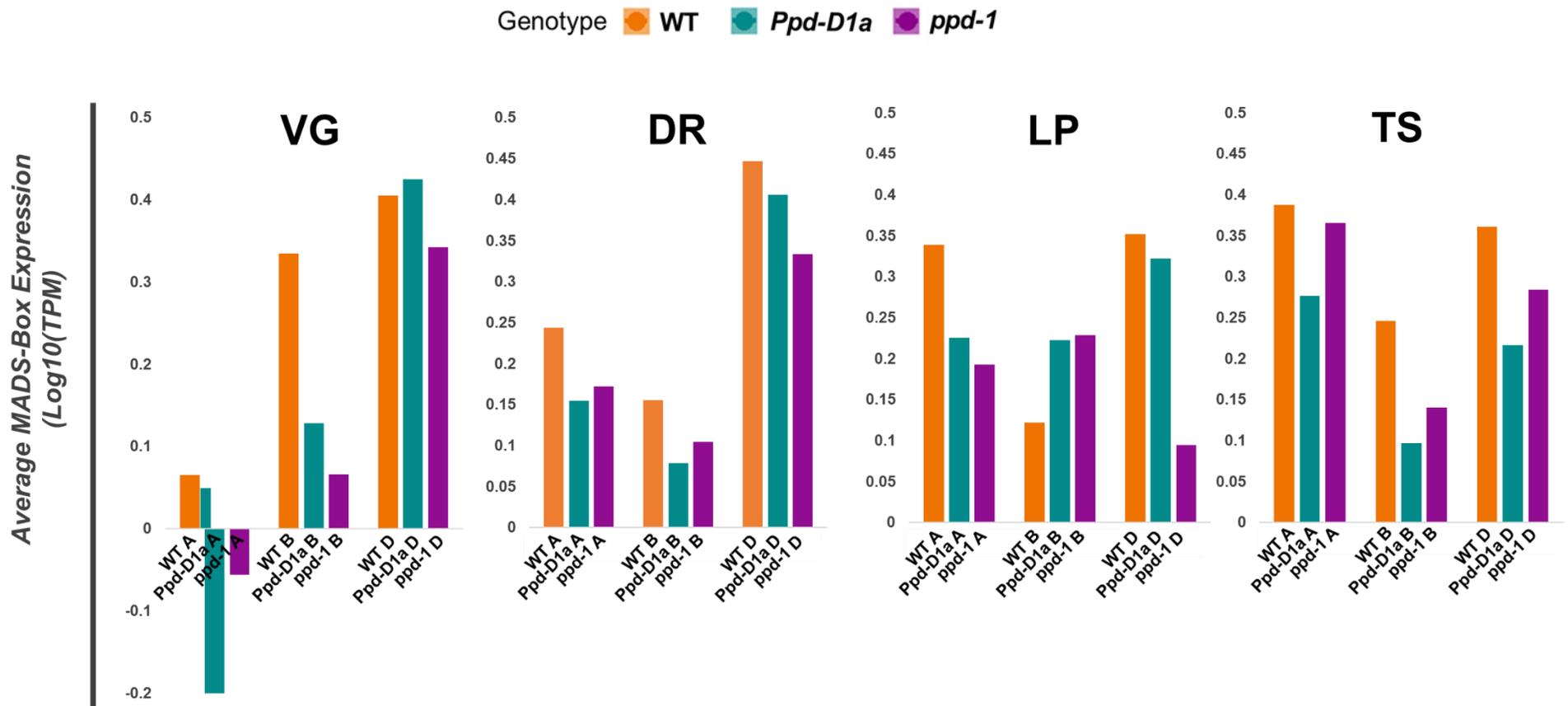
**Figure 3.17 - Average MADS-box gene expression**

The average Log 10(TPM) expression of all MADS-box gene triads expressed in the vegetative (VG), double ridge (DR), lemma primordium (LP) and terminal spikelet (TS) stages. 20 gene triads total. Each graph shows the expression of the A, B and D genomes. Wild-type (orange), *Ppd-D1a* photoperiod-insensitive (cyan) and null *ppd-1* (magenta) NILs under field conditions. Data was logged to normalise gene expression.

### 3.4.5.3 Genome specific expression of auxin-related genes in WT, *Ppd-D1a* and *ppd-1* NILs

To investigate if the trends overserved in the MADS-box genes are present in other important gene families, we examined triad expression patterns in auxin-related genes (Figure 3.18).

Across all genes there is a bias towards the D genome (Figure 3.14, 3.15; Table 3.3). The bias is also seen in the auxin-related gene families, although it is less dominant compared to total gene expression or the MADS-box genes (Figure 3.18; Table 3.3). The B genome is expressed to similar levels to the D genome for the auxin-related genes, particularly in the *Ppd-D1a* line at the VG and DR stages where percentage contribution of the B genome actually exceeds the D genome. In the *ppd-1* NIL, the D genome contributes the greatest proportion of gene expression (35.25 % ± 0.7 SEM), relative to the wild-type (34.86 % ± 0.6 SEM), with a lesser contribution in the *Ppd-D1a* line (34.16 % ± 0.14 SEM).

Based on the average expression there is not a consistent relationship between genomes total expression across development (Figure 3.19). For example, in wild-type at the VG stage, there is a strong bias towards the D genome, with the B genome contributing the least expression. Whereas at the TS stage the expression of the D and B genomes are equal with the A genome showing least expression (Figure 3.19). These trends are different to those observed in the MADS-box genes (Figure 3.17), suggesting the regulatory elements deciding genome specific expression contribution are not consistent across gene families.

At the VG stage, the *Ppd-D1a* line has greater B and D genome expression relative to *ppd-1* or wild-type. This effect may be due to the shifting of expression levels at VG to the same levels as seen in WT at the DR stage, supporting the trends observed previously (Figure 3.7). However, this trend does not continue throughout development, with expression at the TS stage being dominated by the *ppd-1* lines. Expression is higher on average across all stages in both the *Ppd-D1a* and *ppd-1* lines, indicating that the influence of *Ppd-1* on auxin genes is both positive and negative at these stages.

Together, these data illustrate how genome expression biases are not consistent between different families. In addition, the impact of *Ppd-1* regulation on transcription can occur in a genome-specific manner.
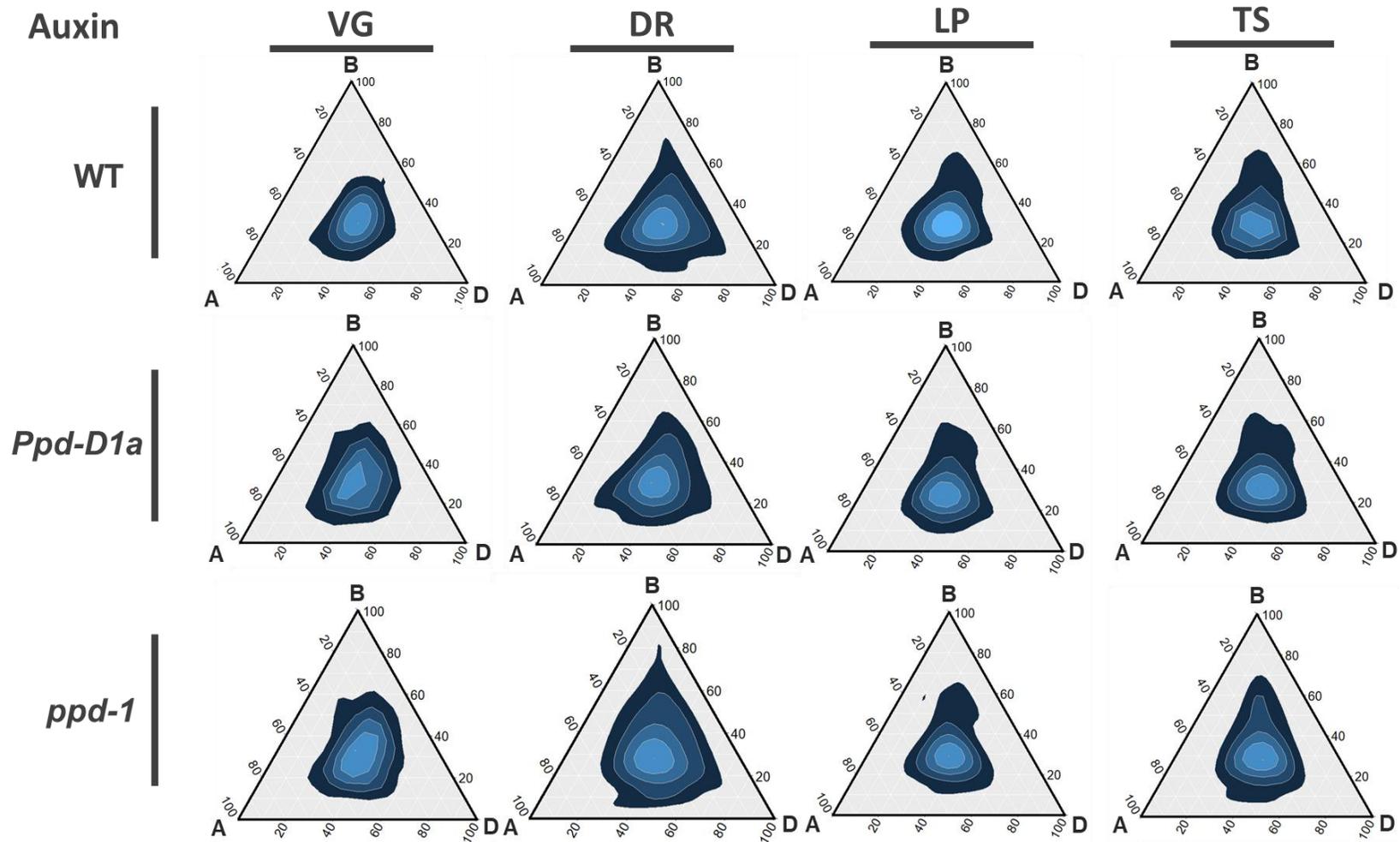
**Figure 3.18 - The percentage contribution of genome expression for every auxin-related gene in wheat**

Ternary plot showing relative expression abundance of 80 auxin related gene triads (240 genes total). Each corner represents the A, B or D genome, with the scale on each side representing the percentage contribution by that genome. Density of data points have been visualised using Kde2d two-dimensional kernel density estimation with increasingly lighter shades of blue representing higher density. Graphs show genome percentage distribution at the vegetative (VG), double ridge (DR), lemma primordium (LP) and terminal spikelet (TS) stages stage in the Wild-type, *Ppd-D1a* and *ppd-1 lines.*
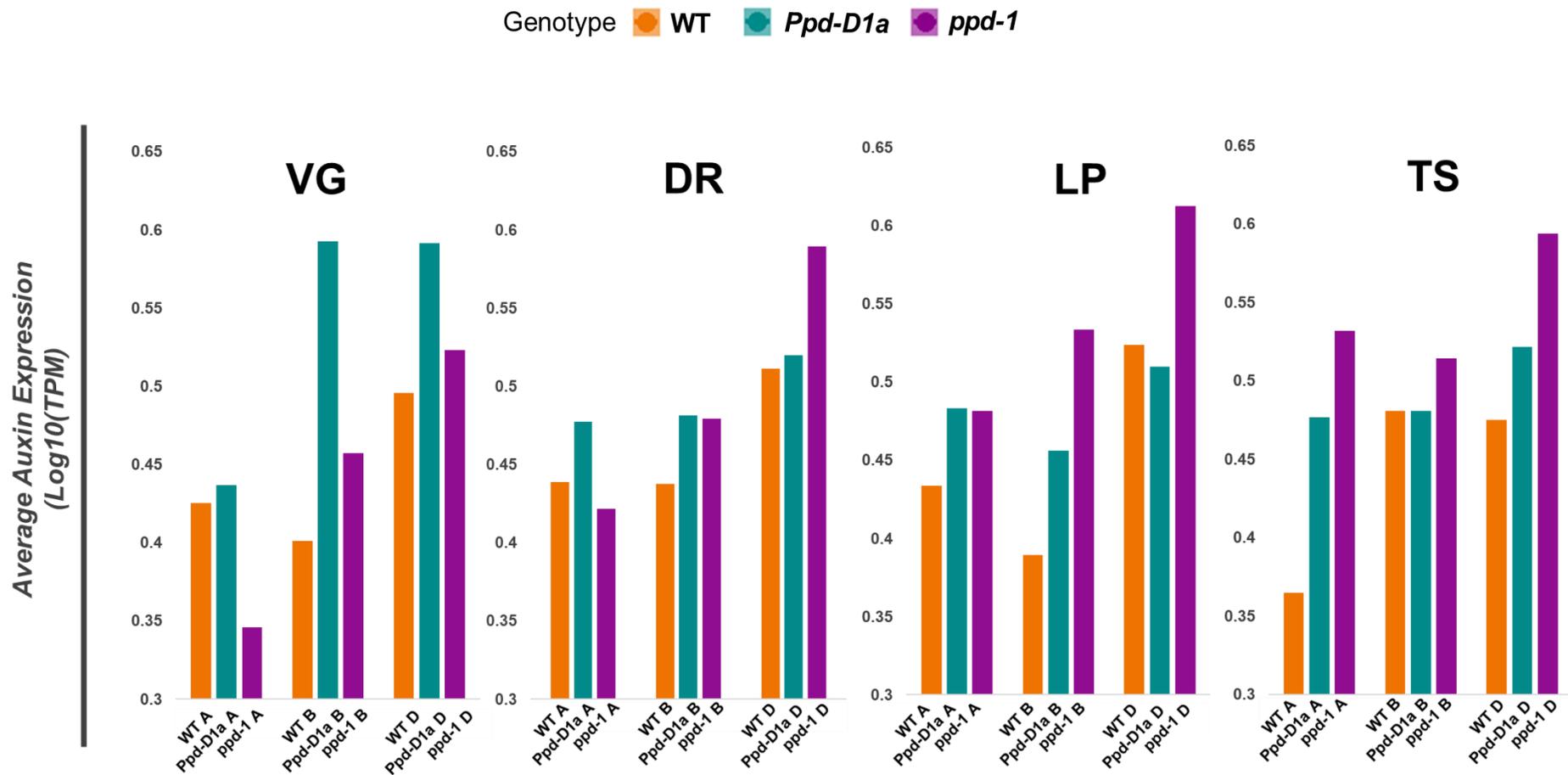
**Figure 3.19 - Average auxin-related gene expression**

The average Log 10(TPM) expression of all Auxin related gene triads expressed in the vegetative (VG), double ridge (DR), lemma primordium (LP) and terminal spikelet (TS) stages. 80 gene triads total. Each graph shows the expression of the A, B and D genomes. Wild-type (orange), *Ppd-D1a* photoperiod-insensitive (cyan) and null *ppd-1* (magenta) NILs under field conditions. Data was logged to normalise gene expression.

**Table 3.3 - Genome expression contribution of key gene families in the inflorescence**

The proportion contribution to expression for each genome. Percentage expression of total gene expression (Total), MADS-box (MADS) and auxin-related (Auxin) gene expression.

| WT Total | A (%) | B (%) | D (%) | WT MADS | A (%) | B (%) | D (%) | WT Auxin | A (%) | B (%) | D (%) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| VG | 33.1518 | 30.6312 | 36.217 | VG | 24.7070512 | 36.22273 | 39.0702 | VG | 32.0109 | 32.1397 | 35.8494 |
| DR | 33.6038 | 30.3312 | 36.0649 | DR | 27.3256146 | 32.14917 | 40.5252 | DR | 31.0791 | 33.9732 | 34.9477 |
| LP | 33.5129 | 30.4752 | 36.0119 | LP | 32.1473003 | 30.18378 | 37.6689 | LP | 31.5573 | 32.9339 | 35.5088 |
| TS | 33.5155 | 30.604 | 35.8806 | TS | 34.5746986 | 32.56581 | 32.8595 | TS | 32.3564 | 34.5201018 | 33.1235 |
| *Ppd-D1a* Total | A (%) | B (%) | D (%) | *Ppd-D1a* MADS | A (%) | B (%) | D (%) | *Ppd-D1a* Auxin | A (%) | B (%) | D (%) |
| VG | 33.4028 | 30.4079 | 36.1893 | VG | 25.3882601 | 33.1995 | 41.4122 | VG | 31.1658 | 34.5343637 | 34.2998 |
| DR | 33.6015 | 30.3876 | 36.0109 | DR | 27.417553 | 30.19173 | 42.3907 | DR | 31.6605 | 34.1791467 | 34.1603 |
| LP | 33.4208 | 30.4455 | 36.1337 | LP | 30.4365285 | 33.59336 | 35.9701 | LP | 34.0936 | 32.1500586 | 33.7563 |
| TS | 33.5378 | 30.516 | 35.9462 | TS | 32.9544974 | 32.76896 | 34.2765 | TS | 32.1272 | 33.4663466 | 34.4064 |
| *ppd-1* Total | A (%) | B (%) | D (%) | *ppd-1* MADS | A (%) | B (%) | D (%) | *ppd-1* Auxin | A (%) | B (%) | D (%) |
| VG | 33.0703 | 30.5549 | 36.3748 | VG | 25.2513809 | 32.488 | 42.2606 | VG | 31.0841 | 32.7320325 | 36.1838 |
| DR | 33.339 | 30.3483 | 36.3127 | DR | 28.2045 | 32.0934 | 39.7022 | DR | 29.8169 | 34.1658686 | 36.0172 |
| LP | 33.3769 | 30.6846 | 35.9384 | LP | 31.5943 | 38.0324 | 30.3733 | LP | 30.8823 | 33.4832568 | 35.6345 |
| TS | 33.3341 | 30.6255 | 36.0404 | TS | 34.3736 | 32.357 | 33.2694 | TS | 32.6139 | 34.220559 | 33.1655 |

## 3.5 Discussion

### 3.5.1 Understanding the role of *Ppd-1* in the developing inflorescence

Whilst there is substantial knowledge of genes regulating inflorescence development in model species relatively little is known in wheat (Boden *et al.*, 2015; Dobrovolskaya *et al.*, 2015; Gauley & Boden, 2019). The work shown here aimed to use RNA-sequencing and bioinformatic analysis to improve our understanding of the genes that control inflorescence development by investigating:

- The transcriptional landscape of early inflorescence development in wheat, both in terms of the global transcriptome and key gene families.

- The influence of genetic diversity for *Photoperiopd-1* on the wheat inflorescence transcriptome.

- The implications of polyploidy on gene expression.

This approach provided a detailed insight into gene regulation during this important period of wheat development. It indicates *Ppd-1* is involved in a broader role in reproductive development beyond the photoperiod-dependent flowering pathway, via regulating gene expression in the inflorescence meristem. I propose a function for *Ppd-1* in influencing gene expression beyond its regulation of *FT1* expression in leaves.

There are several pieces of evidence that suggest this function:

- The scale of the gene expression changes in gene families beyond those usually associated with the photoperiod pathway, including Auxin.

- The greatest transcriptional changes are observed in the *ppd-1* null line not the *Ppd-D1a* line, despite the impact of leaf *FT1* expression in the null line being relatively small compared to the insensitive.

- A key checkpoint of gene expression appears to be the terminal spikelet stage, a point beyond the stepwise induction of *FT1*.

The influence of *Ppd-1* may be through other intermediate genes for example, *FT2*. In addition, I cannot rule out the differences in temperature between these field sample points affecting the expression profiles, although clustered expression profiles don't show trends that would suggest this. Ultimately, whilst these trends are not definitive evidence

for a direct role for *Ppd-1* regulating inflorescence development, it provides preliminary data to direct future studies that may investigate this role. In proposing these roles for *Ppd-1*, I acknowledge that the observed differences in gene expression are based on near isogenic lines that express mutant alleles of *Ppd-1* contributed by other genotypes. While extensive backcrossing was performed to remove genomic content from these genotypes, it is possible that some of the observed changes in gene expression are attributable to residual genome content from the donor genotypes. Nonetheless, the differences in observed flowering times and inflorescence development phenotypes of these lines is consistent with the anticipated effect of the introduced alleles, and identified differentially expressed genes (e.g., genes encoding MADS-box transcription factors) are consistent with those detected in lines that contain mutations in *Ppd-D1* (Boden *et al.*, 2015).

## 3.5.2 Key gene expression clusters mark inflorescence development stages

To determine whether *Ppd-1* has a role in inflorescence development, I first examined its expression over the VG, DR, LP and TS stages in wild-type and *Ppd-1* NILs (Figure 3.1). *Ppd-1* is expressed to high levels in the developing inflorescence, indicating a function for *Ppd-1* in gene regulation during spikelet and floret development (Kirby & Appleyard, 1984). Interestingly, these results also indicate interplay between *Ppd-1* homoeologues. Our previous analysis supported a role for inter- homoeologue regulation in leaves (Figure 2.3). Though, in the meristem it appears to be stage dependent. There is lower expression of *Ppd-B1* at the DR stage and higher expression at the LP and TS stages in the *Ppd-D1a* line, relative to wild-type (Figure 3.1). This varied and stage-dependent expression indicates a role in gene regulation outside the interaction of *Ppd-1* with *FT1*, and a potential role in regulating inflorescence development.

To investigate the molecular effect of *Ppd-1* allelism on the transcriptional landscape of inflorescence development, I identified the genes that are most significantly up- and down-regulated between each major meristem stages in *Ppd-D1a* insensitive and *ppd-1* null lines (Figure 3.2). Because of the considerable morphological changes, the inflorescence experiences between the VG stage and emergence, it may be expected that the activation of gene expression would be favoured in a developing inflorescence (Kirby & Appleyard, 1984). However, a similar number of genes are up- and down-regulated, indicating transcriptional changes are as much about releasing repression on a vegetative state as activating a floral one. In the leaf, *ppd-1* null lines exhibit a slight dampening of *FT1* expression, whilst the insensitive *Ppd-D1a* line significantly promotes *FT1* activity (Figure 2.8). With this trend, I expected a consequently large impact on genes that regulate the

meristem in the *Ppd-D1a* line, and only a comparatively minor effect in the *ppd-1* NIL. In many cases, it appears the insensitive allele has little effect on the total number of genes that are up- and down-regulated between inflorescence stages, relative to wild-type (Figure 3.2). This is likely because of a shifting in expression patterns but not amplitude, as observed previously (Figure 2.14). In the *ppd-1* line, in which leaf *FT1* expression is less disrupted, many more genes are differentially regulated, supporting a role for *Ppd-1* that is independent of *FT1*. Ultimately, there are a large number of genes differentially expressed throughout inflorescence development, illustrating how little we know about the underlying genetic pathways.

There are four key early inflorescence development stages in wheat, during which key yield components such as spikelet number and floret initiation are decided – the vegetative (VG), double ridge (DR), lemma primorida (LP) and terminal spikelet (TS) (Kirby & Appleyard, 1984; Zadoxs *et al.*, 1974).  To investigate if there are characteristic gene expression profiles that represent these key stages, I identified clusters that show stage-specific transcript peaks (Figure 3.3). These gene expression patterns have several key GO terms represented in each of the clusters (Table 3.1). I identified the GO terms of transferase activity and catalytic activity for the VG stage, while the DR cluster contains cyclic compound binding and nucleic acid binding GO terms. The LP cluster highlights GO terms associated with protein heterodimerization activity and amino acid binding, and at the TS stage, the terms of cyclic compound binding and nucleic acid binding are enriched. These GO terms indicate that the expression of genes involved in DNA and protein regulation are modulated throughout early inflorescence development.

The expression of these stage specific gene clusters are vastly disrupted in the *Ppd-D1a* and *ppd-1* lines (Figure 3.3). Scrutinizing the changing expression patterns, I see examples of earlier expression of genes in the *Ppd-D1a* line as hypothesised, likely because of the earlier induction of *FT1* in this genotype (Figure 3.3c). However, I also observe major differences during the lemma primordia stage of development (Figure 3.3b). In the *ppd-1* line, there is an inverse trend of gene expression compared to wild-type, with up-regulation of genes in the *ppd-1* NIL where there is down-regulation in wild-type, especially between the lemma primordia to terminal spikelet stages (Figure 3.3a,d). These patterns indicate that in addition to *Ppd-1's* role as a positive regulator of gene expression through the induction of *FT1*, it has a negative role on gene expression between the LP to TS stages (Shaw *et* al., 2012; Boden *et al.*, 2015). This regulation between the LP and TS transition is essential for spikelet formation. The wheat spike is determinate, meaning that when the terminal spikelet forms, spikelet number is set (Kirby & Appleyard, 1984; Sakuma & Schnurbusch, 2020). The effect of *ppd-1* null lines on the

gene expression during this transition suggests *Ppd-1* is required for normal transition to TS, and without *Ppd-1* the transition is delayed, allowing more spikelets to form.

The representative cluster for LP shows considerable disruption in the *Ppd-D1a* and *ppd-1* lines (Figure 3.3c). To make sense of this disruption, a subsequent clustering analysis was performed on the disordered gene expression patterns. In the *ppd-1* line, I identify clusters that appear to have lost all stage specific regulation after the VG to DR transition, suggesting *Ppd-1* is required to maintain these profiles. In the *Ppd-D1a* line, genes that peaked at LP in wild-type were found to peak at the earlier DR stage (Figure 4a-b,d). Of particular interest was the sub-cluster showing an apparent shift in whole gene expression profile of wild-type, with a peak at the DR and subsequent down-regulation to the LP stage (Figure 4d). This cluster is particularly interesting because it is not just mis-regulation of expression, but an entire frame shift of the profiles. Genes shifted in this manner are being expressed as normal, but at the incorrect developmental time. Profiles such as these have interesting implications for floral organ specification, if for example, genes that are involved in promoting floral organ development at the LP stage are expressed at the DR stage, it could result in malformed structures, or accelerated development through the DR stage (Boden *et al.*, 2015; Dixon *et al.*, 2018; Poursarebani *et al.*, 2015; A. Yoshida *et al.*, 2013).

The genes that shifted in profile in the *Ppd-D1a* NIL, relative to wild-type, are likely to be caused by earlier induction of *FT1* in these lines (Figure 3.4d; Boden *et al.*, 2015; Dixon *et al.*, 2018). To investigate these genes further, the TPM expression profiles of genes represented in this sub-cluster were plotted (Figure 3.5 and Table 3.2). Some key examples of genes with promising profiles are TraesCS1A02G096100.1 (Mediator of RNA polymerase II transcription subunit 28), which is orthologous to the *Arabidopsis* and rice *MEDEATOR 28* (*MED28*). MED28 functions as a bridge between RNA polymerase II and promoter-bound transcriptional regulators, and have shown to be involved in senescence and root development (Shaikhali *et al.*, 2016). In addition, TraesCS5A02G325600.1 (Protein XAP5 CIRCADIAN TIMEKEEPER) is homologous to *XAP5* in Arabidopsis, which influences the circadian clock and photomorphogenesis. *XAP5* inhibits red light input to the circadian clock, with the authors suggesting the gene functions as a light quality integrator that senses blue and red light to regulate plant growth (Martin-Tryon & Harmer, 2008). TraesCS7B02G017600.1 (*Histone H3*) and TraesCSU02G092100.1 (*Histone H3*) both encode proteins orthologous to Histone H3, implicating a role for the epigenetic control of gene expression or cell-cycle processes (Parvathaneni *et al.*, 2020). Together, these genes represent promising targets of study to understand the shift of expression patterns caused by the *Ppd-D1a* insensitive allele in wheat.

### 3.5.3 Cluster relationships between stages and genotypes

Our current understanding of the events determining spikelet number and yield components in wheat mainly focus on the morphological changes that occur between the vegetative and terminal spikelet stages (Boden *et al.*, 2015; Dobrovolskaya *et al.*, 2015; Gauley & Boden, 2019). To further expand our understanding of the events that occur between the key developmental stages and identify genes that contribute to spikelet and floret development, I performed clustering analysis (Figure 3.6). Clustering analysis of the top 30,000 genes expressed in the inflorescence shows peaks and down-regulation at every stage of development. These patterns suggest there are hundreds of genes with expression focused to each stage. For example, the LP stage has emerged as the most important stage for inflorescence development, acting as the focal point for peaks and down-regulation in expression for over 50% of clustered genes. It is at LP that spikelet number is decided, in addition to the formation of floral structures such as the glumes and lemma (Kirby & Appleyard, 1984). The thousands of genes converging at this stage suggest this is a highly regulated process. These patterns are disrupted in both the *Ppd-1* insensitive and null lines, highlighting the influence of *Ppd-1* over inflorescence development, with the disruption likely contributing to the differences in spikelet number observed between these lines. This checkpoint at the LP stage has large implications for wheat breeding. Any mutant line that has more or less spikelets must involve either a delay or accelerated progression from DR to TS, as it is during this window that spikelet number is determined (Kirby & Appleyard, 1984).

Notably, there is equal clustered expression representing down- and up-regulated genes in all three genotypes (Figure 3.6). This patterning supports the idea that repressing gene expression is as important as activating expression in meristem development, and that *Ppd-1* regulates flowering by positively and negatively regulating development. Research into the progression of the IM at the floral transition generally focuses on activation of gene expression, with the FAC inducing the expression of meristem identity genes (Wigge *et al.*, 2005; Abe *et al.*, 2019). Recently, a function for a florigen repression complex (FRC) has been proposed in rice. It is hypothesised that TFL1-like proteins compete with Hd3a for 14-3-3 binding to form a FRC, preventing the formation of a FAC. The balancing of FRC and FAC is hypothesised to optimise inflorescence development (Kaneko-Suzuki *et al.*, 2018).

In the *Ppd-D1a* insensitive line, the effect of earlier *FT1* expression is evident where the expression profiles are shifted in the same pattern to an earlier stage in development, relative to wild-type. This trend supports my earlier conclusions that this is the cause of accelerated development through the DR stage (Figure 3.6).

In the *ppd-1* null line there is evidence for mis-regulation between the LP to TS stages, particularly between the WT C8 and *ppd-1* C11 clusters (Figure 3.6). This mis-regulation is unlikely to be because of *FT1* action, as the major checkpoints of *FT1* regulation in wheat are from the VG to DR and DR to LP stages (figure 2.6) Instead, it points to a checkpoint independent of *FT1*, mediated by *Ppd-1*, and likely downstream regulators to time the progression to TS to optimise the production of reproductive structures initiated at this stage.

## 3.5.4 The role of auxin in inflorescence development

Auxin and auxin-related pathways are involved in virtually every plant developmental process (Lee *et al.*, 2019). This role is ancient, dating back to the first diverging land plants (Mutte *et al.*, 2018). There is strong evidence for a key role of auxin in the formation of spikelets in cereals. In maize, auxin maxima form in spikelet primordia to guide lateral meristem development, with palea and lemma forming important auxin sinks (Gallavotti *et al.*, 2008; Xing *et al.*, 2011). In wheat, due to genomic resources only having recently become available, this family is understudied (Adamski *et al.*, 2020). I propose that auxin-related genes have an important role in inflorescence development in part, mediated by *Ppd-1*, indicated by their expression in the developing inflorescence (Figure 3.7-9).

Examining the clustered expression profiles of auxin-related genes in the inflorescence, I identify peaks and down-regulation of expression at every developmental stage. These patterns illustrate the likely importance of these genes in progression through the key meristem stages where yield traits are decided. Relative to the whole genome analysis (Figure 3.6), there is a greater trend of gene cluster maintenance between genotypes. There is particularly strong conservation between wild-type and *Ppd-D1a,* indicating that this family is more tightly controlled relative to the genome as a whole. However, the NILs still have a strong effect on clustered gene expression. For example, when I examine the impact of the *ppd-1* NIL on expression patterns, I see a shift towards clusters that represent a loss of regulation after the VG stage, whereby the stage specific peaks and down-regulation are lost. This indicates that these auxin genes require *Ppd-1* for stage specific expression.

This chapter has identified 20 ARF genes in wheat (Figure 3.8a). The expression patterns observed indicate that there are diverse roles for these genes in the developing wheat inflorescence (Figure 3.9). Many of these ARF genes fit into the clustered expression profiles observed for the auxin-related gene family (Figure 3.7). From these expression patterns, I can infer what developmental processes these genes may be involved in.

In Arabidopsis, *ARF1* and *ARF2* both perform roles in the control of ageing, the initiation of flowering and floral organ abscission (Ellis *et al.*, 2005). When *ARF2* is deleted, flowering is delayed by five days, with *ARF1* enhancing the function of *ARF2* (Ellis *et al.*, 2005). In wheat inflorescences, there is only one class III (*AtARF1/2*-like) gene expressed, namely *ARF25* (Figure 3.9). *ARF25* clusters into the C1 cluster of wild-type, the C1 of *Ppd-D1a* and C0 of the *ppd-1* null line (Figure 3.7). The C1 cluster is characterised by significant down-regulation of expression at the LP stage, indicating that *ARF25* may be involve in the repression of floral structures initiated at this stage, including glumes and lemma (Kirby & Appleyard, 1984). The clustered expression of *ARF25* in the *ppd-1* null line shows a peak at the DR stage, indicating its role is repressive in wheat development, with upregulation earlier in the *ppd-1* null line possibly delaying progression from DR to LP.

Among the *ARF* transcription factors of Arabidopsis, *ARF5* is arguably the most important in the auxin-mediated control of floral meristem development. Also known as *MONOPTEROS* (*MP*), *ARF5* modulates auxin signalling (Hardtke & Berleth, 1998; Krogan *et al.*, 2016; Lee *et al.*, 2019). ARF5 dimerises with AUXIN/INDOLE-3-ACETIC ACID (Aux/IAA) repressor proteins in the absence of auxin. When auxin is sensed by Aux/IAA proteins, they are degraded. As a result, ARF5 is released to induce the expression of many auxin-regulated genes. It is known to repress *ARR7/ARR15* and activates *AHP6.* In Arabidopsis, *ARR7* and *ARR15* coordinate cytokinin and auxin signalling with the WUS-CLV network (Lee *et al.*, 2019). In wheat, I have identified an orthologue of *AtARF5* as *ARF4.* In this study, I identify clustering of *ARF-A4* into the C0 cluster for both wild-type and *Ppd-D1a*. C0 shows a peak in expression at the DR stage of development, a stage that is particularly important for spikelet meristem development. Whereas, in the *ppd-1* null line, it is transferred to the C2 cluster. The C2 cluster is characterised by a decrease in expression after the VG stage, with an apparent loss of stage specific expression patterns. This mis-regulation of an important regulator of auxin signalling in the developing inflorescence at a time when spikelet meristems are initiated, could contribute to the greater spikelets produced by *ppd-1* null lines.

In tomatoes, silencing of *ARF6* and *ARF8* through higher expression of *miR167a* delayed inflorescence development. The down-regulation resulted in female sterile plants, with a reduced leaf size and stem internode length (Liu *et al.*, 2014). Therefore, I theorised that *ARF6* and *ARF8* orthologues would be promising candidates for analysis in wheat. For example, *ARF9* is a *AtARF6/8-like* gene in wheat. *ARF9* is represented in the C8 cluster in WT, this cluster shows a gradual increase in expression throughout inflorescence development, with an eventual peak in expression at the TS stage. These results suggest a role in progressing inflorescence development, like that observed in tomatoes; however,

here I observe a shift to the C7 cluster in the *Ppd-D1a* lines (Figure 3.7). This cluster has a similar trend of increasing expression as the inflorescence develops, but peaks in expression earlier at the LP stage. This earlier peak may contribute to the accelerated development of the *Ppd-D1a* NILs through these stages, which is essential for spikelet formation (Kirby & Appleyard, 1984). *ARF25*, *ARF4* and *ARF9* represent examples where important auxin response genes that peak at different stages of development have expression profiles altered by *Ppd-1* allelism.

It is notable that there are several genes that are not expressed in maize, but are expressed in wheat, including *ARF11, ARF23, ARF12, ARF25*, *ARF6, ARF15, ARF17* and *ARF19* (Galli *et al.*, 2015). Their expression points to a unique role for these genes in the regulation of inflorescence development in wheat.

Taken together, these data paint an intricate picture for the involvement of *ARF* genes during inflorescence development of wheat. There is not one homogenous expression pattern representative of the *ARF* genes, although, there are peaks at different stages with varying impacts of *Ppd-1.* These peaks support the idea of checkpoints for regulation in wheat inflorescence development, with auxin-mediated gene regulation required to progress through key stages. There are non-consistent effects of *Ppd-1* on *ARF* expression, with *ARF* genes being both positively and negatively regulated. Conversely, the expression of several genes, including *ARF2*, are unaffected by *Ppd-1* allelism. This points towards a system where *Ppd-1* regulates the expression of specific *ARF* genes, both positively and negatively, in a stage-specific manner to time the formation of key development structures, such as spikelets.

## 3.5.5 ABCDE pathway for floral development

The ABCDE pathway for floral development has been well described in *Arabidopsis*, and many of the genes show conservation in cereals; however, in crops there is evidence that the pathway has diverged due to events such as gene duplications (Coen & Meyerowitz, 1991; Schilling *et al.*, 2018, 2020). Using the established model from Arabidopsis, I can predict the expression patterns of homologues in wheat, based on the floral structures that form at each of the inflorescence meristem stages (Figure 3.20).

 According to the ABCDE model, A class genes should peak at early stages of floral development to specify sepals. In wheat, *AP1-1* (*VRN1*) has a well characterised role at the floral transition, and it is highly expressed throughout development, which is surprising as *VRN1* is thought to be induced at the floral transition (Wigge *et al.*, 2005; Abe *et al.*, 2019). Several other *AP1* genes have been characterised in rice. *OsMADS15* is the closest rice orthologue of the wheat *AP1-3* (Schilling *et al.*, 2020). Loss of *OsMADS15*

function perturbs palea development, the equivalent structure to sepals in *Arabidopsis* (Jeon *et al.*, 2000; Yoshida & Nagato, 2011). In wheat, the palea do not form until after the floret primordium stage. Therefore, I would expect a peak of expression at the TS stage (Kirby & Appleyard, 1984). The peak in *AP1-3* at the TS stage indicates a role in palea expression, consistent with rice, whereas expression from the DR stage onwards indicates a possible role earlier in development (Figure 3.12-13).

*OsMADS18* is the closest orthologue to *AP1-2.* In rice, *OsMADS18* overexpression causes earlier flowering and reduced tiller number, while reduced expression delays seed germination and seedling growth (Fornara *et al.*, 2004; Yin *et al.*, 2019). In wheat, I observe an increasingly high expression as the inflorescence develops, correlating with the expected gene expression profiles (Figure 3.12-13). The differences in expression patterns between the *AP1-1, AP1-2* and *AP1-3* suggest diversification of function in wheat for this class, compared to *Arabidopsis*. These differences that will be explored in greater detail in Chapter 4.

The B class family of genes determine petal identity in the second whorl of development in *Arabidopsis*. Whilst both wheat and rice do not produce petals, I would expect an equivalent stage of development to be the early phases of floret production (Guo *et al.*, 2015; Figure 3.20). There are two class B genes in *Arabidopsis, PI* and *AP3*. In rice, *MADS4* and *MADS2* are the only orthologues of *PI*. It is proposed that *MADS4* and *MADS2* were produced through an ancient gene duplication event (Münster *et al.*, 2001). In wheat, as in rice, it appears that the *PI-like* gene was duplicated, with two *PI* genes identified in wheat (Schilling *et al.*, 2020). Interestingly, the expression profiles of both *PI-B1* and *B2* cluster together in the B representative cluster, while *PI-A1* shows a drastic up-regulation at the TS stage falling into the B/C cluster (Figure 3.10). The expression patterns of these genes are generally consistent with each other, with similar profiles for both *PI-1* and *PI-2* genes. The expression appears to indicate that the terminal spikelet stage of development is the most important for *PI* function, which is different for expectations based on the *Arabidopsis* model (Figure 3.20; Jack, Brockman and Meyerowitz, 1992; Goto and Meyerowitz, 1994). Analysis of localised expression in the developing inflorescence may provide further understanding of *PI* gene function in wheat.

The C family of genes specify stamen and carpel identity in *Arabidopsis* (Smaczniak *et al.*, 2012). As these structures do not form in wheat until the TS stage, I would expect these genes to be expressed at the LP and TS stage (Figure 3.20). I see an expected trend in expression for the *AG-A1*, *B1* and *D1* genes, with a general increase in expression as inflorescence development progresses (Figure 3.12,13). *OsMADS66* is the closest orthologue to *AG-1* in wheat, though these genes also show close homology to *OsMADS58* (Yamaguchi *et al.*, 2006). *OsMADS58* is thought to be a consequence of

gene duplication from a progenitor to generate it and *OsMADS3*. Mutants in *OsMADS58* show a more drastic and wide-ranging phenotype compared to *OsMADS3* mutants. Mutants have a severely affected floral determinacy, developing flowers that reiterate lodicules, stamens and carpel floral organs (Dreni *et al.*, 2011). It is therefore likely that these genes have sub-functionalized roles in rice. In wheat, all homoeologues of these genes are up-regulated at the TS stage, correlating with the formation of lodicules. However, floral organs are initiated from the glume primordia stage onwards, as such, the low expression of these genes particularly at the LP stage suggests that the role of these genes may not be so conserved.

The D family of genes in *Arabidopsis* and rice are known to specify ovule identity (Figure 3.20; Favaro *et al.*, 2003; Pinyopich *et al.*, 2003). The primary class D gene in *Arabidopsis* is *STK*. Based on their role in ovule development, I would expect that expression would increase as the inflorescence progresses to the TS stage (Figure 3.20). However, in wheat, *STK-A1* and *D1* show a decrease in expression from DR to TS. In rice, there are two *STK* genes, *OsMADS13* an orthologue of *STK-1* and *OsMADS21* an orthologue of *STK-2* (Dreni *et al.*, 2007). Interestingly, although these genes are paralogues, they have different functions. *OsMADS13* is required for specification of ovule identity, but not *OsMADS21*. Rice *mads13* mutants show ectopic formation of carpels without ovules (Dreni *et al.*, 2007). *MADS21* on the other hand, whilst having homology to class D genes, is expressed similarly to class C genes with expression not restricted to the ovule though having no effect on floral identity (Dreni *et al.*, 2011). My analysis showed that *STK- A1*, *D1* and *B1* all cluster into a D pattern of expression (Figure 3.12,13). This pattern exhibits a peak in expression at the DR stage and subsequent downregulation to the TS stage. This pattern implies a role for these genes in early inflorescence development, possibly during the induction of spikelet meristems and not ovule development like in *Arabidopsis* and rice.

The E class family are known to help specify sepals, petals, stamens and carpels with expression throughout floral development (Figure 3.20; Pelaz *et al.*, 2000; Ditta *et al.*, 2004). The SEP genes are the primary members of the class E genes. This family in wheat are particularly interesting as they have duplicated multiple times, but not in *Arabidopsis* and rice (Schilling *et al.*, 2020). Arabidopsis and rice contain four and five copies of *SEP* genes, respectively (Münster *et al.*, 2001; Smaczniak *et al.*, 2012). Whereas in wheat, there are nine *SEP-like* genes that each have three homoeologues. Of these, a total of 18 *SEP* homoeologues are expressed during the early stages of wheat inflorescence development (Schilling *et al.*, 2020). *SEP1-1*, *2* and *3* are orthologous to *OsMADS1, SEP1- 4* and *5* are orthologous to *OsMADS5* and *SEP1-*6 is orthologous to *OsMADS34* (Figure 3.11). *OsMADS1* in rice is also known as *LEAFY HULL STERILE 1*

(*LHS1*), mutants of which cause transformation of lemma into sterile lemma-like organs (Prasad *et al.*, 2001; Rijpkema *et al.*, 2009; Yoshida & Nagato, 2011). As such, I would expect these genes to peak at the LP stage of wheat development. When examining their expression patterns, I observe many of the *SEP1* genes behaving differently with examples of peaks in expression at DR, LP and TS. These different expression patterns suggest a divergence in gene function between paralogues in wheat. This divergence could be similar to that observed between *OsMADS13* and *OsMADS21* in rice, whereby *SEP1-2* functions earlier in development than *SEP1-1* (Dreni *et al.*, 2007).

The highest expressed MADS-box gene in the developing wheat inflorescence is *TaMADS32*. *MADS32* is interesting because it is monocot specific, which has been shown to play an important role in the regulation of the rice floral meristem and subsequent organ identity. In rice, *OsMADS32* is expressed at the early stages of inflorescence and much later at seed development, through negatively regulating the expression of the *YABBY* gene *DROOPING LEAF* (*DL*) (Wang *et al.*, 2015). Here, I identify high expression of *TaMADS32* at all stages of early wheat inflorescence development, with transcripts accumulating as the stages progress.

In summary, these MADS-box genes do generally fall into the same ABCDE model from *Arabidopsis* with the patterns of expression more closely following the rice pathway (Coen & Meyerowitz, 1991; Murai, 2013). I have identified varied expression patterns of several of paralogues. This variance points to a system where genes may have gained new functions, possibly through spatially different expression, such as that observed in *OsMADS13* and *OsMADS21* (Dreni *et al.*, 2007). The increased copy number in modern wheat presents interesting possibilities for tailoring the effects of mutants or overexpression in these genes. Mutations that would otherwise be deleterious in a diploid may have interesting consequences for yield components, such as spikelet and floret number.

Throughout this chapter I have highlighted the dramatic consequences of *Ppd-1* allelism on gene expression in the developing inflorescence. An exception to this trend appears to be the MADS-box genes. When clustered MADS-box profiles were compared between genotypes (Figure 3.10), I noted a general maintenance of the largest clusters. The plotted TPM of these genes showed some notable exceptions to this trend, whereby several genes were downregulated in the *ppd-1* null line relative to wild-type, including the *AG*, *STK* and *AP1* genes. The trends highlighted in Chapter 2 (Figure 2.14) hold true, with the expression of these being consistent when normalised by stage. The dampening of expression in key MADS-box genes such as *AP1* (*VRN1*) is consistent, with *ppd-1 and ft-*B1 mutant alleles decreasing the expression of meristem identity genes (Boden *et al.*, 2015; Dixon., *et al.*, 2018).
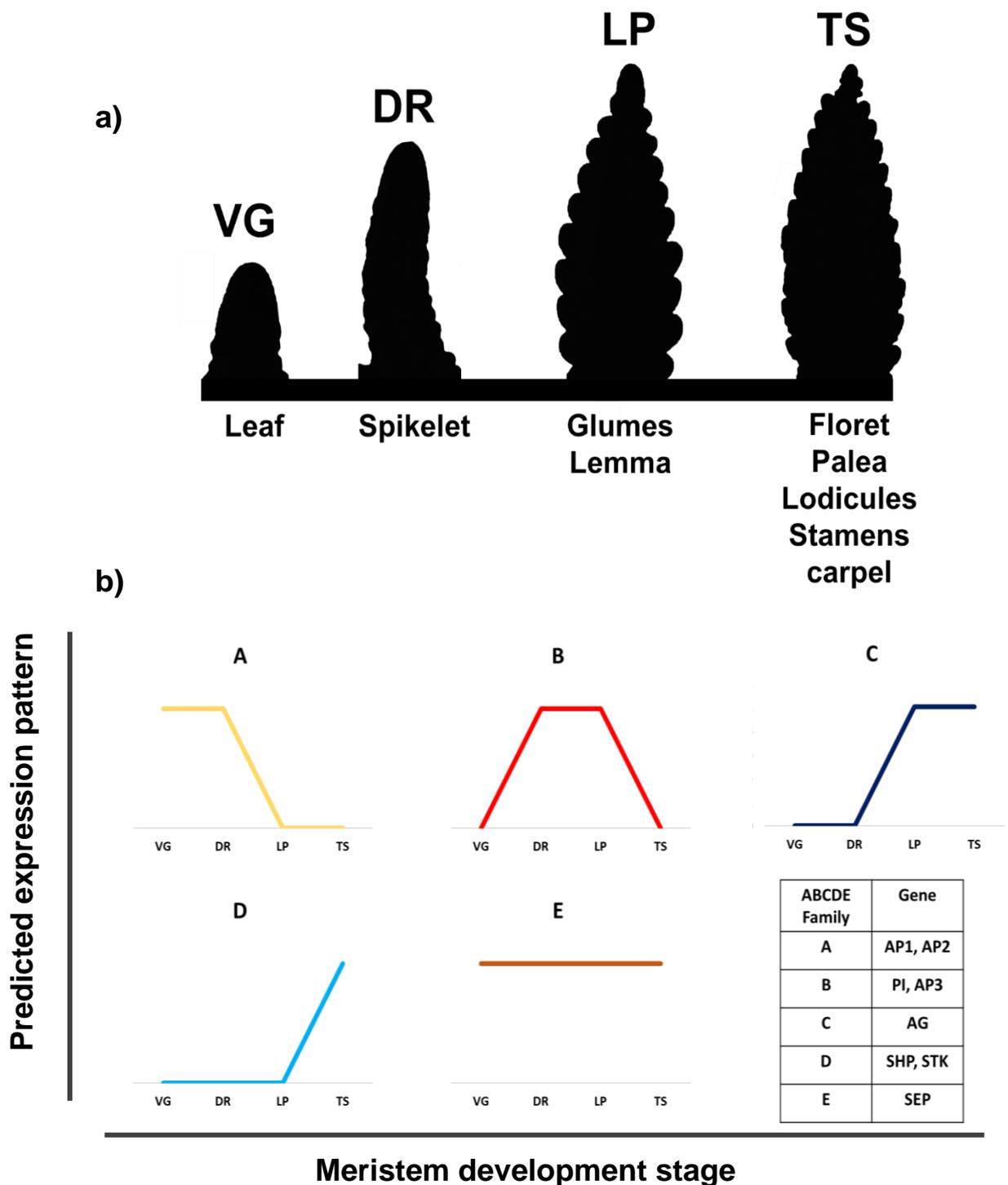
**Figure 3.20 - Predicted expression profiles for the ABCDE flowering pathway genes in wheat**

Based on the reported role of the families of the ABCDE pathway in Arabidopsis we can predict their expression profiles in wheat from where we would expect the corresponding floral organ to develop.  **a)** The floral structures initiated at each of the key early inflorescence development stages. **b)** The predicted expression pattern for each gene class of gene, denoted above profile. A (yellow), B (red), C (dark blue), D (light blue) and E (brown). A table is shown with the corresponding genes for each class.

## 3.5.6 Triplet expression – the three wheat genomes

The key model organisms for plant research, including *Arabidopsis, Brachypodium distachyon* and *Nicotiana benthamiana* have diploid genomes. However, studies predict that between 30 - 80% of flowering plants are polyploid, including bread wheat (Masterson, 1994; Wood *et al.*, 2009). As a result, there is a large gap between our current understanding of effects of polyploidy and its likely importance for gene regulation, and the significance that polyploidy may have for plant development.

This study shows that specific genomes appear more responsive during inflorescence development than others. Overall, the total percentage expression contribution by each genome remains consistent throughout the stages and genomes, with the A genome contributing 33.5%, B genome 30.5% and D genome contributing 36% (Figure 3.14). These contributions reflect a considerable bias towards the D genome. This bias is consistent with the global analysis of 15 tissues, with the D genome on average contributing 33.65% expression, followed by 33.29% for the B and 33.06% for the A genome (Ramírez-González *et al.*, 2018). This study shows the bias towards the D genome is much stronger in the floral meristem compared to other tissues.

Examining triad expression on a gene family level, I see that *Ppd-1* has a preferential effect on genomes in a stage-specific manner. Within the MADS-box gene family, I identify biases of gene regulation dependent on both stage and *Ppd-1* allelism (Figure 3.16-17; Table 3.3). In the wild-type, there is a dominance of the D genome at the VG stage that decreases as the meristem develops, with the A genome gaining influence to increase by 10% contribution from the VG to TS stage (Figure 3.16, 3.17). This trend is shared by the *Ppd-D1a* line. Conversely, the *ppd-1* null line shows shifts in genome bias, particularly at the LP stage, where the genome dominance shifts from the D to B genome. Regarding the auxin-related genes, I see a different trend, the dominance of the D genome is not as strong in wild-type compared to the whole genome or MADS-box genes (Table 3.3). This is particularly evident in the *Ppd-D1a* line, where dominance shifts towards the B genome at the VG and DR stages and the A genome at the LP stage (Figure 3.17). Interestingly, there is no genome bias for gene expression at the terminal spikelet stage, with all genomes showing equal expression (Figure 3.18). The *ppd-1* null line shows a shift away from the A genome towards the B and D genomes. The auxin genes also exhibit stage-dependent regulation via *Ppd-1.* For example, in the *Ppd-D1a* line at VG there is significant up-regulation of auxin genes relative to wild-type or *ppd-1*; however, at the TS stage, expression is highest in the *ppd-1* null line (Figure 3.18). This provides evidence that the role of *Ppd-1* on the auxin family of genes is both positive and negative, with genes capable of being up-regulated in both the *Ppd-D1a* and *ppd-1* lines,

supporting the trends, I observed when examining the expression of the *ARF* genes (Figure 3.7-9).

Based on the results of this study, I conclude that genome-specific gene expression is affected by *Ppd-1*. In general, specific gene families, including MADS-box and auxin-related genes, respond differently to the whole genome trends and other families. This points to a system whereby the regulatory elements for genome bias are consistent within a family. While the mechanism for genome-specific regulation is unknown, chromatin accessibility has been proposed to play a major role (Jordan *et al.*, 2020). It is hypothesised that chromatin accessibility of certain regions of the wheat genome makes genes that are located together in inaccessible chromatin regions less accessible by transcription factors. Gene regulation is known to be greatly influenced by chromatin states, which impacts transcription, replication, recombination and DNA repair (Liu *et al.*, 2013; Gutierrez & Puchta, 2015; Nair *et al.*, 2017; Parvathaneni *et al.*, 2020). Based on the results presented in this chapter, it is likely that not only is this effect based on tissue-specific characteristics, but also temporally within the same tissue as it develops. Crucially, it also varies based on gene family with impact on percentage genome expression changing dynamically.

Here, I identified a pattern of *biased fractionation*, whereby gene copies were lost during or after hybridisation. Studies in polyploid plants have shown that this process isn't completely random, with kinases and transcription factors less likely to be lost (Blanc & Wolfe, 2004; Harper *et al.*, 2016). Genome hybridization can have many effects on gene expression, there is not one established rule for the effect it has on gene regulation. There is evidence for neofunctionalization, where a gene copy will gain a new beneficial function (Hughes, 1994), sub-functionalization, which is when gene copies that are multifunctional divide their functional workload (Force *et al.*, 1999). There are also examples of gene silencing whereby a gene copy loses functionality. There is evidence for certain parts of the genome having gene dominance, a preference for a genome copy expression (Harper *et al.*, 2016). In developing grains, studies have identified cell type and stage specific expression of homoeologues identifying genomic asymmetry and preferential expression of the B and D genomes (Pfeifer *et al.*, 2014). The contribution of genomes to gene expression and regulation has important implications for future plant breeding, allowing gene manipulation methods to be focused on the most important genome copy.

In summary, this study provides a detailed analysis of the transcriptome landscape of wheat inflorescence development, showing distinct gene expression profiles associated with each major stage. In addition, *Ppd-1* has a major impact on the genes expressed as spikelets and florets form, influencing transcript profiles both positively and negatively in three major ways. Firstly, through the induction of genes at the vegetative stage, which

occurs through the induction of *FT1* expression in the leaf, with insensitive alleles of *Ppd-1* showing an induction or repression of genes at the VG to DR transition that in wild-type are induced later. The second involves maintenance of stage specific gene expression, in both the insensitive and null lines. The third regards the LP to TS transition, which shows evidence of both positive and negative regulation mediated by *Ppd-1*. These trends hold true when I examine the profiles and expression levels of both the MADS-box and auxin gene families. To provide a comprehensive analysis I combine observations of gene expression levels and profiles in addition to genome-specific effects. Revealing that gene expression in hexaploid bread wheat undergoes intricate and precise regulation in a stage-specific manner with intergenomic relationships providing a level of fine-tuning, that until now has been undocumented. This new knowledge can help shape gene discovery strategies for the future. There are two major environmental factors that influence inflorescence development in wheat, photoperiod and temperature. This data presented in this chapter have identified significant roles for photoperiod through *Ppd-1* and *FT1,* chapter 4 will explore the role of temperature.

## 3.6 **Methods**

### 3.6.1 Plant material, growth conditions and sampling of *Ppd-1* lines

The plant material and methods for sampling used in this chapter have been characterised previously in Chapter 2.4. In summary, this chapter uses wild-type, photoperiod insensitive *Ppd-D1a* and *ppd-1* null NILs, grown under field conditions. Inflorescences were sampled at the vegetative (VG), double ride (DR), lemma primordium (LP) and terminal spikelet (TS) stages. For RNA-sequencing 5-15 inflorescences were sampled for each biological replicate, dependent on stage.

### 3.6.2 RNA extraction and sequencing

RNA extractions from developing inflorescences were performed using the RNeasy Plant Mini Kit (Qiagen) according to the manufacturer's instructions. A DNase purification was carried out using a TURBO DNA-*free*™ Kit (Invitrogen) according to the manufacturer's instruction. The purity of RNA was verified using a NanoPhotometer spectrophotometer (IMPLEN, CA, USA) RNA QC. Library construction and sequencing were carried out by Novogene (Novogene HK Company Ltd., Hong Kong) on an Illumina NovaSeq 6000. RNA was prepared according to Novogene recommendations.

### 3.6.3 Read alignment and expression analysis

The read alignment and differential expression analysis was carried out in accordance to previously published methods (Brinton *et al.*, 2018). Reads were aligned IWGSC Chinese Spring gene model index v1.1 (Appels *et al.*, 2018). Read alignment in addition to expression quantification were completed using kallisto-0.42.3 (Bray *et al.*, 2016) with default parameters, 30 bootstraps (-b 30) and the ‑pseudobam option as used previously (Borrill *et al.*,2016).

#### 3.6.3.1   Genome specific read alignment

Kallisto default parameters have previously been shown to accurately map homoeolog-specific reads to the Chinese spring RefSeqv1.0+UTR reference (Borrill *et al.*, 2016). This method was further verified by (Ramírez-González *et al.*, 2018) who examined expression of high confidence genes expressed above 0.5 TPM in nulli-terasomic wheat lines lacking an entire chromosome on 1A, 1B or 1D, replaced with a duplicated homoeologous chromosome (Leach *et al.*, 2014; Ramírez-González *et al.*, 2018).

#### 3.6.3.2   Read alignment

To identify spice isoform expression changes between conditions, Sleuth was used. Sleuth uses transcript abundance estimates outputted from Kallisto using pseudo-alignment algorithms to carry out differential expression analysis of gene spice variants. Due to the nature of aligning reads to gene isoforms there is likely to be technical variability in the abundance estimates from alignment, technical replicates are essential to overcome this issue. Sleuth further overcomes this issue using bootstrapping to calculate the estimated abundance for genes using a sub-sample of reads during multiple rounds of bootstrapping. This process allows the estimation of technical variance for each of the genes. Sleuth then combines this technical variance with observed abundance estimates to provide the biological variance used to analyse the differential expression of transcripts (Pimentel *et al.*, 2017; Yi *et al.*, 2018).

Spice variants were identified using the annotations through the Ensembl Plants website (https://plants.ensembl.org/index.html) as of 22/01/20.

#### 3.6.3.3   Differential expression analysis

Differential expression analysis was performed using sleuth-0.28.0 (Pimentel *et al.*, 2017) with default parameters. Transcripts that had a false-discovery rate adjusted $P$-value (q value) < 0.05 and a difference of > 0.5 TPM were considered to be differentially expressed. Transcripts with a mean abundance of < 1 TPM in all four conditions were considered not expressed and were therefore excluded from further analyses. For each

condition, the mean TPM of all three biological replicates was calculated ± SEM. Predicted functional annotation of gene families including auxin, cyclins and histone related genes was carried out using Ensembl Plants Biomarts.

Gene ontology (GO) term enrichment analysis of genes represented between large comparisons was carried out using the R package GOseq v.1.40.0 (Young *et al.*, 2010). For smaller scale analysis of gene subsets the online tool Geneontology was used (Ashburner *et al.*, 2000; Carbon *et al.*, 2019; Mi *et al.*, 2019).

Ribbon plots of TPM values were plotted using the R package ggplot2 (Hadley Wickham, 2016).

### 3.6.4 Clustering analysis

Clustering analysis was carried out using the python package clust (Abu-Jamous & Kelly, 2018). Before clust analysis, gene expression data is normalised to the input data to using a combination of quantile normalisation, log2 and Z-score. The package was run using default parameters.

All auxin, histone and cyclin genes used in clustering analysis were identified based on predicted functional annotation, using Ensembl Plants biomart. MADS-box genes unless otherwise stated were obtained from (Schilling *et al.*, 2020).

### 3.6.5 Phylogenetic analyses

All wheat (*Triticum aestivum*) orthologues of *ARF* genes in maize (*Zea mays*) (Xing *et al.*, 2011) were obtained by BLAST (Basic Local Alignment Search Tool) analysis using Ensembl Plants. The 57 amino acid sequences were aligned using MAFFT v.7.123b alignment algorithm with 4 GUIDANCE v.2.0, 100 bootstrap replicates. An unrooted tree was then generated using MEGAX, with the maximum likelihood method and Dayhoff matrix based model (Kumar *et al.*, 2018; Schwarz & Dayhoff, 1979).

### 3.6.6 Triad analysis

The triad analysis was carried out in accordance with the methods described in (Ramírez-González *et al.*, 2018). Only genes that were represented in triads were analysed, that is genes with a 1:1:1 correspondence across three genomes. A gene triad was deemed to be expressed when total expression was >0.5 TPM. A total of 21,627 gene triads were identified across the wheat genome. 20 MADS-box gene triads and 80 auxin-related gene triads were identified. To standardise expression of all genes, the TPM for each gene was represented as a percentage of total triad expression. The relative triad expressions were then plotted into ternary diagrams using the R package ggtern (Hamilton, 2020).

Significance of whole genome triad expression was carried out using a Kruskal–Wallis test. Average genome specific gene expression was logged to normalise expression, data shows trends in expression.

## 3.6.7 Statistical analysis

Statistical comparisons of gene expression between tissues were carried out using sleuth-0.28.0 (Pimentel *et al.*, 2017) with default parameters. Transcripts that had a false-discovery rate adjusted P-value (q value) < 0.05 were identified as differentially expressed.

# Chapter 4     Effect of ambient temperature on inflorescence development

The results in this chapter concerning temperature differences between the glasshouse and the field and expression of leaf *VRN1* have been published in the following manuscript (Appendix 1):

**Gauley, A., & Boden, S. A. 2020.**. Step-wise increases in *FT1* expression regulate seasonal progression of flowering in wheat (*Triticum aestivum*). *New Phytologist*. https://doi.org/10.1111/NPH.16910

## 4.2 Chapter summary

This chapter investigates the effect of temperature on early stages of the developing inflorescence. By focusing on development after the vegetative to floral transition, I use a detailed phenotype-based approach to gain an understanding of how the spring wheat cultivar Paragon behaves in both field and glasshouse conditions over three growing seasons. This analysis shaped the hypothesis that temperature plays a major role in regulating the key transitions of inflorescence development. To investigate this hypothesis, I performed a detailed examination of key gene families that have a recorded role in temperature mediated development. This study has provided a new understanding of how wheat behaves with regards to both temperature and photoperiod, pointing to an interaction between the two pathways.

## 4.3 Introduction

The sensing of temperature is a major influencer of meristem development with the most documented mechanism being the vernalization pathway (Yan *et al.*, 2003, 2004; Danyluk *et al.*, 2003; Trevaskis *et al.*, 2003, 2006). In the leaf, it releases repression on *FT1*, allowing its expression and subsequent protein translocation to the meristem to trigger flowering (Jaeger *et al.*, 2007; Corbesier *et al.*, 2007; Mathieu *et al.*, 2007; Tamaki *et al.*, 2007). The vernalization process occurs quantitatively, as the plant must experience prolonged periods of cold until the vernalization requirement is satisfied (Hemming *et al.*, 2008). The MADS-box transcription factor *VERNALIZATION1* (*VRN1*) plays a key role as a floral activator in bread wheat, with *VRN1* expression being induced after a period of prolonged cold (Li & Dubcovsky, 2008; Yan *et al.*, 2006). The regulation of *VRN1* in the leaf and meristem helps coordinate flowering to occur under the inductive conditions of spring once temperature and photoperiod conditions have been met. Wheat varieties which have a normally functioning variety of *VRN1* have a winter growth habit.

Conversely, lines that overexpress *VRN1* through a dominant mutation in the A, B or D genomes, or contain a translocation to chromosome 5DS (*VRN-D4*), negate the need for vernalization and flower without a period of cold – these varieties are known as spring wheat (Yan *et al.*, 2003; Danyluk *et al.*, 2003; Trevaskis *et al.*, 2003). The mechanism for *VRN1* regulation is well understood – before vernalization, expression of the repressor *VRN2* is high. A period of prolonged cold epigenetically silences *VRN2*, allowing the induction of the florigen *FT1* (Hemming *et al.*, 2008; Chen & Dubcovsky, 2012). In the SAM, FT1 induces the expression of *VRN1* and other genes that are necessary for floral induction (Moon *et al.*, 2005; Abe *et al.*, 2005; Wigge *et al.*, 2005; Teper-Bamnolker & Samach, 2005; Yoo *et al.*, 2005; Jung *et al.*, 2012; Collani *et al.*, 2019).

Regulation of *VRN1* is an exciting topic of study; it is a focal point for both positive and negative regulation that impacts how much of a vernalization response is required. For example, A glycine-rich RNA-binding protein, GPA2 represses *VRN1* activity by binding to regulatory elements in intron 1 of *VRN1* (Xiao *et al.*, 2014; Kippes *et al.*, 2016). Single nucleotide polymorphisms (SNPs) in exons 4 and 7 are known to affect the required duration of the vernalization response (Eagles *et al.*, 2011; Muterko & Salina, 2018). The SNP in exon 4 causes an amino acid substitution in the conserved k-domain, whereas the SNP in exon 7 reduces VRN1 protein interaction with TaHOX1 (Chen *et al.*, 2009; Eagles *et al.*, 2011; Díaz *et al.*, 2012; Li *et al.*, 2013). In addition, lines carrying an increased copy number of *VRN-A1* have an increased requirement for vernalization, and so flowering is delayed (Díaz *et al.*, 2012). These alleles, alongside SNP variations, present an interesting story for how wheat has been adapted to diverse regions. These examples exhibit how in wheat, which has undergone intense selection pressures, mechanisms for flowering control and inflorescence development can arise uniquely in domesticated wheat, separate even from its wild ancestors.

Post vernalization, there is evidence for regulation of meristem development by temperature. Recently, it has been shown that high temperature treatment after vernalization delay inflorescence development and cause late-flowering in winter wheat (Dixon *et al.*, 2019). The floral repressors *VRN2* and *ODDSOC2* are reactivated before and after vernalization during the high temperature treatment. Whilst the regulation of *ODDSOC2* appears to be independent of photoperiod, *VRN2* regulation is photoperiod dependent (Dubcovsky *et al.*, 2006; Dixon *et al.*, 2019). This provides evidence for an interaction between temperature and photoperiod to regulate winter wheat inflorescence development. There is also significant evidence for an impact of higher ambient temperatures on flowering in barley. High ambient temperatures (28 °C/ 24 °C) delay barley flowering under short days, but accelerate flowering under long days, with evidence pointing towards mediation by *Ppd-H1* in a *VRN1* dependent manner (Hemming *et al.*,

2012; Ejaz & von Korff, 2017). These interactions demonstrate entwinement of the temperature and photoperiod pathways to mediate flowering.

In wheat, the focus of altering temperature responsiveness has concentrated on *VRN1* and *Ppd-1* because of their role in wheat breeding. However, moving forward, it is essential to consider other factors or even independent pathways that have until now been masked by the dominance of *Ppd-1* and *VRN1*. Using newly available resources in wheat, such as advanced transcriptomics and mutant populations, it will become possible to identify these previously hidden factors (Adamski *et al.*, 2020).

The research presented in this chapter stems from a detailed analysis of how the facultative 'spring' wheat Paragon develops under field conditions over three growth seasons. I provide evidence for a role of ambient temperature regulating inflorescence development, not only leading up to the vegetative development, but also during the critical early developmental stages.

## 4.4 **Results**

### 4.4.1 **The floral transition is tightly controlled**

To understand how wheat inflorescence meristems develop within a seasonal context, a detailed phenotyping experiment was performed over three growing seasons. The developing IM was imaged as natural photoperiods extended (Figure 4.1). Within the context of the entire growing season, development is tightly controlled with the DR stage in the 2017 and 2019 growing seasons arriving at the same date (12:15 h) (Figure 4.1b). There was however a notable delay in the 2018 growing season, with the DR stage initiated 8 days later (12:45 h) (Figure 4.1c). Interestingly this delay early in development correlates with a delay in heading date, with heading occurring 6 days later in 2018 (16:36 h), relative to 2017 (16:29 h) and 10 days later compared to 2019 (16:16 h; Figure 4.1c). Publicly available historic temperature data shows this delay correlates with a downward trend in average temperatures before floral initiation (CustomWeather©, 2020). These data illustrate a tightly controlled floral transition in wheat, with it occurring on precisely the same date over two years. However, lower temperatures before the transition are sufficient to cause a delay, correlating with a later heading date.
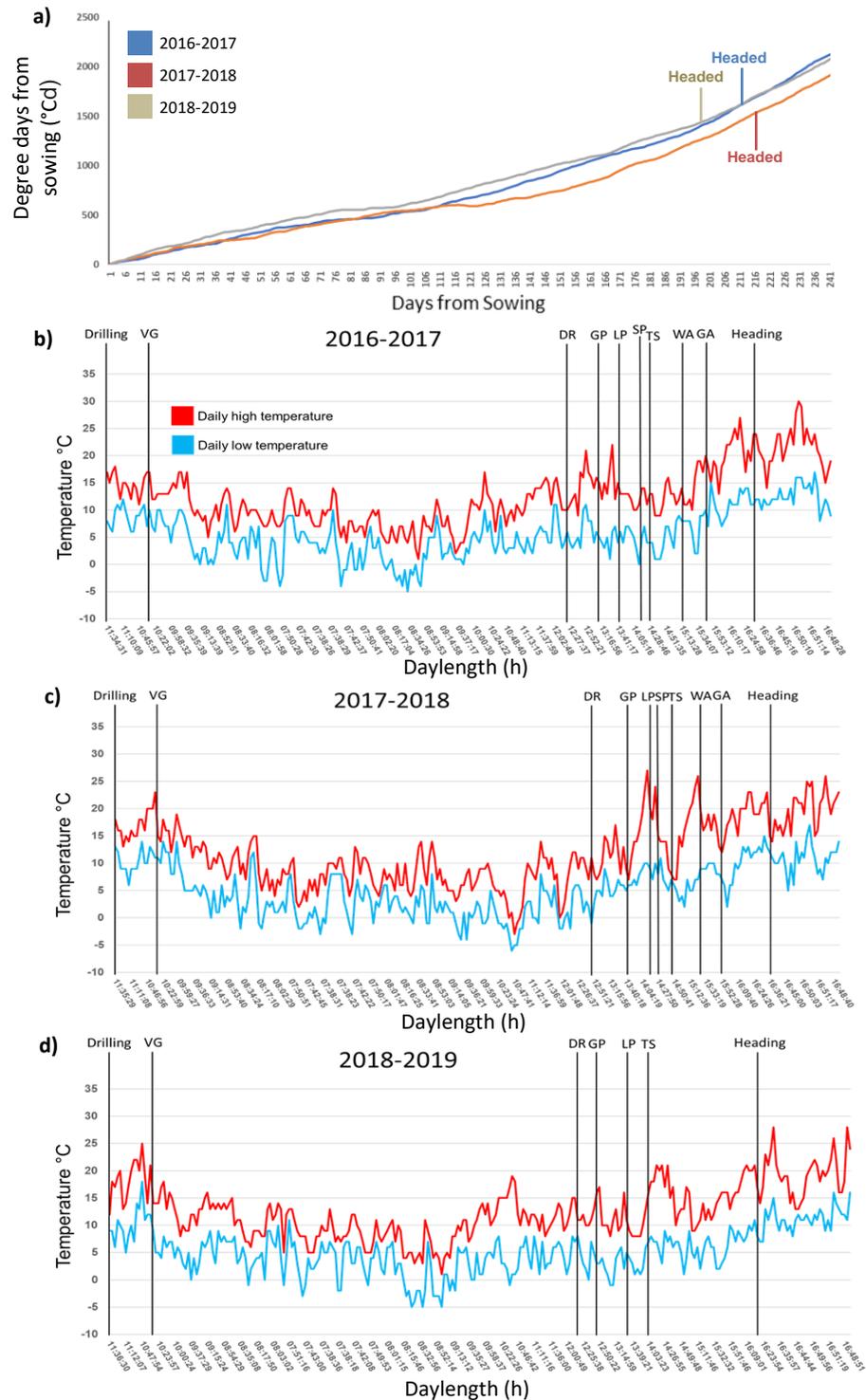
**Figure 4.1 - Seasonal progression of inflorescence development**

**a)** Thermal time plotted against days from sowing for the 2016-2017, 2017-2018, 2018-2019 growing seasons. Thermal time was calculated with a base of 0 °C. Heading date is marked for each growing season. **b-d)** Daily high and low temperature plotted against daylength for the **b)** 2016-2017 growing season **c)** 2017- 2018 **d)** 2018-2019 growing seasons. A growing season is specified as from drilling to heading. The vegetative (VG), double ridge (DR), glume primordium (GP), lemma primordium (LP) and terminal spikelet (TS) stages are marked on each graph and represent when the meristem first reached that stage. The DR stage is marked when the stage is fully formed. All measurements are representative of at least three replicates. High (red) and low (blue) temperature are plotted as lines.

155

## 4.4.2 Lower temperatures correlate with delayed transitions between key meristem stages

The developmental window between DR to TS is short in the context of the growing season; 33, 31 and 28 days for the 2017, 2018 and 2019 seasons respectively. This observation prompted a detailed analysis of the temperature leading up to and during these stage transitions (Figure 4.2). All plants experience the same relative photoperiods; as such, the differences in development are likely due to temperature. During the period from the 9 h daylength until the DR stage, lower temperatures appear to result in delayed progression to DR in the 2018 growing season, relative to the 2017 or 2019 seasons. Importantly, temperatures on average reached below 1 °C every day (maximum temperature of 0.8 °C). Interestingly, when comparing every other transition of the 2017 to 2018 growing seasons, lower average daily minimum temperatures seem to correlate with delayed development (Figure 4.2). Generally, the 2019 season follows this trend, particularly for the GP-LP transition. This data provides evidence for a temperature-based pathway having a considerable role in inflorescence development after the floral transition.
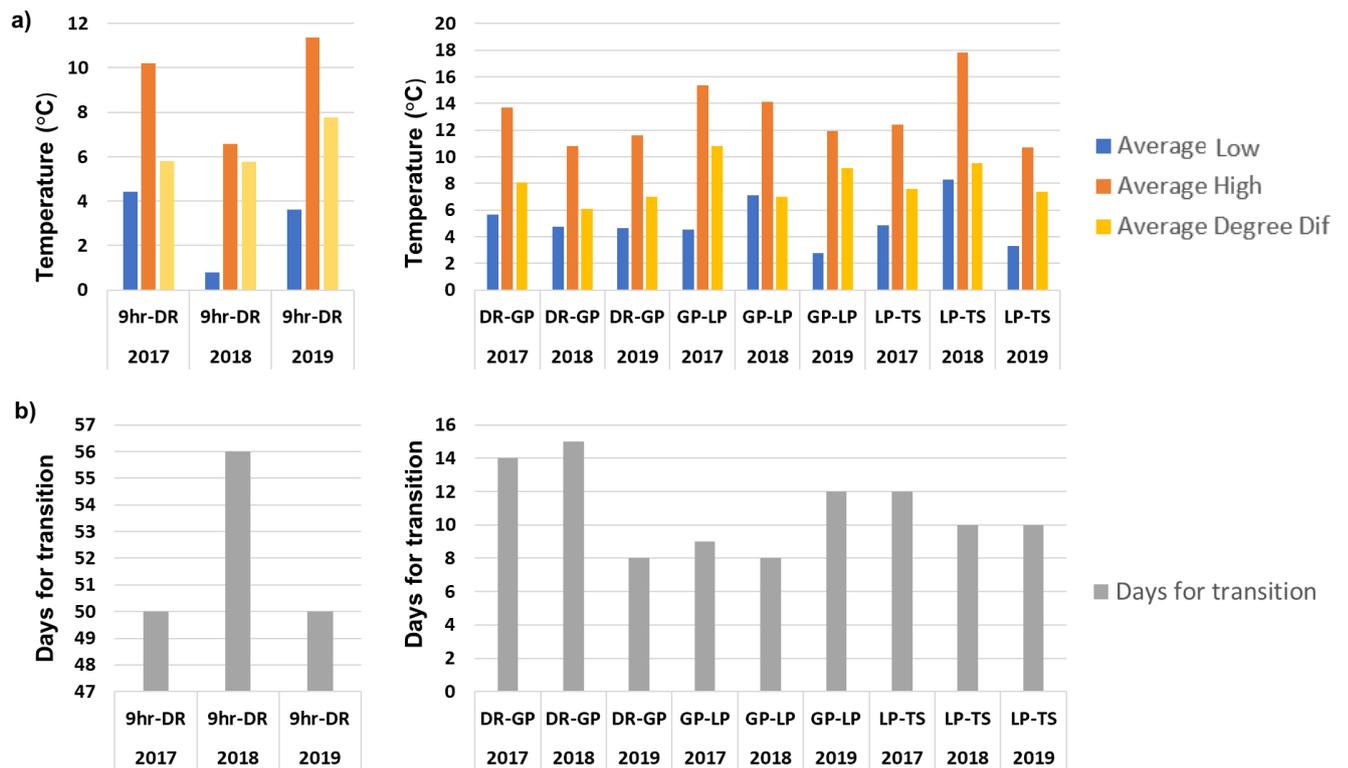
**Figure 4.2 - The relationship between temperature and stage transition rate**

**a)** Average high (Orange), low (blue) and degree difference (yellow) temperatures (°C) throughout each key inflorescence development transition **b)** The calendar days taken for each major transition. The transitions tested are: 9hr photoperiod to DR, DR-GP, GP-LP and LP-TS for the 2017, 2018 and 2019 growing seasons. Each sample point is representative of at least 3 replicates.

### 4.4.3 The glasshouse effect

Because of the temperature-induced delays on floral meristem development in the field, I hypothesised that inflorescence development would accelerate under temperature-controlled conditions. Plants grown under field conditions were compared to GH plants over three growing seasons (Figure 4.3). I observed much earlier and faster development in the GH compared to the field. On average, plants proceeded to DR 21 days earlier in the GH relative to the field (Figure 4.3b-c). The only environmental difference under these conditions is temperature, with the glasshouse being warmer by 1-2 °C (no artificial heating). Differences in environmental temperatures appear to have a heightened effect in the GH, with higher temperatures in the field correlating with accelerated development in the GH (Figure 4.3a). Notably, the 2018 growing season that experienced lower temperatures in February and March has a delayed inflorescence development in both the field and the GH.

**a)**

| Average daily high temperature | | | |
|---|---|---|---|
| | **Jan** | **Feb** | **March** |
| **2019** | 6.3 | 11.8 | 11.61 |
| **2018** | 8.2 | 5.5 | 7.8 |
| **2017** | 6.3 | 8.8 | 12.7 |

**b)**

F 28/3/17   F 28/3/18   F 28/3/19

F 19/4/17   F 17/4/18   F 17/4/19

**c)**

GH 6/3/17   GH 8/3/18   GH 6/3/19

GH 24/3/17   GH 20/3/18   GH 21/3/19

GH 7/4/17   GH 10/4/18   GH 8/4/19
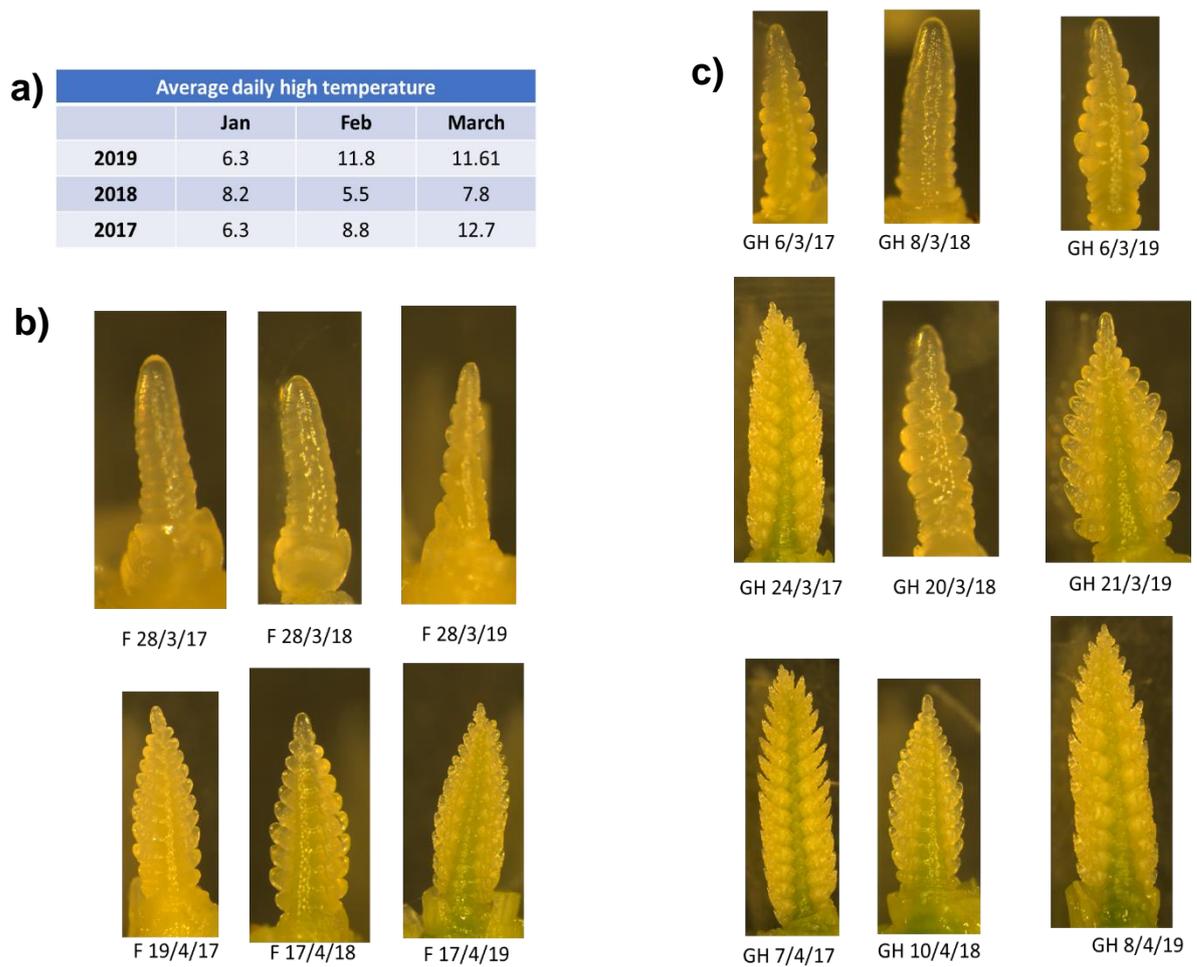
**Figure 4.3 - Comparison of inflorescence phenotypes in the field vs. glasshouse**
**a)** The average daily high temperature (°C) for January, February and March in the field. **b-c)** Representative images of inflorescence meristems in **a)** field (F) and **c)** glasshouse (GH) at the marked dates. Each image is representative of at least three plants.

### 4.4.4 Differences in temperature between the field and the glasshouse

To investigate differences in temperature between the field and glasshouse, the temperature under both conditions was recorded and compared daily at midnight, 08:00 and 16:00 from January to late March. This period captures all key stages of early inflorescence development. On average over the period, the GH is 1.31 °C warmer than the field. However, when the data is analysed segmentally, differences emerge. The midnight sample point is 0.76 °C warmer on average in the GH, and the 16:00 timepoint is 0.3 °C warmer in the GH, relative to the field. The most dramatic difference is at the morning sampling point of 8:00 that is 2.86 °C warmer on average. Importantly, the GH also stopped the temperature dropping below 2 °C. In contrast, the field regularly dropped below 2 °C into sub-zero temperatures during the period leading up to early-mid February. This difference in morning temperature, which correlates with the first peak in diurnal *FT1* expression (Figure 2.7) is likely the cause of accelerated flowering between the GH and field, although further investigation will be required to confirm this.
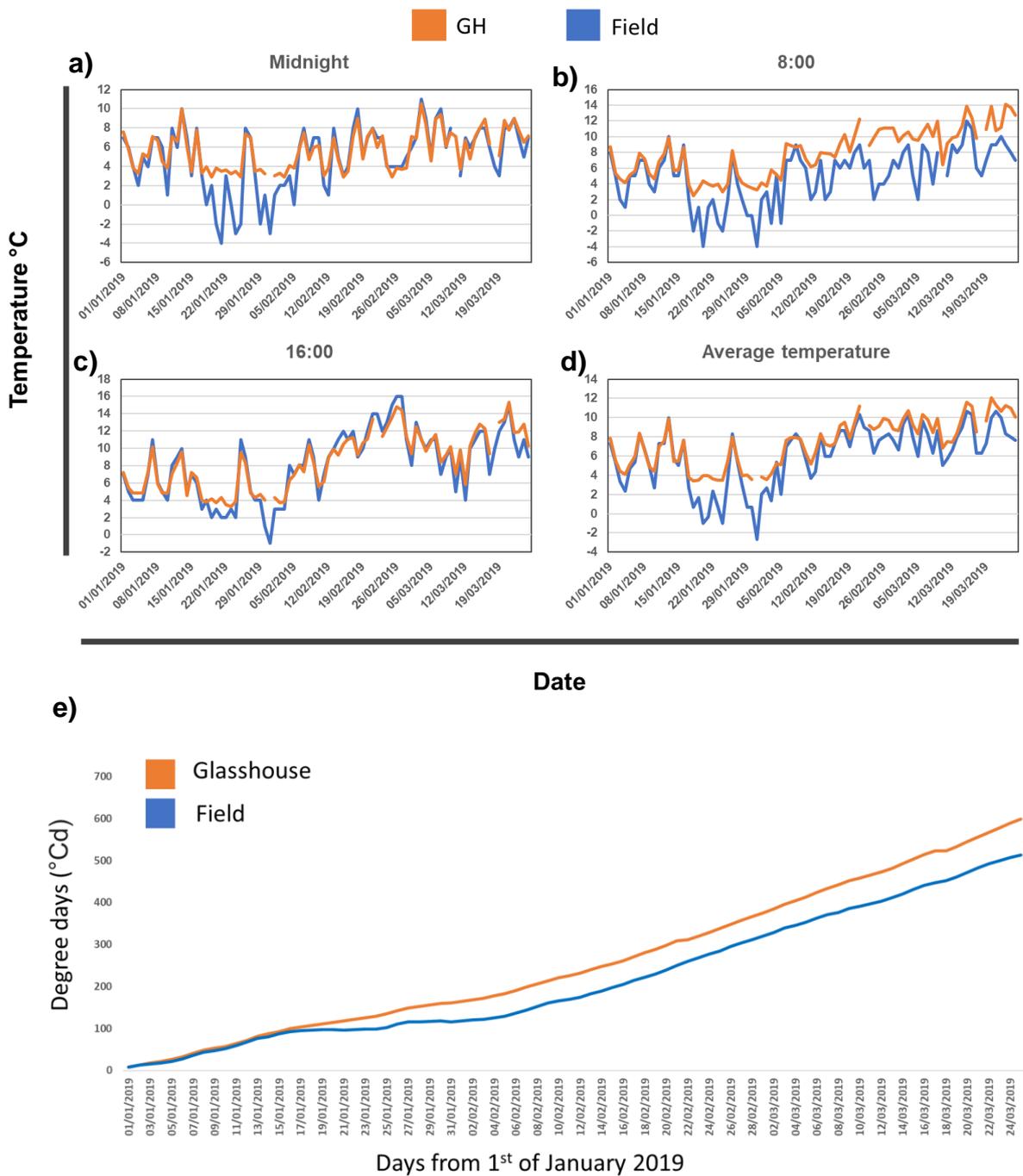
**Figure 4.4 - Glasshouse temperatures relative to the field**

A comparison of daily temperatures (°C) for the glasshouse (GH) compared to the field in the 2019 growing season. The temperature of glasshouse (orange) and field (blue) is plotted against date. Data was collected at 4-time points **a)** Midnight. **b)** 8:00. **c)** 16:00. **d)** Average daily temperature. Figure adapted from (Gauley & Boden, 2020). **e)** Thermal time from the 1/1/2019 in the glasshouse (orange) and the field (blue).

### 4.4.5 Ambient temperature and not thermal time effects meristem development

The data presented in this chapter has pointed to a role for temperature influencing early inflorescence development. However, the effect could be due to either a direct influence of temperature or a product of thermal time. To investigate this, wild-type and *Ppd-D1a* insensitive plants were grown in a GH with unregulated photoperiod and temperature changes. When the plants reached the DR stage at the 10 h photoperiod, the plants were shifted to a GH maintained at 16 °C. After shifting, the comparison is drawn between plants of the same genotype, a population in the unregulated GH and a population maintained at 16 °C (Figure 4.5). Development was accelerated in the plants shifted to 16 °C, relative to GH grown plants, with the wild-type reaching the floret primordium stage when the GH plants were still at the lemma primordium stage. The most substantial difference was observed in the *Ppd-D1a* lines between the two conditions, with the *Ppd-D1a* line reaching a late terminal spikelet stage at 16 °C whilst GH plants were still at the stamen primordia stage. These results indicate that ambient temperature can accelerate floral development post floral transition and the change is compounded by photoperiod insensitivity.

**Figure 4.5 - Effect of a temperature shift on inflorescence development**

Development of wild-type (WT) and *Ppd-D1a* photoperiod insensitive inflorescences in a natural conditions glasshouse (GH) and a glasshouse maintained at 16 °C (16 °C). Sampling occurred when wild-type GH plants reached the lemma primordium stage. Plants are representative of at least 3 inflorescences imaged.

### 4.4.6 Expression of *VRN1* in the leaf

As discussed, *VRN1* is a regulator of flowering. In the leaf, it is a key inducer of *FT1* and is regulated by temperature in winter wheat (Yan *et al.*, 2006; Hemming *et al.*, 2008; Li & Dubcovsky, 2008). In spring wheat, *VRN1* is unregulated and is classically expressed highly. To investigate the role of *VRN1,* its expression was examined in leaf tissue using qPCR from field grown plants. Samples were collected whenever the daylength naturally increased by one hour from the 10 h/ 14 h daylength through to the 13 h/ 11 h daylength. Samples were taken over a 24 hour period to examine its daily rhythms. *VRN1* displays a diurnal expression pattern in the leaf, with a peak during the day and a gradual down-regulation during the night (Figure 4.6). In the *Ppd-D1a* insensitive line, I see a trend of higher *VRN1* expression, relative to wild-type, at the majority of time points, particularly during longer daylengths. During long-days (13 h/11 h) there is a strong trend for *VRN1* to be downregulated in the *ppd-1* null line, relative to wild-type. Conversely, in short days there are some points where expression in the *ppd-1* null line exceeds wild-type expression, particularly at TSD 12 in the 9 h photoperiod. These alternating impacts of daylength and *Ppd-1* suggest a role for Ppd-1 influencing *VRN1* expression.
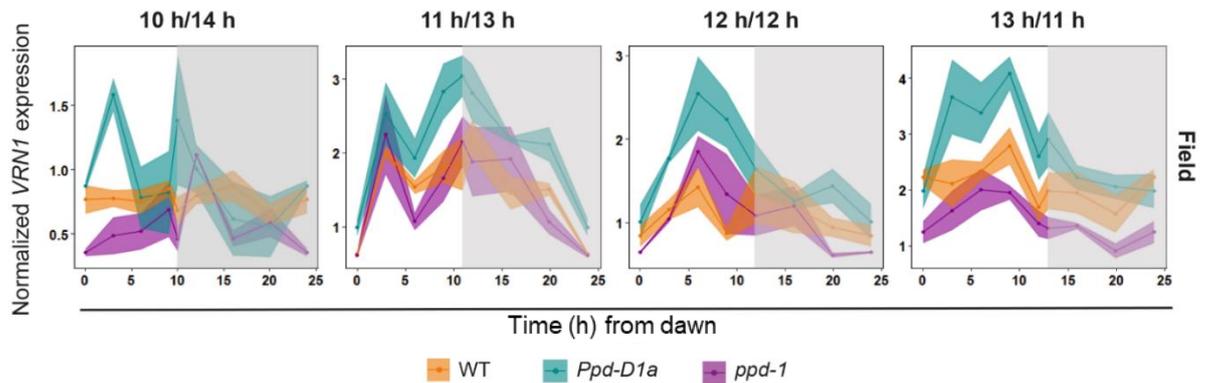
**Figure 4.6 - Leaf expression of *VRN1***

Leaf diel expression profiles of *VRN1* in wild-type (orange), *Ppd-D1a* photoperiod-insensitive (cyan) and null *ppd-1* (magenta) NILs under field conditions. The grey shading highlights night-time data points. All expression profiles are shown over a 24-hour period at hourly incremental increases in daylength, with time represented as time since dawn, with sunrise being 0 h. Graphs are presented as ribbon plots which show the normalised mean transcript levels (solid line with data points) ± SEM (shaded region) of three biological replicates. Figure adapted from (Gauley & Boden, 2020).

## 4.4.7 The expression of the *FLC* genes in wheat

When investigating the role of environmental signals on the floral meristem and possible interaction between the photoperiod and temperature pathways, it is essential to look at the *FLC* and *VRN1* genes due to their documented role in this process (Trevaskis *et al.*, 2007; Alexandre & Hennig, 2008; Fjellheim *et al.*, 2014; Kennedy & Geuten, 2020). The most striking pattern evident in wheat *FLC* expression is that transcript levels are very similar between the genotypes (Figure 4.7). This pattern indicates that, like in the case of other MADS-box genes, the induction of *FLC*-like genes in wheat correlates with developmental stage, not time and the associated temperature differences (Figure 2.14).

There is not one typical expression pattern exhibited by wheat *FLC* genes in the 'Paragon' background (Figure 4.7). The *FLC-A1*, *B1* and *D1* genes all show similar expression profiles. The expression is high at the VG stage with complete down-regulation of all genome copies from the DR stage onwards. Interestingly, *FLC-1* is expressed equally in wild-type and the *Ppd-D1a* line but is not expressed in the *ppd-1 NIL*. The *FLC-A2, B2* and *D2* genes are not expressed in the floral meristem. *FLC-D3* is the only significantly expressed *FLC-3* homoeologue, with consistent high expression throughout meristem development in all three genotypes, and a peak at the TS stage. At the LP stage, there is an increase of *FLC-D3* expression in the *Ppd-D1a* NIL, relative to WT.

There are significant differences between some of the *FLC* genes based on *Ppd-1* allelism. Most notably, the *FLC-D4* transcripts are significantly higher in the *ppd-1* line (14.52 ± 1.05 SEM) and down-regulated in the *Ppd-D1a* line (4.6 ± 0.44 SEM) at the TS stage, relative to wild-type (7.95 ± 0.32 SEM; $p < 0.001$ ). *FLC-4* expression gradually increases throughout inflorescence development, suggesting a role in the later stages of floral meristem development. *FLC-5* is consistently expressed at low levels throughout development. *FLC-6* shows expression in the A and D homoeologues, with *FLC-A6* expressed highest. Expression remains low (< 5 TPM) at the VG and DR stages; however, there is a two-fold increase in expression at the LP stage in the *Ppd-D1a* and *ppd-1* null lines, relative to wild-type. The expression then returns to be below 5 TPM at the TS stage. These expression patterns suggest a unique and stage-specific role for the *FLC* genes in wheat. Future studies may examine if the expression differences between the *Ppd-1* NILs is due to direct influence of Ppd-1, or downstream effectors such as FT, which can influence *FLC* expression (Chen & Penfield, 2018). Ideally, this research will use transgenic lines that over-express *Ppd-1* to negate the possibility that the donor of the *Ppd-D1a* allele contributed variant alleles of *FLC-like* genes to the photoperiod insensitive NIL.
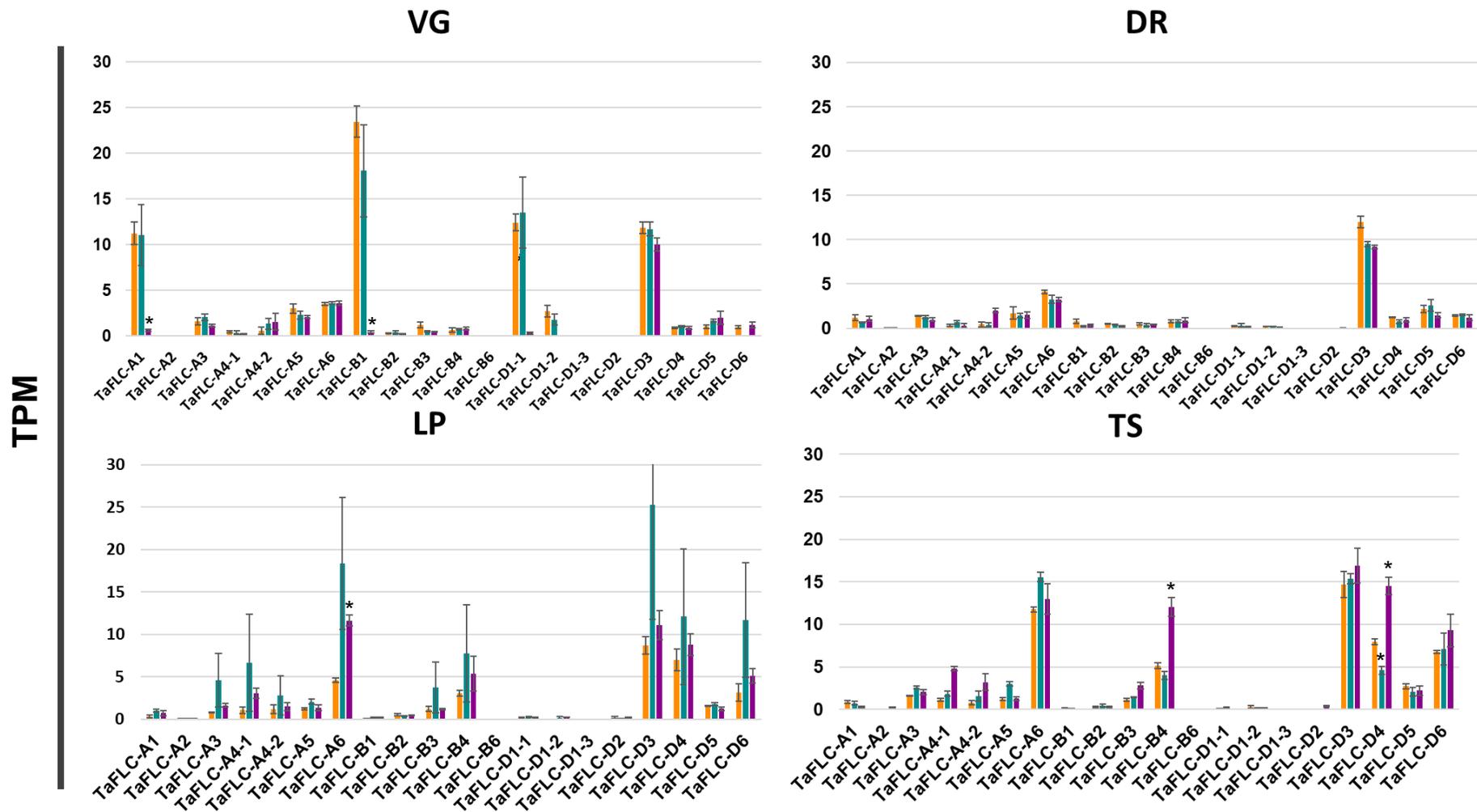
**Figure 4.7 - Expression of FLC genes in the developing inflorescence**

Expression of all *FLC* genes currently identified in wheat from plants grown under field conditions. Data includes expression profiles for wild-type (orange), *Ppd-D1a* photoperiod-insensitive (cyan) and null *ppd-1* (magenta) NILs. The defined stages are vegetative (VG), double ridge (DR), lemma primordium (LP) and terminal spikelet (TS). Data are presented as bar graphs which show the mean TPM ± SEM of three biological replicates. * *P* < 0.001, comparison to wild-type.

### 4.4.8 The expression of the *VRN1* genes during wheat inflorescence development

In wheat, *VRN1* is also known as *AP1* (Schilling *et al.*, 2020). *AP1* has two expressed copies (*AP1-A1* and *AP1-B1*) that are highly expressed throughout inflorescence development. At the DR and TS stages, expression of both genome copies are lower in the *ppd-1* null line, relative to the *Ppd-D1a* line and wild-type. Of the *AP1-2* homoeologues, *AP1-B2* is the only expressed copy with transcripts gradually increasing as the floral meristem develops. All three genome copies of *AP1-3* are expressed, with *AP1-B3* being the highest. Similarly, to *AP1-2,* expression increases as the inflorescence develops, with expression significantly lower in the *ppd-1* null line at TS ($P < 0.001$). Taken together, these *AP1-like* genes are highly expressed and are likely to have diverse roles beyond the vegetative to floral transition.
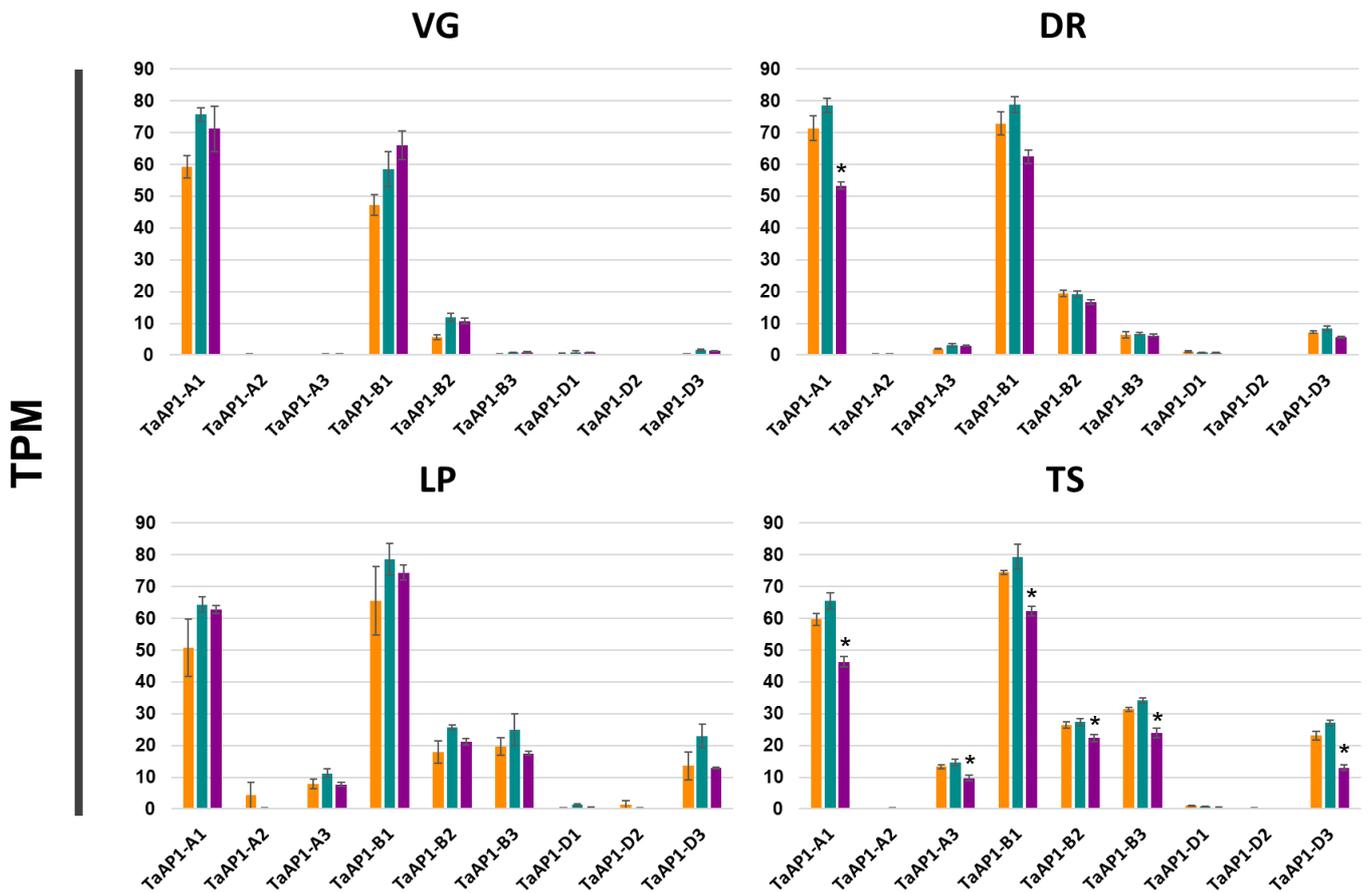
**Figure 4.8 - Expression of *AP1* (*VRN1*) in the inflorescence**

Expression of all *AP1* genes currently identified in wheat from plants grown under field conditions. Data includes expression profiles for wild-type (orange), *Ppd-D1a* photoperiod-insensitive (cyan) and null *ppd-1* (magenta) NILs. The defined stages are vegetative (VG), double ridge (DR), lemma primordium (LP) and terminal spikelet (TS). Data are presented as bar graphs which show the mean TPM ± SEM of three biological replicates. * *P* < 0.001, comparison to wild-type.

## 4.5 **Discussion**

### 4.5.1 Low temperatures delay inflorescence development

The aim of this chapter was to gain insight into how ambient temperature influences the developing inflorescence. Here, I have described precise control over the vegetative to floral transition, with the transition occurring at very similar times for each growing season (Figure 4.1). A detailed analysis of temperature as the inflorescence develops revealed lower temperatures leading up to the vegetative to floral transition appears to be enough to delay development. The temperatures recorded during this phase regularly dropped below 1 °C, likely causing stress sufficient to stall development. There is evidence that the developing inflorescence is particularly resilient to low temperatures when vegetative, as such, the timing of the vegetative to floral transition essential so further development isn't stunted (Limin & Fowler, 2006).

This chapter provides a new resolution of the effect temperature has on these critical early stages of floral meristem development. I provide supporting evidence that lower temperatures delay stage transition rates throughout inflorescence development (Figure 4.2). Our current understanding of ambient temperatures influence on inflorescence defines the optimal temperature range for development as between 17 °C and 23 °C, temperatures beyond this range produce a stress response (Gol *et al.*, 2017). Therefore, studies into the effect of ambient temperature have focused on temperatures above 17 °C. However, by these definitions, wheat progresses through the early stages of inflorescence development under constant temperature stress, as the average high temperature between any stage apart from the LP-TS 2018 transition never exceeded 17 °C in Norwich, UK. These low temperatures do not appear to harm yield, as East Anglia is one of the highest yielding wheat environments on earth with an average yield of 9.9t/ha in 2019, higher than average yield in any other country in Europe (Hannah Ritchie and Max Roser, 2013; Strutt & Parker, 2019). These data indicate that future studies to investigate the effects of ambient temperature on field-grown wheat inflorescences should focus on average temperatures between 4 °C and 17 °C. Previous studies have looked at the effects of temperature on the DR to TS stage, also known as the spikelet ignition phase (Slafer & Rawson, 1994). However, whilst these studies showed small differences in development rate mediated by temperatures, they focused on the 16/9 °C, 23/16 °C and 30/23 °C day/night temperature ranges, concluding the effect on development rates through this phase by temperature were slight. For climates comparable to the UK and northern Europe, another lower temperature testing point may provide interesting results. In the future, when considering the role of temperature under these natural field conditions it will be important to consider if the effects on inflorescence development are due to

optimising temperature based pathways, or due to optimising temperature based stress responses.

Further evidence for the role of ambient temperature influencing inflorescence development is evident by the accelerated development imparted by growing plants in a glasshouse compared to the field. I show differences in the morning temperature correlate with the initial peak in morning *FT1* expression described previously (Figure 4.3; Figure 2.7). These data indicate that it is the early morning time point which is most important in delaying or accelerating flowering. The temperature variations between these two conditions reinforce that an essential temperature range for ambient temperature investigation is between 4 °C and 16 °C. Temperature variation mediating differences in morning *FT* expression has also been recorded in *Arabidopsis* (Song *et al.*, 2018).

To confirm that differences in ambient temperature could influence inflorescence development after the floral transition, I conducted a temperature shift experiment. This shift from low temperatures to 16 °C was sufficient to accelerate development from DR to LP. Surprisingly, the accelerated effects of the temperature shift were much more dramatic in the *Ppd-D1a* line compared to *Ppd-1* wild-type lines. This is likely due to the photoperiod insensitivity of these lines, whereby removing the photoperiod checks on flowering allow the effect of increasing temperature on IM development rates to be more pronounced. This interplay between ambient temperature and photoperiod has been investigated by Hemming *et al.* (2012), who report that a temperature increase of 15 °C to 25 °C was sufficient to delay development under short days and accelerate it under long days in wild-type wheat. Interestingly, *Ppd-D1a* lines were also tested, under short days *Ppd-D1a* NILs show accelerated development at 15 °C, but a decreased rate of development at 25 °C. These findings complement the experiment performed here where 16 °C was sufficient to accelerate *Ppd-D1a* flowering compared to lower temperatures, indicating that these low 15-16 °C temperatures are optimal for accelerated development. These results point to a system whereby variations in temperature as the plant develops are critical to determine the speed at which it progresses through reproductive development. This is particularly pertinent as it is during this short 30-day growing period that essential yield components such as spikelet number and floret initiation are decided (Kirby & Appleyard, 1984). It is also during this period when key yield components of the *Ppd-D1a* insensitive and *ppd-1* null lines are decided, as reflected by accelerated development through these key developmental stages (Coventry *et al.*, 1993; Kirkegaard *et al.*, 2014; Boden *et al.*, 2015; Ochagavía *et al.*, 2018; Prieto *et al.*, 2018). Understanding the temperature pathways that influence this process and how they interact with photoperiod, will be critical to efforts to increase yields in wheat, especially within the context of an increasingly warming climate.

### 4.5.2 *VRN1* in the leaf

In addition to a role regulating *FT1*, our data indicate a role for *Ppd-1* in the regulation of *VRN1* expression. *VRN1* was influenced by *Ppd-1* allelism at specific daily timepoints, with transcripts trending higher in photoperiod-insensitive lines and lower in *ppd-1*. However, *VRN1* was expressed robustly in the absence of *Ppd-1*, especially under short to neutral daylengths, indicating that seasonal regulation of *VRN1* in spring wheat involves more factors than *Ppd-1*. Our data are consistent with the recent analysis in tetraploid wheat, which showed *VRN1* is expressed strongly under short-day photoperiods in *ppd-1* lines at levels comparable or higher than wild-type (Shaw *et al.*, 2020).

### 4.5.3 *FLC* and *VRN1* in the meristem

The role of *FLC* in the regulation of *FT* through vernalization in *Arabidopsis* is well known. However, in wheat, the functional equivalent of a floral repressor like *FLC* is the unrelated *VRN2* (Greenup *et al.*, 2009). I identify high expression of *FLC* orthologues in wheat throughout inflorescence development, pointing towards a potential mechanism for temperature-mediated regulation (Figure 4.7).

In *Arabidopsis*, *FLC* has been shown to bind to many genes throughout developmental stages (Deng *et al.*, 2011). A total of 505 FLC binding sites have been identified, mostly in the promoter regions of genes, including a CArG box motif that is known to be associated with MADS-box genes (Deng *et al.*, 2011). This, combined with the high expression shown here, indicates that *FLC* is likely to act as a regulator of gene expression in the developing inflorescence, possibly in a temperature-dependent way. The molecular function has been characterised for some of these *FLC* copies. The gene encoding *ODDSOC2* has recently been annotated as *FLC-4* (Figure 4.7; Schilling *et al.*, 2020). Interestingly, *ODDSOC2* is upregulated in the *ppd-1* null lines, and *ODDSOC2* has been identified to perform a repressive role on inflorescence development, particularly under high temperatures (Dixon *et al.*, 2019). *ODDSOC2* is expressed higher in the *ppd-1* null lines, particularly at the TS stage. Since these null lines have delayed inflorescence development, it points to *ODDSOC2* possibly acting as an integrator between photoperiod and temperature signals. *ODDSOC1* has been annotated as *FLC-3* (*TaAGL42*) (Figure 4.7; Schilling *et al.*, 2020). *ODDSOC1* is expressed highly throughout inflorescence development, increasing as it develops, suggesting a role during the later stages.

High *VRN1* expression throughout development indicates that *VRN1* is inducted before the VG to DR transition (Figure 4.8). VRN1 may be acting as a priming signal to prepare for the transitioning signal. An essential role for *VRN1* in meristem development has recently been reported in tetraploid wheat. Alongside the MADS-box genes *FUL2* and *FUL3, VRN1*

plays a redundant role in spikelet and spike development in addition to a role suppressing the lower leaf ridge (Li *et al.*, 2019). In the triple mutant, whilst the meristem develops, the spikelet meristem remains vegetative, failing to produce glume and lemma primordia. FUL2 represses *VRN1,* which has an additive effect on its acceleration of flowering when overexpressed. Strikingly, the impact of the triple mutant of *vrn1*, *ful2* and *ful3* does not appear to affect the morphology of the main inflorescence meristem but only the spikelet meristem, limiting floral development. In this study, according to the gene naming conventions recently described, *FUL2* is *AP1-3*, *FUL3* is *AP1-2* and *VRN1* is *AP1-1* (Schilling *et al.*, 2020). *AP1-A1* (*VRN1*) expression remains high throughout inflorescence development without any significant changes. However, *AP1-3* (*FUL2*) exhibits increasing expression levels as the meristem develops, peaking at the TS stage. Likewise, *AP1-2* (*FUL3*) shows an increasing expression as the meristem develops, however only for the B genome, indicating that in any future studies for this gene in hexaploid wheat, only the B genome should be targeted. These expression patterns point to a checkpoint model, whereby *VRN1* is required to initiate development, but its function is mediated by specific spatial expression of *FUL2* and *FUL3*. *VRN1*, *FUL2* and *FUL3* are all significantly lower in the *ppd-1* line at TS, relative to wild-type. These data point to an intricate *Ppd-1*-mediated expression of these genes. These results are an example of the insights RNA-seq transcriptome analysis can provide into many unanswered biological questions.

This chapter uses a detailed field-based phenotypic analysis over three years, quantitative PCR expression analysis in the leaf and high-quality RNA-sequencing analysis, all under field conditions to gain insights into how temperature influences inflorescence meristem development. I have shown that the early stages of inflorescence development occur during a period of relatively low temperature that has largely not been examined. Together, these data help provide a platform for future studies into the interaction between temperature and photoperiod in a seasonal context.

## 4.6 **Methods**

### 4.6.1 Plant material, growth conditions and sampling of *Ppd-1* lines

The plant material, growth conditions and sampling used in this chapter have been characterised in chapter 2.4. Briefly, this chapter uses wild-type, photoperiod insensitive *Ppd-D1a* and *ppd-1* null NILs. Inflorescences were sampled at the vegetative (VG), double ride (DR), lemma primordium (LP) and terminal spikelet (TS) stages from the field and glasshouse. A growing season was regarded as beginning at sowing until flowering the following year. For the temperature shift experiment plants were grown under natural

glasshouse conditions until plants reached the double ridge stage of development. Randomly chosen plants were then shifted to a glasshouse maintained at 16 °C under natural photoperiod conditions (10- 12 h photoperiods occurred during sampling).

## 4.6.2 Temperature data collection

Temperature data for the field represented by temperature data for Norwich, Norfolk, UK was accessed through a publicly available database provided by (CustomWeather© 2020; https://www.timeanddate.com/weather/uk/norwich/historic) originally collected by Norwich Weather Centre. Glasshouse measurements were provided by the John Innes Centre horticultural services using inbuilt temperature sensing equipment.

## 4.6.3 Gene expression comparisons

RT-qPCR analysis of the *VRN1* gene in field leaf tissue was carried out on the same field tissue from the 2019 growing season described in (Chapter 2.5.7). Primers used are identified in (Table 2.1). RNA-sequencing analysis was carried out using the data identified in (chapter 3.5).

## 4.6.4 Statistical analysis

Statistical comparisons of gene expression between tissues were carried out using sleuth-0.28.0 (Pimentel *et al.*, 2017) with default parameters. Sleuth uses the Wald test to test statistical differences between conditions. Transcripts that had a *p* value < 0.05 were considered differentially expressed.

# Chapter 5    A rapid strategy for gene discovery

## 5.2 Chapter summary

In this chapter, I designed a screen to rapidly identify genes involved in inflorescence development. Candidate genes were identified using the RNA-seq data described in Chapter 3 and were subsequently characterised using TILLING and CRISPR/Cas9 mutagenesis. This screen focused of the DR to LP transition and identified and verified four novel genes with a strong involvement in inflorescence development architecture.

## 5.3 Introduction

Novel gene discovery in wheat is time-consuming and laborious, relative to model organisms. This is due to the large complex genome and lengthy generation time restricting genetic analyses (Adamski *et al.*, 2020). Here, I describe a rapid process for gene discovery using transcriptomics, to swiftly identify genes that influence inflorescence development. Similar methods have been used fruitfully in other organisms to identify key developmental genes and guide more traditional mutagenesis-based approaches (Gómez-Ariza *et al.*, 2019).

Mutant screens have been used extensively throughout plant science to identify genes of interest. Commonly, wild-type plants are mutated using chemicals or radiation, and the progeny are screened for desirable phenotypes. This classic method has been altered to provide a greater depth of investigation. For example, an *Arabidopsis* study that aimed to identify C class flowering genes in addition to *AGAMOUS* (*AG*) used the *ag-4* mutant background that exhibited a weak phenotype to screen for mutations that enhanced the phenotype (Chen & Meyerowitz, 1999). From this screen, *HUA1* and *HUA2* were identified as interactors of *AG*, which mediate its function. This study demonstrates the power of intelligently altering the classic mutant screen method to identify novel genes.

Genetic resources in wheat are being developed at a rapid rate, ushering in a period of rapid discovery in this important crop (Adamski *et al.*, 2020). One of the key developments has been the production of a TILLING resource (Krasileva *et al.*, 2017). This resource consists of a population of 2735 hexaploid *cv.* Cadenza and tetraploid c*v.* Kronos lines with a combined 10 million mutations in protein-coding regions (Krasileva *et al.*, 2017). These populations include knockout mutants in almost every gene of hexaploid wheat, allowing mutant screens such as those carried out in other species to be used to great effect in this complex polyploid (Krasileva *et al.*, 2017).

There are two common approaches for identifying novel genes in plants. Forward and reverse genetics. Forward genetics is based on identifying a heritable trait from a large mutant screen, whereas reverse genetics relies on identifying a gene of interest and characterising it through methods such as mutagenesis. In wheat, forward genetics is commonly used to generate phenotypes, involving screening thousands of plants and years of breeding (Peters *et al.*, 2003; Henikoff *et al.*, 2004; Alonso & Ecker, 2006; Jankowicz-Cieslak & Till, 2016).

Both forward and reverse genetic approaches have been used for gene discovery in wheat. For example a forward genetic approach was used recently to identify the molecular basis of awn formation in wheat regulation by *Tipped 1* (*B1*) (Würschum *et al.*, 2020). Using a panel of 1110 winter wheat cultivars, QTL fine-mapping and expression analysis was used to identify a putative zinc finger transcription factor as the candidate gene underlying *B1* with misexpression resulting in a lack of Awn formation.

Reverse genetics has also been used effectively to identify genes of interest in wheat and is becoming more feasible with advances such as TILLING and improvements of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/ CRISPR associated protein 9 (Cas9). For example, a reverse genetics approach was use to discover new alleles within the small heat shock proteins 26 (sHsp26) family (Comastri *et al.*, 2018). A high throughput TILLING screen was used to, isolating and molecularly characterise *sHsp26* homologs in wheat. HSP genes are known to contribute to acquisition of thermotolerance by preventing aggregation of misfolded proteins (Vierling, 1991; Wang *et al.*, 2004; Tyedmers *et al.*, 2010). Similarly, targeted gene mutagenesis with transcription activator–like effector nuclease (TALEN) and CRISPR/Cas9 has been used to mutate all three homoalleles of MILDEW-RESITANCE LOCUS (MLO) providing resistance to powdery mildew infection (Wang *et al.*, 2014).

Here we use a reverse genetic screen to identify genes likely to be involved in wheat inflorescence development. The genes were selected using outputs of the RNA-seq analysis, and I verified their involvement using TILLING and CRISPR/Cas9. Using this method, I have successfully identified four previously uncharacterised genes, illuminating their involvement in wheat inflorescence development.

## 5.4 **Results**

### 5.4.1 Phase 1- Identifying candidate genes from RNA-Seq data

My work highlighted that the DR to LP transition is a key developmental stage transition, and so I initially focused on genes that are differentially expressed during these stages. The first phase of the screen used the previously described expression analysis to examine differentially regulated genes between genotypes (Chapter 3; Figure 3.2). I identified the top 100 genes that were significantly differentially up- and down-regulated in the *Ppd-D1a* insensitive and *ppd-1* null NILs, relative to wild-type. For example, the 4765 and 4311 genes that were up- and down-regulated in *ppd-1* NILs between the DR and LP stages were compared to the 1938 and 1018 genes that were up- and down-regulated in wild-type. The genes not shared between these lines were then ranked according to their significance (q value), and the top 100 genes in this list were selected for further analysis. This process was performed for both the *ppd-1* and *Ppd-D1a* NILs for genes up- and down-regulated, relative to wild-type. This process produced 400 candidate genes that would be used for subsequent analysis (Figure 5.1a).
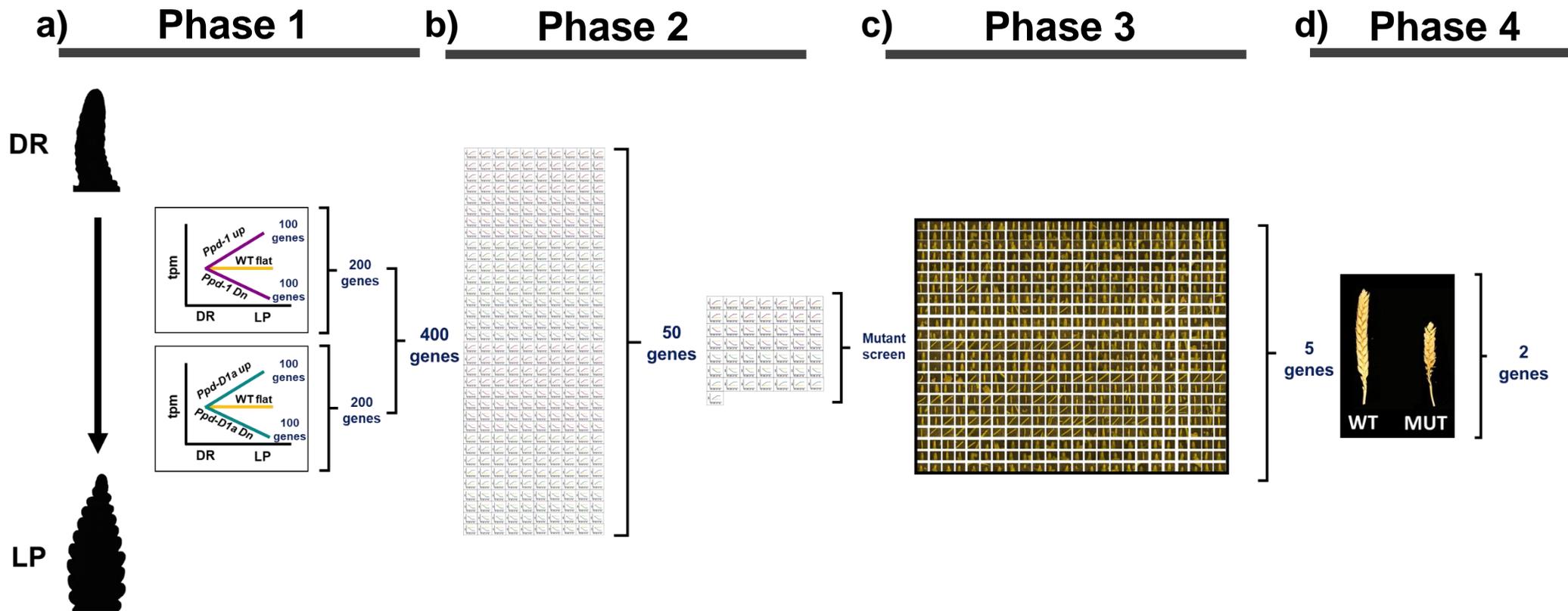
**Figure 5.1 - Mutant screen summary**

A visual summary of the strategy taken in this mutant screen. **a)** Phase 1: The top 100 genes most significantly up- and down-regulated between the double ridge (DR) and lemma primordium (LP) stages in the *Ppd-D1a* and *ppd-1* lines relative to wild-type are selected. Significance= q value <0.05. **b)** Phase 2: TPM values of each gene (400 genes total) were plotted and examined, based on expression pattern characteristics 50 genes are selected. **c)** Phase 3: where possible 3 TILLING mutants of each gene are grown under controlled conditions. Plants are dissected and phenotyped when wild-type plants reach the TS stage. Based on spikelet number, development stage and spike length 5 genes were selected for continuation. **d)** Phase 4: TILLING lines of each gene are grown under controlled GH conditions and phenotyped at maturity. Segregation analysis on two independent mutant lines verifies the candidate gene.

## 5.4.2 Phase 2- Selection based on expression

To perform a mutant screen, the list of candidate genes had to be refined. Ranking genes by statistics alone can be misleading because whilst gene expression across two conditions may be statistically different, the true TPM difference may not be biologically relevant. For example, a difference of < 0.5 TPM could be statistically significant but is not likely to be biologically significant. Therefore, in phase 2 of the screen, the expression values (TPM) for all candidate genes were plotted using RNA-seq data (Figure 5.1b; Figure 5.2). To reduce selection bias, I did not functionally annotate genes at this stage. Based on their expression profile across all stages and between genotypes, genes were shortlisted from 400 to 50 total genes, 25 up-regulation between the DR-LP transition and 25 down-regulated. Genes were selected for progression to the next stage of the screen, based on expression pattern, amplitude of expression, difference between genotypes and standard error of data points.

From this screen, I selected a total of 25 genes most significantly down-regulated between the DR-LP stage relative to wild-type (Figure 5.2a). These genes were predominantly selected based on their *ppd-1* profile. Many of the genes were selected based on very high peaks in expression at the DR stage in the *ppd-1* null line relative to wild-type. Particularly strong candidates were genes that have correspondingly low expression in the *Ppd-D1a* line and high expression in the *ppd-1* null line, relative to wild-type. Other genes were selected based on generally higher expression in the *ppd-1* null line over all stages.

This screen also identified 25 genes most significantly up-regulated between the DR to LP stage in the *Ppd-D1a* and *ppd-1* null line relative to wild-type (Figure 5.2b). These genes were selected under the same criteria as the genes that were down-regulated (Figure 5.2a); however, the focus shifted towards up-regulated genes between DR and LP. The genes with the most interesting profiles in this category tended to be selected based on their *Ppd-D1a* profile.

The expression profiles of these genes identify them as candidates for involvement in a pathway controlling floral meristem development under the influence of *Ppd-1*. To remove bias at this phase, no genes were functionally annotated, and genes were selected based on expression profile alone. These candidate genes will be functionally annotated in subsequent phases.
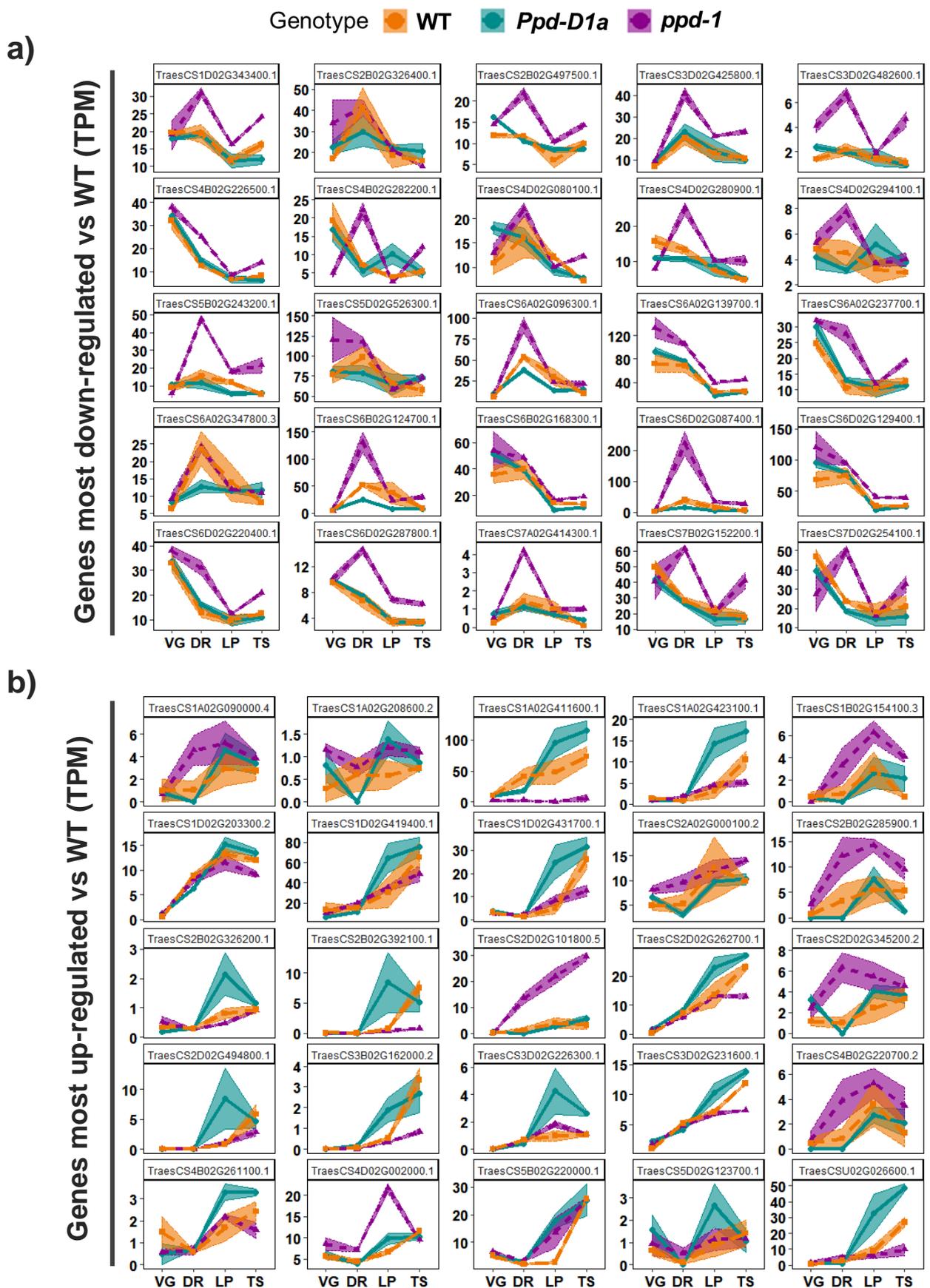
**Figure 5.2 - Phase 2: Top 50 genes selected based on expression profile**

Expression of the top 50 genes **a)** up- and **b)** down-regulated between double ridge (DR) and lemma primordium (LP) stages. Expression is shown at the vegetative (VG), DR, LP and terminal spikelet (TS) stages. Gene ID is shown above each sub-graph. Genes were selected from 400 genes based on expression level, expression profile and SEM of data points. Expression is shown for wild-type (Orange), Photoperiod insensitive *Ppd-D1a* (Cyan)*,* and *ppd-1* (Magenta). Genes expression profiles are mean TPM ± SEM.

### 5.4.2.1 Functional annotation of candidate genes

During the process of narrowing down the gene list from 400 to 50, the functional annotation of gene IDs were not interrogated to eliminate bias based on function from the selection process. However, at this point, we aimed to take the candidate genes forward for a mutant screen using the EMS-induced bread wheat TILLING population (Krasileva *et al.*, 2017). Therefore, a further selection step was performed and the genes were functionally annotated. Of particular interest were genes for which 2-3 homoeologues were identified. For example, all three homoeologues of a bZIP transcription factor (TraesCS6A02G096300.1, TraesCS6B02G124700.1, TraesCS6D02G087400.1 ;*bZIP*) and LIGHT-DEPENDENT SHORT HYPOCOTYLS-like protein (TraesCS6A02G139700.1, TraesCS6B02G168300.1, TraesCS6D02G129400.1 ;*LDSH*) were identified. There were also several genes with two homoeologues represented, including Auxin-induced in root cultures protein 12 (TraesCS7B02G152200.1 and TraesCS7D02G254100.1; *AIRP*), Lipid transfer protein (TraesCS4B02G282200.1 and TraesCS4D02G280900.1; *LTP*), YABBY protein (TraesCS6A02G237700.1 and TraesCS6D02G220400.1; *YABBY*)), ATP-dependent zinc metalloprotease FtsH 1 (TraesCS1A02G423100.1 and TraesCS1D02G431700.1; *FtsH 1*)) and CTP synthase (TraesCS1A02G411600.1 and TraesCS1D02G419400.1; *CTP*). There were also 6 auxin-related genes identified by this analysis, further supporting the conclusion that auxin-related processes play a significant role during early inflorescence development. Based on the putative gene function, and TILLING mutant availability, several genes were excluded from further analysis. These include AT hook motif DNA-binding family protein (TraesCS5B02G243200.1), Auxin responsive SAUR protein (TraesCS2D02G494800.1) and Defensin-like protein (TraesCSU02G026600.1).

To investigate candidate genes, I screened mutants of the corresponding gene from the EMS-induced Cadenza TILLING population (Figure 5.1c). Where possible, two to three TILLING mutants were obtained for each candidate gene. Mutations generating premature stop codons were favoured over missense mutants. Notably, two candidate genes failed to have any deleterious missense or null mutations, which were the three genome copies of the *BZIP transcription factor* and *LIGHT-DEPENDENT SHORT HYPOCOTYLS-like protein*. The lack of deleterious mutations in homoeologues of these genes further highlighted their potential involvement in reproduction, as plants with severally affected flowering time were not likely to make it through the TILLING resource generation phase. Based on these observations, both genes were selected for CRISPR/Cas9 knockout experiments. Optimally, all genes would be subject to CRISPR/Cas9 analysis, due to the ability to knock-out multiple genome copies and lack of background mutations; however,

these screens are resource-heavy and expensive and therefore TILLING mutants were selected to characterise gene function.

# Table 5.1 - Up-regulated genes selected for mutant analysis

The up-regulated genes selected for TILLING mutant analysis with gene ID and functional annotation alongside corresponding TILLING lines used.

| Up | Gene ID | Functional annotation | TILLING line 1 | Type | TILLING line 2 | Type | TILLING line 3 | Type |
|----|---------|----------------------|----------------|------|----------------|------|----------------|------|
| 1 | TraesCS5B02G243200.1 | AT hook motif DNA-binding family protein | | | | | | |
| 2 | TraesCS6D02G287800.1 | Auxin efflux carrier component | Cadenza0129.chr6D.397082360 | missense variant | Cadenza0273.chr6D.397083815 | Stop gained | Cadenza0586.chr6D.397082600 | missense variant |
| 3 | TraesCS4B02G226500.1 | Auxin influx transporter | Cadenza1371.chr4B.473568922 | Stop gained | Cadenza1051.chr4B.473571596 | Stop gained | | |
| 4 | TraesCS2B02G326400.1 | Auxin response factor | Cadenza0127.chr2B.466193624 | Stop gained | Cadenza0188.chr2B.466192620 | Stop gained | | |
| 5 | TraesCS7B02G152200.1 | Auxin-induced in root cultures protein 12 | Cadenza1806.chr7B.202359987 | missense variant | Cadenza1775.chr7B.202359840 | missense variant | Cadenza1409.chr7B.202359891 | missense variant |
| 6 | TraesCS7D02G254100.1 | Auxin-induced in root cultures protein 12 | Cadenza1644.chr7D.230942192 | missense variant | Cadenza0232.chr7D.230942228 | missense variant | Cadenza1511.chr7D.230942306 | missense variant |
| 7 | TraesCS6A02G096300.1 | BZIP transcription factor | Cadenza1842.chr6A.63609863 | missense variant | | | | |
| 8 | TraesCS6B02G124700.1 | BZIP transcription factor | Cadenza1597.chr6B.119848827 | missense variant | | | | |
| 9 | TraesCS6D02G087400.1 | BZIP transcription factor | Cadenza1803.chr6D.52723954 | missense variant | | | | |
| 10 | TraesCS6A02G347800.3 | Diacylglycerol kinase | Cadenza0336.chr6A.580199004 | Stop gained | Cadenza0053.chr6A.580198404 | missense variant | | |
| 11 | TraesCS4D02G080100.1 | Dof zinc finger protein | Cadenza1357.chr4D.54127289 | missense variant | | | | |
| 12 | TraesCS3D02G482600.1 | GRAS family transcription factor containing protein | Cadenza1442.chr3D.580089052 | missense variant | Cadenza0097.chr3D.580088966 | missense variant | | |
| 13 | TraesCS2B02G497500.1 | Homeobox-leucine zipper family protein | Cadenza0423.chr2B.694055848 | Stop gained | Cadenza1701.chr2B.694057145 | missense variant | | |
| 14 | TraesCS1D02G343400.1 | HXXXD-type acyl-transferase family protein | Cadenza0132.chr1D.432406319 | missense variant | Cadenza0353.chr1D.432406244 | missense variant | Cadenza1523.chr1D.432408956 | missense variant |
| 15 | TraesCS5D02G526300.1 | Leucine-rich repeat receptor-like protein kinase | Cadenza0346.chr5D.544231711 | missense variant | Cadenza2014.chr5D.544232157 | missense variant | Cadenza0676.chr5D.544232500 | missense variant |
| 16 | TraesCS6A02G139700.1 | LIGHT-DEPENDENT SHORT HYPOCOTYLS-like protein (DUF640) | Cadenza1785.chr6A.114458746 | missense variant | | | | |
| 17 | TraesCS6B02G168300.1 | LIGHT-DEPENDENT SHORT HYPOCOTYLS-like protein (DUF640) | Cadenza1468.chr6B.178273107 | missense variant | Cadenza1736.chr6B.178273187 | missense variant | Cadenza1576.chr6B.178272551 | missense variant |
| 18 | TraesCS6D02G129400.1 | LIGHT-DEPENDENT SHORT HYPOCOTYLS-like protein (DUF640) | Cadenza1976.chr6D.94602485 | missense variant | Cadenza0196.chr6D.94602507 | missense variant | Cadenza0594.chr6D.94602570 | missense variant |
| 19 | TraesCS4B02G282200.1 | Lipid transfer protein | Cadenza1169.chr4B.565313306 | missense variant | Cadenza1467.chr4B.565313118 | missense variant | Cadenza0643.chr4B.565313168 | missense variant |
| 20 | TraesCS4D02G280900.1 | Lipid transfer protein | Cadenza1265.chr4D.451801982 | missense variant | Cadenza1770.chr4D.451802495 | missense variant | Cadenza0902.chr4D.451801914 | missense variant |
| 21 | TraesCS4D02G294100.1 | Lipoxygenase | Cadenza1586.chr4D.464533453 | Stop gained | Cadenza1675.chr4D.464536313 | Stop gained | | |
| 22 | TraesCS7A02G414300.1 | Myb transcription factor | Cadenza1471.chr7A.604853538 | Stop gained | Cadenza1367.chr7A.604853528 | missense variant | Cadenza2109.chr7A.604853945 | missense variant |
| 23 | TraesCS3D02G425800.1 | Squamosa promoter binding-like protein | Cadenza1763.chr3D.538401404 | missense variant | Cadenza1494.chr3D.538402702 | missense variant | Cadenza0551.chr3D.538402629 | missense variant |
| 24 | TraesCS6A02G237700.1 | YABBY protein | Cadenza0283.chr6A.446901749 | Stop gained | Cadenza0687.chr6A.446903009 | missense variant | | |
| 25 | TraesCS6D02G220400.1 | YABBY protein | Cadenza0246.chr6D.310521429 | missense variant | Cadenza1172.chr6D.310521399 | missense variant | Cadenza0351.chr6D.310521808 | missense variant |

## Table 5.2 - Down-regulated genes selected for mutant analysis

The down-regulated genes selected for TILLING mutant analysis with gene ID and functional annotation alongside corresponding TILLING lines used.

| Down | Gene ID | Functional annotation | TILLING line 1 | Type | TILLING line 2 | Type | TILLING line 3 | Type |
|---|---|---|---|---|---|---|---|---|
| 1 | TraesCS1A02G090000.4 | Cullin-1 | Cadenza0946.chr1A.79702543 | Stop gained | Cadenza1159.chr1A.79702403 | Stop gained | Cadenza1025.chr1A.79701438 | Stop gained |
| 2 | TraesCS1A02G208600.2 | Phosphatidylinositol N-acetyglucosaminlytransferase subunit P-like protein | Cadenza1655.chr1A.370104837 | splice acceptor variant | Cadenza0883.chr1A.370106183 | missense variant | Cadenza0266.chr1A.370107390 | missense variant |
| 3 | TraesCS1A02G411600.1 | CTP synthase | Cadenza1773.chr1A.572208563 | Stop gained | Cadenza0888.chr1A.572209408 | Stop gained | Cadenza0203.chr1A.572209799 | missense variant |
| 4 | TraesCS1A02G423100.1 | ATP-dependent zinc metalloprotease FtsH 1 | Cadenza1713.chr1A.578517767 | Stop gained | Cadenza2085.chr1A.578518837 | Stop gained | | |
| 5 | TraesCS1B02G154100.3 | F-box / LRR-repeat protein | Cadenza0467.chr1B.255506318 | Stop gained | Cadenza0111.chr1B.255510967 | Stop gained | | |
| 6 | TraesCS1D02G203300.2 | MADS-box transcription factor | Cadenza1220.chr1D.287457310 | missense variant | Cadenza0884.chr1D.287467826 | missense variant | | |
| 7 | TraesCS1D02G419400.1 | CTP synthase | Cadenza1289.chr1D.476439440 | Stop gained | Cadenza0105.chr1D.476440573 | Stop gained | | |
| 8 | TraesCS1D02G431700.1 | ATP-dependent zinc metalloprotease FtsH 1 | Cadenza0743.chr1D.481695532 | Stop gained | Cadenza1262.chr1D.481693568 | missense variant | Cadenza0510.chr1D.287457324 | missense variant |
| 9 | TraesCS2A02G000100.2 | Polycomb group protein VERNALIZATION 2 | Cadenza0478.chr2A.251529 | Stop gained | Cadenza0988.chr2A.254521 | Stop gained | | |
| 10 | TraesCS2B02G285900.1 | Tobamovirus multiplication 1 | Cadenza1646.chr2B.393856324 | missense variant | Cadenza1522.chr2B.393856629 | missense variant | Cadenza1735.chr2B.393856701 | missense variant |
| 11 | TraesCS2B02G326200.1 | Protein NRT1/ PTR FAMILY 1.2 | Cadenza0723.chr2B.465826242 | Stop gained | Cadenza1277.chr2B.465828138 | Stop gained | | |
| 12 | TraesCS2B02G392100.1 | GRAM domain-containing protein / ABA-responsive | Cadenza0732.chr2B.556517351 | missense variant | Cadenza2056.chr2B.556517614 | missense variant | Cadenza2110.chr2B.556517677 | missense variant |
| 13 | TraesCS2D02G101800.5 | RNA-binding domain CCCH-type zinc finger protein | Cadenza0638.chr2D.55058720 | Stop gained | Cadenza0638.chr2D.55058720 | Stop gained | | |
| 14 | TraesCS2D02G262700.1 | MADS-box transcription factor | Cadenza1929.chr2D.318916652 | Stop gained | Cadenza1929.chr2D.318916652 | Stop gained | | |
| 15 | TraesCS2D02G345200.2 | Kinase family protein | Cadenza0725.chr2D.441526925 | missense variant | Cadenza1129.chr2D.441527381 | missense variant | | |
| 16 | TraesCS2D02G494800.1 | Auxin responsive SAUR protein | | | | | | |
| 17 | TraesCS3B02G162000.2 | Flowering locus T | Cadenza0352.chr3B.158718848 | Splice | Cadenza0122.chr3B.158718931 | missense variant | Cadenza1655.chr3B.158715883 | missense variant |
| 18 | TraesCS3D02G226300.1 | WRKY transcription factor | Cadenza1491.chr3D.308057589 | missense variant | Cadenza0076.chr3D.308057619 | missense variant | Cadenza0927.chr3D.308058937 | missense variant |
| 19 | TraesCS3D02G231600.1 | XH/XS domain-containing family protein | Cadenza1073.chr3D.317199209 | Stop gained | Cadenza0988.chr3D.317199223 | Stop gained | | |
| 20 | TraesCS4B02G220700.2 | E3 ubiquitin-protein ligase | Cadenza0404.chr4B.464157276 | Stop gained | Cadenza0510.chr4B.464157489 | Stop gained | | |
| 21 | TraesCS4B02G261100.1 | chromatin remodeling factor CHD3 (PICKLE) | Cadenza1313.chr4B.529676961 | missense variant | Cadenza1313.chr4B.529676961 | missense variant | Cadenza0225.chr4B.529677558 | missense variant |
| 22 | TraesCS4D02G002000.1 | Mesoderm induction early response protein 1, putative isoform 1 | Cadenza1181.chr4D.1241952 | Splice | Cadenza1597.chr4D.124201 | missense variant | Cadenza0879.chr4D.1242032 | missense variant |
| 23 | TraesCS5B02G220000.1 | S-adenosylmethionine decarboxylase proenzyme | Cadenza2019.chr5B.394267358 | Stop gained | Cadenza1254.chr5B.394266468 | missense variant | Cadenza2070.chr5B.394266550 | missense variant |
| 24 | TraesCS5D02G123700.1 | Histone-lysine N-methyltransferase setd3 | Cadenza1815.chr5D.184027816 | Stop gained | Cadenza1508.chr5D.184029271 | missense variant | Cadenza0703.chr5D.184039248 | missense variant |
| 25 | TraesCSU02G026600.1 | Defensin-like protein | | | | | | |

### 5.4.3 Phase 3- A mutagenesis screen at TS

To facilitate efficient characterisation of the TILLING mutants, a screen was designed whereby all plants would be germinated at the same time, alongside wild-type controls. Whenever the wild-type plants (*cv.* Cadenza) reached the TS stage, all mutant lines would be dissected to determine relative progression through early inflorescence development. Inflorescences were extracted and imaged for eight plants per TILLING line, to include a total of 850 plants that were processed in this manner. Measurements of spikelet number, meristem stage and length were collected for each mutant. Using this approach, I was able to identify lines that disrupted meristem development or architecture as these traits are determined by the TS stage. Moreover, any delay at the DR or LP stages would be identifiable at the TS stage, relative to wild-type. Lines with strong phenotypes were selected for further analysis, particularly if they behaved consistently between independent TILLING mutants for that gene (Figure 5.3,5.1c). The TILLING lines tend to be a mixture of homozygous, heterozygous and wild-type genotypes for the allele of interest, and so in the data we observe a variety of phenotypes per line. I considered the influence of this heterozygosity by analysing the average of phenotypes per TILLING line. From this analysis, I shortlisted five genes for further investigation (Table 5.3). Mutant lines for these five genes show a wide spread of phenotypes within the 8 representative plants, but generally, the two independent TILLING lines for each gene showed consistent phenotypes.

**Table 5.3 - The top five candidate genes selected for further analysis**

| Gene ID | Functional annotation |
|---|---|
| TraesCS6D02G287800.1 | Auxin efflux carrier component |
| TraesCS7B02G152200.1 | Auxin-induced in root cultures protein 12 |
| TraesCS3D02G425800.1 | Squamosa promoter binding-like protein |
| TraesCS6D02G220400.1 | YABBY protein |
| TraesCS2B02G285900.1 | Tobamovirus multiplication 1 |

The *AUXIN EFFLUX CARRIER COMPONENT* (*AEC*) gene (TraesCS6D02G287800.1) consistently produced fewer spikelets compared to wild-type in the *CAD0129* and *CAD0586* TILLING lines. These lines showed accelerated development relative to wild-type, with a varied spike length. The average decrease in spikelet number is consistent with the accelerated flowering. The *AUXIN-INDUCED IN ROOT CULTURES PROTEIN 12* (*AIR*) gene (TraesCS7B02G152200.1) produced fewer spikelets, on average, relative to wild-type; however, meristem development is delayed, and the spike length of these

mutants is shorter. The phenotypes for all three of the TILLING lines for this gene are consistent with each other (*CAD1775, CAD1806* and *CAD0232*). The gene encoding the *SQUAMOSA PROMOTER BINDING-LIKE PROTEIN* (*SPB*) gene (TraesCS3D02G425800.1) showed consistently higher spikelet number in two independent lines (*CAD1494* and *CAD1763*) with a similar number of spikelets, relative to wild-type and *CAD0551*. Progression to the terminal spikelet stage is delayed in *CAD1494* and *CAD1763* line relative to wild-type; however, progress is dramatically advanced in *CAD0551*. The *TOBAMOVIRUS MULTIPLICATION 1* (*TOM1*) gene (TraesCS2B02G285900.1) was selected because the *CAD1522* line stalled dramatically at the VG stage in 5/8 plants. The remaining 3 plants formed an average of 19 spikelets, whilst the wild-type control produced 17 spikelets. Mutants for the *YABBY* gene (TraesCS6D02G220400.1; *CAD1172* and *CAD351*) produced, on average, more spikelets than wild-type, which correlated with a delay in development.

This mutant screen has identified 5 strong gene candidates for involvement in inflorescence development. The attrition rate in this phase is high, with just 10% (5/50) of the genes examined being continued to the next phase. However, many other TILLING lines had interesting phenotypes; however, because of a restraint on resources, only the top 5 genes were selected. It should also be noted that because TILLING lines were used, a maximum of one genome copy was knocked out in homozygous lines. The fact that we observed so many mutants with phenotypes suggests that redundancy among the genome copies is not prevalent. It also demonstrates an advantage of performing this screen in a hexaploid, as these genes may be vital for plant development and would otherwise be fatal in a diploid species. By knocking out only one genome copy, the phenotype caused by reduced expression may produce a desirable phenotype as opposed to a fatal one.
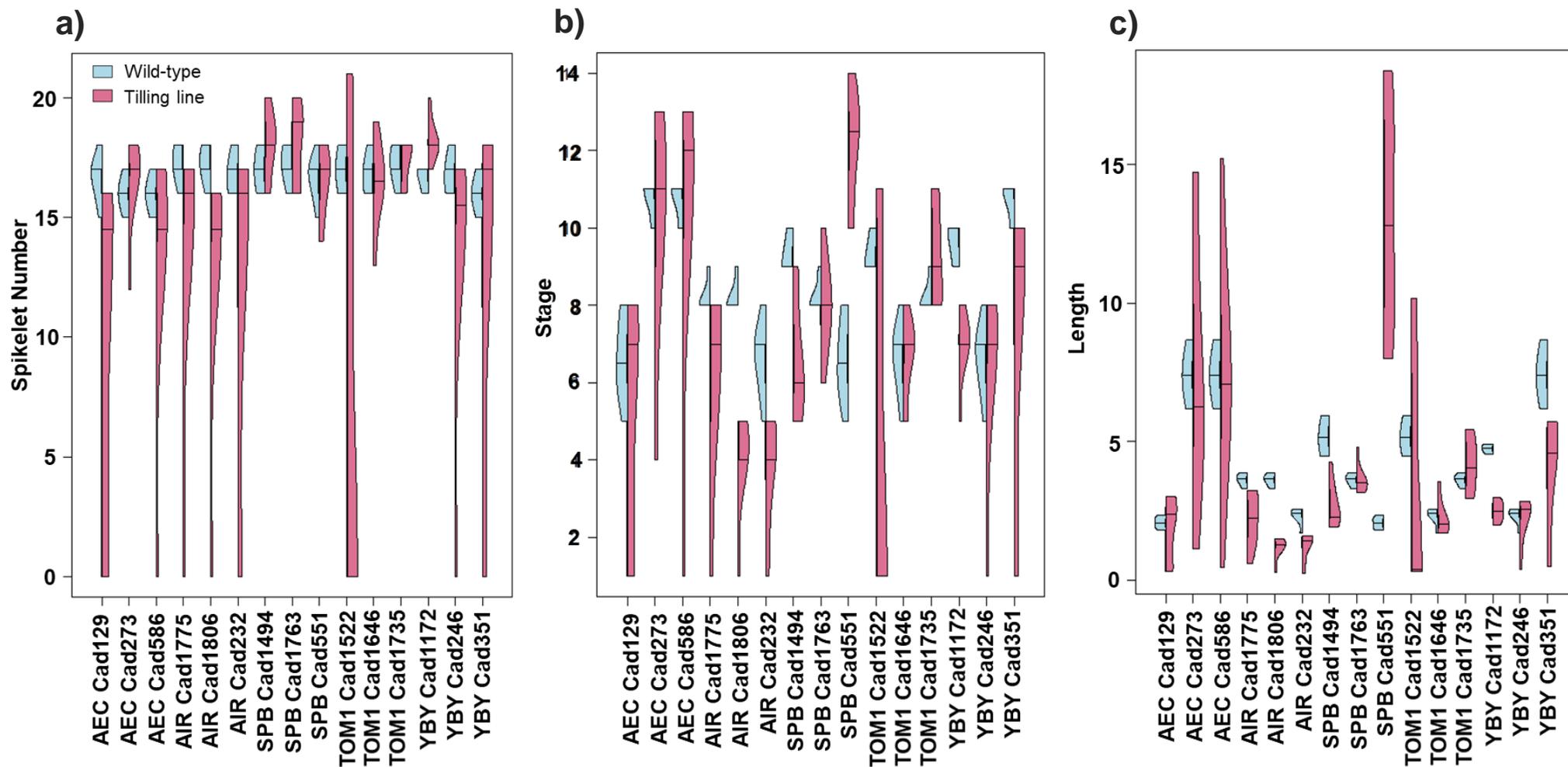
**Figure 5.3 - Phase 3: The phenotypes of genes selected from the screen of dissected inflorescences**

Violin plots summarising the phenotypes of each of the selected TILLING lines. **a)** Spikelet number, **b)** Stage ranging from VG to WA (refer to methods: Chapter 5.6.5), **c)** spike length (mm) for the five genes. Genes tested include: *AUXIN EFFLUX CARRIER COMPONENT* (*AEC*), A*UXIN-INDUCED IN ROOT CULTURES PROTEIN 12* (*AIR*), *SQUAMOSA PROMOTER BINDING-LIKE PROTEIN* (*SPB*), *TOBAMOVIRUS MULTIPLICATION 1* (*TOM1*) and *YABBY protein* (*YBY*). Wildtype is shown in blue and the TILLING line is shown in pink. Data is the summary of 8 samples. Median is denoted by a line across the plot. Width represents density of data points.

## 5.4.4 Phase 4- Segregation analysis

The final phase of this screen is to grow all candidate genes under long-day controlled glasshouse conditions and perform genotype analysis using Kompetitive allele-specific PCR (KASP) (Figure 5.1d). KASP PCR allows rapid large-scale genotyping and is well suited for tracing SNPs in mutant screens. From this phase of the screen, two genes were identified with phenotypes associated with verified homozygous mutations. Importantly, the mutant alleles segregated during this experiment, providing an opportunity to confirm the association of mutations with a given phenotype.

The first gene identified was TraesCS2B02G285900.1, which is functionally annotated as *TOBAMOVIRUS MULTIPLICATION 1* (*TOM1*). I selected TILLING lines that contained mutant alleles with amino acid substitutions. Line *CAD1735* contained a missense allele (V83I) caused by a G > A mutation at base pair 368 bp selected because the predicted SIFT score is 0.03 (i.e. deleterious). Line *CAD1522* expressed a missense allele (V59M) due to a G > A mutation at base pair 296 bp, which was selected because the predicted SIFT score is 0 (i.e. deleterious). I named these alleles *tom-b1.1* and *tom-b1.2,* respectively.

These mutations have strong effects on spike architecture (Figure 5.4). Genotypes containing the *tom-b1.1* or *tom-b1.2* alleles produced fewer spikelets, with the *tom-b1.1* mutant producing significantly less spikelets (16.6 ± 1.5) than wild-type (22.5 ± 0.96; Figure 5.4a, c). The strongest effect was on the number of fertile spikelets, at both the apex and base of the spike, with both the *tom1.2* (9.4 ± 1.122) and *tom1.1* (12.8 ± 0.97) mutants forming significantly less fertile spikelets per spike relative to their respective wild-type siblings, *TM1.2* (20.7 ± 0.33) and *TM1.1* (17.75 ± 0.5; Figure 5.4). These results indicate *TOM1* has a dramatic role in spikelet fertility, likely functioning at the LP stage, where this gene is expressed highly.

The second gene identified encodes an ortholog of *YABBY4* (TraesCS6D02G220400.1). For the *YABBY-like* gene, I selected TILLING lines that contained mutant alleles with either a premature stop codon or an amino acid substitution. Line *CAD1172* contained a premature stop codon (X213*) caused by a G > A mutation at base pair 1241 bp. Line *CAD0351* expressed a missense allele (A196V) due to a G > A mutation at base pair 1191 bp, which was selected because the predicted SIFT (Sorting Intolerant From Tolerant) score is 0 (i.e. deleterious). I named these alleles *yabby-d1.1* and *yabby-d1.2*, respectively.

These *YABBY* mutant lines exhibit similar phenotypes, with both mutants showing a decrease in spikelet number, with the *yabby-d1.1* showing a significant decrease (16 ±

1.6) compared to wild-type siblings (22.3 ± 0.6; Figure 5.5). Both mutant lines show a dramatic decrease in fertile spikelets, *yabby-d1.2* (13.5 ± 0.5) and *yabby-d1.1* (14.7 ± 1.2), compared to wild-type sibling lines,  *YABBY-D1.2* (20.7 ± 0.3) and *YABBY-D1.1* (18.67 ± 0.88).  These data indicate that this *YABBY* gene, like that encoding *TM1*, has an important role regulating floret fertility and spikelet formation.
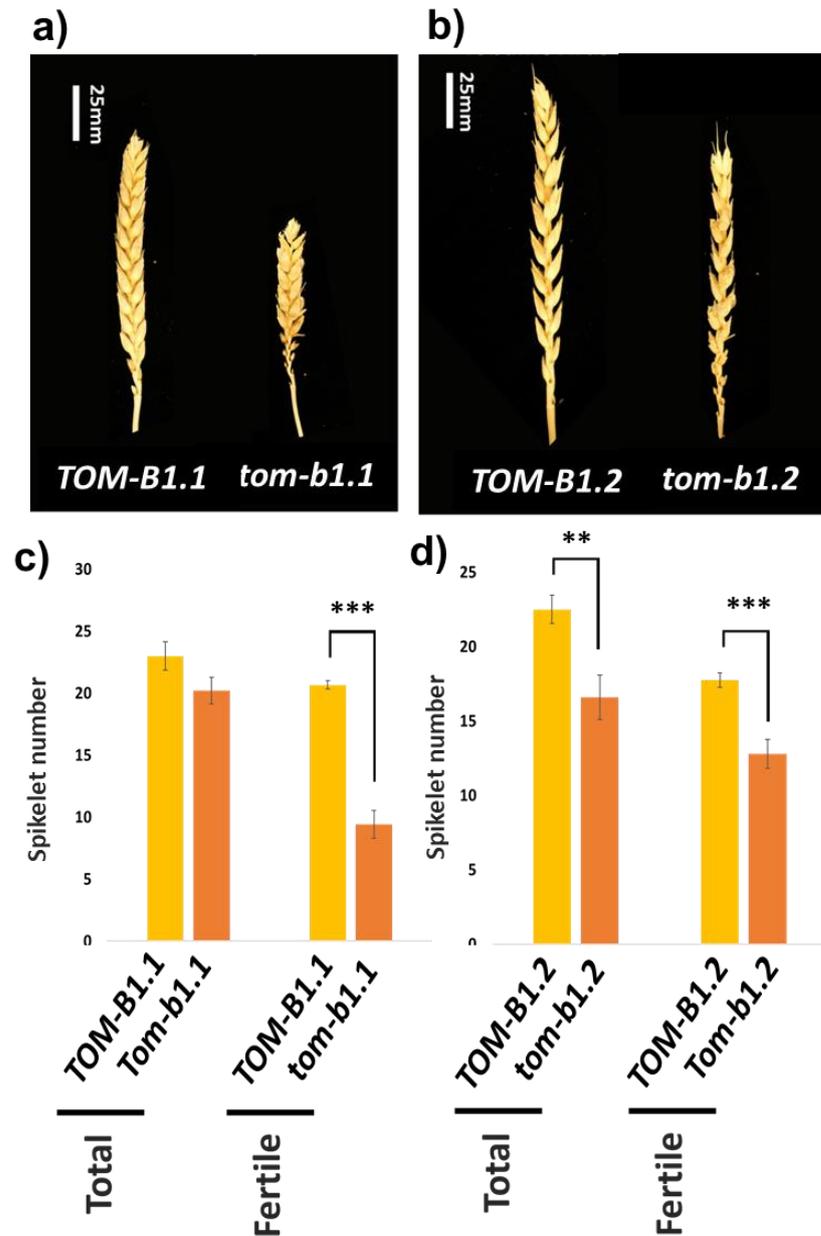
**Figure 5.4 - Spike architecture phenotypes of *tom-b1* TILLING mutants**

**a-b)** Representative inflorescences of the **a)** *tom-b1.1* and **b)** *tom-b1.2* mutants, relative to wild-type. **c-d)** Spikelet number increases on inflorescences of **c)** *tom-b1.1* and **d)** *tom-b1.2* mutants relative to wild-type. Data are the mean ± SEM of at least 5 replicates. Replicates consist of spikes from both primary and secondary tillers of multiple plants. *, $p < 0.05$, **, $p < 0.01$, ***, $< p$ value of 0.005.
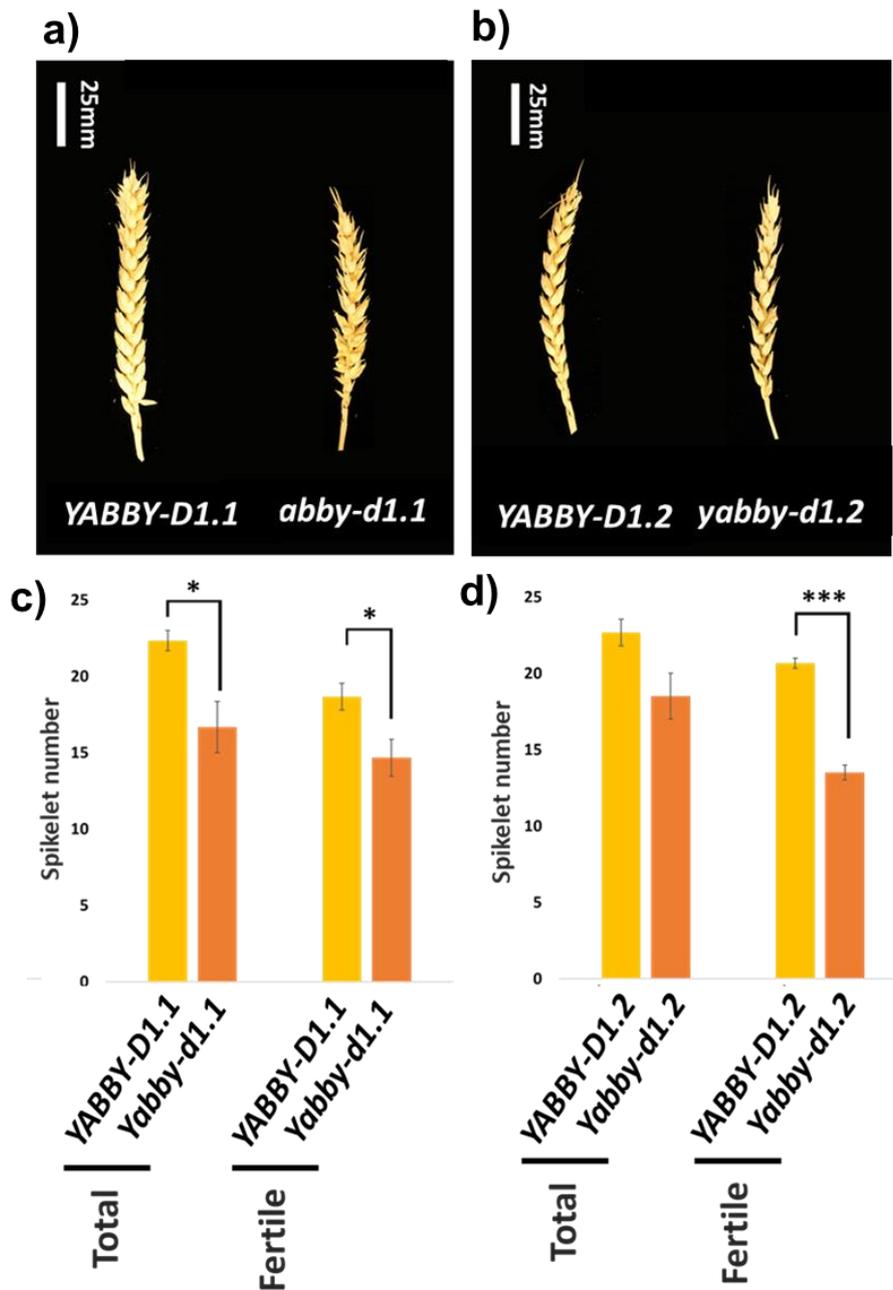
**Figure 5.5 - Spike architecture phenotypes of *yabby-d1* TILLING mutants**

**a-b)** Representative inflorescences of the **a)** *yabby-d1.1* and **b)** *yabby-d1.2* mutants, relative to wild-type. **c-d)** Spikelet number of increases on inflorescences of **c)** *yabby-d1.1* and **d)** *yabby-b1.2* mutants relative to wild-type. Data are the mean ± SEM of at least 5 replicates. Replicates consist of spikes from both primary and secondary tillers of multiple plants. *, *p* < 0.05, **, *p* < 0.01, ***, < *p* value of 0.005.

## 5.4.5 Phylogenetic analysis of YABBY protein

To help characterise the gene encoding the YABBY-like protein identified here, I performed phylogenetic analysis to identify homologs in related species (Figure 5.6). Based on this analysis, the encoded protein is an ortholog of YABBY4 and YABBY4-like proteins in closely related species. Therefore, I named the *YABBY* gene identified here as *YABBY4.*
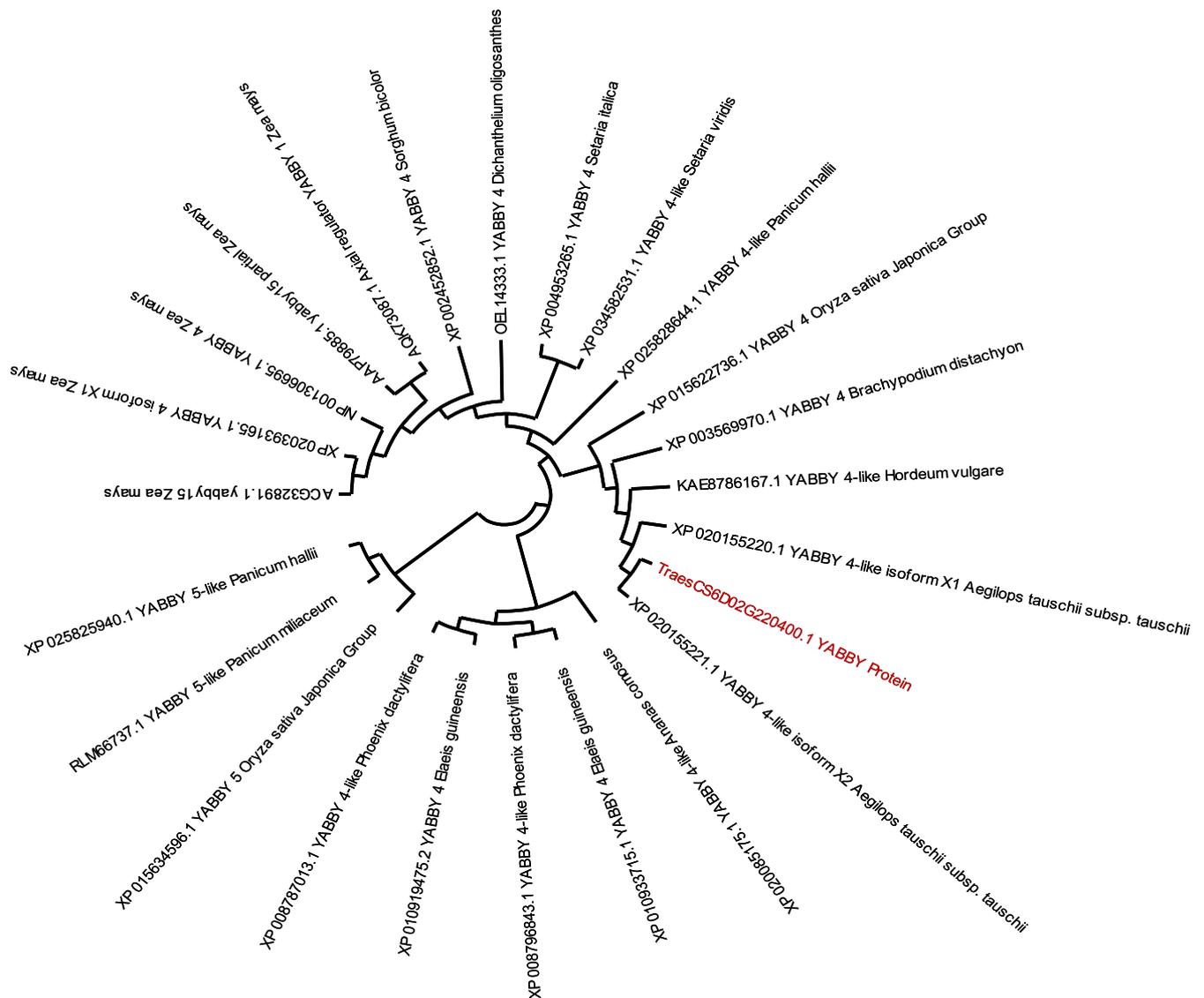


**Figure 5.6 - Phylogenetic tree for YABBY1.**

A phylogenetic analysis for YABBY protein (YABBY1) candidate (red; TraesCS6D02G220400.1) 24 amino acids were compared, obtained through BLAST. An unrooted tree was generated using MEGAX, using the maximum likelihood method and JTT matrix-based model.

## 5.4.6 CRISPR/Cas9 mutagenesis

As part of the screen, I identified two genes of particular interest due to all three homoeologues being identified from phase 2 of the mutant screen (Figure 5.1b; Table 5.1). These genes have homology to bZIP transcription factor (bZIP) proteins and light-dependent short hypocotyl (LDSH) proteins. The genes did not have any deleterious or missense TILLING mutants, so were selected for CRISPR/Cas9 targeted mutagenesis in the hexaploid wheat cultivar Fielder.

### 5.4.6.1  Phylogenetic analysis of *bZIP* and *LDSH*

In wheat, the two candidate genes had not been characterised. Phylogenetic analysis provides clues to their functionality. The *LDSH* gene is 100% identical to protein *G1-like1* in *Aegilops tauschii*, 96.98% to *G1-like1* protein in *Hordeum vulgare* and 73.68% homologous to protein *G1-like1* in *Brachypodium distachyon* (Figure 5.7a). The gene described here is 53.3% protein identical to the annotated rice gene *G1,* this gene has a described role in sterile lemma identity in the rice spikelet as it develops (Yoshida *et al.*, 2009). Based on the functional analysis and homology studies carried out in this study, we name the *LDSH* gene (TraesCS6A02G139700.1, TraesCS6B02G168300.1 and TraesCS6D02G129400.1) as *TaALOG1*.

The *ALOG* (*Arabidopsis LSH1 and Oryza G1*) transcription factors are best described in rice. Therefore, I performed a phylogenetic comparison to the 10 ALOG genes described in rice, alongside *ALOG1* identified in this study (Figure 5.7b; Yoshida *et al.*, 2013). ALOG1 is homologous to *Os*G1L1 in rice, with only a distant relationship to TAWAWA1*,* an *ALOG* gene with significant influence on panicle architecture (Yoshida *et al.*, 2013)*.* This indicates that whilst *ALOG1* is from the same family as *TAWAWA1*, it is not an orthologue.

The *bZIP* gene (TraesCS6A02G096300.1, TraesCS6B02G124700.1 and TraesCS6D02G087400.1) shows 97.28% identity to the predicted *bZIP transcription factor 11-like* in *Aegilops tauschii* and 89.80% to the *Ocs element-binding factor 1* in *Hordeum vulgare* (Figure 5.8)*.* The closest annotated gene is *ZnbZIP11* (*Zea mays*) with 77.62% sequence identity and 74.29% to the *OsbZIP11* (*Oryza sativa* Japonica Group). In *Arabidopsis*, bZIP11 interacts with the auxin pathway via ADA2b adapter proteins to recruit histone acetylation machinery to specific auxin-responsive genes (Weiste & Dröge-Laser, 2014). Based on these similarities, we named the gene identified here as *TabZIP11.*
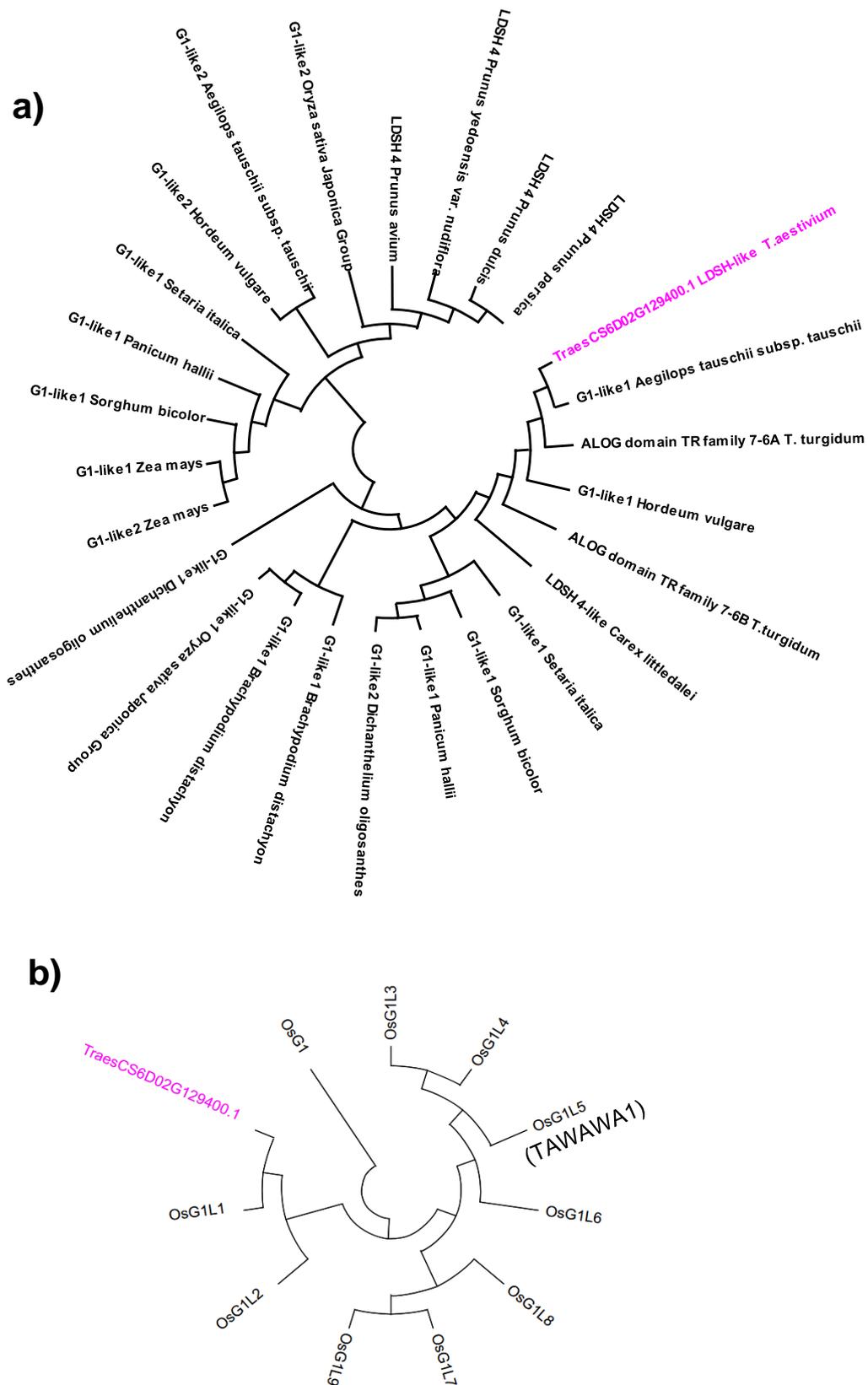
**Figure 5.7 - Phylogenetic analysis of ALOG1 (LDSH)**

**a)** A phylogenetic tree for LDSH (ALOG1; TraesCS6D02G129400.1) protein (magenta). 26 amino acids were compared, obtained through BLAST. **b)** A phylogenetic tree showing the relationship between the *ALOG1* (TraesCS6D02G129400.1) protein (pink) and 10 rice ALOG genes. Both trees are unrooted trees generated using MEGAX, using the maximum likelihood method and Dayhoff matrix based model.
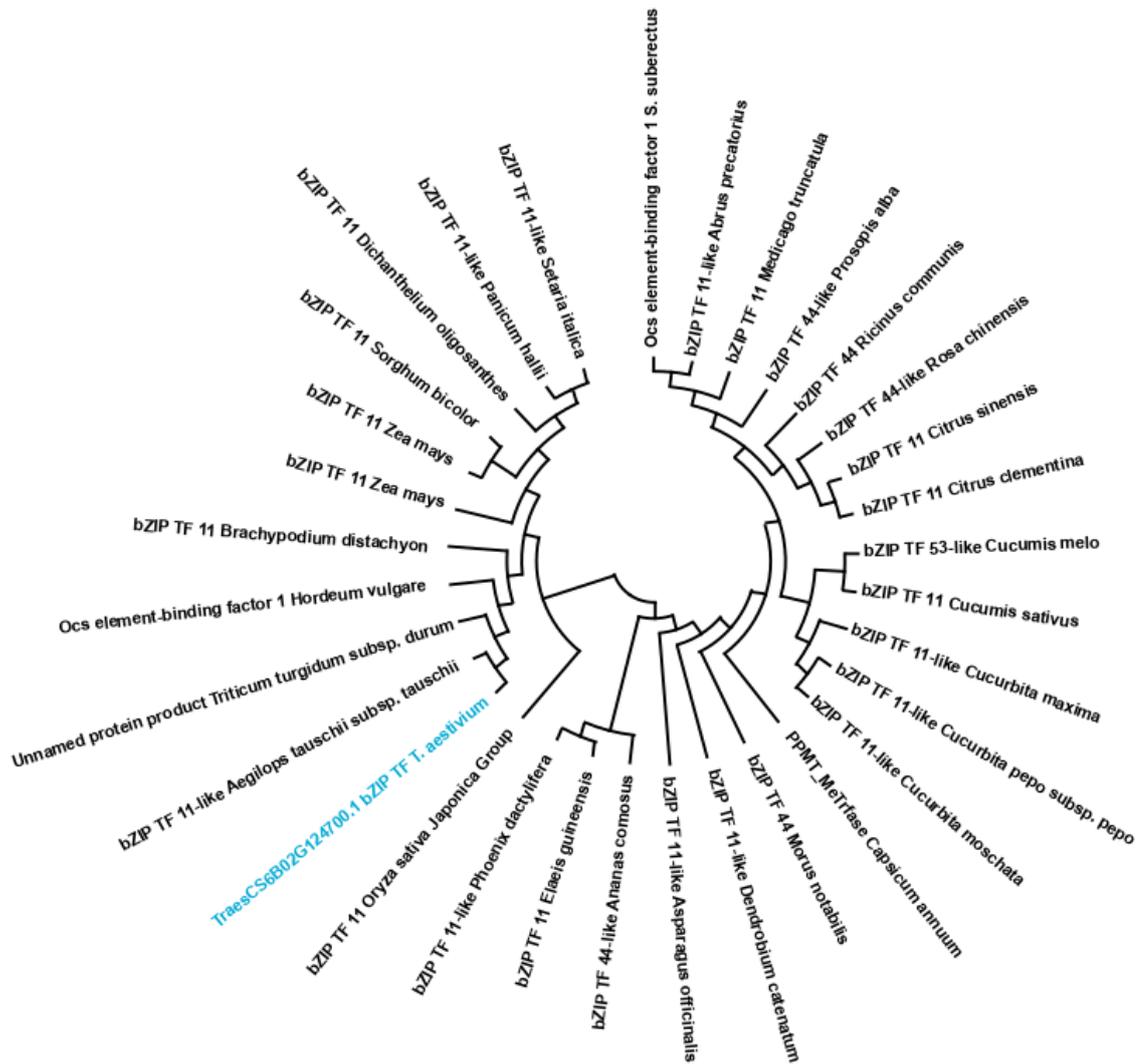
**Figure 5.8 - Phylogenetic analysis of bZIP11 (bZIP)**

A phylogenetic tree for bZIP (TraesCS6B02G124700.1) protein (Cyan). 31 amino acids were compared, obtained through BLAST. An unrooted tree was generated using MEGAX, using the maximum likelihood method and JTT matrix-based model.

### 5.4.6.2  Increased expression of *ALOG1* and *bZIP11* in *ppd-1* NILs.

The *ALOG1* and *bZIP11* genes show striking gene expression profiles (Figure 5.9). *bZIP11* displays high expression throughout inflorescence development with a considerable peak at DR stage (Figure 5.9a). At the DR stage in the *ppd-1* NIL, *bZIP11* is significantly up-regulated relative to both wild-type and the *Ppd-D1a* line, with the gene being down-regulated in the photoperiod-insensitive NIL at all points, relative to wild-type. The peak at the DR stage indicates this stage is the most important for *bZIP11* function. The individual genome copies show very similar expression patterns; however, transcripts of *bZIP-D11* are significantly higher than those of the A and B homoeologues, with peaks of each of the *bZIP-A11*, *bZIP-B11* and *bZIP-D11* copies reaching 91, 131 and 222 TPM in the *ppd-1* NIL, respectively. *ALOG1* transcripts peaks at the VG stage, with a gradual down-regulation to the LP stage, where the expression remains at a comparable level until the TS stage (Figure 5.9b). *ALOG-A1* is the most expressed homoeologue and is significantly higher in the *ppd-1* line, relative to wild-type at DR and TS. This expression profile shows *ALOG1* is negatively regulated during the floral transition.

Based on these profiles, I hypothesised these genes play an important role in *Ppd-1* mediated floral development. I hypothesised that both *bZIP11* and *ALOG1* negatively regulate inflorescence development. As such, when knocked out, the mutant lines would flower faster and produce fewer spikelets. This would reflect the accelerated and delayed inflorescence development observed in the *Ppd-D1a* and *ppd-1* lines, respectively (Figure 2.8).
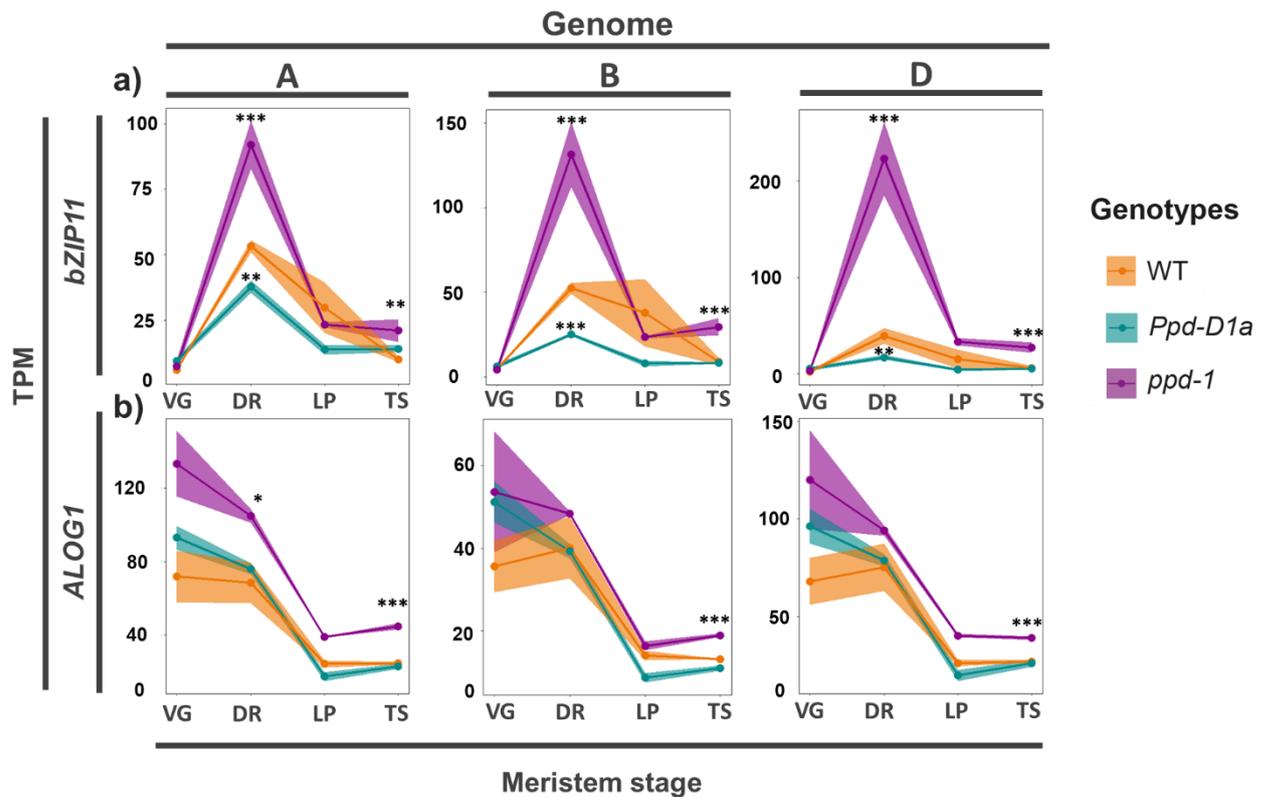
**Figure 5.9 - Expression analysis of *BZIP11* and *ALOG1* during inflorescence development.**

Expression of **a)** *bZIP11* and **b)** *ALOG1*. Data includes expression profiles for wild-type (orange), *Ppd-D1a* photoperiod-insensitive (cyan) and null *ppd-1* (magenta) NILs. Expression shown for the A, B and D genomes for each gene. The stages examined are vegetative (VG), double ridge (DR), lemma primordium (LP) and terminal spikelet (TS). Data are presented as ribbon plots showing mean TPM ± SEM (shaded region) of three biological replicates. *, $P < 0.05$, **, $P < 0.01$, ***, $< P\ 0.005$.

### 5.4.6.3 Protein structure and CRISPR/Cas9 deletion

Clues to the families and functions of *bZIP11* and *ALOG1* can be found based on their amino acid sequence similarities. The ALOG1 protein is 259 aa, with only one predicted domain, an ALOG domain (Figure 5.10). ALOG domains are rich in basic amino acids (e.g. arginine), with four predicted alpha helical domains. ALOG domain containing proteins are plant-specific and highly conserved among land plants, with reported roles in spike architecture, phytochrome-dependent light signalling and the specification of sterile lemma identity (Yoshida *et al.*, 2009, 2013; Takeda *et al.*, 2011). Our CRISPR/Cas9 edit successfully introduced a homogeneous 5 bp deletion at the 30 aa (90bp CDS) site of *ALOG-D1*. The mutation introduced a TAA stop codon, which prematurely ends translation of this gene (Figure 5.10a). bZIP11 (179 aa) also has one predicted domain, the basic-leucine zipper (bZIP) domain. These contain a sequence-specific DNA-binding motif followed by a leucine zipper region, which is required for dimerization. This family has diverse functionality and are found in all eukaryotes. CRISPR/Cas9 mutagenesis introduced a homozygous 2 bp deletion into the B genome at the 62aa (186 CDS) of the gene, introducing a frameshift mutation that encodes a non-functional protein. The altered region of the protein includes a significant portion of the bZIP domain (Figure 5.10b). The CRISPR/Cas9 deletions fall within the core domains of both genes, preventing the production of functional protein.
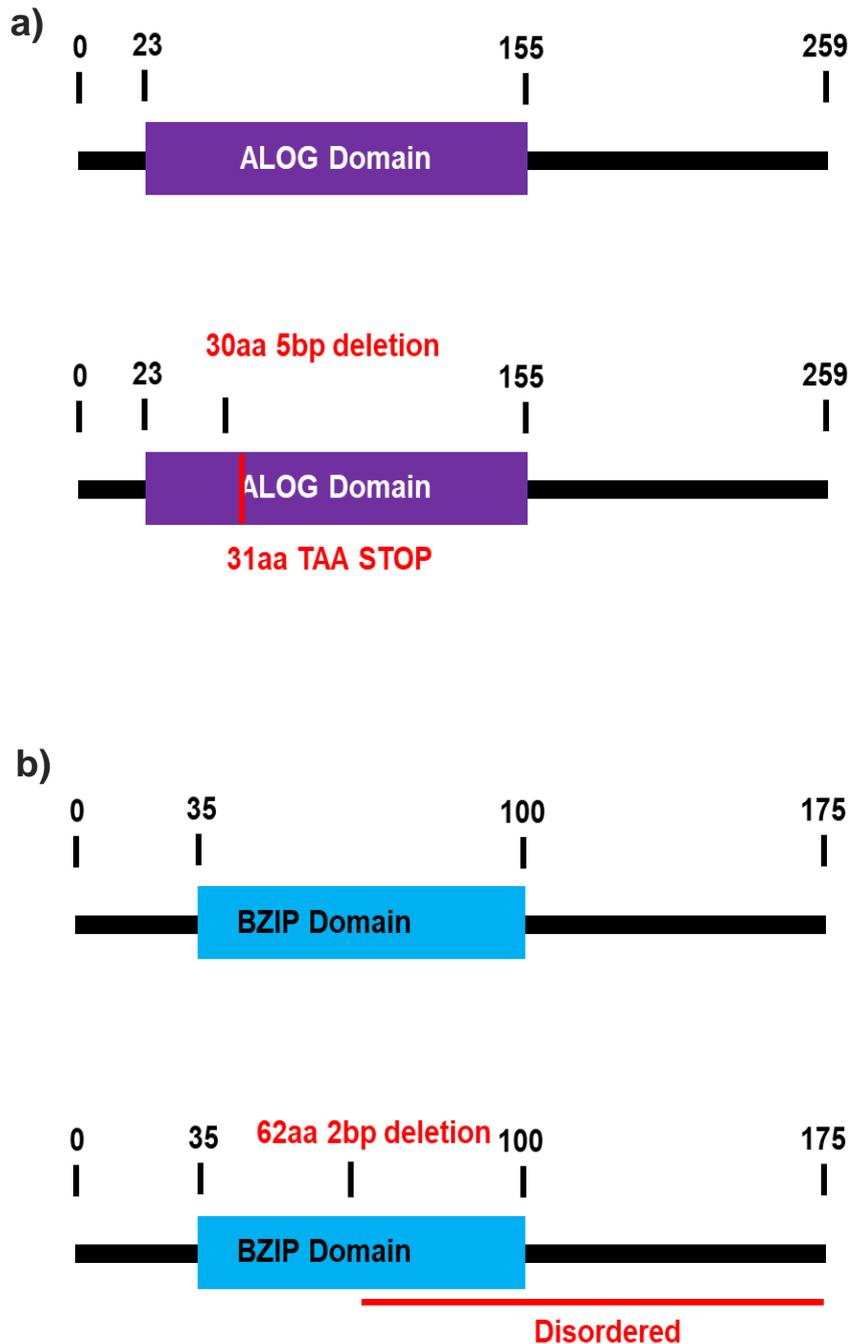
**Figure 5.10 - Predicted protein domains of *ALOG-D1* and *BZIP-B1***

**a)** The protein structure of *ALOG-D1*. Crispr induced homozygous 5bp deletion is shown in red. This mutation causes a premature stop codon at the 31aa position. **b)** The protein structure of *BZIP-B1*. Crispr induced 2bp deletion is shown in red. This mutation causes at the 62aa position causes a disordered protein. Amino acid (aa) positions are denoted in black above each structure.

### 5.4.6.4 Mutant phenotype

To investigate the effect of these CRISPR/Cas9 generated mutants for *bZIP11* and *ALOG1*, the plants were grown under controlled long-day conditions (16 h light/ 8 h dark) alongside a wildtype (Fielder) control. Both mutant lines flowered earlier than wild-type (Figure 5.11)., flowering was accelerated in the *alog-d1* mutant by 11 days, and by 10 days in the *bzip-b11* mutant. Spikelet number is also affected by these mutations, with the wild-type (21.7 ± 0.42 SEM) producing three more spikelets than the *alog-d1* mutant (18 ± 0.4 SEM) and 6 more than the *bzip-b11* mutant (15.3 ± 0.3 SEM). These phenotypes highlight that the function of these genes is dosage-dependent, as disruption of one homoeologue is sufficient to generate flowering and inflorescence development phenotypes. These traits are remarkably similar to those of the photoperiod-insensitive *Ppd-D1a* NIL that displays accelerated flowering and produces fewer spikelets than wild-type (Figure 2.1; 2.8). Future analyses are required to confirm the phenotypes indicated by these preliminary studies – the future work may examine the effect these genes have under short-day conditions, and if these mutants display a photoperiod-insensitive phenotype similar to *Ppd-D1a* lines.
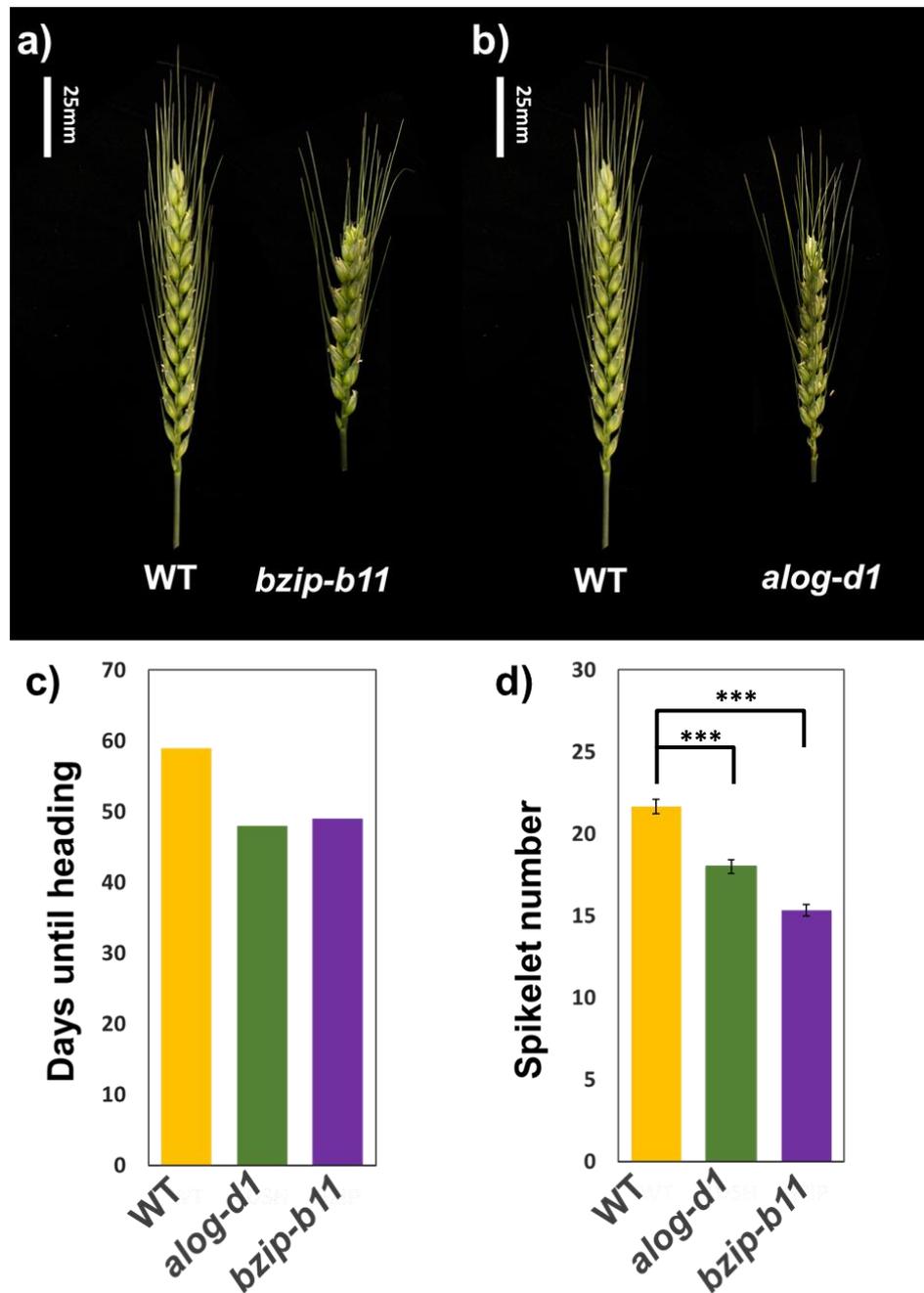
**Figure 5.11 - Phenotypes of *bzip.b11* and *alog-d1* mutants**

**a-b)** Representative inflorescences of the **a)** *bzip-b11* and **b)** *alog-d1* mutants, relative to wild-type (cv. Fielder). (**c-d**) Preliminary analysis of (**c**) flowering-time and (**d**) spikelet number in the *bzip-b11* (purple) and *alog-d1* (green) mutants, relative to wild-type (*cv.* Fielder; yellow).**c)** Heading is recorded when over half of wild-type plants have headed, compared to two confirmed homozygous mutant lines for each gene - all replicates of each line flowered on the same day, so there is no error bars. **d)** Data are the mean $\pm$ SEM of at least 5 inflorescences(multiple spikes per plant, see Chapter 5.6.2). ***, < *P* value of 0.005.

## 5.5 **Discussion**

This chapter presents an unbiased strategy for the characterisation of novel genes in hexaploid wheat where genes are selected based on their expression characteristics and not their annotated function. I use detailed analysis of RNA-sequencing data to find genes most likely to be involved in key meristem transitions, specifically, from the double ridge to lemma primordium stages. Using this method, I have identified 4 novel genes, which until now had no characterised role in inflorescence development.

### *5.5.1 YABBY4*

This study has identified two genes, verified through TILLING mutants. that have a role during inflorescence development, namely *YABBY4* and *TOM1*. The *YABBY* genes are a small family of transcription factors, encoding proteins with two conserved domains, an N-terminal C2C2 zinc-finger motif and a C-terminal YABBY domain (Li *et al.*, 2019). These domains give clues to the importance of these genes, with the zinc finger domain mediating the binding of DNA-motifs in addition to mediating protein-protein interactions (Sawa *et al.*, 1999).  YABBY proteins, including YABBY4 have essential roles in the regulation Arabidopsis ovule polarity. Ovules are precursors to seed and polarity is key to their development. Ovules grow from the abaxial side onward and not the adaxial side, with members of the YABBY family being expressed in the abaxial side of lateral organs. A key role for the YABBY protein, SUPERMAN (SUP), involves polar organ development, with *sup* mutants showing enhanced growth on the adaxial side, to the same degree as the wild-type growth on the abaxial side (Meister *et al.*, 2002; Siegfried *et al.*, 1999). In rice, *YABBY4* plays a key role regulating growth and development through modulation of the gibberellin pathway (Yang *et al.*, 2016). Overexpression of *YABBY4* leads to abnormal development, with a semi-dwarf phenotype and more floral organs. *YABBY* binds the promoter of *GIBBERELLIN 20-OXIDASE 2* gene (*GA20ox2*)*,* itself a direct target of DELLA proteins (Yang *et al.*, 2016). Plants that overexpress *YABBY4* show GA insensitivity, highlighting their importance in the regulation of this key pathway (Yang *et al.*, 2016). These overexpressed *YABBY4* lines show dramatically more stamen and pistils compared to wild-type. This role in rice provides evidence for a role of *TaYABBY4* in floral development, supporting the results we present here.

### *5.5.2 TOM1*

The *TOBAMOVIRUS MULTIPLICATION 1* (*TOM1*) gene is identified to be homologous to *TOM1* genes in many other organisms. However, none of these genes have a reported role in floral meristem development, with mutants in *Arabidopsis* (e.g. *tom1*) identified because they show reduced efficiency of intracellular multiplication of tobamoviruses (Yamanaka *et al.,* 2000; Tsujimoto *et al.,* 2003). In *Arabidopsis, tom1* mutants don't

appear to have a strong impact on plant growth; however, the role of *TOM1* itself has yet to be characterised. The role of this gene in disease resistance presents an interesting opportunity for combining traits associated with disease resistance and developmental traits, a strategy that could be important for increasing overall yields. The balancing of yield traits with resistance genes is an area of active research, with very few defence genes being used in crop breeding because they drastically reduce plant growth (Brown, 2003; Nelson *et al.*, 2018). Interestingly, this goal was achieved in the creation of semi-dwarfing rice. Breeders used a mutation on *GA20ox-2,* involved in *GA* biosynthesis that conferred the dwarfing phenotype (Sasaki *et al.*, 2002; Spielmeyer *et al.*, 2002). However, the GA deficiency caused by this mutation has been shown recently to confer disease resistance (Yang *et al.*, 2008; Qin *et al.*, 2013; de Vleesschauwer *et al.*, 2016).

Taken together, both the *YABBY* and *TOM1* genes identified here have roles in floral development, which until now has not been identified in wheat. These findings demonstrate the effectiveness of the mutant screen described here.

### 5.5.3 bZIP11

This study also identified *bZIP11* and *ALOG1* as two genes with considerable impacts on flowering time and spike architecture, which were analysed genetically using CRISPR/Cas9 mutagenesis. A recent study highlighted the role of two bZIP transcription factors *TGACG-MOTIF BINDING* (*TGA1*) and *TGA4* in inflorescence architecture, flowering, and SAM maintenance of *Arabidopsis*, functioning at the boundary of the developing organs activating 66 genes (Wang *et al.*, 2019). In crops such as maize, a large role for *bZIP* genes have been suggested. For example, *FASCIATED EAR4* (*FEA4*) has a function within the periphery of the meristem to regulate a set of auxin-responsive genes, and it is proposed that *FEA5* may function antagonistically to *KNOTTED1* and *WUSCHEL*, which are two well-described meristem maintenance genes (Pautler *et al.*, 2015). Another study showed through the overexpression of the BZIP transcription factor *ZmbZIP60* in *Arabidopsis*, a gene which is known to have a role in stress tolerance in maize, causes altered expression patterns for *AP1*, *FT* and *CO*. This effect on important floral genes suggests an important role in the flowering pathway, although further characterisation of this gene has not yet been performed (Wang *et al.*, 2012). Some of the most important genes in the flowering process, *FD-like* bZIP transcription factors, assemble into a floral activation complex with FT1 and Hd3a to regulate floral development. A recent study in rice proposes that complexes of bZIP proteins function together to regulate inflorescence development by forming a florigen repressor or activation complexes (Kaneko-Suzuki *et al.*, 2018; Cerise *et al.*, 2020). This is an exciting proposition, as it points to a mechanism where bZIP11 may be part of a florigen repressor complex, whereby deletion of this gene releases its repression and accelerates flowering

(Figure 5.11c). The peak of expression of this gene at the DR stage suggests a stage-specific role in the balancing the promotion of flowering (Figure 5.9). This balancing may ensure the stage transition from DR to LP occurs at the optimal rate for spikelet development.

### *5.5.4 ALOG1*

The ALOG family is named after its earliest identified members (*Arabidopsis LSH1* and *Oryza G1*). The rice genome has a total of 10 ALOG genes, the first identified was *LONG STERILE LEMMA1 (G1)* (Figure 4.7; Li *et al.*, 2019)*.* Mutants in *G1* show large sterile lemma, suggesting that the *G1* is required for the repression of lemma identity in rice and could be involved in morphological modifications during rice evolution (Yoshida *et al.*, 2009). A particularly interesting ALOG gene in rice is *TAWAWA1* (*TAW1*). *TAW1* is a regulator of meristem activity in rice, which promotes inflorescence meristem activity and supresses transition to spikelet meristem identity (Yoshida *et al.*, 2013). In *tawawa1-D* dominant gain-of-function mutants, inflorescence meristem activity is extended, resulting in a delay of spikelet specification. This results in the formation of more branches and consequently an increase in the numbers of spikelets. Reducing the activity of *TAW1* results in abortion of the inflorescence meristem, resulting in smaller inflorescences with less spikelets. The expression of *TAW1* varies based on stage, with the signal gradually decreasing throughout development, and the strongest expression occurring in the branch meristems. Interestingly, it was shown that the higher expression of *TAW1* in the inflorescence correlated with increased severity of phenotype. There are obvious correlations with *TAW1* and the *ALOG1* identified here. In the mutant line developed in this study, I observed a decrease in spikelet number, like that seen in *taw1* null lines. In addition, the expression of *ALOG1* also decreases as the inflorescence develops. Overexpressing *ALOG1* presents interesting possibilities for altering wheat spike architecture in the future. These data not only highlight an effective method for gene discovery, but also shows the influence of *Ppd-1* over gene expression in the developing inflorescence.

Despite the obvious power of the screen, it was limited by time and facilities. In addition, the use of TILLING mutants could eliminate any gene not regulated in a dosage-dependent manner from the screen. The use of CRISPR/Cas9 can bypass this bottleneck by obtaining complete knockout of genes; however, this is very time-consuming and resource intensive. Due to these limitations, many of the genes identified here, but not indicated as positive using TILLING mutants, remain interesting targets for future studies. Technologies are becoming increasingly available in wheat, with recent advances in CRISPR/Cas9 likely paving the way for faster and more efficient mutagenesis. These

advances show promise for mutant screens such as the one identified here (Adamski *et al.*, 2020; Debernardi *et al.*, 2020).

## 5.6 Methods

### 5.6.1 Plant material and growth conditions

The hexaploid wheat used in this study included: The wild-type Paragon, photoperiod insensitive *Ppd-D1a* and *ppd-1* null NILs previously described in (Chapter 2.5.1). Cadenza EMS-induced TILLING lines listed in (Table 5.4; Krasileva *et al.*, 2017). For Phase 3 (Figure 5.1c), all tilling lines listed in (Table 5.4) were grown in a controlled environment room (CER) maintained at 16 h light / 8 h dark, photoperiod at 300 µmol/m2/s (using Plantastar 400-W HQI bulbs (Osram) and Maxim 60-W tungsten bulbs), under a day temperature of 20 °C and a night temperature of 15 °C until wild-type Cadenza plants reached the TS stage at which point plants were dissected and imaged using previous described methods (Chapter 2.5.1). Wild-type *cv.* Fielder, *alog-d1* and *bzip11-b1* plants were also grown under the same CER conditions until fully developed.

### 5.6.2 CRISPR/Cas9 mutant generation

Fielder *alog-d1* and *bzip11-b1* null lines were generated using CRISPR/Cas9 with BRACT (John Innes Centre, Norwich, UK). The guides and guide location are shown in (Figure 5.12). All plants shown are from a T1 generation (one generation from transformed plants). Two homozygous *alog-d1* and *bzip-b11* mutants were identified, data presented is product of multiple spikes. Heterozygous data (unshown) confirms phenotype, future experiments are being carried out to increase replication. Mutants were verified using deep-sequencing in collaboration with (Floodlight Genomics LLC, Tennessee, USA).
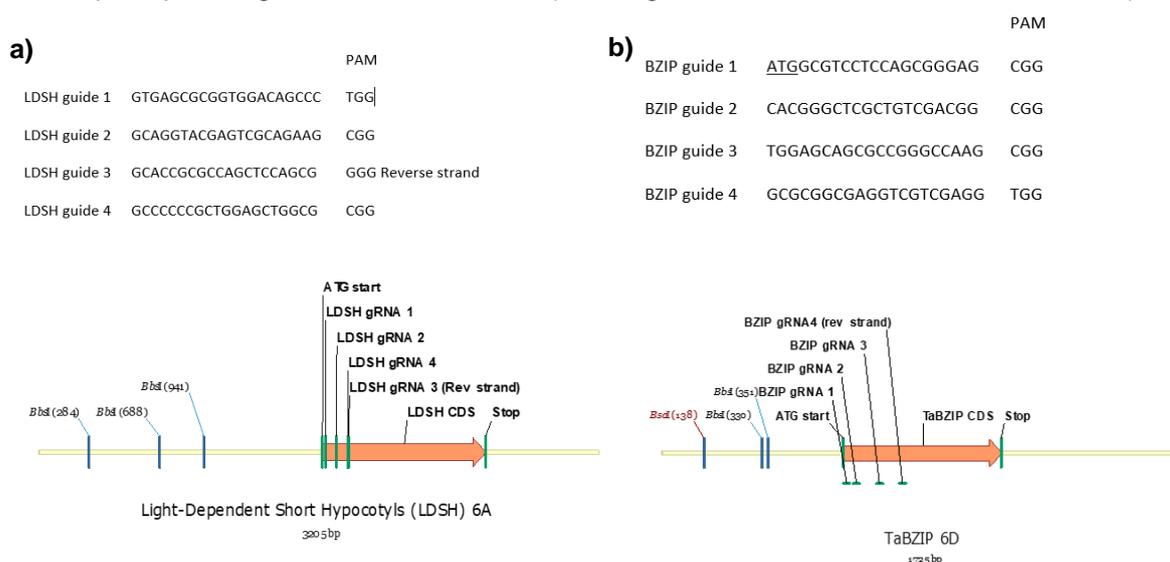


**Figure 5.12 - CRISPR/Cas9 guides for LDSH (*ALOG1*) and BZIP (*bZIP11*).**
Guides and gene guide location for **a)** LDSH (*ALOG1*) and **b)** BZIP (*bZIP11*) *CRISPR/Cas9* screen. Guide location is marked on gene (orange arrow).

## 5.6.3 DNA extractions and sequence analysis.

Genomic DNA for genotyping was extracted as described in in Chapter 2.5.8.

## 5.6.4 Kompetitive Allele-Specific PCR analysis

To analyse TILLING mutant lines, Kompetitive allele-specific PCR (KASP) analysis was performed on extracted gDNA. Oligonucleotides to verify mutant were designed using Polymarker (Ramirez-Gonzalez *et al.*, 2015), these oligos contain wither FAM or HEX compatible tails (FAM tail, 5'-GAAGGTGACCAAGTTCATGCT-3' HEX tail, 5.-GAAGGTCGGAGTCAACGG -ATT-3') (Table 5.5). The KASP assay was then performed as outlined in Chapter 2.6.9.

## 5.6.5 Inflorescence architecture measurements

For fully developed inflorescences all spikelets both viable and inviable spikelets were counted. For developing inflorescences examined in Phase 3 of the screen (Figure 5.1c) spikelet number was counted when first visible at the double ridge stage (DR), at this stage the spikelet meristem ridge (the upper ridge) was counted as a spikelet. As the inflorescence develops the spikelet meristem becomes more pronounced and is visibly defined at the lemma primordium (LP) and terminal spikelet (TS) stages where it is counted as a spikelet. To identify inflorescence meristem stage a detail scale was designed based on the stages defined in (Kirby & Appleyard, 1984) shown in (Figure 5.13). For fully developed inflorescences final spikelet numbers, both viable and inviable were counted. For mature plant phenotyping, data are the average ± SEM of at least 5 replicate inflorescences. Heading date for CRISPR/Cas9 mutant lines was defined when 2 plants per line had emerged 50% from the boot, relative to wild-type. Wild-type (*cv.* Fielder) data included at least 10 replicate plants.
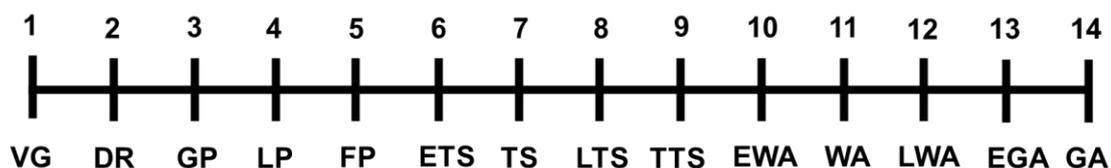


**Figure 5.13 - Meristem development stages for Phase 3 dissection screen**
Schematic representing the stages sampled for the phase 3 screen, stages are presented below the line with the corresponding stage above. 1= vegetative (VG), 2= double ridge (DR), 3= glume primordium (GP), 4= lemma primordium (LP), 5= floret primordium (FP), 6= early terminal spikelet (ETS), 7 = terminal spikelet (TS), 8= late terminal spikelet (LTS), 9= terminal terminal spikelet (meristem is about transition to WA; TTS):, 10= early white anther, 11= white anther, 12= late white anther, 13= early green anther and 14= green anther.

### 5.6.6 Phylogenetic analyses

Homology for ALOG1, bZIP11, YABBY and ALOG family amino acids were obtained by BLAST (Basic Local Alignment Search Tool) analysis using Ensembl Plants. The resulting amino acid sequences were aligned using MAFFT v.7.123b alignment algorithm with 4 GUIDANCE v.2.0, 100 bootstrap replicates. An unrooted tree was then generated using MEGAX, with the maximum likelihood method. The Dayhoff matrix based model (Kumar *et al.*, 2018; Schwarz & Dayhoff, 1979) was used for LDSH and ALOG family genes. The JTT matrix-based model was used for bZIP and YABBY genes (Jones *et al.*, 1992). For all, positions with less than 95% site coverage were eliminated, i.e., fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position.

### 5.6.7 Statistical analysis

Statistical differences between treatments and sample points were determined by two-tailed Students *t* test. The even distribution of data was determined using a Barlett's test. Data in figures are mean ± standard error of the mean (SEM)

## Table 5.4 - Cadenza TILLING lines used in Chapter 5

| Up | Gene ID | TILLING line 1 | TILLING line 2 | TILLING line 3 |
|---|---|---|---|---|
| 1 | TraesCS6D02G287800.1 | Cadenza0129.chr6D.397082360 | Cadenza0273.chr6D.397083815 | Cadenza0586.chr6D.397082600 |
| 2 | TraesCS4B02G226500.1 | Cadenza1371.chr4B.473568922 | Cadenza1051.chr4B.473571596 | |
| 3 | TraesCS2B02G326400.1 | Cadenza0127.chr2B.466193624 | Cadenza0188.chr2B.466192620 | |
| 4 | TraesCS7B02G152200.1 | Cadenza1806.chr7B.202359987 | Cadenza1775.chr7B.202359840 | Cadenza1409.chr7B.202359891 |
| 5 | TraesCS7D02G254100.1 | Cadenza1644.chr7D.230942192 | Cadenza0232.chr7D.230942228 | Cadenza1511.chr7D.230942306 |
| 6 | TraesCS6A02G096300.1 | Cadenza1842.chr6A.63609863 | | |
| 7 | TraesCS6B02G124700.1 | Cadenza1597.chr6B.119848827 | | |
| 8 | TraesCS6D02G087400.1 | Cadenza1803.chr6D.52723954 | | |
| 9 | TraesCS6A02G347800.3 | Cadenza0336.chr6A.580199004 | Cadenza0053.chr6A.580198404 | |
| 10 | TraesCS4D02G080100.1 | Cadenza1357.chr4D.54127289 | | |
| 11 | TraesCS3D02G482600.1 | Cadenza1442.chr3D.580089052 | Cadenza0097.chr3D.580088966 | |
| 12 | TraesCS2B02G497500.1 | Cadenza0423.chr2B.694055848 | Cadenza1701.chr2B.694057145 | |
| 13 | TraesCS1D02G343400.1 | Cadenza0132.chr1D.432406319 | Cadenza0353.chr1D.432406244 | Cadenza1523.chr1D.432408956 |
| 14 | TraesCS5D02G526300.1 | Cadenza0346.chr5D.544231711 | Cadenza2014.chr5D.544232157 | Cadenza0676.chr5D.544232500 |
| 15 | TraesCS6A02G139700.1 | Cadenza1785.chr6A.114458746 | | |
| 16 | TraesCS6B02G168300.1 | Cadenza1468.chr6B.178273107 | Cadenza1736.chr6B.178273187 | Cadenza1576.chr6B.178272551 |
| 17 | TraesCS6D02G129400.1 | Cadenza1976.chr6D.94602485 | Cadenza0196.chr6D.94602507 | Cadenza0594.chr6D.94602570 |
| 18 | TraesCS4B02G282200.1 | Cadenza1169.chr4B.565313306 | Cadenza1467.chr4B.565313118 | Cadenza0643.chr4B.565313168 |
| 19 | TraesCS4D02G280900.1 | Cadenza1265.chr4D.451801982 | Cadenza1770.chr4D.451802495 | Cadenza0902.chr4D.451801914 |
| 20 | TraesCS4D02G294100.1 | Cadenza1586.chr4D.464533453 | Cadenza1675.chr4D.464536313 | |
| 21 | TraesCS7A02G414300.1 | Cadenza1471.chr7A.604853538 | Cadenza1367.chr7A.604853528 | Cadenza2109.chr7A.604853945 |
| 22 | TraesCS3D02G425800.1 | Cadenza1763.chr3D.538401404 | Cadenza1494.chr3D.538402702 | Cadenza0551.chr3D.538402629 |
| 23 | TraesCS6A02G237700.1 | Cadenza0283.chr6A.446901749 | Cadenza0687.chr6A.446903009 | |
| 24 | TraesCS6D02G220400.1 | Cadenza0246.chr6D.310521429 | Cadenza1172.chr6D.310521399 | Cadenza0351.chr6D.310521808 |
| 25 | TraesCS1A02G090000.4 | Cadenza0946.chr1A.79702543 | Cadenza1159.chr1A.79702403 | Cadenza1025.chr1A.79701438 |
| 26 | TraesCS1A02G208600.2 | Cadenza1655.chr1A.370104837 | Cadenza0883.chr1A.370106183 | Cadenza0266.chr1A.370107390 |
| 27 | TraesCS1A02G411600.1 | Cadenza1773.chr1A.572208563 | Cadenza0888.chr1A.572209408 | Cadenza0203.chr1A.572209799 |
| 28 | TraesCS1A02G423100.1 | Cadenza1713.chr1A.578517767 | Cadenza2085.chr1A.578518837 | |
| 29 | TraesCS1B02G154100.3 | Cadenza0467.chr1B.255506318 | Cadenza0111.chr1B.255510967 | |
| 30 | TraesCS1D02G203300.2 | Cadenza1220.chr1D.287457310 | Cadenza0884.chr1D.287467826 | |
| 31 | TraesCS1D02G419400.1 | Cadenza1289.chr1D.476439440 | Cadenza0105.chr1D.476440573 | |
| 32 | TraesCS1D02G431700.1 | Cadenza0743.chr1D.481695532 | Cadenza1262.chr1D.481693568 | Cadenza0510.chr1D.287457324 |
| 33 | TraesCS2A02G000100.2 | Cadenza0478.chr2A.251529 | Cadenza0988.chr2A.254521 | |
| 34 | TraesCS2B02G285900.1 | Cadenza1646.chr2B.393856324 | Cadenza1522.chr2B.393856629 | Cadenza1735.chr2B.393856701 |
| 35 | TraesCS2B02G326200.1 | Cadenza0723.chr2B.465826242 | Cadenza1277.chr2B.465828138 | |
| 36 | TraesCS2B02G392100.1 | Cadenza0732.chr2B.556517351 | Cadenza2056.chr2B.556517614 | Cadenza2110.chr2B.556517677 |
| 37 | TraesCS2D02G101800.5 | Cadenza0638.chr2D.55058720 | Cadenza0638.chr2D.55058720 | |
| 38 | TraesCS2D02G262700.1 | Cadenza1929.chr2D.318916652 | Cadenza1929.chr2D.318916652 | |
| 39 | TraesCS2D02G345200.2 | Cadenza0725.chr2D.441526925 | Cadenza1129.chr2D.441527381 | |
| 40 | TraesCS3B02G162000.2 | Cadenza0352.chr3B.158718848 | Cadenza0122.chr3B.158718931 | Cadenza1655.chr3B.158715883 |
| 41 | TraesCS3D02G226300.1 | Cadenza1491.chr3D.308057589 | Cadenza0076.chr3D.308057619 | Cadenza0927.chr3D.308058937 |
| 42 | TraesCS3D02G231600.1 | Cadenza1073.chr3D.317199209 | Cadenza0988.chr3D.317199223 | |
| 43 | TraesCS4B02G220700.2 | Cadenza0404.chr4B.464157276 | Cadenza0510.chr4B.464157489 | |
| 44 | TraesCS4B02G261100.1 | Cadenza1313.chr4B.529676961 | Cadenza1313.chr4B.529676961 | Cadenza0225.chr4B.529677558 |
| 45 | TraesCS4D02G002000.1 | Cadenza1181.chr4D.1241952 | Cadenza1597.chr4D.124201 | Cadenza0879.chr4D.1242032 |
| 46 | TraesCS5B02G220000.1 | Cadenza2019.chr5B.394267358 | Cadenza1254.chr5B.394266468 | Cadenza2070.chr5B.394266550 |
| 47 | TraesCS5D02G123700.1 | Cadenza1815.chr5D.184027816 | Cadenza1508.chr5D.184029271 | Cadenza0703.chr5D.184039248 |

## Table 5.5 - KASP primers used in Chapter 5

| Gene of interest | Tilling lines | KASP primer |
|---|---|---|
| Auxin Efflux Carrier Com1 | Cadenza0586.chr6D.397082600 | gaaggtgaccaagttcatgctGTCATCACGTGGTAGAAGTC |
| Auxin Efflux Carrier Com SNP | Cadenza0586.chr6D.397082600 | gaaggtcggagtcaacggattGTCATCACGTGGTAGAAGTT |
| Auxin Efflux Carrier Com rev | Cadenza0586.chr6D.397082600 | CCAAGATGATCACGGGCAC |
| Auxin Efflux Carrier Com2 | Cadenza0129.chr6D.397082360 | gaaggtgaccaagttcatgctGCTCATGATGCTCGCCATGC |
| Auxin Efflux Carrier Com SNP | Cadenza0129.chr6D.397082360 | gaaggtcggagtcaacggattGCTCATGATGCTCGCCATGT |
| Auxin Efflux Carrier Com rev | Cadenza0129.chr6D.397082360 | GATGGTCCACTCGAGGCTG |
| CTP synthase | Cadenza1773.chr1A.572208563 | gaaggtgaccaagttcatgctcagtgcttacctttagtagaccC |
| CTP synthase SNP | Cadenza1773.chr1A.572208563 | gaaggtcggagtcaacggattcagtgcttacctttagtagaccT |
| CTP synthase RV | Cadenza1773.chr1A.572208563 | TGCTAAAGAGGTCAGCACG |
| Cullin-1 | Cadenza1159.chr1A.79702403 | gaaggtgaccaagttcatgctccaagaacgtgagggtgaaC |
| Cullin-1 SNP | Cadenza1159.chr1A.79702403 | gaaggtcggagtcaacggattccaagaacgtgagggtgaaT |
| Cullin-1 RV | Cadenza1159.chr1A.79702403 | CTGCCTAAGCCAATCTCAACA |
| F-box / LRR-repeat protein | Cadenza0467.chr1B.255506318 | gaaggtgaccaagttcatgctgctcatgtatcactcaattgtcctC |
| F-box / LRR-repeat protein SNP | Cadenza0467.chr1B.255506318 | gaaggtcggagtcaacggattgctcatgtatcactcaattgtcctT |
| F-box / LRR-repeat protein RV | Cadenza0467.chr1B.255506318 | TGACATATCTAGTGACGCTAACAA |
| Histone-lysine N-methyltransferase setd3 | Cadenza1815.chr5D.184027816 | gaaggtgaccaagttcatgctagcccattcgtgagtagcC |
| Histone-lysine N-methyltransferase setd3 SNP | Cadenza1815.chr5D.184027816 | gaaggtcggagtcaacggattagcccattcgtgagtagcT |
| Histone-lysine N-methyltransferase setd3 RV | Cadenza1815.chr5D.184027816 | TCCTTTCCTCCTCCGGGG |
| Histone-lysine N-methyltransferase setd3 | Cadenza0703.chr5D.184039248 | gaaggtgaccaagttcatgctttggaaccaaagaaagaagttgC |
| Histone-lysine N-methyltransferase setd3 SNP | Cadenza0703.chr5D.184039248 | gaaggtcggagtcaacggattttggaaccaaagaaagaagttgT |
| Histone-lysine N-methyltransferase setd3 RV | Cadenza0703.chr5D.184039248 | TGAAGACAGCCATGCAGCAT |
| Tobamovirus multiplication 1 | Cadenza1646.chr2B.393856324 | gaaggtgaccaagttcatgcttgacctcctcccaccatcC |
| Tobamovirus multiplication 1 SNP | Cadenza1646.chr2B.393856324 | gaaggtcggagtcaacggatttgacctcctcccaccatcT |
| Tobamovirus multiplication 1 RV | Cadenza1646.chr2B.393856324 | AGCTGGTCTCCTCCTCCT |
| Tobamovirus multiplication 1 | Cadenza1735.chr2B.393856701 | gaaggtgaccaagttcatgctccacctcatgaacttcgtcG |
| Tobamovirus multiplication 1 SNP | Cadenza1735.chr2B.393856701 | gaaggtcggagtcaacggattccacctcatgaacttcgtcA |
| Tobamovirus multiplication 1 RV | Cadenza1735.chr2B.393856701 | AATTAACATTGATTACATTGCCCCT |
| Tobamovirus multiplication 1 | Cadenza1522.chr2B.393856629 | gaaggtgaccaagttcatgcttctgcgctggatcctcaC |
| Tobamovirus multiplication 1 SNP | Cadenza1522.chr2B.393856629 | gaaggtcggagtcaacggatttctgcgctggatcctcaT |
| Tobamovirus multiplication 1 RV | Cadenza1522.chr2B.393856629 | CGAGTTCGATTTTGAGGCCG |
| Squamosa promoter binding-like protein | Cadenza1763.chr3D.538401404 | gaaggtgaccaagttcatgctCAAGGTCTGCGAGGCGCACTC |
| Squamosa promoter binding-like protein SNP | Cadenza1763.chr3D.538401404 | gaaggtcggagtcaacggattCAAGGTCTGCGAGGCGCACTT |
| Squamosa promoter binding-like protein RV | Cadenza1763.chr3D.538401404 | GTACCTGCTGCACTGCTG |
| Squamosa promoter binding-like protein | Cadenza1494.chr3D.538402702 | gaaggtgaccaagttcatgctCAACCGGCGCCGCCGGAAGCC |
| Squamosa promoter binding-like protein SNP | Cadenza1494.chr3D.538402702 | gaaggtcggagtcaacggattCAACCGGCGCCGCCGGAAGCT |
| Squamosa promoter binding-like protein RV | Cadenza1494.chr3D.538402702 | CCGTGGTGATTAGCGAACAA |
| Auxin-induced in root cultures protein 12 | Cadenza1775.chr7B.202359840 | gaaggtgaccaagttcatgctGCCGCTCGGCAAGGCCTCCAC |
| Auxin-induced in root cultures protein 12 SNP | Cadenza1775.chr7B.202359840 | gaaggtcggagtcaacggattGCCGCTCGGCAAGGCCTCCAT |
| Auxin-induced in root cultures protein 12 RV | Cadenza1775.chr7B.202359840 | CGTAGAGCCGGATCTTGCC |
| Auxin-induced in root cultures protein 12 | Cadenza1806.chr7B.202359987 | gaaggtgaccaagttcatgctcgaaagcgtgcttgtcgG |
| Auxin-induced in root cultures protein 12 SNP | Cadenza1806.chr7B.202359987 | gaaggtcggagtcaacggattcgaaagcgtgcttgtcgA |
| Auxin-induced in root cultures protein 12 RV | Cadenza1806.chr7B.202359987 | CACAGCGGGATGAAGGCG |
| Auxin-induced in root cultures protein 12 | Cadenza0232.chr7D.230942228 | gaaggtgaccaagttcatgctgcaagatccggctctacgG |
| Auxin-induced in root cultures protein 12 SNP | Cadenza0232.chr7D.230942228 | gaaggtcggagtcaacggattgcaagatccggctctacgA |
| Auxin-induced in root cultures protein 12 RV | Cadenza0232.chr7D.230942228 | CACCGCCTTCATCCCGCT |

| | | |
|---|---|---|
| Lipid transfer protein | [Cadenza1169.chr4B.565313306](#) | CGCCTCCCTCGCGCTCGCCC |
| Lipid transfer protein SNP | [Cadenza1169.chr4B.565313306](#) | CGCCTCCCTCGCGCTCGCCT |
| Lipid transfer protein RV | [Cadenza1169.chr4B.565313306](#) | GGTAGGAATAGGAGGGAGACG |
| Lipid transfer protein | [Cadenza1265.chr4D.451801982](#) | gaaggtgaccaagttcatgctggtcaaggacaaggacgacC |
| Lipid transfer protein SNP | [Cadenza1265.chr4D.451801982](#) | gaaggtcggagtcaacggattggtcaaggacaaggacgacT |
| Lipid transfer protein RV | [Cadenza1265.chr4D.451801982](#) | GGGTCTAGGTAGTAGGTAGGAG |
| YABBY protein | [Cadenza0246.chr6D.310521429](#) | gaaggtgaccaagttcatgctaccagtgggcccatttcC |
| YABBY protein SNP | [Cadenza0246.chr6D.310521429](#) | gaaggtcggagtcaacggattaccagtgggcccatttcT |
| YABBY protein RV | [Cadenza0246.chr6D.310521429](#) | GCAGTGGTAGATAGGCTAAGGA |
| YABBY protein | [Cadenza1172.chr6D.310521399](#) | gaaggtgaccaagttcatgcttgaaggtcctcttgaatccctG |
| YABBY protein SNP | [Cadenza1172.chr6D.310521399](#) | gaaggtcggagtcaacggatttgaaggtcctcttgaatccctA |
| YABBY protein RV | [Cadenza1172.chr6D.310521399](#) | GCTAATCCAAACCTGCTCATG |
| YABBY protein | [Cadenza0351.chr6D.310521808](#) | gaaggtgaccaagttcatgctcggaggagagattttacattctttG |
| YABBY protein SNP | [Cadenza0351.chr6D.310521808](#) | gaaggtcggagtcaacggattcggaggagagattttacattctttA |
| YABBY protein RV | [Cadenza0351.chr6D.310521808](#) | TTTGCGACAGGGATGAGATT |
| Auxin response factor | [Cadenza0188.chr2B.466192620](#) | gaaggtgaccaagttcatgctggcgttttgtgtttcaggaC |
| Auxin response factor SNP | [Cadenza0188.chr2B.466192620](#) | gaaggtcggagtcaacggattggcgttttgtgtttcaggaT |
| Auxin response factor RV | [Cadenza0188.chr2B.466192620](#) | CCTCCGCGTGACAGGATTAT |
| F-box / LRR-repeat protein | [Cadenza0111.chr1B.255510967](#) | gaaggtgaccaagttcatgctatggtcttccccgcatgG |
| F-box / LRR-repeat protein SNP | [Cadenza0111.chr1B.255510967](#) | gaaggtcggagtcaacggattatggtcttccccgcatgA |
| F-box / LRR-repeat protein Generic | [Cadenza0111.chr1B.255510967](#) | GCCTCCTGCATCAGCGTC |
| Auxin response factor | [Cadenza1816.chr2B.466191585](#) | gaaggtgaccaagttcatgctccatgggcagattctgattgG |
| Auxin response factor | [Cadenza1816.chr2B.466191585](#) | gaaggtcggagtcaacggattccatgggcagattctgattgA |
| Auxin response factor | [Cadenza1816.chr2B.466191585](#) | GCACCACTAAAAGCACTCCG |

211

# Chapter 6    General discussion

The aim of this thesis was to understand how the leaf and developing inflorescence detect and respond, both developmentally and at the level of gene expression, to the changing seasons, and to investigate how these factors interact to coordinate flowering. The key findings from this thesis are:

- Discrete photoperiods underpin the initiation of flowering and the control of yield-based traits. This occurs as days become longer from winter into spring and involves a step-wise increase in transcription of *FT1* (**Chapter 2**).

- *Ppd-1* has a dramatic effect on gene expression in the developing meristem (**Chapter 3**).

- Ambient temperature influences inflorescence development beyond the vernalisation pathway (**Chapter 4**).

- Several genes, not previously associated with inflorescence development, have been identified to be involved in *Ppd-1* mediated inflorescence development through a rapid mutant screen (**Chapter 5**).

## 6.1 How *Ppd-1* and *FT1* respond to the changing seasons

In Chapter 2, I examined how *Ppd-1* and *FT1* respond to the changing seasons. I detected a robust gene expression profile for *Ppd-1* across all photoperiods tested. I also noted how the economically important photoperiod insensitive lines confer their earlier flowering characteristic (Beales *et al.*, 2007; Shaw *et al.*, 2012). The deletion in the promoter region causes mis-regulation in the evening, but normal expression during the day. Beyond *Ppd-1*, I have observed a step-wise induction of *FT1* that mediates transition through key meristem development stages. First, *FT1* triggers the vegetative to floral transition during the 10 to 11 h daylengths. The plant then remains at the DR stage until a second induction of *FT1* at the 12-13 h daylength transition which is associated with its development beyond the DR stage to LP. This step-wise induction allows a precise seasonal control over the developing inflorescence, likely allowing sufficient time for spikelet formation and ensuring key developmental structures, such as florets, initiate under optimal environmental conditions (Coventry *et al.*, 1993; Kirkegaard *et al.*, 2014; Boden *et al.*, 2015; Ochagavía *et al.*, 2018; Prieto *et al.*, 2018). In addition, this step-wise induction is only apparent in the field, illustrating the importance of field-based studies.

## 6.2 Regulation of *Ppd-1*

Little is known about the regulation of *Ppd-1.* In Chapter 2, I demonstrate how the diel expression of *Ppd-1* is required for its function, with the photoperiod insensitive lines accelerating flowering through ectopic night expression. I observed consistent *Ppd-1*

expression profiles as the seasons progress both in terms of pattern and amplitude, and yet *Ppd-1* dynamically integrates signals from the environment into the flowering pathway. It is well documented that *FT1* is downstream of *Ppd-1* (Chapter 2; Turner *et al.*, 2005; Yan *et al.*, 2006; Beales *et al.*, 2007; Wilhelm *et al.*, 2009; Kitagawa *et al.*, 2012; Boden *et al.*, 2015; Bratzel & Turck, 2015). However, the pathway upstream of *Ppd-1* is unknown. Studies have suggested the expression of *Ppd-1* is under the control of PHYC, with *phyC* knockouts in tetraploid wheat flowering later under long days and expressing *Ppd-1* significantly less than wild-type (Chen *et al.*, 2014).

In Chapter 2, I have presented evidence for a self-regulatory negative feedback loop regulating *Ppd-1* expression by describing increases of *Ppd-D1* transcripts when there is no functional *Ppd-D1* protein. Further evidence for a negative feedback loop comes in the identification of a potential *Ppd-1* binding site in its own promoter. This presents a promising area for future study, investigating if *Ppd-1* can interact with its own promoter. Mis-regulation of *Ppd-1* in the *Ppd-D1a* photoperiod insensitive line manifests through a 2.09 kb deletion in its promoter region (Beales *et al.*, 2007; Shaw *et al.*, 2013). This deletion contains promoter binding sites for the clock genes *CCA1* and *TOC1*, in addition to *PIF3* and is sufficient to mediate photoperiod insensitive expression. Consequently, this misregulation leads to dramatically increased expression of *FT1* over a 24 h period, and not just at night. Investigating if specifically deleting the promoter region binding sites for the clock genes and *PIF* genes affects *Ppd-1* expression could perhaps tailer the *Ppd-D1a* phenotype. I hypothesise this strategy will uncouple the two primary characteristic of photoperiod insensitive lines, the earlier induction of *FT1*, likely mediated by photoreceptors through *PIF3*, from the increased *FT1* expression levels likely mediated by the clock genes. This strategy may yield earlier flowering photoperiod insensitive plants without the decrease in yields associated with photoperiod insensitivity.

The nature of the *Ppd-D1* promoter deletion allows us to conclude that the difference in *Ppd-1* function between photoperiod insensitive lines and wild-type isn't based on altered protein structure, but instead is because of two possibilities:

1. Presence of *Ppd-1* during the night when it should be restricted to the day.

2. Increased protein abundance during the night that causes increased protein abundance during the day.

Considering the increase of *FT1* transcripts throughout the day in addition to the night, it is likely the protein is capable of accumulating during the night with consequences during the day. Because of the diurnal expression pattern of *Ppd-1*, it is tempting to look for an external coincidence model, such as the model for *CO* regulation in *Arabidopsis* to explain its regulation (Suárez-López *et al.*, 2001; Yanovsky & Kay, 2002; Andrés & Coupland, 2012). Like *CO, Ppd-1* has a diurnal expression pattern with a peak at dusk and a

subsequent drop in transcripts in the evening. If like CO, Ppd-1 is stabilised in the light or is translocated to the nucleus in response to light stimulus, sufficiently long daylengths could be required to bring *Ppd-1* protein over an activating threshold. Long daylengths or ectopic night expression are sufficient to do so. This theory would require future investigation, testing protein abundance during the night in both the photoperiod sensitive and insensitive lines in addition to examining light dependent localisation.

It is likely that *Ppd-1* is being regulated on a protein level in addition to transcriptionally. The predicted structure of Ppd-1 contains a CCT domain, which is likely to be involved in protein-protein interactions (Turner *et al.*, 2005; Beales *et al.*, 2007). I have shown in Chapter 2 that Ppd-1 localises to the nucleus. However, future studies may investigate if this localisation occurs in response to environmental stimulus. Downstream of *Ppd-1* it influences the expression of *FT1.* However, whether this influence is direct, through Ppd-1 binding the promoter region or protein-protein interaction is unclear. The groundwork to answer these questions was laid during my PhD. *Ppd-1* and *FT1* genes were cloned and their protein were expressed and purified to high quality from bacterial cells. In addition, transgenic GFP tagged Ppd-1 lines were obtained. However, the experiments to investigate the core regulation of Ppd-1 have not fallen within the scope of this PhD. Nevertheless, they provide exciting prospects for future studies.

## 6.3 A comprehensive examination of the genes underpinning inflorescence development

In Chapter 2, I characterised the precise control over the developing meristem as coordinated by gene expression pathways in the leaf. In Chapter 3, I investigated the influence of these pathways in the developing inflorescence using *Ppd-D1a* and *ppd-1* NILs. Previous studies have shown how varying *FT1* expression can radically alter inflorescence architecture, through suppression of meristem identity genes (Boden *et al.*, 2015; Dixon *et al.*, 2018). Based on these studies, I hypothesised that the most significant impact on transcription in the meristem would be focused to the VG to DR transition, as this is where previous reports had highlighted the greatest influence of *FT* (Wigge *et al.*, 2005; Abe *et al.*, 2019). However, in Chapter 2, I show that a step-wise induction of *FT1* influences multiple stage transitions, not only the vegetative to floral transition. This result shows an intimate relationship between the leaf and the meristem, beyond the floral transition, whereby the leaf dynamically relays signals from the environment to the inflorescence throughout development. A key aim of Chapter 3 was to understand the transcriptomics underpinning inflorescence development. This analysis has allowed us to capture the stage specific regulation, beyond the VG to DR transition, that is disrupted by varying leaf genes.

Examination of wild-type inflorescence development revealed clusters of gene expression characteristic of each major stage of development (Chapter 3). Examination of *Ppd-1*'s role using *Ppd-D1a* insensitive lines, indicated the primary effect of this line on meristem gene expression is to shift the profiles of genes to earlier in development. In wild-type much of gene expression is likely coordinated by leaf *FT1* expression, focusing on two major transitions, the VG to DR transition, and the DR to LP transition. It appears from gene clustering, that the insensitive allele shifts those clusters focused to the LP stage to the DR, correlating with a much higher induction of *FT1* in leaf at this stage, relative to wild-type (Chapter 2). This demonstrates the level of cross-talk between the leaf and the developing inflorescence. The *ppd-1* null line has a surprisingly vast effect on expression in the developing inflorescence, with double the number of genes differentially regulated between the stages, relative to the *Ppd-D1a* line and wild-type (Chapter 3). In the *ppd-1* null lines, clustering analysis shows a considerable disruption of gene expression profiles, particularly after the VG to DR transition, exhibiting a loss of precise gene regulation mediated by *Ppd-1*. I also identified a considerable loss of gene regulation from the LP to TS stages. This effect of the *ppd-1* null line on gene expression suggests that it is not due to the induction of *FT1* in the leaf as these null lines have a small decrease in gene expression, relative to wild-type, and still experience a step-wise induction of *FT1* expression (Chapter 2). Instead it suggests an additional regulatory mechanism whereby *Ppd-1*, independently of *FT1*, both positively and negatively regulates gene expression in the developing inflorescence, which is consistent with the high expression of *Ppd-1* in the developing inflorescence. This is supported by a recent study investigating the role of *ELF3* and *CO* in the leaf (Shaw *et al.*, 2020). To investigate direct involvement of *Ppd-1* in developing inflorescences, future studies may explore if similar expression trends observed in the *ppd-1* null lines are replicated in *ft* null lines, perhaps focusing on the genes identified in Chapter 5 (Figure 6.1).

Several of the genes identified in Chapter 5, particularly *bZIP11* and *ALOG1*, are up-regulated in the *ppd-1* null line and down-regulated in the *Ppd-D1a* insensitive line. When knocked out, *bzip11* and *alog1* show accelerated flowering times (Chapter 5). This suggests that these genes may be key effectors of inflorescence development, conveying the accelerated development observed in *Ppd-D1a* lines (Figure 6.1). I hypothesise that these genes repress inflorescence development in a quantitative manner, with higher expression delaying development (e.g. *ppd-1* NILs), while down-regulation accelerates development (e.g. *Ppd-D1a* lines). The roles of *YABBY4, TOM1, bZIP11* and *ALOG1* characterised in Chapter 5 provides an insight into the complexity of the pathways underpinning inflorescence development (Figure 6.1). All four of these genes come from different families, with none possessing orthologues with characterised roles in inflorescence development. *TOM1* highlights the possibilities of combining disease resistance and yield traits, or to at least offset negative traits associated with resistance

growth (Brown, 2003; Nelson *et al.*, 2018). Both *bzip-b11* and *alog-d1* mutants accelerate flowering, an extremely rare trait to discover, especially in a mutant screen. This implies roles for these genes in the repression of flowering. *bZIP* transcription factors have key roles in the FAC and the recently characterised FRC at the floral transition (Kaneko-Suzuki *et al.*, 2018; Cerise *et al.*, 2020). The peak in *bZIP11* expression at the DR stage indicates that *bZIP11* may be involved in a FRC, whereby knocking out *bZIP11* prevents FRC mediated repression of inflorescence development. The flowering phenotype of *bZIP11* appears to be quantitative, with a null copy in one genome copy being sufficient to promote flowering. Not only does this highlight *bZIP11* as a potential mediator of the flowering phenotypes caused by *Ppd-1* allelism, it provides a strategy for tailoring the economically important photoperiod insensitive phenotype. Combining overexpressed alleles of *bZIP11* with *Ppd-D1a* lines could produce earlier flowering plants, that develop through the earlier stages of inflorescence development at a slower rate, negating the yield penalties associated with photoperiod insensitivity and increasing yields. It is possible that *ALOG1* functions in a similar way to the *TAWAWA1,* both genes are members of the ALOG family, with *tawawa1* null lines having a delayed inflorescence development (Yoshida *et al.*, 2013). This similarity is exciting because overexpression of *TAWAWA1* has dramatic consequences for rice inflorescence architecture. In the future overexpressing *ALOG1* in wheat has the potential for increasing branching and consequently spikelet number.

Together this thesis points to a complicated regulation of *Ppd-1*, and significant influence on gene expression that extends beyond the leaf.
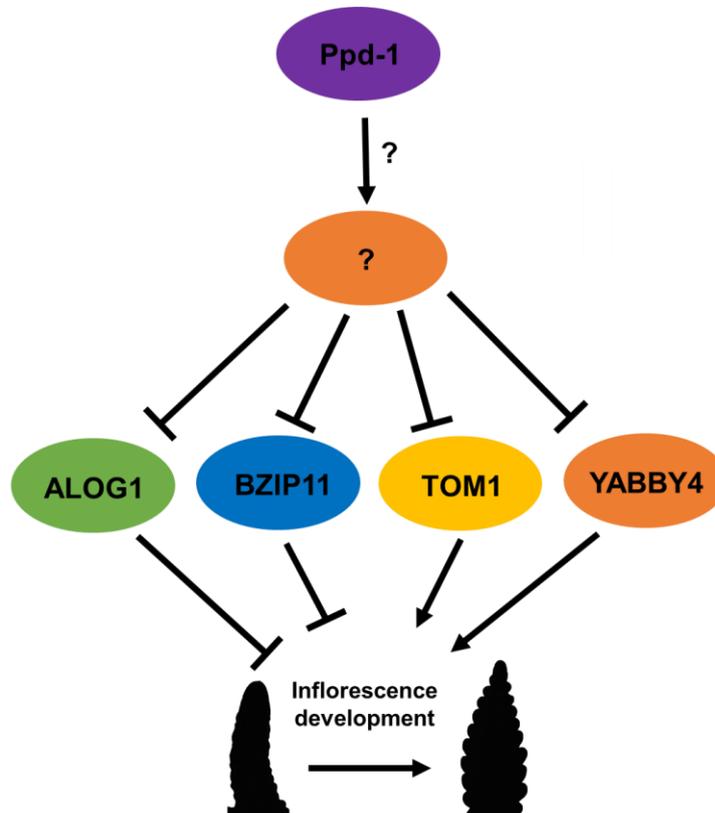
**Figure 6.1 – The role of Ppd-1 in inflorescence development.**

A model summarising the role of Ppd-1 in the regulation of the genes identified in this study. Ppd-1 either directly or through mediator genes, such as the FTs, negatively regulates the expression of *ALOG1*, *BZIP11*, *TOM1* and *YABBY4*. ALOG1 and BZIP11 act to repress inflorescence development whereas TOM1 and YABBY4 promote spikelet formation.

## 6.4 **Influence of polyploidy**

In addition to the effects of *Ppd-1* allelism, this study investigates the influence of polyploidy on gene expression. I find stage and genotype specific effects of genome biases towards expression. On a whole genome level, there is a strong bias towards the D genome, a bias much stronger than that on average across tissues (Ramírez-González *et al.*, 2018). Conversely, I demonstrate that specific gene families, including MADS-box and auxin-related genes, behave much more dynamically as the inflorescence develops, changing expression biases in a stage specific manner. This points to a system whereby gene families can have independent stage-specific biases, offering another layer of understanding to gene expression regulation in wheat. The study of genome biases is important, a major weakness often cited for wheat research is the hexaploid genome, with the need to generate mutants in all three genomes required for a true gene knockout. Understanding how the genomes balance contributions to total gene function, could help inform wheat breeding strategies in the future, targeting the most important homoeologues.

## 6.5 **Regulation of key gene families**

This study has carried out a comprehensive examination of the MICK-type II MADS-box (MADS-box) gene family. I present evidence for a general conservation of gene function from *Arabidopsis* and rice. Importantly, this study has characterised the inflorescence expression patterns of all the MADS-box genes in wheat. I have highlighted how expression of these genes is normalised to stage and not progression through seasons, supporting findings from Chapter 1 (Gauley & Boden, 2020). This tool can help shape future directions of research into these genes, guiding targeted mutagenesis to the highest expressed homoeologue in addition to highlighting some of the most important candidate genes. I carried out a comprehensive comparison of the ABCDE model genes in wheat relative to rice and *Arabidopsis.* Investigating these genes based on expression patterns, revealed they likely have a conserved function to their rice counterparts, with differing expression patterns among paralogues indicating sub-functionalisation. I have also characterised the *ARF* gene family in this study, identifying 20 *ARF* genes in wheat, I have shown how this family of genes show diverse expression patterns across all inflorescence meristem stages, with peaks at different stages, implying stage specific roles for these genes. Together, this data demonstrates the power of high-quality RNA-seq data to rapidly investigate entire gene families.

## 6.6 **Checkpoints in flowering**

In Chapter 2, I determined that a step-wise increase in *FT1* expression mediates the progression of the wheat inflorescence in a checkpoint fashion (Figure 6.2). The first checkpoint correlates with an initial induction of *FT1* mediating the transition from a vegetative meristem to DR. There is a stall in development at this stage until the daylength reaches 11 h, which is followed by a second induction of *FT1* in the leaf, allowing the transition past the DR stage into the LP stage. These inductions are mediated by *Ppd-1*, with insensitive alleles inducing *FT1* earlier and to a greater degree than wild-type levels. The ectopic *FT1* expression is sufficient to accelerate inflorescence development, reducing the delay at the DR stage, resulting in faster flowering. There is no expression of *FT1* in the developing inflorescence, that is until the green anther stage of development, at which point I identified extremely high expression in the rachis, in a *Ppd-1* dependent manner. This points to another checkpoint in development at this GA stage, at which point induction of *FT1* is required for continued floral development.

There is also evidence for a checkpoint in meristem development at the LP stage. I described in Chapter 2 how a short-day photoperiod of 10 hours was sufficient to induce the early stages of flowering. However, without photoperiod extending past 10 h daylengths, development stalls at the LP stage. This stalling of development occurs,

irrespective of photoperiod, in *ft2-b* missense lines. I saw expression of *FT2* is under the influence of *Ppd-1,* indicating a third *Ppd-1* mediated checkpoint of development at the LP stage. To this end, Chapter 3 provides strong evidence for the LP stage being a particularly important stage for gene regulation in the inflorescence, with over 50% of all clustered gene expression focused to this stage with either a peak or dip in expression (Chapter 3).

Together these data point to four checkpoints of *Ppd-1*-mediated control over the developing inflorescence.

1. From VG to DR, mediated by *FT1* from the leaf during short days (10 h).

2. From the DR to LP stage, caused by a second induction of *FT1* in the leaf (12 h).

3. At the LP stage, mediated by *Ppd-1* functioning in the meristem.

4. At the GA stage through induction of *FT1* in the rachis.
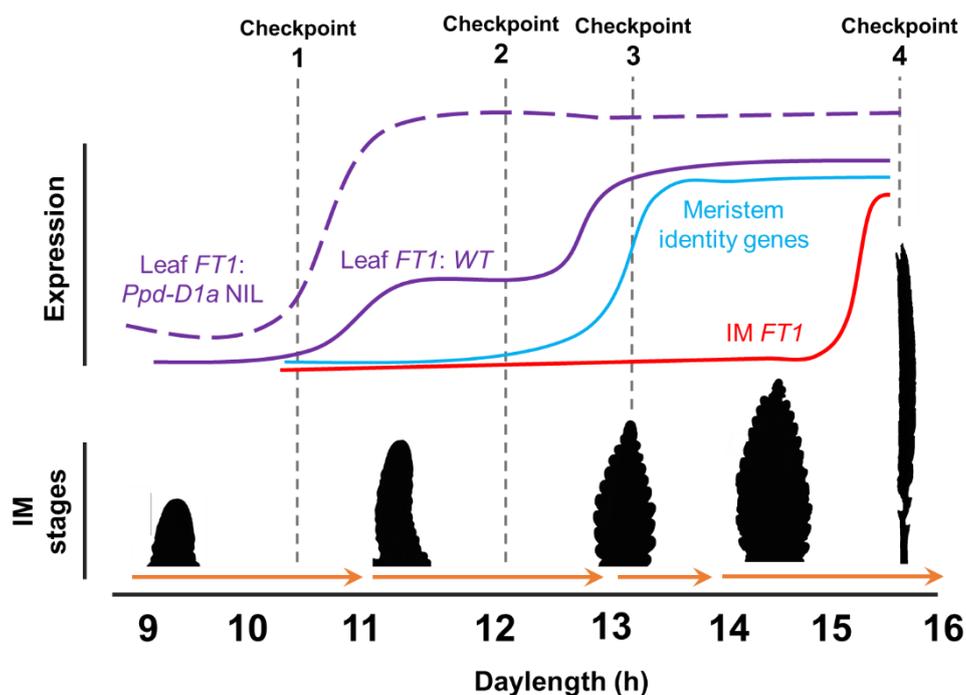


**Figure 6.2 - A checkpoint model for the seasonal regulation of flowering-time pathways and inflorescence development in wheat.**

A model outlining the seasonal regulation of *FT1* expression and resulting checkpoints in development. Checkpoints and inflorescence images are representative of wild-type plants. Checkpoint 1: from VG to DR, mediated by FT1 from the leaf (WT, solid, purple line) during short days (10 h). Checkpoint 2: from the DR to LP stage, caused by a second induction of *FT1* in the leaf (12 h). Both Checkpoint 1 and 2 are overridden by high *FT1* expression in the *Ppd-D1a* photoperiod insensitive line (Dashed, purple line). Checkpoint 3: At the LP stage, mediated by Ppd-1 functioning in the meristem alongside signals from the leaf to induce genes involved in spikelet number (blue). Checkpoint 4- At the GA stage through induction of *FT1* in the rachis (red). The stages shown include vegetative (VG), double ridge (DR), lemma primordium (LP), terminal spikelet (TS) and green anther (GA).

## 6.7 **The role of temperature on inflorescence development.**

In Chapter 4, I described the effect of temperature on inflorescence development. Over three growing seasons, the timing of the floral transition was tightly controlled occurring with remarkable precision every year. However, a period of cold just before the transition is sufficient to delay flowering, previous studies have highlighted the importance of the floral transition occurring under optimal temperatures (Limin & Fowler, 2006). Beyond the transition I show temperature appears to continue to regulate the rate of stage transitions, with lower temperatures correlating with a delay in development. This indicates that temperature regulation of inflorescence continues beyond the floral transition to regulate the formation of key formal structures under optimal conditions. In Chapter 2, I indicate that differences in temperature cause substantial variations in flowering between the field and the glasshouse, with higher temperatures in the glasshouse, overriding the step-wise induction of *FT1.* A role for temperature mediating flowering signals is supported by studies in *Arabidopsis* (Song *et al.*, 2018). These findings point to at least two different points of temperature regulation in wheat, namely temperature-based regulation of *FT1* in the leaf, and temperature-mediated control over inflorescence development. Interaction between the photoperiod and temperature pathways is indicated by photoperiod insensitive lines advancing through inflorescence stages faster under warmer ambient temperatures (Chapter 4).

The genetic mechanisms mediating these processes are relatively unknown. In the leaf, it is possible that it is mediated by the effects of *ELF3* through its recently described prion domains (Jung *et al.*, 2020). In the developing inflorescence the regulation is likely due to one of two reasons.

1. A unique pathway precisely controlling the stage transitions of the inflorescence.

2. A balancing of temperature stress responses.

In reality it is likely a degree of both. Temperature mediated control over the developing inflorescence can occur under a wide range of temperatures, including the vernalisation pathway that can function from <6⁰C to 16⁰C (Dixon *et al.*, 2019; O'Neill *et al.*, 2019). However, one of the primary differences I observed between the glasshouse and the field was the protection against sub-zero temperatures, studies have shown significant biochemical changes occur under these temperatures, likely causing stress and delaying development (Xin & Browse, 2000). There is also evidence for a role for a genuine temperature mediated genetic pathway controlling inflorescence development. Not least that these yield essential development stages are occurring during a 4-16⁰C temperature range, a range considered to induce stress (Gol *et al.*, 2017), and yet produce some of the

highest yield per hectare in the world, 2019 yields in East Anglia are 9.9t/ha, higher than the average yields of any other country in Europe (Ritchie and Roser, 2013; Strutt & Parker, 2019). It is during this short development window that spikelet number is decided, and florets are initiated (Kirby & Appleyard, 1984). This study has already indicated several photoperiod mediated checkpoints of flowering, it stands to reason these would be intertwined with temperature.

The mechanism for this temperature pathway may come in the form of regulation through MADS-box genes, particularly of the *FLC* family, likely because of domestication pressure (Schilling *et al.*, 2018, 2020). For domestication to provide pressure on the family, they must be functional in wheat development, likely through mediation of temperature responses (Theißen *et al.*, 2018; Schilling *et al.*, 2020). The MADS-box genes as a whole, outlined in Chapter 3, are expressed throughout inflorescence development alongside auxin signalling/response genes, experiencing peaks and down-regulation as the meristem develops and have reported roles integrating temperature signals (Chandler, 2016; Chen *et al.*, 2019).

## 6.8 Interaction between photoperiod and temperature pathways

This thesis has outlined examples of the two major environmental factors influencing wheat development in the field. These are the photoperiod pathway, through the function of *Ppd-1* mediated induction of *FT1* and the effects of low and ambient temperature on floral development. In the leaf, both these pathways converge on *FT1*, and there is evidence for interaction between the pathways from studies in Arabidopsis and barley (Kumar *et al.*, 2012; Ejaz & von Korff, 2017). In Chapter 4 I examined the effects of photoperiod insensitivity on development in a glasshouse under natural conditions relative to one maintained at 16 °C, I show an acceleration of development under the higher temperature conditions. This points to interplay between the photoperiod and temperature pathways, whereby removing the photoperiod checks on flowering results in the enhanced temperature mediated acceleration of IM development.

Future work may examine this interaction further, using protein-based assay, to investigate interactors with *FT1* in the leaf. Carrying out similar assays for key photoperiod genes such as *Ppd-1* and temperature genes, including *VRN1* may reveal physical interaction between these important flowering genes. A crucial aspect for these assays will be to perform them under appropriate field conditions. The data I present in Chapter 2 shows the importance of field-based experiments, without carrying out these experiments in the field the step-wise increase in *FT1* would not have been characterised. Therefore, carrying out temperature based assays in the field at the checkpoints identified in this

study; before the floral transition, at the floral transition, at the LP stage and at the GA stage, would undoubtedly reveal seasonal dynamic interactions between these pathways.

## 6.9 **Mutant screens for the future**

Chapter 3 demonstrated the effectiveness of RNA-seq data to investigate the transcriptional landscape of the developing wheat inflorescence, both on a global level and for specific gene families. Based on this analysis, I identified both expected and unexpected trends of gene expression underpinning wheat inflorescence development. However, Chapter 5 took a different approach to understand this key developmental process. A core aim of Chapter 5 was to identify genes I may not expect to be differentially regulated between the inflorescence meristem stages. To this end, I examined the genes most differentially regulated between the DR and LP stages. The steps outlined in Chapter 5 identified 4 mutants with verified roles in wheat inflorescence development and architecture, two of which have accelerated flowering times. A surprising outcome from this screen was that functional redundancy among genome copies was not prevalent. Lines with mutants in one copy could produce strong wheat inflorescence architecture phenotypes. This screen was effective and rapid, allowing me to identify mutants much faster than traditional forward genetics methods. This approach is not only restricted to photoperiod focused investigations of the inflorescence, but could be applied under different conditions, or indeed to different tissues. It demonstrates a method for the rapid identification of candidate genes in wheat. Strategies such as these using the diverse resources now available for wheat research, will pave the way for rapid gene discovery.

## 6.10 **Concluding statement**

Overall, this thesis has described intimate crosstalk between the leaf and the developing inflorescence. In the leaf, *Ppd-1* integrates seasonal signals to mediate a step-wise increase in *FT1* transcription. These leaf signals control the progression of the inflorescence through several stage-specific checkpoints of development. The relationship between the leaf and the inflorescence is dynamic and complex, with *Ppd-1* allelism causing remarkable changes in inflorescence gene expression patterns and amplitude. Investigating these transcriptional effects has revealed four previously uncharacterised genes involved in inflorescence development.

Ultimately this new understanding of the flowering pathway in wheat presents exciting opportunities for exploitation to help safeguard food security for future generations.

# References

Abe M, Kobayashi Y, Yamamoto S, Daimon Y, Yamaguchi A, Ikeda Y, Ichinoki H, Notaguchi M, Goto K, Araki T. 2005. FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. *Science* 309: 1052–1056.

Abe M, Kosaka S, Shibuta M, Nagata K, Uemura T, Nakano A, Kaya H. 2019. Transient activity of the florigen complex during the floral transition in arabidopsis thaliana. *Development (Cambridge)* 146.

Abel S, Theologis A. 1996. Early genes and auxin action. *Plant Physiology* 111: 9–17.

Abu-Jamous B, Kelly S. 2018. Clust: automatic extraction of optimal co-expressed gene clusters from gene expression data. *Genome Biology* 19: 172.

Adamski NM, Borrill P, Brinton J, Harrington SA, Marchal C, Bentley AR, Bovill WD, Cattivelli L, Cockram J, Contreras-Moreira B. 2020. A roadmap for gene functional characterisation in crops with large genomes: Lessons from polyploid wheat. *eLife* 9.

Alabadí D, Oyama T, Yanovsky MJ, Harmon FG, Más P, Kay SA. 2001. Reciprocal regulation between TOC1 and LHY/CCA1 within the Arabidopsis circadian clock. *Science* 293: 880–883.

Alexandre CM, Hennig L. 2008. FLC or not FLC: the other side of vernalization. *Journal of Experimental Botany* 59: 1127–1135.

De Almeida Engler J, De Veylder L, De Groodt R, Rombauts S, Boudolf V, De Meyer B, Hemerly A, Ferreira P, Beeckman T, Karimi M. 2009. Systematic analysis of cell-cycle gene expression during Arabidopsis development. *Plant Journal* 59: 645–660.

Alonso JM, Ecker JR. 2006. Moving forward in reverse: Genetic technologies to enable genome-wide phenomic screens in Arabidopsis. *Nature Reviews Genetics* 7: 524–536.

Alvarez MA, Tranquilli G, Lewis S, Kippes N, Dubcovsky J. 2016. Genetic and physical mapping of the earliness per se locus Eps-A m 1 in Triticum monococcum identifies EARLY FLOWERING 3 (ELF3) as a candidate gene. *Functional and Integrative Genomics* 16: 365–382.

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

Appels R, Eversole K, Feuillet C, Keller B, Rogers J, Stein N, Pozniak CJ, Choulet F, Distelfeld A, Poland J. 2018. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT. 2000. Gene ontology: Tool for the unification of biology. *Nature Genetics* 25: 25–29.

Beales J, Turner A, Griffiths S, Snape JW, Laurie DA. 2007. A Pseudo-Response Regulator is misexpressed in the photoperiod insensitive Ppd-D1a mutant of wheat (Triticum aestivum L.). *Theoretical and Applied Genetics* 115: 721–733.

Bentley AR, Horsnell R, Werner CP, Turner AS, Rose GA, Bedard C, Howell P, Wilhelm EP, Mackay IJ, Howells RM. 2013. Short, natural, and extended photoperiod response in BC2F4 lines of bread wheat with different Photoperiod-1 (Ppd-1) alleles. *Journal of Experimental Botany* 64: 1783–1793.

Bentley AR, Turner AS, Gosman N, Leigh FJ, Maccaferri M, Dreisigacker S, Greenland A, Laurie DA. 2011. Frequency of photoperiod-insensitive Ppd-A1a alleles in tetraploid, hexaploid and synthetic hexaploid wheat germplasm. *Plant Breeding* 130: 10–15.

Blanc G, Wolfe KH. 2004. Functional divergence of duplicated genes formed by polyploidy during arabidopsis evolution. *Plant Cell* 16: 1679–1691.

Block GD, Page TL. 1978. Circadian Pacemakers in the Nervous System. *Annual Review of Neuroscience* 1: 19–34.

Boden SA, Cavanagh C, Cullis BR, Ramm K, Greenwood J, Jean Finnegan E, Trevaskis B, Swain SM. 2015. Ppd-1 is a key regulator of inflorescence architecture and paired spikelet development in wheat. *Nature Plants* 1.

Bommert P. 2005. Genetics and Evolution of Inflorescence and Flower Development in Grasses. *Plant and Cell Physiology* 46: 69–78.

Borlaug NE, Narvaez I, Aresvik O, Anderson RG. 1969. Green revolution yields a golden harvest. *Columbia J World Bus*.

Borrill P, Ramirez-Gonzalez R, Uauy C. 2016. expVIP: A customizable RNA-seq data analysis and visualization platform. *Plant Physiology* 170: 2172–2186.

Brambilla V, Martignago D, Goretti D, Cerise M, Somssich M, de Rosa M, Galbiati F, Shrestha R, Lazzaro F, Simon R. 2017. Antagonistic Transcription Factor Complexes Modulate the Floral Transition in Rice. *The Plant cell* 29: 2801–2816.

Bratzel F, Turck F. 2015. Molecular memories in the regulation of seasonal flowering: From competence to cessation. *Genome Biology* 16: 1–14.

Brinton J, Simmonds J, Uauy C. 2018. Ubiquitin-related genes are differentially expressed

in isogenic lines contrasting for pericarp cell size and grain weight in hexaploid wheat. *BMC Plant Biology* 18: 1–17.

Brown JKM. 2003. A cost of disease resistance: Paradigm or peculiarity? *Trends in Genetics* 19: 667–671.

Calixto CPG, Waugh R, Brown JWS. 2015. Evolutionary Relationships Among Barley and Arabidopsis Core Circadian Clock and Clock-Associated Genes. *Journal of Molecular Evolution* 80: 108–119.

Campoli C, Shtaya M, Davis SJ, von Korff M. 2012. Expression conservation within the circadian clock of a monocot: natural variation at barley Ppd-H1 affects circadian expression of flowering time genes, but not clock orthologs. *BMC Plant Biology* 12: 97.

Cane K, Eagles HA, Laurie DA, Trevaskis B, Vallance N, Eastwood RF, Gororo NN, Kuchel H, Martin PJ. 2013. Ppd-B1 and Ppd-D1 and their effects in southern Australian wheat. *Crop and Pasture Science* 64: 100–114.

Carbon S, Douglass E, Dunn N, Good B, Harris NL, Lewis SE, Mungall CJ, Basu S, Chisholm RL, Dodson RJ. 2019. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Research* 47: D330–D338.

Cerise M, Giaume F, Galli M, Khahani B, Lucas J, Podico F, Tavakol E, Parcy F, Gallavotti A, Brambilla V. 2020. OsFD4 promotes the rice floral transition via florigen activation complex formation in the shoot apical meristem. *New Phytologist*: nph.16834.

Chailakhyan MK. 1937. Hormonal theory of plant development. *Hormonal theory of plant development.*

Chandler JW. 2016. Auxin response factors. *Plant, Cell & Environment* 39: 1014–1028.

Chen Y, Carver BF, Wang S, Zhang F, Yan L. 2009. Genetic loci associated with stem elongation and winter dormancy release in wheat. *Theoretical and Applied Genetics* 118: 881–889.

Chen A, Dubcovsky J. 2012. Wheat TILLING Mutants Show That the Vernalization Gene VRN1 Down-Regulates the Flowering Repressor VRN2 in Leaves but Is Not Essential for Flowering (B Trevaskis, Ed.). *PLoS Genetics* 8: e1003134.

Chen A, Li C, Hu W, Lau MY, Lin H, Rockwell NC, Martin SS, Jernstedt JA, Lagarias JC, Dubcovsky J. 2014. PHYTOCHROME C plays a major role in the acceleration of wheat flowering under long-day photoperiod. *Proceedings of the National Academy of Sciences of the United States of America* 111: 10037–10044.

Chen X, Meyerowitz EM. 1999. HUA1 and HUA2 are two members of the floral homeotic AGAMOUS pathway. *Molecular Cell* 3: 349–360.

Chen M, Penfield S. 2018. Feedback regulation of COOLAIR expression controls seed dormancy and flowering time. *Science* 360: 1014–1017.

Chen D, Wang W, Wu Y, Xie H, Zhao L, Zeng Q, Zhan Y. 2019. Expression and distribution of the auxin response factors in Sorghum bicolor during development and temperature stress. *International Journal of Molecular Sciences* 20.

Choi Y, Chan AP. 2015. PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* 31: 2745–2747.

Chouard P. 1960. Vernalization and its Relations to Dormancy. *Annual Review of Plant Physiology* 11: 191–238.

Chow CN, Zheng HQ, Wu NY, Chien CH, Huang H Da, Lee TY, Chiang-Hsieh YF, Hou PF, Yang TY, Chang WC. 2016. PlantPAN 2.0: An update of Plant Promoter Analysis Navigator for reconstructing transcriptional regulatory networks in plants. *Nucleic Acids Research* 44: D1154–D1164.

Chuck G, Muszynski M, Kellogg E, Hake S, Schmidt RJ. 2002. The control of spikelet meristem identity by the branched silkless1 gene in maize. *Science* 298: 1238–1241.

Coen ES, Meyerowitz EM. 1991. The war of the whorls: Genetic interactions controlling flower development. *Nature* 353: 31–37.

Collani S, Neumann M, Yant L, Schmid M. 2019. FT modulates genome-wide DNA-binding of the bZIP transcription factor FD. *Plant Physiology* 180: 367–380.

Comastri A, Janni M, Simmonds J, Uauy C, Pignone D, Nguyen HT, Marmiroli N. 2018. Heat in wheat: exploit reverse genetic techniques to discover new alleles within the triticum durum shsp26 family. *Frontiers in Plant Science* 9: 1337.

Corbesier L, Vincent C, Jang S, Fornara F, Fan Q, Searle I, Giakountis A, Farrona S, Gissot L, Turnbull C. 2007. FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. *Science* 316: 1030–1033.

Coventry D, Reeves T, Brooke H, Cann D. 1993. Influence of genotype, sowing date, and seeding rate on wheat development and yield. *Australian Journal of Experimental Agriculture* 33: 751.

Danilevskaya ON, Meng X, Selinger DA, Deschamps S, Hermon P, Vansant G, Gupta R, Ananiev E V, Muszynski MG. 2008. Involvement of the MADS-box gene ZMM4 in floral

induction and inflorescence development in maize. *Plant physiology* 147: 2054–69.

Danyluk J, Kane NA, Breton G, Limin AE, Fowler DB, Sarhan F. 2003. TaVRT-1, a putative transcription factor associated with vegetative to reproductive transition in cereals. *Plant Physiology* 132: 1849–1860.

Debernardi JM, Tricoli DM, Ercoli MF, Hayta S, Ronald P, Palatnik JF, Dubcovsky J. 2020. A GRF–GIF chimeric protein improves the regeneration efficiency of transgenic plants. *Nature Biotechnology*: 1–6.

Deng W, Casao MC, Wang P, Sato K, Hayes PM, Finnegan EJ, Trevaskis B. 2015a. Direct links between the vernalization response and other key traits of cereal crops. *Nature Communications* 6.

Deng W, Clausen J, Boden S, Oliver SN, Casao MC, Ford B, Anderssen RS, Trevaskis B. 2015b. Dawn and Dusk Set States of the Circadian Oscillator in Sprouting Barley (Hordeum vulgare) Seedlings (T Wang, Ed.). *PLOS ONE* 10: e0129781.

Deng W, Ying H, Helliwell CA, Taylor JM, Peacock WJ, Dennis ES. 2011. FLOWERING LOCUS C (FLC) regulates development pathways throughout the life cycle of Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America* 108: 6680–6685.

Derbyshire P, Byrne ME. 2013. MORE SPIKELETS1 is required for spikelet fate in the inflorescence of Brachypodium. *Plant physiology* 161: 1291–302.

Díaz A, Zikhali M, Turner AS, Isaac P, Laurie DA. 2012. Copy Number Variation Affecting the Photoperiod-B1 and Vernalization-A1 Genes Is Associated with Altered Flowering Time in Wheat (Triticum aestivum) (SP Hazen, Ed.). *PLoS ONE* 7: e33234.

Digel B, Pankin A, von Korff M. 2015. Global Transcriptome Profiling of Developing Leaf and Shoot Apices Reveals Distinct Genetic and Environmental Control of Floral Transition and Inflorescence Development in Barley. *The Plant cell* 27: 2318–34.

Ditta G, Pinyopich A, Robles P, Pelaz S, Yanofsky MF. 2004. The SEP4 gene of Arabidopsis thaliana functions in floral organ and meristem identity. *Current Biology* 14: 1935–1940.

Dixon LE, Farré A, Finnegan EJ, Orford S, Griffiths S, Boden SA. 2018. Developmental responses of bread wheat to changes in ambient temperature following deletion of a locus that includes FLOWERING LOCUS T1. *Plant Cell and Environment* 41: 1715–1725.

Dixon LE, Greenwood JR, Bencivenga S, Zhang P, Cockram J, Mellers G, Ramm K, Cavanagh C, Swain SM, Boden SA. 2018. TEOSINTE BRANCHED1 regulates

inflorescence architecture and development in bread wheat (Triticum aestivum). *Plant Cell* 30: 563–581.

Dixon LE, Karsai I, Kiss T, Adamski NM, Liu Z, Ding Y, Allard V, Boden SA, Griffiths S. 2019. VERNALIZATION1 controls developmental responses of winter wheat under high ambient temperatures. *Development (Cambridge, England)* 146.

Dixon LE, Knox K, Kozma-Bognar L, Southern MM, Pokhilko A, Millar AJ. 2011. Temporal repression of core circadian genes is mediated through EARLY FLOWERING 3 in Arabidopsis. *Current Biology* 21: 120–125.

Dobrovolskaya O, Pont C, Sibout R, Martinek P, Badaeva E, Murat F, Chosson A, Watanabe N, Prat E, Gautier N. 2015. FRIZZY PANICLE drives supernumerary spikelets in bread wheat. *Plant physiology* 167: 189–99.

Doebley J, Stec A, Hubbard L. 1997. The evolution of apical dominance in maize. *Nature* 386: 485–488.

Dreni L, Jacchia S, Fornara F, Fornari M, Ouwerkerk PBF, An G, Colombo L, Kater MM. 2007. The D-lineage MADS-box gene OsMADS13 controls ovule identity in rice. *The Plant Journal* 52: 690–699.

Dreni L, Pilatone A, Yun D, Erreni S, Pajoro A, Caporali E, Zhang D, Katera MM. 2011. Functional analysis of all AGAMOUS subfamily members in rice reveals their roles in reproductive organ identity determination and meristem determinacy. *Plant Cell* 23: 2850–2863.

Dubcovsky J, Loukoianov A, Fu D, Valarik M, Sanchez A, Yan L. 2006. Effect of photoperiod on the regulation of wheat vernalization genes VRN1 and VRN2. *Plant Molecular Biology* 60: 469–480.

Duncan S, Holm S, Questa J, Irwin J, Grant A, Dean C. 2015. Seasonal shift in timing of vernalization as an adaptation to extreme winter. *eLife* 4.

Eagles HA, Cane K, Trevaskis B. 2011. Veery wheats carry an allele of Vrn-A1 that has implications for freezing tolerance in winter wheats. *Plant Breeding* 130: 413–418.

Ejaz M, von Korff M. 2017. The genetic control of reproductive development under high ambient temperature. *Plant Physiology* 173: 294–306.

Ellis CM, Nagpal P, Young JC, Hagen G, Guilfoyle TJ, Reed JW. 2005. AUXIN RESPONSE FACTOR1 and AUXIN RESPONSE FACTOR2 regulate senescence and floral organ abscission in Arabidopsis thaliana. *Development* 132: 4563–4574.

Emery M, Willis MMS, Hao Y, Barry K, Oakgrove K, Peng Y, Schmutz J, Lyons E, Pires JC, Edger PP. 2018. Preferential retention of genes from one parental genome after polyploidy illustrates the nature and scope of the genomic conflicts induced by hybridization. *PLoS Genetics* 14.

FAO. 2015. *The State of Food Insecurity in the World*.

Faure S, Turner AS, Gruszka D, Christodoulou V, Davis SJ, Von Korff M, Laurie DA. 2012. Mutation at the circadian clock gene EARLY MATURITY 8 adapts domesticated barley (Hordeum vulgare) to short growing seasons. *Proceedings of the National Academy of Sciences of the United States of America* 109: 8328–8333.

Favaro R, Pinyopich A, Battaglia R, Kooiker M, Borghi L, Ditta G, Yanofsky MF, Kater MM, Colombo L. 2003. MADS-Box Protein Complexes Control Carpel and Ovule Development in Arabidopsis. *Plant Cell* 15: 2603–2611.

Feng N, Song G, Guan J, Chen K, Jia M, Huang D, Wu J, Zhang L, Kong X, Geng S. 2017. Transcriptome Profiling of Wheat Inflorescence Development from Spikelet Initiation to Floral Patterning Identified Stage-Specific Regulatory Genes. *Plant physiology* 174: 1779–1794.

Finnegan EJ, Ford B, Wallace X, Pettolino F, Griffin PT, Schmitz RJ, Zhang P, Barrero JM, Hayden MJ, Boden SA. 2018. Zebularine treatment is associated with deletion of FT-B1 leading to an increase in spikelet number in bread wheat. *Plant Cell and Environment* 41: 1346–1360.

Fjellheim S, Boden S, Trevaskis B. 2014. The role of seasonal flowering responses in adaptation of grasses to temperate climates. *Frontiers in Plant Science* 5: 431.

Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J. 1999. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151: 1531–1545.

Ford B, Deng W, Clausen J, Oliver S, Boden S, Hemming M, Trevaskis B. 2016. Barley ( *Hordeum vulgare* ) circadian clock genes can respond rapidly to temperature in an *EARLY FLOWERING 3* -dependent manner. *Journal of Experimental Botany* 67: 5517–5528.

Fornara F, Pařenicová L, Falasca G, Pelucchi N, Masiero S, Ciannamea S, Lopez-Dee Z, Altamura MM, Colombo L, Kater MM. 2004. Functional characterization of OsMADS18, a member of the AP1/SQUA subfamily of MADS box genes. *Plant Physiology* 135: 2207–2219.

Fujiwara S, Wang L, Han L, Suh SS, Salomé PA, McClung CR, Somers DE. 2008. Post-translational regulation of the Arabidopsis circadian clock through selective proteolysis and phosphorylation of pseudo-response regulator proteins. *Journal of Biological*

*Chemistry* 283: 23073–23083.

Gaamouche T, Manes C-L de O, Kwiatkowska D, Berckmans B, Koumproglou R, Maes S, Beeckman T, Vernoux T, Doonan J, Traas J. 2010. Cyclin-dependent kinase activity retains the shoot apical meristem cells in an undifferentiated state. *The Plant Journal* 64: no-no.

Gallavotti A, Yang Y, Schmidt RJ, Jackson D. 2008. The relationship between auxin transport and maize branching. *Plant Physiology* 147: 1913–1923.

Galli M, Liu Q, Moss BL, Malcomber S, Li W, Gaines C, Federici S, Roshkovan J, Meeley R, Nemhauser JL. 2015. Auxin signaling modules regulate maize inflorescence architecture. *Proceedings of the National Academy of Sciences of the United States of America* 112: 13372–13377.

Gauley A, Boden SA. 2020. Step-wise increases in *FT1* expression regulate seasonal progression of flowering in wheat (Triticum aestivum). *New Phytologist.*

Gauley A, Boden SA. 2019. Genetic pathways controlling inflorescence architecture and development in wheat and barley. *Journal of Integrative Plant Biology* 61: 296–309.

Gawroński P, Ariyadasa R, Himmelbach A, Poursarebani N, Kilian B, Stein N, Steuernagel B, Hensel G, Kumlehn J, Sehgal SK. 2014. A Distorted Circadian Clock Causes Early Flowering and Temperature-Dependent Variation in Spike Development in the Eps-3Am Mutant of Einkorn Wheat. *Genetics* 196: 1253–1261.

Gol L, Tomé F, Von Korff M. 2017. Floral transitions in wheat and barley: Interactions between photoperiod, abiotic stresses, and nutrient status. *Journal of Experimental Botany* 68: 1399–1410.

Gómez-Ariza J, Brambilla V, Vicentini G, Landini M, Cerise M, Carrera E, Shrestha R, Chiozzotto R, Galbiati F, Caporali E. 2019. A transcription factor coordinating internode elongation and photoperiodic signals in rice. *Nature Plants* 5: 358–362.

Gómez-Ariza J, Galbiati F, Goretti D, Brambilla V, Shrestha R, Pappolla A, Courtois B, Fornara F. 2015. Loss of floral repressor function adapts rice to higher latitudes in Europe. *Journal of Experimental Botany* 66: 2027–2039.

González-Navarro OE, Griffiths S, Molero G, Reynolds MP, Slafer GA. 2015. Dynamics of floret development determining differences in spike fertility in an elite population of wheat. *Field Crops Research* 172: 21–31.

González FG, Slafer GA, Miralles DJ. 2005. Photoperiod during stem elongation in wheat: is its impact on fertile floret and grain number determination similar to that of radiation?

*Functional Plant Biology* 32: 181.

Goto K, Meyerowitz EM. 1994. Function and regulation of the Arabidopsis floral homeotic gene PISTILLATA. *Genes and Development* 8: 1548–1560.

Greenup A, Peacock WJ, Dennis ES, Trevaskis B. 2009. The molecular biology of seasonal flowering-responses in Arabidopsis and the cereals. *Annals of Botany* 103: 1165–1172.

Griffiths S, Simmonds J, Leverington M, Wang Y, Fish L, Sayers L, Alibert L, Orford S, Wingen L, Herry L. 2009. Meta-QTL analysis of the genetic control of ear emergence in elite European winter wheat germplasm. *Theoretical and Applied Genetics* 119: 383–395.

Guilfoyle T, Hagen G, Ulmasov T, Murfett J. 1998. How does auxin turn on genes? *Plant Physiology* 118: 341–347.

Guo S, Sun B, Looi LS, Xu Y, Gan ES, Huang J, Ito T. 2015. Co-ordination of flower development through epigenetic regulation in two model species: Rice and arabidopsis. *Plant and Cell Physiology* 56: 830–841.

Gutierrez C, Puchta H. 2015. Chromatin and development: a special issue. *The Plant Journal* 83: 1–3.

Hadley Wickham. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Hamilton N. 2020. An Extension to 'ggplot2', for the Creation of Ternary Diagrams [R package ggtern version 3.3.0].

Hannah Ritchie and Max Roser. 2013. Crop Yields. *OurWorldInData.org*.

Hardtke CS, Berleth T. 1998. The Arabidopsis gene MONOPTEROS encodes a transcription factor mediating embryo axis formation and vascular development. *EMBO Journal* 17: 1405–1411.

Harmer SL. 2009. The Circadian System in Higher Plants. *Annual Review of Plant Biology* 60: 357–377.

Harper AL, Trick M, He Z, Clissold L, Fellgett A, Griffiths S, Bancroft I. 2016. Genome distribution of differential homoeologue contributions to leaf gene expression in bread wheat. *Plant Biotechnology Journal* 14: 1207–1214.

Harrop TWR, Ud Din I, Gregis V, Osnato M, Jouannic S, Adam H, Kater MM. 2016. Gene expression profiling of reproductive meristem types in early rice inflorescences by laser microdissection. *The Plant Journal* 86: 75–88.

Hayama R, Sarid-Krebs L, Richter R, Fernández V, Jang S, Coupland G. 2017. PSEUDO RESPONSE REGULATORs stabilize CONSTANS protein to promote flowering in response to day length. *The EMBO Journal* 36: 904–918.

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.

Hedman H, Källman T, Lagercrantz U. 2009. Early evolution of the MFT-like gene family in plants. *Plant Molecular Biology* 70: 359–369.

Hemming MN, Peacock WJ, Dennis ES, Trevaskis B. 2008. Low-temperature and daylength cues are integrated to regulate Flowering Locus T in barley. *Plant Physiology* 147: 355–366.

Hemming MN, Walford SA, Fieg S, Dennis ES, Trevaskis B. 2012. Identification of high-temperature-responsive genes in Cereals. *Plant Physiology* 158: 1439–1450.

Henikoff S, Till BJ, Comai L. 2004. TILLING. Traditional mutagenesis meets functional genomics. *Plant Physiology* 135: 630–636.

Hepworth J, Antoniou-Kourounioti RL, Bloomer RH, Selga C, Berggren K, Cox D, Collier Harris BR, Irwin JA, Holm S, Säll T. 2018. Absence of warmth permits epigenetic memory of winter in Arabidopsis. *Nature Communications* 9: 1–8.

Hsu CY, Liu Y, Luthe DS, Yuceer C. 2006. Poplar FT2 shortens the juvenile phase and promotes seasonal flowering. *Plant Cell* 18: 1856–1861.

Hughes AL. 1994. The evolution of functionally novel proteins after gene duplication. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 256: 119–124.

Hunt J, Trevaskis B, Fletcher A, Peake A, Zwart A, Fettell N. 2015. Novel wheat genotypes for early sowing across Australian wheat production en- vironments. : 20–24.

IDRC. 2017. Facts &amp; Figures on Food and Biodiversity | IDRC - International Development Research Centre.

Jack T, Brockman LL, Meyerowitz EM. 1992. The homeotic gene APETALA3 of Arabidopsis thaliana encodes a MADS box and is expressed in petals and stamens. *Cell* 68: 683–697.

Jacott CN, Boden SA. 2020. Feeling the heat: developmental and molecular responses of wheat and barley to high ambient temperatures. *Journal of Experimental Botany* 71: 5740–5751.

Jaeger KE, Wigge PA, Kim J, Lee JH, Hong SM, Yoo SJ, Yoo SY, Lee JS, Ahn JH,

Turnbull C. 2007. FT Protein Acts as a Long-Range Signal in Arabidopsis. *Current Biology* 17: 1050–1054.

Jankowicz-Cieslak J, Till BJ. 2016. Forward and reverse genetics in crop breeding. In: Advances in Plant Breeding Strategies: Breeding, Biotechnology and Molecular Tools. Springer International Publishing, 215–240.

Jeon JS, Jang S, Lee S, Nam J, Kim C, Lee SH, Chung YY, Kim SR, Lee YH, Cho YG. 2000. leafy hull sterile 1 is a homeotic mutation in a rice MADS box gene affecting rice flower development. *Plant Cell* 12: 871–884.

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

Jones DT, Taylor WR, Thornton JM. 1992. The rapid generation of mutation data matrices from protein sequences. *Bioinformatics* 8: 275–282.

Jordan KW, He F, De Soto MF, Akhunova A, Akhunov E. 2020. Differential chromatin accessibility landscape reveals structural and functional features of the allopolyploid wheat chromosomes. *Genome Biology* 21: 176.

Jung J-H, Barbosa AD, Hutin S, Kumita JR, Gao M, Derwort D, Silva CS, Lai X, Pierre E, Geng F. 2020. A prion-like domain in ELF3 functions as a thermosensor in Arabidopsis. *Nature*: 1–5.

Jung JH, Domijan M, Klose C, Biswas S, Ezer D, Gao M, Khattak AK, Box MS, Charoensawan V, Cortijo S. 2016. Phytochromes function as thermosensors in Arabidopsis. *Science* 354: 886–889.

Jung JH, Ju Y, Seo PJ, Lee JH, Park CM. 2012. The SOC1-SPL module integrates photoperiod and gibberellic acid signals to control flowering time in Arabidopsis. *Plant Journal* 69: 577–588.

Kaneko-Suzuki M, Kurihara-Ishikawa R, Okushita-Terakawa C, Kojima C, Nagano-Fujiwara M, Ohki I, Tsuji H, Shimamoto K, Taoka KI. 2018. TFL1-Like Proteins in Rice Antagonize Rice FT-Like Protein in Inflorescence Development by Competition for Complex Formation with 14-3-3 and FD. *Plant and Cell Physiology* 59: 458–468.

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 (New York, N.Y.)* 286: 1962–5.

Kaufmann K, Melzer R, Theißen G. 2005. MIKC-type MADS-domain proteins: Structural modularity, protein interactions and network evolution in land plants. *Gene* 347: 183–198.

Kennedy A, Geuten K. 2020. The Role of FLOWERING LOCUS C Relatives in Cereals. *Frontiers in Plant Science* 11: 2108.

Kikuchi R, Kawahigashi H, Ando T, Tonooka T, Handa H. 2009. Molecular and functional characterization of pebp genes in barley reveal the diversification of their roles in flowering. *Plant Physiology* 149: 1341–1353.

Kim J, Harter K, Theologis A. 1997. *Protein-protein interactions among the AuxIAA proteins (auxin signalingyeast two-hybrid systemprotein cross-linking-sheet DNA binding motifcombinatorial interactions*.

Kim S, Soltis PS, Wall K, Soltis DE. 2006. Phylogeny and Domain Evolution in the APETALA2-like Gene Family. *Molecular Biology and Evolution* 23: 107–120.

Kippes N, Chen A, Zhang X, Lukaszewski AJ, Dubcovsky J. 2016. Development and characterization of a spring hexaploid wheat line with no functional VRN2 genes. *Theoretical and Applied Genetics* 129: 1417–1428.

Kirby EJM, Appleyard M. 1984. Cereal development guide. 2nd Edition. *Cereal development guide. 2nd Edition.*

Kirkegaard JA, Hunt JR, McBeath TM, Lilley JM, Moore A, Verburg K, Robertson M, Oliver Y, Ward PR, Milroy S. 2014. Improving water productivity in the Australian Grains industry—a nationally coordinated approach. *http://dx.doi.org/10.1071/CP14019*.

Kitagawa S, Shimada S, Murai K. 2012. Effect of Ppd-1 on the expression of flowering-timegenes in vegetative and reproductivegrowth stages of wheat. *Genes and Genetic Systems* 87: 161–168.

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.

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

Komatsu M, Chujo A, Nagato Y, Shimamoto K, Kyozuka J. 2003. Frizzy panicle is required to prevent the formation of axillary meristems and to establish floral meristem identity in rice spikelets. *Development* 130: 3841–3850.

Komiya R, Ikegami A, Tamaki S, Yokoi S, Shimamoto K. 2008. Hd3a and RFT1 are essential for flowering in rice. *Development (Cambridge, England)* 135: 767–74.

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.

Koornneef M, Meinke D. 2010. The development of Arabidopsis as a model plant. *The Plant Journal* 61: 909–921.

Krasileva K V., Vasquez-Gross HA, Howell T, Bailey P, Paraiso F, Clissold L, Simmonds J, Ramirez-Gonzalez RH, Wang X, Borrill P. 2017. Uncovering hidden variation in polyploid wheat. *Proceedings of the National Academy of Sciences of the United States of America* 114: E913–E921.

Krogan NT, Marcos D, Weiner AI, Berleth T. 2016. The auxin response factor MONOPTEROS controls meristem function and organogenesis in both the shoot and root through the direct regulation of PIN genes. *The New phytologist* 212: 42–50.

Kumar SV, Lucyshyn D, Jaeger KE, Alós E, Alvey E, Harberd NP, Wigge PA. 2012. Transcription factor PIF4 controls the thermosensory activation of flowering. *Nature* 484: 242–245.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* 35: 1547–1549.

Kuzay S, Xu Y, Zhang J, Katz A, Pearce S, Su Z, Fraser M, Anderson JA, Brown-Guedira G, DeWitt N. 2019. Identification of a candidate gene for a QTL for spikelet number per spike on wheat chromosome arm 7AL by high-resolution genetic mapping. *Theoretical and Applied Genetics* 132: 2689–2705.

Laurie DA, Pratchett N, Bezant JH, Snape JW. 1995. RFLP mapping of five major genes and eight quantitative trait loci controlling flowering time in a winter X spring barley (Hordeum vulgare L.) cross. *Genome* 38: 575–585.

Leach LJ, Belfield EJ, Jiang C, Brown C, Mithani A, Harberd NP. 2014. Patterns of homoeologous gene expression shown by RNA sequencing in hexaploid bread wheat. *BMC Genomics* 15.

Lee R, Baldwin S, Kenel F, McCallum J, Macknight R. 2013. FLOWERING LOCUS T genes control onion bulb formation and flowering. *Nature Communications* 4.

Lee ZH, Hirakawa T, Yamaguchi N, Ito T. 2019. The roles of plant hormones and their interactions with regulatory genes in determining meristem activity. *International Journal of Molecular Sciences* 20.

Lee J, Lee I. 2010. Regulation and function of SOC1, a flowering pathway integrator.

*Journal of Experimental Botany* 61: 2247–2254.

Lee H, Suh SS, Park E, Cho E, Ahn JH, Kim SG, Lee JS, Kwon YM, Lee I. 2000. The AGAMOUS-LIKE 20 MADS domain protein integrates floral inductive pathways in Arabidopsis. *Genes & development* 14: 2366–76.

Li B, Dewey CN. 2011. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12: 323.

Li C, Dubcovsky J. 2008. Wheat FT protein regulates VRN1 transcription through interactions with FDL2. *Plant Journal* 55: 543–554.

Li Y, Fu X, Zhao M, Zhang W, Li B, An D, Li J, Zhang A, Liu R, Liu X. 2018. A Genome-wide View of Transcriptome Dynamics During Early Spike Development in Bread Wheat. *Scientific Reports* 8: 1–16.

Li Z, Li G, Cai M, Priyadarshani SVGN, Aslam M, Zhou Q, Huang X, Wang X, Liu Y, Qin Y. 2019a. Genome-wide analysis of the YABBY transcription factor family in pineapple and functional identification of AcYABBY4 involvement in salt stress. *International Journal of Molecular Sciences* 20.

Li C, Lin H, Chen A, Lau M, Jernstedt J, Dubcovsky J. 2019b. Wheat VRN1, FUL2 and FUL3 play critical and redundant roles in spikelet development and spike determinacy. *Development (Cambridge)* 146.

Li C, Lin H, Dubcovsky J. 2015. Factorial combinations of protein interactions generate a multiplicity of florigen activation complexes in wheat and barley. 84: 70–82.

Li N, Wang Y, Lu J, Liu C. 2019. Genome-wide identification and characterization of the ALOG domain genes in rice. *International Journal of Genomics* 2019.

Li S-B, Xie Z-Z, Hu C-G, Zhang J-Z. 2016. A Review of Auxin Response Factors (ARFs) in Plants. *Frontiers in plant science* 7: 47.

Li G, Yu M, Fang T, Cao S, Carver BF, Yan L. 2013. Vernalization requirement duration in winter wheat is controlled by TaVRN-A1 at the protein level. *Plant Journal* 76: 742–753.

Limin AE, Fowler DB. 2006. Low-temperature tolerance and genetic potential in wheat (Triticum aestivum L.): Response to photoperiod, vernalization, and plant development. *Planta* 224: 360–366.

Litt A, Irish VF. *Duplication and Diversification in the APETALA1/FRUITFULL Floral Homeotic Gene Lineage: Implications for the Evolution of Floral Development.*

Liu Y, Jiang HY, Chen W, Qian Y, Ma Q, Cheng B, Zhu S. 2011. Genome-wide analysis of

the auxin response factor (ARF) gene family in maize (Zea mays). *Plant Growth Regulation* 63: 225–234.

Liu H, Song S, Xing Y. 2019. Beyond heading time: *FT-* like genes and spike development in cereals. *Journal of Experimental Botany* 70: 1–3.

Liu X, Sun Z, Dong W, Wang Z, Zhang L. 2018. Expansion and functional divergence of the SHORT VEGETATIVe phase (SVP) genes in eudicots. *Genome Biology and Evolution* 10: 3026–3037.

Liu C, Teo ZWN, Bi Y, Song S, Xi W, Yang X, Yin Z, Yu H. 2013. A Conserved Genetic Pathway Determines Inflorescence Architecture in Arabidopsis and Rice. *Developmental Cell* 24: 612–622.

Liu N, Wu S, Houten J Van, Wang Y, Ding B, Fei Z, Clarke TH, Reed JW, Van Der Knaap E. 2014. Down-regulation of AUXIN RESPONSE FACTORS 6 and 8 by microRNA 167 leads to floral development defects and female sterility in tomato. *Journal of Experimental Botany* 65: 2507–2520.

Lv B, Nitcher R, Han X, Wang S, Ni F, Li K, Pearce S, Wu J, Dubcovsky J, Fu D. 2014. Characterization of FLOWERING LOCUS T1 (FT1) Gene in Brachypodium and Wheat (P Hernandez, Ed.). *PLoS ONE* 9: e94171.

Martin-Tryon EL, Harmer SL. 2008. XAP5 CIRCADIAN TIMEKEEPER coordinates light signals for proper timing of photomorphogenesis and the circadian clock in Arabidopsis. *Plant Cell* 20: 1244–1259.

Masterson J. 1994. Stomatal size in fossil plants: Evidence for polyploidy in majority of angiosperms. *Science* 264: 421–424.

Mathieu J, Warthmann N, Küttner F, Schmid M. 2007. Export of FT Protein from Phloem Companion Cells Is Sufficient for Floral Induction in Arabidopsis. *Current Biology* 17: 1055–1060.

Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. 2019. PANTHER version 14: More genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Research* 47: D419–D426.

Moon J, Lee H, Kim M, Lee I. 2005. Analysis of flowering pathway integrators in Arabidopsis. *Plant and Cell Physiology* 46: 292–299.

Moon J, Suh SS, Lee H, Choi KR, Hong CB, Paek NC, Kim SG, Lee I. 2003. The SOC1 MADS-box gene integrates vernalization and gibberellin signals for flowering in Arabidopsis. *Plant Journal* 35: 613–623.

Mulki MA, Bi X, von Korff M. 2018. Flowering locus T3 controls spikelet initiation but not floral development. *Plant Physiology* 178: 1170–1186.

Münster T, Ursula Wingen L, Faigl W, Werth S, Saedler H, Theißen G. 2001. Characterization of three GLOBOSA-like MADS-box genes from maize: Evidence for ancient paralogy in one class of floral homeotic B-function genes of grasses. *Gene* 262: 1–13.

Murai K. 2013. Homeotic genes and the ABCDE model for floral organ formation in wheat. *Plants* 2: 379–395.

Muterko A, Salina E. 2018. Origin and Distribution of the VRN-A1 Exon 4 and Exon 7 Haplotypes in Domesticated Wheat Species. *Agronomy* 8: 156.

Mutte SK, Kato H, Rothfels C, Melkonian M, Wong GKS, Weijers D. 2018. Origin and evolution of the nuclear auxin response system. *eLife* 7.

Nagasawa N, Miyoshi M, Sano Y, Satoh H, Hirano H, Sakai H, Nagato Y. 2003. SUPERWOMAN1 and DROOPING LEAF genes control floral organ identity in rice. *Development* 130: 705–718.

Nair N, Shoaib M, Sørensen CS. 2017. Chromatin dynamics in genome stability: Roles in suppressing endogenous DNA damage and facilitating DNA repair. *International Journal of Molecular Sciences* 18.

Nakamichi N, Kita M, Ito S, Yamashino T, Mizuno T. 2005. PSEUDO-RESPONSE REGULATORS, PRR9, PRR7 and PRR5, Together Play Essential Roles Close to the Circadian Clock of Arabidopsis thaliana. *Plant and Cell Physiology* 46: 686–698.

Nelson R, Wiesner-Hanks T, Wisser R, Balint-Kurti P. 2018. Navigating complexity to breed disease-resistant crops. *Nature Reviews Genetics* 19: 21–33.

Nilsson O, Lee I, Blázquez MA, Weigel D. 1998. *Flowering-Time Genes Modulate the Response to LEAFY Activity*.

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: 398–404.

O'Neill CM, Lu X, Calderwood A, Tudor EH, Robinson P, Wells R, Morris R, Penfield S. 2019. Vernalization and Floral Transition in Autumn Drive Winter Annual Life History in Oilseed Rape. *Current Biology* 29: 4300-4306.e2.

Ochagavía H, Prieto P, Savin R, Griffiths S, Slafer G. 2018. Dynamics of leaf and spikelet

primordia initiation in wheat as affected by Ppd-1a alleles under field conditions. *Journal of Experimental Botany* 69: 2621–2631.

Ochagavía H, Prieto P, Zikhali M, Griffiths S, Slafer GA. 2019. Earliness Per Se by Temperature Interaction on Wheat Development. *Scientific Reports* 9: 1–11.

Otto SP, Whitton J. 2000. Polyploid incidence and evolution. *Annual Review of Genetics* 34: 401–437.

Parvathaneni RK, Bertolini E, Shamimuzzaman M, Vera DL, Lung PY, Rice BR, Zhang J, Brown PJ, Lipka AE, Bass HW. 2020. The regulatory landscape of early maize inflorescence development. *Genome Biology* 21: 165.

Paterson AH, Brubaker CL, Wendel JF. 1993. A rapid method for extraction of cotton (Gossypium spp.) genomic DNA suitable for RFLP or PCR analysis. *Plant Molecular Biology Reporter* 11: 122–127.

Pelaz S, Ditta GS, Baumann E, Wisman E, Yanofsky MF. 2000. B and C floral organ identity functions require SEPALLATTA MADS-box genes. *Nature* 405: 200–203.

Peng FY, Hu Z, Yang R-C. 2015. Genome-Wide Comparative Analysis of Flowering-Related Genes in Arabidopsis, Wheat, and Barley. *International Journal of Plant Genomics* 2015: 1–17.

Pérez-Gianmarco TI, Slafer GA, González FG. 2019. Photoperiod-sensitivity genes shape floret development in wheat. *Journal of Experimental Botany* 70: 1339–1348.

Peters JL, Cnudde F, Gerats T. 2003. Forward genetics and map-based cloning approaches. *Trends in Plant Science* 8: 484–491.

Petersen G, Seberg O, Yde M, Berthelsen K. 2006. Phylogenetic relationships of Triticum and Aegilops and evidence for the origin of the A, B, and D genomes of common wheat (Triticum aestivum). *Molecular Phylogenetics and Evolution* 39: 70–82.

Pfeifer M, Kugler KG, Sandve SR, Zhan B, Rudi H, Hvidsten TR, Mayer KFX, Olsen OA, Rogers J, Doležel J. 2014. Genome interplay in the grain transcriptome of hexaploid bread wheat. *Science* 345.

Pimentel H, Bray NL, Puente S, Melsted P, Pachter L. 2017. Differential analysis of RNA-seq incorporating quantification uncertainty. *Nature Methods* 14: 687–690.

Pin PA, Benlloch R, Bonnet D, Wremerth-Weich E, Kraft T, Gielen JJL, Nilsson O. 2010. An antagonistic pair of FT homologs mediates the control of flowering time in sugar beet. *Science* 330: 1397–1400.

Pingali PL. 2012. Green revolution: Impacts, limits, andthe path ahead. *Proceedings of the National Academy of Sciences of the United States of America* 109: 12302–12308.

Pinyopich A, Ditta GS, Savidge B, Liljegren SJ, Baumann E, Wisman E, Yanofsky MF. 2003. Assessing the redundancy of MADS-box genes during carpel and ovule development. *Nature* 424: 85–88.

Piya S, Shrestha SK, Binder B, Neal Stewart C, Hewezi T. 2014. Protein-protein interaction and gene co-expression maps of ARFs and Aux/IAAs in Arabidopsis. *Frontiers in Plant Science* 5: 1–9.

Pokhilko A, Fernández AP, Edwards KD, Southern MM, Halliday KJ, Millar AJ. 2012. The clock gene circuit in Arabidopsis includes a repressilator with additional feedback loops. *Molecular systems biology* 8: 574.

Poursarebani N, Seidensticker T, Koppolu R, Trautewig C, Gawroński P, Bini F, Govind G, Rutten T, Sakuma S, Tagiri A. 2015. The genetic basis of composite spike form in barley and 'miracle-wheat'. *Genetics* 201: 155–165.

Prasad K, Sriram P, Santhosh Kumar C, Kushalappa K, Vijayraghavan U. 2001. Ectopic expression of rice OsMADS1 reveals a role in specifying the lemma and palea, grass floral organs analogous to sepals. *Development Genes and Evolution* 211: 281–290.

Prieto P, Ochagavía H, Savin R, Griffiths S, Slafer GA. 2018. Dynamics of floret initiation/death determining spike fertility in wheat as affected by Ppd genes under field conditions. *Journal of Experimental Botany* 69: 2633–2645.

Qin X, Liu JH, Zhao WS, Chen XJ, Guo ZJ, Peng YL. 2013. Gibberellin 20-Oxidase Gene OsGA20ox3 regulates plant stature and disease development in rice. *Molecular Plant-Microbe Interactions* 26: 227–239.

Rademacher EH, Möller B, Lokerse AS, Llavata-Peris CI, Van Den Berg W, Weijers D. 2011. A cellular expression map of the Arabidopsis AUXIN RESPONSE FACTOR gene family. *Plant Journal* 68: 597–606.

Ramírez-González RH, Borrill P, Lang D, Harrington SA, Brinton J, Venturini L, Davey M, Jacobs J, Van Ex F, Pasha A. 2018. The transcriptional landscape of polyploid wheat. *Science* 361.

Ramirez-Gonzalez RH, Segovia V, Bird N, Fenwick P, Holdgate S, Berry S, Jack P, Caccamo M, Uauy C. 2015a. RNA-Seq bulked segregant analysis enables the identification of high-resolution genetic markers for breeding in hexaploid wheat. *Plant Biotechnology Journal* 13: 613–624.

Ramirez-Gonzalez RH, Uauy C, Caccamo M. 2015b. PolyMarker: A fast polyploid primer design pipeline. *Bioinformatics* 31: 2038–2039.

Ramsay L, Comadran J, Druka A, Marshall DF, Thomas WTB, MacAulay M, MacKenzie K, Simpson C, Fuller J, Bonar N. 2011. INTERMEDIUM-C, a modifier of lateral spikelet fertility in barley, is an ortholog of the maize domestication gene TEOSINTE BRANCHED 1. *Nature Genetics* 43: 169–172.

Rijpkema AS, Vandenbussche M, Koes R, Heijmans K, Gerats T. 2010. Variations on a theme: Changes in the floral ABCs in angiosperms. *Seminars in Cell and Developmental Biology* 21: 100–107.

Rijpkema AS, Zethof J, Gerats T, Vandenbussche M. 2009. The petunia *AGL6* gene has a *SEPALLATA* -like function in floral patterning. *The Plant Journal* 60: 1–9.

Rockwell NC, Su YS, Lagarias JC. 2006. Phytochrome structure and signaling mechanisms. *Annual Review of Plant Biology* 57: 837–858.

Sakuma S, Golan G, Guo Z, Ogawa T, Tagiri A, Sugimoto K, Bernhardt N, Brassac J, Mascher M, Hensel G. 2019. Unleashing floret fertility in wheat through the mutation of a homeobox gene. *Proceedings of the National Academy of Sciences of the United States of America* 116: 5182–5187.

Sakuma S, Schnurbusch T. 2020. Of floral fortune: tinkering with the grain yield potential of cereal crops. *New Phytologist* 225: 1873–1882.

Salamini F, Ozkan H, Brandolini A, Schäfer-Pregl R, Martin W. 2002. Genetics and geography of wild cereal domestication in the near east. *Nature reviews. Genetics* 3: 429–41.

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.

Sasaki A, Ashikari M, Ueguchi-Tanaka M, Itoh H, Nishimura A, Swapan D, Ishiyama K, Saito T, Kobayashi M, Khush GS. 2002. A mutant gibberellin-synthesis gene in rice: New insight into the rice variant that helped to avert famine over thirty years ago. *Nature* 416: 701–702.

Sawa S, Watanabe K, Goto K, Kanaya E, Morita EH, Okada K. 1999. Filamentous flower, a meristem and organ identity gene of Arabidopsis, encodes a protein with a zinc finger and HMG-related domains. *Genes and Development* 13: 1079–1088.

Schaffer R, Ramsay N, Samach A, Corden S, Putterill J, Carré IA, Coupland G, Kay S.,

Nishimura Y. 1998. The late elongated hypocotyl mutation of Arabidopsis disrupts circadian rhythms and the photoperiodic control of flowering. *Cell* 93: 1219–29.

Schilling S, Gramzow L, Lobbes D, Kirbis A, Weilandt L, Hoffmeier A, Junker A, Weigelt-Fischer K, Klukas C, Wu F. 2015. Non-canonical structure, function and phylogeny of the BsisterMADS-box gene OsMADS30 of rice (Oryza sativa). *The Plant Journal* 84: 1059–1072.

Schilling S, Kennedy A, Pan S, Jermiin LS, Melzer R. 2020. Genome-wide analysis of MIKC-type MADS-box genes in wheat: pervasive duplications, functional conservation and putative neofunctionalization. *New Phytologist* 225: 511–529.

Schilling S, Pan S, Kennedy A, Melzer R. 2018. MADS-box genes and crop domestication: The jack of all traits. *Journal of Experimental Botany* 69: 1447–1469.

Schwarz R, Dayhoff M. 1979. Dayhoff 1979. *Atlas of protein sequences*: 353–358.

Shaikhali J, Davoine C, Björklund S, Wingsle G. 2016. Redox regulation of the MED28 and MED32 mediator subunits is important for development and senescence. *Protoplasma* 253: 957–963.

Sharman BC. 1944. Branched heads in wheat and wheat hybrids [7]. *Nature* 153: 497–498.

Sharman BC. 1978. Morphogenesis of 2,4-D Induced Abnormalities of the Inflorescence of Bread Wheat (Triticum aestivum, L.). *Annals of Botany* 42: 145–153.

Shaw LM, Li C, Woods DP, Alvarez MA, Lin H, Lau MY, Chen A, Dubcovsky J. 2020. Epistatic interactions between PHOTOPERIOD1, CONSTANS1 and CONSTANS2 modulate the photoperiodic response in wheat (S Hake, Ed.). *PLOS Genetics* 16: e1008812.

Shaw LM, Lyu B, Turner R, Li C, Chen F, Han X, Fu D, Dubcovsky J. 2019. FLOWERING LOCUS T2 regulates spike development and fertility in temperate cereals. *Journal of Experimental Botany* 70: 193–204.

Shaw LM, Turner AS, Herry L, Griffiths S, Laurie DA. 2013. Mutant Alleles of Photoperiod-1 in Wheat (Triticum aestivum L.) That Confer a Late Flowering Phenotype in Long Days (DE Somers, Ed.). *PLoS ONE* 8: e79459.

Shaw LM, Turner AS, Laurie DA. 2012. The impact of photoperiod insensitive Ppd-1a mutations on the photoperiod pathway across the three genomes of hexaploid wheat (Triticum aestivum). *The Plant Journal* 71: 71–84.

Slafer GA, Rawson HM. 1994. Sensitivity of wheat phasic development to major environmental factors: A re-examination of some assumptions made by physiologists and modellers. *Australian Journal of Plant Physiology* 21: 393–426.

Smaczniak C, Immink RGH, Angenent GC, Kaufmann K. 2012. Developmental and evolutionary diversity of plant MADS-domain factors: Insights from recent studies. *Development (Cambridge)* 139: 3081–3098.

Song YH, Ito S, Imaizumi T. 2013. Flowering time regulation: Photoperiod- and temperature-sensing in leaves. *Trends in Plant Science* 18: 575–583.

Song YH, Kubota A, Kwon MS, Covington MF, Lee N, Taagen ER, Laboy Cintrón D, Hwang DY, Akiyama R, Hodge SK. 2018. Molecular basis of flowering under natural long-day conditions in Arabidopsis. *Nature Plants* 4: 824–835.

Spielmeyer W, Ellis MH, Chandler PM. 2002. Semidwarf (sd-1), 'green revolution' rice, contains a defective gibberellin 20-oxidase gene. *Proceedings of the National Academy of Sciences of the United States of America* 99: 9043–9048.

Stephenson E, Estrada S, Meng X, Ourada J, Muszynski MG, Habben JE, Danilevskaya ON. 2019. Over-expression of the photoperiod response regulator ZmCCT10 modifies plant architecture, flowering time and inflorescence morphology in maize. *PLoS ONE* 14: e0203728.

Strayer C, Oyama T, Schultz TF, Raman R, Somers DE, Mas P, Panda S, Kreps JA, Kay SA. 2000. Cloning of the Arabidopsis clock gene TOC1, an autoregulatory response regulator homolog. *Science* 289: 768–771.

Strutt & Parker. 2019. *Strutt & Parker Yeild Results*. Cambridge.

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.

Sugiyama SH, Yasui Y, Ohmori S, Tanaka W, Hirano HY. 2019. Rice flower development revisited: Regulation of carpel specification and flower meristem determinacy. *Plant and Cell Physiology* 60: 1284–1295.

Takeda S, Hanano K, Kariya A, Shimizu S, Zhao L, Matsui M, Tasaka M, Aida M. 2011. CUP-SHAPED COTYLEDON1 transcription factor activates the expression of LSH4 and LSH3, two members of the ALOG gene family, in shoot organ boundary cells. 66: 1066–1077.

Takeda T, Suwa Y, Suzuki M, Kitano H, Ueguchi-Tanaka M, Ashikari M, Matsuoka M,

Ueguchi C. 2003. The OsTB1 gene negatively regulates lateral branching in rice. *Plant Journal* 33: 513–520.

Tamaki S, Matsuo S, Hann LW, Yokoi S, Shimamoto K. 2007. Hd3a protein is a mobile flowering signal in rice. *Science* 316: 1033–1036.

Taoka KI, Ohki I, Tsuji H, Furuita K, Hayashi K, Yanase T, Yamaguchi M, Nakashima C, Purwestri YA, Tamaki S. 2011. 14-3-3 proteins act as intracellular receptors for rice Hd3a florigen. *Nature* 476: 332–335.

Teper-Bamnolker P, Samach A. 2005. The flowering integrator FT regulates SEPALLATA3 and FRUITFULL accumulation in Arabidopsis leaves. *Plant Cell* 17: 2661–2675.

Theißen G, Rümpler F, Gramzow L. 2018. Array of MADS-Box Genes: Facilitator for Rapid Adaptation? *Trends in Plant Science* 23: 563–576.

Trevaskis B, Bagnall DJ, Ellis MH, Peacock WJ, Dennis ES. 2003. MADS box genes control vernalization-induced flowering in cereals. *Proceedings of the National Academy of Sciences of the United States of America* 100: 13099–13104.

Trevaskis B, Hemming MN, Dennis ES, Peacock WJ. 2007. The molecular basis of vernalization-induced flowering in cereals. *Trends in Plant Science* 12: 352–357.

Trevaskis B, Hemming MN, Peacock WJ, Dennis ES. 2006. HvVRN2 responds to daylength, whereas HvVRN1 is regulated by vernalization and developmental status. *Plant Physiology* 140: 1397–1405.

Tsujimoto Y, Numaga T, Ohshima K, Yano MA, Ohsawa R, Goto DB, Naito S, Ishikawa M. 2003. Arabidopsis tobamovirus multiplication (TOM) 2 locus encodes a transmembrane protein that interacts with TOM1. *EMBO Journal* 22: 335–343.

Turner A, Beales J, Faure S, Dunford RP, Laurie DA. 2005a. The Pseudo-Response Regulator Ppd-H1 Provides Adaptation to Photoperiod in Barley. *Science* 310.

Tyedmers J, Mogk A, Bukau B. 2010. Cellular strategies for controlling protein aggregation. *Nature Reviews Molecular Cell Biology* 11: 777–788.

United Nations. *World Population Prospects The 2015 Revision.*

Valverde F, Mouradov A, Soppe W, Ravenscroft D, Samach A, Coupland G. 2004. Photoreceptor Regulation of CONSTANS Protein in Photoperiodic Flowering. *Science* 303: 1003–1006.

Vierling E. 1991. The Roles of Heat Shock Proteins in Plants. *Annual Review of Plant*

*Physiology and Plant Molecular Biology* 42: 579–620.

de Vleesschauwer D, Seifi HS, Filipe O, Haeck A, Huu SN, Demeestere K, Höfte M. 2016. The DELLA protein SLR1 integrates and amplifies salicylic acid- and jasmonic acid-dependent innate immunity in rice. *Plant Physiology* 170: 1831–1847.

Wagner GP, Kin K, Lynch VJ. 2012. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory in Biosciences* 131: 281–285.

Wang Y, Cheng X, Shan Q, Zhang Y, Liu J, Gao C, Qiu JL. 2014. Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nature Biotechnology* 32: 947–951.

Wang J-W, Czech B, Weigel D. 2009. miR156-Regulated SPL Transcription Factors Define an Endogenous Flowering Pathway in Arabidopsis thaliana. *Cell* 138: 738–749.

Wang ZY, Kenigsbuch D, Sun L, Harel E, Ong MS, Tobin EM. 1997. A Myb-related transcription factor is involved in the phytochrome regulation of an arabidopsis Lhcb gene. *Plant Cell* 9: 491–507.

Wang D, Pei K, Fu Y, Sun Z, Li S, Liu H, Tang K, Tao Y. 2007. Genome-wide analysis of the auxin response factors (ARF) gene family in rice (Oryza sativa). *Elsevier*.

Wang ZY, Tobin EM. 1998. Constitutive expression of the CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) gene disrupts circadian rhythms and suppresses its own expression. *Cell* 93: 1207–1217.

Wang W, Vinocur B, Shoseyov O, Altman A. 2004. Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends in Plant Science* 9: 244–252.

Wang E, Wang J, Zhu X, Hao W, Wang L, Li Q, Zhang L, He W, Lu B, Lin H. 2008. Control of rice grain-filling and yield by a gene with a potential signature of domestication. *Nature Genetics* 40: 1370–1374.

Wang Y, Yu H, Tian C, Sajjad M, Gao C, Tong Y, Wang X, Jiao Y. 2017. Transcriptome association identifies regulators of wheat spike architecture. *Plant Physiology* 175: 746–757.

Wang H, Zhang L, Cai Q, Hu Y, Jin Z, Zhao X, Fan W, Huang Q, Luo Z, Chen M. 2015a. OsMADS32 interacts with PI-like proteins and regulates rice flower development. *Journal of Integrative Plant Biology* 57: 504–513.

Wang Z, Zhou Z, Liu Y, Liua T, Li Q, Ji Y, Li C, Fang C, Wang M, Wu M. 2015b.

Functional evolution of phosphatidylethanolamine binding proteins in soybean and arabidopsis. *Plant Cell* 27: 323–336.

Weiste C, Dröge-Laser W. 2014. The Arabidopsis transcription factor bZIP11 activates auxin-mediated transcription by recruiting the histone acetylation machinery. *Nature Communications* 5: 1–12.

Wheat Initiative. 2017. Wheat Research for Food Security | Wheat Initiative.

Whitechurch EM, Slafer GA. 2002. Contrasting Ppd alleles in wheat: Effects on sensitivity to photoperiod in different phases. *Field Crops Research* 73: 95–105.

Wigge PA, Kim MC, Jaeger KE, Busch W, Schmid M, Lohmann JU, Weigel D. 2005. Integration of Spatial and Temporal Information During Floral Induction in Arabidopsis. *Science* 309: 1056–1059.

Wilhelm EP, Turner AS, Laurie DA. 2009. Photoperiod insensitive Ppd-A1a mutations in tetraploid wheat (Triticum durum Desf.). *Theoretical and Applied Genetics* 118: 285–294.

Willmann MR, Poethig RS. 2011. The effect of the floral repressor FLC on the timing and progression of vegetative phase change in Arabidopsis. *Development* 138: 677–685.

Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB, Rieseberg LH. 2009. The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences of the United States of America* 106: 13875–13879.

Woods D, Dong Y, Bouche F, Bednarek R, Rowe M, Ream T, Amasino R. 2019. A florigen paralog is required for short-day vernalization in a pooid grass. *eLife* 8.

Worland AJ. 1996. The influence of flowering time genes on environmental adaptability in European wheats. *Euphytica* 89: 49–57.

Worland AJ, Bö Rner A, Korzun V, Li WM, Petrovíc S, Sayers EJ. 1998. *The influence of photoperiod genes on the adaptability of European winter wheats*.

Würschum T, Jähne F, Phillips AL, Langer SM, Longin CFH, Tucker MR, Leiser WL. 2020. Misexpression of a transcriptional repressor candidate provides a molecular mechanism for the suppression of awns by Tipped 1 in wheat (G Rebetzke, Ed.). *Journal of Experimental Botany* 71: 3428–3436.

Xiao J, Xu S, Li C, Xu Y, Xing L, Niu Y, Huan Q, Tang Y, Zhao C, Wagner D. 2014. O-GlcNAc-mediated interaction between VER2 and TaGRP2 elicits TaVRN1 mRNA accumulation during vernalization in winter wheat. *Nature Communications* 5.

Xin Z, Browse J. 2000. Cold comfort farm: The acclimation of plants to freezing

temperatures. *Plant, Cell and Environment* 23: 893–902.

Xing H, Pudake RN, Guo G, Xing G, Hu Z, Zhang Y, Sun Q, Ni Z. 2011. Genome-wide identification and expression profiling of auxin response factor (ARF) gene family in maize. *BMC Genomics* 12: 178.

Yamaguchi A, Kobayashi Y, Goto K, Abe M, Araki T. 2005. TWIN SISTER of FT (TSF) acts as a floral pathway integrator redundantly with FT. *Plant and Cell Physiology* 46: 1175–1189.

Yamaguchi T, Lee DY, Miyao A, Hirochika H, An G, Hirano HY. 2006. Functional diversification of the two C-class MADS box genes OSMADS3 and OSMADS58 in Oryza sativa. *Plant Cell* 18: 15–28.

Yamanaka T, Ohta T, Takahashi M, Meshi T, Schmidt R, Dean C, Naito S, Ishikawa M. 2000. TOM1, an Arabidopsis gene required for efficient multiplication of a tobamovirus, encodes a putative transmembrane protein. *Proceedings of the National Academy of Sciences of the United States of America* 97: 10107–10112.

Yan L, Fu D, Li C, Blechl A, Tranquilli G, Bonafede M, Sanchez A, Valarik M, Yasuda S, Dubcovsky J. 2006. The wheat and barley vernalization gene VRN3 is an orthologue of FT. *Proceedings of the National Academy of Sciences of the United States of America* 103: 19581–19586.

Yan L, Loukoianov A, Blechl A, Tranquilli G, Ramakrishna W, SanMiguel P, Bennetzen JL, Echenique V, Dubcovsky J. 2004. The Wheat VRN2 Gene Is a Flowering Repressor Down-Regulated by Vernalization. *Science* 303: 1640–1644.

Yan L, Loukoianov A, Tranquilli G, Helguera M, Fahima T, Dubcovsky J. 2003. Positional cloning of the wheat vernalization gene VRN1. *Proceedings of the National Academy of Sciences of the United States of America* 100: 6263–6268.

Yang DL, Li Q, Deng YW, Lou YG, Wang MY, Zhou GX, Zhang YY, He ZH. 2008. Altered disease development in the eui mutants and Eui overexpressors indicates that gibberellins negatively regulate rice basal disease resistance. *Molecular Plant* 1: 528–537.

Yang C, Ma Y, Li J. 2016. The rice YABBY4 gene regulates plant growth and development through modulating the gibberellin pathway. *Journal of Experimental Botany* 67: 5545–5556.

Yang J, Yuan Z, Meng Q, Huang G, Périn C, Bureau C, Meunier A-C, Ingouff M, Bennett MJ, Liang W. 2017. Dynamic Regulation of Auxin Response during Rice Development Revealed by Newly Established Hormone Biosensor Markers. *Frontiers in Plant Science* 8: 256.

Yanofsky MF, Ma H, Bowman JL, Drews GN, Feldmann KA, Meyerowitz EM. 1990. The protein encoded by the Arabidopsis homeotic gene agamous resembles transcription factors. *Nature* 346: 35–39.

Yanovsky MJ, Kay SA. 2002. Molecular basis of seasonal time measurement in Arabidopsis. *Nature* 419: 308–312.

Yen C, University JY-J of SA, 1992 undefined. THE ESSENTIAL NATURE OF ORGANS IN GRAMINEAE, MULTIPLE SECONDARY AXES THEORY--A NEW CONCEPT. *en.cnki.com.cn*.

Yi L, Pimentel H, Bray NL, Pachter L. 2018. Gene-level differential analysis at transcript-level resolution. *Genome Biology* 19: 53.

Yin X, Liu X, Xu B, Lu P, Dong T, Yang D, Ye T, Feng YQ, Wu Y. 2019. OsMADS18, a membrane-bound MADS-box transcription factor, modulates plant architecture and the abscisic acid response in rice. *Journal of Experimental Botany* 70: 3895–3909.

Yoo SKSJSY, Chung KS, Kim J, Lee JHJS, Hong SM, Yoo SKSJSY, Yoo SKSJSY, Lee JHJS, Ahn JH. 2005. CONSTANS activates SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 through FLOWERING LOCUS T to promote flowering in Arabidopsis. *Plant Physiology* 139: 770–778.

Yoshida H, Nagato Y. 2011a. Flower development in rice. *Journal of Experimental Botany* 62: 4719–4730.

Yoshida H, Nagato Y. 2011b. *Flower development in rice*. Oxford Academic.

Yoshida A, Sasao M, Yasuno N, Takagi K, Daimon Y, Chen R, Yamazaki R, Tokunaga H, Kitaguchi Y, Sato Y. 2013. TAWAWA1, a regulator of rice inflorescence architecture, functions through the suppression of meristem phase transition. *Proceedings of the National Academy of Sciences of the United States of America* 110: 767–72.

Yoshida A, Suzaki T, Tanaka W, Hirano HY. 2009. The homeotic gene long sterile lemma (G1) specifies sterile lemma identity in the rice spikelet. *Proceedings of the National Academy of Sciences of the United States of America* 106: 20103–20108.

Young MD, Wakefield MJ, Smyth GK, Oshlack A. 2010. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biology* 11: R14.

Zadoks JC, CHang TT, Konzak CF. 1974. A decimal code for the growth stages of cereals. *Weed Research* 14: 415–421.

Zakhrabekova S, Gough SP, Braumann I, Mu̇ler AH, Lundqvist J, Ahmann K, Dockter C,

Matyszczak I, Kurowska M, Druka A. 2012. Induced mutations in circadian clock regulator Mat-a facilitated short-season adaptation and range extension in cultivated barley. *Proceedings of the National Academy of Sciences of the United States of America* 109: 4326–4331.

Zhao Y. 2010. Auxin biosynthesis and its role in plant development. *Annual Review of Plant Biology* 61: 49–64.

Zhao J, Chen H, Ren D, Tang H, Qiu R, Feng J, Long Y, Niu B, Chen D, Zhong T. 2015. Genetic interactions between diverged alleles of *Early heading date 1* ( *Ehd1* ) and *Heading date 3a* ( *Hd3a* ) */ RICE FLOWERING LOCUS T1* ( *RFT1* ) control differential heading and contribute to regional adaptation in rice ( *Oryza sativa* ). *New Phytologist* 208: 936–948.

Zhao T, Ni Z, Dai Y, Yao Y, Nie X, Sun Q. 2006. Characterization and expression of 42 MADS-box genes in wheat (Triticum aestivum L.). *Molecular Genetics and Genomics* 276: 334–350.

Zhu QH, Hoque MS, Dennis ES, Upadhyaya NM. 2003. Ds tagging of BRANCHED FLORETLESS 1 (BFL1) that mediates the transition from spikelet to floret meristem in rice (Oryza sativa L). *BMC Plant Biology* 3.

Zikhali M, Wingen LU, Griffiths S. 2016. Delimitation of the Earliness per se D1 (Eps-D1) flowering gene to a subtelomeric chromosomal deletion in bread wheat (Triticum aestivum). *Journal of Experimental Botany* 67: 287–299.

# Appendices

# Appendix 1

Gauley, A., & Boden, S. A. 2020. Step-wise increases in *FT1* expression regulate seasonal progression of flowering in wheat (*Triticum aestivum*). *New Phytologist.* https://doi.org/10.1111/NPH.16910

# Stepwise increases in *FT1* expression regulate seasonal progression of flowering in wheat (*Triticum aestivum*)

## Adam Gauley[1] and Scott A. Boden[1,2] (iD)

[1]Department of Crop Genetics, John Innes Centre, Colney Lane, Norwich, NR4 7UH, UK; [2]School of Agriculture, Food and Wine, Waite Research Institute, University of Adelaide, Glen Osmond, SA 5064, Australia

Author for correspondence
*Scott A. Boden*
*Email: scott.boden@adelaide.edu.au*

## Summary

- Flowering is regulated by genes that respond to changing daylengths and temperature, which have been well studied using controlled conditions; however, the molecular processes underpinning flowering in nature remain poorly understood.
- Here, we investigate the genetic pathways that coordinate flowering and inflorescence development of wheat (*Triticum aestivum*) as daylengths extend naturally in the field, using lines that contain variant alleles for the key photoperiod gene, *Photoperiod-1* (*Ppd-1*).
- We found flowering involves a stepwise increase in the expression of *FLOWERING LOCUS T1* (*FT1*), which initiates under day-neutral conditions of early spring. The incremental rise in *FT1* expression is overridden in plants that contain a photoperiod-insensitive allele of *Ppd-1*, which hastens the completion of spikelet development and accelerates flowering time. The accelerated inflorescence development of photoperiod-insensitive lines is promoted by advanced seasonal expression of floral meristem identity genes. The completion of spikelet formation is promoted by *FLOWERING LOCUS T2*, which regulates spikelet number and is activated by *Ppd-1*.
- In wheat, flowering under natural photoperiods is regulated by stepwise increases in the expression of *FT1*, which responds dynamically to extending daylengths to promote early inflorescence development. This research provides a strong foundation to improve yield potential by fine-tuning the photoperiod-dependent control of inflorescence development.

## Introduction

Plants use environmental signals including daylength and temperature to determine the optimal time to flower. By monitoring changes in daylength, plants flower with remarkable seasonal precision despite variable environmental conditions such as fluctuating daily temperatures. Temperate cereals such as bread wheat (*Triticum aestivum*) perceive the extending daylengths of spring to promote flowering so that seed production occurs under favourable conditions (Worland *et al.*, 1998; Fjellheim *et al.*, 2014). Alleles that modify the plant's response to these seasonal cues have been used to expand the geographical range of cultivation and improve productivity in marginal environments – these alleles typically accelerate flowering by reducing the requirement for long days or low temperatures (Danyluk *et al.*, 2003; Trevaskis *et al.*, 2003; Turner *et al.*, 2005; Beales *et al.*, 2007).

In wheat, the responsiveness to daylength (photoperiod) is largely determined by allelic diversity for *Photoperiod-1* (*Ppd-1*) (Laurie *et al.*, 1995; Turner *et al.*, 2005; Beales *et al.*, 2007). *Ppd-1* influences flowering by modifying the expression of *FLOWERING LOCUS T1* (*FT1*), which is a conserved activator of flowering in plants (Turner *et al.*, 2005; Yan *et al.*, 2006; Beales *et al.*, 2007; Wilhelm *et al.*, 2009; Kitagawa *et al.*, 2012;

Boden *et al.*, 2015; Bratzel & Turck, 2015). FT1 protein is expressed in leaves and transported to the shoot apical meristem (SAM), where it forms a complex with FLOWERING LOCUS D-LIKE (FDL) and 14-3-3 proteins (Corbesier *et al.*, 2007; Tamaki *et al.*, 2007; Li & Dubcovsky, 2008; Taoka *et al.*, 2011). The complex activates expression of meristem identity genes, which promote reproductive development of the inflorescence meristem (IM) (Corbesier *et al.*, 2007; Tamaki *et al.*, 2007; Li & Dubcovsky, 2008). In hexaploid wheat, photoperiod-insensitive alleles of *Ppd-1* activate *FT1* expression in the absence of long-day photoperiods – these alleles carry deletions in the *cis*-regulatory regions of *Ppd-1* or additional copies of the gene on the A, B and D genomes (termed *Ppd-A1a*, *Ppd-B1a* and *Ppd-D1a*, respectively) (Beales *et al.*, 2007; Diaz *et al.*, 2012; Kitagawa *et al.*, 2012; Seki *et al.*, 2013). Under constant short daylengths, the *cis*-regulatory mutations alter the daily rhythms of *Ppd-1* expression, causing it to be expressed in the evening when it is otherwise suppressed. While photoperiod-insensitive alleles confer early flowering phenotypes under all growth conditions, our understanding of the molecular function of *Ppd-1* comes from experiments performed in controlled growth environments (Beales *et al.*, 2007; Wilhelm *et al.*, 2009; Diaz *et al.*, 2012; Kitagawa *et al.*, 2012; Boden *et al.*, 2015). These experiments used

extreme short (8–9 h) or long (15–16 h) daylengths, which are different from the photoperiods when plants initiate flowering naturally in the field (11–13 h). These analyses have also focused on the role of *Ppd-1* in leaves, with little attention given to the impact of photoperiod-insensitive alleles on expression of meristem identity genes in the developing inflorescence. Given that photoperiod-insensitive alleles reduce yield potential by significantly decreasing spikelet number and the survival of floret primordia (i.e. floret fertility), it is vital that we learn more about the role of *Ppd-1* in inflorescences of field-grown plants under natural daylengths to devise strategies for improving grain production (González *et al.*, 2005; Fischer *et al.*, 2014; González-Navarro *et al.*, 2015; Prieto *et al.*, 2018; Perez-Gianmarco *et al.*, 2019).

To understand how flowering is regulated in a seasonal context, we investigated the photoperiod-dependent flowering pathway and early inflorescence development under natural photoperiods in the field. We used near-isogenic lines (NILs) containing photoperiod-insensitive, sensitive and null alleles of *Ppd-1* to genetically alter the ability of wheat to perceive changes in daylength – the insensitive *Ppd-D1a* allele represents the majority of wheat grown in spring-type mega-environments (Shaw *et al.*, 2013). Our work demonstrates that the floral-promoting pathway of wheat dynamically responds to increasing daylengths to regulate spikelet and floret development, and presents new genetic targets for yield improvement.

## Materials and Methods

### Plant material and growth conditions

Hexaploid wheat (*Triticum aestivum* L.) used here included: wild-type photoperiod-sensitive cv Paragon; Paragon NILs containing the *Ppd-D1a* photoperiod-insensitive allele (Shaw *et al.*, 2013) or null *ppd-1* alleles on the A, B and D genomes (Shaw *et al.*, 2013); and two missense *ft-B2* mutants (*Cad0122* and *Cad1655*) obtained from the hexaploid wheat TILLING population (Krasileva *et al.*, 2017). The *ft-B2* mutations were verified using segregation analysis – mutant NILs were compared to cv Cadenza and wild-type sibling lines.

Plants were grown at field sites of the John Innes Centre, Norwich, UK (52°62′25.7″N, 1°21′83.2″E) in 1 m$^2$ plots, and in glasshouses under natural temperature and photoperiod conditions (Supporting Information Fig. S1). Seeds were sown in week 2–3 of October for the field experiments, and at the end of October for the glasshouse experiment – we sowed later in the glasshouse because the plants established faster under these conditions. Phenotypic data were collected over two growing seasons (2017 and 2018), and molecular data over two growing seasons (2018 and 2019). The *ft-B2.1* mutant and wild-type controls were grown in a glasshouse under 16 h : 8 h, light : dark. The *ft-B2.2* and wild-type controls were grown under extralong daylengths (22 h : 2h).

For the moving bench experiment, *Ppd-D1a*, *ppd-1* and wild-type lines were grown in a glasshouse under natural photoperiods until the daylength reached 10 h. At 10 h, plants were shifted to a glasshouse with moving benches that transferred plants to a dark chamber after 10 h of natural daylight (Fig. S2). The experiment was repeated for two seasons.

### RNA extractions and expression analysis

Leaves from wild-type, *Ppd-D1a* and *ppd-1* NILs were sampled every 3–4 h during the day and night, and at dusk (sunset). Sample time-points are expressed in terms of Zeitgeber time (ZT), with sunrise being 0 h. Leaf samples were harvested at photoperiods defined by 1 h increases in daylength, commencing at a photoperiod of 9 h and ending at 13 h (Fig. S1). Three biological replicates were collected per time-point, each from the most recently emerged leaf of the primary tiller. For inflorescence expression analyses, samples were collected at four stages: vegetative (VG), double-ridge (DR), lemma primordia (LP) and terminal spikelet (TS). Inflorescences for each genotype were sampled at ZT 6–8 h. Each sample included pools of five to 15 inflorescences per replicate, dependent on stage. Three biological replicates were collected per stage.

Leaf RNA extractions were performed using the Spectrum Plant Total RNA Kit (Sigma-Aldrich). RNA extractions from developing inflorescences were performed using the RNeasy Plant Mini Kit (Qiagen). cDNA synthesis and reverse transcriptase quantitative PCR (RT-qPCR) were performed as described previously, using a LightCycler480 Instrument II (Roche Life Science) (Boden *et al.*, 2014). RT-qPCR oligonucleotides are listed in Table S1 – the oligonucleotides amplify all three homoeoalleles for *FT2*, *VRN1* and the meristem identity genes; *Ppd-1* sequences are homoeoallele-specific. Candidate gene expression from leaf and inflorescence was normalized using TraesCS6D02G145100 (Traes_6DS_ BE8B5E56D.1), which we previously verified to be stably expressed in leaves and inflorescences across different photoperiods (Borrill *et al.*, 2016; Dixon *et al.*, 2018b). RT-qPCR data are the average of at least three biological replicates and two technical replicates per reaction. Expression ribbon plots were created using the R package GGPLOT2 (Wickham, 2016).

For the MADS-box transcription factors analysed here, we use the proposed nomenclature of Schilling *et al.* (2020), which are different from those we have described previously (Dixon *et al.*, 2018b).

### Inflorescence architecture measurements

Spikelet number was counted for the developing inflorescence at sequential stages. At the double-ridge stage, the spikelet meristem ridge was counted as a spikelet. From the lemma primordium stage onwards, spikelet meristems were clearly visible and counted as a spikelet. For fully emerged inflorescences, both viable and nonviable spikelets were counted. For *Ppd-D1a*-insensitive and *ppd-1* lines, data of early inflorescence development are the average ± SEM of at least three replicates. For final spikelet numbers, data are the average ± SEM of at least 10 replicates. For *ft-B2* mutants, spikelet data are the average ± SEM of 5–7 replicates.

### DNA extractions and sequence analysis

Genomic DNA extractions were carried out as described previously (Dixon *et al.*, 2018b). All gene sequences were obtained by BLAST searches (Ensembl Plants). The primers used for *ft-b2* mutant alleles are listed in Table S1. DNA fragments were amplified using Phusion DNA polymerase (New England Biolabs, Ipswich, MA, USA). Amplicon sequencing was carried out with Mix2Seq Kit (Eurofins, Luxembourg).

### Kompetitive allele-specific PCR analysis

Kompetitive allele-specific PCR (KASP) analysis was used to analyse TILLING mutant lines, before verification of mutations using sequence analysis. Oligonucleotides were designed using POLYMARKER (Ramirez-Gonzalez *et al.*, 2015) and they contained either the standard FAM or HEX compatible tails (Table S2). The assay was performed as described previously (Dixon *et al.*, 2018b).

Additional methods are provided in Methods S1.

## Results

### Seasonal and genetic regulation of *Photoperiod-1*

To investigate the seasonal regulation of the flowering pathway, we first measured the response of *Ppd-1* to increasing daylengths in the field. Specifically, we analysed expression of *Ppd-B1* and *Ppd-D1* in photoperiod-sensitive wild-type plants (cv Paragon), as these homoeologues contribute the major photoperiod-insensitive alleles that confer early-flowering phenotypes of hexaploid wheat in global breeding programmes (*Ppd-B1a* and *Ppd-D1a*, respectively) (Beales *et al.*, 2007; Diaz *et al.*, 2012; Shaw *et al.*, 2013). Transcripts were measured over a series of photoperiods defined by hourly increases in daylength from winter (9 h : 15 h light : dark) until late spring (13 h : 10 h light : dark), and diurnal patterns were analysed to precisely detect the daily peak(s) in gene expression (Figs 1, S1; Dataset S1). *Ppd-D1* and *Ppd-B1* were expressed at comparable levels to each other and they displayed very similar daily expression profiles (Fig. 1a–c,e–f). The diurnal rhythm of *Ppd-D1* and *-B1* was maintained across all photoperiods, with transcripts peaking during the day (ZT 3–6 h) and at dusk, and dipping during the night (ZT 16–24 h). The consistent diel pattern in relation to dawn and dusk indicates *Ppd-1* expression adjusts to the changing daylengths. The amplitude of *Ppd-D1* expression was stable across all photoperiods tested (Fig. 1a,b,e). *Ppd-B1* transcript levels were highest under 9 h photoperiods, before maintaining a normalized range of transcript peaks between 0.04 and 0.09, which were slightly higher at 10 and 13 h photoperiods (Fig. 1c,f). These results suggest that the seasonal regulation of flowering in field-grown wheat is not determined by quantitative changes in *Ppd-1* expression.

To determine how photoperiod-insensitive alleles modify *Ppd-1* expression under field conditions, we used an NIL expressing the early-flowering *Ppd-D1a* allele that contains a 2.09 kb promoter deletion (Beales *et al.*, 2007; Shaw *et al.*, 2013). The

photoperiod-insensitive *Ppd-D1a* allele altered the diurnal expression of *Ppd-D1* from the 10–13 h photoperiods by promoting higher expression late at night (ZT 20 and 24 h), relative to wild-type (Fig. 1a; Dataset S1). This difference was particularly significant at daylengths of 11 and 12 h; the difference in night-time expression was not detected in the 9 h photoperiod. There were also minor changes in expression of *Ppd-D1* during the day in the 13 h photoperiod (Fig. 1a). The insensitive *Ppd-D1a* allele did not affect the amplitude of *Ppd-D1* expression, relative to wild-type, especially during the daytime when *Ppd-D1* peaked between 3 and 6 h after dawn in both genotypes (Fig. 1a, e). *Ppd-B1* expression was also affected in the photoperiod-insensitive NIL, particularly during the 10 and 11 h photoperiods, where it modified the timing of transcript peaks during the day (10 h) and increased the amplitude of *Ppd-B1* expression at certain time-points (Fig. 1c). No significant effects on *Ppd-B1* expression were detected during the 12 and 13 h photoperiods. Based on these results, we conclude that the insensitive *Ppd-D1a* allele mis-regulates *Ppd-D1* expression during late hours of the night, particularly under day-neutral photoperiods of early spring, and there may be an interaction between *Ppd-1* homoeologues.

To genetically investigate the contribution of *Ppd-1* on seasonal regulation of flowering-time genes, we analysed NILs that contain nonfunctional alleles for all three homoeoalleles (i.e. *ppd-1* NILs). These lines carry deletions for *Ppd-A1* and *Ppd-B1* and a nonsense allele of *Ppd-D1*, which produces a transcript with a premature stop codon (Shaw *et al.*, 2013). *Ppd-D1* transcript levels were significantly higher during the day of the 10–13 h photoperiods at multiple time-points, relative to wild-type, and, as expected, no *Ppd-B1* transcripts were detected (Figs 1b,c,e, S3; Dataset S1).

An unexpected outcome of the field-based *Ppd-D1* expression analysis was that misregulated expression of *Ppd-D1* in the photoperiod-insensitive line was limited to late hours of the night of the 10–12 h photoperiods, as previous analyses using controlled environments showed insensitive *Ppd-D1a* alleles alter expression during all hours of the evening, particularly under short-day photoperiods (9 h : 15 h) (Fig. 1a; Dataset S1) (Beales *et al.*, 2007; Shaw *et al.*, 2012; Boden *et al.*, 2015). To investigate whether there is a difference between controlled-environment and field conditions, we compared the expression of *Ppd-D1* from field- and glasshouse-grown plants (Figs 1d, S3). We found that the photoperiod-insensitive allele promoted significantly higher expression of *Ppd-D1* during all hours of the night in all photoperiods for glasshouse-grown plants, relative to wild-type, consistent with previous reports that used controlled environments (Fig. 1d) (Beales *et al.*, 2007; Shaw *et al.*, 2012; Boden *et al.*, 2015). These data indicate that controlled growth conditions exaggerate the impact of photoperiod-insensitive alleles on *Ppd-D1* expression, relative to field-grown plants. In wild-type, the differences in *Ppd-D1* expression between field- and glasshouse-grown plants were less dramatic – the transcript peaks were slightly higher in glasshouse-grown plants during the day and the troughs were moderately lower during the night, relative to the field-grown plants (Fig. 1d). In the *ppd-1* NIL, we observed a
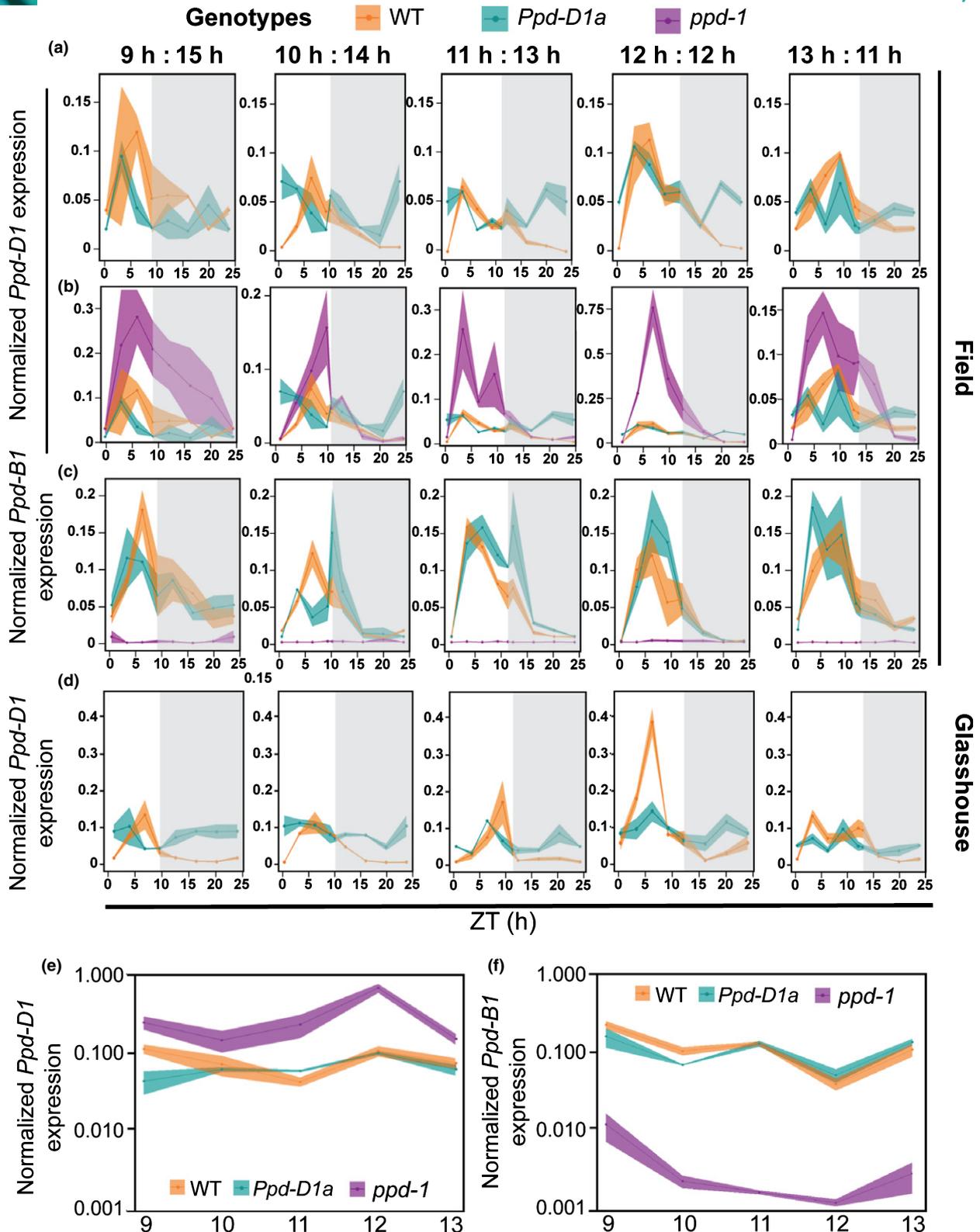
New
Phytologist



**Fig. 1** Seasonal regulation of *Ppd-1* under field- and glasshouse-based conditions in wheat. (a, b) Diurnal expression profiles of *Ppd-D1* in wild-type (WT, orange), *Ppd-D1a* photoperiod-insensitive (cyan) and null *ppd-1* (magenta) near-isogenic lines (NILs) under field conditions. In (a), the *ppd-1* data have been removed to show differences in transcript levels between wild-type and *Ppd-D1a* NILs. (c) Diurnal expression profiles of *Ppd-B1* in wild-type (orange), *Ppd-D1a* photoperiod-insensitive (cyan) and null *ppd-1* (magenta) NILs under field conditions. (d) Expression of *Ppd-D1* in WT and *Ppd-D1a* NILs under glasshouse conditions. (e, f) Data summarizing the peak of (e) *Ppd-D1* and (f) *Ppd-B1* expression for each daylength, plotted on a logarithmic scale. The grey shading highlights night-time data. All expression profiles are shown over a 24 h period at hourly incremental increases in daylength, with time represented as Zeitgeber time (ZT), with sunrise being 0 h. Graphs are presented as ribbon plots which show the normalized mean transcript levels (solid line with data points) ± SE (shaded region) of three biological replicates.

similar dramatic increase in *Ppd-D1* expression that was detected in the field, supporting the suggestion that *Ppd-1* forms a self-regulatory feedback loop (Fig. S3).

## Stepwise increases in *FT1* expression underpin seasonal regulation of flowering

To investigate the seasonal progression of major flowering-time genes that act downstream of *Ppd-1*, we analysed expression of *FT1* and *VERNALIZATION1* (*VRN1*) (Figs 2, S4; Dataset S1). In wild-type, *FT1* expression was very low under 9 h and 10 h daylengths and moderately induced between the 10 h and 11 h photoperiods of late winter, with a threefold increase in transcripts (Fig. 2a,b,d). Under these daylengths, the diurnal expression pattern of *FT1* included peaks during the day and at dusk. This pattern continued under 12 h photoperiods with a prominent peak in the morning (ZT 3 h) and a minor peak after dusk (ZT 16 h). *FT1* transcripts increased significantly in spring between 12 and 13 h daylengths, with expression at 13 h photoperiods being *c.* 10-fold higher than that at 12 h photoperiods (Fig. 2a,d). Under 13 h daylengths, *FT1* transcripts peaked during the day at ZT 6 h and again in the evening (ZT 13–16 h). The *ppd-1* lines unexpectedly displayed a similar diurnal pattern of *FT1* expression and near-identical responsiveness to the changes in daylength, relative to wild-type; however, the amplitude of expression was significantly lower than wild-type at the majority of time-points of the 11–13 h photoperiods (Fig. 2a,d). Conversely, the insensitive *Ppd-D1a* plants showed dramatically different *FT1* transcript activity compared to the wild-type, with expression detected much earlier than for the wild-type during the 9–10 h photoperiods and showing much higher amplitudes (Fig. 2b,d). In the *Ppd-D1a* NILs, we detected low levels of *FT1* transcripts under 9 h and 10 h photoperiods of winter, before expression increased significantly as daylengths extended at the beginning of spring to 11 h, with levels comparable to that detected in wild-type under 13 h (Fig. 2b,d). *FT1* expression then settled at 12 h daylengths, before increasing again at 13 h – under photoperiods of 12 and 13 h, the diurnal expression pattern included peaks during the morning (ZT 3–6 h) and evening (ZT 16–20 h). These results show that *Ppd-1* is required for robust expression of *FT1*, and that photoperiod-insensitive alleles promote rapid induction of *FT1* under shorter daylengths of winter. In wild-type, the increase in *FT1* expression occurs in a step-wise process, with an initial rise at the beginning of spring (11 h) followed by a stronger induction in late spring (13 h; Fig. 2d). The relative differences in *FT1* expression associated strongly with flowering-time phenotypes, as the photoperiod-insensitive lines flowered 13 d earlier than wild-type, while the *ppd-1* NILs flowered 11 d later (Fig. 2e).

As we detected a difference in *Ppd-1* expression in field-grown plants relative to glasshouse conditions, we hypothesized that *FT1* would display an altered transcriptional profile under the two growth regimes. In contrast to the field-grown plants, wild-type glasshouse-grown plants expressed *FT1* under the 11 h photoperiod, which stabilized as daylengths increased to 12 and 13 h, indicating that 11 h daylengths are sufficient to induce flowering

in wheat (Figs 2c, S5; Dataset S1). The *ppd-1* lines displayed a similar trend with considerable expression at 11 h, which is maintained through to 13 h; however, at 12 and 13 h there were significantly fewer transcripts than wild-type throughout the day and night (Figs 2c, S5). In the photoperiod-insensitive lines, *FT1* was induced at the 10 h daylength, with a dramatic increase in expression as daylengths extend to 11 h, which is maintained through to 13 h (Fig. 2c). At all time-points and daylengths, the amplitude of *FT1* transcripts is higher in the *Ppd-D1a* lines than wild-type and *ppd-1* NILs (Fig. 2c). A similar diurnal expression pattern of *FT1* was detected in all three lines, with one peak detected in the morning (ZT 0-6 h) and another at dusk. The relative differences in *FT1* expression are reflected in the flowering-time phenotypes, as photoperiod-insensitive lines flowered 20 d earlier than the wild-type, while the *ppd-1* NILs flowered 11 d later (Fig. 2e).

To investigate floral promoting genes that respond to temperature, we examined the transcript levels of *VRN1*, expression of which increases following exposure to cold and influences the timing of flowering for winter wheat treated under higher temperatures (Danyluk *et al.*, 2003; Trevaskis *et al.*, 2003; Yan *et al.*, 2003; Dixon *et al.*, 2019) (Fig. S4; Dataset S1). *VRN1* was strongly expressed in leaves under all photoperiods, consistent with Paragon containing a spring allele for *VRN-A1*. While *VRN1* did not display a particular pattern of diurnal expression across all photoperiods, transcript levels were significantly affected by *Ppd-1* at certain time-points of the day and night. *VRN1* expression was significantly higher in photoperiod-insensitive NILs in the late evening/early morning (ZT 0, 3 and 16 h), relative to wild-type, and significantly lower in *ppd-1* NILs in the morning (ZT 3 h) (Fig. S4).

## Inflorescence development responds dynamically to changes in *FT1* activity

To investigate the connection between the floral-promoting pathway in the leaves with IM development, we first examined developmental progression of inflorescences from field-grown plants in the context of changes in *FT1* expression (Fig. 3). In wild-type, the SAM remained vegetative until the beginning of spring at the 11 h photoperiod, at which point it transitioned towards the double-ridge stage (Figs 3a–c, S6). The timing of this transition coincided with the three- to five-fold increase in *FT1* expression that occurs between the 10 and 11 h photoperiods (Fig. 2a,d). The IM remained at the double-ridge stage until daylengths extended to 12.5 h, when it transitioned to the glume and lemma primordium stages (Figs 3a, S6). The timing of this second transition coincided with the 10-fold increase in *FT1* expression that occurs in leaves between the 12 and 13 h photoperiods of April (Fig. 2a,d). The IMs remained at the lemma primordium stage until daylengths reached 13.5 h. (Figs 3a, S6). Growth and development of the IM then proceeded rapidly beyond this point, with spikelets at the lemma primordium stage forming floret primordia and reaching the terminal spikelet stage when daylengths were 14 h (Figs 3a, S6). Interestingly, transition of the wild-type IM from glasshouse-grown plants followed a very similar
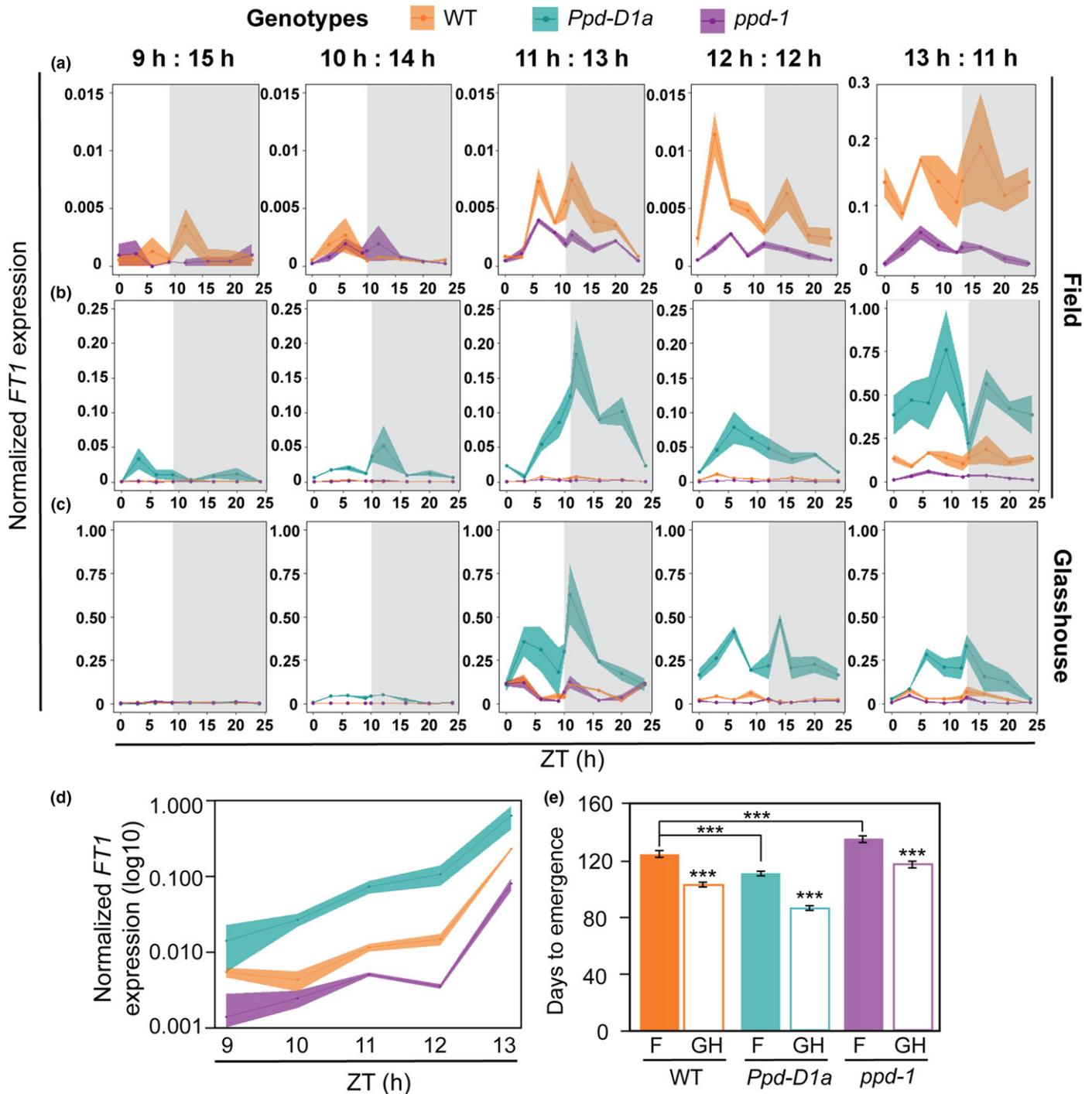
**Fig. 2** Seasonal regulation of *FT1* under field- and glasshouse-based conditions in wheat. Diurnal expression profile of *FT1* in wild-type (WT, orange), *Ppd-D1a* photoperiod-insensitive (cyan) and null *ppd-1* (magenta) near-isogenic lines (NILs) under field (a, b) and glasshouse (c) conditions. In (a), the data from the *Ppd-D1a* NIL have been removed to show differences in transcript levels between wild-type and *ppd-1* NILs. (d) Data summarizing the peak of *FT1* expression for each daylength, plotted on a logarithmic scale. (e) Field and glasshouse flowering time phenotypes for the three *Ppd-1* NILs, normalized as days to emergence from the 9 h daylength. The grey shading highlights night-time data. (a–d) Data are presented as ribbon plots which show the normalized mean transcript levels (solid line with data points) ± SE (shaded region) of three biological replicates. (e) Data are mean ± SE, with field replicates being five independent plots and glasshouse replicates being 10–15 plants. ***, $P < 0.001$.

progression, except that key events occurred at a photoperiod *c.* 1 h earlier, relative to the field, to coincide with the earlier induction of *FT1* expression (Fig. 3b). The stage-specific progression of IM development was consistent within each genotype,

with almost no variation observed among replicates in both the field and the glasshouse. The IMs transitioned to double-ridge and lemma primordium stages when daylengths were 10 and 11.75 h, respectively (Fig. 3b). These results show that the timing
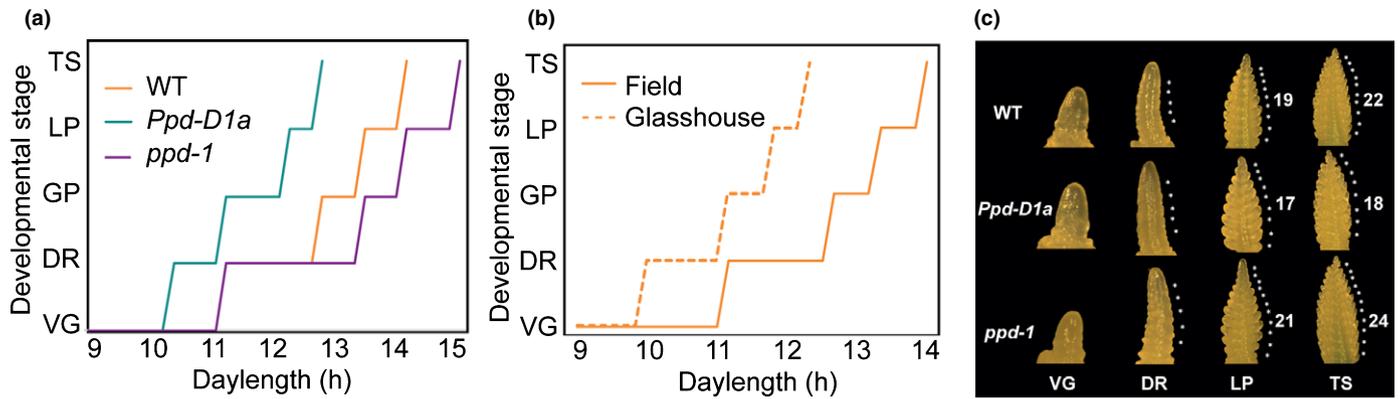
**Fig. 3** Seasonal progression of inflorescence meristem development in wheat *Ppd-1* near-isogenic lines (NILs). (a) Progression of inflorescence meristem development throughout the season in wild-type (WT, orange), *Ppd-D1a* photoperiod-insensitive (cyan) and null *ppd-1* (magenta) NILs under field conditions, at the vegetative (VG), double-ridge (DR), glume primordia (GP), lemma primordia (LP) and terminal spikelet (TS) stages. (b) Comparative progression of wild-type inflorescence meristems grown under field (solid) and glasshouse (dashed) conditions. (c) Representative images of inflorescence meristems from each genotype at the four developmental stages. In (a, b) data are the average of four or five replicates per developmental stage.

of the transition for wild-type IMs to the double-ridge stage in the field coincides with an initial rise in *FT1* expression. The IMs then stall at the double-ridge stage until the second increase in *FT1* expression, when the IMs proceed to the lemma primordium and terminal spikelet stages. In the *Ppd-D1a* insensitive NILs, IM development closely tracked changes in *FT1* expression. The IM transitioned to the double-ridge stage between the 10 and 11 h daylengths of late winter, coinciding with the photoperiods when *FT1* transcripts increased significantly (Figs 3a, S6). The IMs then proceeded rapidly to the glume and lemma primordium stages when daylengths were 11 and 12 h, respectively, without the developmental pause observed in wild-type, which coincided with the high expression of *FT1* under these photoperiods (Figs 2, 3a). The IMs arrived at the terminal spikelet stage when daylengths were 12.75 h. The IMs of *ppd-1* NILs transitioned to the double-ridge stage at the 11 h photoperiod, proceeding to the lemma primordium and terminal spikelet stages when daylengths were 14 and 15 h, respectively (Figs 3a, S6). Development of the *ppd-1* inflorescences therefore closely followed changes in *FT1* expression in leaves, which initiated between the 10 and 11 h photoperiods, before increasing again at 13 h. In both the *Ppd-D1a* photoperiod-insensitive and the *ppd-1* NILs, the influence of *Ppd-1* on spikelet number was detected between glume primordium and terminal spikelet stages, with fewer spikelets forming in the insensitive NIL and more developing in *ppd-1*, relative to the wild-type (Fig. 3c). This trend in spikelet number for each genotype was also observed at maturity, with photoperiod-insensitive lines producing shorter inflorescences with fewer spikelets than the wild-type, while *ppd-1* NILs formed longer inflorescences with more spikelets (Fig. S7).

Our previous analysis showed *FT1* is required for robust expression of meristem identity genes and timely progression to the terminal spikelet stage (Boden *et al.*, 2015; Dixon *et al.*, 2018a). Based on the *FT1* expression analyses and modified development of IMs in *Ppd-D1a* and *ppd-1* NILs shown here, we hypothesized that expression of meristem identity genes occurs earlier and to a higher level in insensitive NILs, relative to wild-

type, and is delayed and lower in the *ppd-1* NILs. To test this, we analysed expression of transcription factors that have key roles in early inflorescence development, including the regulation of spikelet architecture, floral organ identity and fertility that are affected by *Ppd-1* allelism (Boden *et al.*, 2015; Schilling *et al.*, 2020). The genes include MADS-box transcription factors *APETALA1-like* (*AP1-2*, *AP1-3* and *VRN1/AP1-1*), *AGAMOUS-LIKE6* (*AGL6*), *SEPALLATA1-6* (*SEP1-6*) and *SUPPRESSOR OF CONSTANS1* (*SOC1*), as well as the homeobox gene, *HOX2*, and *FLOWERING LOCUS T2* (*FT2*), which regulate floret fertility and spikelet number in wheat (Fig. 4; Dataset S1) (Boden *et al.*, 2015; Sakuma *et al.*, 2019; Shaw *et al.*, 2019; Schilling *et al.*, 2020). Transcripts of *AP1-2*, *AP1-3*, *SEP1-6* and *SOC1* were detected in *Ppd-D1a* NILs earlier than in the wild-type, as determined by time since germination (daylength) (Fig. 4a). These genes were expressed between the 10 and 11 h photoperiods in *Ppd-D1a* NILs, but not in the wild-type until the 12–13 h photoperiods, coinciding with daylengths of midspring when *FT1* expression increased significantly in leaves and the IMs transitioned to the double-ridge stage. Expression of *FT2*, *AGL6* and *HOX2* occurred later in the season with transcripts detected earlier in *Ppd-D1a* NILs relative to the wild-type. For the *ppd-1* NILs, expression of these genes was inverted relative to changes observed in the insensitive NILs, with transcripts detected at later photoperiods than in the wild-type, at daylengths corresponding to the delay in *FT1* induction and IM development observed in these plants. The amplitude of *AP1-2*, *AP1-3*, *VRN1*, *SEP1-6*, *SOC1*, *MADS6* and *FT2* transcripts unexpectedly reached the same maximum level in the wild-type as in *Ppd-D1a* NILs, demonstrating that meristem identity genes are not expressed at higher levels in photoperiod-insensitive plants (Fig. 4; Dataset S1). By contrast, transcripts were significantly lower in the *ppd-1* NILs at certain time-points, indicating that *Ppd-1* is required for robust expression of genes that promote spikelet and floret development (Fig. 4; Dataset S1). *HOX2* transcripts spiked to higher levels in *Ppd-D1a* NILs, relative to wild-type and *ppd-1*, at the green anther stage. Based on the seasonal shift in expression peak detected for these genes in the insensitive and *ppd-1*
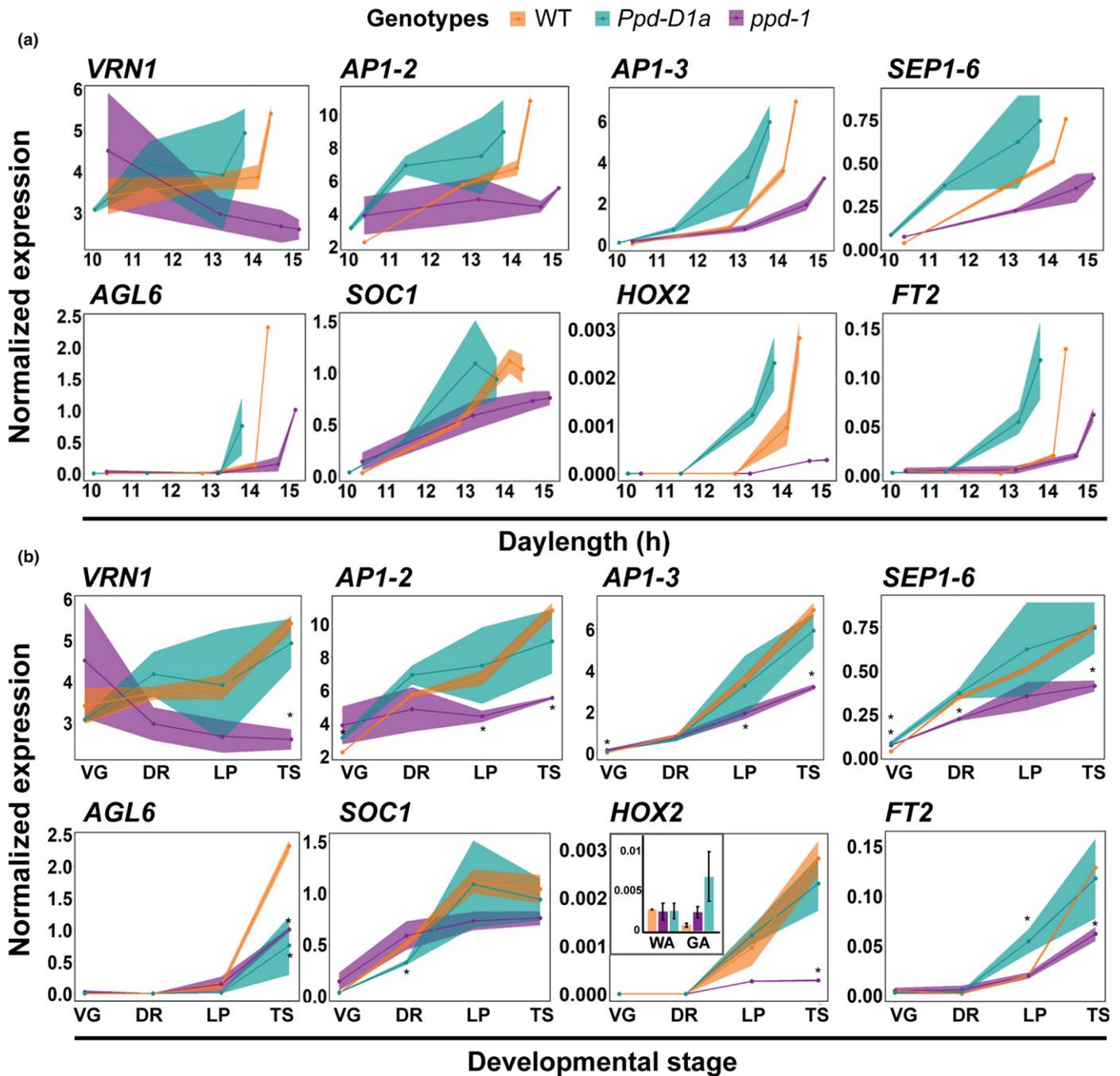
**Fig. 4** Seasonal and stage-specific expression analysis of meristem identity genes in wheat. Expression of key meristem identity genes plotted by daylength (a) and developmental stage (b) from plants grown under field conditions. Data include expression profiles for wild-type (WT, orange), *Ppd-D1a* photoperiod-insensitive (cyan) and null *ppd-1* (magenta) near-isogenic lines (NILs). The defined stages are vegetative (VG), double-ridge (DR), lemma primordium (LP), terminal spikelet (TS), white anther (WA) and green anther (GA). Data are presented as ribbon plots which show the normalized mean transcript levels (solid line with data points) ± SE (shaded region) of three-four biological replicates. Inset for *HOX2* (b) shows normalized mean transcript levels ± SE. *, $P < 0.05$.

NILs, relative to the wild-type, we hypothesized that their expression was changing in relation to developmental stage rather than daylength. To test this hypothesis, we normalized gene expression according to developmental stages including vegetative, double-ridge, lemma primordium, terminal spikelet, white anther and green anther (Fig. 4b). Following normalization, transcript levels were almost identical in *Ppd-D1a* NILs, relative to the wild-type, supporting the conclusion that photoperiod-insensitive alleles

advance expression of meristem identity genes without increasing transcript levels (Fig. 4b). An exception to this trend was *FT2*, which was higher in *Ppd-D1a* NILs at the lemma primordium stage, relative to the wild-type, in which transcript levels remained low until the terminal spikelet stage (Fig. 4b). In *ppd-1* NILs, transcripts of meristem identity genes, except for *VRN1*, were much lower and peaked at later developmental stages, relative to the wild-type − *VRN1* was highly expressed in all three

genotypes from the earliest stage (Fig. 4b). These field-based results show insensitive alleles provoke accelerated but equal expression of meristem identity genes in IMs, relative to the wild-type, and that *Ppd-1* is required for timely expression of genes that promote spikelet and floret development.

## *FT2* contributes to the termination of spikelet development

To further investigate the role of the initial induction of *FT1* expression in leaves and its relationship to IM development, we analysed the molecular and physiological effects of maintaining plants at 10 h daylengths of late winter. We hypothesized that the inceptive rise in *FT1* expression in the wild-type promotes the initial stages of inflorescence development, and the second stronger induction is required to proceed to later reproductive stages. To test this hypothesis, plants were grown in a glasshouse under natural photoperiods until the daylength was 10 h, before being shifted to a moving bench that maintained a fixed short-day 10 h photoperiod (Fig. S2). Wild-type plants maintained at 10 h progressed to the lemma primordium stage on the same day and produced the same amount of spikelets as plants maintained under natural photoperiods, by which time the daylength had surpassed 12 h (Fig. 5a). However, these plants stalled at the lemma primordium stage for an extra 30 d before arriving at the terminal spikelet stage and produced more spikelets compared to plants maintained under natural photoperiods ($29 \pm 0.5$ vs $24 \pm 0.3$ spikelets, respectively; Fig. 5a–c). The delay in inflorescence development coincided with the developmental stage at which the second stronger induction of *FT1* expression occurred (Fig. 2a,d). This delay was also observed in *ppd-1* NILs, which produced more spikelets and transitioned to the terminal spikelet stage significantly later than plants maintained under natural photoperiods ($28.3 \pm 0.3$ vs $25.4 \pm 0.2$ spikelets; Fig. 5b,c). The *Ppd-D1a* NILs transitioned to the terminal spikelet stage more rapidly than the wild-type, but there was a slight developmental delay relative to natural photoperiods, as plants maintained at 10 h produced more spikelets than glasshouse-grown plants ($25.8 \pm 0.4$ vs $21.5 \pm 0.2$; Fig. 5b,c). Wild-type and *ppd-1* NILs flowered later under the fixed 10 h photoperiods than plants grown under natural daylengths, but the photoperiod-insensitive NILs flowered at the same time (Fig. S2). In addition to increasing spikelet number, the 10 h photoperiods altered inflorescence architecture of wild-type and the *ppd-1* NIL by forming elongated basal internodes immediately before ear emergence, as occurs in *ppd-1* mutants grown under constant long-days (Figs 5d, S2) (Shaw *et al.*, 2013; Boden *et al.*, 2015). The insensitive *Ppd-D1a* lines did not produce these elongated internodes. These results indicate that the initial induction of *FT1* is sufficient to promote transition of the IM to the lemma primordium stage; however, the second higher induction of *FT1* is required to progress development of the IM to the terminal spikelet and later stages. This delay is overridden in *Ppd-D1a*-insensitive NILs, probably due to the higher expression of *FT1* (Fig. 2).

To investigate genes that contribute to the progression of inflorescence development beyond the lemma primordium stage, we investigated the role of *FT2*. We selected *FT2* because its expression increased strongly in wild-type inflorescences between the lemma primordium and terminal spikelet stages, relative to leaves and stems at the same developmental stages, and it was expressed earlier in *Ppd-D1a* NILs and later in *ppd-1* plants (Figs 4, S8). At
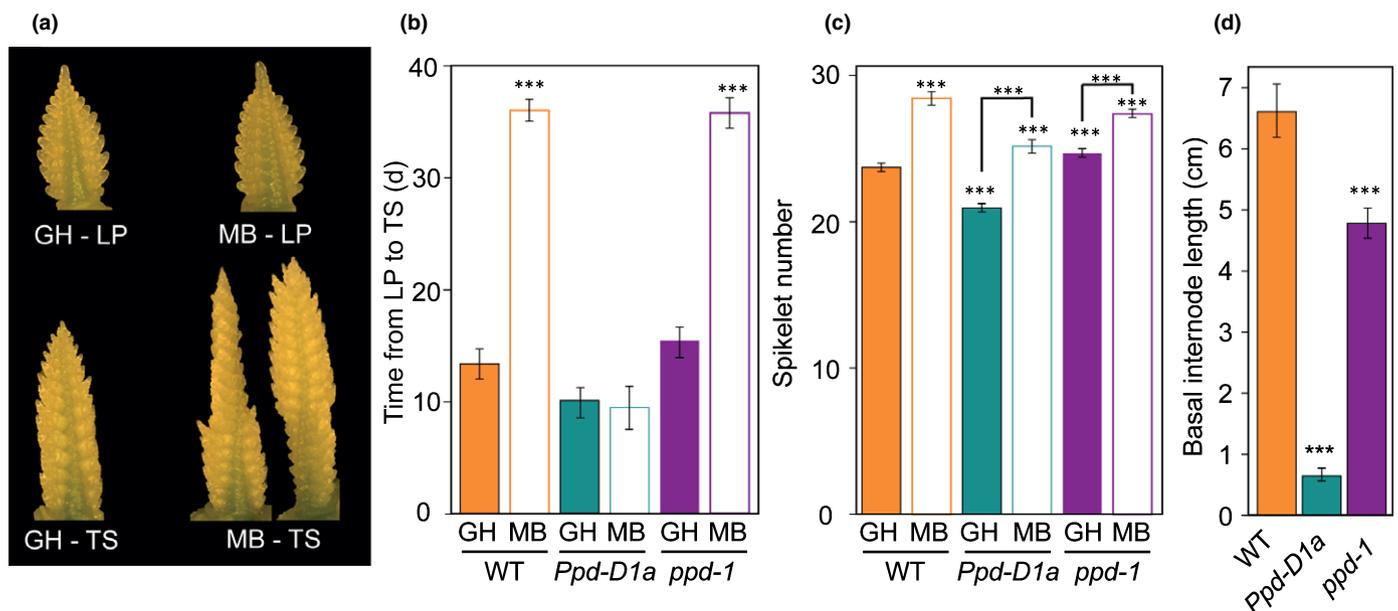


**Fig. 5** Inflorescence phenotypes of wheat plants maintained under short-day photoperiods. (a) Inflorescences of wild-type (WT) plants grown under natural glasshouse (GH) and 10 h moving bench (MB) conditions develop at the same rate until the lemma primordia (LP) stage, but develop more spikelets by the terminal spikelet stage (TS). (b) IM development of WT and *ppd-1* near-isogenic lines (NILs) plants grown under MB conditions is significantly delayed between the LP and TS stages. (c) Spikelet numbers for all three genotypes grown under natural GH or MB conditions. (d) WT and *ppd-1* NIL plants show elongated basal rachis internodes, but not *Ppd-D1a* NILs. Data are the average $\pm$ SEM of five biological replicates. ***, $P < 0.001$.

later developmental stages, *FT2* was also expressed strongly in flag leaves and the pre-booting inflorescence. Expression of *FT2* was much lower in IMs at the lemma primordium stage of plants shifted to the 10 h photoperiod, relative to plants maintained under natural photoperiods, suggesting development of the IM to the terminal spikelet stage is associated with robust expression of *FT2* (Fig. 6a). To test the role of *FT2* genetically, we analysed two independent lines containing missense mutations in *FT2* of the B genome (*FT-B2*; G45D and R150C), which are predicted to be deleterious for protein function (PROVEAN scores of −5.1 and −6.8) (Choi & Chan, 2015). The *ft-B2* mutants produced more spikelets than their wild-type NILs, indicating that progression of the IM to the terminal spikelet stage was delayed (Fig. 6b–d). The *ft-B2* mutants flowered later than wild-type plants, even under extreme long-day photoperiods (Fig. 6e).

## Discussion

Flowering is a crucial process in the life cycle of an annual plant. Flowering-time genes respond to seasonal cues including daylength and temperature to coordinate seed production with favourable environmental conditions. Our understanding of the genes that regulate flowering of wheat mostly stems from work performed under controlled conditions different from those experienced by field-grown plants (Worland *et al.*, 1998; González *et al.*, 2005; Beales *et al.*, 2007; Boden *et al.*, 2015; González-Navarro *et al.*, 2015; Prieto *et al.*, 2018; Perez-Gianmarco *et al.*, 2019). Here, we provide new evidence of the seasonal regulation of flowering using natural photoperiods of a standard growing season.

The induction of *FT-like* genes is core to the floral transition of angiosperms, which is conserved in wheat with *FT1* being a key activator of flowering (Yan *et al.*, 2006; Bratzel & Turck, 2015; Finnegan *et al.*, 2018; Dixon *et al.*, 2018a). Our analysis of field-grown plants unexpectedly showed that *FT1* induction occurs in a stepwise process, with an initial rise in transcripts detected under 11 h daylengths before increasing again as they extend to 13 h (Fig. 7). The induction of *FT1* under 11 h photoperiods demonstrates that long daylengths are not necessary to induce *FT1* expression in photoperiod-sensitive wheat. This is consistent with inflorescences transitioning to early reproductive stages at the beginning of spring, when daylengths are shorter than 12 h – a process promoted by *FT1* (González-Navarro *et al.*, 2015; Dixon *et al.*, 2018a). The rise in *FT1* expression as daylengths extend to 13 h is consistent with the significant increase in *FT1* transcripts when plants are shifted to extreme long-days (22 h : 2 h), and with analysis of Arabidopsis grown under natural summer photoperiods, where *FT1* transcripts were higher under 16 h than under 14 h photoperiods (Song *et al.*, 2018; Dixon *et al.*, 2018b). A second unexpected outcome of the *FT1* expression analysis regards the diurnal expression pattern under natural photoperiods, which showed peaks in transcript levels in the morning and at dusk. These patterns are different from those reported for wheat and barley in controlled conditions under constant long-days, where *FT1* peaks at dusk (Turner *et al.*, 2005; Campoli *et al.*, 2012; Chen *et al.*, 2014; Boden *et al.*, 2015). Our data are consistent with comparisons of Arabidopsis grown under laboratory vs natural photoperiod conditions, which showed temperature and light quality signals are responsible for a predominant peak of *FT* expression during the morning that was not prevalent in the laboratory (Song *et al.*, 2018). These data show that the regulation of flowering-time processes in the field is potentially more complex than indicated by work performed using controlled conditions.
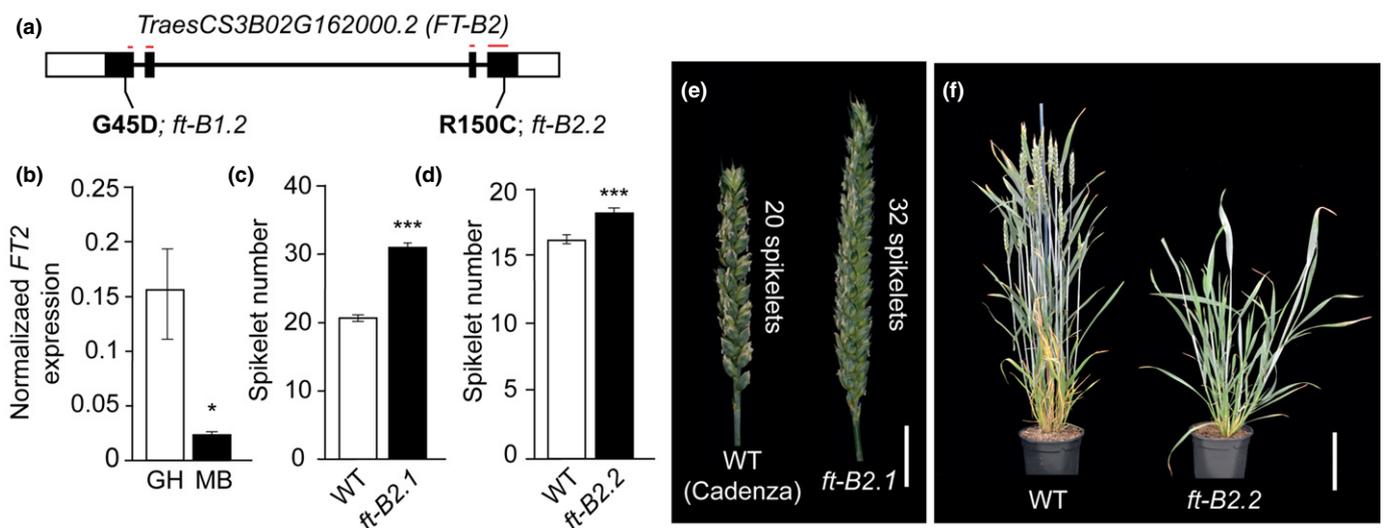


**Fig. 6** *FT2* influences inflorescence development and spikelet number in wheat. (a) A schematic of *FT-B2* showing sites of the two missense mutations, exons (black boxes), introns (black lines), untranslated regions (white boxes) and the PEBP domain (red line). (b) *FT2* expression is perturbed in developing IMs of wild-type (WT) plants shifted to 10 h photoperiods using a moving bench (MB), relative to plants maintained under natural photoperiods. (c–e) Spikelet number increases on inflorescences of (c) *ft-B2.1* mutants and (d) *ft-B2.2* mutants, relative to wild-type, (e) including images of representative inflorescences. (f) *ft-B2.2* mutant lines flower later than WT under long daylengths. (b) Data are the mean ± SE of three biological replicates, and five replicates for (c, d). Bars: (e) 1 cm; (f) 10 cm. *, $P < 0.05$; ***, $P < 0.001$.
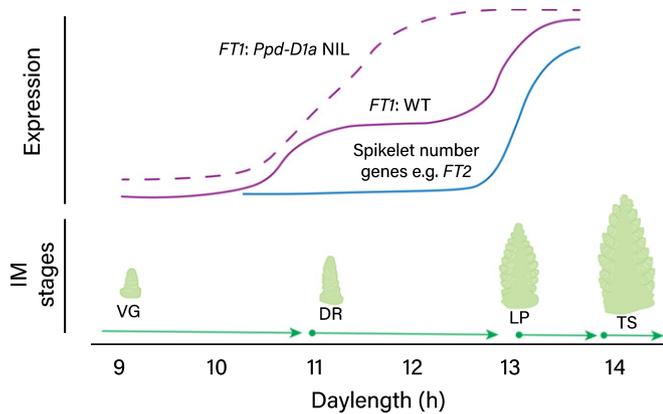
**Fig. 7** Model for seasonal regulation of flowering-time pathways and inflorescence development in wheat. A model outlining the seasonal regulation of *FT1* expression in wheat, defined by hourly increases in daylength, in wild-type (WT, solid, purple line) and *Ppd-D1a* photoperiod-insensitive lines (dashed purple line). The second induction of *FT1* promotes expression of meristem identity genes that determine spikelet number, such as *FT2* (blue line), which helps transition the IM to the lemma primordium and terminal spikelet stages. The regulation of floral promoting pathways coincides with developmental stages (cartoon images) according to the seasonal progression of inflorescence development, shown here for wild-type plants. The stages include vegetative (VG), double-ridge (DR), lemma primordium (LP) and terminal spikelet (TS).

Genetic variation for *Ppd-1* has a major effect on *FT1* activity and flowering time, with photoperiod-insensitive alleles promoting higher *FT1* expression and earlier flowering, relative to wild-type plants (Beales *et al.*, 2007; Wilhelm *et al.*, 2009; Boden *et al.*, 2015). Here, we have shown that photoperiod-insensitive alleles cause *FT1* to be expressed earlier in the season and to higher levels, relative to the wild-type, with the initial levels comparable to those detected in the wild-type later in the year (Fig. 7). The photoperiod-insensitive allele significantly affects *FT1* expression during early spring, which corresponds to photoperiods when inflorescence development initiates in the field and spikelet number is determined (González-Navarro *et al.*, 2015). The pronounced effect of the *Ppd-D1a* allele on *FT1* expression during the shorter daylengths of winter is consistent with the exaggerated acceleration of flowering that occurs in photoperiod-insensitive lines under constant 9 h photoperiods, where flowering occurs 60–190 d earlier than photoperiod-sensitive lines; photoperiod-insensitive alleles only accelerate flowering by 6-12 d under field conditions (Fig. 2e) (Worland *et al.*, 1998; Beales *et al.*, 2007; Diaz *et al.*, 2012; Cane *et al.*, 2013). The direct mechanism by which photoperiod-insensitive alleles promote higher *FT1* expression remains unknown. As *Ppd-1* encodes a transcription factor that interacts with other CCT (CONSTANS, CONSTANS-like, TIMING OF CAB1) domain containing-proteins, the high evening expression may facilitate ectopic protein interactions that promote *FT1* transcription (Li *et al.*, 2011; Shaw *et al.*, 2020). Alternatively, evening expression of *Ppd-1* may activate an as yet unknown gene that subsequently promotes *FT1* transcription. Interestingly, the mis-regulated expression of *Ppd-D1* in

photoperiod-insensitive lines was less dramatic in field-grown plants, relative to the glasshouse and laboratory conditions, with the high evening expression not detected under 9 h photoperiods and limited to late hours of the night as photoperiods extended (Beales *et al.*, 2007; Boden *et al.*, 2015). Given that both glasshouse- and field-grown plants were grown under natural photoperiods, the altered pattern of *Ppd-1* expression may be attributed to differences in temperature and/or light quality, which would be consistent with *Ppd-1* being regulated by phytochromes that are sensors for changes in light quality and temperature in plants (Rockwell *et al.*, 2006; Chen *et al.*, 2014; Jung *et al.*, 2016). Indeed, the glasshouse was warmer than the field (*c.* 1°C, on average), particularly in the hours surrounding sunrise (Fig. S9). Regarding the *ppd-1* lines, *FT1* transcript levels were significantly lower than in the wild-type, but *FT1* expression still responded to increasing daylengths and displayed the same diurnal pattern as in the wild-type. These results indicate that *Ppd-1* is not the only factor regulating seasonal activity of *FT1*.

In addition to a role regulating *FT1*, our data indicate that *Ppd-1* regulates *VRN1* and its own expression. *VRN1* was influenced by *Ppd-1* allelism at certain daily time-points, with transcripts tending to be higher in photoperiod-insensitive lines and lower in *ppd-1*. However, *VRN1* was expressed robustly in the absence of *ppd-1*, especially under short to neutral daylengths, indicating that seasonal regulation of *VRN1* in spring wheat involves more factors than *Ppd-1*. Our data are consistent with recent analysis in tetraploid wheat, which showed *VRN1* is expressed strongly under short-day photoperiods in *ppd-1* lines at levels comparable or higher than in the wild-type (Shaw *et al.*, 2020). Regarding *Ppd-1*, the higher levels of *Ppd-D1* transcripts in *ppd-1* lines indicates that this pseudo-response regulator may influence its own expression during the day through a self-regulatory feedback loop. This potential function of *Ppd-1* is supported by the interaction detected between *Ppd-1* homoeologues in the photoperiod-insensitive *Ppd-D1a* NILs, and is consistent with *Ppd-B1* expression being higher in *Ppd-D1a* mutants containing splice site mutations (Boden *et al.*, 2015).

The process of flowering involves communication of FT1 protein from the leaves to the SAM, from which reproductive development is promoted through expression of meristem identity genes. In Arabidopsis and rice, overexpression of *FT*/*Hd3a* (*Heading date 3a*) leads to hyper-activation of meristem identity genes (Kardailsky *et al.*, 1999; Yoo *et al.*, 2005; Taoka *et al.*, 2011; Kaneko-Suzuki *et al.*, 2018). Based on *FT1* transcripts being higher in photoperiod-insensitive lines and spikelet identity genes being expressed at lower levels in *ft-B1* mutants, we hypothesized that meristem identity genes would be more highly expressed in *Ppd-D1a* NILs, relative to the wild-type (Beales *et al.*, 2007; Shaw *et al.*, 2013; Boden *et al.*, 2015). Surprisingly, while photoperiod-insensitive lines accelerated meristem identity gene expression to occur earlier in the season, the overall amplitude of transcripts was identical to that in the wild-type. In *ppd-1*, induction of meristem identity genes was delayed, and transcript levels were lower than in wild-type plants, consistent with the reduced activity of *FT1* in these lines and the decreased

expression of the same genes in *ft-B1* mutants (Boden *et al.*, 2015). Curiously, *VRN1* did not follow this trend, instead being expressed highly in all three genotypes before induction of the floral transition, before falling to lower levels in *ppd-1* at the terminal spikelet stage. These results indicate that the *VRN-A1a* allele of spring wheat promotes high expression of *VRN1* in the meristem before the induction of flowering, but the stepwise increases in *FT1* transcripts are required for sustained *VRN1* activity. The stepwise increase in *FT1* expression in leaves aligned strongly with the seasonal upregulation of other meristem identity genes and progression of IM development, with the first induction of *FT1* facilitating the vegetative to double-ridge transition and the second rise promoting advancement to later stages (Fig. 7). The accelerated peak in meristem identity gene expression in photoperiod-insensitive lines coincided with arrival of IMs at the terminal spikelet stage earlier in the season. These data indicate that the regulation of inflorescence development and spikelet number by *Ppd-1* is not determined by the absolute level of meristem identity gene expression, but by the day during the season when the peak occurs. The advanced induction of *FT1* in the photoperiod-insensitive line potentially explains its ability to advance to terminal spikelet when daylengths were maintained at 10 h, relative to the wild-type and *ppd-1* that stalled at the lemma primordium stage. In addition to regulating spikelet number, photoperiod-insensitive alleles also reduce floret fertility; the increased expression of *HOX2* at later stages in *Ppd-D1a* NILs, relative to wild-type plants, may explain the decrease in fertile florets (Prieto *et al.*, 2018; Sakuma *et al.*, 2019). Taken together with the seasonal analysis of *FT1* expression, these results indicate that inflorescence development is intimately connected with the activity of floral signals generated in leaves, which dynamically respond to increasing daylengths.

Our analysis identified *FT2* as a key regulator of spikelet development in hexaploid wheat. Based on the *ft-B2* mutants producing more spikelets, and the significant increase in *FT2* expression between the lemma primordium and terminal spikelet stages, we propose that *FT2* helps determine spikelet number by promoting transition of the IM to the terminal spikelet stage. The delayed ear emergence of *ft-B2* mutants suggests *FT2* advances flowering time, potentially through its role in developing IMs or in the flag leaf and emerging inflorescence, where it is expressed highly. This conclusion is consistent with analysis performed in tetraploid wheat, which showed *FT2* is expressed strongly in the developing IM and that loss-of-function *ft2* alleles delay flowering and increase spikelet number (Shaw *et al.*, 2019). A role for *FT2* in spikelet termination is supported by transcript analysis in *ppd-1* and photoperiod-insensitive *Ppd-D1a* NILs, in which *FT2* expression was inversely proportional to spikelet number and the rate of inflorescence development. The early rise of *FT2* transcripts in photoperiod-insensitive *Ppd-D1a* NILs may explain why IMs from this line could progress to the terminal spikelet stage when plants were maintained at 10 h photoperiods, relative to the wild-type, in which *FT2* expression was significantly lower than in plants grown under natural photoperiods. Interestingly, the rise in *FT2* transcripts between the lemma primordium

and terminal spikelet stages and its genetic association with spikelet number contrasts with the profile of genes such as *AP1*, *SEP1*, *VRN1* and *WFZP*, for which expression increases earlier in IM development and is associated with supernumerary spikelet formation (Boden *et al.*, 2015; Dobrovolskaya *et al.*, 2015). These results indicate that genes induced between the lemma primordium and terminal spikelet stages contribute to determination of spikelet number, while those expressed at earlier stages influence inflorescence architecture. Together with studies that have investigated the inflorescence transcriptome of wheat, these results allude to a strategy for identifying genes that regulate spikelet number (Wang *et al.*, 2017; Li *et al.*, 2018). For example, expression of *TERMINAL FLOWER 1* (*TFL1*) and *WHEAT ORTHOLOGUE OF APO1* correlates with spikelet number, with increased transcripts during the glume and floret primordium stages being associated with extra spikelets (Wang *et al.*, 2017; Kuzay *et al.*, 2019).

In summary, our work provides new insights into the regulation of molecular processes controlling flowering and inflorescence development in the field. We show that the early stages of inflorescence development are coordinated by stepwise inductions of *FT1* expression, which are overridden by photoperiod-insensitive *Ppd-D1a* alleles. The results highlight the importance of complementing laboratory-based analysis with experiments performed in the field – a concept echoed by studies performed in oilseed rape, Arabidopsis and rice that have used field-grown plants to uncover new information about the molecular events controlling seasonal regulation of flowering (Duncan *et al.*, 2015; Gomez-Ariza *et al.*, 2015; Hepworth *et al.*, 2018; Song *et al.*, 2018; O'Neill *et al.*, 2019). Our work provides an important foundation for understanding the mechanisms controlling yield-related traits of wheat, which is vital given our increasing need to improve global food security by generating superior yielding cultivars (Fischer *et al.*, 2014).

## Author contributions

AG and SAB contributed to conceptualization, formal analysis, methodology, investigation, validation, visualization and writing (original draft preparation and review/editing). SAB also contributed to supervision, project administration and funding acquisition.

## ORCID

Scott A. Boden (iD) https://orcid.org/0000-0001-5297-4067

# References

Beales J, Turner A, Griffiths S, Snape JW, Laurie DA. 2007. A pseudo-response regulator is misexpressed in the photoperiod insensitive *Ppd-D1a* mutant of wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics* **115**: 721–733.

Boden SA, Cavanagh C, Cullis BR, Ramm K, Greenwood J, Finnegan EJ, Trevaskis B, Swain SM. 2015. *Ppd-1* is a key regulator of inflorescence architecture and paired spikelet development in wheat. *Nature Plants* **1**: 14016.

Boden SA, Weiss D, Ross JJ, Davies NW, Trevaskis B, Chandler PM, Swain SM. 2014. EARLY FLOWERING3 regulates flowering in spring barley by mediating gibberellin production and *FLOWERING LOCUS T* expression. *The Plant Cell* **26**: 1557–1569.

Borrill P, Ramirez-Gonzalez R, Uauy C. 2016. expVIP: a customizable RNA-seq data analysis and visualization platform. *Plant Physiology* **170**: 2172–2186.

Bratzel F, Turck F. 2015. Molecular memories in the regulation of seasonal flowering: from competence to cessation. *Genome Biology* **16**: 192.

Campoli C, Shtaya M, Davis SJ, von Korff M. 2012. Expression conservation within the circadian clock of a monocot: natural variation at barley *Ppd-H1* affects circadian expression of flowering time genes, but not clock orthologs. *BMC Plant Biology* **12**: 97.

Cane K, Eagles HA, Laurie DA, Trevaskis B, Vallance N, Eastwood RF, Gororo NN, Kuchel H, Martin PJ. 2013. *Ppd-B1* and *Ppd-D1* and their effects in southern Australian wheat. *Crop and Pasture Science* **64**: 100–114.

Chen A, Li C, Hu W, Lau MY, Lin H, Rockwell NC, Martin SS, Jernstedt JA, Lagarias JC, Dubcovsky J. 2014. Phytochrome C plays a major role in the acceleration of wheat flowering under long-day photoperiod. *Proceedings of the National Academy of Sciences, USA* **111**: 10037–10044.

Choi Y, Chan AP. 2015. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* **31**: 2745–2747.

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.

Danyluk J, Kane NA, Breton G, Limin AE, Fowler DB, Sarhan F. 2003. *Ta*VRT-1, a putative transcription factor associated with vegetative to reproductive transition in cereals. *Plant Physiology* **132**: 1849–1860.

Diaz A, Zikhali M, Turner AS, Isaac P, Laurie DA. 2012. Copy number variation affecting the *Photoperiod-B1* and *Vernalization-A1* genes is associated with altered flowering time in wheat (*Triticum aestivum*). *PLoS ONE* **7**: e33234.

Dixon LE, Farre A, Finnegan EJ, Orford S, Griffiths S, Boden SA. 2018a. Developmental responses of bread wheat to changes in ambient temperature following deletion of a locus that includes *FLOWERING LOCUS T1*. *Plant, Cell & Environment* **41**: 1715–1725.

Dixon LE, Greenwood JR, Bencivenga S, Zhang P, Cockram J, Mellers G, Ramm K, Cavanagh C, Swain SM, Boden SA. 2018b. TEOSINTE BRANCHED1 regulates inflorescence architecture and development in bread wheat (*Triticum aestivum*). *The Plant Cell* **30**: 563–581.

Dixon LE, Karsai I, Kiss T, Adamski NM, Liu Z, Ding Y, Allard V, Boden SA, Griffiths S. 2019. *VERNALIZATION1* controls developmental responses of winter wheat under high ambient temperatures. *Development* **146**: dev172684.

Dobrovolskaya O, Pont C, Sibout R, Martinek P, Badaeva E, Murat F, Chosson A, Watanabe N, Prat E, Gautier N et al. 2015. FRIZZY PANICLE drives supernumerary spikelets in bread wheat. *Plant Physiology* **167**: 189–199.

Duncan S, Holm S, Questa J, Irwin J, Grant A, Dean C. 2015. Seasonal shift in timing of vernalization as an adaptation to extreme winter. *eLife* **4**: e06620.

Finnegan EJ, Ford B, Wallace X, Pettolino F, Griffin PT, Schmitz RJ, Zhang P, Barrero JM, Hayden MJ, Boden SA et al. 2018. Zebularine treatment is associated with deletion of *FT-B1* leading to an increase in spikelet number in bread wheat. *Plant, Cell & Environment* **41**: 1346–1360.

Fischer T, Byerlee D, Edmeades G. 2014. *Crop yields and global food security*. Canberra, ACT, Australia: Australian Centre of International Agricultural Research.

Fjellheim S, Boden S, Trevaskis B. 2014. The role of seasonal flowering responses in adaptation of grasses to temperate climates. *Frontiers in Plant Science* **5**: 431.

Gomez-Ariza J, Galbiati F, Goretti D, Brambilla V, Shrestha R, Pappolla A, Courtois B, Fornara F. 2015. Loss of floral repressor function adapts rice to higher latitudes in Europe. *Journal of Experimental Botany* **66**: 2027–2039.

González FG, Slafer GA, Miralles DJ. 2005. Photoperiod during stem elongation in wheat: is its impact on fertile floret and grain number determination similar to that of radiation? *Functional Plant Biology* **32**: 181–188.

González-Navarro OE, Griffiths S, Molero G, Reynolds MP, Slafer GA. 2015. Dynamics of floret development determining differences in spike fertility in an elite population of wheat. *Field Crops Research* **172**: 21–31.

Hepworth J, Antoniou-Kourounioti RL, Bloomer RH, Selga C, Berggren K, Cox D, Collier Harris BR, Irwin JA, Holm S, Sall T et al. 2018. Absence of warmth permits epigenetic memory of winter in Arabidopsis. *Nature Communications* **9**: 639.

Jung JH, Domijan M, Klose C, Biswas S, Ezer D, Gao M, Khattak AK, Box MS, Charoensawan V, Cortijo S et al. 2016. Phytochromes function as thermosensors in Arabidopsis. *Science* **354**: 886–889.

Kaneko-Suzuki M, Kurihara-Ishikawa R, Okushita-Terakawa C, Kojima C, Nagano-Fujiwara M, Ohki I, Tsuji H, Shimamoto K, Taoka KI. 2018. TFL1-like proteins in rice antagonize rice FT-like protein in inflorescence development by competition for complex formation with 14-3-3 and FD. *Plant and Cell Physiology* **59**: 458–468.

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.

Kitagawa S, Shimada S, Murai K. 2012. Effect of Ppd-1 on the expression of flowering-time genes in vegetative and reproductive growth stages of wheat. *Genes & Genetic Systems* **87**: 161–168.

Krasileva KV, Vasquez-Gross HA, Howell T, Bailey P, Paraiso F, Clissold L, Simmonds J, Ramirez-Gonzalez RH, Wang X, Borrill P et al. 2017. Uncovering hidden variation in polyploid wheat. *Proceedings of the National Academy of Sciences, USA* **114**: E913–E921.

Kuzay S, Xu Y, Zhang J, Katz A, Pearce S, Su Z, Fraser M, Anderson JA, Brown-Guedira G, DeWitt N et al. 2019. Identification of a candidate gene for a QTL for spikelet number per spike on wheat chromosome arm 7AL by high-resolution genetic mapping. *Theoretical and Applied Genetics.* **132**: 2689–2705.

Laurie DA, Pratchett N, Snape JW, Bezant JH. 1995. RFLP mapping of five major genes and eight quantitative trait loci controlling flowering time in a winter x spring barley (*Hordeum vulgare* L.) cross. *Genome* **38**: 575–585.

Li C, Distelfeld A, Comis A, Dubcovsky J. 2011. Wheat flowering repressor VRN2 and promoter CO2 compete for interactions with NUCLEAR FACTOR-Y complexes. *The Plant Journal* **67**: 763–773.

Li C, Dubcovsky J. 2008. Wheat FT protein regulates *VRN1* transcription through interactions with FDL2. *The Plant Journal* **55**: 543–554.

Li Y, Fu X, Zhao M, Zhang W, Li B, An D, Li J, Zhang A, Liu R, Liu X. 2018. A genome-wide view of transcriptome dynamics during early spike development in read wheat. *Scientific Reports* **8**: 15338.

O'Neill CM, Lu X, Calderwood A, Tudor EH, Robinson P, Wells R, Morris R, Penfield S. 2019. Vernalization and floral transition in autumn drive winter annual life history in oilseed rape. *Current Biology* **29**: 4300–4306.

Perez-Gianmarco TI, Slafer GA, Gonzalez FG. 2019. Photoperiod-sensitivity genes shape floret development in wheat. *Journal of Experimental Botany* **70**: 1339–1348.

Prieto P, Ochagavia H, Savin R, Griffiths S, Slafer GA. 2018. Dynamics of floret initiation/death determining spike fertility in wheat as affected by *Ppd* genes under field conditions. *Journal of Experimental Botany* **69**: 2633–2645.

Ramirez-Gonzalez RH, Uauy C, Caccamo M. 2015. PolyMarker: a fast polyploid primer design pipeline. *Bioinformatics* **31**: 2038–2039.

Rockwell NC, Su YS, Lagarias JC. 2006. Phytochrome structure and signaling mechanisms. *Annual Review of Plant Biology* **57**: 837–858.

Sakuma S, Golan G, Guo Z, Ogawa T, Tagiri A, Sugimoto K, Bernhardt N, Brassac J, Mascher M, Hensel G et al. 2019. Unleashing floret fertility in

wheat through the mutation of a homeobox gene. *Proceedings of the National Academy of Sciences, USA* **116**: 5182–5187.

Schilling S, Kennedy A, Pan S, Jermiin LS, Melzer R. 2020. Genome-wide analysis of MIKC-type MADS-box genes in wheat: pervasive duplications, functional conservation and putative neofunctionalization. *New Phytologist* **225**: 511–529.

Seki M, Chono M, Nishimura T, Sato M, Yoshimura Y, Matsunaka H, Fujita M, Oda S, Kubo K, Kiribuchi-Otobe C *et al.* 2013. Distribution of photoperiod-insensitive allele *Ppd-A1a* and its effect on heading time in Japanese wheat cultivars. *Breeding Science* **63**: 309–316.

Shaw LM, Li C, Woods DP, Alvarez MA, Lin H, Lau MY, Chen A, Dubcovsky J. 2020. Epistatic interactions between PHOTOPERIOD1, CONSTANS1 and CONSTANS2 modulate the photoperiodic response in wheat. *PLoS Genetics* **16**: e1008812.

Shaw LM, Lyu B, Turner R, Li C, Chen F, Han X, Fu D, Dubcovsky J. 2019. FLOWERING LOCUS T2 regulates spike development and fertility in temperate cereals. *Journal of Experimental Botany* **70**: 193–204.

Shaw LM, Turner AS, Herry L, Griffiths S, Laurie DA. 2013. Mutant alleles of Photoperiod-1 in wheat (*Triticum aestivum* L.) that confer a late flowering phenotype in long days. *PLoS ONE* **8**: e79459.

Shaw LM, Turner AS, Laurie DA. 2012. The impact of photoperiod insensitive *Ppd-1a* mutations on the photoperiod pathway across the three genomes of hexaploid wheat (*Triticum aestivum*). *The Plant Journal* **71**: 71–84.

Song YH, Kubota A, Kwon MS, Covington MF, Lee N, Taagen ER, Laboy Cintron 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.

Tamaki S, Matsuo S, Wong HL, Yokoi S, Shimamoto K. 2007. Hd3a protein is a mobile flowering signal in rice. *Science* **316**: 1033–1036.

Taoka K, Ohki I, Tsuji H, Furuita K, Hayashi K, Yanase T, Yamaguchi M, Nakashima C, Purwestri YA, Tamaki S *et al.* 2011. 14-3-3 proteins act as intracellular receptors for rice Hd3a florigen. *Nature* **476**: 332–335.

Trevaskis B, Bagnall DJ, Ellis MH, Peacock WJ, Dennis ES. 2003. MADS box genes control vernalization-induced flowering in cereals. *Proceedings of the National Academy of Sciences, USA* **100**: 13099–13104.

Turner A, Beales J, Faure S, Dunford RP, Laurie DA. 2005. The pseudo-response regulator *Ppd-H1* provides adaptation to photoperiod in barley. *Science* **310**: 1031–1034.

Wang Y, Yu H, Tian C, Sajjad M, Gao C, Tong Y, Wang X, Jiao Y. 2017. Transcriptome association identifies regulators of wheat spike architecture. *Plant Physiology* **175**: 746–757.

Wickham H. 2016. *ggplot2: elegant graphics for data analysis*. New York, NY, USA: Springer-Verlag. [WWW document] URL https://ggplot2.tidyverse.org [accessed 1 November 2019].

Wilhelm EP, Turner AS, Laurie DA. 2009. Photoperiod insensitive *Ppd-A1a* mutations in tetraploid wheat (*Triticum durum* Desf.). *TAG. Theoretical and Applied Genetics.* **118**: 285–294.

Worland A, Borner A, Korzun V, Petrovic S, Sayers E. 1998. The influence of photoperiod genes on the adaptability of European winter wheats. *Euphytica* **100**: 385–394.

Yan L, Fu D, Li C, Blechl A, Tranquilli G, Bonafede M, Sanchez A, Valarik M, Yasuda S, Dubcovsky J. 2006. The wheat and barley vernalization gene VRN3 is an orthologue of *FT*. *Proceedings of the National Academy of Sciences, USA* **103**: 19581–19586.

Yan L, Loukoianov A, Tranquilli G, Helguera M, Fahima T, Dubcovsky J. 2003. Positional cloning of the wheat vernalization gene VRN1. *Proceedings of the National Academy of Sciences, USA* **100**: 6263–6268.

Yoo SK, Chung KS, Kim J, Lee JH, Hong SM, Yoo SJ, Yoo SY, Lee JS, Ahn JH. 2005. CONSTANS activates *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* through FLOWERING LOCUS T to promote flowering in Arabidopsis. *Plant Physiology* **139**: 770–778.

## Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

**Dataset S1** Normalized expression data and statistical analyses of transcript data.

**Fig. S1** Seasonal events of photoperiod experiment in Norwich, UK.

**Fig. S2** Experimental design and phenotypes of plants from the photoperiod shift experiment.

**Fig. S3** Seasonal regulation of *Ppd-D1* in the glasshouse.

**Fig. S4** Seasonal regulation of *VRN1* in the field.

**Fig. S5** Seasonal regulation of *FT1* in the glasshouse.

**Fig. S6** Seasonal progression of inflorescence meristem development.

**Fig. S7** Inflorescence architecture phenotype of *Ppd-1* lines.

**Fig. S8** Tissue-specific analysis of *FT2* expression at early and late developmental stages.

**Fig. S9** Temperature profile of field- and glasshouse-based experiments.

**Methods S1** Supplementary methods, outlining phenotype analysis, tissue-specific expression analysis and statistical analyses.

**Table S1** Sequences of oligonucleotides used in qRT-PCR assays.

**Table S2** Sequences of oligonucleotides used for genotype analysis of *FT-B2* mutants.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

# Appendix 2

Statistics for Chapter 2 gene expression comparisons. See attached file: Appendix 2

Data previously published in:

Gauley, A., & Boden, S. A. 2020. Step-wise increases in *FT1* expression regulate seasonal progression of flowering in wheat (*Triticum aestivum*). *New Phytologist*. https://doi.org/10.1111/NPH.16910