Understanding genetic and environmental factors in wheat vernalization

Kathryn Teresa O'Connor

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds

School of Biology

Centre for Plant Sciences

October 2023

I confirm that the work submitted is my own, except where work which has formed part of jointly authored publications has been included. My contribution and the other authors to this work has been explicitly indicated below. I confirm that appropriate credit has been given within the thesis where reference has been made to the work of others.

Work from jointly authored publications forms part of **Chapter Two** of this thesis. The citation is as follows:
Authors: Jin-Kyung Cha, Kathryn O'Connor, Samir Alahmad, Jong-Hee Lee, Eric Dinglasan, Hyeonjin Park, So-Myeong Lee, Dominique Hirsz, Soon-Wook Kwon, Youngho Kwon, Kyeong-Min Kim, Jong-Min Ko, Lee T. Hickey, Dongjin Shin, Laura E. Dixon,

Title: Speed vernalization to accelerate generation advance in winter cereal crops,
Journal: Molecular Plant, Volume 15, Issue 8, 2022, Pages 1300-1309.
https://doi.org/10.1016/j.molp.2022.06.012.

Kathryn O'Connor investigated flowering time of wheat (alongside J.-K.C, D.H., and K.-M.K.), analyzed the data (alongside J.-K.C. and S.-W.K.), generated figures, and discussed the manuscript along with all other authors. The other authors contributed as follows: L.T.H., D.S., and L.E.D. conceived and supervised the project. J.-H.L., L.T.H., D.S., and L.E.D. designed the experiments. S.A. and E.D. investigated flowering time of barley. J.-K.C., H.P., S.-M.L., Y.K., and J.-M.K. developed bread wheat wheat-breeding materials. L.T.H., D.S., and L.E.D. wrote the manuscript.

Abstract

Prolonged exposure to low temperatures facilitates the flowering response in plants. This process is known as vernalization and regulates the transition from vegetative to reproductive development during winter. In cereals, vernalization centres around the floral activator *VERNALIZATION 1* (*VRN1*) and its interaction with *VRN2*, *VRN3*, *VRT2* and the *ODDSOC* genes, *ODDSOC1* and *ODDSOC2*. Vernalization is temperature sensitive; changes in temperature before, during, and after vernalization and the duration of exposure can affect the vernalization response. In light of climate change, an understanding of how fluctuating temperatures influence the vernalization response is necessary to maintain flowering in cereals.

Here, the impact of ambient temperature on the vernalization response in the major global crop bread wheat (*Triticum aestivum*) is examined. The results of this thesis reveal that winter wheat is successfully vernalized at higher temperatures of 10°C and 14°C. Vernalization was also achieved over winter in the field, under fluctuating temperature. Furthermore, extending the photoperiod to 22 h light/2 h dark was also sufficient for vernalization to proceed in an array of winter wheat cultivars. In addition, interpretation of the role of floral activator *ODDSOC1* and floral suppressor *ODDSOC2* in vernalization of wheat was explored. These findings indicated that *ODDSOC* genes may have a role in fine-tuning the vernalization requirement of cultivars in response to their environment and are potential targets for the development of climate-robust wheat.

Acknowledgements

This thesis would not be possible without the gargantuan encouragement and support that I've been fortunate enough to receive from so many individuals. I would like to thank the University of Leeds for funding and providing me with the opportunity to undertake this PhD project in the School of Biology. I would also like to thank Dr Andrea Harper at York for contributing the eQTL that kickstarted the *ODDSOC* journey, alongside various other collaborators over the past few years.

I am also incredibly grateful to all the Dixon lab for their unwavering support, including my supervisor Dr Laura Dixon for all her encouragement and wisdom.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

3-AT	3-amino-1,2,4-triazole
AGL	*AGAMOUS*-LIKE
AP1	*APETALA1*
CLV2	*CLAVATA2*
CDS	Coding sequence
CO1	*CONSTANS1*
dH$_2$O	deionised H$_2$O
DR	Double ridge cereal apex stage of development
DDF	*DWARF AND DELAYED FLOWERING*
ELF3	*EARLY FLOWERING3*
ELF4	*EARLY FLOWERING4*
eQTL	Expression Quantitative Trait Loci
FDR	False Discovery Rate
FDL	*FD-LIKE*
FLC	*FLOWERING LOCUS C*
FT1	*FLOWERING LOCUS T1*
FPF1	*FLOWERING PROMOTING FACTOR 1*
FYF	*FOREVER YOUNG FLOWER*
FPKM	Fragments per kilobase million
FRI	*FRIGIDA*
FUL	*FRUITFULL*
GEM	Gene Expression Marker

GO	Gene Ontology
GOI	Geno of interest
GRU	Germplasm Resource Unit
GI	*GIGANTEA*
GRP2	GLYCINE RICH PROTEIN 2
GRP2	GLYCINE-RICH RNA-PROTEIN 2
IWGSC	International Wheat Genome Sequencing Consortium
IPTG	Isopropyl β-D-1-thiogalactopyranoside
JIC	John Innes Centre
LB	Liquid Broth
LD	Long Day photoperiod
LLD	Long Long-Day photoperiod
lncRNA	long non-coding RNA
MIKC	MADS-box, Intervening, Keratin-like, C-terminal
MADS	*MINICHROMOSOME MAINTENANCE 1 (MCM1*), *AGAMOUS*, *DEFICIENS*, Serum-Response Factor (*SRF*)
MSA	Multiple Sequence Alignment
MEF2	Myocyte Enhancer Factor 2
NF	No flowering
OS1	*ODDSOC1*
OS2	*ODDSOC2*
PEP1	*PERPETUAL FLOWERING 1*
PHYC	*PHYTOCHROME C*

PIF4	*PHYTOCHROME INTERACTING FACTOR 4*
PRC2	*Polycomb Repressive Complex 2*
PEG	polyethylene glycol
Ppd-1	*Photoperiod-1*
PRR	*PSEUDO-RESPONSE REGULATOR*
qPCR	Quantitative Polymerase Chain Reaction
RG	Regular glasshouse conditions
SEP3	*SEPALLATA*
SD	Short Day photoperiod
SVP	*SHORT VEGETATIVE PHASE*
SNP	Single Nucletoide Polymorphism
SSD	Single Seed Descent
SB	Speed breeding conditions
SGV	Speed green vernalization (surface) conditions
SV	Speed vernalization (buried) conditions
SPL	SQUAMOSA-promoter binding protein-like
SOC1	*SUPPRESSOR OF OVEREXPRESSION OF CO1*
TOC1	*TIMING OF CAB EXPRESSION1*
TF	Transcription Factor
TFBS	Transcription Factor Binding Site
TSS	Transcription Start Site
TPM	Transcripts per kilobase million
TRP1	*TRYPTOPHAN BIOSYNTHESIS1*
VG	Vegetative cereal apex stage of development

VRT2 *VEGETATIVE TO REPRODUCTIVE TRANSITION 2*

VRN1 *VERNALIZATION1*

VRN2 *VERNALIZATION2*

VAL1 *VIVIPAROUS1/ABI3-LIKE1*

wRV *Warm regular vernalization conditions*

YPDA *yeast peptone dextrose adenine*

ZT *Zeitgeber*

ZCCT1 *ZINC FINGER CO, CO-LIKE, TOC1-DOMAIN 1*

ZCCT2 *ZINC FINGER CO, CO-LIKE, TOC1-DOMAIN 2*

Chapter 1: Introduction to this thesis

1.1 The significance of wheat as a global crop

In 1970, Norman Borlaug, recipient of the Nobel Peace Prize for leading the Green Revolution, opened his Nobel Lecture with the following statement:

> "Civilization as it is known today could not have evolved, nor can it survive, without an adequate food supply."

Borlaug was acutely aware of both the burgeoning strain of global food insecurity and the necessity of increasing yield in major crops such as wheat and rice. The decision to look beyond traditional agricultural practices and incorporate technology into plant breeding was a breakthrough that resulted in new cultivars such as the dwarf wheat of the Green Revolution that helped win Borlaug the 1970 Nobel Peace Prize (Hedden, 2003). 50 years later, agriculture still faces many challenges. Borlaug's opening address remains true; without an adequate food supply, civilization is in trouble. Based on the most recent FAO assessment, 45 countries worldwide require external assistance for food (FAO, 2023). It's estimated that between 691 and 783 million people worldwide experienced hunger in 2022; by 2030, projections indicate that 600 million people will have surpassed hunger to face chronic malnourishment (I. FAO, UNICEF, WFP, WHO, 2023). The expanding population will put severe strain on food availability, requiring increased food production (Godfray et al., 2010).

Wheat is a major global crop, consumed daily by some degree in every country in the world. Approximately 50 countries consumed more than 800 kcal per capita per day from wheat and wheat-based products in 2020 (FAOSTAT, 2020). Based on the most recent calculations from 2020, daily global wheat consumption exceeds that of rice, maize, cassava, and sorghum (**Figure 1.1**). In 2022, 801.8 million tonnes of wheat was produced worldwide (FAO, 2023). This production is uneven across the globe and highly sensitive to weather conditions. The top wheat producing region in 2022 was China, closely followed by the European Union (encompassing top wheat growing countries Germany, France, and Poland) and India (**Figure 1.2**) Wheat is a long-day crop: this means it preferentially flowers and produces grain when the length of the day exceeds the length of the night, such as

during summer. A defining feature of wheat is the option to plant the crop during late autumn or the following spring – in both cases, the crop flowers and can be harvested the following autumn. This has enabled farmers to grow two generations of wheat per year, boosting yield. Although winter wheat (i.e., wheat planted in late autumn that grows throughout winter) has a longer growth cycle compared to spring wheat (planted in spring and harvested in autumn), it is generally higher yielding and therefore favourably grown in temperate regions.

Winter wheat is the most dominant crop grown in Europe, with considerably more area dedicated to the production of winter wheat compared to spring (Le Gouis et al., 2020; Palosuo et al., 2011). The European Union experienced an above average winter wheat crop of 140 million tonnes for the 2022/2023 season, 6% higher than the previous three years (FAO, 2023). This level of production exceeded that of the United States (45.3 million tonnes) for the 2022/2023 growing season and is associated with favourable weather conditions in Europe at the end of 2022 (FAO, 2023).

In the UK, 15.7 million tonnes of wheat were produced in 2022, ranking 12th worldwide for wheat production (FAO, 2022; OECD et al., 2022). This amount is reflected in the total area dedicated to wheat production, with UK wheat areas consistently surpassing other key UK crops such as barley, oats, and oilseed rape (**Figure 1.3**). The mild temperate climate experienced by the UK is particularly favourable to growing winter wheat, making it more extensively grown than its spring counterpart (Cho et al., 2012). Considering that the producer price for UK wheat was just under £190 per tonne in 2022, winter wheat is a lucrative crop that shows no signs of falling out of popularity (OECD et al., 2022). Even so, there is always room for improvement: UK-grown winter wheat is photoperiod sensitive with a narrow flowering window (Sheehan & Bentley, 2021). This has increased the demand for greater range in flowering time for winter wheat varieties (Sheehan & Bentley, 2021).

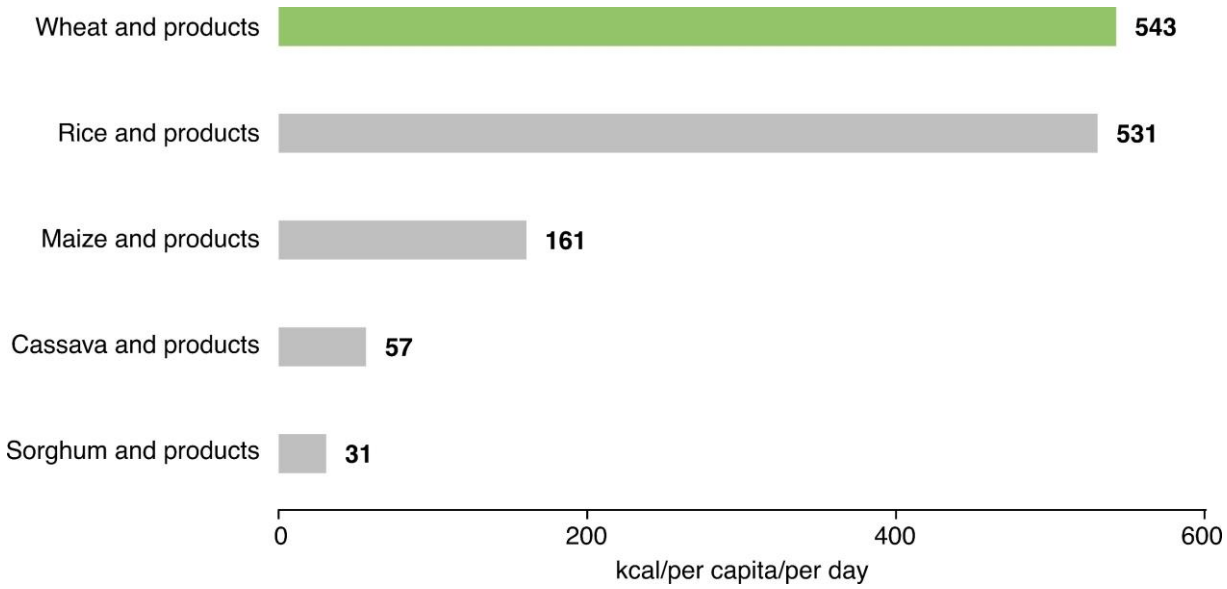


Figure 1.1: Global consumption of key crops. Source: FAO. Data sourced from 2020.

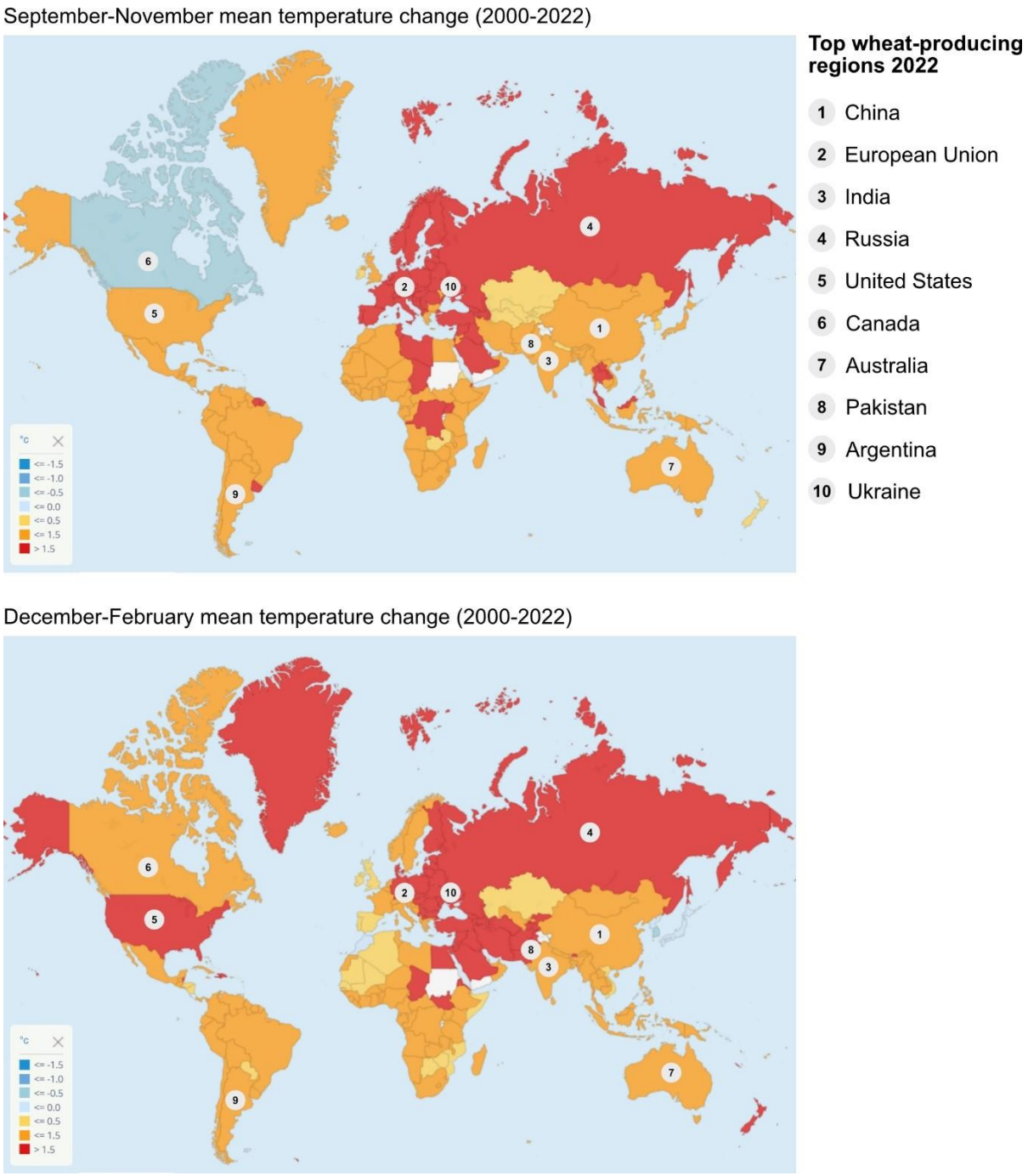


Figure 1.2: Map of the mean temperature change on land by country between the years 2000 and 2022. The top 10 wheat-producing regions for the 2022-2023 period are listed and labelled on the map. Mean temperature change data was sourced from the FAO (https://www.fao.org/faostat/en/#data/ET); top wheat-producing region data was sourced from OECD-FAO Agricultural Outlook (https://stats.oecd.org/).

1.2 An introduction to modern bread wheat

Modern bread wheat, also referred to as common wheat or *Triticum aestivum* (*T. aestivum*), is both deceptively young and deceptively complex. This species formed between 8,500 and 9,000 years ago (Levy & Feldman, 2022). *T. aestivum* is an allohexaploid, consisting of three sub-genomes (AABBDD) and 21 chromosome pairs in the format 2n = 6x = 42 (Feldman et al., 1972). This results in homoeologous sets of duplicated genes occurring on the A, B, and D sub-genome (**Figure 1.4**). The most recent annotation of the *T. aestivum* reference sequence revealed a staggeringly large 16,000 Mbp genome (Appels et al., 2018). In comparison, the complete human genome is approximately 3.05 Mbp (Nurk et al., 2022). A wheat genome of this size provides plant breeders with both challenges and opportunities; large genomes often contain vast regions of repetitive elements but also contain plenty of coding and regulatory genomic material, allowing the freedom to adapt and improve the species. To contextualize how *T. aestivum* developed this expansive genome, we can examine the evolutionary history of *Triticeae*, the tribe to which wheat and other cereals like barley and rye belong.

Although bread wheat is grown worldwide today, it originated in the south-west region of Asia known as the fertile crescent (Brown et al., 2009). This region contains the major rivers Jordan, Tigris, and Euphrates, culminating in highly fertile agricultural lands that facilitated the domestication of milestone crops such as barley (*Hordeum vulgare*), flax (*Linum usitatissimum*), chickpea (*Cicer arietinum*), and two wheat species: the tetraploid emmer wheat (*Triticum turgidum* subsp. *dicoccum*) and the diploid einkorn wheat (*Triticum monococcum*) (Brown et al., 2009). Emmer wheat (genome AABB; 2n = 4x = 28) supplied the A and B sub-genomes found in *T. aestivum*, whereas the wild grass *Aegilops tauschii* supplied the D sub-genome (Feldman & Levy, 2012; Levy & Feldman, 2022; Luo et al., 2017). Genome sequencing of these has suggested the most probable origin of each sub-genome, with the A sub-genome and the B sub-genome diverging from a common ancestral lineage approximately 7 million years ago, and subsequent hybridisation of the A and B sub-genomes giving rise to the D sub-genome approximately 5.5 million years ago (Marcussen et al., 2014).

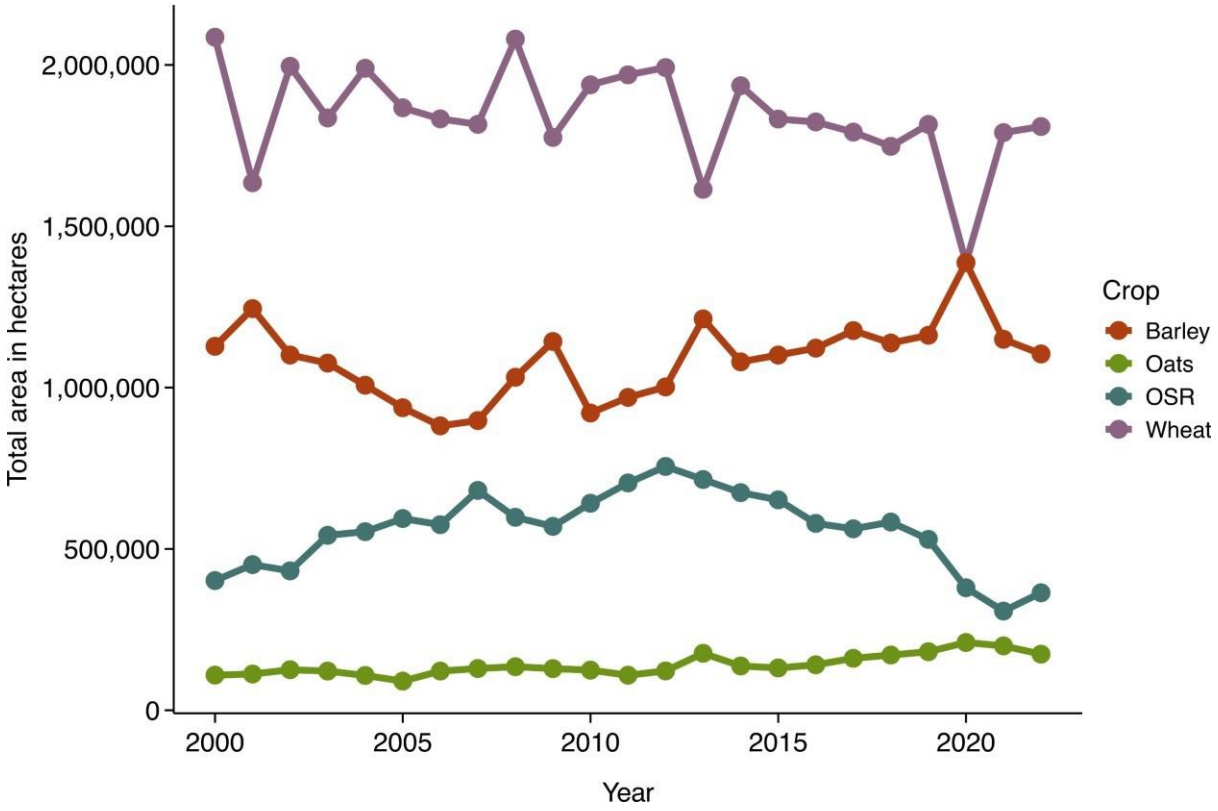


Figure 1.3: Total area in hectares of the UK devoted to crop production between the period of 2000-2022. Crops are wheat, barley, oilseed rape (OSR), and oats.

Since these ancient divergence and hybridisation events occurred millions of years ago, the formation of the AABBDD genome found in *T. aestivum* is relatively modern, as summarised in **Figure 1.4A**, and was likely assisted by domestication efforts. Following its inception in the Fertile Crescent, the cultivation of *T. aestivum* has expanded significantly, enabling the crop to be grown in a wide range of northern and southern latitudes. As a result, exposure to varied environments and subsequent breeding selection has created the explosion of different wheat cultivars that exist today. Therefore, an understanding of how bread wheat responds to environmental conditions is imperative in selecting which cultivar to grow.

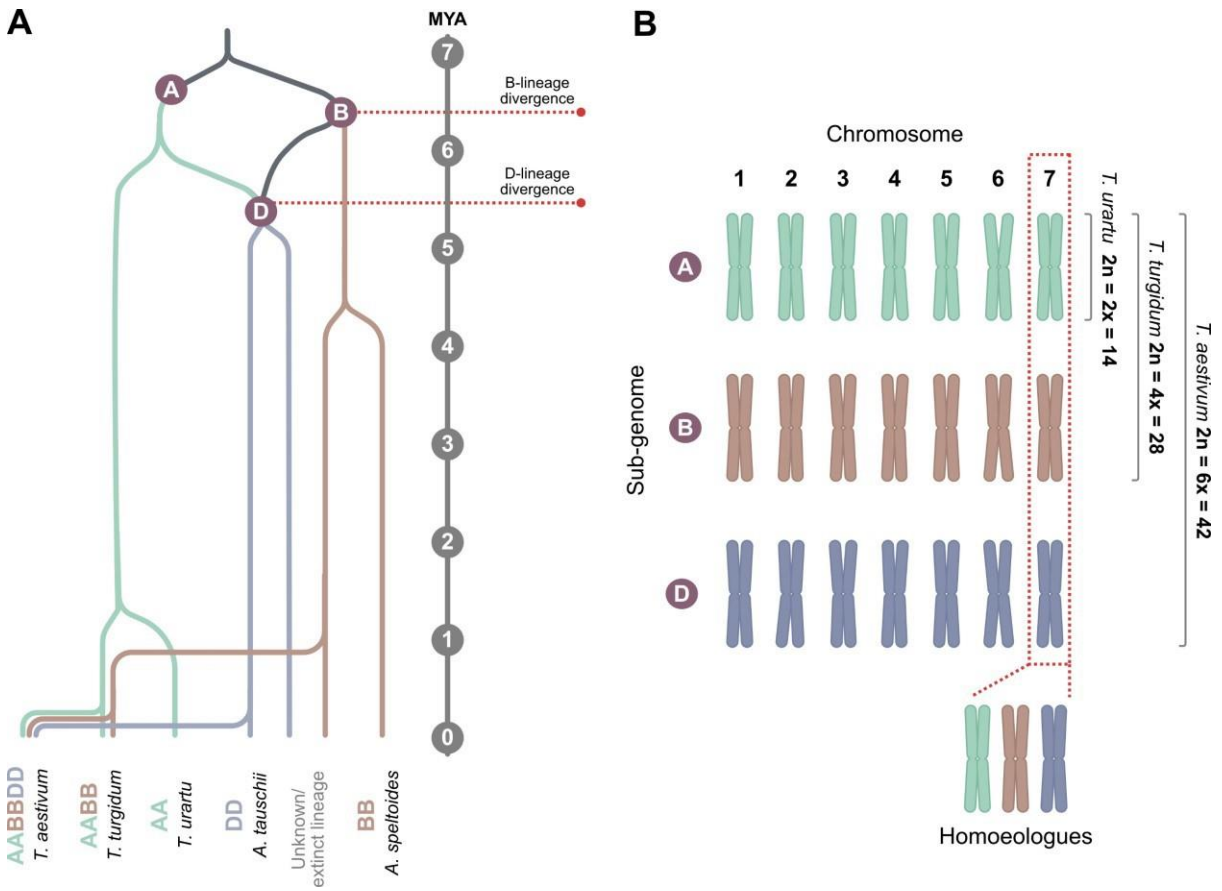


Figure 1.4: Schematic representation of the hexaploid wheat genome. (**A**) Evolution of the wheat lineages; phylogenetic tree adapted from Levy and Feldman (2022). (**B**) Breakdown of the chromosome number and multiple sub genomes in hexaploid bread wheat, including the concept of homoeologues referred to throughout this thesis.

1.3 Environmental factors that affect wheat development

Already, the impact of a warming climate on environmental responses is being reported – for example, flowering data indicates that UK plants now flower substantially earlier than 300 years ago, with earlier flowering observed in the south of the UK compared to the north (Büntgen et al., 2022). The most prominent shift was detected in herb species that now flower 32 days earlier, however all species had an earlier flowering phenotype (Büntgen et al., 2022). This corresponds with the index developed by Amano et al. (2010) which estimated flowering 5 days earlier for every 1 °C increase in temperature.

In cereals, drought and heat are regularly cited as the most severe environmental impacts on yield. For example, the combination of drought and heat stress during grain filling resulted in a 76% reduction in winter wheat yield (Balla et al., 2011). Temperatures between 32 °C and 35 °C during flowering and grain filling are sufficient to reduce yield – particularly concerning as the UK is expected to experience at least one day >32 °C during the crucial grain filling period in the near future (Jones et al., 2020). In addition to reducing yield, drought and temperature extremes also affect grain quality, with excessively high and low temperatures during grain filling reported to impact the level of starch and protein in the grain (Dupont & Altenbach, 2003; Farooq et al., 2011; Hurkman et al., 2003; Labuschagne et al., 2009). Climatic conditions that boost crop production in one region can have detrimental effects in another. High levels of rainfall in spring contributed to a forecasted above- average level of cereal production in parts of Asia; in contrast, erratic rainfall in East Africa caused flash flooding, resulting in significant crop losses (FAO, 2023).

Crop simulation models are being implemented to predict the impact of future climatic scenarios on wheat production. European winter wheat yield is predicted to be negatively impacted in the future due to increases in heat and drought stress (Webber et al., 2018). In central Europe, simulation models estimate a reduction in winter wheat yield due to as little as 2 °C increase in air temperature (Thaler et al., 2012). This is projected to add additional strain to water usage, with up to 37 mm more water required per season (Thaler et al., 2012).

1.4 Flowering in wheat is a finely tuned process

In their lifetime, plants face many crucial decisions. One such major decision: when is the most appropriate time to flower? Optimally timing flowering requires a recognition of seasonal cues, so that plants complete flowering under favourable conditions. In wheat, flowering is completed under warmer, longer days such as those associated with summer. This recognition of temperature and day length is essential and underpins the ability of wheat to flower. This process requires a complex network of genes which are controlled through the plants innate circadian mechanism. However, there are key regulators within the process and in wheat, recognition/responsiveness to day length (photoperiod) is regulated by *Photoperiod-1* (*Ppd-1*) (Beales et al., 2007). *Ppd-1* is a *PSEUDO-RESPONSE REGULATOR* (*PRR*), with closest homology across *PRR9*, *PRR7*, and *PRR5* genes in Arabidopsis (Nakamichi et al., 2005). The Arabidopsis *PRR* genes, which also include *TIMING OF CAB 1* (*TOC1*) and *PRR3*, are involved in circadian and photoperiod responsiveness in Arabidopsis (Para et al., 2007).

Ppd-1 has been extensively studied in Barley (*Hordeum vulgare*) and wheat (Beales et al., 2007; Díaz et al., 2012; Nishida et al., 2013; Turner et al., 2005; Wilhelm et al., 2009). Allelic variation of *Ppd-1* can confer changes in photoperiod sensitivity; a notable example of this is a mutation in *Ppd-D1* located on chromosome 2D that results in photoperiod insensitivity, transforming wheat from a long day to a day neutral plant (Beales et al., 2007). This discovery was implemented into the development of wheat as part of the Green Revolution and facilitated adaptation to a broader range of environments than would be possible with a photoperiod sensitive line (Beales et al., 2007; Borlaug, 1983). *Ppd-1* regulates flowering in cereals through interaction with *FLOWERING LOCUS T1* (*FT1*) (Beales et al., 2007; Turner et al., 2005). In wheat cultivars containing the photoperiod insensitive version of *Ppd- 1*, gradual stepwise expression of *FT1* is overridden and the plant experiences an accelerated flowering time (Gauley & Boden, 2021).

However, this central flowering pathway doesn't account for all flowering responses. Notably, for species that are adapted to temperate regions that experience moderate winters, plants that grow throughout winter have an additional pathway – the vernalization pathway.

1.5 Vernalization or learning to embrace "winter"

Vernalization is a cold-induced response that up-regulates expression of floral activators. The term was originally coined in the 1930s to describe a process that transformed a winter cereal into a spring cereal (Chouard, 1960). It is now understood that the vernalization process does not exclusively affect cereals, with many species having some form of vernalization requirement (Michaels & Amasino, 2000). Vernalization is an adaptive feature in temperate climates where cold winters are followed by a warmer spring; in climates that don't experience colder temperatures and seasonality, vernalization can delay flowering, hindering development and survival (McKeown et al., 2016). Vernalization therefore doesn't transform a winter habit into a spring habit, but rather facilitates overwintering in unfavourable cold conditions. Vernalization is quantitative. The length of cold exposure determines the rate at which subsequent flowering occurs; as the duration of vernalization increases, the flowering time decreases. This continues until the vernalization requirement is saturated, after which prolonged cold exposure doesn't further reduce flowering time.

The response to vernalization can be either facultative or obligate. A facultative response means that vernalization is not essential for the plant to flower, however it enables rapid flowering once the cold treatment is over (Michaels & Amasino, 2000). Many winter annuals germinate in autumn and flower the following spring – vernalization accelerates the rate of flowering, but is not essential, making it a facultative response. In biennials, flowering will not occur without vernalization – therefore, the response is obligate (Amasino & Sung, 2004; Michaels & Amasino, 2000). In cereals such as wheat, the vernalization requirement distinguishes winter varieties from spring varieties (Yan et al., 2003).

In *Arabidopsis thaliana* (referred to as Arabidopsis from here onwards), the vernalization pathway is one of the four major pathways required for flowering to occur (Boss et al., 2004). The late flowering phenotype present in winter varieties of Arabidopsis is regulated by the dominant *FLOWERING LOCUS C* (*FLC*) and *FRIGIDA* (*FRI*) loci (Michaels & Amasino, 1999). The vernalization process downregulates expression of these repressors, facilitating accelerated flowering in Arabidopsis once vernalized (Michaels & Amasino, 1999; Sheldon et al., 1999).

For the vernalization requirement to be satisfied, the optimum temperature varies depending on the plant. In species that are adapted to warmer climates, vernalization can occur at 13°C and higher; in contrast, certain cereals are vernalized at -6°C (Michaels & Amasino, 2000). The general range of temperatures in which vernalization occurs is between 1-7 °C (Michaels & Amasino, 2000). Similarly, the optimum duration of cold treatment also varies. For example, celery can undergo vernalization after 8 days of exposure to cold treatment, whereas other plants require up to 12 weeks (Michaels & Amasino, 2000). The wide range in vernalization requirements is particularly important in agriculture; an understanding of the vernalization pathway has enabled the development of varieties that flourish when grown in conditions that match their vernalization requirement.

Both the well-studied and characterised model plant Arabidopsis and temperate grasses can respond to vernalization, however the vernalization pathways differ between the groups; there are similarities in the regulatory profiles of key genes, but the genes involved are distinct (Yan et al., 2004). My research project focuses on the vernalization response in temperate grasses, with particular attention paid to the major crop wheat.

1.6 Physiology of vernalization

As mentioned, vernalization catalyses flowering in cereals by facilitating the transition from vegetative to reproductive growth. This transition is achieved by the regulation of an array of genes; namely those that control the identity of the shoot apical meristem which regulates whether the shoot apex generates vegetative or reproductive structures. When a plant is in the vegetative stage of development, the apical meristem produces leaf primordia that will eventually form the leaves. These primordia appear as ridges along the apex as it develops and extends. If plants that require vernalization to flower aren't vernalized, they will remain in this vegetative stage and continue to produce leaf primordia.

During vernalization, expression of floral activators is up regulated until a threshold is reached; at which point, the apex transitions to the reproductive stage (A. Distelfeld et al., 2009). Once the transition to reproductive development is initiated, the shoot apex starts to produce secondary meristems, known as floral primordia. The floral primordia are situated above each leaf primordia; this creates a distinct double-ridge appearance and are an indicative feature that inflorescence initiation is occurring. As the apex develops, the floral primordia split into inflorescence branches containing the florets or flowers of a cereal (Zadoks et al., 1974). For the majority of cereals, the later stages of apex development (i.e., spike elongation and spikelet formation) occur when day length increases (A. Distelfeld et al., 2009). Therefore, the colder, shorter days initiate the transition to reproductive growth in the apex and the reproductive structures mature in warmer, longer days. The timing of this process determines whether flowering is successful and highlights the importance of both temperature and photoperiod in the vernalization response.

1.7 Molecular mechanism behind vernalization in cereals

In cereals, three main *VERNALIZATION* (*VRN*) genes are thought to be involved in the vernalization process: *VRN1*, *VRN2*, and *FT1* (**Figure 1.5**). In some vernalization models, *VRN4* is also incorporated into the pathway; however, this gene wasn't focussed on within this thesis. Of the three other genes, *VRN1* is the cornerstone of the vernalization pathway. *VRN1* is the major locus identified in classic QTL studies for cereal vernalization and its characterisation has identified that it is a cold-induced floral activator (Oliver et al., 2009). Despite sharing a name

with *VRN1* in Arabidopsis, the *VRN1* cereal gene is functionally and genetically different; this is also the case with *VRN2* in cereals (Yan et al., 2004). Cereal *VRN1* is a MADS-box gene and a homologue of the *APETALA1* (*AP1*)/*FRUITFULL* (*FUL*) genes found in Arabidopsis (Shimada et al., 2009; Yan et al., 2003). *VRN1* is expressed in both the apical meristem and the leaves of cereals (Chen & Dubcovsky, 2012). Through mutant studies it has been identified that *VRN1* up-regulates the *FT* homologue *FT1* (Shimada et al., 2009) (**Figure 1.5**). The FT protein acts as a signalling molecule in plants, activating floral meristem identity genes and promoting flowering (Corbesier et al., 2007; Kobayashi et al., 1999; Shimada et al., 2009). Shimada et al. (2009) reported that the upregulation of *VRN1* under long-day conditions resulted in accumulation of *FT1* transcripts, indicating that *VRN1* is upstream of FT and is a regulator of its expression. Furthermore, the MADS-box transcription factor encoded by *VRN1* binds to CArG-box in the promoter region of *FT*, up-regulating expression of the gene (Tanaka et al., 2018).

Most of the variation in vernalization responses in wheat is attributed to allelic differences in *VRN1* (Chen & Dubcovsky, 2012). Winter wheat carries recessive alleles of *VRN1* on each genome; in contrast, spring wheat contains at least one dominant version of *VRN1* and therefore the vernalization requirement is greatly or completely reduced (Trevaskis et al., 2003). Of the three alleles, the version on the A genome (*VRN1*-A) has the strongest impact on flowering time and individuals with this dominant allele don't require any vernalization to flower (Trevaskis et al., 2003). This dominant allele is conferred by a deletion in either the promoter or first intron of *VRN-A1* (Fu et al., 2005). *VRN-B1* and *VRN-D1* also confers a spring habit but to a weaker effect; if *VRN-A1* is missing, the plant will still flower earlier than winter varieties, however vernalization further accelerates the flowering process. These varieties are called semi-spring or facultative wheat (Trevaskis et al., 2003).

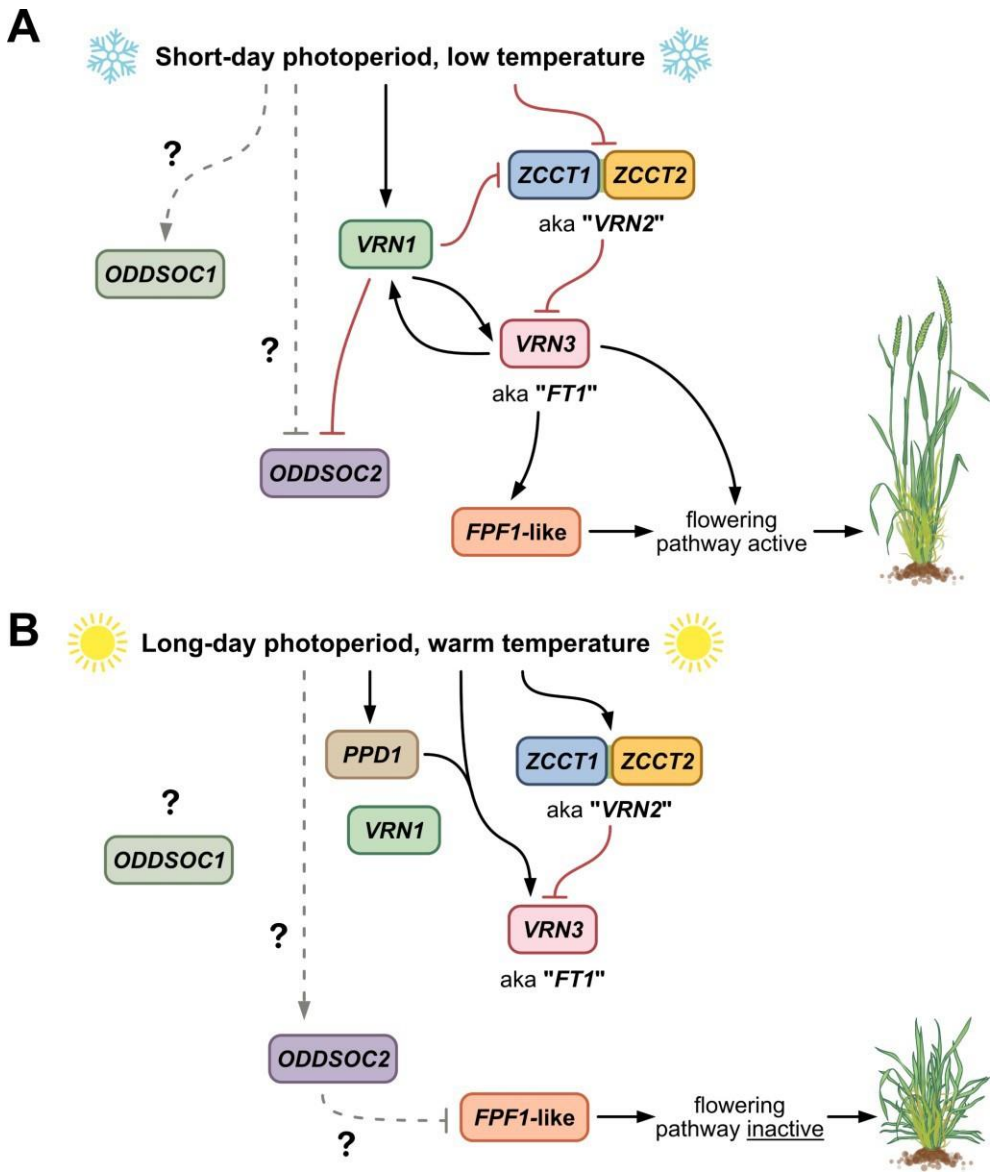


Figure 1.5: Schematic illustrating the vernalization pathway in cereals. (A) Vernalization pathway active, and (B) vernalization pathway inactive.

It is thought that the evolution of spring varieties arose from mutations or modifications of *VRN1* in ancestral winter varieties. Initially, this was attributed to mutations in the promoter region of *VRN1*, with Yan et al. (2004); Yan et al. (2003) reporting that deletions in the promoter region of *VRN1* in diploid wheat was consistent with a spring growth habit. However, further research indicated that a substantial deletion in the first intron of the gene can also be responsible for spring varieties of wheat and barley (Fu et al., 2005). The deleted region of the first intron contains regulatory elements for vernalization – therefore, without this region vernalization cannot occur. By removing the vernalization requirement, *VRN1* is constituently expressed and the plant flowers rapidly, independent of prolonged cold exposure (Fu et al., 2005; Trevaskis et al., 2007). This results in a spring growth habit in cereals. *VRN1* was shown to be strongly expressed in spring wheat but not expressed in winter wheat that hadn't been vernalized (Trevaskis et al., 2003). Upon receiving vernalization treatment, *VRN1* was shown to be greatly up regulated in winter wheat (Trevaskis et al., 2003).

As well as up-regulating floral activator *FT1*, *VRN1* also down-regulates expression of *VRN2*. *VRN2* is understood to act as a predominant repressor of flowering (Yan et al., 2004). The *VRN2* locus contains the linked zinc finger-CCT domain genes *ZCCT1* and *ZCCT2* (A. Distelfeld et al., 2009; Yan et al., 2004). These repressors inhibit *FT1* when day length reduces at the end of autumn; this prevents initiation of flowering in winter (Chen & Dubcovsky, 2012; Kippes et al., 2014). Since FT1 is a mobile protein that is expressed in the leaves and travels to the apical meristem, when *VRN2* is expressed in the leaves it prevents this journey from happening (Chen & Dubcovsky, 2012). To counteract this response, expression of *VRN1* in the leaves down-regulates *VRN2*, inducing *FT1* synthesis to accelerate flowering (Chen & Dubcovsky, 2012). This interplay between *VRN1* and *VRN2* is an important aspect of flowering time. In diploid barley and wheat (*Triticum monococcum*), mutations in redundant versions of the *VRN2* allele results in a spring habit (Kippes et al., 2016). Similarly, non-functional versions of *ZCCT1* and *ZCCT2* in tetraploid wheat (*Triticum durum*) also results in a spring growth habit that doesn't require vernalization (Assaf Distelfeld et al., 2009).

In addition to the *VRN1*-*VRN2* regulatory loop, *VRN1* and *FT1* also interact with each other to up-regulate their expression. Of the multiple *FT* genes in cereals, *FT1* shares the greatest similarity with Arabidopsis FT (Bennett & Dixon, 2021; Yan et al., 2006). Like Arabidopsis *FT*, wheat *FT1* is a major component of the florigen signal. Florigen is a mobile signalling molecule that is stimulated by day length to promote flowering (Kobayashi & Weigel, 2007). Under long day conditions, *FT1* is expressed in the leaves and its expression pattern correlates with that of *VRN1* (Li et al., 2015). In wheat, up-regulation of *VRN1* by *FT1* is indirect; instead, the FT1 protein interacts with two bZIP proteins, FDL2 and FDL6. Of the two, FDL2 binds to a binding site in the promoter region of *VRN1*, up-regulating expression (Li & Dubcovsky, 2008).

Finally, *VRN4* has also been shown to be part of the vernalization response and plants carrying this gene have a reduced vernalization requirement (Kippes et al., 2015). *VRN4* arose from a duplication event which included the region of chromosome 5AL containing *VRN1* – for this reason, *VRN4* is sometimes characterised as a copy of *VRN1* (Kippes et al., 2015). The *VRN4* copy of *VRN1* contains three single nucleotide polymorphisms (SNPs) within the first intron that prevent the binding of GLYCINE-RICH RNA-PROTEIN 2 (GRP2), a suppressor of *VRN1* (Kippes et al., 2015). Since GRP2 is unable to bind to the site, expression of *VRN4* isn't downregulated. However, investigation into the role of *VRN4* in the vernalization response is limited; it's significance in vernalization signalling is not yet clear.

As well as the *VRN* genes, additional candidates of the vernalization pathway have been identified – *ODDSOC1* (*OS1*), *ODDSOC2* (*OS2*), and *VEGETATIVE TO REPRODUCTIVE TRANSITION 2* (*VRT2*). *ODDSOC1* and *ODDSOC2* are MADS-box gene involved in the activation and repression of flowering in monocots (Greenup et al., 2010; Ruelens et al., 2013). Recent findings indicate that *ODDSOC2* is either a homologue or highly homologous to *FLC* in Arabidopsis (Ruelens et al., 2013; Sharma et al., 2017). Furthermore, transcript levels of *ODDSOC2* prior to vernalization is a good indicator of the required length of vernalization in *Brachypodium distachyon* (Sharma et al., 2017). Vernalization represses expression of *ODDSOC2*, and further down-regulation by *VRN1* in

vernalized plants to enable flowering (Greenup et al., 2010). The binding of *VRN1* to the promoter of *ODDSOC2* to down-regulate expression supports the hypothesis that *ODDSOC2* is part of the vernalization signalling pathway (Deng et al., 2015; Sharma et al., 2017). In *ODDSOC2* overexpression studies, inflorescence initiation and flowering was significantly delayed (Greenup et al., 2010).

In addition to *ODDSOC2*, *VRT2* is also potentially involved in the vernalization pathway. *VRT2* is a homologue of *SHORT VEGETATIVE PHASE* (*SVP*) in Arabidopsis, and the *VRT2* protein shares 51% amino acid sequence similarity with SVP (Kane et al., 2005). SVP suppresses expression of *FT* in Arabidopsis, therefore inhibiting flowering (Lee et al., 2007). Initially, it was thought that, as an SVP homologue, *VRT2* must act as a floral repressor (Kane et al., 2005). This was supported by the discovery that *VRT2* encodes a transcription factor from a MADS-box group associated with suppression of flowering and has an expression pattern similar to *VRN2* (Kane et al., 2005). However, recent research indicates that *VRT2* is a floral promoter that is induced by vernalization, much like *VRN1* (Xie et al., 2021). In-depth transcriptomic analyses revealed that *TaVRT2* binds to the promoter of Ta*VRN1*, up-regulating expression (Xie et al., 2021). *VRT2* was shown to facilitate a positive feedback loop of *VRN1* expression, promoting flowering post-vernalization (Xie et al., 2021).

1.8 The role of antisense transcripts in vernalization in Arabidopsis

In Arabidopsis, *FLC* is regulated by a population of long non-coding antisense transcripts collectively known as *COOLAIR* (Liu et al., 2010; Swiezewski et al., 2009). These *COOLAIR* transcripts are involved in the epigenetic silencing of *FLC* and facilitate co-ordination of switching chromatin states (Csorba et al., 2014). Low temperature upregulates *COOLAIR* (Rosa et al., 2016; Swiezewski et al., 2009). Additional antisense transcripts, *COLDAIR* and *COLDWRAP*, are also involved in *FLC* silencing, with the transcripts and the polycomb repressive complex PRC2 forming a repressive intragenic chromatin loop during vernalization (Kim & Sung, 2017).

1.9 The effect of ambient temperature on vernalization of cereals

Since vernalization is influenced by both temperature and day length, the effect of ambient temperature on a plant is key to understanding the vernalization pathway and subsequent flowering. Ambient temperature affects the epigenetic memory of vernalization; exposure to high temperatures can remove epigenetic markers and undo the effect of vernalization (Bouché et al., 2015). This reversal effect is known as devernalization and emphasizes the importance of temperature on vernalization- induced flowering. Although vernalization is associated with winter and colder temperatures, the process is affected by a wider range of temperatures than initially thought. While it is recognised that floral activator genes respond to the cold temperatures associated with vernalization, there is now evidence to suggest that floral activators and repressors also respond to warm temperatures (Dixon et al., 2019).

In winter wheat, exposure to higher temperatures during and after vernalization affected the promotion of flowering (Dixon et al., 2019). Vernalization down-regulates the floral repressors *VRN2* and *ODDSOC2*, facilitating activation of the flowering pathway. However, Dixon et al. (2019) reported that these genes were re-activated when plants were exposed to higher temperatures. This treatment affected flowering time and floret number (Dixon et al., 2019). In addition, different cultivars had different temperature responses. Genetically similar cultivars had a similar vernalization response to consistent cold exposure, but when this treatment was interrupted or followed up with periods of high temperature, the flowering response between cultivars was significantly different (Dixon et al., 2019).

As well as ambient temperatures during and after vernalization, the temperature immediately before cold exposure also affects the vernalization response. For example, *FLC* is down regulated during vernalization to facilitate spring flowering. However, *FLC* expression levels in winter annual oilseed rape (*Brassica napus*) have been shown to decrease during the warmer autumnal period preceding winter (O'Neill et al., 2019). This enables the vegetative to reproductive transition to occur at the start of winter, and for the inflorescence meristem to overwinter before successfully flowering in spring (O'Neill et al., 2019). Increasing the autumnal

temperature by as little as 5°C reversed the effect, up-regulating *FLC* expression and delaying the reproductive transition (O'Neill et al., 2019). This temperature increase also slightly delayed flowering in the spring.

1.10 Summary and thesis aims

The main aim of this thesis has been to investigate how vernalization in cereals proceeds under conditions that would not typically be associated with vernalization. This includes warmer temperatures and longer photoperiods than those that have previously been examined. Moreover, this thesis aims to look beyond the usual suspects that are integral to vernalization in flowering in an attempt to understand the unaccountable differences in vernalization responses that have been observed in cereal cultivars.

Chapter 2: Challenging the vernalization response in bread wheat

2.1 Introduction

2.1.1 Methods to improve the rate of plant development: temperature as a variable

Temperature adaptability underpins both sowing time and geographical range of crops. For example, rice is especially susceptible to cold temperatures, which has limited the geographical spread of the crop (Bin Rahman & Zhang, 2018). As a result, rice domestication has taken place in tropical regions and no vernalization pathway has been identified (Shrestha et al., 2014). In contrast, bread wheat is tolerant to low temperatures and can be grown in a wide range of temperate regions, with the *VRN-A1* locus believed to control tolerance to low temperature (Dhillon et al., 2010; Laudencia-Chingcuanco et al., 2011). Tolerance to low temperature is intrinsically linked to vernalization response. For example, in cultivars with a spring growth habit, neither a high tolerance to freezing temperatures nor a high vernalization requirement is necessary, as these cultivars aren't typically grown throughout winter (Dhillon et al., 2010). Conversely, for a winter growth habit, tolerance to low temperatures conveys hardiness and survival through winter, while also facilitating activation of the vernalization pathway. The range of temperatures associated with low temperature tolerance is diverse, with some winter wheat cultivar having a tolerance level of up to -22 °C (Båga et al., 2007). However, the range of low temperatures associated with vernalization is narrower, with optimal temperatures believed to be between 0-10 °C (Chouard, 1960).

With this in mind, examination of the molecular mechanism driving cereal vernalization is typically investigated under artificial vernalization temperatures of 2-6 °C . This is in contrast with increasing winter temperatures; for example, winter UK temperatures exceeded 20 °C for the first time on record in February 2019 (Kendon et al., 2020). Despite this increase in winter temperature, total yield of winter cereals in the UK was higher in 2019 compared to 2018 (https://www.gov.uk/government/statistics/farming-statistics-provisional-crop-areas-yields-and-livestock-populations-at-1-june-2019-united-kingdom). Furthermore,

maximum daily winter temperature in the UK is generally between 16-17 °C whereas minimum daily temperature is between 6-8 °C (Kendon et al., 2020). This would suggest that previous investigation into the vernalization response has been carried out under temperatures closer to the lower range of winter UK temperatures. Therefore, this raises the question of how the genetic elements involved in vernalization behave at the higher temperature range.

In Arabidopsis, vernalization was still completed at temperatures of 14 °C (Duncan et al., 2015). Furthermore, this study indicated that vernalization proceeded in autumn in Sweden, well before the first snowfall (Duncan et al., 2015). In bread wheat, higher ambient temperatures of 11-25 °C post-vernalization have been shown to up regulate the floral repressor *VRN2* locus after vernalization (Dixon et al., 2019). In this study, winter wheat vernalization was also examined under low temperatures of 6 °C and long day (16 hr light/8 hr dark) conditions after 12 weeks vernalization (Dixon et al., 2019). This resulted in flowering for winter cultivars Buster and Charger, indicating that vernalization was completed even under a long day photoperiod (Dixon et al., 2019). Historically, investigation into vernalization has been carried out under short day (8 hr light/16 hr dark) conditions as this is most representative of winter. Furthermore, a short day photoperiod down regulates expression of the *VRN2* locus (Dubcovsky et al., 2006). This short day requirement can be considered a crucial variable of vernalization, excluding it from other methods of accelerated plant growth due to extended photoperiod, such as speed breeding.

2.1.2 Photoperiod as a variable to improve the rate of plant development

Historically, plant growth under an extended photoperiod – including continuous light – can be beneficial for certain crops (Velez-Ramirez et al., 2011). For example, plant growth under a long day photoperiod exhibited traits such as increased leaf expansion (Adams & Langton, 2005). Plant growth under 24-hr continuous light is also feasible, however this can induce undesirable effects including leaf chlorosis and necrosis (Sysoeva et al., 2010; Velez-Ramirez et al., 2011). Despite these potential side effects, growth under continuous light has been shown to hasten growth and has been used to develop wheat suitable for growth under constant light on space stations (Bugbee

& Koerner, 1997). This research influenced the development of the speed breeding protocol to accelerate wheat breeding (Watson et al., 2018). Here, an extended photoperiod of 22 hours at 22 °C was used with 2 hours dark at 17 °C; this inclusion of a dark period was incorporated to support function of circadian gene expression (Watson et al., 2018). Currently, this method has been shown to be effective in a variety of crops (Ghosh et al., 2018; Watson et al., 2018). The method is most effective for long day and day neutral species that do not require vernalization, otherwise winter cultivars with a vernalization requirement need preliminary standard vernalization to benefit from speed breeding post-floral transition (Ghosh et al., 2018). However, there is evidence that exposing short day species to an extended photoperiod prior to vernalization can hasten flowering; for example, growing the short day species *Amaranthus* spp. L. under 16 hr light/8 hr dark at 35 °C for 10 weeks before transfer to 8 hr light/16 hr dark at 30 °C rapidly accelerated flowering (Stetter et al., 2016). This suggests that there are opportunities to hasten development even in short day species.

2.1.3 Duration of vernalization as a variable to improve the rate of plant development

Finally, duration of vernalization is also a variable that has the potential to be optimised. The current standard protocol for vernalization involves exposing plants to low temperatures for 6-12 weeks, with cultivars with a high vernalization requirement believed to need between 10-12 weeks of cold for vernalization to be completed (Andrés & Coupland, 2012; Xu & Chong, 2018). As with temperature, this duration aims to replicate field conditions, with plants experiencing several weeks of low temperature when grown in the field over winter. However, field studies in Brassicaceae have shown that vernalization is completed in autumn, after which the plant continues to grow slowly throughout winter (Duncan et al., 2015; O'Neill et al., 2019). This indicates that the vernalization requirement was saturated by winter and that further cold exposure beyond this point wasn't essential.

Since vernalization in Brassicaceae can be initiated under warmer autumnal conditions, I considered that vernalization in bread wheat could also proceed under warmer temperatures than those typically used to study vernalization. In addition autumnal daylength exceeds the 8 hr short day photoperiod used in typical vernalization protocols; in combination with the findings from Dixon et al. (2019) that showed that winter wheat can be vernalized under a 16 hr photoperiod, I decided to investigate whether an extended photoperiod reduced time until flowering. Furthermore, I speculated whether increasing the temperature and/or photoperiod would impact the central genetic components of the vernalization pathway in cereals.

2.2 Materials and methods

2.2.1 Plant material

The hexaploid winter *Triticum aestivum* cvs. Claire, Buster, Charger, and Hereward were used to investigate the speed vernalization protocol on European winter wheat. These four cultivars have similar genetic background but varying vernalization requirement, attributed to differences in *VRN-A1* copy number. This difference is as follows: Claire, a single haploid copy of *VRN-A1* and a low vernalization requirement; Buster, two copies of *VRN-A1* and a moderate vernalization requirement; Charger and Hereward, three copies of *VRN-A1* and a high vernalization requirement. *VRN-A1* copy number identification and characterisation has been previously reported by Díaz et al. (2012) and Dixon et al. (2019).

2.2.2 Growth conditions

The conditions used to evaluate wheat (*Triticum aestivum*) lines were as follows:

wRV-RG: 8 h light:16 h dark 8°C followed by transfer to 16 h light:8 h dark 22°C wRV–RG: 8 h light:16 h dark 10°C followed by transfer to 16 h light:8 h dark 22°C wRV-RG: 8 h light:16 h dark 14°C followed by transfer to 16 h light:8 h dark 22°C wRV–SB: 8 h light:16 h dark 10°C followed by transfer to 22 h light:2 h dark 22°C
SV (and SGV)–SB: 22 h light:2 h dark 10°C followed by transfer to 22 h light:2 h dark 22°C

Where wRV = warm regular vernalization; SV = speed vernalization; SGV = speed green vernalization; RG = regular glasshouse; and SB = speed breeding glasshouse conditions.

Warmer conditions were used in wRV than are classically used in RV to enable direct comparison of photoperiods for phenotype and gene expression analysis.

Seeds were germinated for 2 days in darkness at 4°C in 9-cm Petri dishes that had a layer of filter paper saturated with 5 ml deionised H_2O (dH_2O). The germinated seeds were transferred to 3 × 3 cm cell pots of John Innes Centre (JIC) cereal mix

(40% medium grade peat, 40% sterilised soil, 20% horticultural grit, 1.3 kg/m3 PG mix 14-16-18 + Te base fertiliser, 1 kg/m3 osmocote mini 16-8-11 2 mg + Te 0.02% B, Wetting agent, 3 kg/m3 maglime, 300g/m3 exemptor). For the wRV-RG experiments, plants were vernalized in a Sanyo MLR-352 plant growth chamber under 200-250 µmol/m^2/s light intensity before transferral to regular glasshouse conditions. For the wRV-SB, SV-SB, and SV-SB experiments, plants were placed under vernalization conditions in Snijders MICROCLIMA MC1000 plant growth cabinets with a light intensity of approximately 900 µmol/m^2/s. Plants were watered when required, and no additional nutrients were added. At 1- or 2-week intervals, plants were sampled for gene expression and apex analysis, and three plants were transferred to glasshouse growth conditions for RG (PhytoLux Plessey; model ATTIS-7) and SB (Heliospectra; model MITRA) treatments. Plants were placed in cereal mix (as above) in 9 × 9 cm pots and tissue was sampled 1 and 2 weeks after transfer to SB conditions.

For the SGV treatment, seeds were placed in 9-cm Petri dishes containing filter paper and 5 ml dH_2O and kept in complete darkness at 4°C for 48 h. Seeds were then placed on the soil surface (JIC cereal mix) in a P24 seedling tray. Care was taken to press the seed into the soil surface while ensuring that the seed remained uncovered and exposed to the light. Plant trays were watered from the base, and a spray bottle of dH_2O was used to mist the soil surface; particular care was taken when misting the soil during the 1st week of growth, when the roots were anchoring into the soil. At set weekly intervals, plants were moved into SB conditions and flowering was recorded.

2.2.3 Plant phenotyping

2.2.3.1 Phenotyping the flowering response

Flowering time was recorded as half-ear emergence (Zadok's scale 55). Plants that did not flower after 170 days were recorded as "non-flowering". A control group consisting of the same four genotypes (n = 10; s = 40) was grown under constant SB conditions, and a representative for each genotype was imaged once the plant had reached maturity. Total spikelet number per ear was counted after ear emergence. Flowering time and spikelet number were plotted as violin plots using the R packages 'ggplot2' and 'ggpubr' (https://ggplot2.tidyverse.org/). Statistical

analysis was performed using Student's t-test or Mann Whitney U test.

2.2.3.2 Phenotyping apex development during vernalization

For apex samples, three plants from winter wheat cv. Buster, Charger, Hereward and Claire were dissected for each apex sample (n = 3) at 2-week intervals for wRV and 1-week intervals for SV. Apices were dissected with a FEATHER® incision micro scalpel and imaged using a Keyence VHX-6000 microscope. Using a fixed magnification and scale, images were captured and apex length was measured using ImageJ (Schneider et al., 2012). Measurements were taken in micrometres (µm) from the base to the tip of the apex. The apex length was the standard error of the mean (SEM) for 3 biological replicates.

In addition to measuring the length of the apex, the point at which vegetative to reproductive tissue growth was observed was also recorded. This is defined as the transition from the vegetative (VG) stage to the double ridge (DR) stage of growth (Kirby & Appleyard, 1987). Here, floral primordia form above the leaf primordia along the apex, creating a distinct "double ridge" appearance and indicating that the apex now can produce reproductive tissue (Kirby & Appleyard, 1987; Trevaskis et al., 2007).

2.2.4 Expression analysis

2.2.4.1 Quantitative PCR (qPCR) analysis

To examine expression in leaf tissue, leaf samples from three plants for each of three biological replicates (n = 3) were taken at each sampling stage. Samples were taken from the base of the youngest leaf at zeitgeber time ZT1 and flash frozen in liquid nitrogen. To investigate gene expression during SV and SGV, leaf tissue was sampled at 1, 2, 3, 4, 5, 6, 7, and 8 weeks of growth under vernalization conditions. Additionally, leaf tissue was sampled after 1 and 2 weeks in SB conditions post-SGV to monitor changes in gene expression between conditions. The tissue was lysed using the TissueLyserLT (Qiagen) with 3-mm steel ball bearings, and total RNA was extracted using the Spectrum™ Plant Total RNA Kit (Sigma-Aldrich) following the manufacturer's recommended protocol.

In addition to sampling leaf tissue across the timecourse, apical tissue from cv.

Hereward was sampled at zeitgeber time ZT1 after 3 weeks of growth under SGV conditions. Three biological replicates were sampled, with each replicate containing >15 individual apices. Apices were dissected using a FEATHER® incision micro scalpel with the assistance of a Keyence VHX-6000 microscope and incised at the base of the apex. After dissection, the apex was immediately flash frozen using a 1.5 mL microcentrifuge tube. The microcentrifuge tube was partially submerged in a liquid nitrogen bath using a floating tube rack until all apices had been sampled, after which the tube was stored at -80 °C. To extract RNA, the tissue was first disrupted using a sterile porcelain pestle and mortar. Approximately 50 mL of liquid nitrogen was added to the mortar to ensure the apices didn't defrost as the sample was being ground to a fine powder. Total RNA was extracted using the RNeasy® Plant Mini Kit (Qiagen) following the manufacturer's instructions and eluted in 50 µL dH_2O.

Prior to first-strand cDNA synthesis, leaf and apex RNA samples were treated using the RQ1 RNase-Free DNase kit (Promega). Briefly, 5 µL of RNA was added to a 0.1 mL nuclease-free PCR tube containing 1 µL RQ1 RNase-Free DNase and 1.5 µL RNase-Free DNase 10X Reaction Buffer, with a final reaction volume of 7.5 µL. The reaction was incubated at 37 °C for 30 mins, after which 1 µL RQ1 DNase Stop Solution was added, and the sample was incubated for a further 10 mins at 65 °C to terminate the DNase reaction. Next, 1 µL of 50 µm oligo(dT) and 1.5 µL 10 mM dNTP mix (equal ratio 10 mM dATP, dCTP, dGTP, and dTTP) was added to the sample and incubated at 65 °C for 5 mins, preceded by 5 mins incubation on ice before proceeding to the cDNA synthesis step. First-strand cDNA synthesis was performed using SuperScript™ III Reverse Transcriptase (Invitrogen) and RNaseOUT™ (Invitrogen) as per the manufacturer's protocol.

The cDNA was diluted to 50 ng/µL and 2.5 µL was combined with 5 µL GoTaq® qPCR Master Mix (Promega) and 2.5 µL primer mix (1 µL 100 µm forward primer, 1 µL 100 µm reverse primer, 46 µL nuclease-free water), with a final reaction volume of 10 µL. Two technical replicates were used for each sample, and each reaction was also duplicated to receive primers to amplify the gene of interest (GOI) and the reference gene *TraesCS5A02G015600* (Borrill et al., 2016). Primers were designed from the gene of GOI coding sequence (CDS) using the Primer-BLAST tool (https://www.ncbi.nlm.nih.gov/tools/primer-blast/; Ye et al. (2012)); all primers

used are listed in **Table 2.1**. Reaction components were pipetted into a sterile 96-well non- skirted qPCR plate, sealed with optically clear plate sealing film, and centrifuged briefly using a PCR Plate Centrifuge II (Avantor Sciences) to collect the reaction in the bottom of the plate well. Quantitative PCR (qPCR) was performed using the CFX96 Thermal Cycler (Bio-Rad) with the following conditions: 95 °C for 5 min, 39 cycles of 95 °C for 10 s, and 60 °C for 30 s, followed by a melt starting at 65 °C for 5 s, increasing in 0.5 °C increments to 95 °C. Expression levels of the genes of interest were calculated relative to *TraesCS5A02G015600* following the $2^{-\Delta Cq}$ format (where ΔC_q = GOI quantification cycle (C_q) value − *TraesCS5A02G015600* C_q value). Statistical analysis of expression level was performed using either Student's t-test (normal distribution) or Mann-Whitney U test (non-normal distribution) when comparing two groups, and ANOVA followed by the Tukey post-hoc test for data with three or more groups to compare.

2.2.4.2 Preparation of wheat leaf tissue for RNA-seq

In addition to sampling leaf tissue and extracting RNA for qPCR analysis of genes of interest, RNA samples from leaf tissue during 10°C wRV, 14°C wRV, 10 °C SV, 10 °C SGV, and non-vernalized low-light 10 °C SGV control conditions were analysed using RNA-seq analysis. For all conditions, 4-weeks of treatment was selected for comparison. RNA was extracted following the method detailed in section **2.2.4.1**. Post-extraction, RNA was treated with RQ1 RNase-Free DNase (Promega) and Monarch® RNA Cleanup Kit (New England Biolabs) as an additional purification step. RNA samples were analysed by Novogene UK (Cambridge Science Park). Three biological replicates per condition were sent. Sample concentration was within the range of 40-60 ng/µL and a final sample volume of 40 µL.

Table 2.1: Complete list of primers used in this project.

Gene	Description	Sequence	Orientation
TraesCS5A02G015600	qPCR reference	TCTAAATGTCCAGGAAGCTGTTA	Sense
TraesCS5A02G015600	qPCR reference	CCTGTGGTGCCCAACTATT	Antisense
VRN1	Generic qPCR	GAACAAGATCAACCGGCAGGTGAC	Sense
VRN1	Generic qPCR	GGAGAAGATGATGAGGCCGACCTC	Antisense
ZCCT1 (VRN2)	Generic qPCR	GCCCACATCGTGCCATTTTACGGA	Sense
ZCCT1 (VRN2)	Generic qPCR	GCTCTCTCCTGCATTGTGGGATA	Antisense
ZCCT2 (VRN2)	Generic qPCR	CATCGTGCCATTCTGCGGG	Sense
ZCCT2 (VRN2)	Generic qPCR	CCCTGTACCTCATCACCTTCGCCT	Antisense
FT-B1	Specific qPCR	GTCGTTCGGGCAGGAG	Sense
FT-B1	Specific qPCR	TGGAAGAGTACGAGCACGA	Antisense
GIGANTEA	Generic qPCR	TTCATTTCTTGCGTGCGATT	Sense
GIGANTEA	Generic qPCR	CTTCAACTCCTTCAGCATGC	Antisense
PHYTOCHROME C	Generic qPCR	TCTCAGGTATGCTTGCGAAT	Sense
PHYTOCHROME C	Generic qPCR	GTAACACAATGCTGCACCAT	Antisense
PIF4	Generic qPCR	GGCACAGGGTCTAAATCAGA	Sense
PIF4	Generic qPCR	CCTCAGCTGAATGTTTTGCA	Antisense
CONSTANS1	Generic qPCR	CTTCCATCAGCAATGACATATC	Sense
CONSTANS1	Generic qPCR	GAAGTGAATGGCCTGAGAG	Antisense
TraesCS3A02G432900	Generic qPCR	TGCTGCCTGGTCTCTCGATA	Sense
TraesCS3A02G432900	Generic qPCR	TGCTCCCGCCACTATTTTGT	Antisense
TraesCS3A02G435000 *TraesCS3B02G470000* *TraesCS3D02G428000*	Generic qPCR	GGGCTCTTCAAGAAGGCCTT	Sense

Table 2.1: Complete list of primers used in this project (continued).

Gene	Description	Sequence	Orientation
TraesCS3A02G435000	Generic qPCR	GCAAATGCCTGATAGC	Antisense
TraesCS3B02G470000		GG TC	
TraesCS3D02G428000			
TraesCS3A02G435000	Generic qPCR	GCAAATGCCTGATAGC	Antisense
TraesCS3B02G470000		GG TC	
TraesCS3D02G428000			
TraesCS3A02G434400 *TraesCS3B02G469700* *TraesCS3D02G427700*	Generic qPCR	TCAAGAAGGCGTTCGA GC TC	Sense
TraesCS3A02G434400 *TraesCS3B02G469700* *TraesCS3D02G427700*	Generic qPCR	ACTTGCATCGCCTCCA TTC A	Antisense
TraesCS3A02G435000	Specific qPCR		Sense
TraesCS3A02G435000	Specific qPCR		Antisense
TraesCS3B02G470000	Specific qPCR	CAAGGCTTGAAGAGATT AC TA	Sense
TraesCS3B02G470000	Specific qPCR	TGCACCACTATTTCGTT GC G	Antisense
TraesCS3D02G428000	Specific qPCR		Sense
TraesCS3D02G428000	Specific qPCR		Antisense
TraesCS3A02G432900 downstream antisense	Specific qPCR	TTCCCCCAACACACTG TTC C	Sense
TraesCS3A02G432900 downstream antisense	Specific qPCR	ACGCTTTCGCTAAAATT TC CTCC	Antisense
TraesCS3A02G435000 downstream antisense	Specific qPCR	ACACTCTCTCTCCCTGA CA CC	Sense
TraesCS3A02G435000 downstream antisense	Specific qPCR	TTCTCCTCTCCTCCGCT TC TGT	Antisense
TraesCS3A02G432900 intron antisense	Specific qPCR	ACATCGTCGCCACAAA AAG C	Sense
TraesCS3A02G432900 intron antisense	Specific qPCR	GTCTGTACTTTGCCTC GGC T	Antisense

Table 2.1: Complete list of primers used in this project (continued).

Gene	Description	Sequence	Orientation
TraesCS3A02G432900	Yeast plasmid cloning	CGCGAATTCATGGCGC GG CG	Sense
TraesCS3A02G432900	Yeast plasmid cloning	GCGGATCCTCAAGCCT TC CTCCC	Antisense
TraesCS3A02G434400	Yeast plasmid cloning	CGCGAATTCATGGCGC GG CG	Sense
TraesCS3A02G434400	Yeast plasmid cloning	GCGGATCCTCAGGTCC TT CCTCG	Antisense
TraesCS3A02G435000 *TraesCS3B02G470000* *TraesCS3D02G428000*	Yeast plasmid cloning	CGCGAATTCATGGCGC GG CGCG	Sense
TraesCS3B02G470000	Yeast plasmid cloning	GCGGATCCTCAAGTCCT TC CTTC	Antisense
TraesCS3D02G428000	Yeast plasmid cloning	GCGGATCCTCAAGTCCT TC TG	Antisense
TraesCS3A02G435000	Yeast plasmid cloning	GCGGATCCTCAAGTCCT TC CTTCCTTC	Antisense
SPL-associated TFBS	Dimer formation	CGCGAATTCTGTTCGAA CA CGAGCTCGCG	Sense
SPL-associated TFBS	Dimer formation	CGCGAGCTCTGTTCGA AC ACGAATTCGCG	Antisense
pGAD424	Plasmid sequencing	AATACCACTACAATGGA T	Sense
pGBT9	Plasmid sequencing	GAGTAGTAACAAAGGTC AA	Sense
pHIS-2	Plasmid sequencing	CTAGACCGCGGATCGA TT CG	Sense

2.2.4.3 RNA-seq analysis and soft clustering

Abundance of transcripts was quantified using the 'kallisto' programme (https://github.com/pachterlab/kallisto; Bray et al. (2016)) and undertaken on ARC4, part of the High Performance Computing facilities at the University of Leeds, UK. Using kallisto, a transcriptome index was produced from the RNA-seq transcriptome output from Novogene. Pseudoalignments were generated using the IWGSC Chinese Spring reference genome (https://www.ebi.ac.uk/ena/browser/view/GCA_900519105.1; Appels et al. (2018)). Kallisto-quantified abundance of transcripts were reported in transcripts per kilobase million (TPM) and were analysed using the R package sleuth (https://github.com/pachterlab/sleuth; Pimentel et al. (2017)). Using sleuth, differential expression analysis was performed to identify genes that were differentially expressed between conditions. Differentially expressed genes were determined based on q-value < 0.05. Soft clustering of differentially expressed genes was carried out using the mfuzz package (https://github.com/a-velt/Mfuzz_RNAseq; Kumar and Futschik (2007)). An additional pre-processing step was carried out to standardise gene expression to have a mean value of zero and a standard deviation of one, to ensure that genes with similar changes in expression were close in Euclidean space. A fuzzifier ("m") was calculated to prevent clustering of random data and an optimal number of clusters ("c") was selected based on the rate of decline of the minimum distance ("dmin") of the cluster centroid. Soft clustering calculated the cluster centroid, which represents the overall expression pattern within clusters. All genes were assigned a membership value between 0-1 to determine which cluster best represents the expression change observed in the gene. Plots of clusters were visualized using the function 'mfuzz.plot2'. The ShinyGO v0.77 platform (http://bioinformatics.sdstate.edu/go/; Ge et al. (2020)) was used to perform gene ontology (GO) enrichment analysis on genes and identify enriched pathways that were overrepresented by the input gene list. Significant pathways were isolated by filtering by enrichment FDR and sorting by fold enrichment value. Dotplots representing GO terms were produced using the R packages 'ggplot2' and 'viridis' (https://github.com/sjmgarnier/viridis)

2.3 Results

2.3.1 Vernalization of European winter wheat is achieved under warmer temperatures

2.3.1.1 Winter wheat is successfully vernalized under temperatures of 10 °C and 14 °C

Winter wheat cultivars Claire, Buster, Charger, and Hereward had varying responses when grown under 8, 10, and 14 °C short-day (8 hr light/16 hr dark) vernalization conditions (**Figure 2.1**). Overall, 6 weeks of growth under these temperatures was the most effective to satisfy vernalization. Despite all cultivars flowering under 10 °C and 14 °C, there was a statistically significant ($p<0.05$; Student's t-test) difference in flowering time, with faster flowering observed after 6 weeks at 10 °C (**Figure 2.1B**). In addition, 8 weeks of growth under 10 °C and 14 °C was also sufficient for vernalization in all cultivars; flowering occurred faster under 10 °C compared to 14 °C (statistically significant at $p<0.05$) (**Figure 2.1B**; **Figure 2.1C**).

Two weeks of treatment was the least efficient method of vernalization; no cultivars flowered after receiving two weeks of treatment at 8 °C (**Figure 2.1A**), only cultivars Claire and Buster flowered at 10 °C, and a single cv. Claire plant flowered after two weeks at 14 °C. Although vernalization was completed after two weeks for 10 °C and 14 °C, plants flowered relatively late compared to other timepoints. Under 10 °C, cv. Buster flowered after 162 days of growth and cv. Claire flowered after 176 days (**Figure 2.1B**). At an increased temperature of 14 °C, cv. Claire flowered later after 192 days (**Figure 2.1C**). After 4-weeks of temperature treatment, only cv. Claire was vernalized under 8 °C and 10 °C, flowering after 132 and 159 days respectively (**Figure 2.1A**; **Figure 2.1B**). However, 4-weeks of treatment at 14 °C was sufficient to vernalize cultivars Claire, Buster, and Charger, with cv. Claire flowering first (134 days) followed by cv. Charger (162 days) and cv. Buster (163 days) (**Figure 2.1C**). Despite successful flowering of these cultivars at a warmer temperature, the total number of days between flowering of the first and final individual was variable (34, 15, and 22 days for Claire, Charger, and Buster, respectively). This would indicate that 4-weeks growth at 14 °C isn't the most efficient method of vernalization.

With this in mind, the most consistent duration of treatment under each temperature was 6-weeks of growth. At 8 °C, vernalization was satisfied in cultivars Claire, Charger, and Buster after 6 weeks, and flowering was recorded after 102 days for cultivars Claire and Charger and 124 days for cv. Buster (**Figure 2.1A**). Only a single plant for cv. Buster was vernalized after 6 weeks at 8 °C, however an increase to 10 °C led to flowering for all plants (**Figure 2.1B**). Furthermore, plants flowered only 4 days later at this increased temperature. Another difference observed after 6-weeks growth at 10 °C was the flowering of cv. Hereward, indicating vernalization was completed under these conditions for this cultivar (**Figure 2.1B**). Hereward flowered after 154 days of growth and the final plant flowered after 158 days, signifying the vernalization was completely saturated at this stage. Moreover, vernalization was completed in all cultivars after 6 weeks of growth at 14 °C (**Figure 2.1C**). The flowering response was slightly delayed at 14 °C compared to 10 °C.

Interestingly, increasing the duration and exposure to temperature to 8 weeks didn't necessarily result in an earlier flowering response. Under 8 °C, only a single plant for cultivars Claire and Charger were vernalized after 8 weeks growth (**Figure 2.1A**). Similarly, after 8 weeks growth at 10 °C, cultivars didn't flower earlier than after 6 weeks growth (**Figure 2.1B**). Here, the staggered flowering response was comparable to 6 weeks growth, with cv. Charger flowering first after 135 days, followed by cv. Buster (144 days), cv. Hereward (155 days), and finally cv. Claire (168 days). Since this flowering response range after 8 weeks (168-190 days) in cv. Claire overlaps with 2 weeks (176-179 days), vernalization after 8 weeks isn't more efficient than vernalization after 2 weeks at 10 °C for this cultivar. Similar to 6 weeks, vernalization was completed for all four cultivars after 8 weeks growth at 14 °C (**Figure 2.1C**). Cv. Charger flowering first after 146 days, then cultivars Buster and Claire after 160 days, then cv. Hereward after 174 days. Cultivars flowered at a relatively similar rate, with the range of flowering dates overlapping for all cultivars.

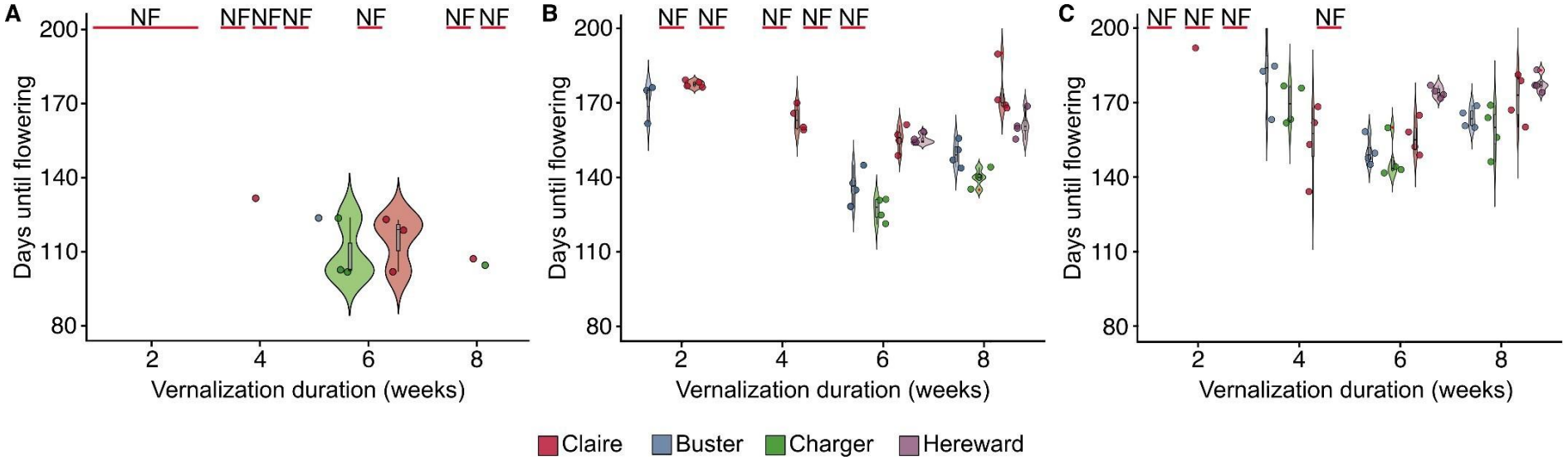


Figure 2.1: Heading of winter wheat cultivars under varying vernalization temperatures. Plants were grown under 8 hr light/16 hr dark for 8 weeks at temperatures of (A) 8 °C, (B) 10 °C, and (C) 14 °C. Wheat cultivars are ordered from low to high vernalization requirement as follows: cv. Claire (red), Buster (blue), Charger (green), and Hereward (purple). Vernalization duration represents the point at which plants were transferred from vernalization conditions into florally inducive (22 °C, 16 hr light/8 hr dark) conditions. Days until flowering = total days from germination until half-ear emergence. NF = no flowering after 200 days; n ≤ 4 plants.

2.3.1.2 Warmer temperature during vernalization corresponds with an increase in apex length

Since a variable flowering response was observed across cultivars and temperatures, rate of apex development was investigated as the stage of development is often used as an indicator of successful vernalization in winter cultivars. Therefore, apex development was recorded during growth under 8 °C, 10 °C, and 14 °C short-day (8 hr light/16 hr dark) conditions. From the length of apices across the timecourse, the rate of development generally corresponded with temperature, with the slowest rate of growth observed at 8 °C and the fastest rate at 14 °C (**Figure 2.2**). After 2 weeks of growth, there was no significant difference in apex length between temperatures for any of the cultivars; however, by 4 weeks of growth, differences between apex length were observed. The greatest difference in length after 4 weeks was reported in cv. Hereward, with the average apex length under 14 °C approximately 57% longer than under 8 °C (**Figure 2.2C**). In contrast, the least difference was recorded for cv. Charger, with an 8% increase in length at 14 °C compared to 8 °C (**Figure 2.2B**). However, it's noteworthy that cultivars Hereward and Charger had the smallest average apex length when grown at 10 °C (70.2 µm and 82.3 µm, respectively).

After 6 weeks of growth, differences in apex length across all cultivars became more apparent. When comparing average apex length under 8 °C and 14 °C, all cultivars exhibited an increase in length. The greatest percentage change was in cv. Charger, with an 174% increase in length under 14 °C, a substantial increase compared to the 8% increase observed at 4 weeks growth (**Figure 2.2B**). Cv. Buster had the smallest percentage change (52.7% difference in average apex length between 8 °C and 14 °C), however the average apex length after 6 weeks growth at 8 °C (227.7 µm) was longer than all other cultivars (**Figure 2.2A**).

The greatest average apex length overall occurred after 8 weeks growth at 14 °C, with cv. Claire having the highest average length, 684.8 µm, an increase of 121% compared to growth under 8 °C (**Figure 2.2D**). Other cultivars were as follows: cv. Buster, 593.33 µm (106% increase); Charger, 489.7 µm (120.9%); Hereward, 476.01 µm (119.5%). Surprisingly, this increase in apex length didn't correlate with a transition from vegetative to reproductive tissue. This was deduced based on the absence of double-ridge formation in the apices (**Figure 2.3**). When grown under 10 °C, after 2- and 4-weeks of growth, apices for all cultivars were at the early vegetative stage of development. By 6 weeks, apices of cultivars Buster and Claire had progressed to late vegetative stage, but still hadn't transitioned to the double ridge stage that indicates the generation of reproductive apical tissue (**Figure 2.3A**; **Figure 2.3D**). By 8 weeks, cultivars Charger and Hereward had also reached the late vegetative stage. Intriguingly, cv. Claire had an elongated apex by 8 weeks, but lacked a double ridge and was still only forming leaf primordia along the apex (**Figure 2.3D**). Under growth at 14 °C, the absence of double ridge was more pronounced. Here, all cultivars had reached the late vegetative stage by 6 weeks of growth (**Figure 2.3E- Figure 2.3H**). By 8 weeks, elongated apices were observed for all cultivars, with the most distinctive elongation reported in cv. Claire. Similar to growth under 10 °C, the elongated apices were still in the late vegetative stage of development (**Figure 2.3D**; **Figure 2.3H**). This would indicate that these apices are still in the process of transitioning to reproductive tissue formation even after 8 weeks. This is compelling as the flowering data suggests that vernalization was complete before this point.

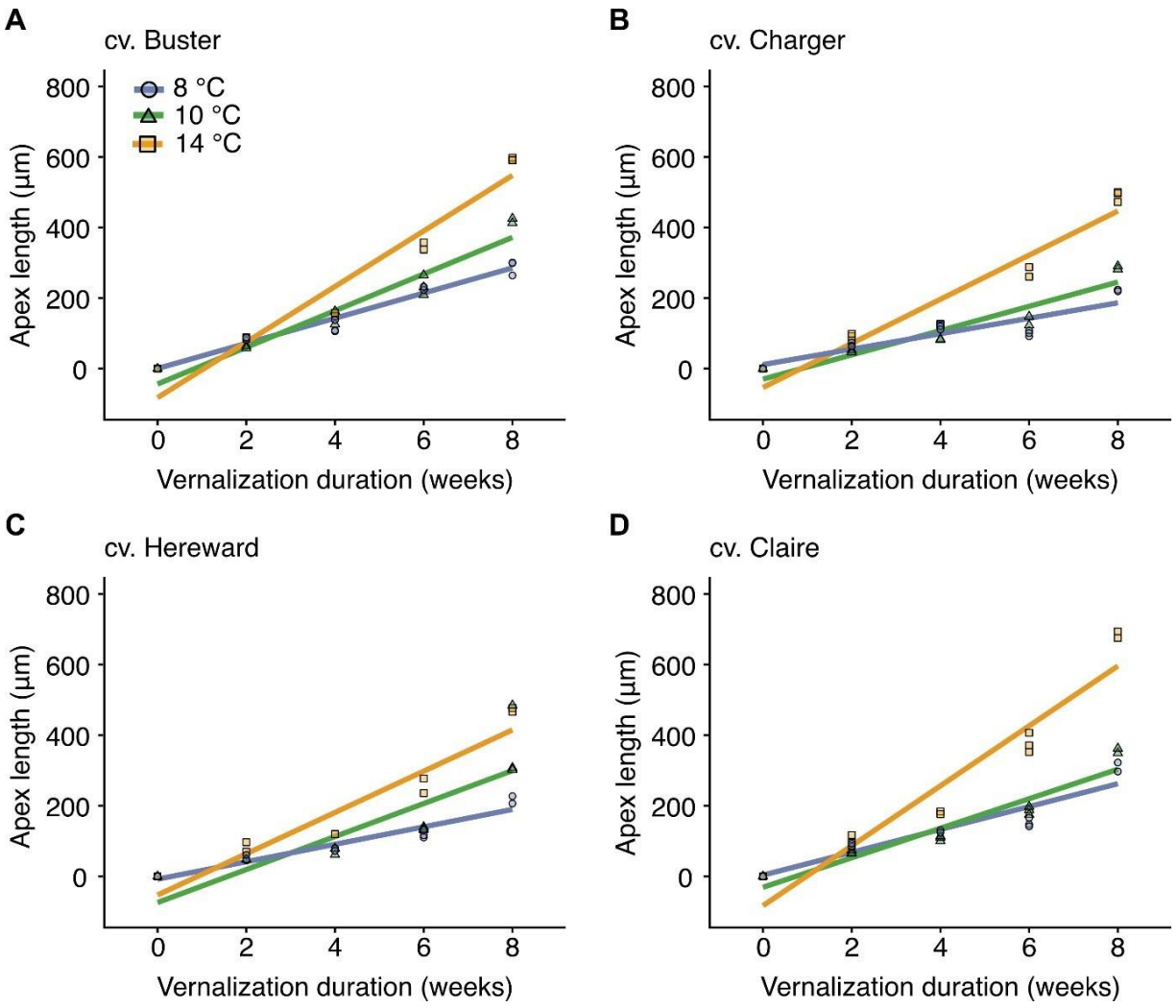


Figure 2.2: Development of winter wheat apices under varying vernalization temperatures. Apex length (µm) was measured at 2-week intervals for 8 weeks in cultivars (A) Buster, (B) Charger, (C) Hereward, and (D) Claire. Plants were grown under 8 hr light/16 hr dark for 8 weeks at temperatures of 8 °C (blue circle), 10 °C (green triangle), and 14 °C (orange square). Coloured lines represent fitted linear regression lines.

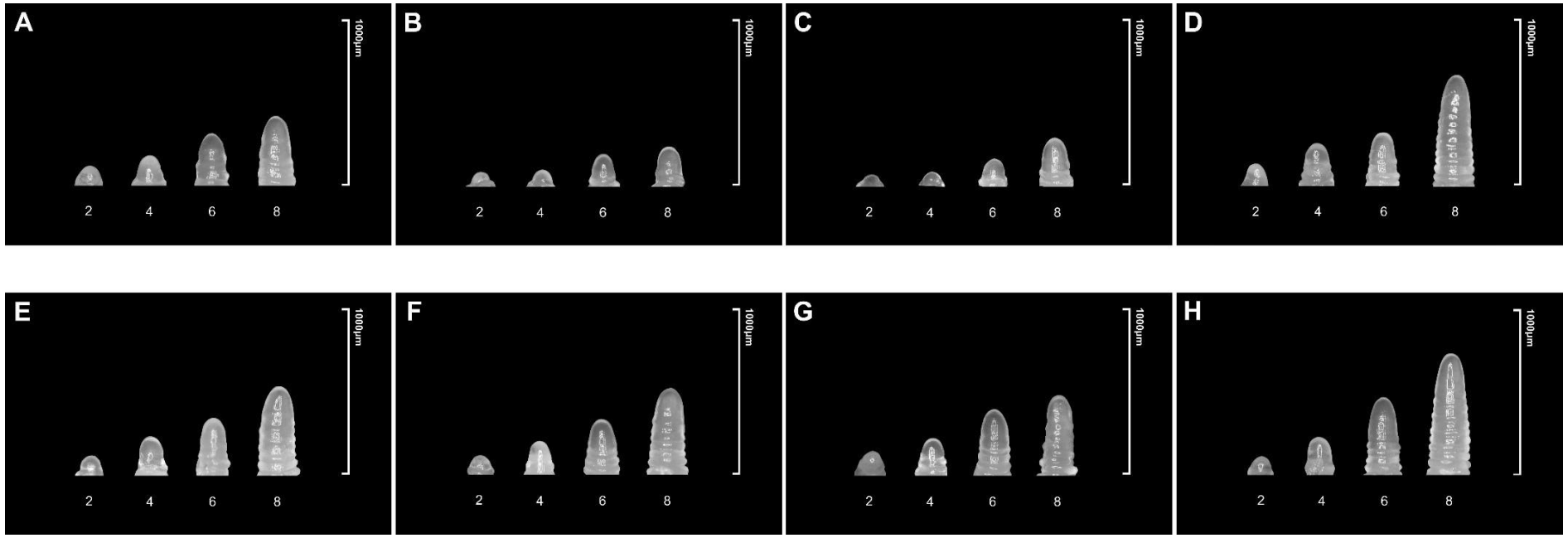


Figure 2.3: Apex development of winter wheat over 8 weeks of growth under varying vernalization temperatures. (A-D) Growth under 10 °C short day conditions; (E-H) Growth under 14 °C short day conditions. Numbers beneath apices signify duration of vernalizing conditions. Cultivars are (A,E) Buster, (B,F) Charger, (C,G) Hereward, and (D,H) Claire. A scale bar (1000 µm) is included to demonstrate scale.

2.3.1.3 Key wheat vernalization genes are active under warmer vernalization temperatures

To investigate the molecular mechanism underpinning the vernalization under warmer temperatures, expression analysis was performed on leaf tissue from plants grown under 8 °C, 10 °C, and 14 °C short day conditions. First, expression of floral activator *VRN1* was examined in cultivars Claire and Charger (**Figure 2.4A**-**Figure 2.4B**). As the 4-week and 6-week timepoint marked the completion of vernalization for these two cultivars, *VRN1* expression after 4 and 6 weeks was investigated. At a lower temperature of 8 °C, there was no difference in expression level of *VRN1* after 4 weeks between cultivars Claire and Charger. Similarly, under 10 °C, and 14 °C, there was no significant difference in *VRN1* in cv. Charger, although expression was higher under 10 °C (**Figure 2.4B**). An increased temperature of 14 °C corresponded with reduced expression of *VRN1* for both cultivars.

At 6 weeks, *VRN1* expression under 8 °C was significantly higher ($p<0.05$; Tukey post-hoc test) than 10 °C and 14 °C for cv. Claire. In contrast, cv. Charger had high *VRN1* level under both 8 °C and 10 °C, with the highest expression level at 10 °C. Expression levels for both these temperatures were significantly higher than at 14 °C (**Figure 2.4**). When comparing cultivars, expression level of *VRN1* under 10 °C was significantly higher for cv. Charger at both 4 and 6 weeks.

Repression of the *VRN2* locus by *VRN1* is understood to facilitate activation of the vernalization pathway (Yan et al., 2004; Yan et al., 2003). Therefore, *VRN2* expression (i.e., cumulative expression of genes *ZCCT-A1*, *ZCCT-B1*, *ZCCT-D1*, *ZCCT-A2*, and *ZCCT-D2*) under these same conditions in cultivars Claire and Charger was similarly examined. In cv. Claire, there was no difference in *VRN2* expression between 4 and 6 weeks at 8 °C. At an increased temperature of 10 °C, expression of *VRN2* was completely repressed by 4 weeks and remained off after 6 weeks (**Figure 2.4C**). Under a temperature of 14 °C, *VRN2* expression was higher at 4 weeks and down regulated to a significant degree ($p<0.05$; Student's t-test) by 6 weeks in cv. Claire. However, in cv. Charger, *VRN2* was down regulated at the 4- week

timepoint across all 3 temperatures (**Figure 2.4D**).

By 6 weeks of growth at 8 °C, *VRN2* was sharply up regulated compared to 4 weeks; however, this result was not statistically significant ($p>0.05$; ANOVA). Nonetheless, increased expression level of *VRN2* at 8 °C aligned with later flowering under this temperature; this is compelling as it indicates that down regulation of *VRN2* occurs at a higher temperature, leading to vernalization being completed.

Expression analysis for *VRN1* and *VRN2* follows a typical vernalization route under 8 °C, but the pattern of expression for these genes deviates under warmer temperatures. Furthermore, the low level of expression of floral activator *VRN1* at 14 °C doesn't explain the vernalization and subsequent flowering response observed under these conditions. Therefore, it was necessary to investigate other candidate genes that could be involved in vernalization under a warmer temperature.

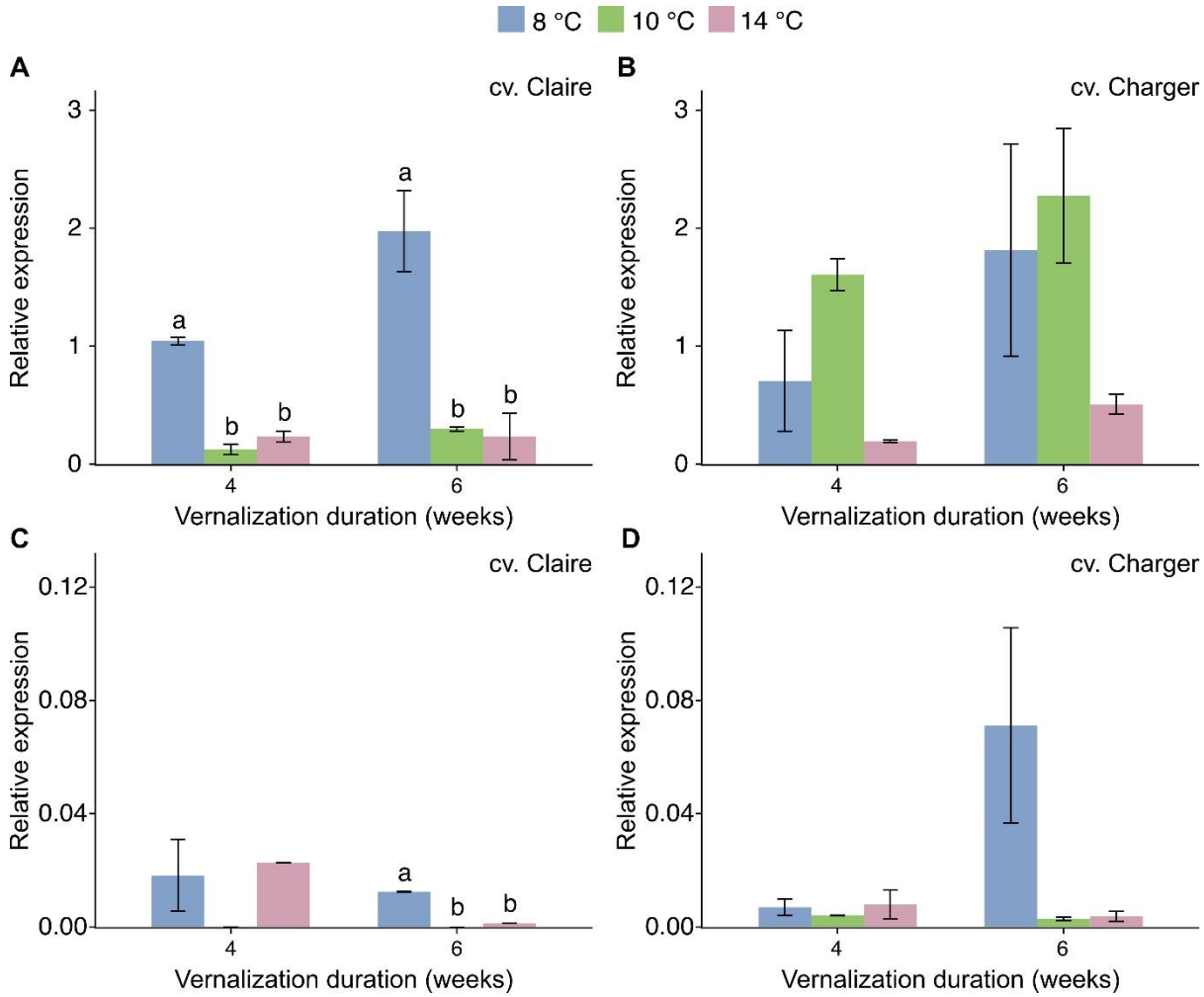


Figure 2.4: Expression of *VRN1* and *VRN2* in winter wheat under varying vernalization temperatures. (**A**-**B**) *VRN1* expression in cv. Claire and cv. Charger under 8 °C (blue), 10 °C (green), and 14 °C (purple) short day conditions; (**C**-**D**) *VRN2* expression in cv. Claire and cv. Charger under 8 °C (blue), 10 °C (green), and 14 °C (purple) short day conditions. Plants were exposed to the conditions for 4 and 6 weeks. Error bars = standard error of the mean for 3 biological replicates. Different letters above bars represent statistically significant difference ($p<0.05$; ANOVA and Tukey post-hoc test) in expression between temperature conditions at the respective timepoint.

2.3.1.4 RNA-seq reveals contrast in differential gene expression

In the previous section, the results of flowering indicated that 4 weeks of vernalization at 14 °C effectively vernalized cultivars Buster and Charger, whereas these cultivars were yet to be vernalized after 4 weeks at 10 °C (section **2.3.1.1; Figure 2.1**). Therefore, the 4-week timepoint was investigated in more depth to determine why vernalization proceeded at a higher temperature instead of a lower temperature in winter cultivars that have a moderate and high vernalization requirement. To do this, differential expression analysis was performed on leaf tissue from cv. Buster grown for 4 weeks under either 10 °C or 14 °C short day conditions.

Sleuth analysis of the RNA-seq dataset identified 3,279 transcripts and 3,239 genes that were being differentially expressed between a temperature of 10 °C and 14 °C. In total, 1,502 transcripts were up regulated and 1,777 were down regulated between the different temperatures (**Figure 2.5**). Four genes associated with vernalization were differentially expressed; these were *VRN-A1*, *VRT-A2*, *VRT-B2*, and *VRT-D2*. These genes were all down regulated under 14 °C short day compared to 10 °C short day (**Figure 2.5**). Under 10 °C, *VRN-A1* had an expression level of 4.729 TPM compared to 0.764 TPM under 14 °C (**Figure 2.6**). Expression level varied when comparing the homoeologues of *VRT2*; the highest level of expression was observed in *VRT-D2* under 10 °C (42.144 TPM) and 14 °C (3.175 TPM), followed by *VRT-A2* (20.969 and 3.098 TPM), and *VRT-B2* (1.032 and 0.147 TPM) (**Figure 2.6**). None of the *ZCCT* genes or *FT-B1* were differentially expressed between temperatures. Furthermore, *FT-B1* and *ZCCT1* (*ZCCT-A1*, *ZCCT-B1*, and *ZCCT-D1*) were not expressed under either temperature (**Figure 2.6**). *ZCCT2* had a low level of expression and there was no significant difference in expression of *ZCCT-A2* (0.127 TPM at 10 °C; 0.036 TPM at 14 °C) or *ZCCT-D2* (0.141 TPM at 10 °C; 0.000 TPM at 14 °C) in response to temperature.

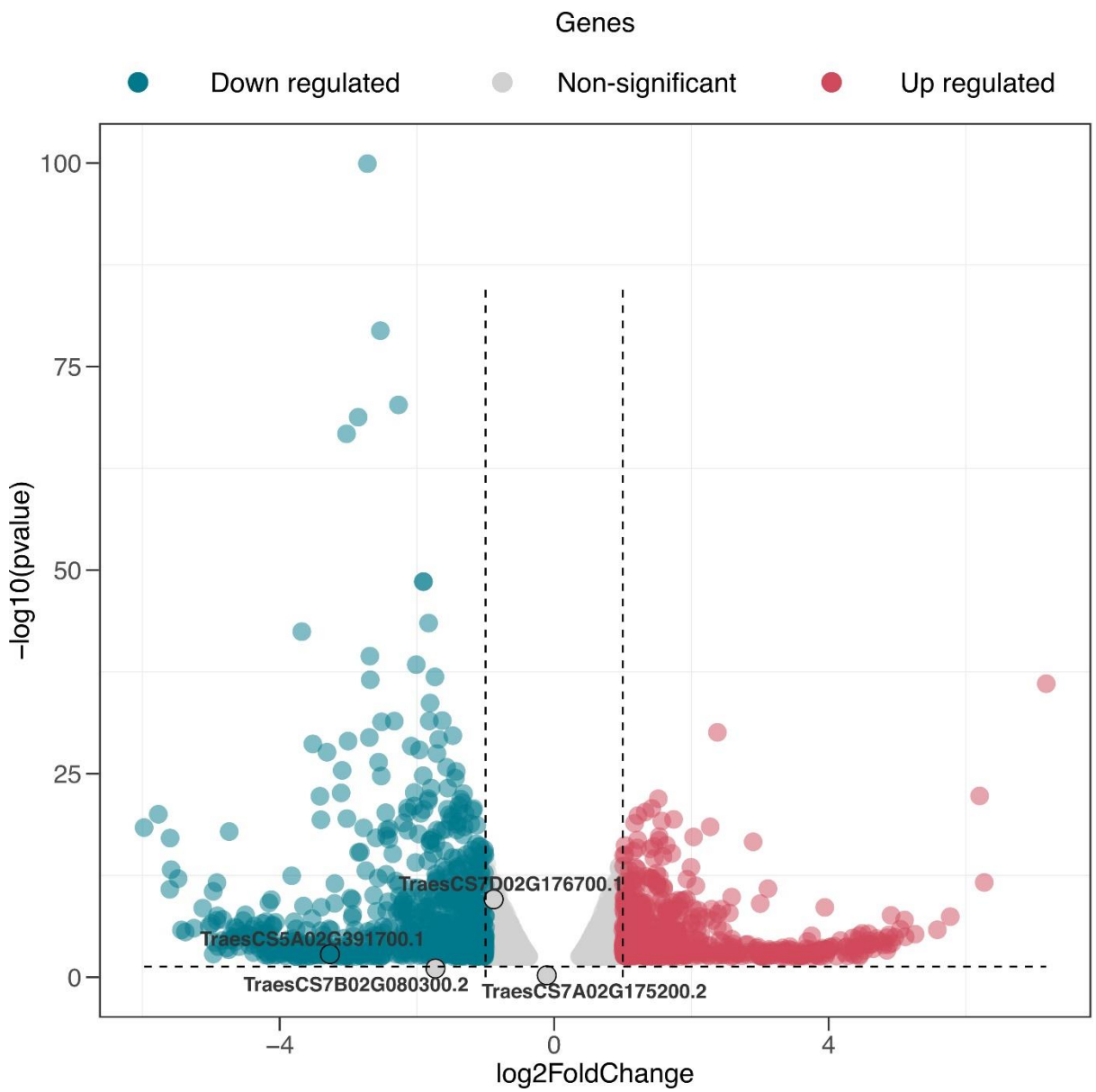


Figure 2.5: Volcano plot displaying differentially expressed genes in cv. Buster after 4 weeks of growth under 10 °C and 14 °C short day conditions. Differentially expressed genes were determined from comparing RNA-seq data for 10 °C vs 14 °C. Vernalization genes are annotated as follows: *TraesCS5A02G391700* (*VRN-A1*), *TraesCS7A02G175200* (*VRT-A2*), *TraesCS7B02G080300* (*VRT-B2*), and *TraesCS7D02G176700* (*VRT-D2*). Down regulated genes = blue; up regulated genes = red; non-significant genes = grey.

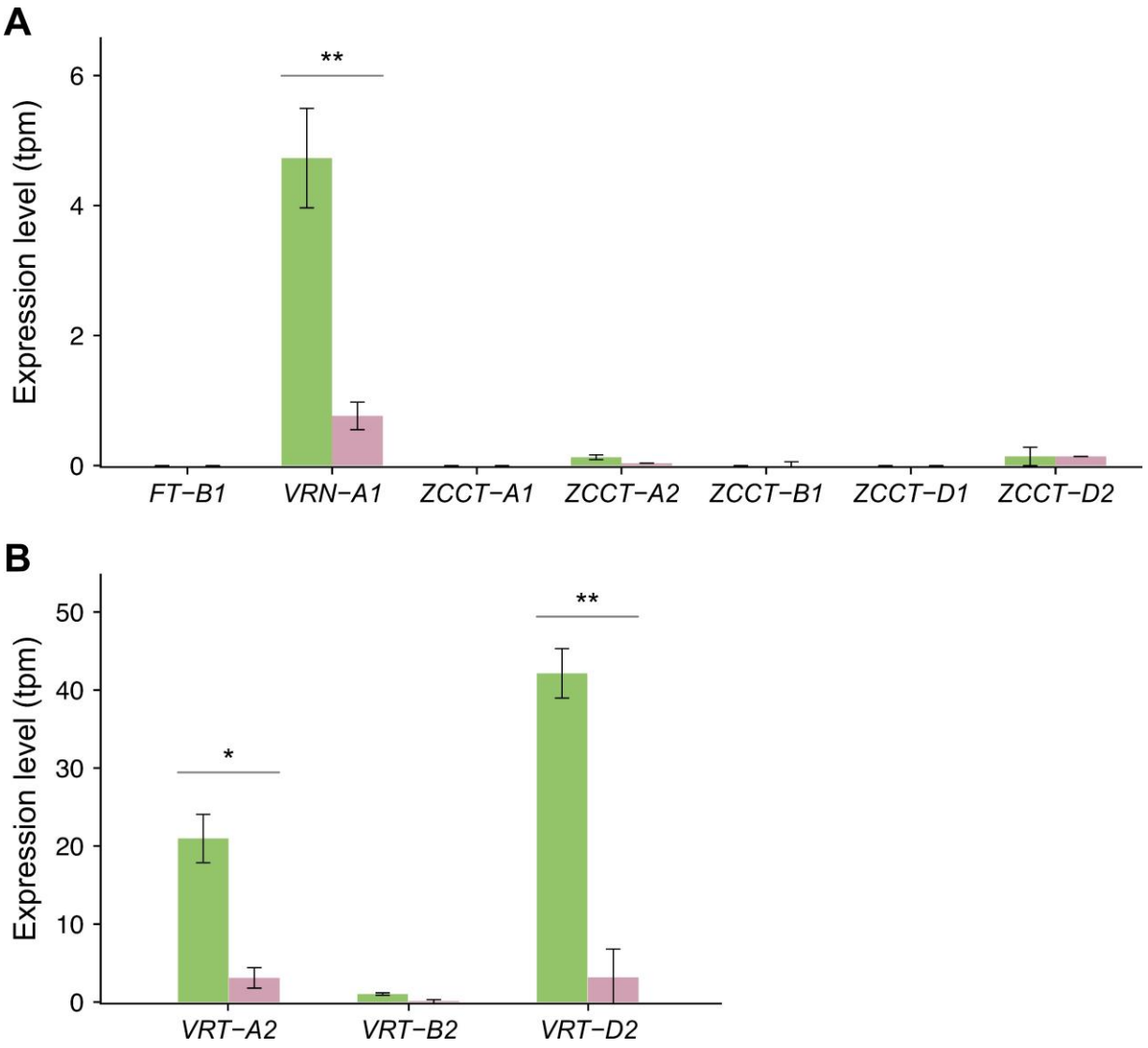


Figure 2.6: Expression of vernalization associated genes in cv. Buster after 4 weeks of 10 °C and 14 °C short day conditions. (**A**) Floral activator genes *FT-B1* and *VRN-A1*, and floral repressor *VRN2* – i.e., *ZCCT-1, ZCCT-A2, ZCCT-B1, ZCCT-D1,* and *ZCCT-D2*. (**B**) *VRT2* genes. Green = 10 °C; purple = 14 °C. * = p<0.05; ** = p<0.01.

Since there was no difference in the expression level of central floral repressor *ZCCT* genes at 10 °C and 14 °C, I considered the involvement of another floral repressor that was supressing flowering in cv. Buster under 10 °C conditions. To identify potential candidates, gene ontology (GO) enrichment analysis was carried out on all differentially expressed genes. This analysis identified the top 19 GO Biological Process pathways associated with the 3,279 differentially expressed transcripts (**Table 2.2**). Within the top pathways, 3 were particularly relevant: response to cold (fold enrichment = 4.3), response to temperature stimulus (fold enrichment = 4), and response to abiotic stimulus (fold enrichment = 2.8) pathways.

The response to cold pathway contained 21 differentially expressed transcripts, of which 14 genes were down regulated and 7 were up regulated at 14 °C. All of these transcripts were also identified in the 'response to temperature stimulus' pathway; in total, this pathway group contained 39 differentially expressed transcripts, including 18 transcripts unique to the response to temperature stimulus pathway. Similarly, the response to abiotic stimulus pathway contained all transcripts from the response to cold and response to temperature stimulus pathways. In addition, this pathway also contained 60 other differentially expressed transcripts not directly related to temperature. Since all the transcripts identified in the response to cold pathway also were present in the other two relevant pathways, this group was examined in more detail.

2.3.2 Vernalization in European winter wheat is achieved under an extended day length

Results from section **2.3.1** highlights that increased temperatures of 10 °C and 14 °C are sufficient for vernalization to proceed, with successful vernalization of winter cultivars Claire, Buster, Charger, and Hereward achieved by 6 weeks of growth under these conditions. These findings are consistent with evidence from Arabidopsis in which vernalization proceeds under warmer autumn temperatures (Duncan et al., 2015; Hepworth et al., 2018). As this suggests that vernalization under artificial temperatures of 4 °C-6 °C might not be the most efficient approach, we investigated if adjusting the photoperiod from the standard 8 hr light/16 hr dark would further optimise vernalization. Since section **2.3.1** indicated that vernalization proceeded more efficiently under 10 °C compared to 14 °C, this

temperature was selected as a starting point. Typically, a short-day photoperiod is used for artificial vernalization of cereals, as this reduces expression of the floral repressive locus *VRN2*. However, the results of expression analysis in section **2.3.1.4** revealed that the *ZCCT* genes that represent the *VRN2* locus were completely repressed in cv. Buster after 4 weeks growth at 10 °C, despite the vernalization requirement yet to be met. In conjunction with examples that suggest that repression of flowering pre-vernalization is a multigenic response (Greenup et al., 2010; Xie et al., 2021), we considered that *VRN2* is not a limiting factor for vernalization to proceed. To investigate this, the vernalization response was examined in cultivars Claire, Charger, and Hereward grown at 10 °C under an extended photoperiod of 22 hr light/2 hr dark and increased light intensity (**Figure 2.7A**; **Figure 2.7C**). From here onwards, these conditions will be referred to as speed vernalization (SV) (**Figure 2.7B**). For comparison, these cultivars were also grown under 10 °C short day conditions and increased light intensity, referred to as warm regular vernalization (wRV). To identify when vernalization was completed, plants were moved into speed breeding conditions (22 °C, 22 hr light/17 °C, 2 hr dark) and flowering date was recorded (**Figure 2.7D**).

Table 2.2: Enriched biological pathways in winter wheat cv. Buster under 10 °C and 14 °C vernalization temperatures. DEGs = differentially expressed genes.

Enrichment FDR	Fold Enrichment	DEGs	Total Pathway Genes	Pathway
1.67E-14	6.917323	30	144	Chloroplast organization
3.49E-08	6.06596	19	104	Glycosyl compound metabolic proc.
1.75E-14	5.941616	34	190	Plastid organization
5.54E-13	5.142081	35	226	Sulfur compound biosynthetic proc.
5.65E-08	5.037719	22	145	Pyrimidine-containing compound metabolic proc.
6.77E-07	4.812051	20	138	Pyrimidine-containing compound biosynthetic proc.
2.55E-06	4.251623	21	164	Response to cold
2.17E-12	3.969942	44	368	Protein folding
5.58E-11	3.960009	39	327	Response to temperature stimulus
3.93E-16	3.210483	76	786	Sulfur compound metabolic proc.
1.84E-11	3.201732	54	560	Alpha-amino acid metabolic proc.
2.00E-18	2.944265	101	1139	Small molecule biosynthetic proc.

1.11E-12	2.835856	72	843	Carboxylic acid biosynthetic proc.
3.73E-17	2.814308	99	1168	Response to abiotic stimulus
4.20E-07	2.705442	44	540	Response to inorganic substance
2.00E-18	2.396795	143	1981	Oxoacid metabolic proc.
2.27E-17	2.354713	139	1960	Carboxylic acid metabolic proc.
4.39E-09	2.314346	75	1076	Cellular amino acid metabolic proc.
3.20E-06	2.285412	53	770	Monocarboxylic acid metabolic proc.

2.3.3 Winter wheat is satisfactorily vernalized using SV conditions

Under an extended photoperiod of 22 hr light at 10 °C, vernalization was successfully completed in cultivars Claire, Charger, and Hereward. Comparative to vernalization under 10 °C short day conditions, vernalization was not fully saturated after 2 weeks. For cv. Claire, plants that received 2 weeks of SV flowered after 82 days, compared to 118 days for cv. Charger (**Figure 2.8**). This was earlier than for plants that received 2 weeks of wRV for both cultivars. In contrast, 2 weeks of SV conditions was not sufficient for vernalization to proceed in cv. Hereward (**Figure 2.10**). However, after 4 weeks of SV, vernalization was satisfied in cv. Hereward, with uniform flowering after 100 days (**Figure 2.10**). Similarly, cultivars Claire and Charger were vernalized after 4 weeks of SV. Claire flowered after 65 days under these conditions, but there wasn't a statistically significant difference in flowering time when compared to wRV (**Figure 2.8**). Conversely, vernalization was completed in Charger (**Figure 2.9**) and Hereward (**Figure 2.10**) under SV but not wRV conditions.

Interestingly, 4 weeks of SV satisfied vernalization for all cultivars with no statistically significant difference in flowering time between 4 and 6 weeks of SV (**Figure 2.8**; **Figure 2.9**; **Figure 2.10**). Flowering of cultivars Claire and Hereward after 6 weeks of SV was not significantly different to 6 weeks of wRV. Furthermore, cv. Charger flowered significantly later ($p<0.01$; Student's t-test) after 6 weeks SV compared to 6 weeks wRV (**Figure 2.9**). Based on total spikelet number per ear, there was a statistically significant difference in spikelet number when comparing wRV and SV (**Figure 2.11**), with SV resulting in decreased spikelet number (**Figure 2.11A**; **Figure 2.11B**). No statistically significant difference in spikelet number was reported under SGV conditions in cv. Charger (**Figure 2.11C**). For cv. Hereward; there was no statistically significant difference in spikelet number under SGV compared to either wRV or SV (**Figure 2.11F**). Conversely, a statistically significant difference ($p<0.01$; Mann- Whitney U test) was reported between 6 weeks wRV and SV, with increased spikelet number under wRV conditions (**Figure 2.11D**; **Figure 2.11E**). For cv. Claire, there was a statistically significant difference when comparing 4 weeks of wRV and SV ($p<0.01$; Student's t-test) (**Figure 2.11G**; **Figure 2.11H**). Additionally, a statistically significant difference was observed when comparing 6 weeks of wRV and SV ($p<0.01$; Mann-Whitney U test).

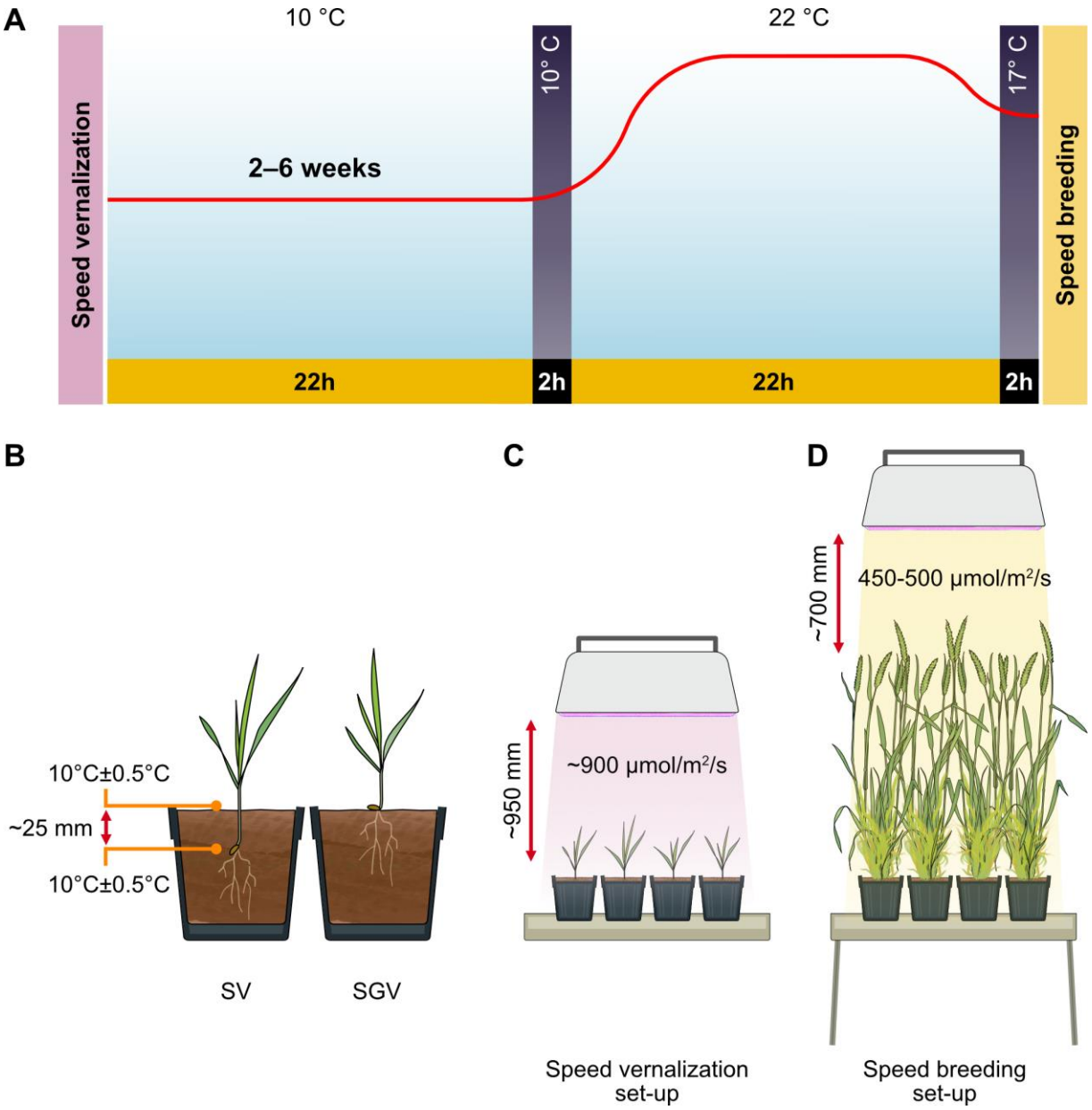


Figure 2.7: Schematic of the speed vernalization protocol. (**A**) Schematic combining speed vernalization (either SV or SGV) with SB to maximise effects of these protocols; (**B**) Illustration depicting speed vernalization set up with the seed covered (SV) and uncovered (SGV). For SV, temperature was measured both above and below the soil surface; (**C**) Illustration showing plants growing under high light intensity during speed vernalization in a growth cabinet; (**D**) Illustration showing plants flowering in speed breeding post-speed vernalization in a glasshouse.

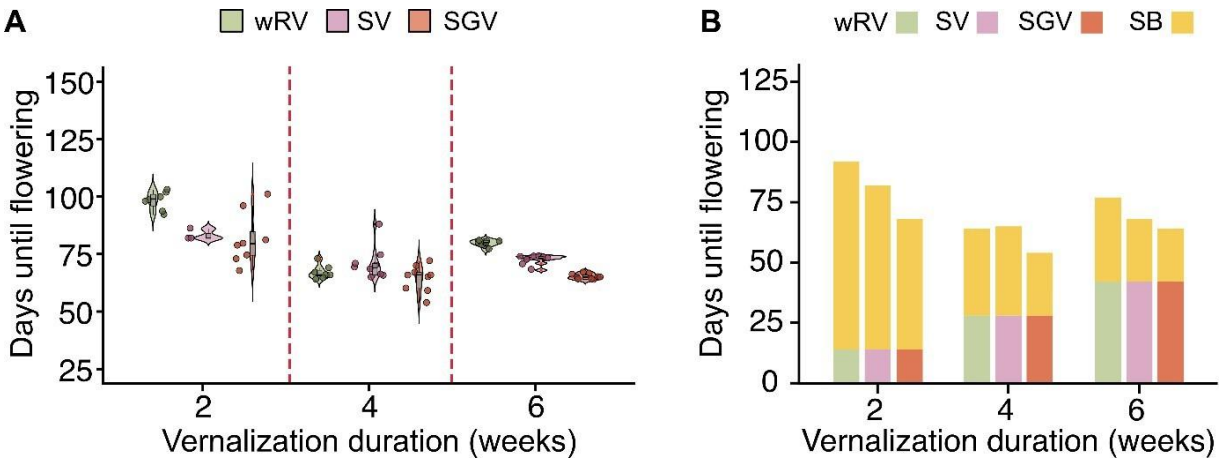


Figure 2.8: Heading of winter wheat cv. Claire under the different vernalization conditions. (**A**) Days until flowering from germination after 2, 4, and 6 weeks of vernalization type; warm regular vernalization (wRV, green), speed vernalization (SV, purple), speed green vernalization (SGV, red). (**B**) Flowering of the first individual plant, including total days of vernalization treatment and total days of speed breeding (SB, yellow) post-vernalization treatment.

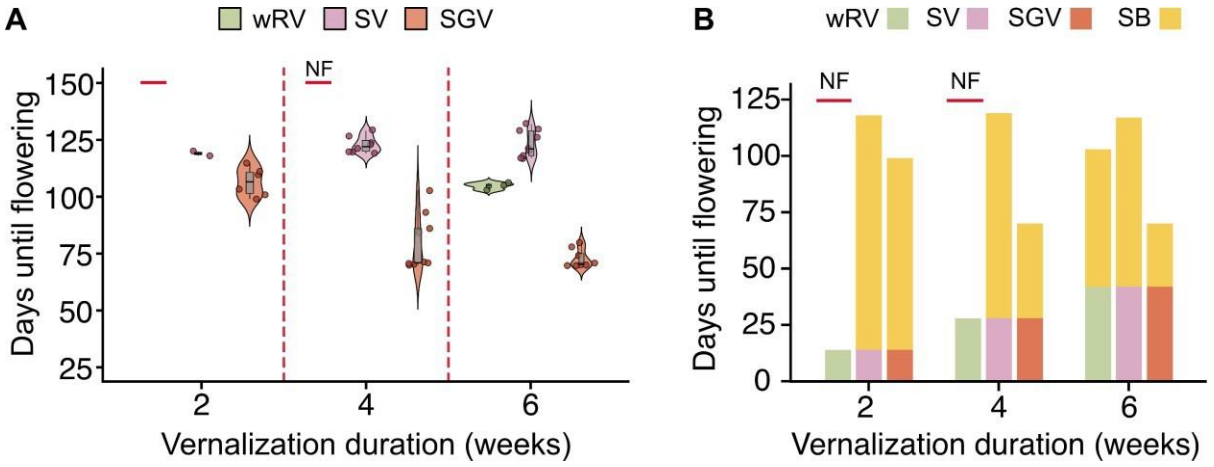


Figure 2.9: Heading of winter wheat cv. Charger under the different vernalization conditions. (**A**) Days until flowering from germination after 2, 4, and 6 weeks of vernalization type; warm regular vernalization (wRV, green), speed vernalization (SV, purple), speed green vernalization (SGV, red). (**B**) Flowering of the first individual plant, including total days of vernalization treatment and total days of speed breeding (SB, yellow) post-vernalization treatment.

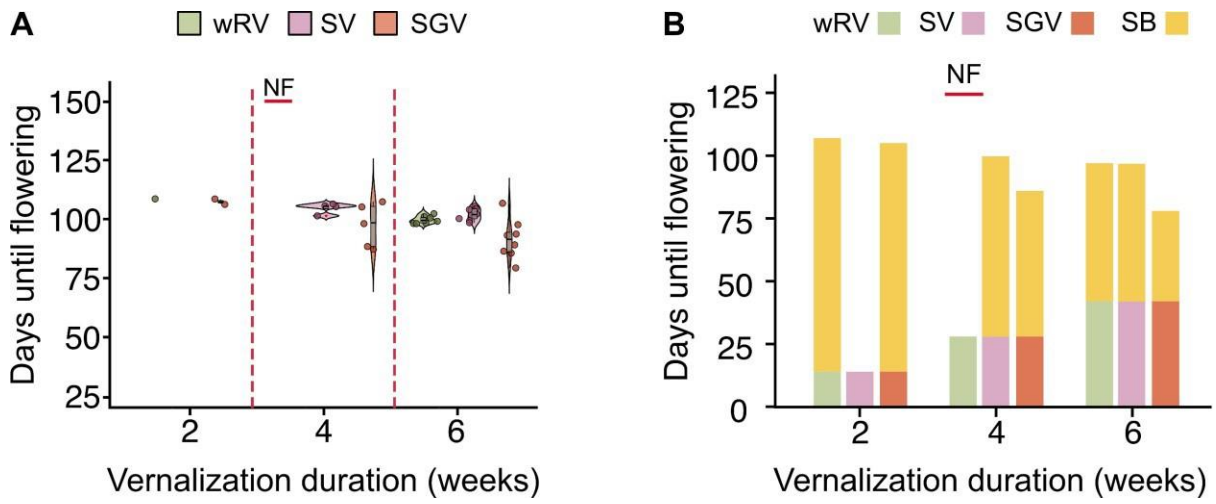


Figure 2.10: Heading of winter wheat cv. Hereward under the different vernalization conditions. (**A**) Days until flowering from germination after 2, 4, and 6 weeks of vernalization type; warm regular vernalization (wRV, green), speed vernalization (SV, purple), speed green vernalization (SGV, red). (**B**) Flowering of the first individual plant, including total days of vernalization treatment and total days of speed breeding (SB, yellow) post-vernalization treatment.

2.3.4 Expression analysis indicates complexity of SV process

Surprisingly, *VRN1* expression wasn't considerably up regulated over the course of SV. Here, expression of *VRN1* was examined as the difference in vernalization for the winter cultivars tested was partially attributed to this gene. Using cultivars Claire and Charger to investigate *VRN1* expression across 2- to 6-weeks SV, expression remained low (**Figure 2.12**). After 2, 4, and 6 weeks of SV, there was no significant difference in *VRN-A1* when comparing cultivars Claire and Charger.

2.3.5 Vernalization on the surface (SGV) is an effective method of vernalization

Speed vernalization adequately vernalized cultivars Claire, Charger, and Hereward, as demonstrated in section **2.4.1**. However, we considered whether we could optimise the protocol further to reduce generation time and develop our understanding of the pathway active during speed vernalization. Expression analysis under SV conditions revealed that genes known to be involved in cereal vernalization, including *VRN1*, *ZCCT* genes, and *FT1*, were not behaving as expected. This indicates a separate pathway or element is involved in SV. Germinating seeds uncovered on the surface of the soil resulted in successful vernalization of the tested winter wheat cultivars Claire, Charger, and Hereward. Furthermore, this form vernalization (referred to as 'SGV' from here onwards) resulted in earlier flowering when compared to SV and wRV (**Figure 2.8**; **Figure 2.9**; **Figure 2.10**). Prior to flowering, plants grown under SGV differed in appearance when compared to those grown under SV (**Figure 2.13**). When grown under SGV, plants had a reduction in plant height and leaf length when compared to SV (**Figure 2.13**). Nonetheless, SGV sufficiently vernalized plants grown under these conditions. This response varied across timepoints, with 2 weeks of SGV observed to be the least effective duration of vernalization for all cultivars. For cv. Claire, there was a statistically significant difference ($p<0.05$; Mann-Whitney U test) in flowering time between 2 weeks of wRV and SGV, with plants grown under SGV flowering after 68 days compared to 94 days for wRV. In contrast, there was no difference in flowering between wRV and SGV after 4 weeks for cv. Claire. By 6 weeks, vernalization was fully saturated in cv. Claire, similar to what was observed under wRV and SV, indicating that 6 weeks is more than a sufficient amount of time for vernalization to be completed for this cultivar. Furthermore, there was a

statistically significant difference ($p<0.01$; Mann-Whitney U test) between SGV and wRV at this timepoint, with SGV also resulting in earlier flowering (**Figure 2.8**).

In cv. Charger, the difference in flowering time between wRV and SGV was more pronounced (**Figure 2.9**). Plants were vernalized after 2 weeks of SGV, with the first plant flowering after 99 days. This was significantly earlier than under wRV, where vernalization wasn't satisfied. Moreover, after 4 weeks vernalization wasn't achieved under wRV conditions, and flowering under SV was significantly slower ($p<0.01$; Mann-Whitney U test) than under SGV. Vernalization was almost completed saturated by 6 weeks; here, plants flowered between 70 and 80 days and this result was significantly different when compared to SV ($p<0.01$; Mann-Whitney U test) and wRV ($p<0.01$; Mann-Whitney U test).

A similar result was observed for cv. Hereward (**Figure 2.10**). After 2 weeks of SGV, vernalization was completed; however, compared to cultivars Claire and Charger, this duration wasn't as effective at achieving vernalization, as only two plants went on to flower under these conditions. Nonetheless, longer duration of SGV resulted in a higher number of vernalized plants, with 4 weeks of SGV leading to flowering after 86 days (**Figure 2.10**). In contrast, cv. Hereward wasn't successfully vernalized after 4 weeks of wRV, with no flowering after 120 days growth. After 6 weeks of SGV, plants flowered after 84 days. There was no statistically significant difference in flowering when comparing 4 weeks and 6 weeks of SGV; however, flowering was slightly more uniform after 6 weeks (**Figure 2.10**). Conversely, there was a statistically significant difference ($p<0.05$; Student's t-test) in flowering between SGV and wRV after 6 weeks, with SGV resulting in earlier flowering. When comparing SV and SGV, there was also a statistically significant difference ($p<0.01$; Student's t- test).

In addition to flowering, total spikelet number per ear was also investigated under SGV conditions. For cv. Claire, there was no difference in total spikelet number per ear when comparing wRV and SGV conditions after 2 weeks growth (**Figure 2.11G**; **Figure 2.11I**). However, spikelet number was lower when plants were grown for 4 weeks under SGV conditions compared to wRV. This result was statistically significant ($p<0.01$; Mann-Whitney U test). Similarly, a statistically significant difference ($p<0.01$; Mann-Whitney U test) was also observed after 6 weeks, with fewer spikelets under SGV conditions. There was a statistically significant difference

($p<0.01$; Mann-Whitney U test) between spikelet number after 6 weeks of wRV compared to SGV for cultivar Charger, with reduced spikelets under SGV conditions (**Figure 2.11A**; **Figure 2.11C**). When comparing spikelet number for SV and SGV, there was also a statistically significant difference ($p<0.01$; Mann-Whitney U test) at 6 weeks. For cv. Hereward, there was a statistically significant difference ($p<0.01$; Mann-Whitney U test) between spikelet number after 4 weeks of SV compared to SGV; here, higher spikelet number was observed under SGV conditions compared to SV (**Figure 2.11D**; **Figure 2.11F**). However, a statistically significant difference ($p<0.01$; Mann- Whitney U test) was observed for cv. Hereward between spikelet number across vernalization conditions, with less spikelets produced under SGV conditions compared to wRV. When comparing spikelet number at 6 weeks of SV and SGV, a statistically significant difference ($p<0.01$; Mann-Whitney U test) was also reported, with a wider range of total spikelet number per ear observed under SGV conditions (**Figure 2.11D**; **Figure 2.11F**).

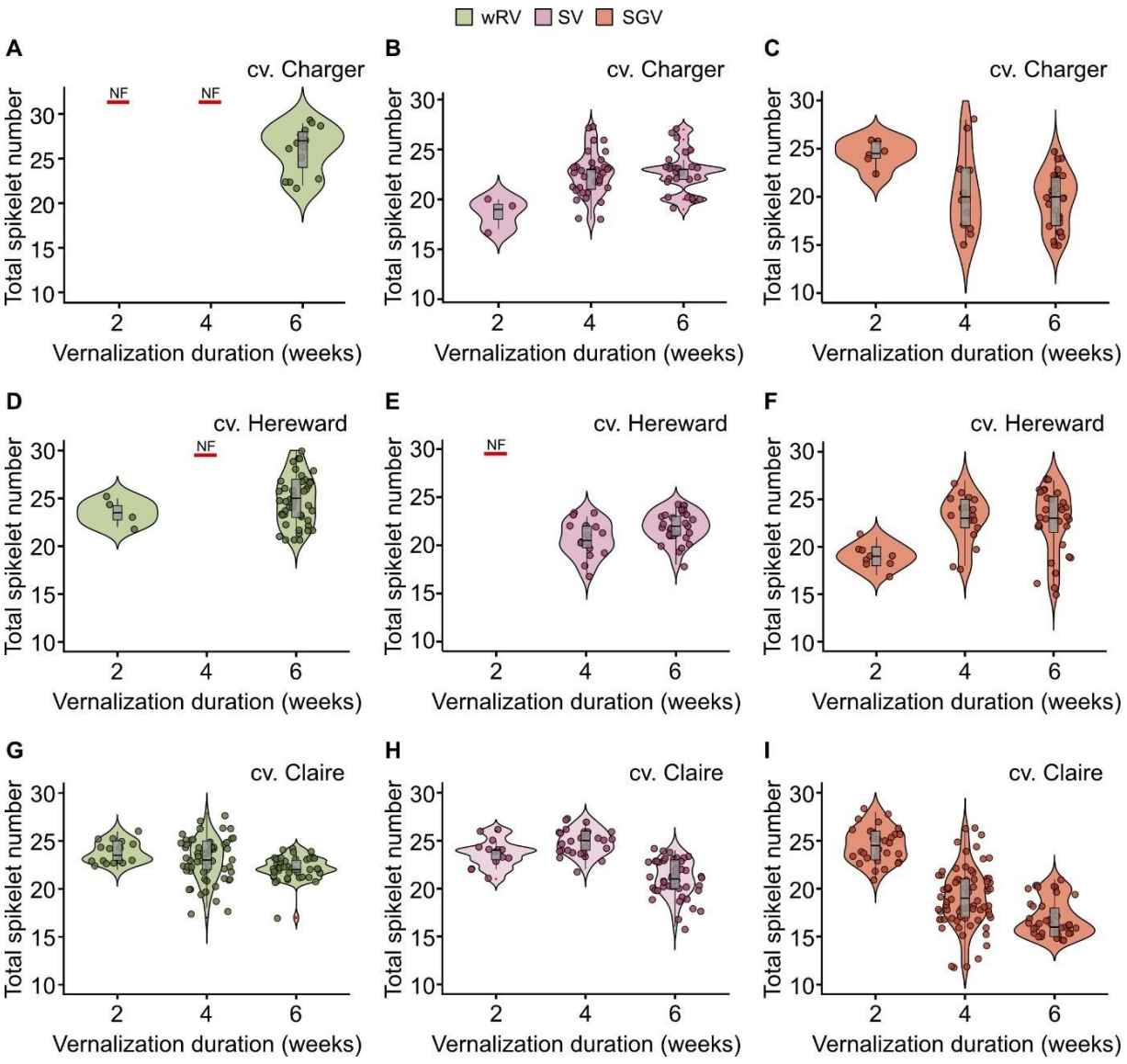


Figure 2.11: Spikelet number per ear under different vernalization conditions. (**A**-**C**) Total spikelet number per ear of winter cv. Charger under wRV (**A**), SV (**B**), and SGV (**C**). (**D**-**F**) Total spikelet number per ear of winter cv. Hereward under wRV (**D**), SV (**E**), and SGV (**F**). (**G**-**I**) Total spikelet number per ear of winter cv. Claire under wRV (**G**), SV (**H**), and SGV (**I**). Each vernalization treatment was followed by SB conditions to facilitate flowering.

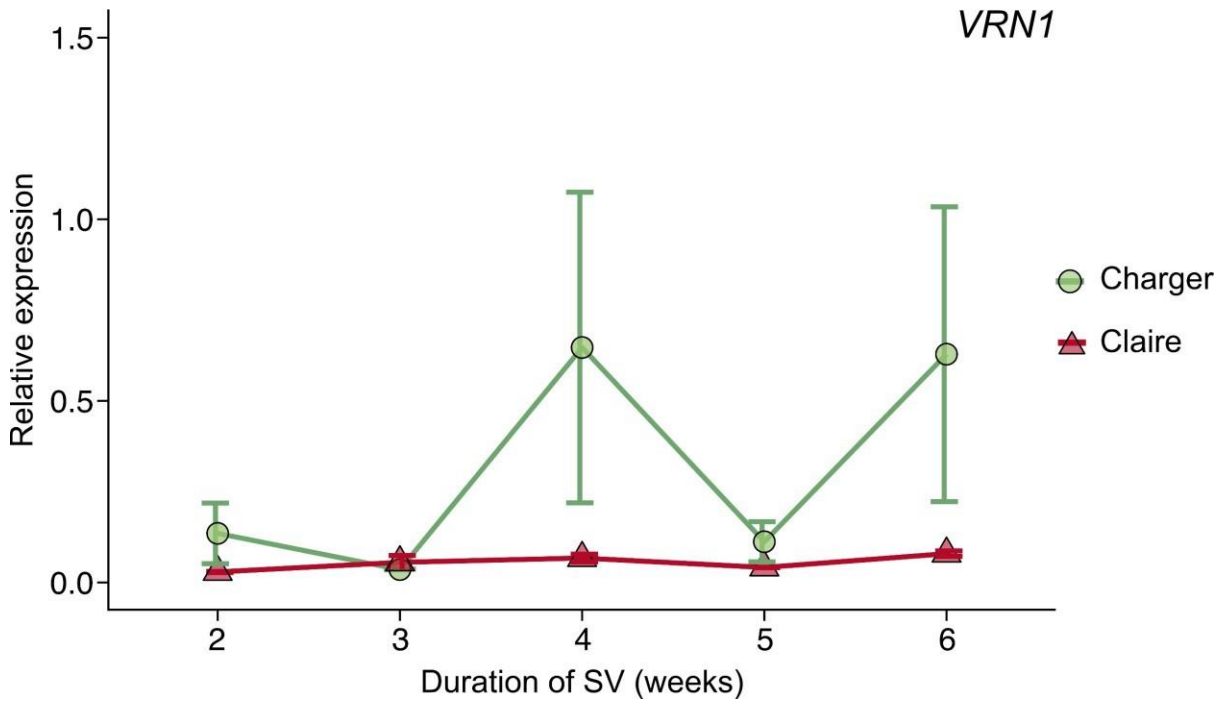


Figure 2.12: Expression of *VRN1* throughout speed vernalization. The level of *VRN1* expression relative to the reference gene is shown for cultivars Claire (red triangle) and Charger (green circle) across 6 weeks of growth under speed vernalization (SV). Error bars represent the standard error of the mean for 3 biological replicates.

2.3.6 Expression analysis reveals differences in SV vs. SGV as a method of vernalization

2.3.7 Central cereal vernalization gene expression doesn't explain the flowering response observed in SV and SGV

Flowering of winter wheat cultivars indicated that vernalization proceeded under the extended photoperiod present in both SV and SGV conditions. However, based on the flowering response, vernalization under SGV was more effective than SV. This introduced the question of whether classical vernalization genes were involved in SV and SGV, and whether differences in gene expression could explain the differences in vernalization observed for SV and SGV. To investigate this, leaf tissue from plants undergoing SV and SGV was analysed to measure expression of vernalization genes.

In cv. Hereward, vernalization genes appeared to behave differently under SV compared to SGV (**Figure 2.14**). *VRN1* was up regulated after 4 weeks of growth under SV (**Figure 2.14A**). *ZCCT1* was sharply down regulated between 2 and 3 weeks of SV but not completely repressed, with expression of the *ZCCT1* genes remaining stable between 3 and 6 weeks of SV. The counterpart of the *VRN2* locus, *ZCCT2*, was slightly up regulated between 2 and 3 weeks but down regulated between 3 and 4 weeks (**Figure 2.14A**). The final level of *ZCCT2* expression after 6 weeks SV was comparative to that of *ZCCT1*. Interestingly, flowering promoter *FT-B1* was completely switched off across the 6-week course of SV. Under SGV conditions, these genes behaved slightly differently (**Figure 2.14B**). Here, expression level of *VRN1* was higher than SV after 2 weeks SGV and by 3 weeks, expression of the gene was equal to highest expression at 6 weeks for SV (**Figure 2.14B**). Expression reduced moderately after this timepoint but remained stable between 5 and 6 weeks of SGV. Surprisingly, floral repressors *ZCCT1* and *ZCCT2* were both up regulated over the course of SGV (**Figure 2.14B**). This didn't correlate with increased expression of *VRN1*; instead, *VRN1* was down regulated after 3 weeks SGV. Overall, *ZCCT1* had the greatest rate of up regulation and steadily increased over the course of SGV (**Figure 2.14B**).

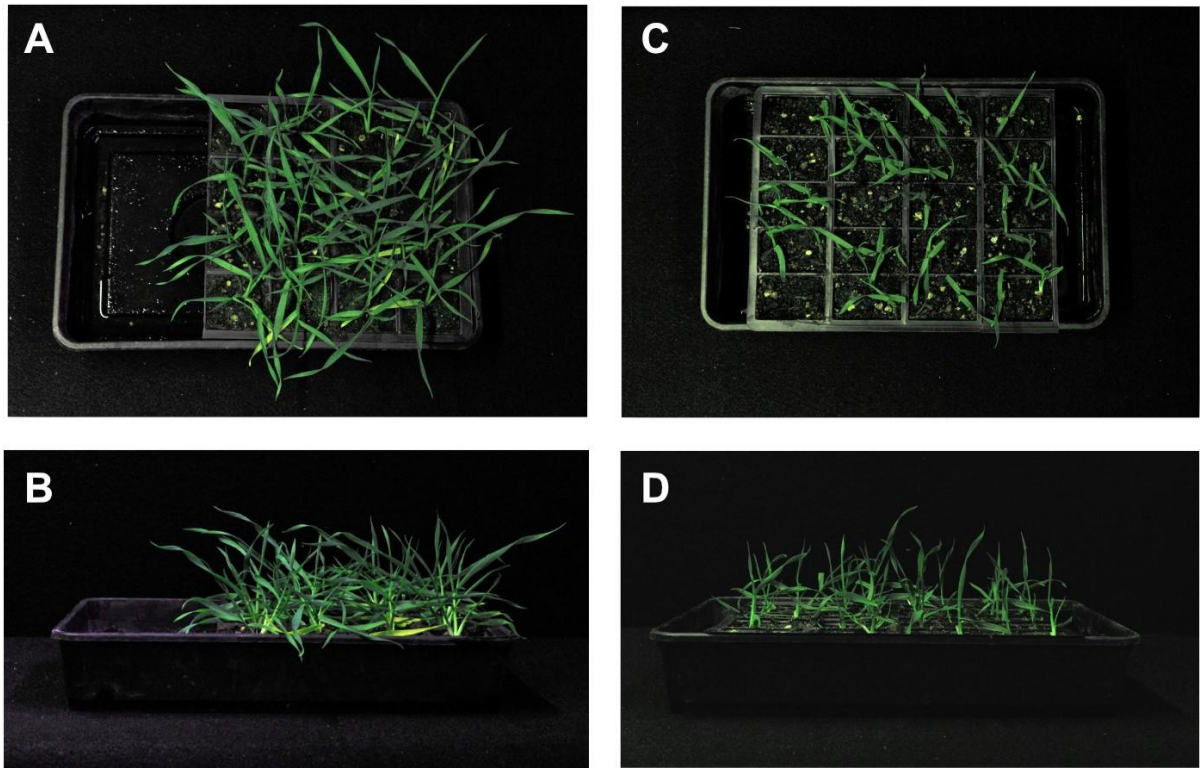


Figure 2.13: Comparison of plant growth for cv. Hereward under speed vernalization conditions. (**A**-**B**) Aerial and side profiles of Hereward plants after 4 weeks growth under SV conditions. (**C**-**D**) Aerial and side profiles of Hereward plants after 4 weeks growth under SGV conditions.

This analysis strongly suggested that the process of SV and SGV was following a different molecular route than standard vernalization but concluding with the same developmental output. This raised the question of which genes could be involved in the process. Using literature mostly based on the Arabidopsis flowering pathway, genes that were known to be involved in both light and temperature-related promotion of flowering were identified and incorporated into further analysis. These genes were *PHYTOCHROME INTERACTING FACTOR 4 (PIF4)* (Jin & Zhu, 2019; Xu & Zhu, 2021), *GIGANTEA* (*GI*) (Martin-Tryon et al., 2007; Park et al., 2020), *PHYTOCHROME C* (*PHYC*), and *CONSTANS1* (*CO1*) (Kinmonth-Schultz et al., 2016).

Under SV, both *PHYC* and *CO1* were expressed to a lower degree in cv. Hereward than under SGV (**Figure 2.15A**; **Figure 2.15B**). Here, *CO1* spiked after 4 weeks under SGV whereas expression level of *PHYC* was highest after 5 weeks of SGV (**Figure 2.15B**). For *GI*, expression decreased after 4 weeks under SV; conversely, *GI* was up regulated over 4 weeks and 5 weeks of SGV (**Figure 2.15C**). In cv. Hereward, expression of *PIF4* steadily increased over the course of vernalization under SV and SGV (**Figure 2.15C**; **Figure 2.15D**). Up until 4 weeks of growth, expression level of *PIF4* was similar under SV and SGV; beyond this point, expression level deviated and *PIF4* was expressed to a higher level under SGV (**Figure 2.15D**). This initial analysis demonstrated that flowering genes were being expressed differently under the SV and SGV conditions compared to wRV. This highlighted that the process identified was most likely following a different molecular pathway to regular vernalization. To understand the global gene expression changes with the aim of identifying genes which are involved in the newly identified molecular pathway, RNA-seq analysis was conducted.

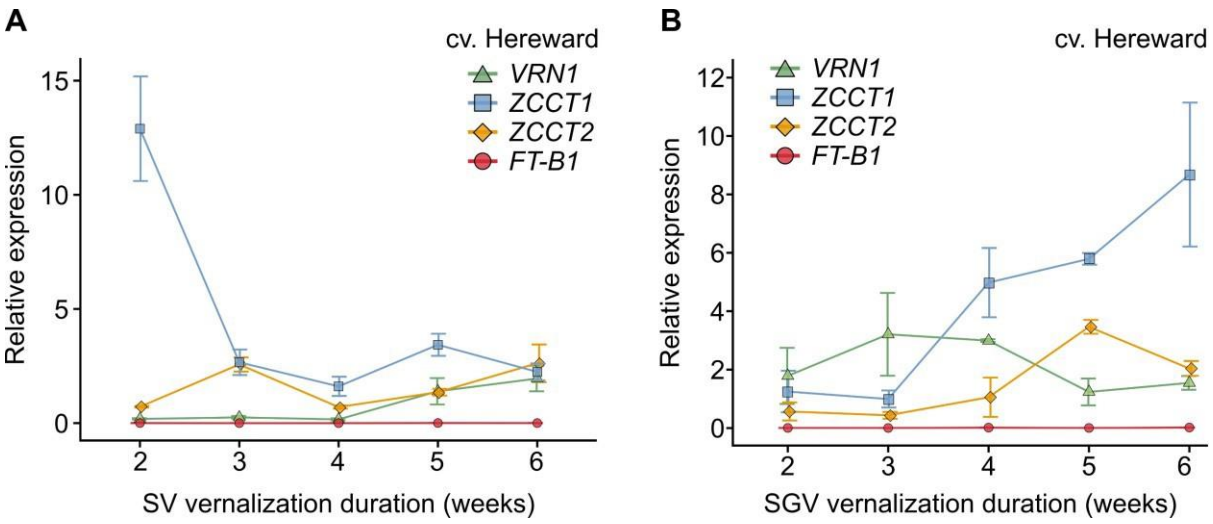


Figure 2.14: Expression of genes involved in the cereal vernalization pathway in cv. Hereward under speed vernalization conditions. (**A**) Relative expression of *VRN1* (green triangle), *ZCCT1* (blue square), *ZCCT2* (yellow diamond), and *FT-B1* (red circle) under SV conditions. (**B**) Relative expression of *VRN1*, *ZCCT1*, *ZCCT2*, and *FT-B1* under SGV conditions. Error bars represent the standard error of the mean for 3 biological replicates.

A control condition was included of leaf tissue from Hereward plants that had received SGV but the vernalization requirement wasn't met. Although the flowering results indicated that vernalization, and by extension flowering, proceeded faster under SGV, this didn't necessarily correspond with successful vernalization for all individuals. Therefore, tissue from Hereward that hadn't undergone vernalization during SGV was an appropriate control as it reduced confounding variables and enabled me to focus solely on the difference in vernalization response. RNA-seq analysis of total transcripts per kilobase million (TPM) revealed expression level of certain genes of interest under the three different vernalization conditions. First, *VRN-A1* was examined: this indicated that *VRN-A1* was expressed during SV, SGV, and the control condition (**Figure 2.16A**). The highest level of expression was under control conditions, followed by SGV, then SV. However, these differences in expression weren't statistically significant ($p>0.05$; ANOVA).

Expression of *FT-B1* wasn't detected under any of the vernalization conditions (**Figure 2.16A**). In contrast, *VRT2* had high expression across all conditions and for all homoeologues, with the highest level of expression overall in *VRT-D2* (**Figure 2.16B**). Finally, expression of the 5 different *ZCCT* genes varied: lowest expression level was in *ZCCT-A1*, followed by *ZCCT-A2* with consistent low expression across all conditions (**Figure 2.16C**). *ZCCT-B1* had slightly higher level of expression, with expression under SV higher ($p<0.05$; ANOVA and Tukey post-hoc test) than both SGV and the control condition. Highest expression overall was in *ZCCT-D1*, with a range of 12.7-16.4 TPM (**Figure 2.16C**). Furthermore, *ZCCT-D1* expression was greater under SV compared to the control condition, although this difference was not statistically significant ($p>0.05$; ANOVA).

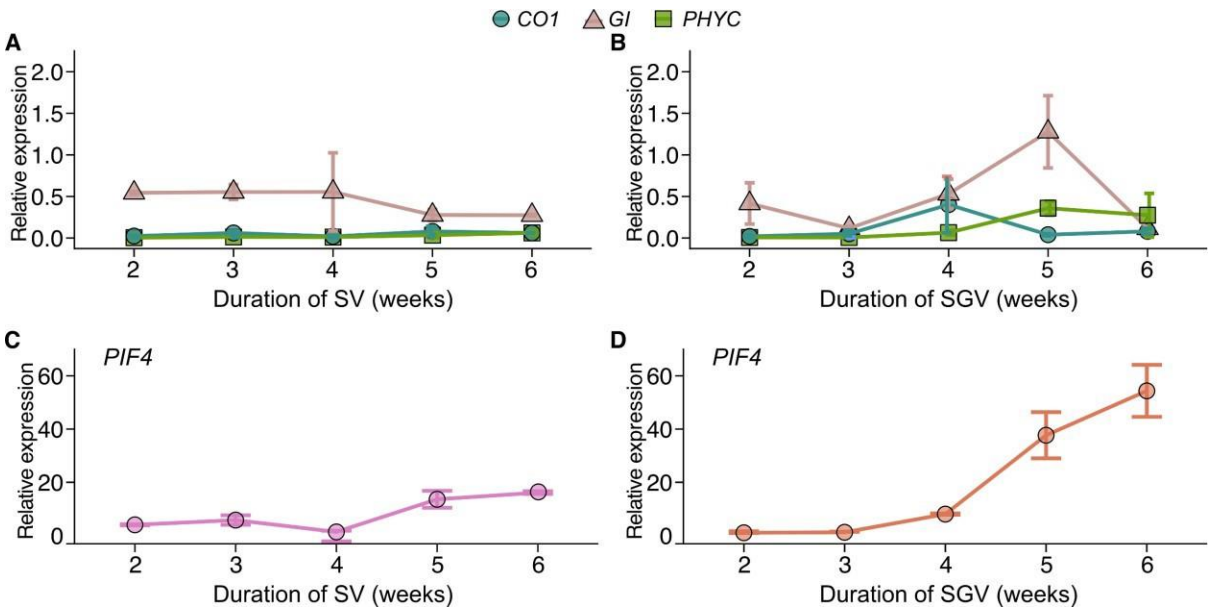


Figure 2.15: Expression of light and temperature-responsive genes in cv. Hereward under speed vernalization conditions. (**A**) Relative expression of *CONSTANS1* (*CO1*, blue circle), *GIGANTEA* (*GI*, purple triangle), and *PHYTOCHROME C* (*PHYC*, green square) under 6 weeks of SV conditions. (**B**) Relative expression of *CO1*, *GI*, and *PHYC* under 6 weeks SGV conditions. (**C**) Relative expression of *PIF4* under 6 weeks of SV conditions. (**D**) Relative expression of *PIF4* under 6 weeks of SGV conditions. Error bars represent the standard error of the mean for 3 biological replicates.

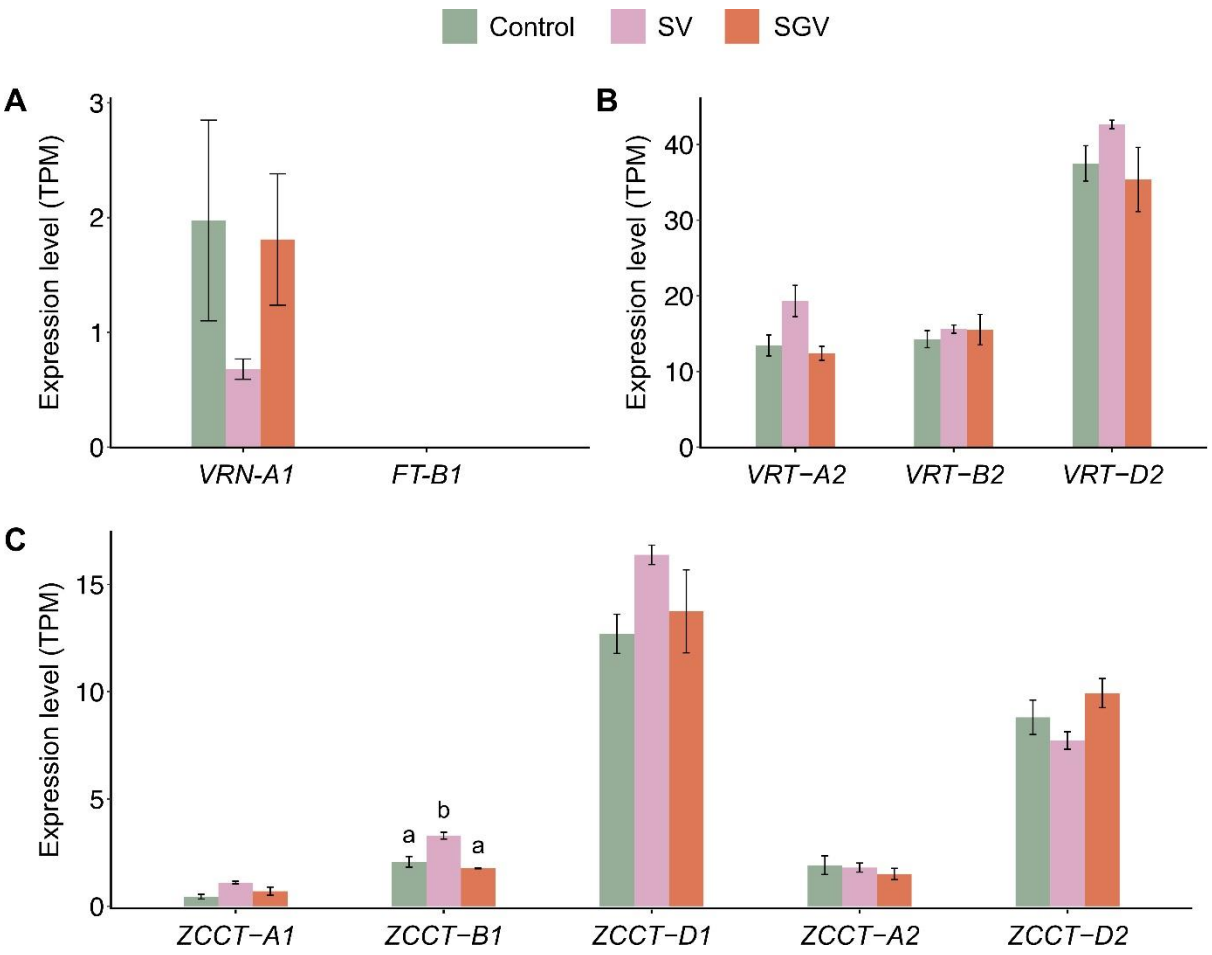


Figure 2.16: Expression of genes involved in the cereal vernalization pathway in cv. Hereward under different speed vernalization conditions. Expression measured by transcripts per kilobase million (TPM) for genes (**A**) *VRN-A1* and *FT-B1*, (**B**) *VRT2*, and (**C**) *ZCCT1* and *ZCCT2*. The vernalization conditions are the control group (grey), buried seed speed vernalization (SV, purple), and uncovered seed speed vernalization (SGV, red). Error bars represent the standard error of the mean for 3 biological replicates; letters above bars indicate statistically significant differences in expression level based on p-value, where $p<0.05$ using ANOVA and the Tukey post-hoc test.

Looking beyond known vernalization genes, expression of circadian clock and photoperiod genes was examined further. Using cv. Hereward, transcript abundance under SV, SGV, and the SGV control (incomplete vernalization under SGV conditions) was quantified for *PHYC*, *LUX*, *GI*, *CLAVATA2* (*CLV2*), *EARLY FLOWERING3* (*ELF3*), *ELF4*, *CONSTANS*-like (CO-like), and *TIMING OF CAB EXPRESSION1* (*TOC1*). Since expression was anticipated to differ between homoeologues, these were included in the analysis. For genes *CLV2* (**Figure 2.17D**), *ELF4* (**Figure 2.17F**), and *CO*-like (**Figure 2.17G**), there were no statistically significant differences in expression level between SV, SGV, and the control across any of the homoeologues. In contrast, *LUX* (**Figure 2.17B**) and *TOC1* (**Figure 2.17H**) had significant difference in one of their homoeologues. *LUX- D* had a higher level of expression ($p<0.05$; ANOVA and Tukey post-hoc test) under SV compared to both the control conditions and SGV conditions (**Figure 2.17B**). A similar result was observed in *TOC-B1*, with a greater expression ($p<0.05$; ANOVA and Tukey post-hoc test) reported under SV conditions compared to the control condition (**Figure 2.17H**). Overall, the prominent differences in expression were in *PHYC* and *ELF3*. In *PHYC-B*, there were statistically significant differences ($p<0.05$; ANOVA and Tukey post-hoc test) in expression level between all three conditions, with highest expression under SV, followed by SGV, and the control conditions (**Figure 2.17A**). Likewise, *ELF-D3* had differences in expression level. This statistically significant difference ($p<0.05$; ANOVA and Tukey post-hoc test) in expression level occurred between the control and SV conditions, as well as SV and SGV conditions, with highest overall expression under SV, followed by SGV and the control conditions.

2.3.8 Genes are differentially expressed between speed vernalization conditions

Analysis with the sleuth package identified various genes that were being differentially expressed between speed vernalization conditions. Comparison of transcripts per kilobase million (TPM) for the total number of aligned transcripts (133,346) in samples from SV and the control (i.e., non-vernalized) conditions identified 2,606 transcripts that were differentially expressed. Of this list, 2,587 were individual genes. Based on the sleuth analysis, the calculated fold-change values indicated that there were 811 transcripts up-regulated (fold-change value > 0) and 1,795 transcripts down-regulated (fold-change value < 0) between SV and the control condition.

GO enrichment analysis was performed on this list of 2,606 transcripts; this identified 801 enriched biological process pathways associated with the transcript list. Filtering by enrichment FDR and sorting by fold enrichment value ranked these pathways to indicate which pathways, and by extension genes within these pathways, were overrepresented in the enrichment analysis. The top 30 most enriched pathways are displayed in **Figure 2.18**. Within the top 30 pathways, there were no pathways associated with vernalization, temperature response, or reproductive processes; instead, pathways involved in chlorophyll biosynthesis and metabolism, pigment, and amino acid synthesis were highly enriched. The photosynthesis pathway was also highly enriched (fold enrichment = 4.2) with a total of 72 differentially expressed transcripts. From the complete enriched pathway list, certain pathways that could potentially elucidate genes of interest involved in vernalization under SV are listed in **Table 2.3**.

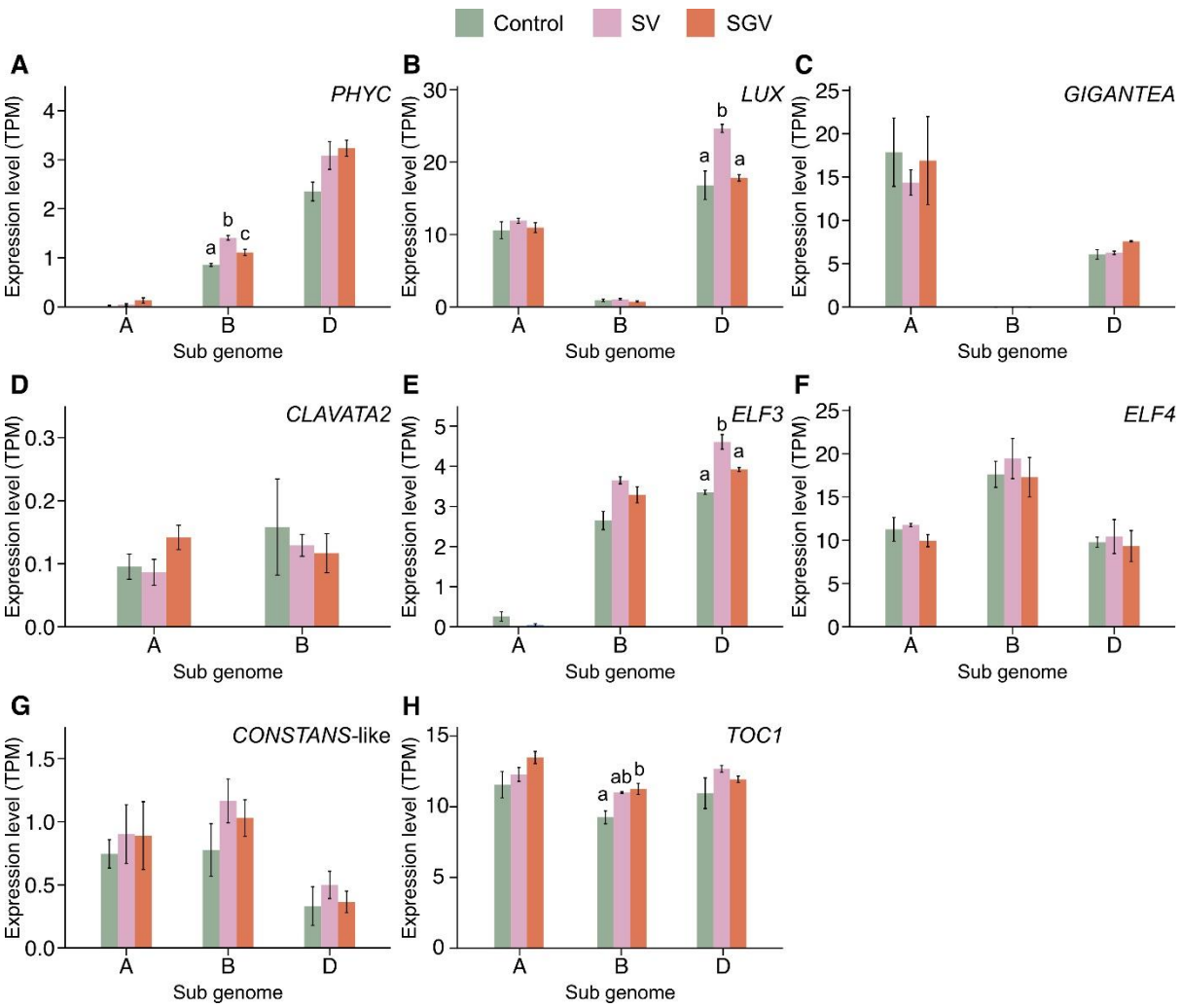


Figure 2.17: Expression of selected light and temperature-responsive genes in cv. Hereward under different speed vernalization conditions. Expression measured by transcripts per kilobase million (TPM) for genes (**A**) *PHYC*, (**B**) *LUX*, (**C**) *GIGANTEA*, (**D**) *CLAVATA2*, (**E**) *ELF3*, (**F**) *ELF4*, (**G**) *CONSTANS*-like, and (**H**) *TOC1*. The vernalization conditions are the control group (grey), buried seed speed vernalization (SV, purple), and uncovered seed speed vernalization (SGV, red). Leaf tissue was sampled after 4 weeks of growth. Error bars represent the standard error of the mean for 3 biological replicates; letters above bars indicate statistically significant differences in expression level based on ANOVA followed by the Tukey post-hoc test.

Table 2.3: Enriched pathways of interest from GO enrichment analysis of transcripts differentially expressed between SV and the control sample.

Pathway	transcripts	Enrichment FDR	Fold enrichment
Response to temperature stimulus	26	0.000002	3.31
Response to cold	13	0.001	3.30
Reproductive development	22	0.047	1.65
Response to light stimulus	30	0.000004	2.91
Response to light intensity	10	0.0001	5.26
Blue light signalling pathway	4	0.002	11.09

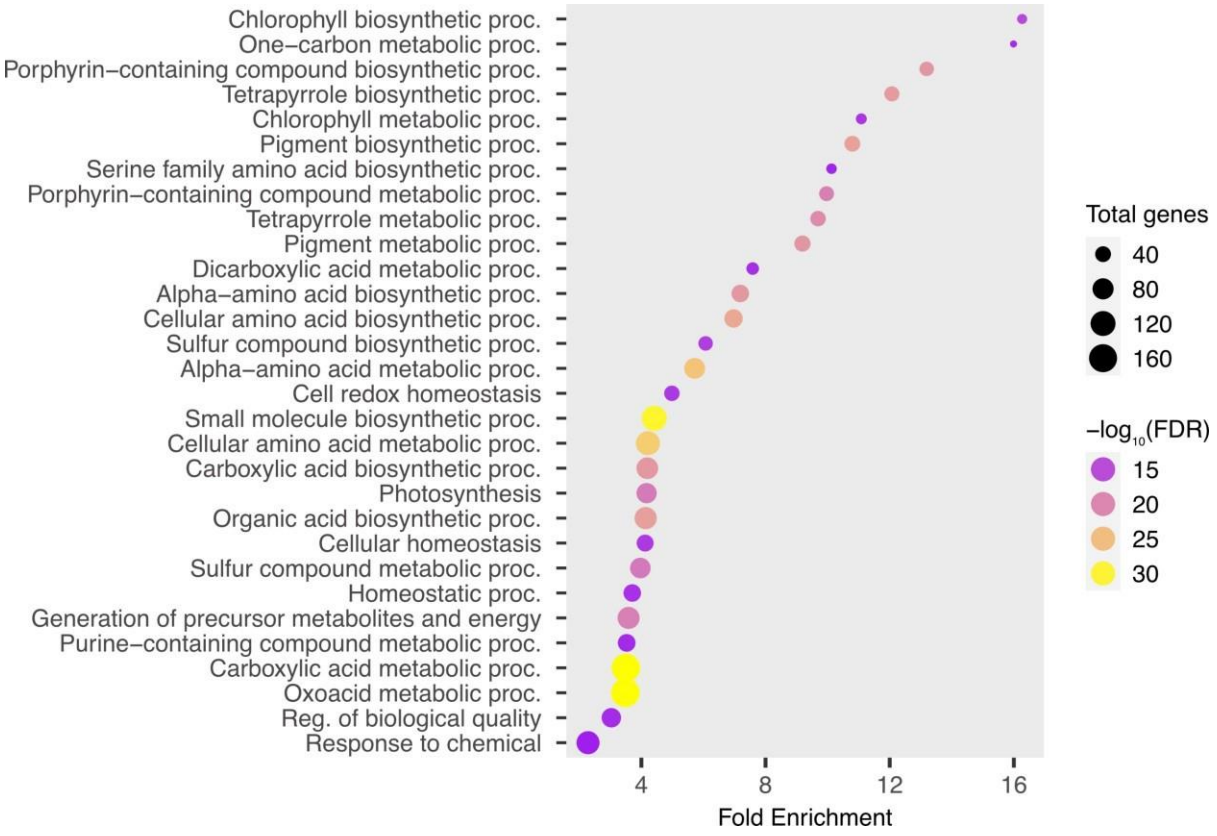


Figure 2.18: Enriched gene pathways under SV and the control vernalization conditions. The top 30 pathways were identified based on GO enrichment analysis of gene that were differentially expressed between vernalization conditions.

Comparison of TPM values for Speedvern samples grown beneath and on the surface of the soil identified 1,437 transcripts that were differentially expressed, 1,425 of which were individual genes. Based on the sleuth analysis, the calculated fold-change values indicated that there were 463 genes up-regulated (fold-change value > 0) and 974 genes down-regulated (fold-change value < 0). GO enrichment analysis assigned these differentially expressed transcripts to the most appropriate biological process pathways. The most enriched pathways were alpha-amino acid, cellular amino acid, and sulphur compound biosynthetic processes (**Figure 2.19**). Of the top pathways, the photosynthetic pathway (fold enrichment = 4.9; total differentially expressed transcripts = 47) was the most relevant pathway that could elucidate the differences observed between SV and SGV conditions. Within the total list of enriched pathways, relevant pathways included response to cold, cold acclimation, response to temperature stimulus, and response to light stimulus and intensity.

Comparative analysis of TPM values of the control, SV, and SGV samples revealed 417 transcripts which were differentially expressed between the three different conditions. Within this list, 412 genes were differentially expressed. To explore this set of differentially expressed transcripts further, soft clustering analysis was carried out to further understanding of expression patterns between these conditions. In preparation for this step, the expression values for all genes were standardised so that the mean value was equal to zero and the standard deviation was equal to 1. The optimal cluster number was set to 6 (**Figure 2.20**) and the total number of genes per cluster is listed in **Table 2.4**; these soft clusters are shown in **Figure 2.21**. Of these, clusters 2, 4, and 5 had expression profiles that were selected for further analysis as these followed the observed vernalization response – i.e., no vernalization in the control group, vernalization partially satisfied in SV, and vernalization saturated in SGV. Cluster 2 contained 78 transcripts, cluster 4 contained 70 transcripts, and cluster 5 contained 81 transcripts. Each of these clusters were subjected to GO enrichment analysis to identify enriched pathways. However, none of these transcripts were enriched in pathways associated with vernalization, temperature response, cold response or acclimation, or floral and/or reproductive development.

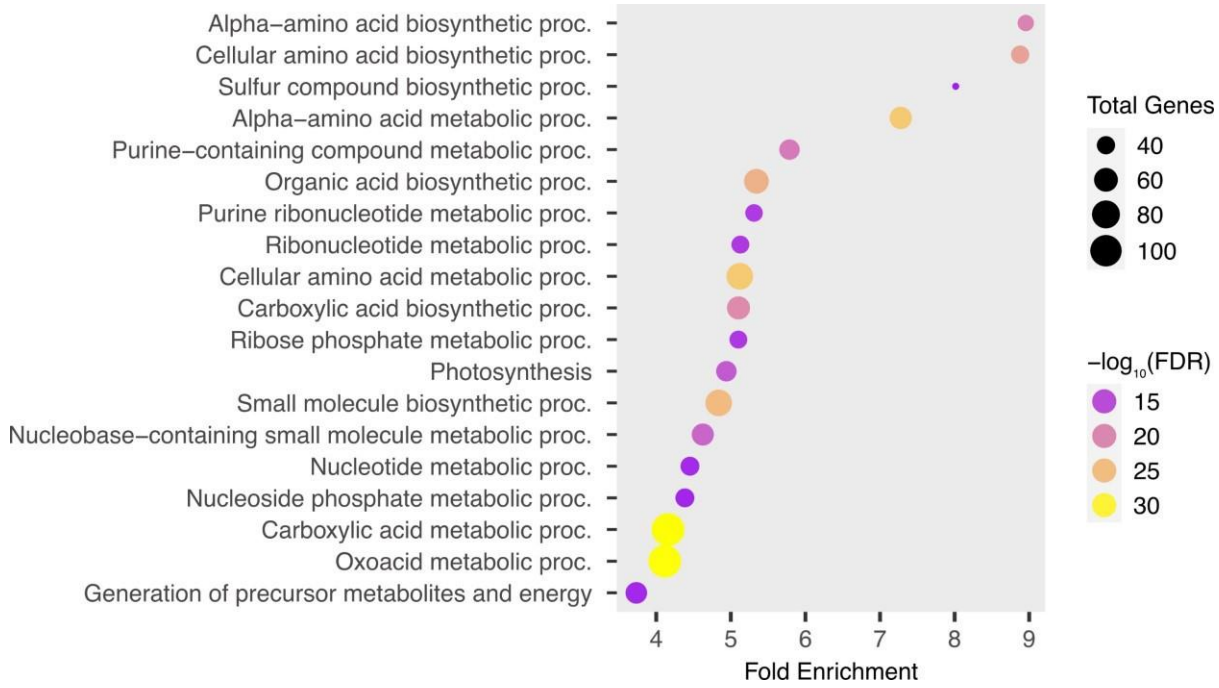


Figure 2.19: Enriched gene pathways under SV and the SGV vernalization conditions. The top 19 pathways were identified based on GO enrichment analysis of gene that were differentially expressed between vernalization conditions.

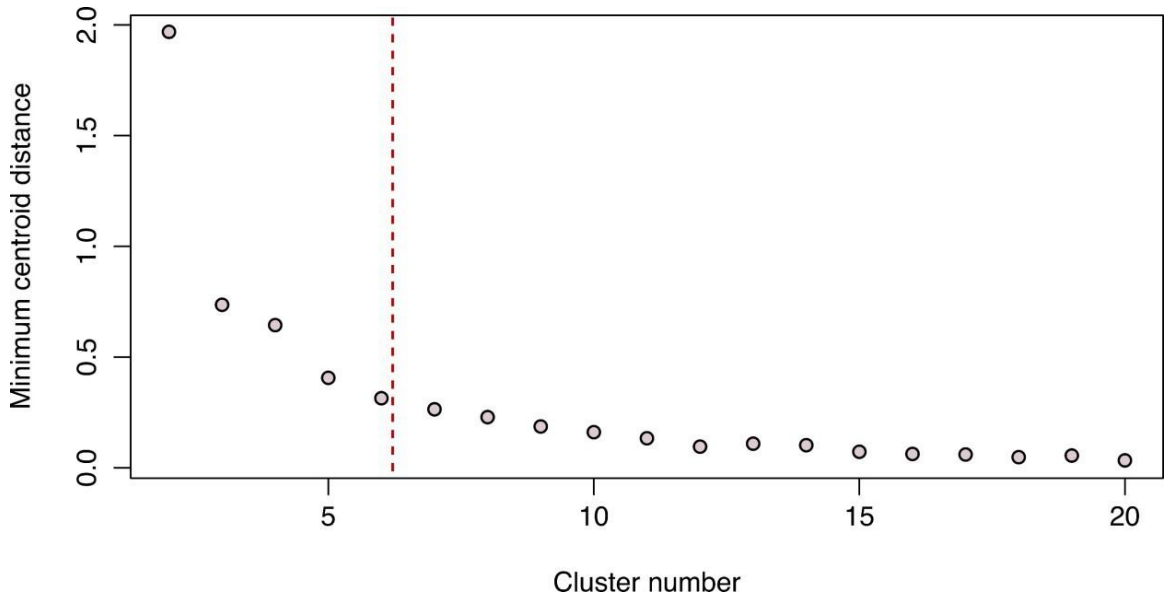


Figure 2.20: Selection of most appropriate cluster number for clustering analysis based on minimum centroid distance. Minimum centroid distance is calculated from the minimum distance between the centers of two clusters determined by c-means clustering. The point after which cluster number stabilises is represented by a broken line; therefore, n = 6 was selected as the optimal cluster number.

Table 2.4: Total number of genes per cluster based on RNA-seq analysis of the three different speed vernalization conditions; n = number of genes.

Cluster	n
1	37
2	78
3	73
4	70
5	81
6	78

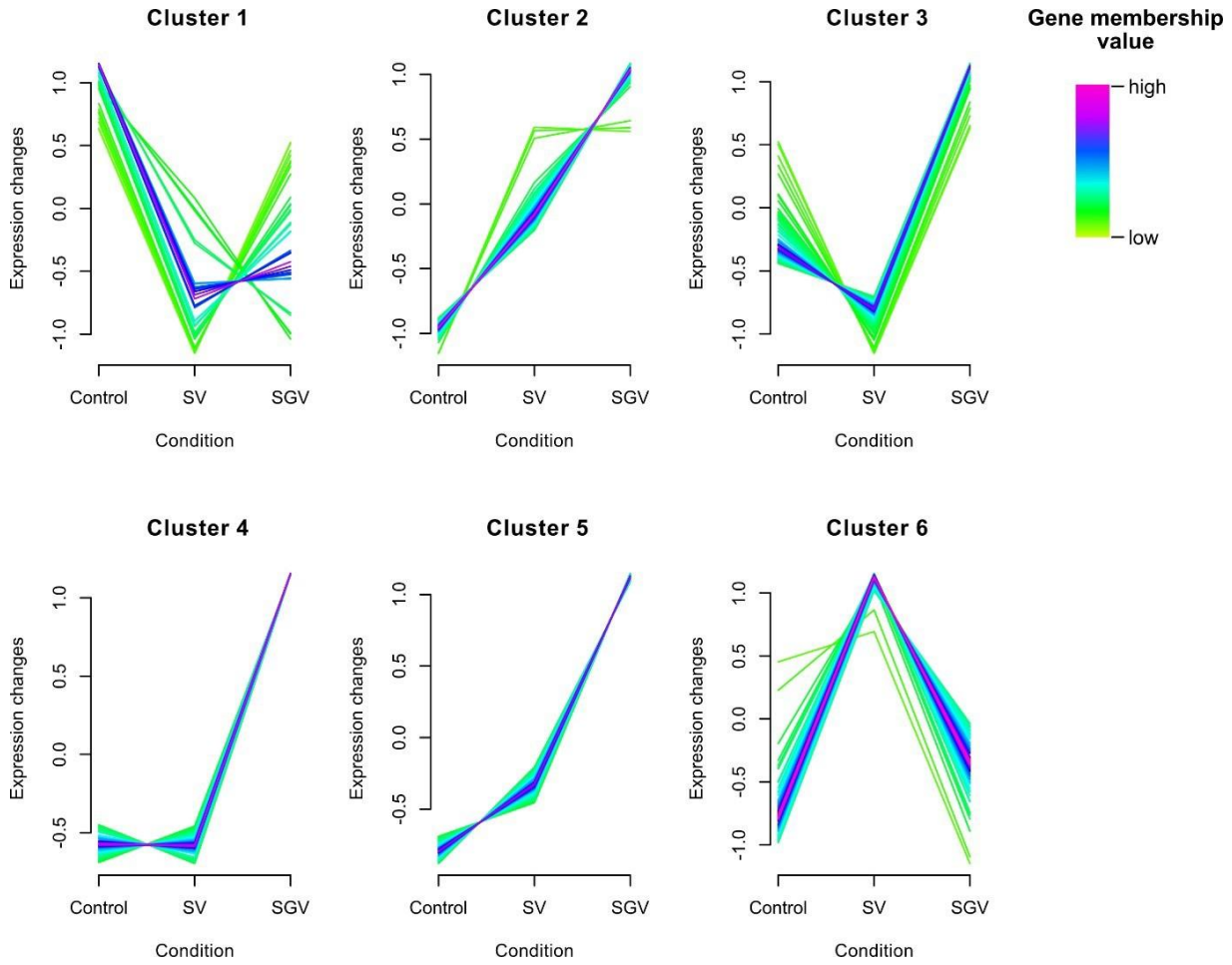


Figure 2.21: Soft clusters of speed vernalization expression data. Soft clusters were determined based on differentially expressed genes in cv. Hereward RNA-seq analysis data for the control, SV, and SGV conditions. Yellow and green lines represent low gene membership values; magenta and red lines represent high gene membership values.

2.3.9 Apex gene expression

Since vernalization facilitates the transition to reproductive tissue development in the apex, expression level in cv. Hereward apical tissue compared to leaf tissue during SGV. Hereward flowered after both 2- and 4-weeks of SGV, however flowering was more efficient and consistent by 4-weeks; this suggests that vernalization was in the process of completion between 2- and 4-weeks. For this reason, expression analysis of tissue after 3-weeks of SGV was carried out. Vernalization genes *VRN1*, *ZCCT1*, and *ZCCT2* were examined in addition to *PIF4*.

Greater *VRN1* expression level was observed in leaf tissue compared to apex tissue after 3-weeks SGV (**Figure 2.22A**). This difference in expression level was more marked for *ZCCT1*, with higher expression ($p<0.05$; Student's t-test) reported in the leaf compared to the apex (**Figure 2.22C**). Conversely, expression of *ZCCT2* was higher in the apex compared to the leaf (**Figure 2.22D**). Similar expression levels of *PIF4* were present in the apex and leaf, with slightly higher expression reported in the leaf (**Figure 2.22B**).

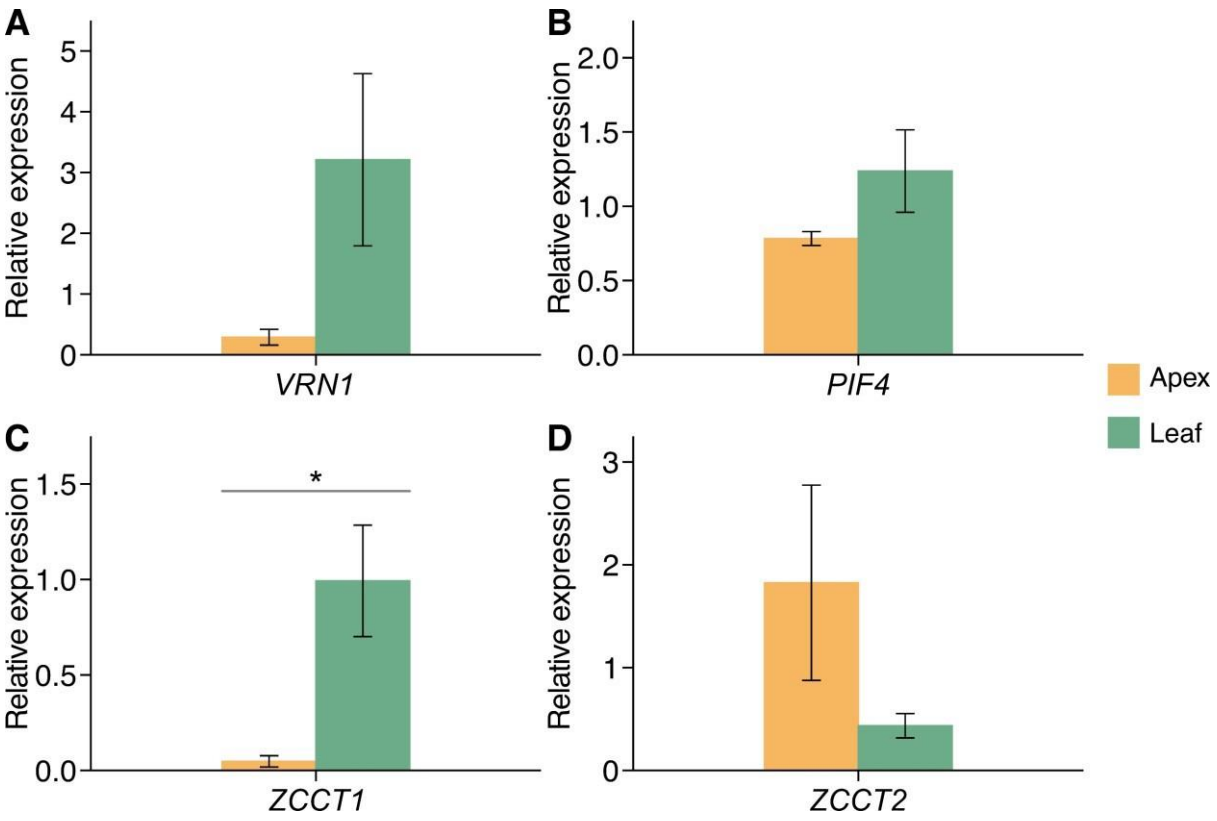


Figure 2.22: Expression of select genes in leaf and apex tissue under SGV conditions. (**A**) *VRN1*, (**A**) *PIF4*, (**A**) *ZCCT1*, and (**A**) *ZCCT2* expression in cv. Hereward in leaf (green) and apex (yellow) tissue after 3 weeks of growth under SGV conditions. Expression analysis was carried out using qPCR. Error bars represent the standard error of the mean for 3 biological replicates; asterisks above bars indicate statistically significant differences in expression level based on p-value (Student's t-test), * ≤ 0.05.

2.4 Discussion

2.4.1 Vernalization is completed under warmer ambient temperatures and is cultivar specific

Previous research has shown that vernalization can proceed under warmer temperatures than would classically be expected for winter varieties. Here, vernalization is completed under warmer temperatures of 10 °C and 14 °C short day conditions for winter cultivars Claire, Buster, Charger, and Hereward. Despite their similar genetic background, these cultivars have varying vernalization requirements attributed to *VRN-A1* copy number (Dixon et al., 2019). Previously, it was reported that cv. Charger (3 copies of *VRN-A1*, high vernalization requirement) had a delayed vegetative to reproductive transition of the apical meristem when exposed to higher ambient temperatures post-standard vernalization (Dixon et al., 2019). This indicates that floral repressors can be up regulated in response to higher temperatures post- vernalization.

In this chapter, winter cultivars Claire, Buster, Charger, and Hereward were successfully vernalized at temperatures between 8-14 °C; surprisingly, vernalization was more efficient and uniform under temperatures of 10-14 °C compared to a lower temperature of 8 °C. Since cv. Claire has a low vernalization requirement and can be considered to have a facultative growth habit, therefore it's vernalization response under warmer ambient temperatures isn't unexpected. However, cultivars Charger and Hereward have a high vernalization requirement; this is especially intriguing as their vernalization response is either more accelerated or comparative to cv. Claire.

Vernalization was completed in all cultivars after 6 weeks and further exposure to 10 °C or 14 °C didn't correspond with earlier flowering once plants had been transferred into 22 °C long day conditions post-vernalization. The responsiveness of these cultivars with varying vernalization requirements could suggest that *VRN-A1* is not the only element dictating vernalization under warmer temperatures. *VRN-A1* expression in these cultivars supports this. For example, *VRN-A1* was higher in cv. Claire under 8 °C, whereas in cv. Charger, *VRN-A1* was up regulated under 8-10° C. Interestingly, *VRN-A1* in cultivars Claire, Buster, and Charger was lowest at 14 °C. Despite low expression of *VRN-A1*, *ZCCT* genes were sufficiently

repressed. This could indicate that either repression of the *VRN2* locus via *VRN1* is up regulated earlier (i.e., before 4 weeks) under warmer temperatures, or that an additional floral activator which is yet to be identified is also involved.

If a winter cultivar goes on to flower, it is confirmation that the vernalization requirement has been met. Completion of vernalization can be identified at an earlier stage than this through the appearance of the double ridge stage of apex development. When grown under a low temperature, plants develop slowly and apices can remain in the vegetative stage for many weeks before temperature and photoperiods increase, facilitating the transition to double ridge stage. In this chapter, apices developed faster when grown under warmer temperatures during vernalization. The elongated apices still had a single ridge after 8 weeks of growth under warmer temperatures. As these lines had been vernalized after as little as 2 weeks, the absence of double ridge was especially interesting. This could potentially signify that when vernalization is completed under warmer temperatures, prolonged exposure to these temperatures can have a priming effect on the developing apex in preparation for the apical transition that occurs under florally-inducive conditions.

Since standard methods of vernalization aim to mimic winter, exposure to low temperatures (e.g., ≤8 °C) over a long period of time (e.g., >8 weeks) is typically used to ensure that vernalization is fully satisfied. However, unlike photoperiod, temperature throughout winter can naturally fluctuate daily and even hourly. Successful vernalization of the winter cultivars examined in this chapter support previous findings that vernalization can proceed under a wider range of temperatures than originally thought (Duncan et al., 2015; O'Neill et al., 2019). This supports looking beyond the standard vernalization pathway to understand how vernalization operates in cereals.

2.4.2 SV and SGV proceed in a manner that deviates from the classic cereal vernalization pathway

Building upon vernalization under warmer temperatures, we examined whether we could optimise the vernalization response further. As part of the speed breeding protocol, extending the photoperiod to 22 hr light accelerates growth. Although this adjustment can dramatically reduce generation time, cultivars with a moderate to high vernalization requirement still need vernalization prior to speed breeding. Under standard vernalization, this can add up to 8-12 weeks onto the speed breeding timeline. We were interested in optimizing this vernalization step to reduce the total vernalization to speed breeding timeline. Under 22 hr light/2hr dark at 22 °C, we would anticipate that cultivars with a high vernalization requirement that didn't receive vernalization would exhibit accelerated growth but would remain vegetative. In addition to a higher temperature, photoperiod has also been considered a limiting factor here. It is well documented that standard vernalization proceeds under a short- day photoperiod (Chouard, 1960). Central to this, floral repressor *VRN2* is down regulated under a short-day photoperiod (Dubcovsky et al., 2006). Supplemented by suppression of *VRN2* by *VRN1*, vernalization can proceed under these conditions.

Unexpectedly, vernalization was satisfied in winter cultivars Claire, Charger, and Hereward under the 22 hr light/2 hr dark photoperiod SV and SGV conditions (**Figure 2.8**; **Figure 2.9**; **Figure 2.10**). This indicates that, despite an extended photoperiod, the *VRN2* locus is not actively repressing *FT1* and the activation of flowering under these conditions. Under SV, expression analysis revealed that the *VRN2* was down regulated over time – however, this didn't correspond with significant up regulation of its repressor, *VRN1*. Even more surprising, *VRN2* was up regulated under SGV, contrasting with the down regulation of *VRN1* after 3 weeks of growth. This deviation from the standard process of vernalization suggests that a separate pathway that activates vernalization is operating under these conditions. Furthermore, this indicates that perception of light levels could be involved in this additional vernalization pathway, as plants were vernalized on the surface under an extended photoperiod. It was particularly interesting to observe that, although the classic vernalization genes were not being recruited for the signalling, variation in the response between cultivars was present (demonstrated

through different end flowering time). This is further supported as the cultivars Hereward and Charger showed different flowering times despite having the same *VRN-A1* copy number (**Figure 2.8**; **Figure 2.9**). If the SV and SGV response are using the classical vernalization pathway then we would anticipate that the cultivars with the same copy number of *VRN-A1* would respond in the same way. This indicated that the response was under genetic regulation and therefore we can anticipate identifying molecular components to that pathway.

In addition to extended photoperiod, increased light intensity paralleling Speed Breeding methods was also incorporated into the SV and SGV protocol. Since the key difference between the SV and SGV protocol was that seeds were buried in soil for SV in contrast to being placed uncovered on the soil surface for SGV, recognition of light level would be expected to vary between these conditions. In plants, the phytochrome family of photoreceptors function as red and far-red light sensors that facilitate development under varying light levels. Phytochromes can be further broken down into three groups: *PHYTOCHROME A* (*PHYA*), *PHYB*, and *PHYC* (Mathews & Sharrock, 1997). Unlike in Arabidopsis, *PHYC* in wheat is functional and stable when other phytochromes are absent, with *PHYC* promoting flowering under florally inductive conditions (Chen et al., 2014). In cereals, *PHYC* is linked to *VRN1*, contextualising its role in the flowering pathway (Chen & Dubcovsky, 2012). *PHYC* was more active under SGV compared to SV, supporting the hypothesis that this separate pathway is linked to light sensing.

In addition to the photoreceptor *PHYC*, clock genes were also examined. The circadian clock plays an integral role in regulation of plant development in response to changes in the environment. The evening complex, comprised of *ELF3*, *ELF4*, and *LUX*, facilitates the expression of *PIF4* and *PIF5* (Nusinow et al., 2011). Given that the components identified as showing differential expression between the SV, SGV and control samples were related to light and circadian signalling, as well as the identification of a number of pathways linked to light signalling (**Table 2.3**) it suggests that the alternative pathway may relate to high light stress signalling. This would explain why SV/SGV is able to bypass the vernalization signalling as it is triggering a more innate survival response.

In addition, it may also indicate why the pathway has not been identified before as it relates to stress responses rather than photoperiodic or vernalization responses. This is an area of interest which would be interesting to further understand. This is especially the case as if we were able to select for the SV/SGV response but not impact vernalization genes then breeding companies may be able to select for SV/SGV as a response to accelerate population development without impacting how the plants grow in the field.

For this reason, genes beyond the usual candidates should potentially be considered. In Arabidopsis, regulation of flowering under high light intensity involves multiple complex signals, including silencing of floral repressor *FLC* (Feng et al., 2016).

Chapter 3: Characterization of *ODDSOC* gene diversity

3.1 Introduction

3.1.1 Introducing MADS-box MIKC genes in plants

The role of MADS box genes in plant evolution and development is well recognised, occupying an essential role in plant survival and success through their involvement in many biological processes ranging from embryo to root to fruit development (Becker & Theißen, 2003; Gramzow & Theißen, 2010; Theißen et al., 2000). Of particular interest for research into reproductive development and flowering, MADS box genes control floral organ identity, ultimately driving floral development (Coen & Meyerowitz, 1991). MADS-box proteins are believed to function as transcription factors and bind to double-strand DNA as dimers. These proteins preferentially bind to AT-rich regions of DNA containing a highly conserved 10 bp core following the format of $CC(A/T)_6GG$, known as the CArG box (Messenguy & Dubois, 2003). Variants of this CArG box, such as $CC(A/T)_7G$ and $C(A/T)_8G$, also function as binding sites for some MADS-box proteins (Aerts et al., 2018; Folter & Angenent, 2006).

MADS box genes derive their name from *MINICHROMOSOME MAINTENANCE 1 (MCM1*), *AGAMOUS*, *DEFICIENS*, and serum-response factor (*SRF*) genes, as the domain common to these genes was originally identified in yeast, Arabidopsis, Antirrhinum, and humans respectively (Riechmann & Meyerowitz, 1997). Despite their ubiquity across eukaryotes, MADS box genes occupy an especially important role in plant development; this is supported by the presence of hundreds of MADS box genes in flowering plants, compared to only a handful in animals, fungi, and protists (Gramzow & Theißen, 2010; Pařenicová et al., 2003; Riechmann et al., 2000). A defining feature of all MADS-box genes is the MADS domain, a highly conserved sequence of approximately 180 bp that drives DNA-binding of MADS-domain proteins (Kaufmann et al., 2005).

MADS-box genes have been classified into two types: type I and type II (Alvarez-Buylla et al., 2000). Type I MADS-box genes encompass *SRF*-like genes and lack the keratin-like domain present in the type II group (De Bodt et al., 2003; Shore & Sharrocks, 1995). However, relatively little is known about the function of type I

genes in plants, with the vast majority of key developmental genes belonging to the type II group (Alvarez-Buylla et al., 2000; Becker & Theißen, 2003; De Bodt et al., 2003). The type II group include Myocyte Enhancer Factor 2 (MEF2)-like transcription factors in fungi and animals and MIKC-type genes in plants (Becker & Theißen, 2003).

MIKC-type genes are made up of four discrete domains: MADS-box (M), intervening (I), keratin-like (K), and C-terminal (C) (Pařenicová et al., 2003). These domains have specific functions, with the M and I domain facilitating DNA binding and dimerization, the K domain guiding protein dimerization, and the C domain acting as a transcriptional activation domain as well as contributing to protein interaction specificity (Egea-Cortines et al., 1999; Masiero et al., 2002; van Dijk et al., 2010). In total, the bread wheat genome contains 201 MIKC-type genes that are distributed equally across the chromosomes (Schilling et al., 2020). This number is substantially higher than many other plant species, with flowering plants previously reported to have between 24 (*Medicago trunculata*) and 113 (*Malus domestica*) MIKC-type genes (Gramzow & Theißen, 2013). It's speculated that the high number of these genes in wheat can partly be attributed to expansion and duplication events that have enhanced adaptability of wheat to a wide range of environments (Schilling et al., 2020).

3.1.2 The role of MADS-Box genes in vernalization

Since MADS-box genes are central components in flowering processes such as floral induction, meristem development, and reproductive organ formation, this group contains many well-characterised flowering time genes (Kaufmann et al., 2005). This includes genes from the vernalization pathway, which will be expanded on in the following sections.

In *Brassicaceae*, such as Arabidopsis and Brassica species, *FLOWERING LOCUS C* (*FLC*) is the major locus determining the vernalization requirement (Alexandre & Hennig, 2008; Michaels & Amasino, 1999; Sheldon et al., 1999). *FLC* is a well-characterised MADS-box gene that functions as a floral repressor through downstream/subsequent suppression of floral activators *FT* and MIKC-type MADS-box gene *SOC1* (Helliwell et al., 2006; Lee et al., 2000; Michaels & Amasino, 1999).

FLC functions as a central component of the vernalization pathway, with the level of *FLC* transcript corresponding with the extent of vernalization requirement; alternately, duration of cold exposure corresponds with downregulation of *FLC* (Sheldon et al., 2000). However, *FLC* is also responsive to temperature cues that are not associated with vernalization: for example, increased ambient temperature downregulates *FLC* (Gan et al., 2014), whereas short-term exposure to cold temperature upregulates *FLC* (Jung et al., 2013).

An interesting feature of *FLC* is its downregulation via a collection of *FLC*-associated non-coding antisense transcripts (Liu et al., 2010; Swiezewski et al., 2009). Of these antisense transcripts, the most extensively characterised is *COOLAIR* (Swiezewski et al., 2009). *COOLAIR* long non-coding RNA (lncRNA) antisense transcripts are transcribed downstream from *FLC* and terminate beyond the sense strand transcription start site for *FLC* (Liu et al., 2010; Marquardt et al., 2014; Swiezewski et al., 2009). A second group of lncRNA transcripts associated with *FLC* is *COLDAIR* (Heo & Sung, 2011). *COLDAIR* is transcribed from the first intron of *FLC* and suppresses *FLC* activity through enrichment of POLYCOMB REPRESSION COMPLEX 2 (PRC2) (Kim et al., 2017).

As previously mentioned, the identification of *FLC* outside of Brassicaceae has been challenging (Kennedy & Geuten, 2020; Ruelens et al., 2013). This is partly due to the large number of highly similar MADS-domain proteins in wheat, and also as the vernalization pathway in cereals was identified to be operating independently of *FLC*. Instead, the MIKC-type MADS-box gene *VRN1* is the main locus involved. *VRN1* has been characterised as an orthologue of MIKC-type MADS-box gene *APETALA1* (*AP1*) in Arabidopsis (Yan et al., 2003). As the MIKC-type classification suggests, *VRN1* is comprised of a 60 amino acid MADS-box domain and a 90 amino acid K- domain, in addition to I and C domains (Danyluk et al., 2003). In barley, VRN-H1 protein binds sites involved in flowering, including *FT1*, *CONSTANS*-like genes, and *VRN2* (Deng et al., 2015). Interestingly, VRN-H1 also binds to the promoter of a MIKC-type MADS-box floral repressor unique to cereals: *HvODDSOC2* (Deng et al., 2015).

3.1.3 Discovery of *ODDSOC1* and *ODDSOC2*

ODDSOC genes were first named and characterised in barley as *HvOS1* and *HvOS2* (Greenup et al., 2010). However, these genes were first identified by Trevaskis et al. (2003) based on difference in expression level of MADS box genes in spring versus winter wheat seedlings. Here, *MADS box gene 23* (*TaMX23*) was more strongly expressed in the non-vernalized winter cultivar compared to spring; subsequent vernalization repressed *TaMX23* in the same winter line. Later, Winfield et al. (2009) reported that *TaMX23* had high sequence similarity (95%) to the previously annotated gene *TaAGL42*. In addition, *TaMX23* also had high sequence similarity (85%) to *TaAGL33* (Winfield et al., 2009). *TaAGL33* and *TaAGL42* have the closest amino acid sequence similarity (72% and 61%, respectively) to *OsMADS51* in rice (Winfield et al., 2009; Zhao et al., 2006). The transcript profiles of *TaAGL33* and *TaAGL42* led Winfield et al. (2009) to describe *TaAGL33* as a repressor of flowering and *TaAGL42* as a promoter in winter wheat cultivars.

Returning to *TaMX23*, two homologues in barley were identified by Greenup et al. (2010). No equivalent was found in Arabidopsis, although both genes shared weak similarity with *AtSOC1* – therefore, the genes were assigned the names *HvODDSOC1* (*HvOS1*) and *HvODDSOC2* (*HvOS2*) (Greenup et al., 2010). Based on amino acid sequence similarity, *HvOS1* shared greatest similarity with *TaAGL42* and *HvOS2* was most similar to *TaAGL33* (both 95% match), corresponding with what was reported by Winfield et al. (2009). For *HvOS1*, the next closest amino acid sequence was that of *BdOS1* in *B. distachyon*, whereas *HvOS2* shared greatest similarity with *TaAGL41* and *BdOS2* after *TaAGL33* (Greenup et al., 2010). Beyond this point, *HvOS1* and *HvOS2* converged in sequence similarity, with their next closest genes being *OsMADS51*, *SbOS-like* in *Sorghum bicolor*, and *ZmOS-like* in *Zea mays* (Greenup et al., 2010). Despite the additional identification of these *ODDSOC*-like genes in Brachypodium, sorghum, maize, and rice, so far these genes appear to be unique to grasses (Greenup et al., 2010).

Despite *HvOS1* and *HvOS2* sharing 88.9% sequence similarity (Kapazoglou et al.,

2012), Greenup et al. (2010) observed that *HvOS1* was upregulated during vernalization, whereas *HvOS2* was downregulated. Furthermore, expression of *HvOS2* was decreased in a barley line containing a *HvVRN1* allele that conferred reduced vernalization requirement and higher HvVRN1 levels when compared to the wild-type version of the *HvVRN1* allele; this suggests that *HvVRN1* downregulates *HvOS2* (Greenup et al., 2010). In contrast, overexpression of *HvOS2* resulting in delayed flowering and inhibited leaf and apex development (Greenup et al., 2010). Additionally, overexpression of *HvOS2* downregulated the equivalent of *FLOWERING PROMOTING FACTOR 1* (*FPF1*), further supporting the predicted role of *HvOS2* as a repressor of flowering (Greenup et al., 2010).

Furthering the link between *HvOS2* and vernalization, *HvOS2* was downregulated in winter barley grown under long-day conditions at 25°C and short-day conditions at 15°C (Hemming et al., 2012). Conversely, under short-day conditions at 25°C *HvOS2* was upregulated (Hemming et al., 2012). In conjunction with the increased *HvOS2* transcript levels in non-vernalized barley plants that had yet made the transition from vegetative to reproductive growth (Hemming et al., 2012), it would suggest that *HvOS2* inhibits floral transition under unfavourable flowering conditions such as short-day photoperiod warm temperature. Interestingly, transcript levels of *HvOS1* were lowest under long-day conditions at 25°C (Hemming et al., 2012). This is unexpected as *OS1* has been considered to function as a floral activator in a similar manner to *VRN1*, potentially suggesting that *OS1* has a role in the vernalization pathway at an earlier stage of development.

In addition to being expressed during vegetative and early reproductive development, *OS1* and *OS2* have also reported to be involved in other developmental processes. For example, *HvOS1* and *HvOS2* are differentially expressed in the endosperm during seed development of facultative barley cultivars (Kapazoglou et al., 2012). *HvOS2* transcripts have also been detected in roots, leaves, and immature flowers (Kapazoglou et al., 2012). Furthermore, in the tetraploid spring wheat cv. Kronos, the homologue of *HvOS2*, *TaAGL33*, was upregulated in *phytochromeB*-null and *phytochromeC*-null wheat TILLING mutants, alongside central flowering gene *CONSTANS1* (Pearce et al., 2016). This could

suggest that the role of the *ODDSOC* gene family in cereals extends beyond repression and activation of flowering.

3.1.4 Parallels between *ODDSOC* genes and *FLC*

ODDSOC genes were first characterised as part of the *FLC*-like gene group by Ruelens et al. (2013). Here, a putative *FLC* subfamily in cereals was identified based on syntenic markers including tandem repeats in close proximity to *SEPALLATA* and *SQUAMOSA* genes across a range of genera such as *Oryza*, *Brachypodium*, *Populus*, *Vitis*, and *Solanum* (Ruelens et al., 2013). Using this approach, *OsMADS37* and *OsMADS51* were identified as strong candidates for the closest *FLC*-like gene in rice. This links *FLC*-like genes with the orthologues of *OsMADS51*, *HvOS2* (Greenup et al., 2010) and *BdOS2* (Ruelens et al., 2013).

In *Brachypodium*, Ruelens et al. (2013) reported that *BdOS2* was downregulated when exposed to a prolonged period at 4 °C. In contrast, *BdOS2* was upregulated when exposed to 28°C conditions (Ruelens et al., 2013). These findings were explored further by (Sharma et al., 2017), confirming that *BdOS2* is downregulated in Brachypodium during cold exposure in a manner similar to *FLC*. Furthermore, overexpression of *BdOS2* in a facultative winter barley cultivar that had received 2 weeks of vernalization at 4 °C resulted in delayed flowering and increased vernalization requirement (Sharma et al., 2017). Conversely, when *BdOS2* was knocked out, this led to accelerated flowering when plants were grown without initial vernalization at 4°C, effectively removing the vernalization requirement (Sharma et al., 2017). Finally, differences in H3K27me3 modifications of *BdOS2* was observed for spring and winter Brachypodium, with the winter line exhibiting H3K27me3 enrichment across the entire locus during vernalization (Sharma et al., 2017). This is reminiscent of the regulation observed across the *FLC* locus in Arabidopsis (De Lucia et al., 2008; Heo & Sung, 2011; Swiezewski et al., 2009).

In addition, equivalents of the lncRNA *COOLAIR* transcript that regulates *FLC* has been identified in Brachypodium (Jiao et al., 2019). These transcripts were named *BdCOOLAIR1* and *BdCOOLAIR2* and were associated with *BdOS1* and *BdOS2*, respectively (Jiao et al., 2019). In a winter Brachypodium accession, these antisense transcripts were induced by the cold (Jiao et al., 2019). Intriguingly, *BdCOOLAIR1* and *BdCOOLAIR2* were both upregulated during vernalization; this resulted in both *BdOS1* and *BdOS2* being suppressed, indicating interaction between transcripts (Jiao et al., 2019). Both *BdCOOLAIR* lncRNA were well conserved across 54 *Brachypodium* lines (Jiao et al., 2019). Additionally, similar antisense were identified in silico for two bread wheat *OS2* genes, *TaAGL33* and *TaAGL41* (Jiao et al., 2019).

3.1.5 The role of *ODDSOC* genes in cereal vernalization

The characterisation from barley and *Brachypodium* has identified that the *ODDSOC* genes *OS1* and *OS2* are MIKC-type MADS-box genes speculated to be involved in the vernalization pathway in cereals (Deng et al., 2015; Greenup et al., 2010; Sharma et al., 2017). The exact molecular mechanism and nature of this role has yet to be determined, with more occurrences in the literature of characterisation of *OS2* compared to *OS1*. Furthermore, most findings have been reported in the model cereal diploid *B. distachyon* and the diploid barley, with no in-depth analysis of the existence and role of *ODDSOC* genes in hexaploid wheat.

In barley, *HvOS2* is down regulated by low temperature, with the prolonged low temperature associated with vernalization up regulating *HvVRN1* which further down regulates *HvOS2* (Greenup et al., 2010). A similar result has been reported for *BdOS2* in *B. distachyon* (Ruelens et al., 2013). Studies using cultivars with varying *VRN1* alleles identified that *HvOS2* is repressed independently of *HvVRN1* but sustained repression of HV*OS2* by *HvVRN1* is necessary in vernalized plants (Greenup et al., 2010). Similarly, *BdOS2* in *B. distachyon* is down regulated prior to up regulation of *BdVRN1* (Sharma et al., 2017). This indicates that *OS2* is not exclusively repressed by *VRN1* and that other unidentified elements are involved in this mechanism. Photoperiod also plays a role in this process, with *HvOS2* up regulated in short-day warm temperature but down regulated in long-day warm temperature conditions (Hemming et al., 2012).

Additionally, *HvOS1* in barley has been shown to be up regulated in response to low temperature (Greenup et al., 2010). However, in *B. distachyon BdOS1* was reportedly down regulated during vernalization (Jiao et al., 2019). This demonstrates how the connection between *OS1*, *OS2*, and other components of the vernalization pathway, including *VRN1* and floral repressor *VRN2* remains unclear. To investigate this further, I aimed to identify the relationship between *ODDSOC* genes in wheat and investigate their role in temperature responsiveness.

3.2 Material and Methods

3.2.1 Growth conditions and phenotyping

3.2.1.1 Germplasm and experimental set up

Winter wheat cv. Claire was germinated in cereal mix as described in section **2.2.2**. Once established, plants were grown under 10 °C short-day (8 hr light/16 hr dark) photoperiod conditions in a Sanyo MLR-352 plant growth chamber. After 4 weeks of growth under these conditions, leaf tissue for 3 biological replicates was sampled in 1.5 mL microcentrifuge tubes containing sterile ball bearings and flash frozen in liquid nitrogen.

To examine response to photoperiod independent to temperature, cultivars Claire, Buster, and Charger were germinated in cereal mix (described in **2.2.2**) and grown in a Sanyo MLR-352 plant growth chamber under 22 °C short day (8 hr light/16 hr dark), long day (16 hr light/8 hr dark), and long-long day (22 hr light/2 hr dark) conditions. These conditions were not inducive to vernalization. Leaf tissue was sampled after 3 weeks of growth.

For investigation of expression over a 24-hr period, cv. Buster was germinated in cereal mix and grown at either 10 °C or 16 °C short day (8 hr light/16 hr dark) photoperiod conditions, or 10 °C or 16 °C long-long day (22 hr light/2 hr dark) photoperiod conditions in a Snijders MICROCLIMA MC1000 plant growth cabinets. Leaf tissues was sampled at zeitgeber ZT0, ZT5, ZT8, ZT10, ZT15, and ZT20 timepoints and flash-frozen in liquid nitrogen. Experimental setup and sample processing was carried out by Dominique Hirsz.

Winter wheat cv. Julius and cv. Buster were germinated in cereal mix and grown under 10°C short-day (8 hr light/16 hr dark) photoperiod conditions in a Sanyo MLR- 352 plant growth chamber for 3 weeks, at which point leaf tissue was sampled at the zeitgeber ZT20 timepoint, i.e., 20 hours post-dawn. Tissue from 3 biological replicates was sampled in 1.5 mL microcentrifuge tubes containing sterile ball bearings and flash-frozen in liquid nitrogen.

3.2.2 Gene and transcript expression

3.2.2.1 YoGI expression data

Expression data from the YoGI panel dataset was provided by Dr Andrea Harper at the University of York. The experimental design is described in (Barratt, Reynolds, et al., 2023), but briefly, 342 wheat lines from the YoGI accession panel were grown under long day glasshouse conditions (16 hr light/8hr dark, 20°C/14°C). Leaf tissue was sampled at GS13 (i.e., 3-leaf) stage for expression analysis. Total transcripts were normalised as Fragments Per Kilobase of transcript per Million mapped reads (FPKM).

3.2.2.2 RNA extraction

Frozen leaf tissue from cv. Claire, Buster, and Julius was disrupted using the TissueLyser II (Qiagen). RNA from disrupted tissue was extracted using the Spectrum™ Plant Total RNA Kit (Sigma-Aldrich) following the manufacturer's instructions and eluted in 50 µL dH_2O. The concentration and purity of extracted RNA was measured using the NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific). RNA was stored at -80 °C.

3.2.2.3 cDNA synthesis

RNA from cv. Claire, Buster, and Julius was used as a template for cDNA synthesis following the method already described in section **2.2.4.1**. Once synthesised, cDNA was stored at -20 °C.

3.2.2.4 qPCR analysis

Expression level of genes was quantified using qPCR. First, primers were designed within the CDS of the gene for quantification of gene expression level and within the antisense transcript for quantification of the lncRNA transcript. Primers amplifying the normalisation gene *TraesCS5A02G015600* were also used; complete list of primers used are listed in **Table 2.1**. The qPCR reaction and cycling conditions were previously described in section **2.2.4.1**.

3.2.3 In silico analysis of *ODDSOC* genes and gene structure

3.2.3.1 Identification of *ODDSOC* genes within a flowering eQTL

Expression quantitative trait locus (eQTL) dataset was generated and kindly provided by Dr Andrea Harper at the University of York. Gene expression marker (GEM) analysis was performed to identify markers with a significant ($p < 0.05$) false discovery rate (FDR) score. Expression was scored against the early vegetative trait.

3.2.3.2 Using pathway networks/phylogenetics to infer additional *ODDSOC* genes

Involvement of genes of interest in genetic pathways was analysed using the Reactome Pathway Browser (www.reactome.org). Knetminer (www.knetminer.com) was used to identify involvement of genes in additional pathways and processes. IWGSC RefSeq v1.0 (2018) gene I.Ds were obtained from Ensembl Plants, whereas IWGSC RefSeq v2.1 (2021) gene I.Ds were obtained from GrainGenes (https://wheat.pw.usda.gov/GG3/genome_browser). The online platform PLAZA (https://bioinformatics.psb.ugent.be/plaza/versions/plaza_v5_monocots/; Van Bel et al. (2022)) was used to build a phylogenetic tree. A network was generated using the original *ODDSOC* gene *TraesCS3A02G435000* (see **Table 3.1** for all IWGSC IDs for this gene) as a seed gene to generate a network by combining MUSCLE multiple sequence alignment (https://www.ebi.ac.uk/Tools/msa/muscle/; Edgar (2004)) across multiple genes in *T. aestivum* with the FastTree construction algorithm ("https://github.com/PavelTorgashov/FastTree"). This produced a phylogenetic tree of 188 *T. aestivum* genes that were the closest related to the seed gene, based on amino acid sequence alignment. To determine the presence of homoeologues, genes were input into Ensembl Plants (https://plants.ensembl.org/index.html). Domain identification within the amino acid sequence of each gene was performed using ScanProsite (https://prosite.expasy.org/scanprosite/). To identify potential K-domains, position of coiled coils was first predicted using DeepCoil2 (https://toolkit.tuebingen.mpg.de/tools/deepcoil2; Gabler et al. (2020)), after which K- domains were identified manually. Secondary structures in the amino acid sequence were predicted using Quick2D

(https://toolkit.tuebingen.mpg.de/tools/quick2d; Gabler et al. (2020)).

3.2.3.3 Analysis of *ODDSOC* gene structure and sequence motif identification

To verify the amino acid sequence provided by EnsemblPlants, the CDS was input into the Expasy translation tool (https://web.expasy.org/translate/) to confirm that the sequence was in frame. Next, the list of *ODDSOC* genes was input into BioMart (https://plants.ensembl.org/biomart/) to collect additional gene features. The 2.5 kb region upstream of the transcription start site (TSS) was analysed using the PlantPan 3.0 promoter analysis tool (http://plantpan.itps.ncku.edu.tw/; Chow et al. (2018)). Based on the output of the promoter analysis, predicted transcription factor binding sites (TFBS) were filtered depending on transcription factor (TF) family and prioritised with respect to genes involved in the flowering and/or vernalization response. Promoter analyses were compared between *ODDSOC* genes to identify predicted TFBS that were conserved throughout the group. The first intron of each *ODDSOC* was analysed using the same approach.

Predicted epigenetic modifications in genes were visualised using the JBrowse wheat browser, available via the WheatOmics 1.0 platform (http://wheatomics.sdau.edu.cn/; Ma et al. (2021)). This browser uses cv. Chinese Spring as a reference genome. Histone modification tracks for H3K27ac, H3K27me1, H3K4me1, H3K4me3, and H3K9ac were selected to identify regions of the chromatin associated with up-regulation. In addition, the long non-coding RNA (lncRNA) track was selected and filtered by antisense transcripts enriched at the 5’ end of the gene and the first intron.

3.2.4 Cloning of *ODDSOC* genes

3.2.4.1 Isolation of genes from wheat DNA

The complete set of *ODDSOC* genes cloned is listed in **Table 3.2**. To clone *TraesCS3A02G432900* and *TraesCS3A02G434400*, template cDNA from the winter cv. Claire was used. In preparation, the CDS for each *ODDSOC* gene in the reference genome for cv. Chinese Spring was aligned to cv. Claire scaffold (release ver. Elv1.1) using the BLASTn feature of EnsemblPlants (https://plants.ensembl.org/Multi/Tools/Blast) and specific forward and reverse primers were designed to amplify the complete CDS. The full list of primers is listed in **Table 3.3**. To clone the CDS, cDNA of cv. Claire was used. From previous investigation in chapter 2 section **2.3.1.3**, leaf cDNA from 4 weeks of growth under short day/10°C conditions was used as qPCR analysis indicated high levels of transcript for several *ODDSOC* genes at this stage of development (see **Figure 3.19**). RNA was extracted from leaf tissue using the Spectrum™ Plant Total RNA Kit as per the manufacturer's instructions. Next, cDNA was synthesized from the extracted RNA as described in section **2.2.4.1**. *TraesCS3A02G435000*, *TraesCS3B02G470000* and *TraesCS3D02G428000* were synthesised using the service provided by Genewiz (https://www.genewiz.com/en-GB), now part of Azenta (https://www.azenta.com/), as gene-specific primers couldn't de designed due to identical coding sequences at the start and end of the gene. All genes were amplified using Q5® High-Fidelity DNA Polymerase (NEB) and the following annealing temperatures: *TraesCS3A02G432900*, 64 °C; *TraesCS3A02G434400*, 65 °C. Polymerase Chain Reaction (PCR) conditions described in detail in **Table 3.4**.

Following PCR, reactions were ran on a 1% agarose gel. The gel was comprised of molecular-grade agarose and 1x TAE (Invitrogen; 40 mM Tris-acetate, 10 mM EDTA) buffer. GelRed ® Nucleic Acid Stain (Biotium) at 1X final concentration (i.e., 5 µL of 10,000X GelRed in 50 mL TAE buffer) was used to visualise the product on the gel. Band size was inferred using a 100 bp DNA ladder (NEB) and bands of correct size were excised from the gel and the product was extracted using Monarch ® DNA Gel Extraction Kit (NEB) following the manufacturer's instructions and eluted in 20 µL dH_2O.

Table 3.1: *ODDSOC* genes cloned in this project.

Gene	Source	Final vector	Vector description
TraesCS3A02G432900	cv. Claire cDNA	pGAD424; pGBT9	yeast
TraesCS3A02G434400	cv. Claire cDNA	pGAD424; pGBT9	yeast
TraesCS3A02G435000	pUC57 plasmid	pGAD424; pGBT9	yeast
TraesCS3B02G470000	pUC57 plasmid	pGAD424; pGBT9	yeast
TraesCS3D02G428000	pUC57 plasmid	pGAD424; pGBT9	yeast

3.2.4.2 Cloning genes into the pGEM®-T Easy Vector

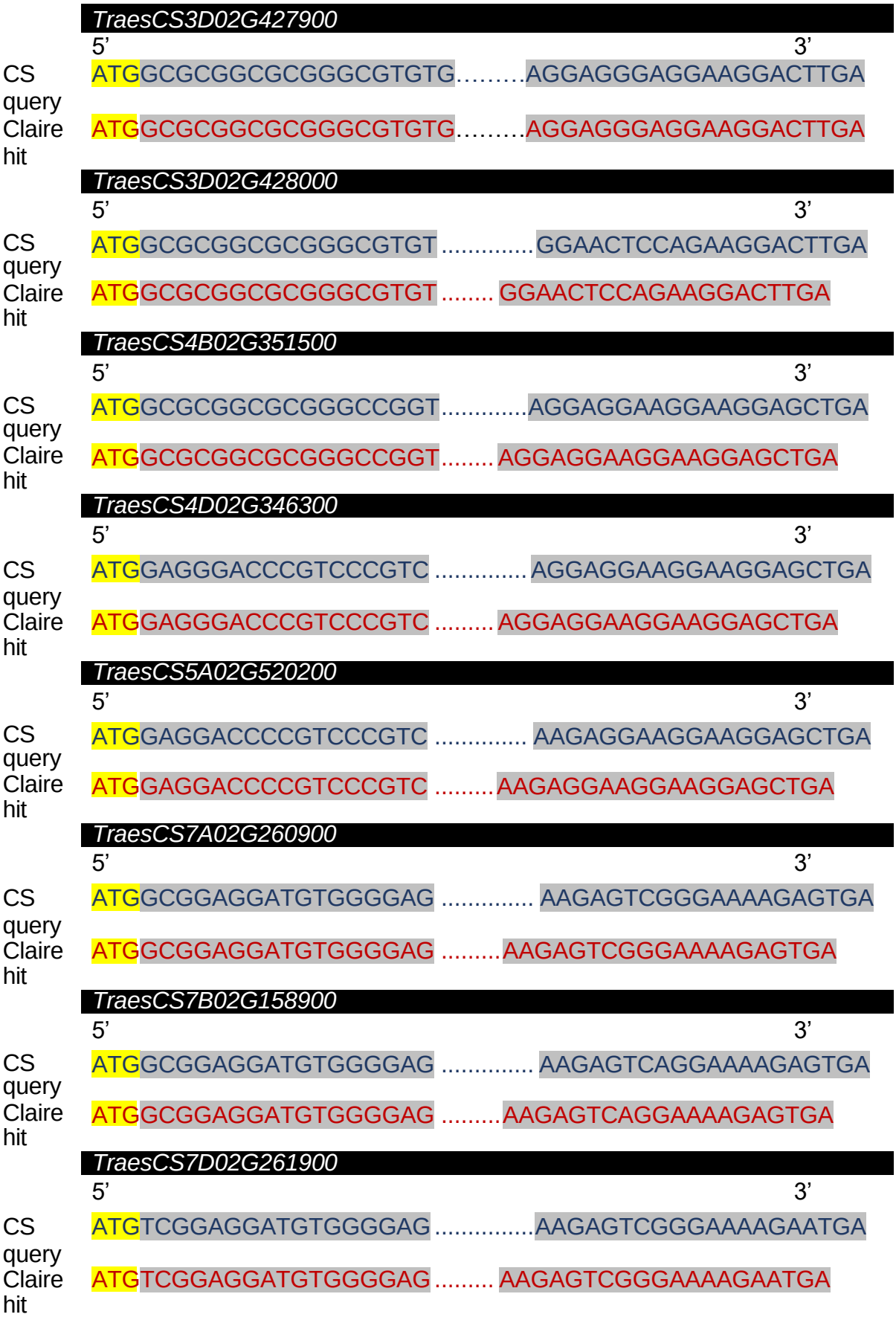
Once cloned, *TraesCS3A02G432900*, *TraesCS3A02G434400*, and *TraesCS3D02G428000* genes were transferred into a vector to ensure stability and facilitate ease of transfer to subsequent vectors. The pGEM®-T Easy Vector Systems (Promega) was used. Prior to ligation into the pGEM vector, a pre-cursor step was carried out to add extra 3' adenines to the blunt-end PCR fragment. Taq polymerase (Jena Biosciences), 10X Crystal buffer (Jena Bioscience), and 10 mM dATPs (NEB) were added to a 0.2 ml PCR tube containing the purified PCR fragment and incubated for 30 mins at 72 °C. This is a variation of the protocol described in **Table 3.4**. Next, a ligation reaction was performed. The reaction containing the PCR fragment, now with additional 3' adenines, was briefly centrifuged to gather the contents at the bottom of the tube. A ligation reaction was set up following the NEB T4 DNA Ligase protocol, maintaining a 1:3 ratio of pGEM® vector to PCR product. To ensure maximum efficiency, the ligation reaction was incubated at 4 °C overnight.

Next, the ligated product and vector was transformed into NEB® 10-beta Competent *E. coli* (High Efficiency) cells. Competent cells were gently thawed on ice and once thawed, 25 µL of cells were carefully transferred to a sterile 1.5 ml microcentrifuge tube. The ligation reaction was centrifuged and 1 µL was added to the tube containing the cells. The mixture was then kept on ice for 30 mins. After incubation on ice, the tube was placed in a water bath for 47 secs at 42 °C to heat-shock the cells. The tube was returned to ice for 2 mins. Next, 450 µL of NEB® 10-beta/Stable Outgrowth Medium was added to the transformation. This was incubated at 37 °C with shaking (~200 rpm) for 60 mins. To aid selection of successfully transformed bacterial colonies, a blue-white screening technique was applied. This technique is compatible with the pGEM®-T Easy Vector System, as the pGEM® vector contains a multiple cloning site within the α-peptide coding region of the enzyme β-galactosidase. Therefore, successful recombination with the GOI within this region interrupts the enzyme activity and enables selection of white colonies.

Table 3.2: Alignment of *ODDSOC* CDS in the cv. Chinese Spring (CS) reference genome with the cv. Claire scaffold Elv1.1. Only the region at the start and end of the CDS alignment (i.e., where the primers bind) is depicted.

	TraesCS3A02G432900
	5’ 3’
CS query	ATGGCGCGGCGCGGACGGGT..... AGAATGGGAGGAAGGCTTGA
Claire hit	ATGGCGCGGCGCGGGCGGGT ..... AGAATGGGAGGAAGGCTTGA
	TraesCS3A02G434400
	5’ 3’
CS query	ATGGCGCGGCGCGGGCGTGT..... AGAAGCGAGGAAGGACCTGA
Claire hit	ATGGCGCGGCGCGGGCGTGT ..... AGAAGCGAGGAAGGACCTGA
	TraesCS3A02G434900
	5’ 3’
CS query	ATGGCGCGGCGCGGGCGTGT..... CTGGAAGGAGGAAGGACTTGA
Claire hit	ATGGCGCGGCGCGGGCGTGT .....CTGGAAGGAGGAAGGACTTGA
	TraesCS3A02G435000
	5’ 3’
CS query	ATGGCGCGGCGCGGGCGTGT..... AGAAGGAAGGAAGGACTTGA
Claire hit	ATGGCGCGGCGCGGGCGTGT ..... AGAAGGAAGGAAGGACTTGA
	TraesCS3B02G469700
	5’ 3’
CS query	ATGGCGCGGCGCGGGCGTGT..... AGAAGCGAGGAAGGACTTGA
Claire hit	ATGGCGCGGCGCGGGCGTGT ..... AGAAGCGAGGAAGGACTTGA
	TraesCS3B02G470000
	5’ 3’
CS query	ATGGCGCGGCGCGGGCGTGT..... AGAAGGAAGGAAGGACTTGA
Claire hit	ATGGCGCGGCGCGGGCGTGT ..... AGAAGGAAGGAAGGACTTGA
	TraesCS3D02G427700
	5’ 3’
CS query	ATGGCGCGGCGCGGGCGTGT..... AGAAGCGAGGAAGGACCTGA
Claire hit	ATGGCGCGGCGCGGGCGTGT ..... AGAAGCGAGGAAGGACCTGA

Table 3.2: Alignment of *ODDSOC* CDS in the cv. Chinese Spring (CS) reference genome with the cv. Claire scaffold Elv1.1, continued.



To select on plates, 250 µL of Isopropyl β-D-1-thiogalactopyranoside (IPTG, 100 mM) and 250 µL X-Gal solution (20 mg/mL) was added to 25 mL LB agar to make a single 25 mL plate. After incubation, 100 µL of the transformation was plated on a LB/ampicillin/IPTG/X-Gal agar plate. The plate was incubated upside-down overnight at 37 °C and when colonies appeared, white colonies were selected. To isolate the vector containing the GOI from the cells, individual colonies were pricked and cultured overnight in LB and selective antibiotic (ampicillin or kanamycin), where the final concentration of ampicillin was 100 µg/ml and kanamycin was 50 µg/ml. In parallel, glycerol stocks were made from the same culture: 250 µL of an 80% glycerol stock was added to 750 µL of the culture so that the final concentration was preserved in 20% glycerol. Stocks were rapidly cooled in a dry ice ethanol bath and stored at -80 °C.

Colonies were isolated using the NEB® Plasmid Miniprep kit following the manufacturer's instructions. Plasmid purity and concentration was determined using a NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific). To confirm the presence of the GOI in the pGEM® vector, the plasmid was diluted if necessary to a concentration of 30-100 ng/µL and sequenced using Sanger sequencing (Eurofins; Genewiz) and the appropriate sequencing primer (**Table 2.1**).

Table 3.3: Primers used to amplify target *ODDSOC* CDS.

Gene	Orientation	Sequence	Q5 annealing tm (°C)	CS product size (bp)
TraesCS3A02G432900	Forward	ATGGCGCGGCGCGGACGGGT		
	Reverse	TCAAGCCTTCCTCCCATTCT	67	501
TraesCS3A02G434400	Forward	ATGGCGCGGCGCGGGCGTGT		
	Reverse	TCAGGTCCTTCCTCGCTTCT	69	504
TraesCS3A02G434900	Forward	ATGGCGCGGCGCGGGCGTGT		
	Reverse	TCAAGTCCTTCCTCCTTCCAG	67	504
TraesCS3A02G435000	Forward	ATGGCGCGGCGCGGGCGTGT		
	Reverse	AGTCCTTCCTTCCTTCTGTC	64	504
TraesCS3B02G469700	Forward	ATGGCGCGGCGCGGGCGTGT G		
	Reverse	TCAAGTCCTTCCTCGCTTCTTC	68	504
TraesCS3B02G470000	Forward	ATGGCGCGGCGCGGGCGTGT		
	Reverse	TCAAGTCCTTCCTTCCTTCT	63	504
TraesCS3D02G427700	Forward	ATGGCGCGGCGCGGGCGTGT		
	Reverse	TCAAGTCCTTCCTCGCTTCT	66	501
TraesCS3D02G427900	Forward	ATGGCGCGGCGCGGGCGT GT		
	Reverse	TCAAGTCCTTCCTCCCTCCTG	69	504

Table 3.3: Primers used to amplify target *ODDSOC* CDS (continued).

Gene	Orientation	Sequence	Q5 annealing tm (°C)	CS product size (bp)
TraesCS3D02G428000	Forward	ATGGCGCGGCGCGGGCGTGT		
	Reverse	AGTCCTTCTGGAGTTCCCG	67	480
TraesCS4B02G351500	Forward	ATGGCGCGGCGCGGGCCGGT		
	Reverse	TCAGCTCCTTCCTTCCTCC	67	495
TraesCS4D02G346300	Forward	ATGGAGGGACCCGTCCCGTC		
	Reverse	TCAGCTCCTTCCTTCCTCCTG	70	624
TraesCS5A02G520200	Forward	ATGGAGGACCCCGTCCCGTC		
	Reverse	TCAGCTCCTTCCTTCCTCTT	66	609
TraesCS7A02G260900	Forward	ATGGCGGAGGATGTGGGGAG		
	Reverse	TCACTCTTTTCCTGACTCTT	61	654
TraesCS7B02G158900	Forward	ATGTGGGGAGGAGGAGGATG		
	Reverse	TCACTCTTTTCCCGACTCTTG	65	618
TraesCS7D02G261900	Forward	ATGTCGGAGGATGTGGGGAG		
	Reverse	TCATTCTTTTCCCGACTCTTGG	65	660

Table 3.4: PCR reaction set up and cycling conditions to amplify ODDSOC genes from wheat cDNA. Asterisk indicates range in annealing temperature based on specific *ODDSOC* gene.

Reaction set up:

Component	50 µL reaction volume	Final concentration
5X Q5 Reaction Buffer	10 µL	1X
10 mM dNTPs	1 µL	200 µM
10 µM Forward Primer	2.5 µL	0.5 µM
10 µM Reverse Primer	2.5 µL	0.5 µM
Template cDNA	1.5 µL	~500 ng
Q5 High-Fidelity DNA Polymerase	0.5 µL	0.02 U/µL
5X Q5 High GC Enhancer	10 µL	1X
Nuclease-Free Water	5 µL	-

Thermocycling conditions:

Step	Temperature	Duration
Denaturation	98 °C	30 secs
Annealing	98 °C	5 secs
	*64-65 °C	30 secs
	72 °C	45 secs
Extension	72 °C	2 mins

3.2.5 Yeast two-hybrid analysis

3.2.5.1 Gene transformation into yeast

To investigate the binding of *ODDSOC* proteins to other proteins, a GAL4-based two-hybrid system was used to detect protein interactions *in vivo* in yeast. Genes were cloned into the yeast vectors pGAD424 and pGBT9 (both Clontech) for expression in a yeast two-hybrid system (Fields & Song, 1989). Here, the Matchmaker™ GAL4 Two-Hybrid System 2 (Clontech) was used. This system uses a plasmid containing a yeast GAL4 activation domain (pGAD424, target) and a yeast GAL4 DNA-binding domain (pGBT9, bait). When genes expressed in the target and the bait plasmids interact, the activation and DNA-binding domains are brought in proximity and transformed yeast can be screened on selective media. The yeast plasmids contain the ampicillin resistance gene *bla* and either the *LEU2* (pGAD424) or the *TRP1* (pGBT9) nutritional genes, enabling selection when grown on media supplemented with ampicillin and lacking specific amino acids. Full description of this protocol is described in the manufacturer's manual.

To transport the coding region from either the pGEM vector or the pUC-57 construct into the yeast vectors, a double restriction digest was performed. Primers were designed to attach the restriction enzyme digest sites for *Eco*RI (GAATTC) and *Bam*HI (GGATCC) along with random spacer flanking base pairs to the 5' and 3' flanking ends of the gene coding region and a PCR was carried out. PCR cycling conditions and primers are outlined in **Table 3.4**. Fragments were amplified, the product was run on a gel, and extracted as described in section **3.2.3.1**. Once extracted, the fragment was digested with *Eco*RI-HF® and *Bam*HI-HF® (both NEB) as following the NEB double-digest manufacturer's protocol. The yeast vectors pGAD424 and pGBT9 were digested in the same fashion using the same restriction enzymes.

To ensure successful ligation of the digested fragment into the vector, the digested vector was run on an agarose gel and the linearised plasmid was extracted. Then, both the digested fragment and linearised plasmid were ligated together following the ligation reaction method described in section **3.2.3.2**. Subsequently, the ligated product was transformed into NEB® 10-beta Competent *E. coli* (High Efficiency)

cells as described in section **3.2.3.2**. The transformant was plated on LB agar plates supplemented with ampicillin (100 µg/mL). Plates were incubated upside-down overnight at 37 °C. Colonies were selected, cultured overnight in LB and ampicillin (100 µg/mL), and plasmids isolated using the Monarch® Plasmid Miniprep Kit (NEB) and eluted in 50 µL elution buffer. Plasmid concentration was determined using the NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific). The presence of the correct insert in the vector was confirmed by sequencing, as described in section **3.2.3.2**. Sequencing primers are described in **Table 2.1**.

Next, pGAD424 and pGBT9 vectors containing the inserted *ODDSOC* CDS were transformed into the AH109 *Saccharomyces cerevisiae* yeast strain. Prior to transformation, AH109 yeast was cultured on yeast peptone dextrose adenine (YPDA) media plates for several days to ensure viable yeast. 1L YPDA media contained the following: 10g Bacto yeast extract, 20 g Bacto peptone, 20g Glucose monohydrate, 40 mg Adenine hemisulfate, 20 g Bacto Agar, and 800 ml dH_2O. Media was autoclaved at 121°C for 15 mins minimum before use. The yeast transformation protocol used is adapted from the method described by Amberg et al. (2006). Briefly, deoxyribonucleic acid sodium salt type III from salmon testes was denatured by boiling at 100 °C for 5 mins; next, 1M sterile lithium acetate, sterile polyethylene glycol (PEG)-3350 50% w/v, and β-mercaptoethanol were combined to make a transformation master mix. The denatured deoxyribonucleic acid sodium salt, transformation master mix, plasmid to transform, and a large AH109 yeast colony were thoroughly mixed in a sterile 1.5 mL microcentrifuge tube. These steps were carried out on ice, under sterile conditions. Once sufficiently mixed, the transformant was incubated with shaking at 37 °C for 45 mins. After incubation, the transformant was plated on yeast synthetic drop-out media lacking either the amino acid leucine (pGAD424) or tryptophan (pGBT9). Plates were incubated upside-down without shaking at 29 °C for up to 7 days. To test interaction between ODDSOC proteins and VRN-A1, the *VRN-A1* and the *VRN-A1* exon 4 SNP genes were also cloned into yeast plasmids and generously provided by Dominique Hirsz.

3.2.5.2 Validation of yeast co-transformation to determine strength of interaction

Vectors containing the GOI were co-transformed following the adapted yeast transformation described above (section **3.2.3.1**). To test for the formation of heterodimers, an AH109 yeast strain transformed with pGAD424:gene-1 was co-transformed with the pGBT9:gene-2, with gene-1 and gene-2 being the protein-protein interaction tested. For testing the formation of homodimers, the co-transformation follows the format pGAD424:gene-1 and pGBT9:gene-1. Once co-transformed, the transformant was plated on yeast synthetic drop-out media lacking the amino acids leucine and tryptophan for low stringency screening of protein-protein interaction. Plates were incubated upside-down at 29 °C until colonies appeared.

To eliminate any false-positives from the low-stringency screening, the co-transformed yeast colonies were plated on synthetic drop-out media lacking leucine, tryptophan, adenine, and histidine. This medium is auxotrophic for yeast lacking pGAD424 (LEU1), pGBT9 (TRP2), and a protein-protein interaction that activates expression of adenine-synthesising gene *ADE2* and histidine-synthesising gene *HIS3*. Growth of yeast colonies on this media confirmed protein-protein interaction and the strength of the interaction was further tested by supplementing the media lacking leucine, tryptophan, adenine, and histidine with 3-amino-1,2,4-triazole (3-AT). 3-AT is a competitive inhibitor of the HIS3 protein and was added to media in concentrations of 2.5 mM, 5 mM, and 10 mM.

3.2.6 Watkins collection

To investigate *ODDSOC* gene allelic diversity in a large wheat population, the Watkins bread wheat landrace collection was used (Winfield et al., 2018; Wingen et al., 2014; Wingen et al., 2017). Access to this short-read sequenced collection was provided pre-publication by the John Innes Centre, Norwich and facilitated by Dr Simon Griffiths. Information for each line, including the Germplasm Resource Unit (GRU) ID and source country, is available from the Watkins Historic Collection of Landrace Wheats portion of the SeedStor site (https://www.seedstor.ac.uk/search-browseaccessions.php?idCollection=4). Haplotype maps for each of the *ODDSOC* genes were provided by Shifeng Chen (Agricultural Genomics Institute at Shenzhen,

Chinese Academy of Agri Sciences) and Dr Simon Griffiths as part of the Wat-Seq initiative.

3.2.7 Identification of *ODDSOC*-associated non-coding transcripts

Identification of long non-coding RNA (lncRNA) transcripts from the 10 °C and 14 °C RNA-seq dataset described in section **2.3.1.4** was carried out by Wilfried Haerty at the Earlham Institute, Norwich, UK. Each transcript was submitted to the BLASTn to identify which *ODDSOC* gene it was associated with. The 2.k kb region downstream of the associated gene was processed using MEME (https://meme-suite.org/meme/) to identify potential CArG-box sites from which the transcripts were transcribed. To investigate the presence of transcripts in samples, qPCR primers were designed to amplify a region common to each lncRNA transcript population (**Table 2.1**). qPCR using cDNA was performed as described in section **2.2.4.1**. To confirm the presence of the *OS1*-associated lncRNA transcripts, the same qPCR primers were used to amplify a band from cDNA via PCR assay and subsequent gel electrophoresis. The excised band was extracted, ligated into pGEM® and transformed into NEB® 10-beta Competent *E. coli* (High Efficiency) cells. The transformant was plated on LB agar supplemented with ampicillin, X-GAL, and IPTG. Positive transformed colonies were selected and cultured overnight in LB supplemented with ampicillin. Plasmids were isolated from the culture using Monarch® Plasmid Miniprep Kit and sequenced using Sanger sequencing. Full details of this cloning process are described in section **3.2.3.2**.

3.2.8 Statistical analysis

Statistical analysis was performed using either Student's t-test (normal distribution) or Mann-Whitney U test (non-normal distribution) when comparing two groups, and ANOVA followed by the Tukey post-hoc test for data with three or more groups to compare.

3.3 Results

3.3.1 Multiple *OS2* genes are identified in the wheat genome

The investigation into *OS2* genes was initiated by results from a flowering eQTL analysis (Barratt, He, et al., 2023; Barratt, Reynolds, et al., 2023) which had been performed to infer genes involved in the vegetative to reproductive transition. This analysis was conducted in bread wheat that had a spring background, linking gene expression with the floral development phenotype. A panel of 95 spring wheat cultivars was used and GEM analysis was performed. This analysis was carried out by Dr Andrea Harper at the University of York.

GEM analysis identified seven genes out of >45,000, indicating these candidates were involved in the floral development response. BLASTx results indicated that six of the seven genes were predicted to be MADS-box genes (*MADS-box transcription factor 51*-like and *TaAGL33*, respectively) which had not previously been well-characterised (**Table 3.5**). These genes were: *TraesCS3A02G435000*, *TraesCS3B02G470000*, *TraesCS3D02G428000*, *TraesCS4B02G351500*, *TraesCS4D02G346300*, and *TraesCS5A02G520200*. Ensembl Plants had annotated *TraesCS3A02G435000*, *TraesCS3B02G470000*, *TraesCS3D02G428000* as homoeologues and *TraesCS4B02G351500*, *TraesCS4D02G346300*, and *TraesCS5A02G520200* as homoeologues. Alignment of the protein sequences of these genes revealed a high level of similarity within homoeologous triads: *TraesCS3A02G435000* shares 96.4% sequence similarity with *TraesCS3B02G470000* and 89.9% with *TraesCS3D02G428000*, whereas *TraesCS4B02G351500* shares 90.9% similarity with *TraesCS4D02G346300* and 92.1% with *TraesCS5A02G520200*. A summary of these results in relation to *OS1* and *OS2* in cereals is shown in **Table 3.6**.

3.3.2 Pathway networks and phylogenetics can be used to identify putative additional *OS2* genes

To uncover more information about these genes, all six were input into the Reaction Pathway Browser to investigate whether these genes had been categorised as part of a plant reaction pathway. This uncovered the involvement of the six genes in a ‘short-day regulated expression of florigen’ pathway, occupying the role of

orthologues of rice gene *OsMADS51*. This sub-group within the pathway consisted of six additional genes: *TraesCS3A02G432900*, *TraesCS3A02G434400*, *TraesCS3A02G434900*, T r a e s C S *3B02G469700*, *TraesCS3D02G427700*, and *TraesCS3D02G427900*. When these gene were analysed using Ensembl Plants, it was uncovered the following homologous groups: *TraesCS3A02G434400, TraesCS3B02G469700*, and *TraesCS3D02G427700*; *TraesCS3A02G434900* and *TraesCS3D02G427900*.

To validate the relationship between these 12 genes (6 original genes derived from the flowering eQTL and the 6 additional genes from the 'short-day regulated expression of florigen' pathway), the gene list was input into Knetminer. This revealed the involvement of all 12 genes in 4 discrete pathways: Growth and developmental processes, Reproductive structure development, Inflorescence development, and, as previously mentioned, Short-day regulated expression of florigen. A condensed version of the overlapping relationships of the putative *OS2* genes in these four pathways is presented in **Table 3.7**.

Next, the translated protein sequences of all 12 genes were scanned using BLASTn. This confirmed that all 12 genes had the closest match with each other and that no additional genes with high amino acid sequence similarity had been omitted from the list. As well as the top 12 most similar genes, 4 additional genes were within 89.8% sequence similarity. These genes were *TraesCS7A03G0604100*, *TraesCS7B03G0432700*, *TraesCS7D03G0591000*, and *TraesCS3A02G434700*. To determine whether to include these 4 genes in further *ODDSOC* analysis and to explore how closely these genes were related over an evolutionary timescale, a phylogenetic tree was generated, as described in materials and methods section (**Figure 3.1**). Using *TraesCS3A02G435000* as a seed gene, the tree re-affirmed the remaining 11 genes as the closest neighbours of *TraesCS3A02G435000* (70% confidence node representing linkage between branches).

Table 3.5: Summary table of the *OS2* candidate genes identified in GEM and eQTL analysis.

Marker (IWGSC RefSeq v1.1)	FDR score (GEM)	FDR-corrected P-value (eQTL)	BLASTx functional annotation best match
TraesCS3D02G428000.1	0.003888937	7.12E-05	MADS-box transcription factor 51-like [Aegilops tauschii subsp. tauschii]
TraesCS3A02G435000.1	0.004277599	0.001361929	MADS-box transcription factor TaAGL33 [Triticum aestivum]
TraesCS4D02G346300.1	0.003807903	1.39E-13	MADS-box transcription factor 51-like [Aegilops tauschii subsp. tauschii]
TraesCS5A02G520200.1	0.008759601	1.19E-07	MADS-box transcription factor 51-like [Aegilops tauschii subsp. tauschii]
TraesCS3B02G470000.1	0.014998818	0.002318561	MADS-box transcription factor TaAGL33 [Triticum aestivum]
TraesCS4B02G351500.1	0.01287821	7.74E-09	MADS-box transcription factor 51-like [Aegilops tauschii subsp. tauschii]

Table 3.6: Summary of *ODDSOC* genes named in the literature.

Gene	*T. aestivum*	*B. distachyon*	*H. vulgare*	*O. sativa*
OS1	*TaMX23* (Trevaskis et al., 2003), *TaAGL42* (Winfield et al., 2009)	*BdOS1* (Ruelens et al., 2013)	*HvOS1* (Greenup et al., 2010)	*OsMADS51* (Zhao et al., 2006)
OS2	*TaAGL33* (Winfield et al., 2009)	*BdOS2* (Ruelens et al., 2013)	*HvOS2* (Greenup et al., 2010)	*OsMADS51* (Zhao et al., 2006)

Next, the translated protein sequences of all 12 genes were scanned using BLASTn. This confirmed that all 12 genes had the closest match with each other and that no additional genes with high amino acid sequence similarity had been omitted from the list. As well as the top 12 most similar genes, 4 additional genes were within 89.8% sequence similarity. These genes were *TraesCS7A03G0604100*, *TraesCS7B03G0432700*, *TraesCS7D03G0591000*, and *TraesCS3A02G434700*. To determine whether to include these 4 genes in further *ODDSOC* analysis and to explore how closely these genes were related over an evolutionary timescale, a phylogenetic tree was generated, as described in materials and methods section (**Figure 3.1**). Using *TraesCS3A02G435000* as a seed gene, the tree re-affirmed the remaining 11 genes as the closest neighbours of *TraesCS3A02G435000* (70% confidence node representing linkage between branches). Branches represent the level of relatedness, with homoeologues sharing the closest level of similarity. Based on the phylogenetic tree, the 4 additional genes (*TraesCS7A03G0604100*, *TraesCS7B03G0432700*, *TraesCS7D03G0591000*, and *TraesCS3A02G434700*) had close proximity to the already identified 12 genes based on an 82% confidence level of similar phylogeny. Of the additional 4 genes, *TraesCS7A03G0604100*, *TraesCS7B03G0432700*, and *TraesCS7D03G0591000* were identified as homoeologues via Ensembl Plants. No known homoeologues were identified for *TraesCS3A02G434700.* The complete set of 16 genes is listed in **Table 3.7**.

3.3.3 Analysis of *ODDSOC* gene structure and sequence motif identification

The identification of a potentially expanded *ODDSOC* gene family raised the question of how closely related these genes were and how many shared high similarity to the previously identified *OS1* and *OS2* in barley and *B. distachyon*.

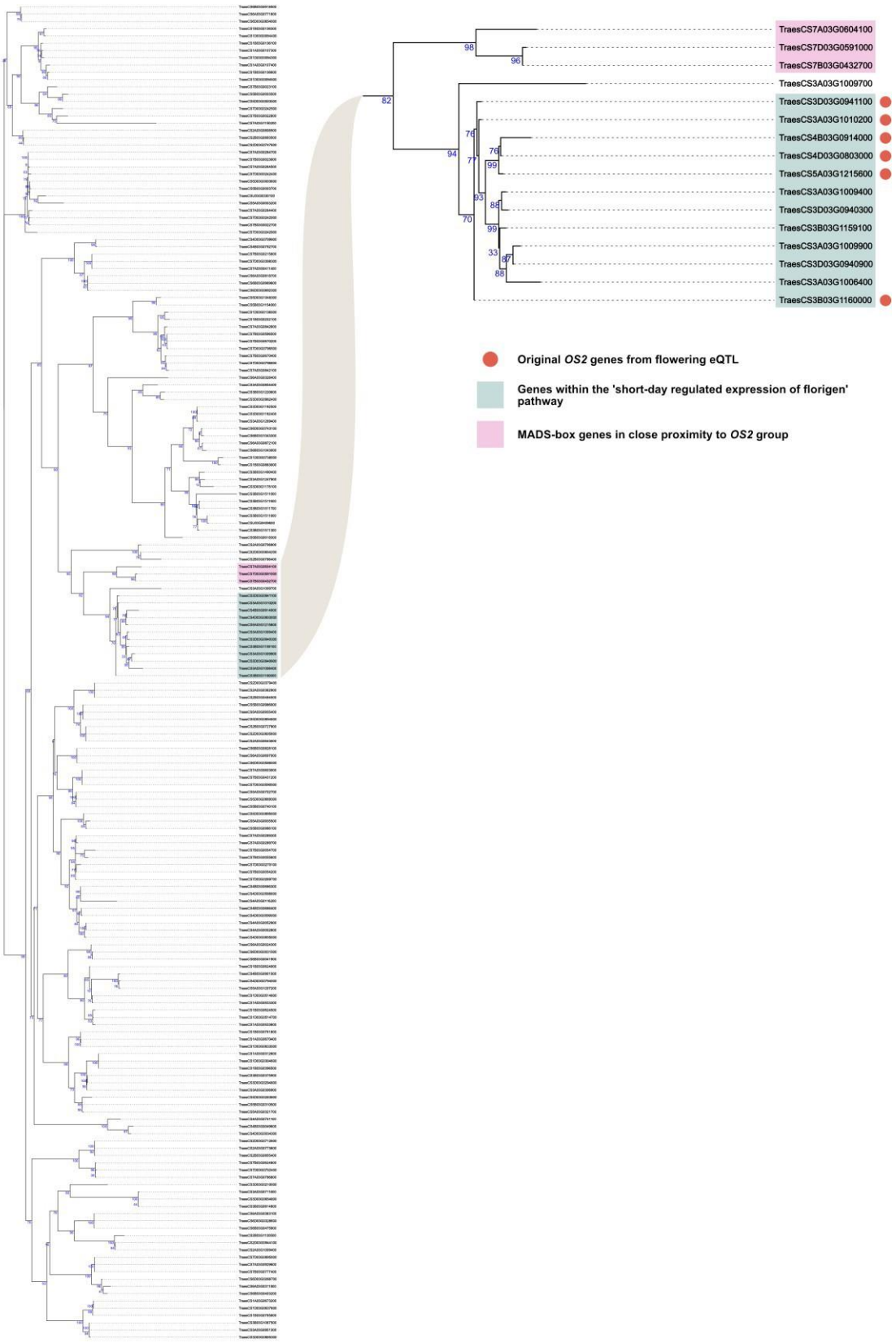


Figure 3.1: Phylogenetic tree based on *ODDSOC* input seed genes. Tree generated using MUSCLE MSA of protein sequences and FastTree (https://github.com/PavelTorgashov/FastTree) and Monocots PLAZA 5.0.

Table 3.7: *ODDSOC* candidate genes based on in-silico investigation.

IWGSC RefSeq v2.1 I.D.	IWGSC RefSeq v1.1 I.D.	Position (v2.1)	Homoeologues
TraesCS3A03G1006400	*TraesCS3A02G432900*	Chr3A:674409622..674433638 (+ strand)	
TraesCS3A03G1009400	*TraesCS3A02G434400*	Chr3A:676393319..676429534 (- strand)	
TraesCS3A03G1009900	*TraesCS3A02G434900*	Chr3A:677066707..677087044 (- strand)	
TraesCS3A03G1009700	*TraesCS3A02G434700*	Chr3A:677032258..677035388 (+ strand)	
TraesCS3A03G1010200	*TraesCS3A02G435000*	Chr3A:677322033..677352530 (- strand)	
TraesCS3B03G1159100	*TraesCS3B02G469700*	Chr3B:731424430..731456570 (- strand)	
TraesCS3B03G1160000	*TraesCS3B02G470000*	Chr3B:731996693..732027796 (- strand)	
TraesCS3D03G0940300	*TraesCS3D02G427700*	Chr3D:541989070..542021642 (- strand)	
TraesCS3D03G0940900	*TraesCS3D02G427900*	Chr3D:542195220..542227445 (- strand)	
TraesCS3D03G0941100	*TraesCS3D02G428000*	Chr3D:542421172..542455539 (- strand)	
TraesCS4B03G0914000	*TraesCS4B02G351500*	Chr4B:642788532..642791635 (+ strand)	
TraesCS4D03G0803000	*TraesCS4D02G346300*	Chr4D:500441774..500444389 (+ strand)	
TraesCS5A03G1215600	*TraesCS5A02G520200*	Chr5A:682786658..682789182 (+ strand)	
TraesCS7B03G0432700	*TraesCS7B02G158900*	Chr7B:229626646..229663264 (+ strand)	
TraesCS7A03G0604100	*TraesCS7A02G260900*	Chr7A:258977308..259018791 (+ strand)	
TraesCS7D03G0591000	*TraesCS7D02G261900*	Chr7D:240960460..240990049 (+ strand)	

Table 3.8: Abridged table of key genes of interest involved in Growth and developmental processes, Reproductive structure development, Inflorescence development, and Short-day regulated expression of florigen pathways. Sourced from Knetminer (https://knetminer.com/Triticum_aestivum/).

IWGSC RefSeq v1.1 I.D. (2018)	Gene	Pathway seed
Traescs3A02G435000	*Traescs3A02G435000*	yes
Traescs3B02G470000	*Traescs3B02G470000*	yes
Traescs3D02G428000	*Traescs3D02G428000*	yes
Traescs4B02G351500	*Traescs4B02G351500*	yes
Traescs4D02G346300	*Traescs4D02G346300*	yes
Traescs5A02G520200	*Traescs5A02G520200*	yes
Traescs3A02G434700	*AGL14*	no
Traescs3B02G612500	*AGL29*	no
Traescs7D02G388600	*AGL57*	no
Traescs3B02G612800	*AGL61*	no
Traescs2A02G174300	*AGL79*	no
Traescs3B02G612400	*AGL91*	no
Traescs3B02G612300	*AGL91*	no
Traescs2D02G181400	*AP1*	no
Traescs2B02G200800	*AP1*	no
Traescs6A02G214900	*BHLH128*	no
Traescs2A02G486100	*BZIP18*	no
Traescs1B02G091500	*BZIP53*	no
Traescs7A02G211300	*CO*	no
Traescs2D02G563400	*DOF3.4*	no
Traescs5D02G318000	*DREB1A*	no
Traescs5A02G310500	*DREB1B*	no
Traescs5D02G317800	*DREB1B*	no
Traescs5A02G310600	*DREB1C*	no
Traescs5D02G318100	*DREB1F*	no
Traescs1B02G235100	*DREB2B*	no

Table 3.8: Abridged table of key genes of interest involved in Growth and developmental processes, Reproductive structure development, Inflorescence development, and Short-day regulated expression of florigen pathways (continued).

IWGSC RefSeq v1.1 I.D. (2018)	Gene	Pathway seed
Traescs5A02G376500	*PIF4*	no
Traescs5B02G380200	*PIF5*	no
Traescs5B02G422000	*PIF5*	no
Traescs5A02G420200	*PIF5*	no
Traescs3A02G436100	*TCP16*	no
Traescs3B02G470700	*TCP6*	no
Traescs3D02G428700	*TCP6*	no
Traescs4D02G364500	*ZCCT1*	no

To investigate this, the CDS of each gene was aligned (**Figure 3.2**). In general, the complete CDS for all genes was between 309- and 660 bp in length. Conserved regions could be observed across all 16 genes; unsurprisingly, the greatest level of conservation occurred between homoeologues. The CDS alignment revealed several outliers – for example, *TraesCS3A02G434700* was considerably shorter than all other genes, with a CDS of 309 bp compared to an CDS between 480-660 bp for the other genes. *TraesCS3A02G434700* shared the highest level of similarity (93.7%) with *TraesCS3B02G470000* and *TraesCS3D02G428000*. However, *TraesCS3A02G434700* is in close proximity (32,140 bp) upstream of *TraesCS3A02G434900* (93% sequence similarity), indicating *TraesCS3A02G434700* possibly originated from a duplication event of *TraesCS3A02G434900*. On the other end of the scale, the homoeologues *TraesCS7A03G0604100*, *TraesCS7B03G0432700*, and *TraesCS7D03G0591000* were larger in length (618-660 bp) than all other genes, containing an additional regional of approximately 120 bp preceding the ATG start codon of most other genes. Interestingly, *TraesCS4D02G346300* and *TraesCS5A02G520200* also had a precursory region of approximately 120 bp, despite aligning with the ATG start site of the majority of the other genes. This possibly suggests a conserved regulatory region within this area.

A further indication of conserved regulatory regions was an unusually long intron of >17,000 bp common to 12 of the 16 putative *ODDSOC* genes (**Figure 3.3**). To investigate the level of conservation that was potentially present, an intron alignment for all genes was carried out. For *TraesCS3A02G432900*, *TraesCS3A02G434400*, *TraesCS3B02G469700*, *TraesCS3D02G427700*, *TraesCS3A02G434900*, *TraesCS3D02G427900*, *TraesCS3A02G435000*, *TraesCS3B02G470000*, and *TraesCS3D02G428000*, the first intron was the longest intron (**Figure 3.4**). However, for TraesCS7A02G260900, TraesCS7B02G158900, and *TraesCS7D03G0591000*, the largest intron was the second intron – therefore, this intron was selected for alignment. When the first intron of *TraesCS7A02G260900*, *TraesCS7B02G158900*, and *TraesCS7D02G261900* was analysed using BLASTx, the only matches were these same genes. This indicated that the first and second introns in *TraesCS7A02G260900*, *TraesCS7B02G158900*, and *TraesCS7D02G261900* weren't the result of division of the first intron of the other genes.

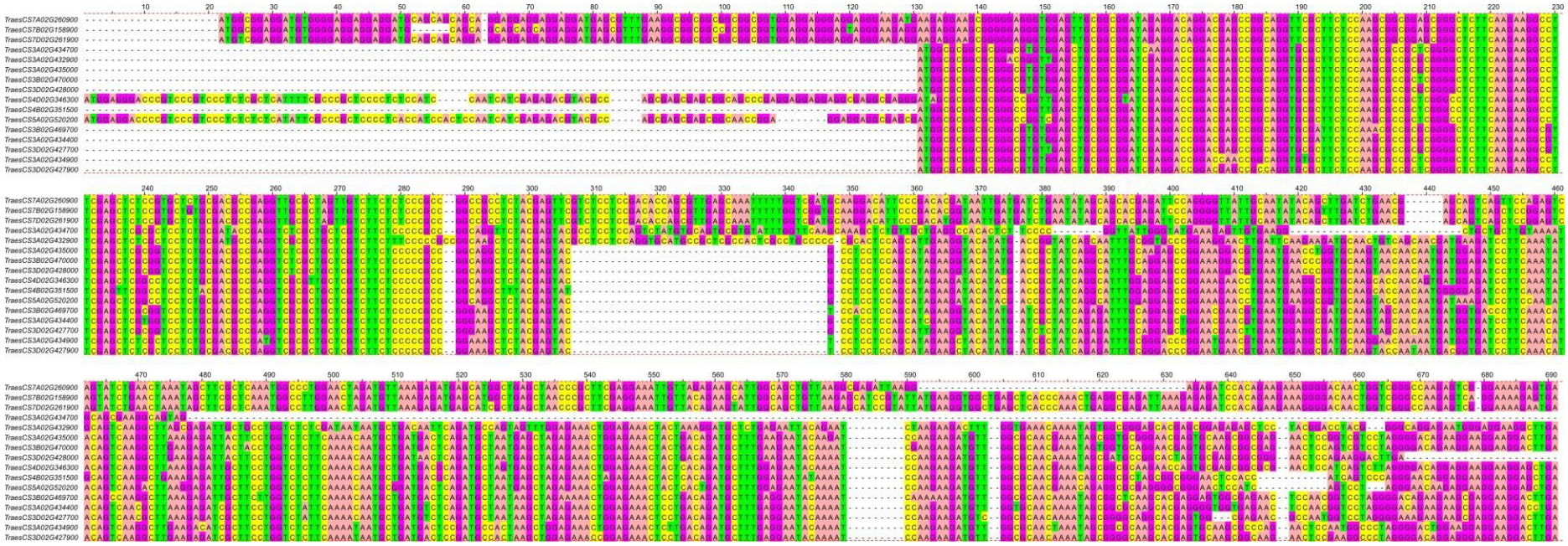


Figure 3.2: Alignment of the 16 *ODDSOC* genes CDS. The CDS were retrieved from Ensembl Plants and aligned using Clustal Omega.

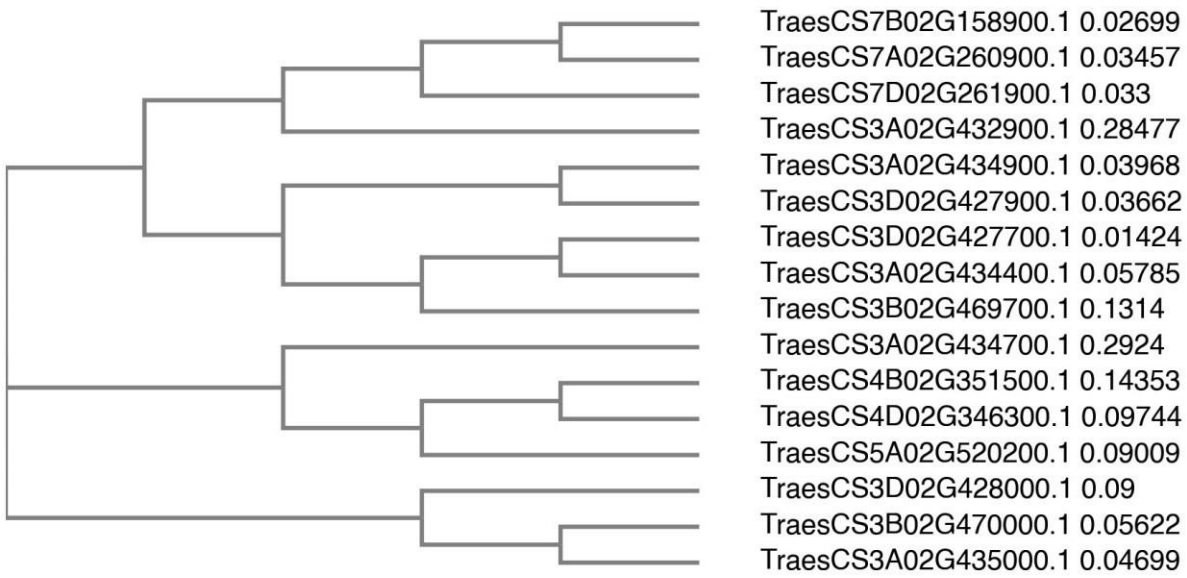


Figure 3.3: Phylogenetic tree of the alignment of introns for the 16 *ODDSOC* genes. The tree was generated using Clustal Omega; scores represent the sequence distance score calculated from the multiple sequence alignment (MSA).

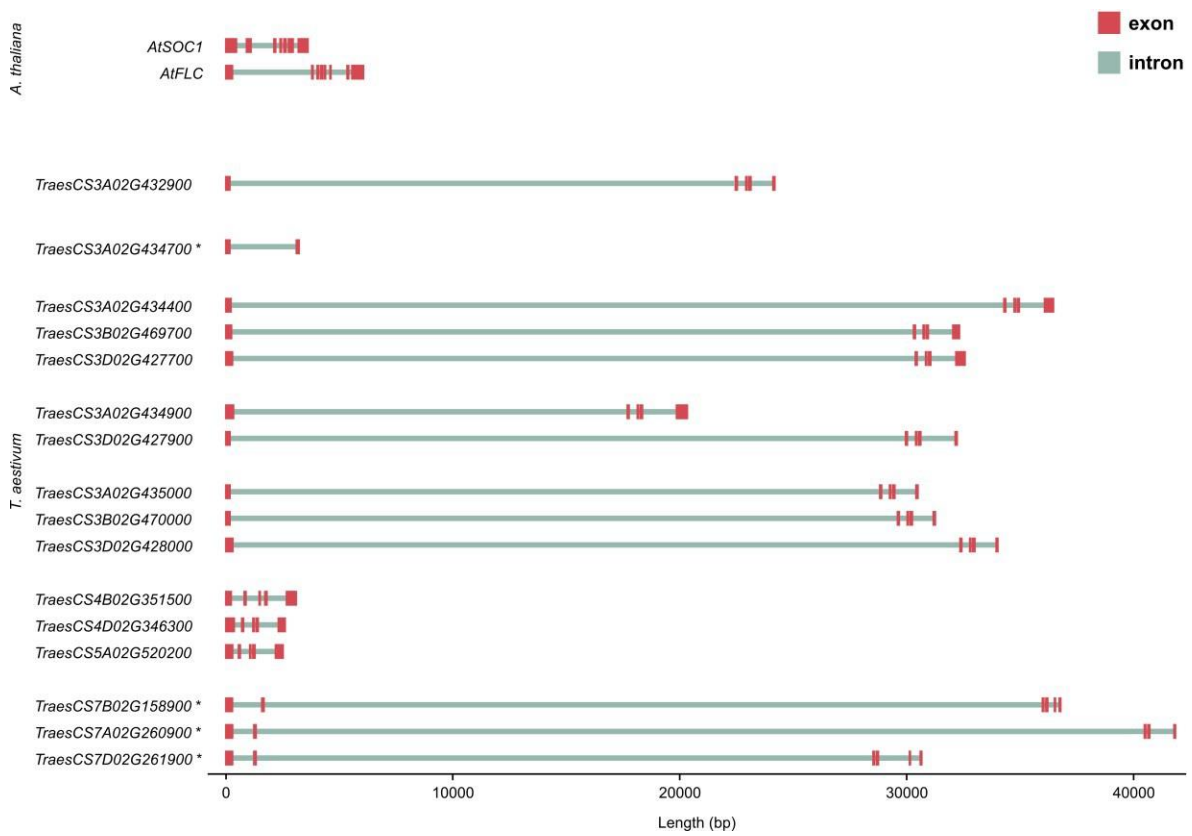


Figure 3.4: Gene structure of *ODDSOC* genes in bread wheat and *SOC1* and *FLC* in *Arabidopsis* (*A. thaliana*). Genes are scaled appropriately in relation to length in base-pairs. Red sections represent exons; green sections represent introns. Asterisks represent anomalies: *TraesCS3A02G434700* only contains a single intron, whereas *TraesCS7A02G260900*, *TraesCS7B02G158900*, and *TraesCS7D02G261900* have a large second intron instead of first intron.

Of the putative identified *ODDSOC* genes, *TraesCS3A02G434700*, *TraesCS4B02G351500*, *TraesCS4D02G346300*, and *TraesCS5A02G520200* didn't have an extended intron (**Figure 3.4**). In this case, the first intron of these genes was between 254 bp and 570 bp; these introns were also included in the alignment. Once aligned, a phylogenetic tree was generated to represent level of similarity between introns (**Figure 3.3**). The phylogenetic tree generated showed that the intron sequence was conserved to varying degrees between genes. Furthermore, conservation of introns corresponds with conservation of the CDS (**Figure 3.2**); this suggests that the CDS and first intron of each gene were subjected to the same selection pressure. This is most apparent for homoeologues, with the highest level of similarity observed between homoeologues for both the CDS and the first intron. Genes with the highest distance score (i.e., most dissimilar based on the output of the multiple sequence alignment) were *TraesCS3A02G434700* (0.2924) and *TraesCS3A02G432900* (score = 0.28477). *TraesCS3A02G434700* didn't share a significant level of similarity in CDS and intron sequence, didn't share structural similarity, and appeared to be a partial duplication of *TraesCS3B02G470000*; for this reason, it wasn't included in putative *ODDSOC* gene investigation beyond this point.

Next, presence of a MIKC-type MADS-domain were identified in all 15 putative *ODDSOC* genes (**Figure 3.5**). This MADS-domain spans approximately the first 60 amino acids and is highly conserved (**Figure 3.5**). Several coiled coils were predicted roughly between amino acid position 100 and 130 for *TraesCS3A02G432900*, and amino acid position 115 and 145 for the remaining *ODDSOC* genes. Therefore, putative K1, K2, and K3 domains were identified within this region (**Figure 3.5**). When compared with MADS- and K-domains in notable MIKC-type genes *AtSOC1*, *AtFLC*, *AtAP1*, and *TaVRN-A1*, positions of the domains were highly similar (**Figure 3.6**).

When analysing the MADS-box domain in all *ODDSOC* genes, multiple transmembrane, α-helix, and β-strand secondary structure features were identified from the amino acid sequence (**Figure 3.7**). Interestingly, the *TraesCS3A02G432900* MADS-box domain was truncated compared to the other *ODDSOC* genes, with a length of 51 amino acids instead of 60-62 amino acids. This truncated domain was lacking α-helices and/or β-strands spanning 5-7 amino acids towards the end of the

domain, which was conserved in all other *ODDSOC* genes (**Figure 3.7**). Next, the list of *ODDSOC* genes was input into BioMart (https://plants.ensembl.org/biomart/) to collect additional gene features; these results are summarised in **Table 3.8**. The associated GO terms were similar for all genes, and included terms related to DNA binding and transcription as would be expected for MADS-domain proteins.

To explore regulation mechanisms of these genes, the genomic region upstream of each gene was analysed to identify the potential transcription factor binding sites. Since MADS domain proteins preferentially bind to CArG box sites (Aerts et al., 2018; Tilly et al., 1998), the upstream region was scanned for CArG binding sites ($CC(A/T)_6GG$, $CC(A/T)_7G$, $C(A/T)_8G$). Potential CArG binding sites were discovered in the region upstream for all genes. Additionally, this same upstream region was analysed using PlantPan: this identified potential transcription factor binding sites associated with *SOC1*-like (*AT4G24540, AT4G22950*), *AP1* (*AT1G26310*), *SVP* (*AT3G57230*), and *FLC* (*AT3G57230*, *AT2G45650*) genes in Arabidopsis (**Table 3.9**). The identification of an expanded family of *ODDSOC* genes that contain conserved protein motifs, regulatory regions, CArG-boxes, and an elongated first intron strongly suggests that at least some of these may have additional features of regulation as observed in *FLC* – for example, regulatory lncRNA.

3.3.4 In-silico investigation and RNA-seq analysis reveals lncRNA transcripts associated with *ODDSOC* genes

Initial exploration into epigenetic modifications associated with *ODDSOC* genes utilised pooled online datasets that uncovered intriguing histone modification peaks in region in close proximity to both the first and final exon (Figure 3.8). Histone modification tracks linked to up-regulation (H3K27ac, H3K27me1, H3K4me1, H3K4me3, and H3K9ac) were selected, there modification peaks localised to the first exon were presumably associated with transcription of the gene. In contrast, peaks localised to the final exon suggests additional transcription in the region directly downstream of the gene. Given that lncRNA is a known mechanism of regulation of *FLC* (Csorba et al., 2014; Liu et al., 2010; Swiezewski et al., 2009), this was investigated in the putative *ODDSOC* genes. The lncRNA track displayed numerous antisense transcripts aligning with the latter half of the gene. This pattern was consistent for all *ODDSOC* genes, and the results are summarised in **Table 3.10**.

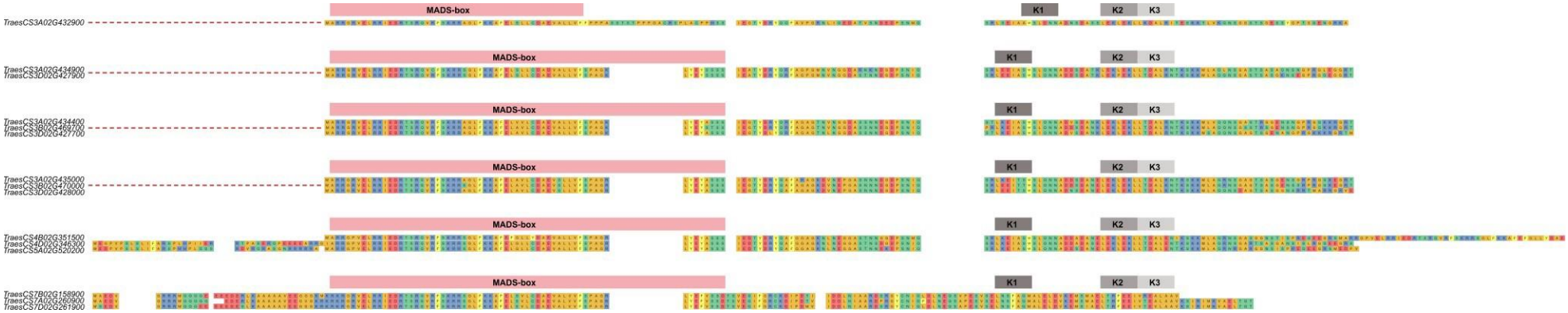


Figure 3.5: Domains associated with MIKC-genes identified within *ODDSOC* amino acid sequences.

MADS-box domains for each cluster was identified using PROSITE (https://prosite.expasy.org/) and are represented with a pink box. K- domains are denoted by K1, K2, and K3 within a grey box; position of predicted K-domains were identified via the identification of coiled coils using DeepCoil2 (https://toolkit.tuebingen.mpg.de/tools/deepcoil2).

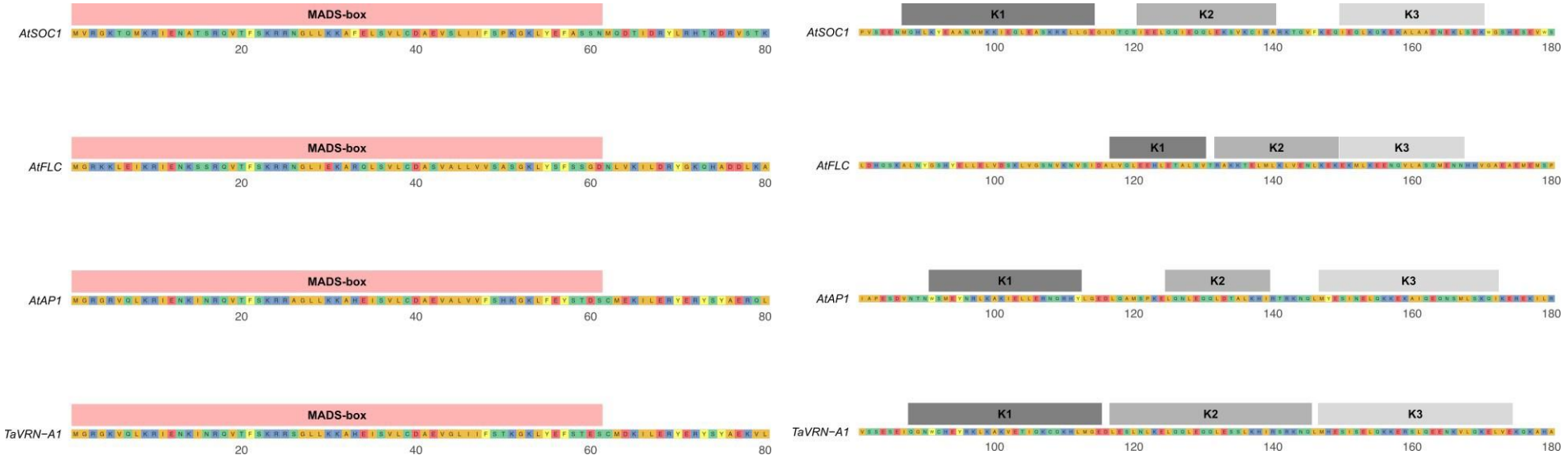


Figure 3.6: Domains within relevant MIKC-type genes amino acid sequences. (**A**) MADS-Box domains within the amino acid sequence for Arabidopsis genes *AtSOC1*, *AtFLC*, *AtAP1*, and bread wheat gene *TaVRN-A1*. (**B**) K-Box domains within *AtSOC1*, *AtFLC*, *AtAP1*, and *TaVRN-A1*; K-Boxes are denoted as K1, K2, and K3.

Table 3.9: Summary of the BioMart output for the *ODDSOC* genes.

Gene	Transcription start site	GO term accession	GO term name
TraesCS3A02G432900	674,255,517	GO:0003677 GO:0046983	DNA binding, protein dimerization activity
TraesCS3A02G434400	676,243,866	GO:0006351 GO:0005634 GO:0003677 GO:0046983 GO:0045944 GO:0000977 GO:0006355	DNA-templated transcription nucleus DNA binding protein dimerization activity positive regulation of transcription by RNA polymerase II RNA polymerase II transcription regulatory region sequence-specific DNA binding regulation of DNA-templated transcription
TraesCS3A02G434900	676,906,060	GO:0006351 GO:0005634 GO:0003677 GO:0046983 GO:0045944 GO:0000977	DNA-templated transcription nucleus DNA binding protein dimerization activity positive regulation of transcription by RNA polymerase II RNA polymerase II transcription regulatory region sequence-specific DNA binding
TraesCS3A02G435000	677,172,170	GO:0006351 GO:0005634 GO:0003677 GO:0046983 GO:0045944 GO:0000977	DNA-templated transcription nucleus DNA binding protein dimerization activity positive regulation of transcription by RNA polymerase II RNA polymerase II

			transcription regulatory region sequence-specific DNA binding
TraesCS3B02G469700	716,741,914	GO:0006351 GO:0005634 GO:0003677 GO:0046983 GO:0045944 GO:0000977	DNA-templated transcription nucleus DNA binding protein dimerization activity positive regulation of transcription by RNA polymerase II RNA polymerase II transcription regulatory region sequence-specific DNA binding

Table 3.9: Summary of the BioMart output for the *ODDSOC* genes, continued.

Gene	Transcription start site	GO term accession	GO term name
TraesCS3B02G470000	717,310,147	GO:0006351 GO:0005634 GO:0003677 GO:0046983 GO:0045944 GO:0009977 GO:0006355 GO:0009978 GO:0006357 GO:0009981	DNA-templated transcription nucleus DNA binding protein dimerization activity positive regulation of transcription by RNA polymerase II RNA polymerase II transcription regulatory region sequence-specific DNA binding regulation of DNA-templated transcription RNA polymerase II cis-regulatory region sequence-specific DNA binding regulation of transcription by RNA polymerase II DNA-binding transcription factor activity, RNA polymerase II-specific

TraesCS3D02G427700	541,197,902	GO:0006351 GO:0005634 GO:0003677 GO:0046983 GO:0045944 GO:0000977 GO:0006355	DNA-templated transcription nucleus DNA binding protein dimerization activity positive regulation of transcription by RNA polymerase II RNA polymerase II transcription regulatory region sequence-specific DNA binding regulation of DNA-templated transcription
TraesCS3D02G427900	541,403,704	GO:0006351 GO:0005634 GO:0003677 GO:0046983 GO:0045944 GO:0000977	DNA-templated transcription nucleus DNA binding protein dimerization activity positive regulation of transcription by RNA polymerase II RNA polymerase II transcription regulatory region sequence-specific DNA binding

Table 3.9: Summary of the BioMart output for the *ODDSOC* genes, continued.

Gene	Transcription start site	GO term accession	GO term name
TraesCS3D02G428000	541,632,751	GO:0006351 GO:0005634 GO:0003677 GO:0046983 GO:0045944 GO:0009977 GO:0006355 GO:0009978 GO:0006357 GO:0009981	DNA-templated transcription nucleus DNA binding protein dimerization activity positive regulation of transcription by RNA polymerase II RNA polymerase II transcription regulatory region sequence-specific DNA binding regulation of DNA-templated transcription RNA polymerase II cis-regulatory region sequence-specific DNA binding regulation of transcription by RNA polymerase II DNA-binding transcription factor activity, RNA polymerase II-specific

TraesCS3A02G432900 0	643,648,509	GO:0006351 GO:0005634 GO:0003677 GO:0046983 GO:0045944 GO:0000977	DNA-templated transcription nucleus DNA binding protein dimerization activity positive regulation of transcription by RNA polymerase II RNA polymerase II transcription regulatory region sequence-specific DNA binding
TraesCS3A02G432900 1	500,311,363	GO:0006351 GO:0005634 GO:0003677 GO:0046983 GO:0045944 GO:0000977	DNA-templated transcription nucleus DNA binding protein dimerization activity positive regulation of transcription by RNA polymerase II RNA polymerase II transcription regulatory region sequence- specific DNA binding
TraesCS3A02G432900 2	681,001,943	GO:0006351 GO:0005634 GO:000367	DNA-templated transcription nucleus DNA binding protein dimerization activity positive regulation of transcription by RNA polymerase II

	7 GO:004698 3 GO:004594 4 GO:000097 7	RNA polymerase II transcription regulatory region sequence-specific DNA binding

Table 3.9: Summary of the BioMart output for the *ODDSOC* genes, continued.

Gene	Transcription start site	GO term accession	GO term name
TraesCS3A02G432900 3	254,390,155		
TraesCS3A02G432900 4	215,932,303		
TraesCS3A02G432900 5	238,974,776		

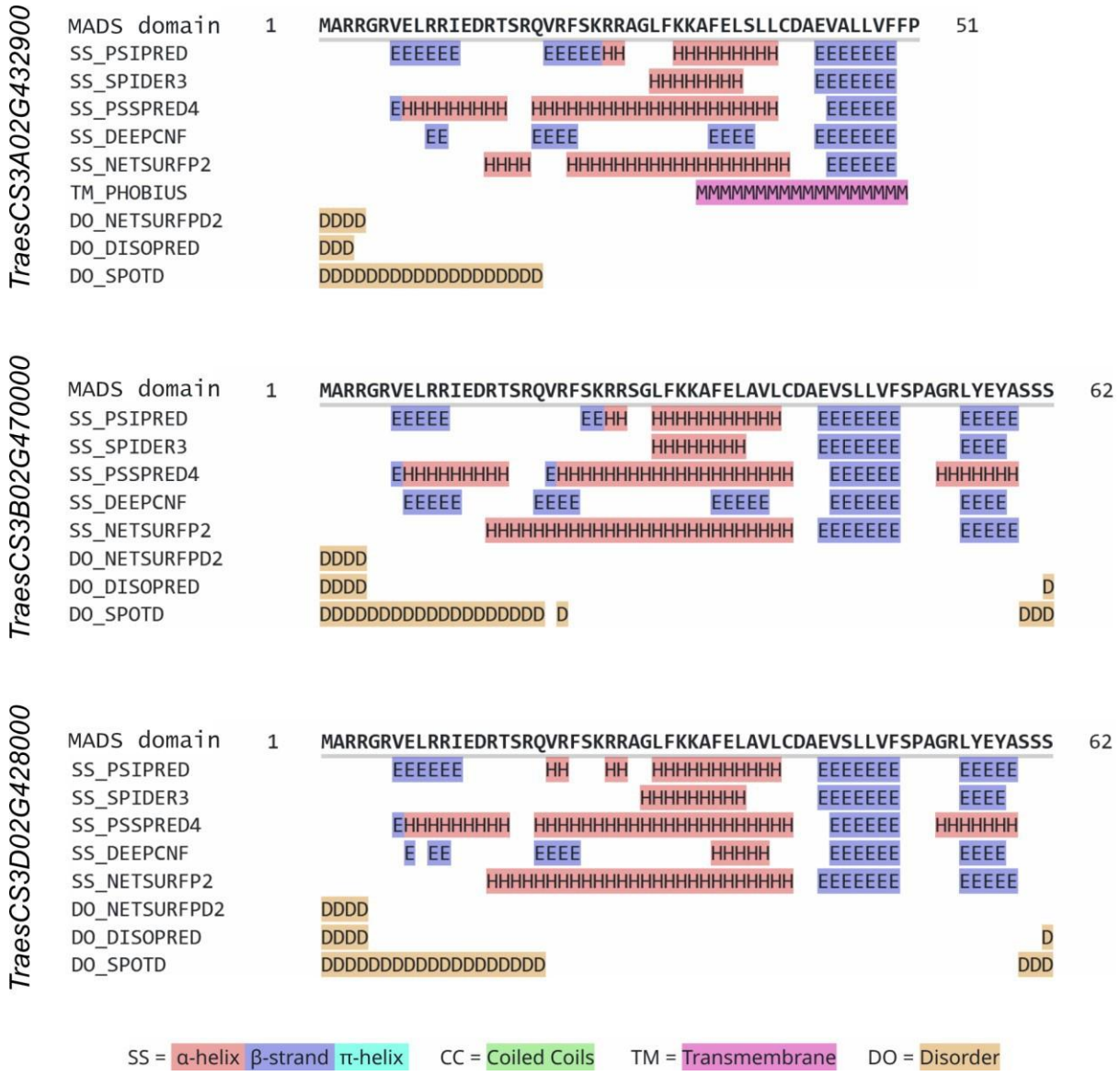


Figure 3.7: Secondary structure features of the MADS domain of *ODDSOC* genes *TraesCS3A02G432900*, *TraesCS3B02G470000*, and *TraesCS3D02G428000*.

Structures were predicted by Quick2D (https://toolkit.tuebingen.mpg.de/tools/quick2d) and compile predictions from a range of secondary structure feature prediction tools and are listed on the left.

Table 3.10: Predicted MADS-box TFBS within the 2.5 kb region upstream of the *OS2* genes. Motifs were predicted using PlantPan 3.0 multiple promoter analysis tool (http://plantpan.itps.ncku.edu.tw/promoter_multiple.php).

OS2 gene	Associated gene	Gene description
TraesCS3A02G432900		
TraesCS3A02G434400		
TraesCS3A02G435000		
TraesCS3B02G469700		
TraesCS3B02G470000		
TraesCS3D02G427900	*AT1G60920*	
TraesCS3D02G428000	(*A. thaliana*)	
TraesCS4B02G351500		
TraesCS4D02G346300		
TraesCS5A02G520200		
TraesCS7A02G260900		
TraesCS7D02G261900		
TraesCS3A02G435000	*AT4G24540*	Encodes a MADS-box protein involved in flowering; regulates the expression of *SOC1* and is also upregulated by *SOC1*; maintenance of inflorescence meristem identity, regulation of timing of transition from vegetative to reproductive transition positive regulation of vernalization response
TraesCS3B02G470000		
TraesCS3D02G428000		
TraesCS4D02G346300		
TraesCS5A02G520200		
TraesCS7A02G260900		
TraesCS3A02G435000	*AT4G22950*	MADS-box protein AGL19; *FYF*-like gene that belongs to the *SOC1* group; regulation of timing of transition from vegetative to reproductive phase, involved in vernalization response
TraesCS3B02G470000		
TraesCS3D02G428000		
TraesCS4D02G346300		
TraesCS5A02G520200		
TraesCS7A02G260900		

Table 3.10: Predicted MADS-box TFBS within the 2.5 kb region upstream of the *OS2* genes, continued.

OS2 gene	Associated gene	Gene description
TraesCS3A02G434400 *TraesCS4B02G351500* *TraesCS7B02G158900* *TraesCS7D02G261900*	*AT5G62165*	*FYF*; Involved in floral transition; regulation of flower development
TraesCS3A02G435000 *TraesCS3B02G470000* *TraesCS3D02G428000* *TraesCS4D02G346300* *TraesCS5A02G520200* *TraesCS7A02G260900*	*AT1G26310*	Floral homeotic gene encoding a MADS domain protein homologous to *AP1*; floral meristem determinacy
TraesCS3A02G435000 *TraesCS3B02G470000* *TraesCS3D02G428000* *TraesCS4D02G346300* *TraesCS5A02G520200* *TraesCS7A02G260900*	*AT2G45650*	Negatively regulates the *FLC/MAF* clade genes and positively regulates *FT*; positive regulation of flower development, vegetative to reproductive phase transition of meristem
TraesCS3A02G435000 *TraesCS3B02G470000* *TraesCS3D02G428000* *TraesCS4D02G346300* *TraesCS5A02G520200* *TraesCS7A02G260900*	*AT3G57230*	Directly interacts with *SVP* and indirectly interacts with *FLC*; long-day photoperiodism, flowering
TraesCS3A02G435000 *TraesCS3B02G470000* *TraesCS3D02G428000* *TraesCS4D02G346300* *TraesCS5A02G520200* *TraesCS7A02G260900*	*AT3G57390*	Negative regulation of flower development; Negative regulation of short-day photoperiodism; flowering

Table 3.10: Predicted MADS-box TFBS within the 2.5 kb region upstream of the *OS2* genes, continued.

OS2 gene	Associated gene	Gene description
TraesCS3A02G435000 *TraesCS3B02G470000* *TraesCS3D02G428000* *TraesCS4D02G346300* *TraesCS5A02G520200* *TraesCS7A02G260900*	*AT4G11880*	Maintenance of floral meristem identity; vegetative to reproductive phase transition of meristem

3.3.5 Validation of findings using the available literature

To determine whether the collated list of *ODDSOC* genes had been previously described, I cross-referenced the list with the literature. Since domain analysis had characterised all the genes as MADS-box genes based on their predicted/conserved

MADS-box binding domain, I investigated whether any of these genes had been characterised in widescale analyses of wheat MADS-box genes. Schilling et al. (2020) had previously annotated all 15 genes as forms of *TaFLC* in their genome-wide analysis of MIKC-type MADS-box wheat genes. In support of their findings, Raza et al. (2021) identified these same genes in their analysis of MADS-box genes in wheat. Both these in-silico findings are distilled with respect to the *OS2* gene list and represented in **Table 3.12**. In addition, distilling the in-silico/bioinformatics findings of both the literature and this project, the genomic position of these *ODDSOC* genes is represented in **Figure 3.9**, highlighting the migration of these 15 genes across all 3 sub genomes.

3.3.6 Are all *ODDSOC* genes created equal?

The identification of the extended *ODDSOC* gene family and the related analysis indicated that 4 genes arose from duplications in chromosome 3A, 2 genes from chromosome 3B, and 3 genes from chromosome 3D. This conclusion was based on close proximity and high sequence similarity; positions are described in **Table 3.1**. Furthermore, the gene family was comprised of 5 discrete homoeologous groups that were distributed across the A, B, and D sub genome, with the exception of *TraesCS3A02G434900* and *TraesCS3D02G427900*, which had no homoeologue from the B sub genome. Position and movement in relation to the sub genome is represented in **Figure 3.9**. This level of relatedness between these genes raised the question of which *ODDSOC* genes were involved in the regulation of floral transition in wheat. Since *TraesCS3A02G432900* was most likely the equivalent of *OS1* in wheat, this gene was selected for further investigation. Interestingly, no homoeologues were identified for *TraesCS3A02G432900*. *TraesCS3A02G435000*, *TraesCS3B02G470000*, and *TraesCS3D02G428000* were studied in more detail as these homoeologues were highlighted in eQTL analysis as having a role in the vegetative to reproductive transition (see section **3.3.1**).

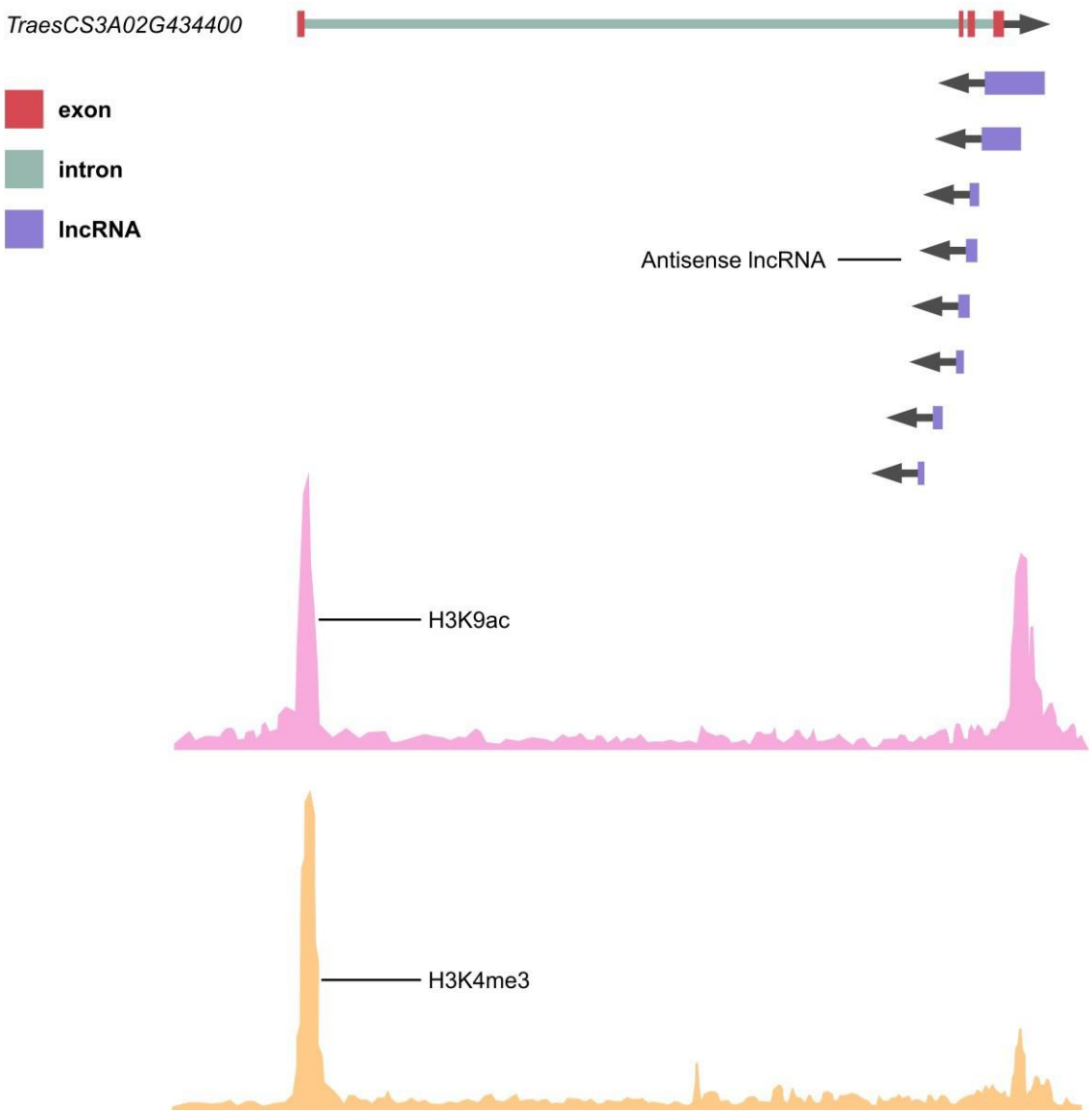


Figure 3.8: Schematic of histone modifications and lncRNA transcripts associated with the genomic region containing *TraesCS3A02G434400* (chr3A:676200201..676247350). Tracks were available from the JBrowse wheat browser provided by WheatOmics (http://wheatomics.sdau.edu.cn/).

Table 3.11: Long non-coding antisense transcripts associated with *ODDSOC* genes. (Epigenome browser last accessed 09/04/23).

Gene	Homoeologues	Transcript number	Tissue type
TraesCS3A02G432900		13	Seedling, root, spike, stem
TraesCS3A02G434400		24	Leaf, seedling, root, spike, stem
TraesCS3A02G434900		10	Leaf, seedling, root, spike
TraesCS3A02G435000		4	Leaf, seedling, root
TraesCS3B02G469700		22	Leaf, seedling, root, spike, stem
TraesCS3B02G470000		5	Leaf, seedling, root
TraesCS3D02G427700		16	Leaf, seedling, root, spike, stem
TraesCS3D02G427900		13	Leaf, seedling, root, spike
TraesCS3D02G428000		3	Leaf, root
TraesCS4B02G351500		1	Root
TraesCS4D02G346300		1	Root
TraesCS5A02G520200		1	Root
TraesCS7B02G158900		6	Seedling, spike
TraesCS7A02G260900		0	-
TraesCS7D02G261900		1	Seedling

Table 3.12: Nomenclature of *ODDSOC* in wheat from previous studies.

Schilling et al. (2020)	Ma et al. (2017)	Ruelens et al. (2013)	Zhao et al. (2006)	NCBI GenBank I.D.	*H. vulgare* BLASTx hit
TaFLC-A5			*TaAGL42*		
TaFLC-A3			*TaAGL41*		
TaFLC-A6					
TaFLC-A4-1				*TaAGL33*,	
TaFLC-B3	*TaMADS58*				
TaFLC-B4				*TaAGL33*,	HORVU.MOREX.
TaFLC-D3					
TaFLC-D6					
TaFLC-D4				*TaAGL33*,	
TaFLC-B2			*TaAGL33*		
TaFLC-D2					
TaFLC-A2					
TaFLC-B1	*TaMADS157*	*TaAGL12*			
TaFLC-A1	*TaMADS51*	*TaAGL12*		*TaAGL12*	
TaFLC-D1-1		*TaAGL12*		*TaAGL12*	

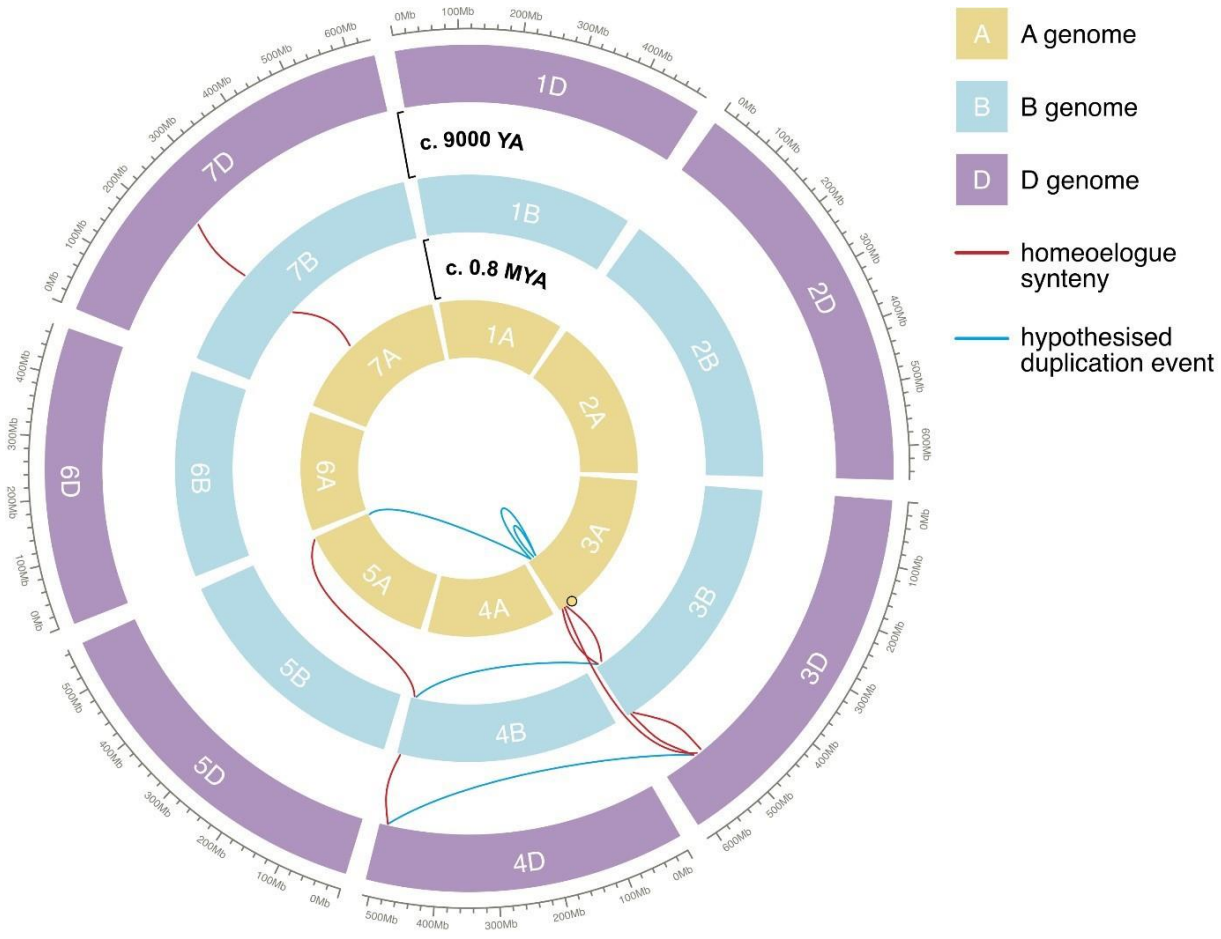


Figure 3.9: Chromosomal position of *ODDSOC* genes across the three wheat sub genomes. Red lines depict the relationship between homoeologues across genomes; blue lines depict hypothesised duplication events between genes that are not part of the same homoeologues triad. The position of *TraesCS3A02G432900*, which lacks homoeologues and hypothesised duplication events, is represented by a yellow circle on chromosome 3A.

3.3.7 *ODDSOC* genes interact with each other

MADS-domain proteins are understood to form homodimers and heterodimers with each other, as exemplified by the ABC model of floral development (Coen & Meyerowitz, 1991; Riechmann et al., 1996), as well as participating in larger multi-protein complexes that can determine functional specificity. This specificity involves association with CArG box domains and therefore the transcriptional activity of MADS-box proteins (Shore & Sharrocks, 1995). Since the *ODDSOC* genes examined here are characterised as MADS-box genes, this indicates the possibility of the formation of ODDSOC protein complexes. Furthermore, protein complex formation could elucidate understanding of how these proteins function in floral development. As the examination of protein complex formation is not yet feasible in transgenic wheat, protein-protein interactions were examined within a yeast two-hybrid system.

When examining the interactions of several *ODDSOC* proteins, a range of interactions were observed. Here, *OS1* (*TraesCS3A02G432900*) and *OS2* (*TraesCS3A02G434400*, *TraesCS3A02G435000*, *TraesCS3B02G470000*, and *TraesCS3D02G428000*) were cloned into the plasmids pGAD424 and pGBT9 to test homo- and heterodimerization in yeast. This revealed that TraesCS3B02G470000 didn't form homodimers with itself (**Figure 3.10**). In contrast, TraesCS3D02G428000 was observed to form weak (2.5 mM 3-AT) and moderate (5 mM 3-AT) homodimers (**Figure 3.11**). As *TraesCS3B02G470000* and *TraesCS3D02G428000* are homoeologues that shared 95.6% amino acid sequence similarity, this difference in homodimer formation is intriguing.

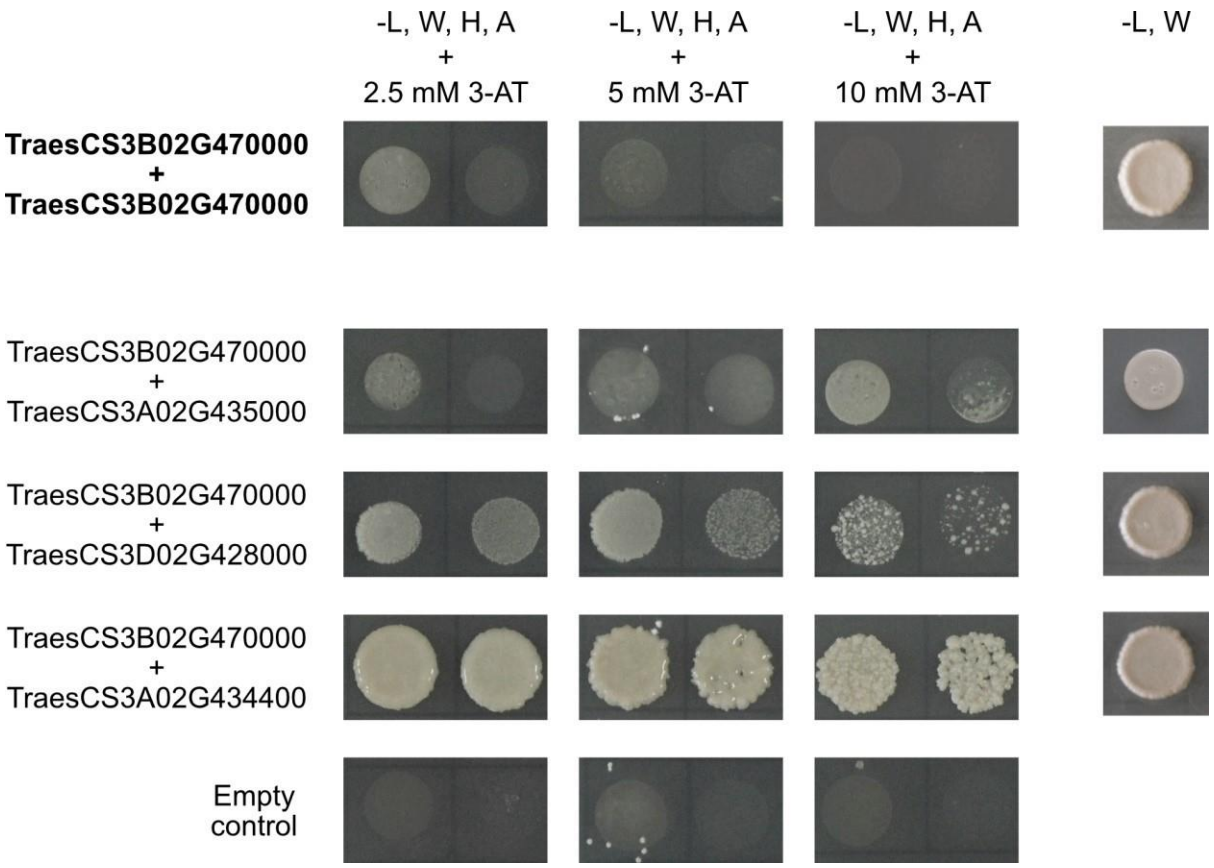


Figure 3.10: Protein-protein interactions for TraesCS3B02G470000. Capital letters indicate amino acids lacking in synthetic yeast media, i.e., L (leucine), W (tryptophan), H (histidine), and A (adenine). Empty plasmids = pGAD424-TraesCS3B02G470000 co-transformed with empty pGBT9 plasmid; 3-AT = 3-aminotriazole concentration.

Next, heterodimer formation was investigated. TraesCS3B02G470000 did not interact with its homoeologues TraesCS3D02G428000 or TraesCS3A02G435000 (**Figure 3.10**). However, TraesCS3A02G435000 and TraesCS3D02G428000 did form strong heterodimers (94.9% amino acid sequence similarity), indicating that the absence of heterodimer formation with TraesCS3B02G470000 may be meaningful. Strong heterodimer formation was observed between TraesCS3A02G434400 and TraesCS3B02G470000; in contrast, only a weak heterodimer formation was recorded between TraesCS3A02G434400 and TraesCS3D02G428000 (**Figure 3.11**). This indicates that despite high amino acid conservation and conservation of the MADS-box domains between *ODDSOC* genes, specific interactions between genes were observed. To further explore this variation in dimer formation between homoeologues TraesCS3B02G470000 and TraesCS3D02G428000, interaction between these *OS2* proteins and *OS1*, TraesCS3A02G432900, was tested. Interestingly, TraesCS3B02G470000 and TraesCS3A02G432900 formed strong heterodimers (**Figure 3.12**). However, no heterodimer formation occurred between TraesCS3D02G428000 and TraesCS3A02G432900 (**Figure 3.12**). This furthers the suggestion that *TraesCS3B02G470000* and *TraesCS3D02G428000* inhabit different roles in vernalization/development despite their high level of structural and sequential similarity.

Since *VRN-A1* has been reported to supress *OS2*, interactions between *OS2* proteins, TraesCS3B02G470000 and TraesCS3D02G428000, and VRN-A1 was investigated. For both OS2 proteins, an interaction with VRN-A1 was observed (**Figure 3.13**). TraesCS3B02G470000 and VRN-A1 formed strong (under 10 mM 3-AT) heterodimers (**Figure 3.13**). Similarly, TraesCS3D02G428000 and VRN-A1 formed strong heterodimers with the inclusion of 3-AT. Next, heterodimer formation between TraesCS3B02G470000 and a variant of VRN-A1 with a SNP in exon 4 conferring a stronger vernalization requirement was tested (Díaz et al., 2012; Dixon et al., 2019). In contrast to WT VRN-A1, the exon 4 SNP VRN-A1 variant didn't form heterodimers with TraesCS3B02G470000 (**Figure 3.13**). Considering that the difference between these forms of VRN-A1 is a single nucleotide difference, it demonstrates the varying specificity of the MADS-Box domain interactions. This potentially suggests an important role for TraesCS3B02G470000 in mediating a longer vernalization requirement in cultivars that contain the VRN-A1 exon 4 variant.

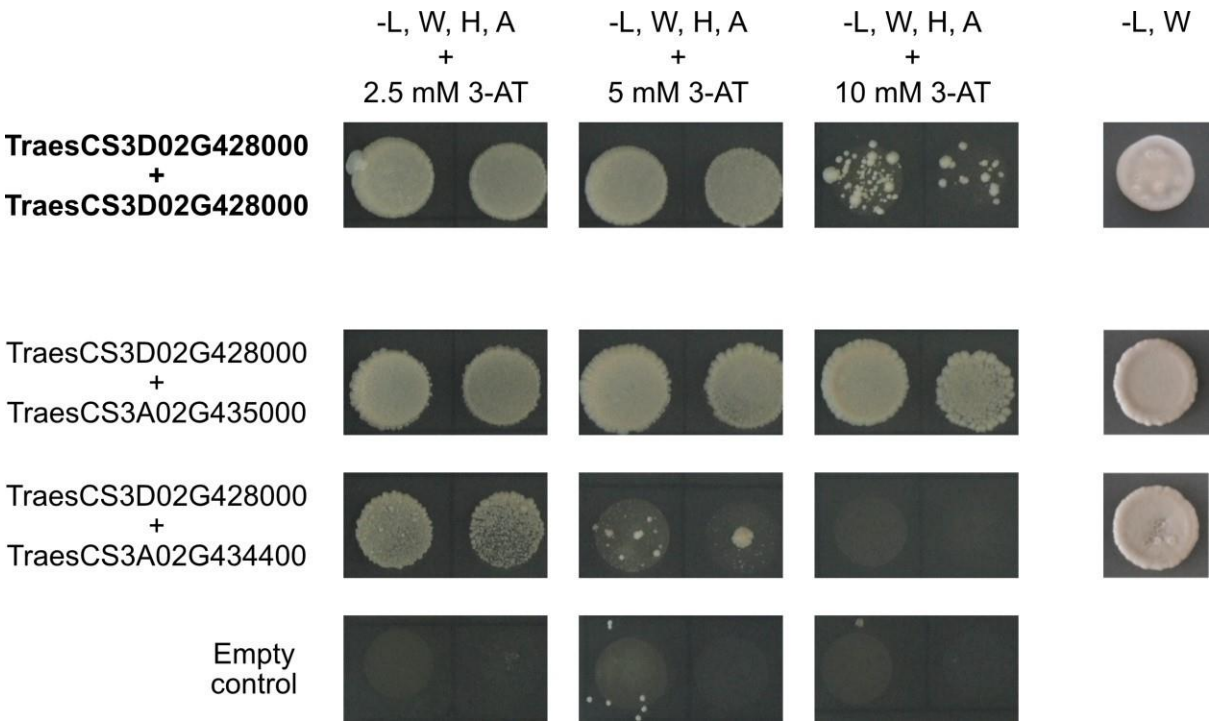


Figure 3.11: Protein-protein interactions for TraesCS3D02G428000. Capital letters indicate amino acids lacking in synthetic yeast media, i.e., L (leucine), W (tryptophan), H (histidine), and A (adenine). Empty plasmids = pGAD424- TraesCS3D02G428000 co-transformed with empty pGBT9 plasmid; 3-AT = 3-aminotriazole concentration.

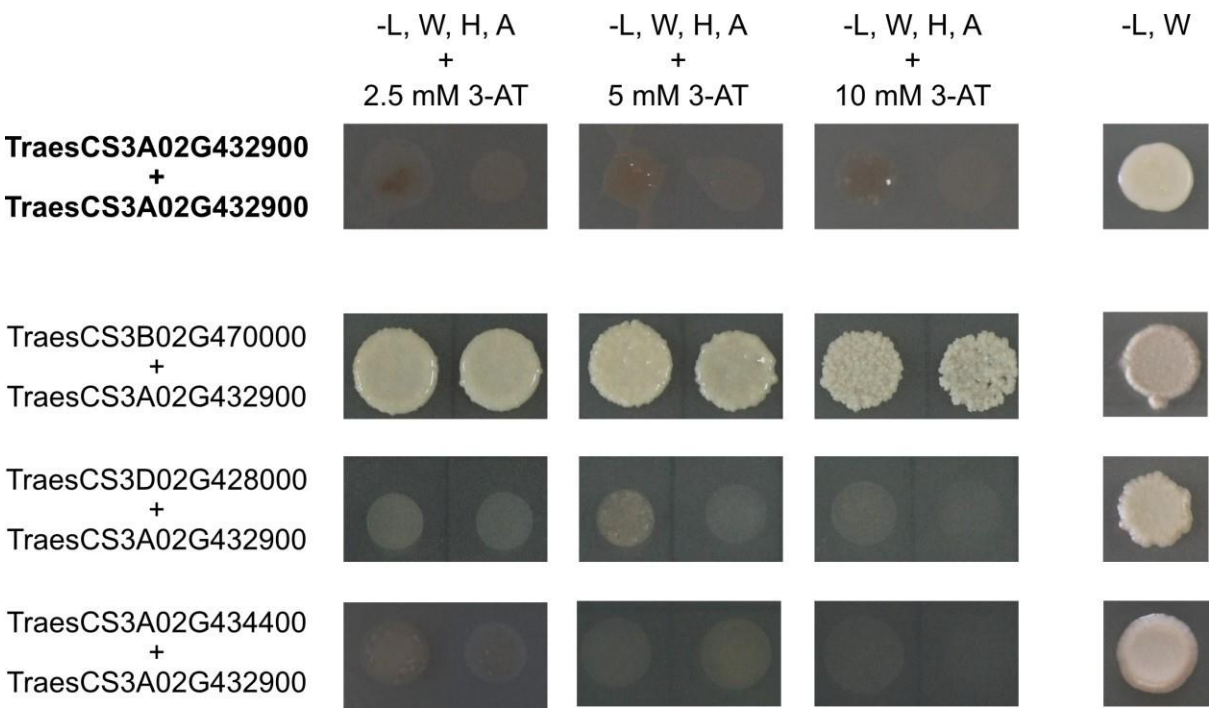


Figure 3.12: Protein-protein interactions for TraesCS3A02G432900. Capital letters indicate amino acids lacking in synthetic yeast media, i.e., L (leucine), W (tryptophan), H (histidine), and A (adenine). Empty plasmids = pGAD424- TraesCS3A02G432900 co-transformed with empty pGBT9 plasmid; 3-AT = 3-aminotriazole concentration.

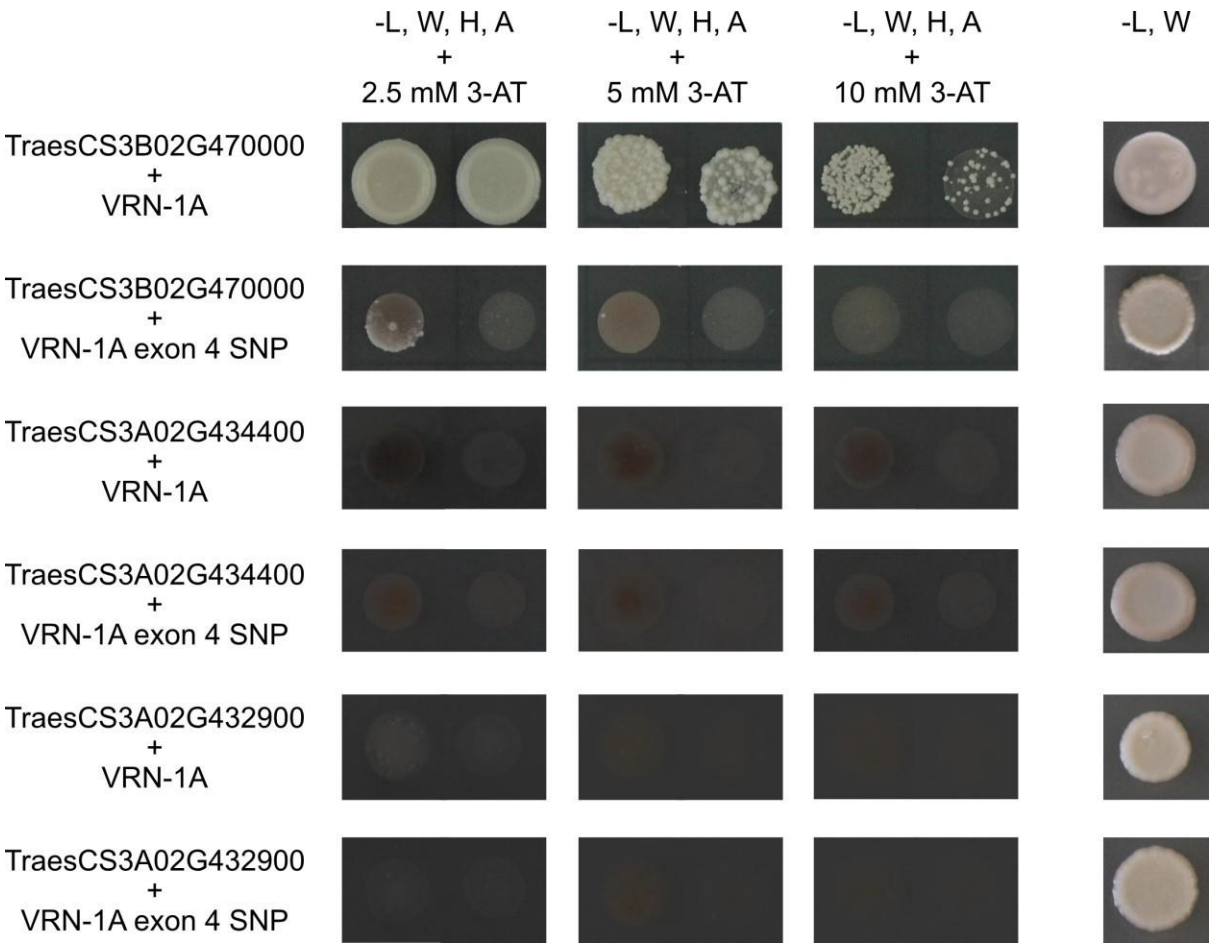


Figure 3.13: Protein-protein interactions for ODDSOC and VRN-A1. Capital letters indicate amino acids lacking in synthetic yeast media, i.e., L (leucine), W (tryptophan), H (histidine), and A (adenine). 3-AT = 3-aminotriazole concentration.

3.3.8 Watkins diversity panel, YoGI expression panel, and WatSeq uncovers variation in *ODDSOC* gene expression

As the interaction analysis supported that hexaploid wheat *OS2* did associate with *VRN1*, I wanted to further investigate whether there was any association between the expression of the *ODDSOC* genes and growth habit.

To explore whether variation in *OS1* and *OS2* expression is present between spring and winter wheat, expression from the YoGI accession of 339 lines was analysed. This revealed trends in expression between genes. *OS1*/*TraesCS3A02G432900* had a relatively equal level of expression between spring and winter wheat lines, with both the highest (19.74812856 FPKM) and lowest (0.084473841 FPKM) level of expression recorded in spring lines (**Figure 3.14A**). Similarly, *OS2* gene *TraesCS3D02G427900* didn't display a statistically significant trend towards higher expression in spring compared to winter lines, although the overall level of expression was lower than *TraesCS3A02G432900* (0.391076725 and 4.165609224 FPKM, respectively) (**Figure 3.14C**). Likewise, *OS2* genes and homoeologues *TraesCS3A02G434400*, *TraesCS3B02G469700*, and *TraesCS3D02G427700* don't show a significant trend to a particular habit (**Figure 3.15A-C**).

Conversely, *TraesCS3A02G434900* (**Figure 3.14B**), *TraesCS3A02G435000*, *TraesCS3B02G470000*, *TraesCS3D02G428000* (**Figure 3.15D-F**), *TraesCS4B02G351500*, *TraesCS4D02G346300*, *TraesCS5A02G520200* (**Figure 3.16A-C**), *TraesCS7B02G158900*, *TraesCS7A02G260900*, and *TraesCS7D02G261900* (**Figure 3.16D-F**) all show a trend towards higher expression levels in winter wheat lines. Variation in gene expression existed within this subset: the average expression levels were *TraesCS3A02G434900* (1.634 FPKM), *TraesCS3A02G435000* (1.876 FPKM), *TraesCS3B02G470000* (6.101 FPKM), glasshouse conditions; therefore, it was necessary to determine whether the observed higher expression in winter lines for certain *ODDSOC* genes was consistent under different conditions. To begin, expression of *TraesCS3A02G435000*, *TraesCS3B02G470000*, *TraesCS3D02G428000*, *TraesCS3A02G434400*, *TraesCS3B02G469700*, and *TraesCS3D02G427700* was investigated.

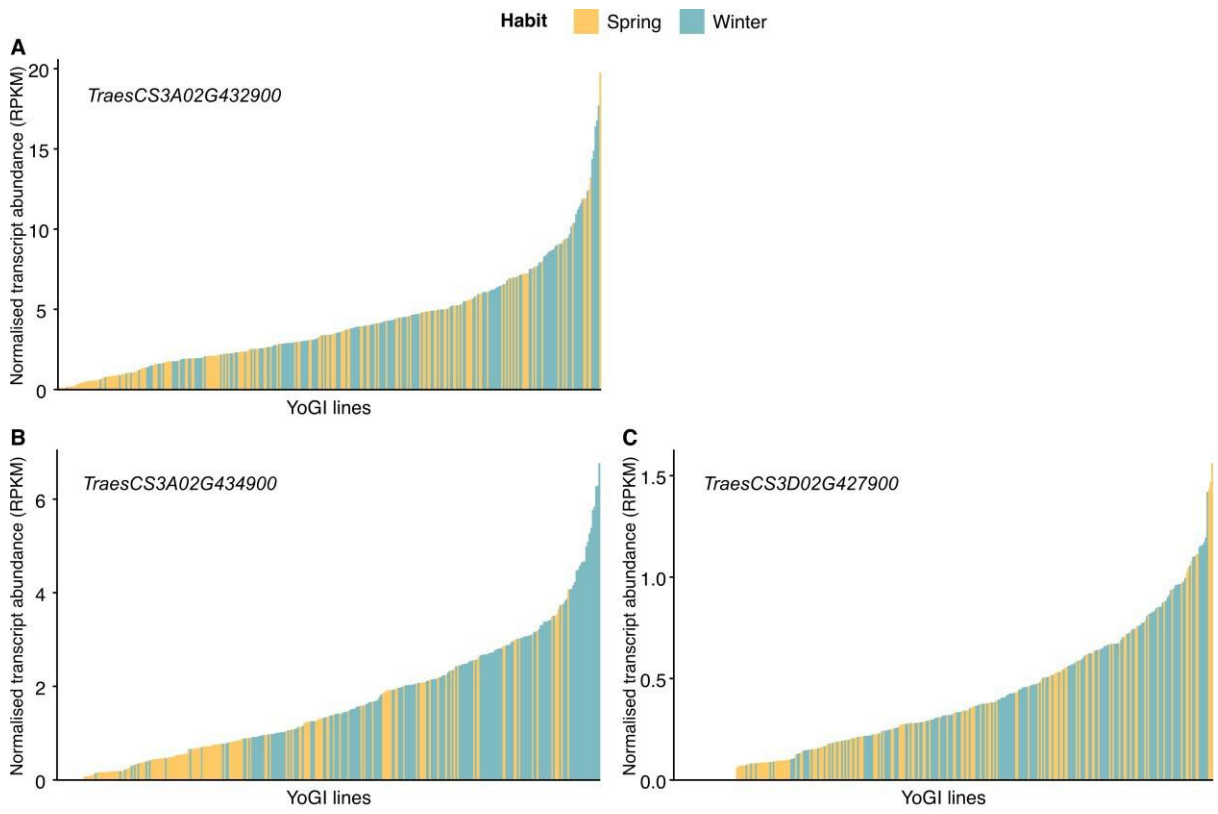


Figure 3.14: *ODDSOC* gene expression in YoGI lines for the first set of genes. Expression data was acquired from the YoGI accession lines (n = 339) and contains winter (blue) and spring (yellow) habit wheat. (**A**) = *TraesCS3A02G432900*; Figure (**B-C**) = homoeologues *TraesCS3A02G434900* and *TraesCS3D02G427900*. Leaf tissue was sampled at the 3-leaf stage of development by Dr Andrea Harper's team, University of York.

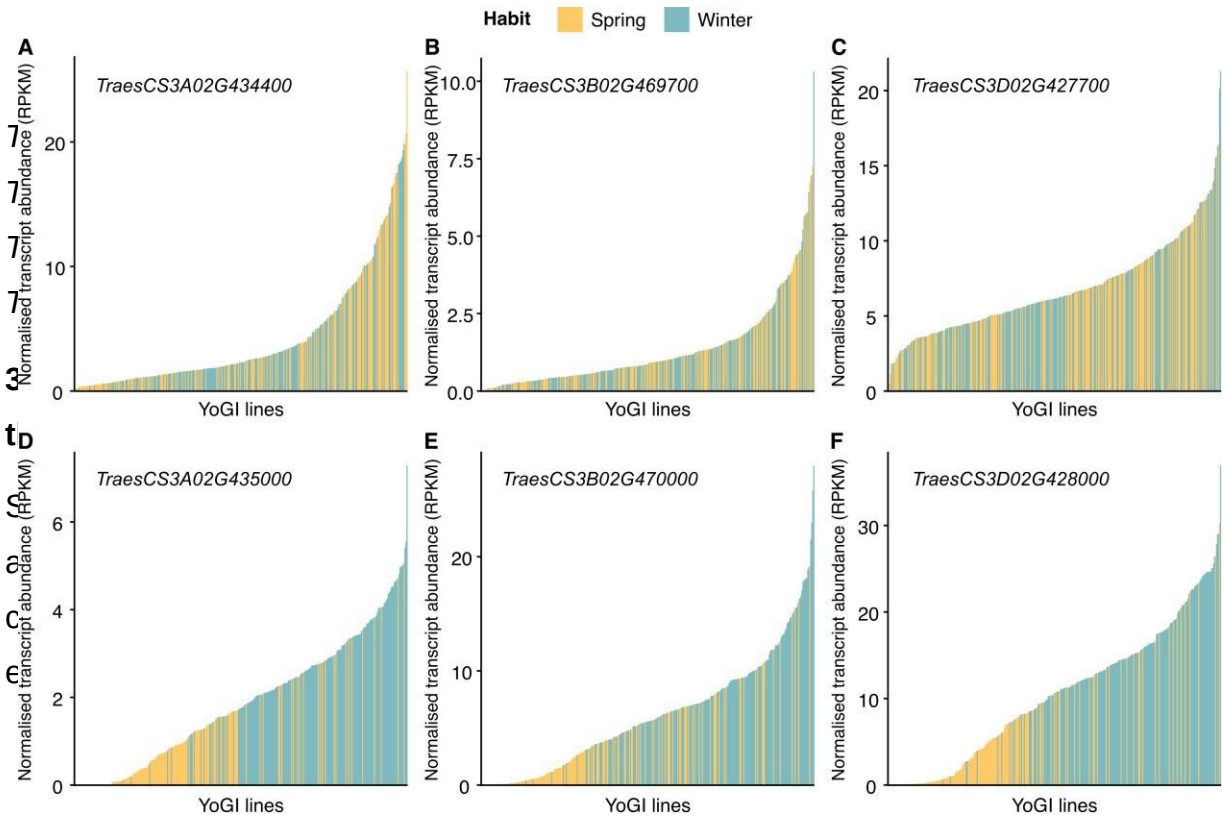


Figure 3.15: *ODDSOC* gene expression in YoGI lines for the second set of genes. Expression data was acquired from the YoGI accession lines (n = 339) and contains winter (blue) and spring (yellow) habit wheat. (**A**-**C**) = homoeologues *TraesCS3A02G434400*, *TraesCS3B02G469700*, and *TraesCS3D02G427700*; (**D**-**F**) = *TraesCS3A02G435000*, *TraesCS3B02G470000*, and *TraesCS3D02G428000*. Leaf tissue was sampled at the 3-leaf stage of development by Dr Andrea Harper's team, University of York.

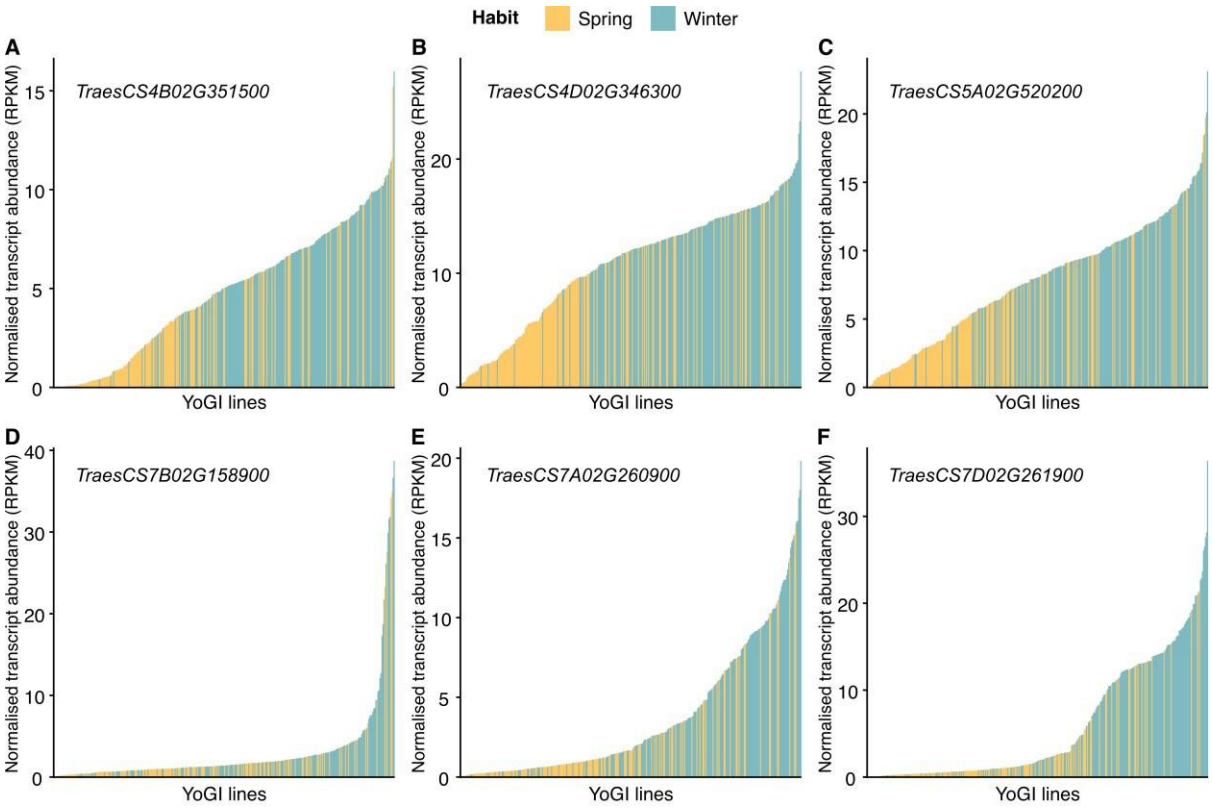


Figure 3.16: *ODDSOC* gene expression in YoGI lines for the third set of genes. Expression data was acquired from the YoGI accession lines (n = 339) and contains winter (blue) and spring (yellow) habit wheat. (**A**-**C**) = homoeologues *TraesCS4B02G351500*, *TraesCS4D02G346300*, and *TraesCS5A02G520200*; (**D**-**F**) = homoeologues *TraesCS7B02G158900*, *TraesCS7A02G260900*, and *TraesCS7D02G261900*. Leaf tissue was sampled at the 3-leaf stage of development by Dr Andrea Harper's team, University of York.

As described in the section **3.3.8**, expression of homoeologues *TraesCS3A02G435000*, *TraesCS3B02G470000*, and *TraesCS3D02G428000* was higher in winter lines, whereas homoeologues *TraesCS3A02G434400*, *TraesCS3B02G469700*, and *TraesCS3D02G427700* had no preference for growth habit. Therefore, expression of these genes in winter lines under different photoperiod length was investigated further. This revealed that expression of these genes under 22 °C short day, long day, and long-long day conditions was variable (**Figure 3.17**). Highest expression for both homoeologous groups was observed under 22 °C short day conditions. Interestingly, homoeologues *TraesCS3A02G435000*, *TraesCS3B02G470000*, and *TraesCS3D02G428000* had higher expression in cultivars Buster and Charger, which have moderate to high vernalization requirements. Conversely, *TraesCS3A02G434400*, *TraesCS3B02G469700*, and *TraesCS3D02G427700* had elevated expression in facultative winter cv. Claire, which has a low vernalization requirement. When day length was extended to 16 hr light/8 hr dark, all 6 genes were down regulated in the cultivars, with no statistically significant difference in expression level (**Figure 3.17B**). Strikingly, increasing the photoperiod further to 22 hr light/2 hr dark resulted in up regulation of all genes (**Figure 3.17C**).

In cultivars Claire and Charger, expression of *TraesCS3A02G434400*, *TraesCS3B02G469700*, and *TraesCS3D02G427700* was slightly higher than *TraesCS3A02G435000*, *TraesCS3B02G470000*, and *TraesCS3D02G428000* (**Figure 3.17C**). However, these differences in expression level were not statistically significant ($p>0.05$; ANOVA). Since highest expression overall was in cv. Buster with homoeologues *TraesCS3A02G435000*, *TraesCS3B02G470000*, and *TraesCS3D02G428000*, this could indicate that these genes are responsive to photoperiod in this cultivar. To determine whether temperature is also involved in this response, expression of *TraesCS3B02G470000* in cv. Buster was examined under either 10 °C or 16 °C short day conditions. *TraesCS3B02G470000* was selected as results of the protein-protein interactions in section **3.3.7** demonstrated that TraesCS3B02G470000 interacted with putative floral activator TraesCS3A02G432900 and VRN-A1, possibly indicative of involvement in the vernalization pathway. To pinpoint at which time of the day *TraesCS3B02G470000* transcript level was highest, expression was investigated over a 24-hr period.

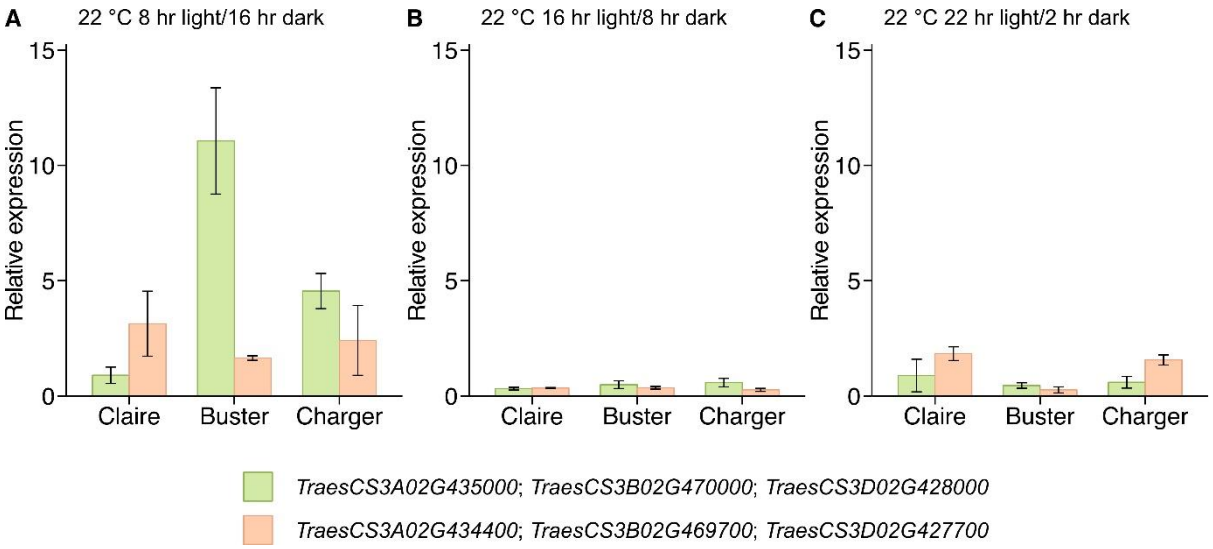


Figure 3.17: Expression of *ODDSOC* homoeologous groups under different photoperiods. (**A**) Expression of homoeologues TraesCS3A02G435000, TraesCS3B02G470000, TraesCS3D02G428000 (green) and TraesCS3A02G434400, TraesCS3B02G469700, TraesCS3D02G427700 (orange) under 22 °C short day (8 hr light/16 hr dark conditions). (**B**) Expression of homoeologues under 22 °C long day (16 hr light/8 hr dark) conditions. (**C**) Expression of homoeologues under 22 °C long-long day (22 hr light/2 hr dark). Leaf tissue from winter wheat cultivars Claire, Buster, and Charger was sampled at 3 weeks of growth. Error bars represent standard error of the mean for 3 biological replicates.

The expression analysis revealed that *TraesCS3B02G470000* was more up regulated under 16 °C compared to 10 °C short day conditions (**Figure 3.18A**). Furthermore, expression under 16 °C was consistently higher than 10 °C across the timecourse. Intriguingly, under 16 °C, *TraesCS3B02G470000* had a rhythmic pattern of expression, indicating potential circadian recognition (**Figure 3.18A**). Expression remained steading during the initial 5 hours post-dawn, before increasing at dusk (ZT8) and continuing to increase until ZT15. After this point, expression started to decrease once more, aligning with the arrival of dawn at ZT24/ZT0. Since expression continued to rise and fall in total darkness, it was important to establish whether the observed rhythmic expression pattern was independent to photoperiod recognition. Therefore, *TraesCS3B02G470000* expression for these same temperatures but under a 22-hr photoperiod was examined next. This revealed that under a 22-hr photoperiod, expression level in response to both temperatures was significantly lower than under short day (i.e., 8 hr photoperiod) conditions (**Figure 3.18B**). Under a temperature of 16 °C, *TraesCS3B02G470000* was still slightly higher compared to 10 °C. However, the rhythmic pattern of expression that was observed under a short- day photoperiod was absent under an extended photoperiod of 22 hrs light (**Figure 3.18**). This would suggest that *TraesCS3B02G470000* is responsive to both temperature and photoperiod.

Considering that *TraesCS3B02G470000* is up regulated under conditions not suitable for flowering, such as the warmer shorter days associated with late autumn/beginning of winter, and under conditions suitable for flowering only if the vernalization requirement has been met, such as the warm long days associated with summer, this would support characterisation of *TraesCS3B02G470000* as a floral repressor. With this in mind, the next step was to determine whether any of the other putative *ODDSOC* genes were responsive to different temperatures under a short-day photoperiod. Therefore, expression of the *ODDSOC* genes in cv. Buster was investigated under 10 °C and 14 °C short-day conditions, using the RNA-seq dataset introduced in the previous chapter (section **2.3.1.4**).

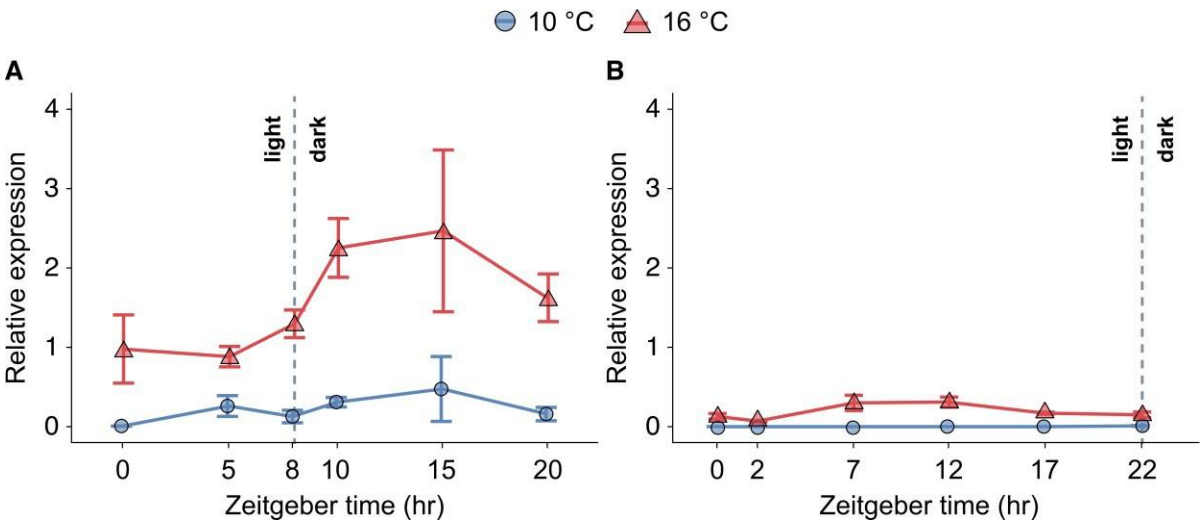


Figure 3.18: Expression of *TraesCS3B02G470000* in winter wheat cv. Buster under varying temperature and photoperiod. (**A**) TraesCS3B02G470000 expression under 10 °C (blue circle) and 16 °C (red triangle) short day (8 hr light/16 hr dark) conditions. Expression was measured across a 24 hr period, with zeitgeber (ZT) 24 equatable to ZT0. (**B**) TraesCS3B02G470000 expression under 10 °C and 16 °C long-long day (22 hr light/2 hr dark) conditions; ZT24 equatable to ZT0. Error bars represent standard error of the mean for 3 biological replicates.

The results of this expression analysis revealed variation in both expression level and responsiveness to temperature (**Figure 3.19**). As observed under 16 °C short-day conditions, *TraesCS3B02G470000* had a higher level of expression under 14 °C compared to 10 °C. Interestingly, the *TraesCS3B02G470000* homoeologue *TraesCS3D02G428000* was up regulated under both 10 °C and 14 °C; however, expression was slightly higher under 14 °C. This level of expression was significantly greater ($p<0.05$; Student's t-test) than *TraesCS3B02G470000*. Furthermore, the third homoeologue of these genes, *TraesCS3A02G435000*, was not expressed under 10 °C short-day conditions, and was only minimally expressed under 14 °C. This difference in expression level was statistically significant ($p<0.01$; Student's t-test) when compared to expression of *TraesCS3B02G470000* and *TraesCS3B02G470000*. In addition to *TraesCS3A02G435000*, the genes *TraesCS3D02G427900* and *TraesCS3A02G434900* also had low levels of expression under both temperatures (**Figure 3.19**). Genes *TraesCS3A02G432900*, *TraesCS4B02G351500*, *TraesCS5A02G520200*, *TraesCS3A02G434900*, *TraesCS7A02G260900*, *TraesCS7B02G158900*, and *TraesCS7D02G261900* all had higher levels of expression under 10 °C. However, a statistically significant difference ($p<0.05$; Student's t-test) in expression under different temperatures was only reported in *TraesCS3A02G432900*. This aligns with this gene being the most likely candidate for *TaAGL42*, predicted to be the equivalent of floral promoter *OS1* in bread wheat.

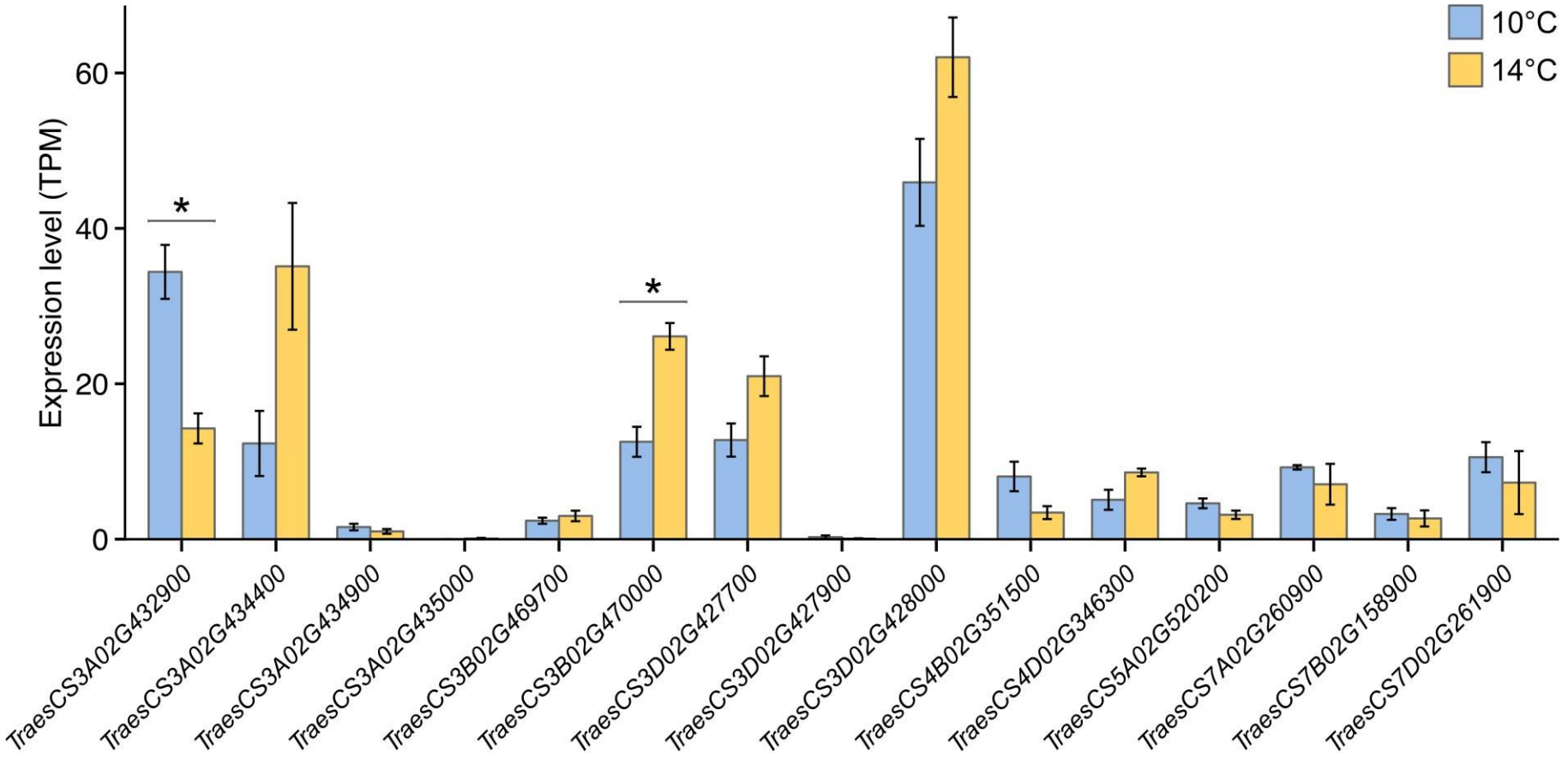


Figure 3.19: *ODDSOC* gene expression in winter wheat cv. Buster under different temperatures during vernalization. Plants were grown under 10 °C (blue) and 14 °C (yellow) short day (8 hr light/16 hr dark) conditions. Expression was measured after 4 weeks of growth. Error bars represent standard error of the mean for 3 biological replicates; asterisk (*) indicates statistically significant difference ($p<0.01$).

Since *TraesCS3B02G470000* was down regulated under 10 °C long long-day conditions, we would anticipate that this gene is repressed during the SV and SGV conditions introduced in sections **2.3.3** and **2.3.4**. To examine how the other *ODDSOC* genes responded to these conditions, *ODDSOC* gene expression in cv. Hereward was investigated after 4 weeks of growth in SV, SGV, or control (i.e., unsuccessful SGV). Strikingly, this revealed that *OS2* genes were not exclusively repressed under any of these conditions, despite a 10 °C/22 hr photoperiod being common to all conditions (**Figure 3.20**). Furthermore, expression was shown to vary even between groups of homoeologues. Expression of *TraesCS3A02G434900* was significantly higher than its homoeologue *TraesCS3D02G427900*; *TraesCS3A02G434400* and *TraesCS3D02G427700* were both considerably higher than their homoeologue *TraesCS3B02G469700*, and expression of *TraesCS3A02G435000* was not detected, unlike its homoeologues *TraesCS3B02G470000* and *TraesCS3D02G428000* (**Figure 3.20**).

Markedly, expression level was highest under SV conditions for 10 out of the 15 genes. This difference was most pronounced in *TraesCS4D02G346300*, with highest expression under SV conditions ($p<0.05$; ANOVA and Tukey post-hoc test) (**Figure 3.20**). Conversely, there was a statistically significant difference ($p<0.05$; ANOVA and Tukey post-hoc test) in the expression level of *TraesCS7B02G158900,* with higher expression under the control and SGV conditions compared to SV conditions (**Figure 3.20**). Since the key difference between these conditions is that SGV and the control conditions involved an uncovered seed on the soil surface, this could potentially suggest that elements involvement in light sensing directly or indirectly down regulate certain *ODDSOC* genes, while up regulating others.

Interestingly, the expression of *TraesCS3D02G427700* and *TraesCS7D02G261900* under each vernalization condition was most comparable to *OS1* gene *TraesCS3A02G432900* (**Figure 3.20**). The CDS of *TraesCS3A02G432900* shares the highest level of similarity with *TraesCS3D02G427700* (95.9%) whereas the protein sequences share 72.6% similarity; therefore, *TraesCS3D02G427700* could potentially operate as a floral activator, similar to what has been shown for *OS1* in other species.

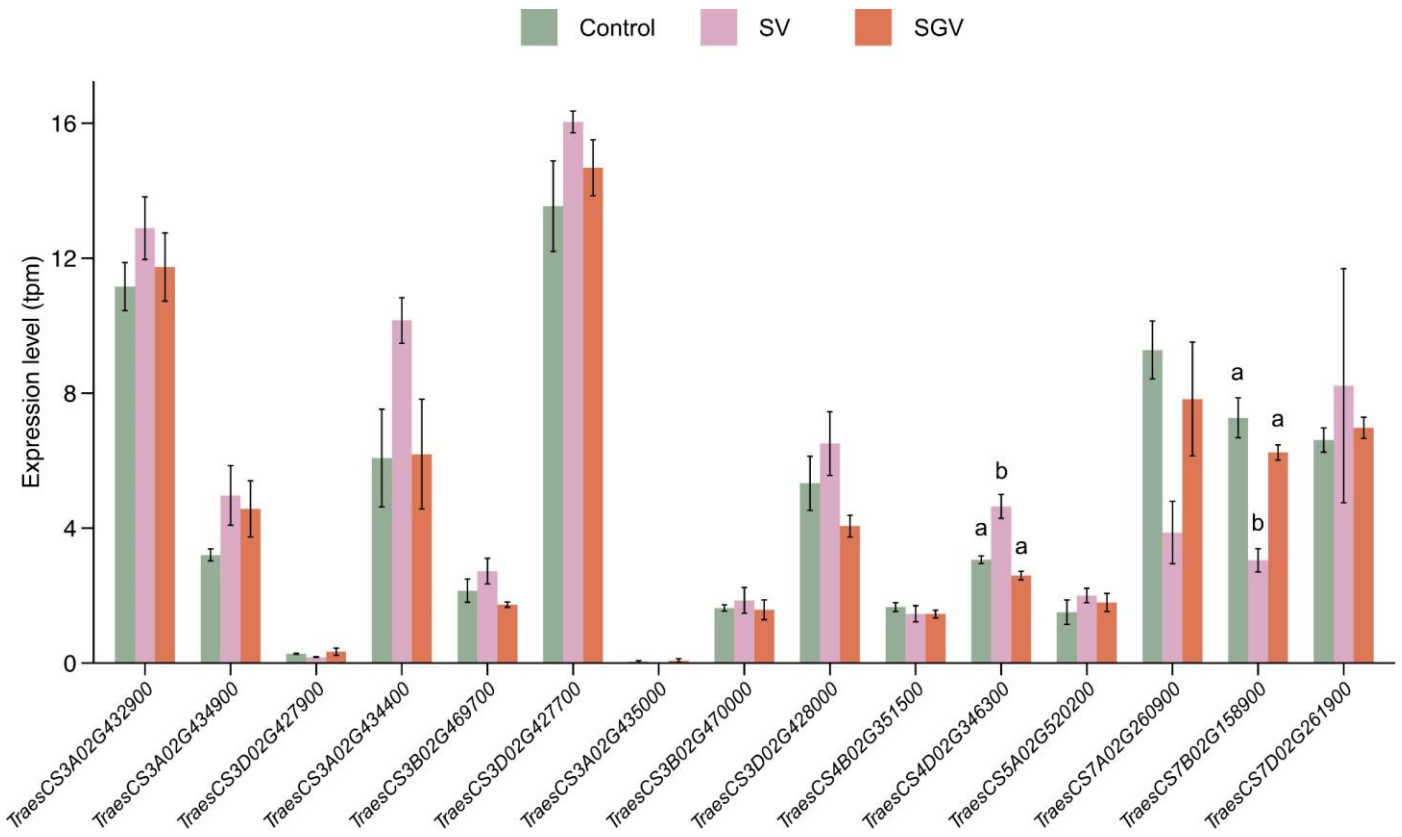


Figure 3.20: Expression of *ODDSOC* genes under the varying speed vernalization conditions. Control = no vernalization under SGV (grey); SV = speed vernalization (purple); SGV = seed green vernalization (red). Error bars represent standard error of the mean for 3 biological replicates, letters above bars indicate statistically significant differences in expression level ($p<0.05$; ANOVA and Tukey post-hoc test).

3.3.9 *OS1* and *OS2* have long non-coding antisense transcripts involved in regulation

RNA-seq analysis of winter wheat cv. Buster grown under 10°C and 14°C vernalization conditions (see section **2.3.1.4**) revealed that several *ODDSOC* genes had non-coding associated transcripts. More individual non-coding transcripts were identified at 10°C (10 unique transcripts) compared to 14°C (7 unique transcripts), however – excluding *TraesCS3A02G434400*, which was only transcribed at 10°C – the same genes were transcribed under both temperatures. These genes were *TraesCS3A02G432900*, *TraesCS3A02G435000*, *TraesCS3D02G428000*, *TraesCS4B02G351500*, *TraesCS4D02G346300*. Notably, *TraesCS3A02G432900* had 4 discrete transcripts present at 10°C (624, 2003, 1582, and 595 bp in length) compared to a single transcript (603 bp) at 14°C. Based on sequence hits, all non-coding transcripts were associated with the antisense strand and were more than 200 bp in length; for this reason, these transcripts were categorised as long non-coding RNA (lncRNA) antisense transcript (Kapranov et al., 2007). Transcripts partially aligned with both exons and either introns and/or untranslated regions immediately downstream of the gene. To validate this bioinformatics approach, a region of transcript common to all *TraesCS3A02G432900* lncRNA transcripts was confirmed with sequencing, confirming the presence of these transcripts in addition to the genes. To further investigate, the region downstream of the final exon was examined for binding sites. This uncovered 3 potential CArG box sites (motifs CTAAATATGG, CATTTTTTGG, and CATTTTTTGG) 350 bp downstream of the final exon and 122 bp downstream of the start of the associated lncRNA antisense transcript (**Figure 3.21**).

Preliminary expression analysis suggested a role for lncRNA in regulation of *TraesCS3A02G432900*; therefore, other variants of *TraesCS3A02G432900* associated lncRNA were examined. Of interest, 4 of the 5 lncRNA transcripts contained regions that aligned with the elongated first intron of the gene (**Figure 3.22**). Under colder, 10°C short-day conditions, this included an antisense transcript 1065 bp in length. A much smaller, truncated version (59 bp) of this transcript was identified at 14°C short-day conditions. The sequence of this 59 bp transcript didn't

contain potential CArG box sites, however several motifs associated with TFBS were identified – notably, the motifs TGTAC and ATGTACTTTGT associated with the *SQUAMOSA PROMOTER-LIKE 3* (*SPL3*) and *SPL11* genes in Arabidopsis. These genes have been reported to be involved in vegetative to reproductive transition and interact with *FT* and floral meristem identify genes *APETALA1* (*AP1*) (Jung et al., 2016; Xu et al., 2016). Other predicted TFBS identified included the motifs TGTCGACATC and TCAGTGTCGACATCA, which are associated with *DWARF AND DELAYED FLOWERING 1* (*DDF1*) and *DDF2* in Arabidopsis. *DDF1* and *DDF2* have been shown to be involved in regulation of timing of transition from vegetative to reproductive phase in Arabidopsis (Magome et al., 2004; Vyse et al., 2022).

Since the RNA-seq analysis only provided one timepoint at two different temperatures, additional expression analysis was carried out. To understand how these genes and their associated transcripts are temporally expressed, the expression was examined over a 24-hr period. First, expression of *TraesCS3A02G435000* was investigated in cv. Buster, as the RNA-seq analysis had identified the lncRNA transcripts in this cultivar. Under 10 °C short day (8 hr light/16 hr dark) conditions, both *TraesCS3A02G435000* and its associated lncRNA had a minimal level of expression (**Figure 3.23**). Slight peaks in expression were observed at zeitgeber ZT15 for *TraesCS3A02G435000* and at ZT8 and ZT15 for the lncRNA transcript. In contrast, growth under warmer 16 °C short day conditions led to an increase in expression of *TraesCS3A02G435000*. Here, expression was up regulated between ZT0 and ZT8; expression decreased between ZT8 and ZT10 but was up regulated once again between ZT10 and ZT15, before expression diminished by ZT20. Expression of the lncRNA transcript was low, with only a slight increase expression between ZT15 and ZT20 (**Figure 3.23**). In comparison to growth under 10 °C, the expression level of the lncRNA transcript was lower under 16 °C.

ACTTTGGTGAAACAAAATAGTGGCGGGAGCACGAGCGGAGAGAGCTCCTACGGACCTACG
GGGCAGGAGAATGGGAGGAAGGCTTGAGCAGAGCAAAGTTCCCCTGTGCTGTGTATGCATT
TGGTACTATGTTGCTTCCCCCAACACACTGTTCCCCAGTTAGGCTGTTAATGGGCAAGAATTA
GTTGCCAGGGCTGCTACTTGTAAGTTGTAAGGGAGGAAATTTTAGCGAAAGCGTGGCTTAATA
ATACTAATAATAATACAGTGTCCGTGCTATAAAGAGACCCCGTTGTACCTAGGACTGTGTGAG
TTGGCTAATCTAATTACGTATGCTTGTTCCTGGAGTAACTGTTGTTTTGATGTTGTGCCGATGC
TCATCTCTTATGATCTGCAGCTGAATGGCAAACTTTGGTGGGTTCAGAAGGGTCCCTCTGAAA
CTAAATATGGCTGTTATCGAGTGGCATGCTGCTGTATCTAGCAAGAGTTCGCGAGCCTGACC
TGACACGCATTTTTTGGAACCTC ACTTGGCTGCTTGCATTTTTTGGAGAAAAATGAAA

Figure 3.21: Putative CArG-boxes for the *OS1*-associated lncRNA antisense transcript. Nucleotides highlighted yellow represent potential reverse-complementary CArG-boxes for the antisense transcript. Regions highlighted grey represent the antisense transcript and bold red text represent final exon 5.

Figure 3.22: Schematic of *TraesCS3A02G432900* representing the position of the associated lncRNA antisense transcript in relation to this gene. Boxes represent exons (red), introns (green), upstream and downstream regions (turquoise), alignment of antisense transcript identified in RNA-seq analysis (purple), and the region amplified by qPCR primers from cDNA (orange).

Next, the lncRNA associated with *TraesCS3A02G432900* was examined. Expression of both *TraesCS3A02G432900* and the lncRNA localised to the end of the gene was investigated. This revealed that over a 24-hr period, contrasting expression profiles were observed under 10°C and 16°C short-day conditions in the winter cv. Buster (**Figure 3.24**). Under 10°C short-day conditions, both *TraesCS3A02G432900* and its associated lncRNA antisense transcript were expressed. The greatest level of expression was observed at ZT15, i.e., 15 hours after dawn (**Figure 3.24**). Overall, the lncRNA was expressed to a higher degree than *TraesCS3A02G432900*; this was most apparent at the ZT15 timepoint. In contrast, under 16°C short-day conditions, both *TraesCS3A02G432900* and its associated lncRNA was expressed to a significantly lower degree, suggesting that both *TraesCS3A02G432900* and lncRNA are upregulated under cold short-day conditions (**Figure 3.24**). This would align with putative characterisation of this gene as the *OS1* equivalent in bread wheat, as we would anticipate that a floral activator such as *OS1* would be up regulated under cold short-day conditions.

To investigate further, expression of both *TraesCS3A02G432900* and the associated lncRNA was studied in wheat cultivars containing deletions within the region of the first intron that matched with the antisense transcript. This gene had two main antisense transcripts associated with it: transcripts localised to the first intron region and transcripts localised around the final exon (**Figure 3.25A**). First, expression of the transcript localised to the first intron was investigated. Expression of this transcript was not detected in either cv. Buster or cv. Julius, indicating that an intronic deletion in Julius within the transcribed region of the antisense transcript didn't impact expression level of the transcript (**Figure 3.25B**). Next, the transcript localised to the final exon was examined. In stark contrast to the intronic transcript, the transcript associated with the final exon was up regulated in the cv. Buster (**Figure 3.25B**). Expression in cv. Julius that contained a downstream deletion was significantly lower ($p<0.01$; Student's t-test). Furthermore, there was a statistically significant difference ($p<0.01$; Student's t-test) in expression level of the antisense transcript compared to *TraesCS3A02G432900* in cv. Buster (**Figure 3.25B**). Conversely, there was no significant difference in expression of the antisense transcript and *TraesCS3A02G432900* in cv. Julius (**Figure 3.25B**).

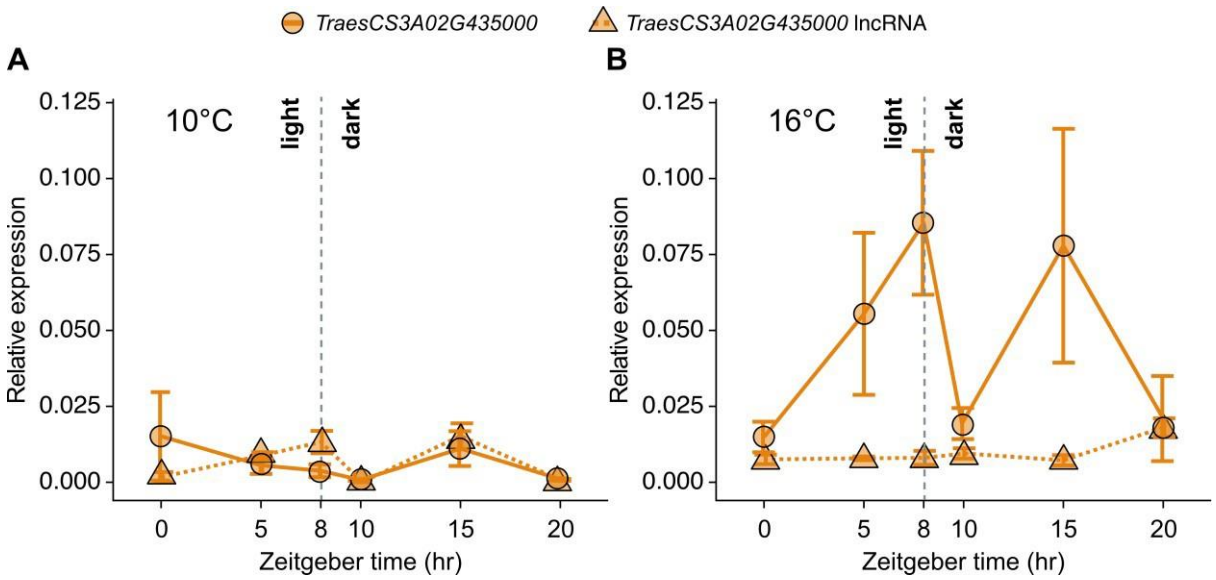


Figure 3.23: Expression of *OS2* gene *TraesCS3A02G435000* and the associated lncRNA transcript. Expression of *TraesCS3A02G435000* (circle) and lncRNA transcript (triangle) was analysed under (**A**) 10 °C and (**B**) 16 °C short day (8 hr light/16 hr dark) conditions over a 20-hr cycle. Error bars represent standard error of the mean for 3 biological replicates.

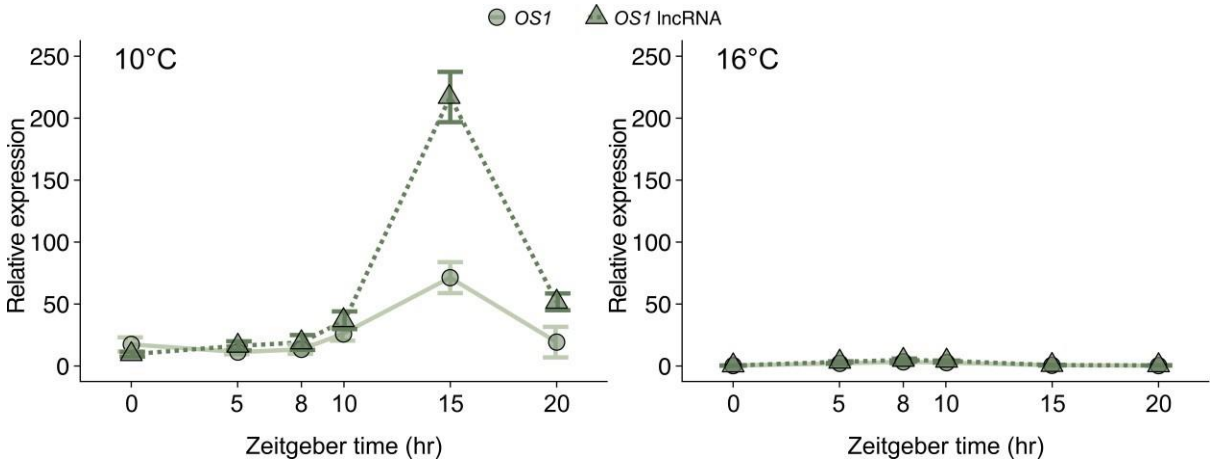


Figure 3.24: Expression of *OS1* and its associated lncRNA transcript under cold (10°C) and warm (16°C) short day conditions. Expression of *OS1/TraesCS3A02G432900* (circle, solid line) and the associated lncRNA transcript (triangle, broken line) was measured over a 20-hr cycle under either 10 °C or 16 °C 8 hr light/16 hr dark conditions.

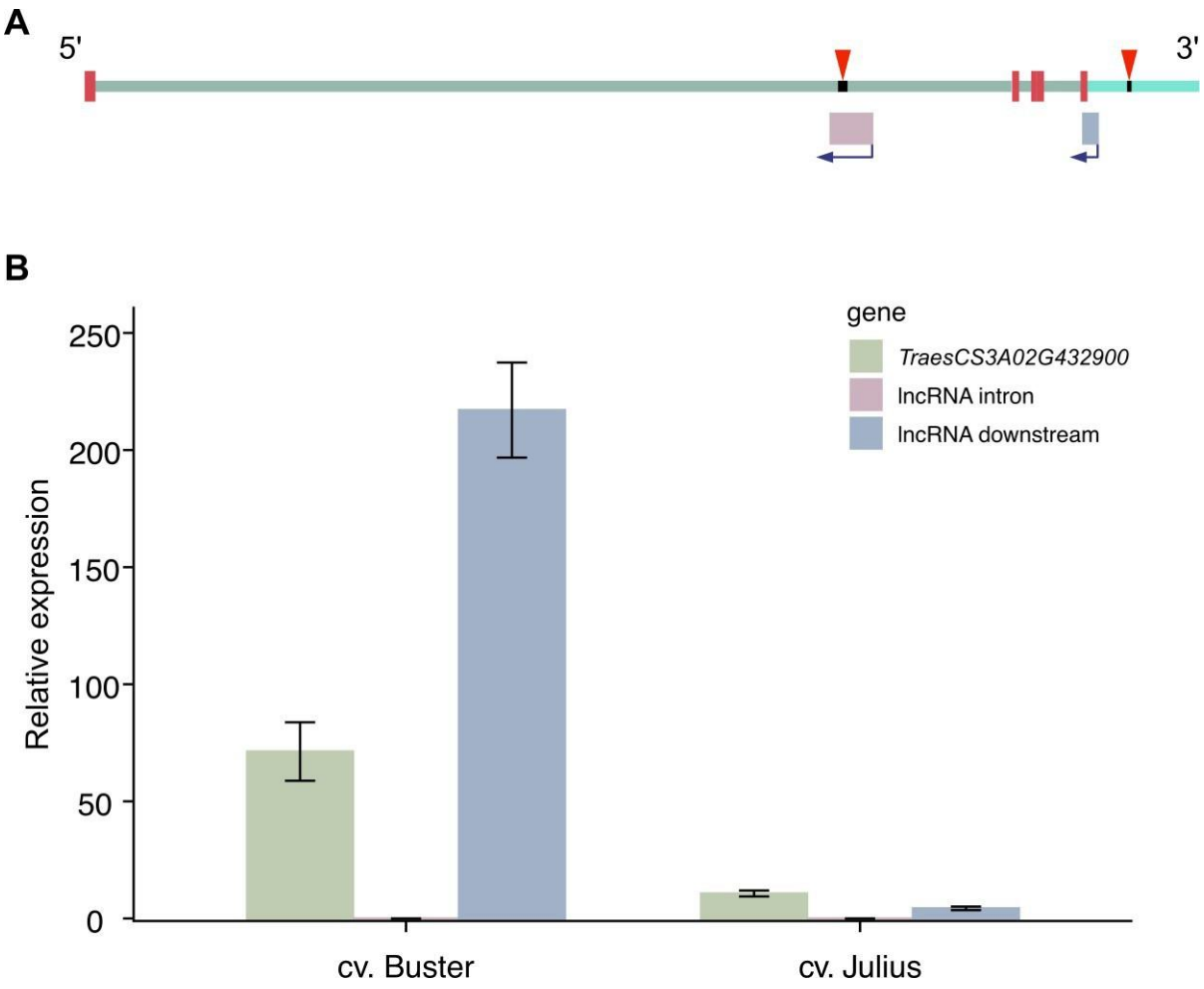


Figure 3.25: Regulation of *TraesCS3A02G432900* via the associated lncRNA antisense transcript. (**A**) Schematic showing the position of the intronic-associated lncRNA transcript (purple) and the downstream-associated lncRNA transcript (blue) relative to the *TraesCS3A02G432900* genomic region. Red boxes represent exons, introns (green), downstream regions (turquoise), and the associated intronic (purple) and downstream (blue) antisense transcript identified in RNA-seq analysis. (**B**) *TraesCS3A02G432900* expression in winter wheat cultivars Buster (no deletions) and Julius (deletions within the intronic and downstream genomic region that overlaps with the antisense transcripts).

3.4 Discussion

3.4.1 Multiple putative *ODDSOC* genes are present in bread wheat

In section 3.3, we explore the parallels between *FLC*-like genes and *OS2* genes in bread wheat and identify 14 putative *OS2* gene candidates. Previously, these 14 genes had been identified in MADS-box gene classification analysis in bread wheat (Raza et al., 2021; Schilling et al., 2020). In these analyses, the *OS2* candidate genes were categorised either as a *FLC*-like subfamily (Schilling et al., 2020) or as part of a broader MIKC*-subfamily (Raza et al., 2021). Furthermore, we also uncovered a gene that is strong candidate for *OS1* in bread wheat. To our knowledge, *OS1* has yet to be defined in bread wheat; similarly, this *OS1* candidate gene, *TraesCS3A02G432900*, has not been reported to be involved in temperature or vernalization studies in wheat. For this reason, we aimed to characterise these potential *ODDSOC* genes in the context of their role as floral activators (*OS1*) or floral repressors (*OS2*) in vernalization.

The list of 14 *OS2* possible candidate genes consisted of 5 groups of homoeologues. This suggests that these genes underwent positive selection pressure, with homoeologues from the 3 sub genomes being conserved.

3.4.2 The expanded family of *ODDSOC* genes show features of neofunctionalisation

Since 6 potential *OS2* genes were identified to have differential expression out of over 45,000 genes in GEM analysis, this strongly indicates that these genes have a role in the vegetative to reproductive transition in bread wheat. It's worth noting that other genes known to be involved in the transition were not identified within the top 10 candidates; this could potentially be due to the developmental stage at which tissue was sampled for expression, suggesting that these 6 *OS2* genes had a particularly relevant function at this early (i.e., GS13 or 3-leaf) stage of development. Surprisingly, despite the high level of similarity between amino acid sequences, the additional 8 putative OS2 genes were not identified in the GEM analysis. This is compelling as it could indicate that even with shared features, not all these genes have the same function. This could also provide an explanation for why these 14 genes have been well-conserved in the bread wheat genome. Expansion of gene number isn't uncommon in cereals; for example, 12 discrete

FT proteins have been identified in monocots and have been reported to have different roles, with some members of this group facilitating flowering and others repressing flowering (Bennett & Dixon, 2021; Halliwell et al., 2016; Pieper et al., 2021; Shaw et al., 2019).

In contrast, no homoeologues were identified for potential *OS1* candidate *TraesCS3A02G432900*. From an evolutionary perspective, the A sub genome to which *TraesCS3A02G432900* belongs can be considered the original genome, with the subsequent hybridisations contributing the B and D sub genomes. This would suggest that the B and D homoeologues of *TraesCS3A02G432900* weren't conserved/didn't undergo the same selection pressure as this gene. In bread wheat, expression across homoeologues has been reported to be relatively balanced (Alabdullah et al., 2019). However, biased expression patterns across gene homoeologues is possible (Alabdullah et al., 2019; Leach et al., 2014). Therefore, it's possible that *TraesCS3A02G432900* was the dominant gene, and this contributed to its conservation compared to its other homoeologues. However, this raises the question as to why only one floral activator (i.e., *TraesCS3A02G432900*/*OS1*) was identified compared to 14 floral repressors (i.e., *OS2*). Despite the lack of identified homoeologues, *TraesCS3A02G432900* shared similarities such as conserved domains and a large first intron with the 14 potential *OS2* genes. Furthermore, prior analysis had grouped *TraesCS3A02G432900* with the putative *OS2* genes (Raza et al., 2021; Schilling et al., 2020). In addition, it's probable that although phylogenetic analysis has grouped *TraesCS3A02G432900* (as *TaAGL42*) with *BdOS1* and *HvOS1*, and *OS2* in bread wheat (as *TaAGL33*, *TaAGL41*) with *BdOS2* and *HvOS2*, division in function isn't clear cut. For this reason, the 14 potential *OS2* genes could also contain genes that function as floral activators. This is highly reminiscent of what has been observed with the *FT* gene family in cereals; studies in barley and wheat have identified *FT* genes, such as *FT3*, that function as floral promoters despite an amino acid sequence that would suggest a function as a floral repressor (Zikhali et al., 2014).

A further possibility for the single identified *OS1* is that the function of this gene is highly similar to that of the dominant floral activator *VRN1*, reducing the efficacy of 3 homoeologous *OS1* genes. For example, both *OS1* and *VRN1* are active very early

in floral transition, are MADS-box proteins that potentially interact with the same targets, and both show a clear sensitivity to temperature. In this regard, it raises the question of what unique features does OS1 have in comparison to VRN1? One possibility could be exact stage at which OS1 functions or the environmental adaptation that it provides.

However, there was an intriguing difference in the MADS-box domain of *TraesCS3A02G432900* compared to the other genes analysed. The *TraesCS3A02G432900* MADS-box domain was truncated and lacked α-helices and/or β-strands (**Figure 3.5; Figure 3.7**); since the MADS-box domain is involved in protein interactions, it would be informative to further investigate whether this variation in the domain influences specific protein or DNA-binding. This, in combination with genetic characterisation, would ultimately contribute to defining the role of this gene as a floral activator instead of a floral repressor. Similarly, the putative OS2 proteins TraesCS7A02G260900, TraesCS7B02G158900, and TraesCS7D02G261900 had a unique region directly upstream of the MADS-box domain spanning approximately 40 amino acids (**Figure 3.6**). Since this region was only conserved in these 3 genes, this could indicate that this contributes to function of these genes.

3.4.3 *ODDSOC* genes contain a number of regulatory features previously linked to environmental adaptation

Conversely, a feature common to 12 of the 15 potential *ODDSOC* genes was an especially long intron over 17 kb in length (**Figure 3.4**). A large intron would indicate that it has undergone positive selection pressure and therefore potentially contains regulatory features. Since large introns are involved in the regulation of key vernalization genes *VRN1* in cereals and *FLC* in Brassicaceae, these introns could also be involved in the regulation of the genes described in this chapter. Analysing the introns of *ODDSOC* genes for the presence of regulatory features would be informative to elucidate whether these genes are functioning in a similar manner to *VRN1* and *FLC*. Further fortifying the link between the classification of these genes as part of the *FLC*-like family (Schilling et al., 2020), the presence of antisense transcripts is especially compelling. As previously mentioned, lncRNA antisense

transcripts are involved in the regulation of *FLC*; therefore, the ubiquity of predicted lncRNA antisense transcripts localised to the end of multiple putative *ODDSOC* genes could indicate a similar regulatory role.

3.4.4 *ODDSOC* genes show specificity in homo- and hetero-dimerization

Using yeast two-hybrid assays, protein-protein interactions between the ODDSOC proteins revealed variation in proteins that interacted with each other (**Figure 3.10**-**Figure 3.13**). This is interesting as the amino acid sequences shared a high level of similarity; for this reason, we would anticipate to see these proteins either interacting or not interacting with each other. However, this wasn't what we observed. Protein- protein interactions varied even within homoeologues; for example, TraesCS3B02G470000 didn't form homodimers with itself or with its homoeologue TraesCS3A02G435000 but did with its other homoeologue TraesCS3D02G428000 (**Figure 3.12**). This could indicate that despite the similarity that these proteins share, they occupy individual roles. Building upon this, there was variation in interactions with different versions of the VRN-A1 protein. Here, we observed several prospective OS2 proteins forming heterodimers with the wild type VRN-A1 protein (**Figure 3.15**). However, these same OS2 proteins did not interact with a version of the VRN-A1 protein associated with higher vernalization requirement. This is compelling as it could suggest that not only are these potential OS2 genes involved in vernalization, but also can preferentially bind based on the level of vernalization requirement.

Since it has been characterised in other species as a floral activator like *VRN-A1*, it's surprising that TraesCS3A02G432900 didn't interact with either variant of VRN-A1. One possible explanation for this is that *TraesCS3A02G432900* and *VRN-A1* function independently and potentially at different stages of development. Additionally, it is interesting to note that certain OS2 proteins didn't interact with either form of VRN-A1, possibly suggesting that these genes aren't directly involved in vernalization or that their role in vernalization doesn't involve interaction with the VRN-A1 protein.

3.4.5 *ODDSOC* gene expression varied between genes and with respect to growth habit

As we had observed specificity in protein interaction as well as a number of unique features in the *ODDSOC* genes, we decided to explore this further using expression data across a diverse wheat panel. YoGI expression analysis provided further indication that the potential *ODDSOC* genes occupied varying roles in vernalization (**Figure 3.16**-**Figure 3.18**). For example, the 6 *OS2* genes originally identified in the vegetative to reproductive transition eQTL all had higher expression in winter bread wheat lines compared to spring. Combined with expression under different temperature and photoperiod, we observed that potential *OS2* gene *TraesCS3B02G470000* had higher expression at a temperature of 16 °C in a winter bread wheat cultivar (**Figure 3.20**). However, increasing the photoperiod length resulting in the gene being down regulated even at this same temperature. This potentially indicates that this gene is most active when days are getting shorter and temperature is lower than in the height of summer but still not the typical daily temperature experienced throughout winter – that is, at the end of autumn/start of winter when floral transition is still repressed and vernalization has yet to be completed.

Chapter 4: Exploring regulation of *ODDSOC* genes in wheat

4.1 Introduction

To further understand how *ODDSOC* genes in wheat respond to their environment, I decided to investigate expression of these genes under natural vernalizing conditions in the field. In addition, as mentioned in the previous chapter, a common feature of *ODDSOC* genes is an abnormally long first intron. To further our understanding of how these genes are regulated, I chose to next investigate whether the large first intron of *ODDSOC* genes played a role in gene regulation.

4.1.1 Looking to classical vernalization to understand *ODDSOC*

Today, artificial vernalization is standard practice in both winter breeding programmes and vernalization research. However, the core vernalization pathway in both Brassicaceae and cereals evolved under natural conditions, reflected in the adaptability of cultivars to a wide range of winter conditions worldwide (Chouard, 1960). In recent years, exploration into how vernalization proceeds in the field has contributed to understanding of how the pathway operates, including the predicted impact climate change will have on the development of winter crops in the field. Since the previous chapter introduced a set of *FLC*-like genes that were potential *ODDSOC* candidates, *FLC* activity under field conditions is an intriguing area to pursue.

In winter *B. napus* cultivars grown in the field over winter, orthologues of *FLC* were down regulated during October when daily temperatures were between 10 °C-15 °C (O'Neill et al., 2019). Following this, floral transition proceeded during November, indicating that the vernalization pathway is initiated earlier than anticipated and before exposure to lower daily temperatures of 5 °C-10 °C (O'Neill et al., 2019). Furthermore, when the temperature of field plots was artificially raised by 5 °C in October, *FLC* transcript levels increased and floral transition was delayed (O'Neill et al., 2019). This didn't have a significant impact on completion of vernalization, with plants still observed to flower the following spring; however, field plots that had received a 5 °C increase in temperature during October flowering 5-7 days later than unheated field plots (O'Neill et al., 2019). When plots were heated to increase

temperature by 5 °C-6 °C shortly after floral transition (i.e., during flower bud development) of winter *B. napus*, this resulted in delayed bolting and pod and seed set (Lu et al., 2022). Furthermore, floral initiation in the field removed the epigenetic mark H3K4me3 from the *FLC* locus and this modification remained subdued for the remainder of winter; however, winter warming of the field plot interrupted this process and the H3K4me3 mark reappeared, effectively reversed the epigenetic silencing of *FLC* (Lu et al., 2022).

In Arabidopsis, autumnal germination in the field had a lower rate of success in a line with an *FLC* allele linked to low-expression and a non-functional *FRI* allele, compared to a near isogenic line with a high-expressing *FLC* allele (Chiang et al., 2009). This would suggest that in addition to mediating premature floral transition in autumn, *FLC* is also involved in promoting germination for winter lines sown in autumn. Arabidopsis field studies across three climatically discrete sites (temperate oceanic with mild winter, warm-summer continental with a cold winter, and subarctic, respectively) revealed that variation in *FLC* underpinned the vernalization response in different environments (Hepworth et al., 2020).

Field studies from these same three field sites was also informative in deciphering activity of the *FLC*-associated antisense transcript *COOLAIR* (Zhao et al., 2021). Here, *COOLAIR* transcription was up regulated when the field site experienced the first frost (Zhao et al., 2021). Furthermore, *COOLAIR* was not exclusively up regulated when temperatures dropped below 0 °C; instead, a period of fluctuating temperatures, including freezing temperatures, strongly up regulated *COOLAIR* (Zhao et al., 2021).

4.1.2 The impact of introns on gene regulation

In addition to upstream promoters and enhancers, introns can also have a substantial impact on gene expression (Rose, 2008). In some cases, alternative promoter or enhancer elements within introns regulate gene expression; in other cases, a process known as intron-mediated enhancement (IME) is involved in boosting mRNA accumulation (Rose, 2008; Shaul, 2017). Although intron-driven regulation is not as well-characterised as its 5’ promoter counterpart, it can still occupy an important role in regulation. For example, deletion of the promoter region

of Arabidopsis gene *TRYPTOPHAN BIOSYNTHESIS1* (*TRP1*) had no effect on expression level, as long as the IME-associated intron remained intact (Gallegos & Rose, 2017). In Arabidopsis, computational analysis estimates that based on an IME score, approximately 15% of genes are regulated by introns, with a similar percentage projected in higher plant species (Gallegos & Rose, 2015).

4.1.3 MIKC-type MADS-box genes can be regulated by their introns

Structural analysis of MADS-box genes indicates that MIKC-type MADS-box genes in plants contain multiple introns, with a trend towards the first intron being the longest in length (Pařenicová et al., 2003). A good example of this is *FLC*. *FLC* is reported to be regulated by its intron: as previously described in Chapter 1, the long intronic non-coding RNA *COLDAIR* is transcribed from the first intron of *FLC* and in response to vernalization epigenetically represses the gene (Heo & Sung, 2011). Furthermore, as part of a multi-step epigenetic approach, the modification H3K27me3 spreads over the entire *FLC* locus including the introns to promote long-term silencing of the gene (Yang et al., 2017). In addition to histone modification, variations within the intron sequence have also been shown to affect regulation of *FLC*. For example, a cytosine to thymine nucleotide change within the first intron and 585 bp downstream of the TSS for *FLC* affects binding of *VIVIPAROUS1/ABI3-LIKE1* (*VAL1*) to the RY motif associated with this nucleotide change (Qüesta et al., 2016). Consequently, *VAL1* cannot bind, PHD-PRC2 is unable to nucleate the region, and *FLC* isn't silenced when this intronic variation is present (Qüesta et al., 2016). Conversely, FLC protein binds to a region within the first intron of Arabidopsis *FT* (Helliwell et al., 2006; Searle et al., 2006). A putative CArG box site the consensus sequence TTTTCCTTTTTTGGGGTA within this intronic region facilitates this interaction, resulting in the repression of *FT* (Helliwell et al., 2006). The *FLC* homolog *BoFLC2* in *Brassica oleracea* (*B. oleracea*) can also be regulated by the first intron; specifically, the presence of a 215 bp deletion in the first intron reduces the rate of silencing of *BoFLC2*, resulting in a higher vernalization requirement and later flowering (Li et al., 2022).

4.1.4 The role of intron regulation in flowering genes in cereals

MIKC-type MADS-box genes involved in cereal flowering have been shown to be regulated by their introns, similar to Brassicaceae. However, other cereal flowering genes can also be regulated by their introns, indicating that this mechanism isn't restricted to MIKC-type MADS-box genes in cereals. For example, *Ppd-1* isn't a MADS-box gene, but its intron can play role in gene functionality. This can been in allelic variation for *Ppd-D1*; in hexaploid wheat cv. Paragon, the first intron of *Ppd-D1* contains a large 4.8 kb transposon insertion (Beales et al., 2007; Shaw et al., 2012). This resulted in frameshift mutations and premature stop codons within the transcript, ultimately producing a non- or weakly-functional protein lacking a CCT domain (Shaw et al., 2012).

To complement this, MIKC-type MADS-box genes *VRT2* and *VRN1* in cereals can be regulated by their introns. Rearrangements within the first intron of *SVP*-like gene *VRT-A2* in both tetraploid (*T. turigidum, T. polonicum*) and hexaploid wheat (*T. aestivum*, *T. petropavlovskyi*) is associated with increased expression of the gene (Adamski et al., 2021; Liu et al., 2021). This increase in expression level was linked to elongated outer glumes and grains (Adamski et al., 2021; Liu et al., 2021).

VRN1 in cereals has regulatory sites outside of the promoter and within the first intron. *VRN-A1*, *VRN-B1*, and *VRN-D1* in hexaploid wheat, as well as *VRN-H1* barley all have large deletions within their first intron (Fu et al., 2005). In barley, the large first intron has a role in regulation of *VRN-H1*, playing a role in vernalization requirement and growth habit (Cockram et al., 2007; Oliver et al., 2009; Szűcs et al., 2007; von Zitzewitz et al., 2005). Large deletions (e.g. 5.2 kb in spring barley cv. Morex; 6.4 kb in spring barley cv. OWB-D) in the *VRN-H1* first intron have been linked to determining growth habit, with intron deletions associated with spring habit in barley (Fu et al., 2005; Oliver et al., 2009).

Similar to barley, *VRN1* in wheat also has an intronic deletion associated with regulation. Deletions within the first intron of *VRN-A1*, *VRN-B1*, and *VRN-D1* overlap in a 4 kb region, within which there is a highly conserved 2.8 kb portion containing regulatory elements involved in vernalization (Fu et al., 2005). The

intronic deletion, in combination with the deletion in the *VRN1* promoter, drastically reduces or removes the vernalization requirement, resulting in a spring growth habit (Fu et al., 2005). This intron contains a binding site for glycine-rich RNA-binding protein 2 (TaGRP2) which inhibits *VRN1* expression (Kippes et al., 2015; Xiao et al., 2014). Cultivars with a low vernalization requirement, such as cv. Claire and winter cv. Jagger, contain a SNP within the TaGRP2 intron binding site (Kippes et al., 2015). In contrast, winter cultivar Hereward has a high vernalization requirement and no SNP in this site (Kippes et al., 2015). Exposure to prolonged low temperatures interrupts TaGRP2 from binding this site, up regulating *VRN1* expression (Xiao et al., 2014; Xu et al., 2019).

4.1.5 The introns of MIKC-type MADS-box genes can contain regulatory elements that elucidate their function

A common feature shared by *FLC* and *VRN1* in cereals is a large first intron (3,493 bp and 4,389-9,988 bp, respectively) that contains crucial regulatory elements. Since the majority of *ODDSOC* also have a large first intron (17,345-34,055 bp; if including the second intron of *TraesCS7A02G260900* that aligns with the first intron of most other *ODDSOC* genes, maximum intron size is 39,138 bp), I considered that this region could also be involved in the regulation of these genes. Furthermore, the deletion within the first intron of *VRN1* in cereals has a key role in determining vernalization requirement. For this reason, the introns of the *ODDSOC* genes previously characterised in chapter 4 were investigated to establish whether they have a role in *ODDSOC* regulation.

4.2 Materials and Methods

4.2.1 RNA-seq of winter field trial

4.2.1.1 Establishment of winter field trial

The winter wheat cv. Hereward were grown over the winter of 2019 at the University of Leeds Spen Farm field site, Tadcaster, UK. Seeds were drilled 20/10/2019 and seedlings emerged 31/10/2019; 6 replicate plots were grown in a randomized design. Meteorological data for the duration of the trial was obtained from the COSMOS-UK (Stanley et al., 2023) datalogger at the field site. Daylength in hours was calculated from the field site latitude and date using the function 'daylength' in the R package 'geosphere' (https://github.com/rspatial/geosphere). Daylength and average daily temperature was plotted using the R package 'ggplot2'.

Leaf tissue was sampled for expression analysis after 5, 7, 10, 13, and 16-weeks of growth in the field at approximately 3 hours post-dawn. RNA extraction was carried out alongside Matthew Turner. Tissue was disrupted using the TissueLyser II (Qiagen) and RNA was extracted using the Spectrum™ Plant Total RNA Kit (Sigma-Aldrich) as described in section **2.2.4.1** and eluted in 50 µL dH_2O. Post-extraction, RNA was purified using the Monarch® RNA Cleanup Kit (NEB) following the manufacturer's instructions. Purified RNA was treated with the TURBO DNA-free™ Kit (Invitrogen) as per the manufacturer's protocol to remove residual DNA from the sample. RNA concentration was evaluated using the NanoDrop 1000 Spectrophotometer (ThermoFisher) and diluted to approximately 120 ng/µL and a final volume of 40 µL. Library preparation and Illumina-based sequencing of RNA samples was carried out by Novogene, UK.

In addition to sampling tissue, 3 plants were dissected and the meristems imaged using a Keyence VHX-7000 digital microscope and apex length was measured using ImageJ. To investigate the flowering response, 4 plants were transferred from the field plot to 20 °C long-day (16 hr light/8 hr dark) glasshouse conditions after 5, 7, 10, 13, and 16-weeks of growth in the field. Flowering was scored at half-ear emergence (Zadok's growth stage Z55) and total spikelet number per ear was recorded.

4.2.1.2 Bioinformatic analysis of RNA-seq dataset

Transcripts from the winter field trial RNA-seq datasets were pseudoaligned to the Chinese Spring reference genome following the method outlined in section **2.2.4.3**. Furthermore, differential expression, soft clustering, and GO analysis was carried out as previously described in section **2.2.4.3**.

4.2.2 Identification of *ODDSOC* variation within the Watkins collection

WatSeq haplotype maps displaying variation within the gene and the regions upstream and downstream across 2,101 wheat lines were provided by Professor Shifeng Cheng and Dr Simon Griffiths (see chapter 3 section **3.2.5**). Haplotype maps included habit (spring, winter, or unknown) and location of origin of each line. Average monthly temperature for source location was obtained from Berkeley Earth (https://berkeleyearth.org/data/). Boxplots to represent data were produced using the R packages 'ggplot2' and 'ggpubr'.

4.2.3 Geographical mapping of lines

Base map of the world was generated using the R package 'maps' (https://github.com/adeckmyn/maps). Site locations for wheat lines of interest were obtained from latitude and longitude values included in the haplotype maps, as were habit types for each line.

4.2.4 Motif identification within the first intron of *ODDSOC* genes

Intron sequences for all 15 potential *ODDSOC* genes were obtained from Ensembl Plants (http://plants.ensembl.org/index.html). For genes *TraesCS7A02G260900*, *TraesCS7B02G158900*, and *TraesCS7D02G261900*, the sequence of the second intron was used as this was the largest intron for these genes. Full intron sequences were analysed using the PlantPAN 4.0 Promoter Analysis tool (http://plantpan.itps.ncku.edu.tw/plantpan4/promoter_analysis.php) to identify predicted TFBS within the introns. As this PlantPAN release doesn't yet include a wheat-specific transcription factor database, binding motifs from the *A. thaliana*, *B. distachyon*, *G. max*, *O. sativa*, *S. bicolor*, and *Z. mays* databases were used. Predicted TFBS were filtered to isolate motifs that were present in all intron sequences.

4.2.5 Motif identification within *TraesCS3B02G470000* intron deletions

To investigate binding motifs within the intronic deletions, the SNP call location from the *TraesCS3B02G470000* haplotype map was used to infer the general position of the deletions within the intron. Based on the general position, the exact sequence of the deletions was determined using the Ensembl Cactus alignment feature (Paten et al., 2011) to align wheat cv. Julius which contains both intronic deletions and cv. Chinese Spring, the reference genome used in the WatSeq initiative. The SNP call location was used to select the Cactus alignment blocks that covered the deleted regions.

Next, the complete sequences of both deletions were individually scanned for TFBS using the PlantPAN 4.0 Promoter Analysis tool. Predicted motifs of interest were cross-referenced using the Tomtom Motif Comparison tool (https://meme-suite.org/meme/tools/tomtom) that uses an Arabidopsis TFBS database, and the JASPAR transcription factor database (https://jaspar.genereg.net/) (Castro-Mondragon et al., 2022; Gupta et al., 2007).

4.2.6 Yeast one-hybrid assay

4.2.6.1 Cloning motif fragment into yeast vector

The DNA motif selected to test in the yeast one-hybrid system was too short (11 bp) to extract from an electrophoresis gel post-PCR amplification. For this reason, an oligonucleotide was designed comprised of the DNA motif flanked either side by the restriction digest sites for EcoRI and SacI as well as several random spacer nucleotides to facilitate effective digestion by the restriction enzymes. A complementary oligonucleotide was also designed, and the sequences of these primers were CGCGAATTCTGTTCGAACACGAGCTCGCG and CGCGAGCTCTGTTCGAACACGAATTCGCG. The oligonucleotides were annealed together using PCR. Crucially, no template DNA was added to the reaction to prevent the amplification of non-target products; instead, 5 µL of each primer (100 µM concentration) and 10 µL of 10X Crystal Buffer (Jena Biosciences) was added to a 0.2 mL PCR tube. To encourage formation of double-strand primer dimers, a Touchdown PCR was carried out with an initial annealing temperature of 70 °C that was reduced by 0.5 °C every cycle.

Once annealed together, the primer dimer consisting of the motif and flanking restriction enzymes sites was digested with the restriction enzymes *Eco*RI-HF and *Sac*I-HF (NEB). The restriction double-digest protocol was previously described in section **3.2.4.1**. In parallel, the vector pHIS-2 (Clontech) was digested with *Eco*RI-HF and *Sac*I-HF in the same fashion. The linearised plasmid was run on a 1% agarose gel, excised, and extracted using the Monarch® DNA Gel Extraction Kit (NEB) following the manufacturer's instructions and eluted in 30 µL dH_2O. The linearised pHIS-2 vector and digested motif fragment were ligated together using the Quick Ligation™ Kit (NEB) as per the manufacturer's instructions. Once ligated, the reaction was transformed into NEB® 10-beta Competent *E. coli* (High Efficiency) cells as previously described in section **3.2.4.1**. The transformant was plated on media supplemented with IPTG, X-Gal, and the antibiotic kanamycin (final concentration 50 µg/ml). After incubating upside-down at 37 °C overnight, colonies were selected and cultured overnight at 37 °C with shaking in 5 mL LB supplemented with 7 µL of 35 mg/mL kanamycin (final concentration 50 µg/ml). Plasmids were isolated from the culture using the Monarch® Plasmid Miniprep Kit (NEB) and eluted in 50 µL elution buffer. The presence of the correct insert in the vector was confirmed by sequencing, as described in section **3.2.3.2**.

4.2.6.2 Testing protein-DNA interaction with yeast one-hybrid screening

To examine interaction between the DNA motif and proteins of interest, a yeast one- hybrid system was used. Briefly, a DNA target sequence is cloned into the multiple cloning site (MCS) of the pHIS yeast vector (Clontech) and the CDS of a DNA- binding protein is cloned into the pGAD424 vector (Clontech). As with the yeast two- hybrid system detailed in section **3.2.4.1**, proteins are expressed as fusions to the GAL4 AD in pGAD424. The pHIS2 MCS is upstream of *HIS3*, therefore interaction between the protein and the target sequence (here, the motif TGTTCGAACAC) facilitates expression of this reporter gene. Therefore, yeast growth on media lacking leucine, tryptophan, adenine, and histidine confirms protein-DNA sequence interaction.

The pHIS2 vector containing the binding motif was transformed into the Y187 *S. cerevisiae* yeast strain. Transformation was performed following the method described in section **3.2.4.1**. Once transformed, the transformant was plated on

synthetic drop-out media lacking the amino acid tryptophan and incubated upside-down at 29 °C until colonies appeared. Following this, a co-transformation was performed between the yeast transformed with the pHIS2-DNA motif and pGAD424- gene, and the transformant was plated on synthetic drop-out media lacking the amino acids leucine and tryptophan. After upside-down incubation at 29 °C, colonies were selected and plated on synthetic drop-out media lacking leucine, tryptophan, adenine, and histidine and supplemented with 3-AT to counteract leaky *HIS3* expression. The addition of 3-AT followed the description in section **3.2.4.1**; however concentrations of 10 mM, 15 mM, and 20 mM 3-AT were used as the Y187 yeast strain has a higher basal level of His compared to the AH109 strain used in the yeast two-hybrid system.

4.2.7 In-silico expression analysis of *TraesCS3B02G470000* within diverse populations

The WatSeq initiative haplotype map for *TraesCS3B02G470000* provided the subset of lines containing both deletions in the first intron, the habit, and the Watkins Stabilised Collection of Tetraploid and Hexaploid Landrace Wheats ID code for each line. The ID code was cross-referenced with the YoGI expression dataset provided by Dr Andrea Harper to match presence or absence of the intron deletions from the WatSeq haplotype map with the expression level of *TraesCS3B02G470000* from the YoGI dataset. Expression barplots were generated with R packages ‘ggplot2’, with lines ordered from lowest to highest expression. To establish whether data were normally distributed, a Shapiro-Wilk test was performed. Based on the result of the Shapiro-Wilk test, either a Student’s t-test (normal distribution) or Mann-Whitney U test (non normal distribution) was used.

4.2.8 Intron deletion experiment

4.2.8.1 Germplasm

To test regulation via the deleted regions of the *TraesCS3B02G470000* first intron in spring cultivars, the following spring lines from the Watkins collection were grown: WATDE0005, WATDE0070, WATDE0018 containing the deletions within the first intron, and WATDE0104 containing the wild-type version of the first intron.

For the winter experiment, the following winter lines were grown: WATDE0848, WATDE0638, WGED0223 containing the deletions within the first intron, and wild-type line WATDE0063. Selection of all lines was made based on the WatSeq initiative haplotype map for *TraesCS3B02G470000*.

4.2.8.2 Experimental conditions

Lines were grown under either 10 °C or 16 °C short-day (8 hr light/16 hr dark) in Sanyo MLR-352 plant growth chamber. Meristems were dissected and imaged as described in Chapter 3 at Zadoks cereal growth stage Z12, Z14, Z15, and Z16. Leaf tissue was sampled and RNA was extracted, cDNA was synthesised, and qPCR analysis was carried out as previously detailed in section **2.2.4.1**. Sample processing and qPCR analysis was partly carried out by Harry Taylor.

4.2.9 Development and analysis of transgenic *TraesCS3B02G470000* lines

4.2.9.1 Growth conditions and experimental set up

Transgenic *TraesCS3B02G470000* wheat lines were generated by NIAB, UK. To summarise, wheat calli of the spring bread wheat cv. Cadenza were transformed with a construct with the transgene *TraesCS3B02G470000* CDS and 2.5 kb upstream region containing the native promoter elements, a HA (hemagglutinin) tag, and terminator sequence. Control lines were transformed with empty plasmids, while maintaining their original copy of *TraesCS3B02G470000*. The complete list of transgenic lines and their copy number is listed in **Table 4.1**.

After transformation, the calli were grown in sterile conditions and delivered to the University of Leeds as small plant plugs. These plants were transferred to P24 cell trays containing cereal mix and grown at 20 °C long-day (16 hr light/8 hr dark) glasshouse conditions. Leaf tissue for expression analysis was sampled after 7 days of growth in these conditions. RNA extraction, cDNA synthesis, and qPCR analysis was carried out as previously described in section **2.2.4.1**.

Once lines had flowered, spikelet number per ear and overall wheat ear appearance was phenotyped. To transfer the transgene into a winter line, transgenic plants were crossed into winter cultivars Charger (moderate vernalization requirement) and Hereward (strong vernalization requirement). The transgenic line (T0) was the pollen donor, whereas Charger and Hereward were the receptor lines. To identify successful winter crossed with transgenic lines, crosses were grown and leaf tissue was sampled. DNA was extracted from leaf tissue using the chloroform:isoamyl alcohol method adapted from (Paterson et al., 1993). To summarise, 500 µL of cotton-lysis buffer (1 M Tris-Cl pH 8, 5 M NaCl, 0.5 M EDTA, 2% w/v CTAB, 2% w/v PVP40) was added to each sample and incubated for 45 mins at 65 °C. Samples were cooled on ice for 5 mins before a volume of 375 µL of chloroform:isoamyl alcohol was added. The samples were centrifuged for 5 mins at 13,000 rpm. The 300 µL supernatant was transferred to a new 1.5 ml tube containing an equal volume of isopropanol. The tube was inverted to homogenise the contents and incubated for 5 mins at 21°C. Samples were centrifuged for 30 mins at 8,000 rpm. The supernatant was disposed and 800 µL of 70% ethanol was added. The samples were centrifuged for a further 5 mins at 13,000 rpm. The supernatant was removed, and the remaining DNA was left to air dry before being resuspended in 50 µL dH_2O.

A PCR was performed to amplify the transgene using the forward (GCCTCCTCCAGCATAGAAGGTA) and reverse primer (TGCGCCAACATCTTCTTCGATT). Primers were designed to span exon-exon boundaries and in gDNA containing the transgene, the primers amplify a 250 bp section of the transgene. Subsequently, the PCR product was run on a 1% agarose gel and visualised using the Syngene G:BOX F3 imager and GeneSys software to determine plants containing the transgene.

In addition, transgenic lines were also bagged to encourage self-pollination and to produce the first backcross (BC_1). When plants had sufficiently senesced and grains had dried, ears were harvested and seeds were threshed. BC_1 seeds from lines with high expression of the *TraesCS3B02G470000* transgene and the control line were sown and grown under 20 °C long-day (16 hr light/8 hr dark) glasshouse conditions. Half-ear emergence (Zadok's growth stage Z55) and spikelet number per ear was scored for all plants.

Table 4.1: Transgenic *TraesCS3B02G470000* wheat lines used in this study.

Plant number	Copy number	Line I.D.
CTA58.1	4+	T-1
CTA58.2	2	T-2
CTA58.3	2	
CTA58.4	4+	T-3
CTA58.6	4+	
CTA58.7	4+	T-4
CTA58.8	1	
CTA58.9	1	
CTA58.10	1	
CTA58.11	3	T-5
CTA58.12	4+	
CTA58.13	4+	
CTA58.14	1	T-6
CTA59.1	4+	
CTA59.3	1	
CTA59.4	4	
CTA59.5	4+	
CTA59.6	4+	
CTA59.7	4	
CTA59.8	2	
CTA59.9	1	

Table 4.1: Transgenic *TraesCS3B02G470000* wheat lines used in this study (continued).

Plant number	Copy number	Line I.D.
CTA59.10	4+	
CTA60.1	2	
CTA60.2	1	
CTA60.3	4+	T-7
CTA60.4	4	T-8
CTA60.5	3	
CTA 61.1	4	
CTA 61.2	4	
CTA 61.3	4+	
CTA 61.4	4	
CTA 61.5	4+	
CTA62.1	2	
CTA62.2	3	
CTA63.1	3	
CTA63.2	4+	
CTA63.3	4+	
CTA63.4	2	
CTA63.5	4+	
Control	0	

4.3 Results

4.3.1 Field flowering

In the previous chapter, identified *ODDSOC* genes had different interactions and expression activity. Therefore, we next wanted to explore how these *ODDSOC* genes behave under a realistic vernalization scenario. To do this, vernalization in winter cv. Hereward grown over winter in the field was investigated. As mentioned, Hereward has three copies of *VRN-A1* and a high vernalization requirement, therefore it was selected for analysis during growth in the field over winter as it is typically understood to need more time exposed to winter conditions for vernalization to be fully satisfied.

The results of cv. Hereward flowering revealed that vernalization was completed in the field after 7, 10, 13, and 16 weeks of growth. In Hereward, vernalization was completed by 7 weeks, with plants flowering after an average of 184 days (**Figure 4.1**). As expected, the number of days until flowering reduced steadily after 7 weeks of growth in the field. After 10 weeks in the field and removal to glasshouse conditions, flowering was observed after an average of 152 days. However, further growth in the field beyond 10 weeks resulted in gradually delayed flowering (181 days after 13 weeks; 195 days after 16 weeks). This could possibly suggest that vernalization in cv. Hereward was completely saturated by 10 weeks, and additional exposure to winter field conditions initiated de-vernalization.

4.3.2 Genes involved in vernalization are expressed throughout winter in the field

To explore the observed vernalization and potential de-vernalization, RNA-seq analysis was carried out on Hereward leaf tissue from across all timepoints to examine the temporal expression of genes involved in vernalization, including the 15 *ODDSOC* genes introduced in the previous chapter. Since vernalization was completed after 7 weeks of growth in the field, we first chose to investigate expression of the *ZCCT* genes that comprise the *VRN2* locus. This revealed that all *ZCCT* genes had low expression (i.e., <0.25 TPM) and were down regulated further over the time course, indicating that *VRN2* was already at least partially supressed by 5 weeks (**Figure 4.2**).

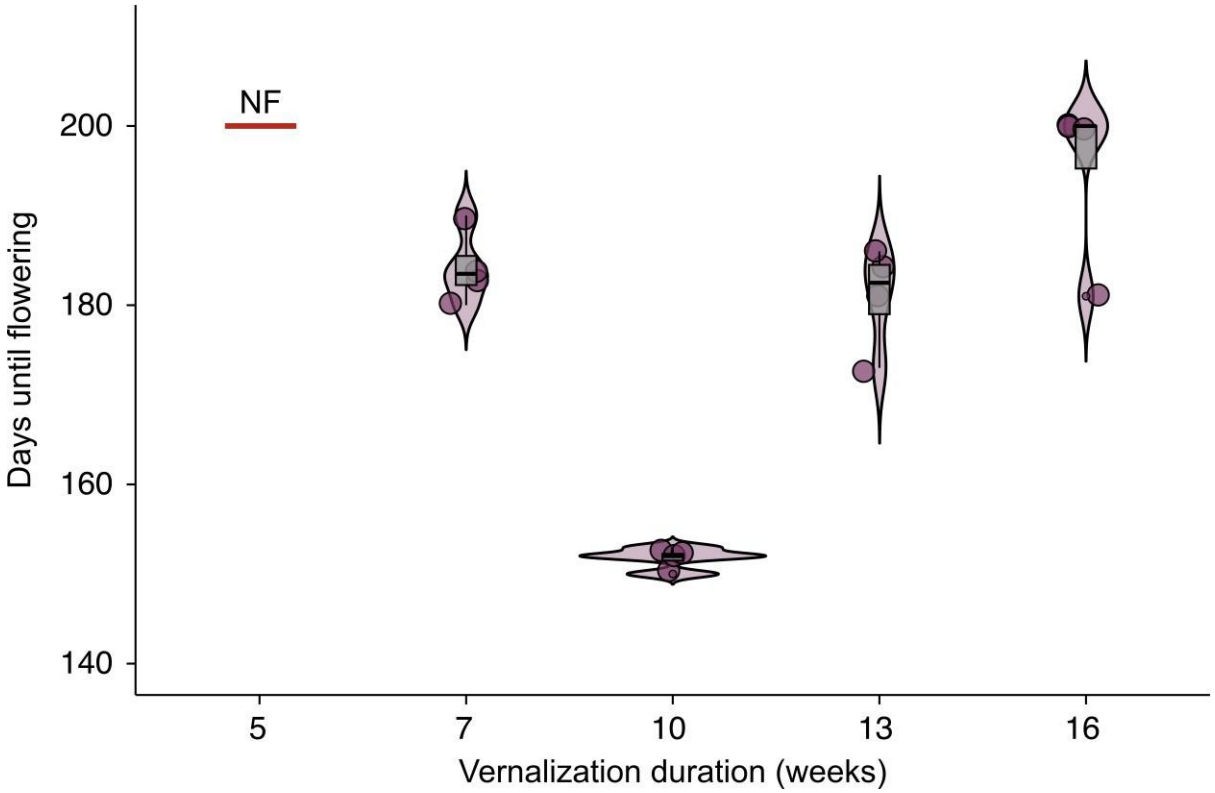


Figure 4.1: Days until flowering of winter wheat cv. Hereward grown in the field over winter. NF = no flowering after 200 days; n = 4; vernalization duration = amount of time in weeks of growth in the field over winter.

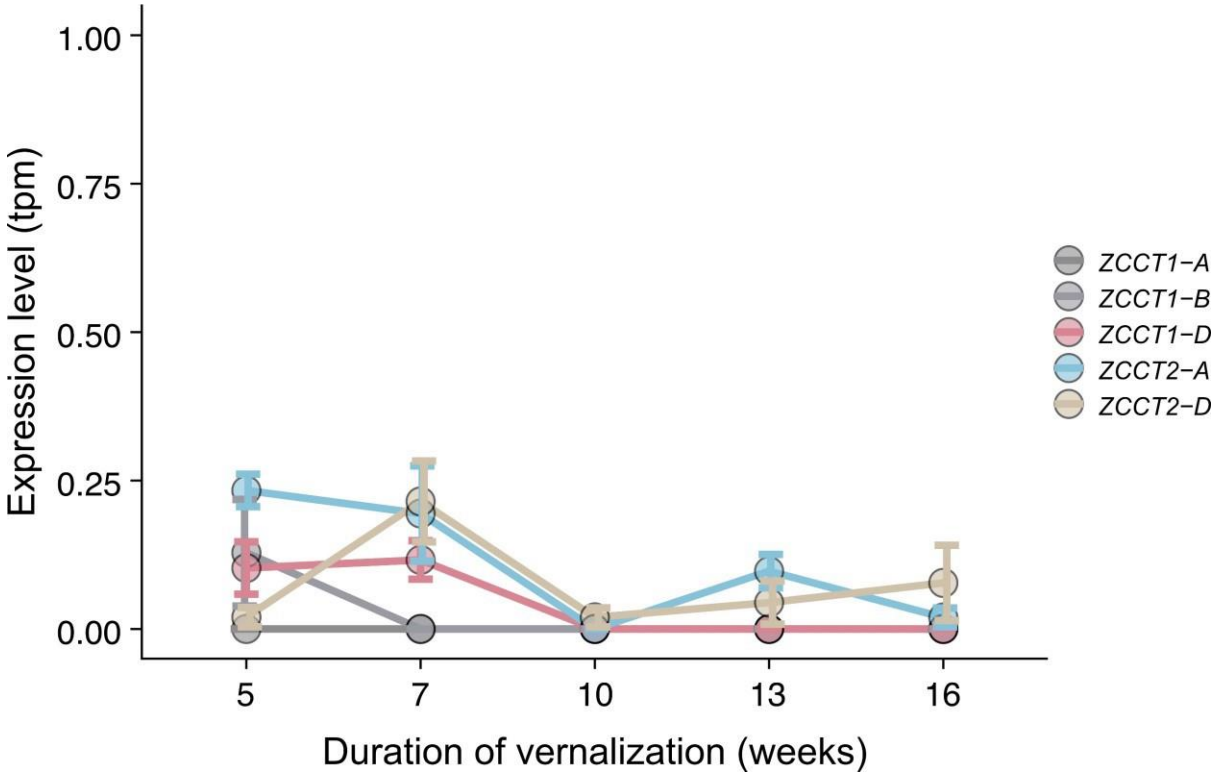


Figure 4.2: Expression of *ZCCT1* and *ZCCT2* genes in the field over winter. Error bars represent standard error of the mean for 3 biological replicates.

Interestingly, expression of *ZCCT-D2* was up regulated between 5 weeks and 7 weeks; expression level was still below 0.25 TPM and it's unclear whether this shift in expression is noteworthy. In contrast, *VRN-A1* was up regulated over the time course. Expression level dropped slightly between 7 and 10 weeks but was sharply up regulated between 10 weeks and 13 weeks and continued to increase between 13 and 16 weeks of growth in the field. Floral activator *FT-B1* was not expressed across the timecourse.

Next, expression of all 15 identified *ODDSOC* genes was investigated. Since *TraesCS3A02G432900* was determined to most likely be floral activator *OS1*, expression of this gene was examined separately to the other *OS2* genes that are predicted to function as floral repressors. Here, expression analysis revealed that *TraesCS3A02G432900* up regulated between 5, 7, and 10 weeks of growth in the field. After 10 weeks, *TraesCS3A02G432900* was down regulated, although expression remained around a level of 10 TPM from 13 weeks onwards. Notably, the expression level of *TraesCS3A02G432900* at 5 weeks was significantly higher than *VRN-A1* ($p<0.05$; Student's t-test) (**Figure 4.3**). Furthermore, expression of *TraesCS3A02G432900* remained higher than *VRN-A1* for the first 10 weeks of growth. Conversely, expression of the 14 putatively identified *OS2* genes indicated that 12 genes were down regulated over the time course. The only *OS2* genes that didn't follow this pattern were *TraesCS7A02G260900*, *TraesCS7B02G158900*, and *TraesCS7D02G261900* (**Figure 4.4**); instead, these homoeologous genes were steadily up regulated and finally stabilised between 13 weeks and 16 weeks.

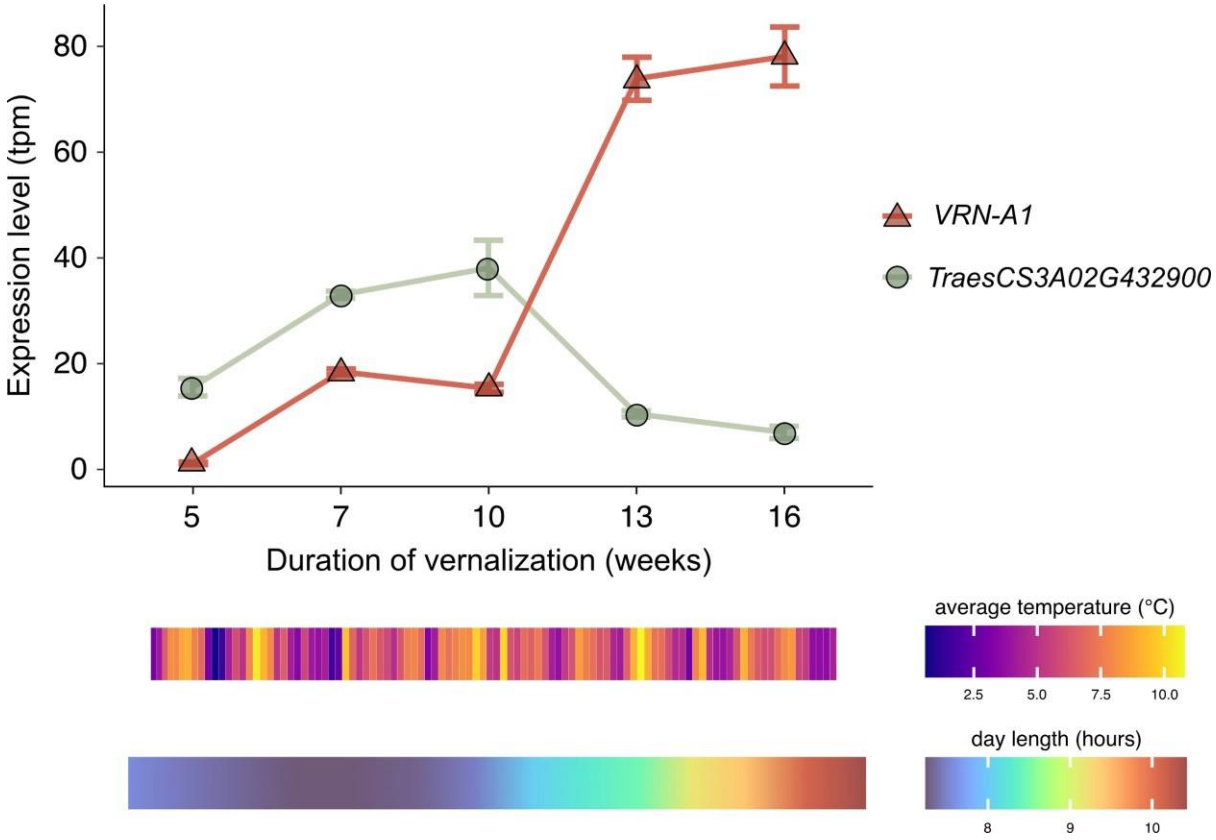


Figure 4.3: Expression of *VRN-A1* and *TraesCS3A02G432900* in the field over winter. Expression was examined in winter wheat cv. Hereward. *VRN-A1* = red triangle, *TraesCS3A02G432900* = green circle; error bars represent standard error of the mean for 3 biological replicates. Panels below the expression plot represent average daily temperature (°C) and day length (hours) over the time course.

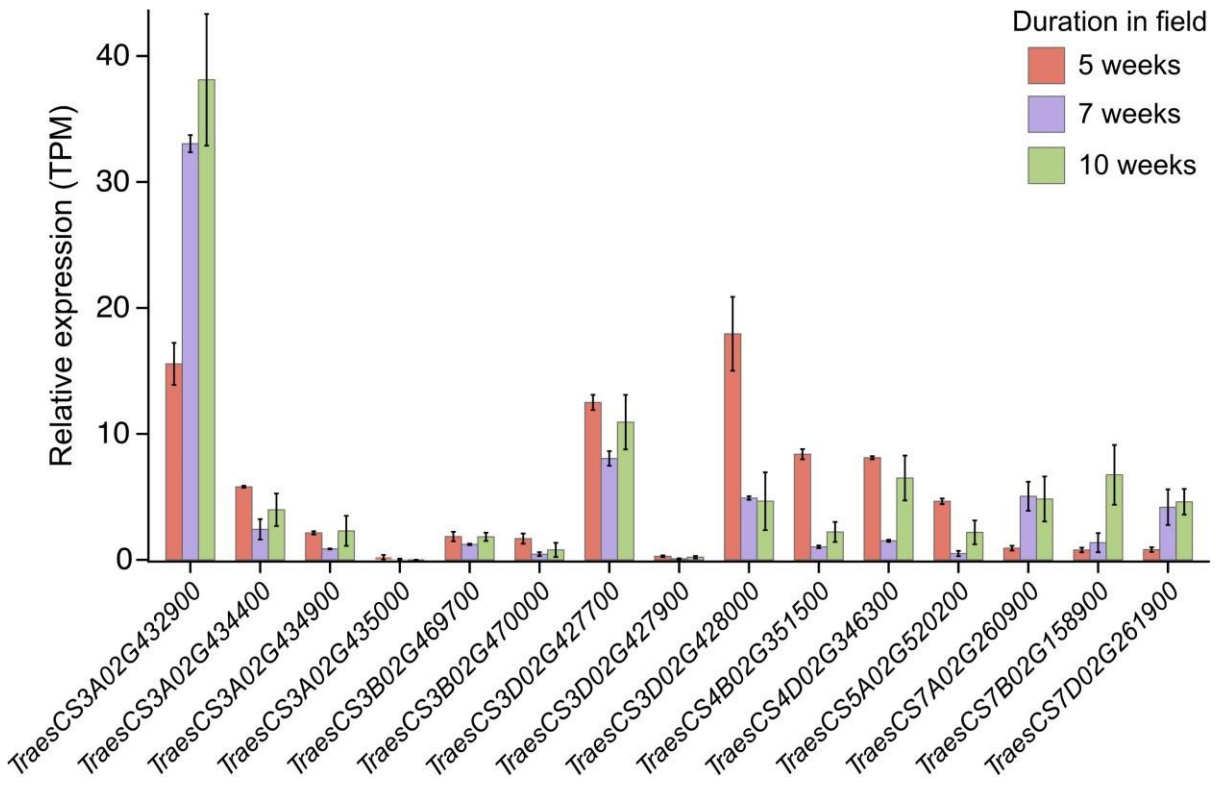
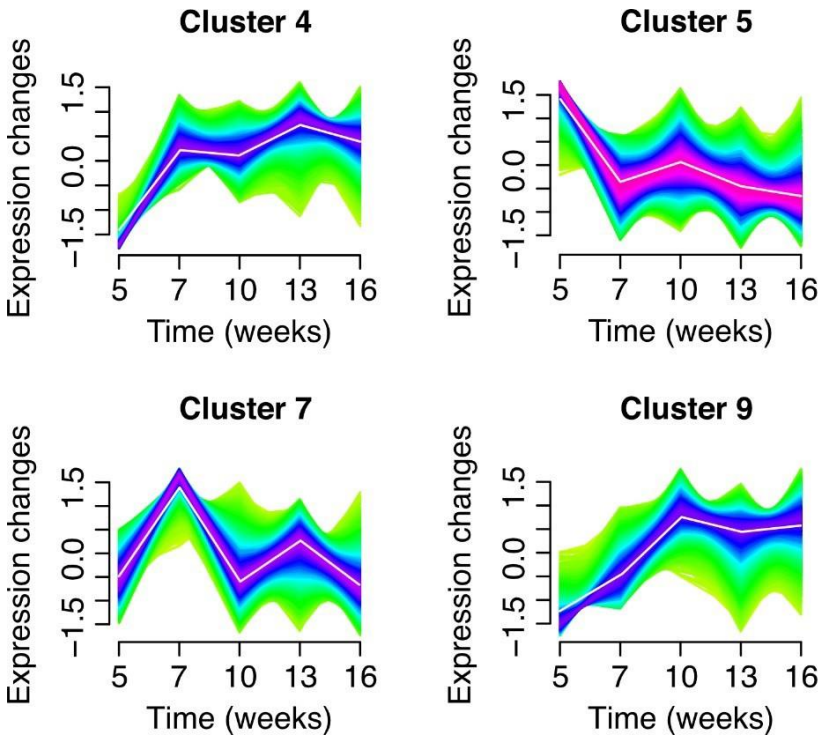


Figure 4.4: Expression of *ODDSOC* genes in the field over winter. Expression was analysed at 5 weeks (red), 7 weeks (purple), and 10 weeks (green) of growth in the field for winter wheat cv. Hereward. Error bars represent standard error of the mean for 3 biological replicates.

4.3.3 Global clustering reveals trends in dynamic *ODDSOC* expression during vernalization

To investigate this further, soft clustering analysis was performed on 7,054 genes. Based on the results of the soft clustering analysis, the only classical vernalization gene identified was *VRN-A1*; this gene was shown to belong to cluster 9 (**Figure 4.5**). Additionally, 12 of the 15 *ODDSOC* genes were differentially expressed. Furthermore, these genes could be grouped into 4 discrete soft clusters based on expression patterns across the time course. The homoeologues *TraesCS7A02G260900*, *TraesCS7B02G158900*, and *TraesCS7D02G261900* were up regulated across 16 weeks in the field (**Figure 4.5**). *TraesCS7A02G260900* and *TraesCS7D02G261900* were sharply up-regulated between 5 and 7 weeks (cluster 4), whereas *TraesCS7B02G158900* was up-regulated slightly later between 5 and 10 weeks (cluster 9). After the initial up regulation, expression remained high for all three genes. Next, *TraesCS3A02G432900* was attributed to cluster 7; the expression profile of this cluster involved sharp up regulation between 5 and 7 weeks, but unlike cluster 4, expression decreased between 7 and 10 weeks, followed by up regulation at 13 weeks and down regulation at 16 weeks.

The majority of *ODDSOC* genes were found in cluster 5 (**Figure 4.5**). This cluster contained genes that initially had high expression after 5 weeks of growth in the field, followed by a rapid down regulation by 7 weeks (**Figure 4.5**). After a slight increase in expression between 7 and 10 weeks, these genes were steadily down regulated from 10 weeks onwards. Considering that *OS2* genes have been characterised as floral repressors in barley (Greenup et al., 2010), down regulation of this set of genes throughout vernalization-inducive conditions is unsurprising. The characterisation of *TraesCS3A02G432900* as a putative floral activator corresponds with up regulation after 7 weeks of vernalization-inducive conditions. Interestingly, the expression pattern of *TraesCS3A02G432900* doesn't match that of a traditional floral activator, such as *VRN-A1* in cluster 9, which is upregulated over time. Instead, *TraesCS3A02G432900* sharply upregulated between 5, 7, and 10 weeks, preceded by a swift decrease between 10, 13, and 16 weeks (**Figure 4.3**; **Figure 4.5**). In contrast, *VRN-A1* was continually upregulated across the time course (**Figure 4.3**).



Cluster 4
TraesCS7A02G260900
TraesCS7D02G261900

Cluster 5
TraesCS3A02G434400
TraesCS3A02G434900
TraesCS3B02G469700
TraesCS3D02G427700
TraesCS3D02G428000
TraesCS4B02G351500
TraesCS4D02G346300
TraesCS5A02G520200

Cluster 7
TraesCS3A02G432900

Cluster 9
TraesCS7B02G158900

Figure 4.5: Soft clusters containing *ODDSOC* genes that are temporally expressed in the field. Time represents the duration in weeks of growth in the field over winter. Expression changes are derived from differentially expression between timepoints; white lines represent the best fit of clustering expression progression.

TraesCS7A02G260900 and *TraesCS7D02G261900* don't follow this trend; these genes belong to a cluster that is sharply up regulated between 5 and 7 weeks in the field and remain at a relatively high level of expression until 16 weeks of growth (**Figure 4.5**). This indicates that these genes are up regulated earlier than their homoeologue, *TraesCS7B02G158900*, which belongs to cluster 9 and exhibits slightly delayed up regulation compared to *TraesCS7A02G260900* and *TraesCS7D02G261900*. The *OS2* genes *TraesCS3D02G427900*, *TraesCS3A02G435000*, and *TraesCS3B02G470000* weren't assigned to a cluster. This is likely because these genes had low expression (i.e., <1 TPM) and differences between timepoint weren't substantial enough to classify as these genes being differentially expressed.

Since vernalization in Hereward was saturated by 10 weeks, expression of *ODDSOC* genes was compared for 5, 7, and 10 weeks of growth (**Figure 4.4**). After 5 weeks of growth in the field, *TraesCS3A02G432900*, *TraesCS3D02G427700*, and *TraesCS3D02G428000* had the highest level of expression. By 7 weeks, *TraesCS3A02G432900* expression had increased, whereas all *OS2* genes (excluding *TraesCS7A02G260900*, *TraesCS7B02G158900*, and *TraesCS7D02G261900*) were down regulated. *TraesCS3D02G428000* experienced the most significant difference in decrease of expression level ($p<0.05$; Student's t-test) (**Figure 4.4**). Intriguingly, the homoeologues of *TraesCS3D02G428000*, *TraesCS3A02G435000* and *TraesCS3B02G470000*, had a much lower expression level in comparison ($p<0.05$; Student's t-test). This potentially suggests that under these winter conditions, *TraesCS3D02G428000* is the most prominent floral repressor. To explore this further, WatSeq haplotype maps were examined to identify how well-conserved homoeologues *TraesCS3A02G435000*, *TraesCS3B02G470000*, and *TraesCS3D02G428000* were within a diverse range of wheat lines.

4.3.4 Deletions within the first intron of *OS2* gene *TraesCS3B02G470000*

Analysis of the WatSeq haplotype maps for *TraesCS3A02G435000*, *TraesCS3B02G470000*, and *TraesCS3D02G428000* revealed that within a diverse wheat accession, *TraesCS3A02G435000* and *TraesCS3D02G428000* were well conserved with no observable major changes in gene sequence that was significant enough to be grouped as a cluster. In contrast, *TraesCS3B02G470000* had two deletions that were well conserved within the first intron. The larger deletion was approximately 5.6 kb (position chr3B: 717300001-717305617) and the smaller deletion 0.4 kb (position chr3B: 717291188- 717291579). Intriguingly, both deletions appeared together, potentially suggesting a linked deletion event. These deletions were present in 381 spring, winter, and facultative habit wheat lines, making up 18.1% of all Watkins lines examined in this study. In total, 238 spring lines (16.5%), 117 (19.5%) winter lines, and 26 (52%) unknown lines contained both intronic deletions (**Figure 4.6**).

4.3.5 Motifs within first intron deletions suggest potential regulatory role of this region

Since both deletions were in the first intron of *TraesCS3B02G470000* (i.e., outside the gene coding region), this indicated that these deletions would not be affecting protein structure. Instead, I suspected that these deleted regions could contain important regulatory features. As mentioned in the introduction to this chapter (sections **4.1.2** and **4.1.3**), the major intron 1 deletions in *VRN-1* and *FLC* contain key regulatory domains. With this in mind, I decided to investigate these *TraesCS3B02G470000* intronic deletions further as they had been conserved in 381 lines.

To examine the significance of these conserved intronic deletions, motif analysis was carried out on both deleted regions. This identified multiple predicted binding sites, some of which were relevant to regulation of this gene, flowering, and vernalization. In particular, motifs associated with binding of members of the MIKC-type MADS-Box family were identified. Within the smaller deletion (0.4 kb), 3 discrete MIKC-type MADS-Box binding motifs associated with Arabidopsis *FLC* were identified (**Table 4.2**). These motifs were the only MIKC-type MADS-Box binding sites identified in this

deletion. In the larger deletion, 12 individual MIKC-type MADS-Box were identified, 3 of which were associated with *FLC* but were different to the binding sites identified in the smaller deletion (**Table 4.2**). The remaining MIKC-type MADS-Box bindings sites were associated with Arabidopsis *SOC1* and *SOC1*-affiliated genes, including *AGL* genes *AGL8* (Lee & Lee, 2010), *AGL24* (Liu et al., 2008), *AGL42* and *AGL72* (Dorca-Fornell et al., 2011), as well as the *SEPALLATA* (*SEP*) gene *SEP3 (Lee & Lee, 2010)*. Position of these predicted binding sites within the intronic deletions are represented in **Figure 4.7**.

Of the predicted MIKC-type MADS-Box binding sites, the motifs ACAAAATTGGC and GCAAAAAAGGC were selected for further investigation. These motifs are variations of the same binding site (**Figure 4.8A**). Alignment of the first intron for the *ODDSOC* genes indicated similar motifs aligning with ACAAAATTGGC; the highest level of similarity occurred between the motif in *TraesCS3B02G470000* and *TraesCS3D02G427900*, followed by *TraesCS3A02G435000* and *TraesCS3D02G428000*, and finally *TraesCS3A02G434400*, *TraesCS3B02G469700*, and *TraesCS3D02G427700* (**Figure 4.8B**). Furthermore, a variation of this binding site is present in the first intron of *VRN-A1* (position chr 5A:587420437-587420456), with the *VRN-A1* intron motif (GCCAAAAAGGG) sharing greatest similarity with motif GCAAAAAAGGC (**Figure 4.8C**). These motifs don't follow the standard CArG consensus sequence of $CC(A/T)_6GG$ but may still function as CArG box binding sites.

In addition to scanning the intronic deletions for predicted binding sites of interest, the appearance of specific binding sites motifs was also investigated. MIKC-type MADS-Box genes such as *VRN-A1* preferentially binds CArG sites (Deng et al., 2015; Dubcovsky et al., 2006; Yan et al., 2003). For this reason, enrichment of CArG box sites within the deletions could suggest binding of MIKC-type MADS-Box elements to this intronic region. Standard CArG box sites have the consensus sequence $CC(A/T)_6GG$, however based on the sequence variation for Arabidopsis *FLC*, *SVP*, *SOC1*, *SEP3*, and *AP1* CArG binding sites, the deletions were scanned for motifs that matched the CArG box consensus MYWWWWDWRG (Aerts et al., 2018).

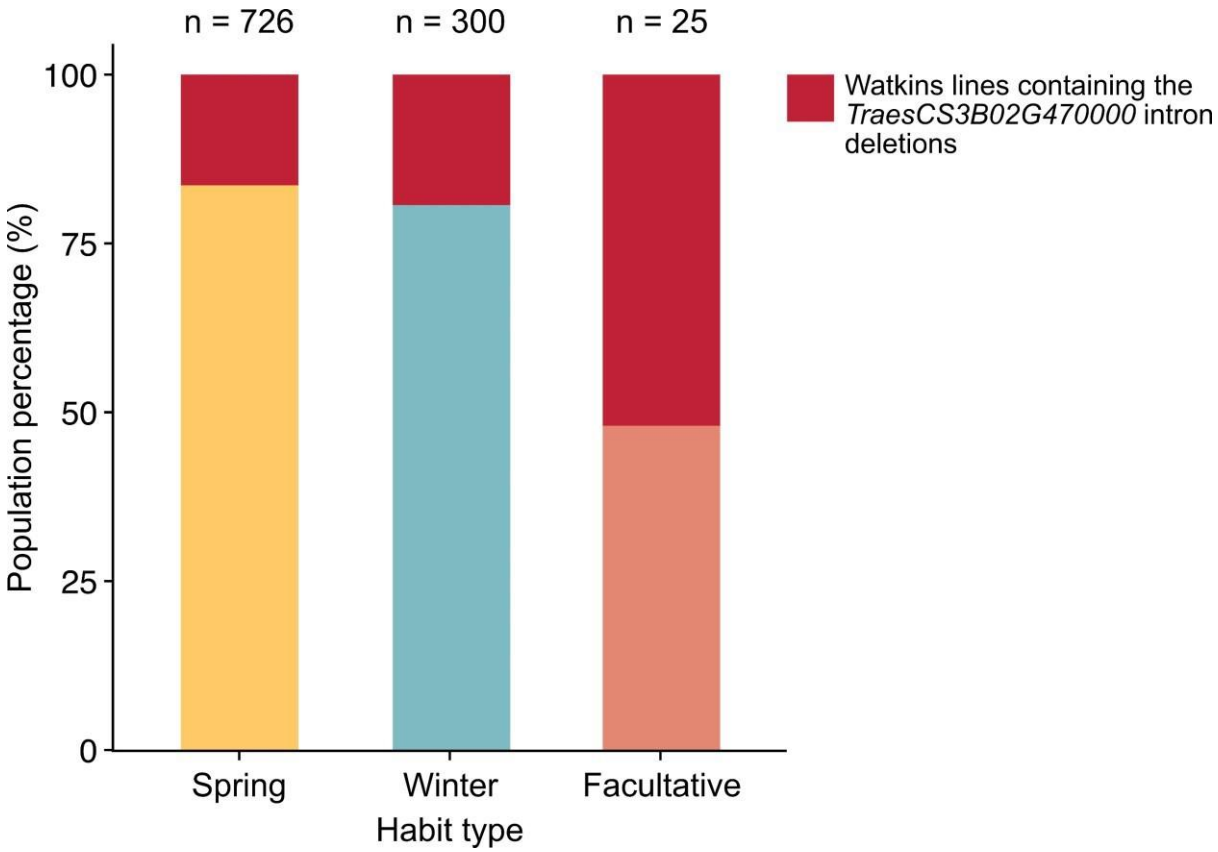


Figure 4.6: Watkins bread wheat lines containing the wild-type and the deletion version of the first intron of *TraesCS3B02G470000*. Growth habits are spring (yellow, n = 726), winter (blue, n = 300), and facultative (pink, n = 25); red regions of the bar represent percentage of total lines that contain both intronic deletions.

Table 4.2: Predicted MIKC-type MADS-Box binding sites within *TraesCS3B02G470000* intronic deletions.

Gene	Deletion	Motif
SOC1	big	CCCTCTTTTGT
SOC1	big	ACAAAATTGGC
SOC1	big	GCAAAAAAGGC
SOC1	big	AGCTTTTTTGG
SOC1	big	CGACCCATAGGTAGAATT
FLC	big	TCGACAACGATGGAAAGC
FLC	big	CATATAAACATGGAAATA
FLC	big	TATAAACATGGAAAT
FLC	small	GATTTCCTTTTGGACATC
FLC	small	CCTTTCCCTTCGCGTACA
FLC	small	ATTTCCTTTTGGACA
MADS (other)	big	CCTTTGAATTGG
MADS (other)	big	TTAGCTTTTTTGGCA
MADS (other)	big	CCCCAAATATCGAAA
MADS (other)	big	TTTTAGCTTTTTTGGCAC

Analysis uncovered that these deleted regions contained multiple CArG box sites matching the MYWWWWDWRG sequence. Within the smaller deletion, there were 38 CArG box sites with a statistical threshold of $p < 0.05$ (**Figure 4.9A**). Within the larger deletion, 481 matching CArG box sites with a threshold of $p < 0.05$ were identified; for conciseness, the 111 sites with a p-value < 0.01 are represented in **Figure 4.9B.**. Identified CArG box sites were distributed evenly across both deletions.

As well as CArG box sites, the deletions were analysed for the presence of the EVENING ELEMENT (EE) MYB motif AAAATATCT (Michael & McClung, 2002). The EE motif is involved in circadian-regulated expression and induces peak expression later in the day (Harmer et al., 2000; Harmer & Kay, 2005). In Arabidopsis, the EE has been reported to play a role in cold-induced expression of key clock genes, such as *CONSTANS-like 1* (*COL1*) (Bieniawska et al., 2008; Mikkelsen & Thomashow, 2009; Seo et al., 2012). Based on the previous expression analysis reported in Chapter 3, *TraesCS3B02G470000* expression followed a circadian rhythm over a 24-hr period which was temperature-responsive (Section **3.3.9**). For this reason, the MYB binding motif was examined in more detail. Sequence analysis identified 81 MYB-associated motifs with a p-value < 0.01 within the larger deletion; additionally, 7 motifs with a p-value < 0.01 were identified within the smaller deletion. These motifs were matched to NAWAWATCH to account for base pair differences indicated by the Tomtom motif comparison tool; there were no identical matches to the AAAATATCT, however one site within the larger deletion matched the AAATATC core motif. The identification of numerous predicted MYB, CArG, and other MIKC-type MADS-Box binding sites within both the small and large deletion could indicate a regulatory role for this region of the intron of *TraesCS3B02G470000*.

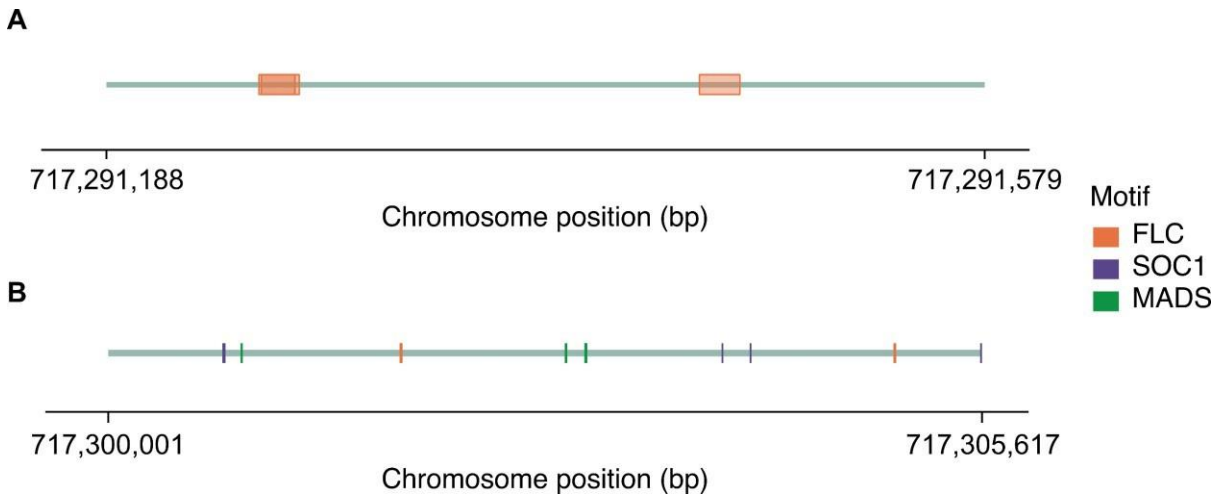


Figure 4.7: Location of relevant predicted MIKC-type MADS-Box binding sites within *TraesCS3B02G470000* intronic deletions. (**A**) Small deletion; (**B**) large deletion. Motifs selected were associated with *FLC* (orange), *SOC1* (purple), and MADS- related genes (green) in Arabidopsis.

Finally, SPL motifs were identified in the intronic deletions. Interestingly, this included the motif TGTTCGAACAC; variations of this motif, TGTAC and ATGTACTTTGT (**Figure 4.10A**), were identified in the *TraesCS3A02G432900* intronic region associated with the lncRNA antisense transcript (section **3.3.10**). This motif was highly conserved in the first intron of 656 Watkins lines, suggesting that this predicted binding site could have a regulatory role. As mentioned in section **3.3.10**, *SPL* genes in Arabidopsis are involved in the vegetative to reproductive transition. For example, the floral promoter SPL15 can form a complex with SOC1 that activates expression of target genes involved in floral induction (Hyun et al., 2016). Furthermore, both *SPL15* and *SOC1* are targeted by *FLC* in Arabidopsis and the *FLC* orthologue *PERPETUAL FLOWERING 1* (*PEP1*) in *Arabis alpina*, repressing their expression and subsequently floral induction (Hyun et al., 2019; Mateos et al., 2017). Since this demonstrates a link between *FLC*, *SOC1*, and *SPL* genes, this indicates that this could be a relevant binding site. Therefore, interaction between this motif and ODDSOC proteins was examined further using a yeast one-hybrid assay.

4.3.6 The TraesCS3B02G470000 protein binds to a *SPL*-associated motif within the intronic deletion

The results of the yeast one-hybrid assay revealed that TraesCS3B02G470000 interacted with the SPL-associated TFBS (**Figure 4.10B**). Furthermore, this was indicated to be a strong interaction (**Figure 4.10B**). The lack of interaction between this binding site and the empty pGAD424 plasmid validates the interaction between TraesCS3B02G470000 and the binding site. This suggests that the first intron of *TraesCS3B02G470000* contains well-conserved binding sites that are involved in regulation of the gene. Moreover, the TraesCS3B02G470000 protein can potentially bind to this site within its intron, indicating an auto-regulatory function. Whether binding at this site promotes up regulation, down regulation, or a dual-regulatory feature is still unclear. In addition to TraesCS3B02G470000, interaction between this predicted binding site and the VRN-A1 exon 4 SNP variant of the protein that is associated with a higher vernalization requirement. This uncovered that this version of the VRN-A1 protein did not interact with the binding site (**Figure 4.10C**). This corresponded with the lack of interaction between the TraesCS3B02G470000 protein and this variant of the VRN- A1 protein reported in section **3.3.7**.

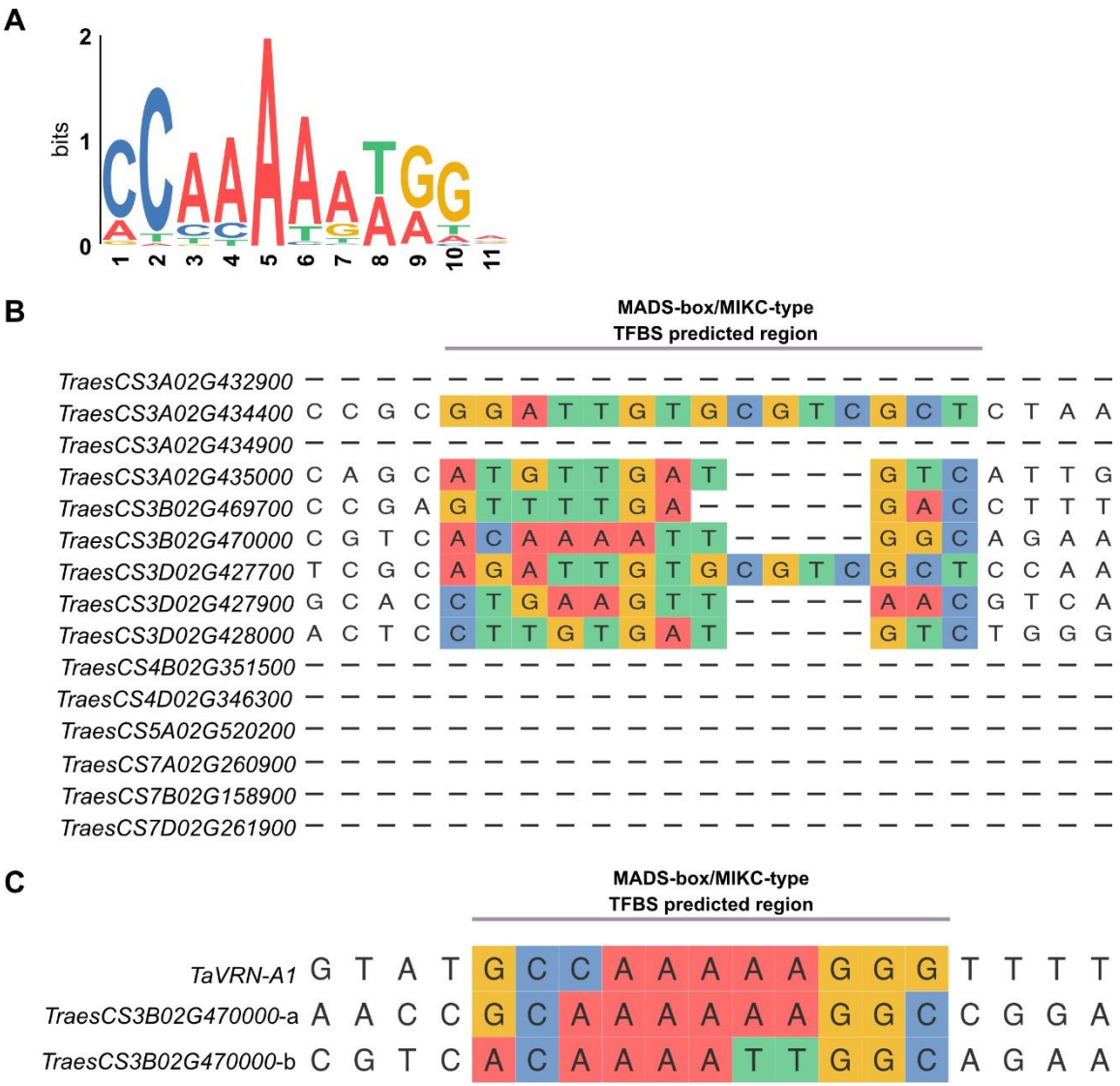


Figure 4.8: Schematic of the MADS-box predicted binding domain within the intronic deletion. (**A**) Binding motif; (**B**) Alignment of motif within *ODDSOC* introns; (**C**) Alignment of motif within *VRN-A1* and *TraesCS3B02G470000.*

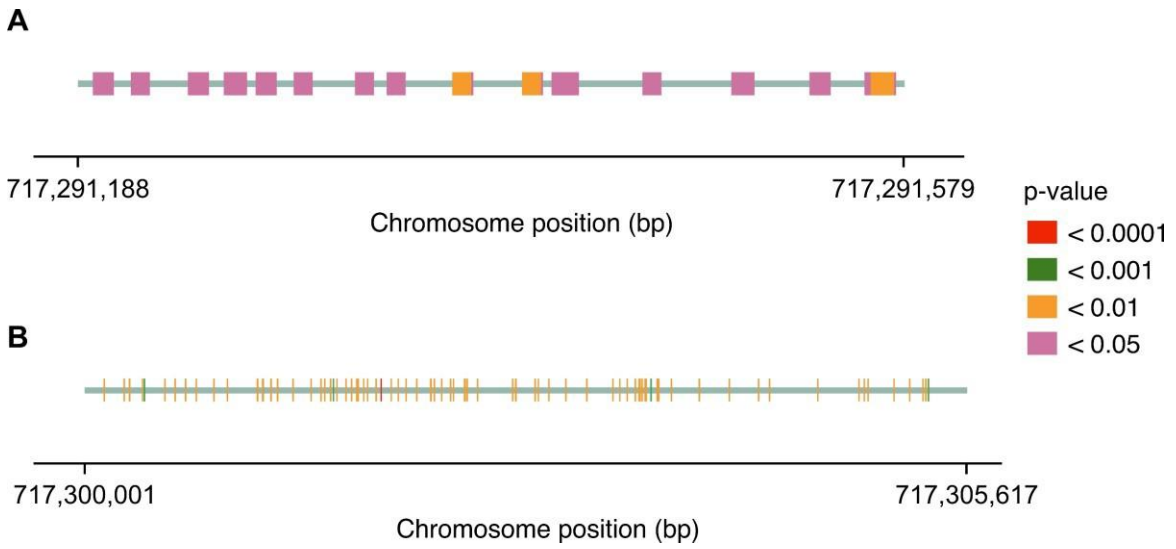


Figure 4.9: Location of predicted CArG box motifs within *TraesCS3B02G470000*. (**A**) predicted CArG box binding sites within the small deletion, and (**B**) within the large deletion. P-values are indicated by the panel on the right.

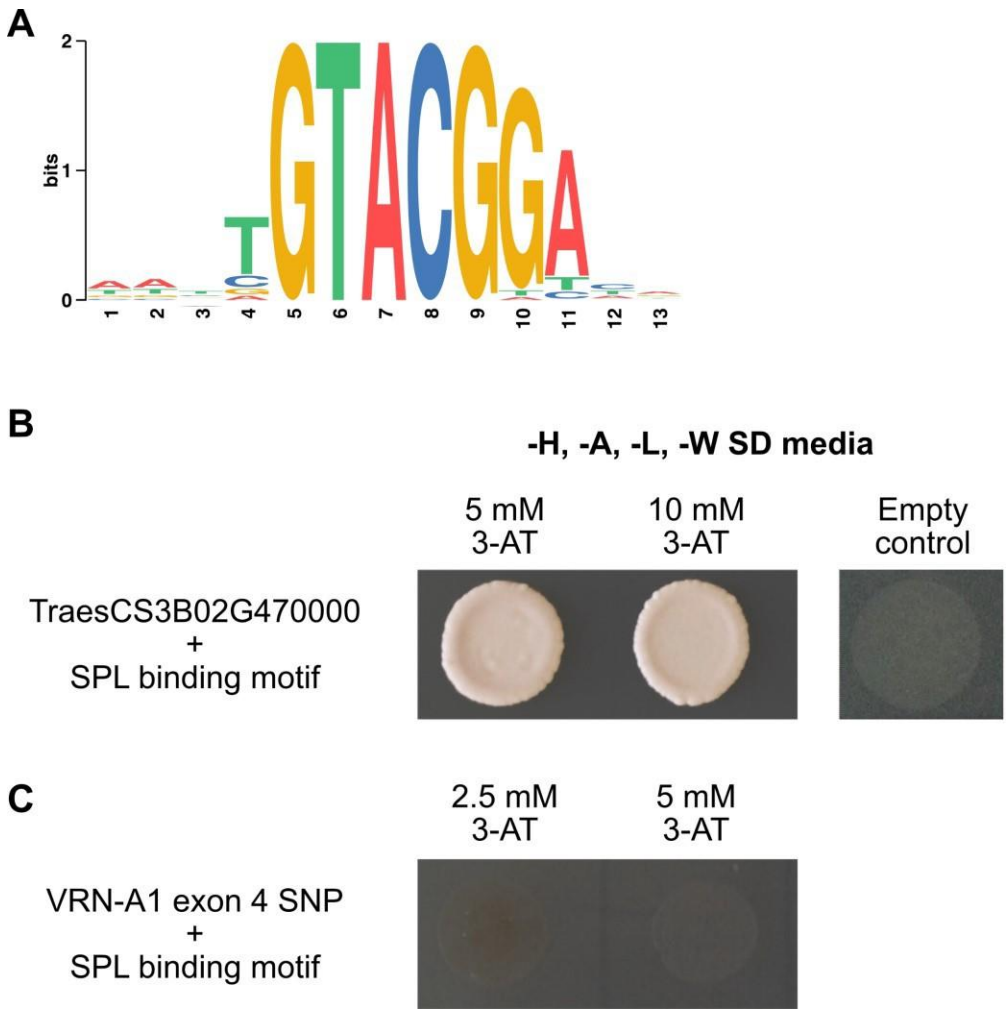


Figure 4.10: SPL motif common to the first intron of *TraesCS3A02G432900* and *TraesCS3B02G470000*. (**A**) Amino acid sequence logo for the NNNNGTACGDHHD motif based on motif occurrences in the ArabidopsisDAPv1 database. (**B**) Yeast one-hybrid assay showing interaction between TraesCS3B02G470000 protein and the *SPL*-associated DNA-binding motif. Interactions tested on moderate (5 mM) and high (10 mM) 3-AT supplemented media. (**C**) Yeast one-hybrid assay showing interaction between VRN-A1 exon 4 SNP protein and the *SPL*-associated DNA-binding motif. Interactions tested on low (2.5 mM) and moderate (5 mM) 3-AT supplemented media. The empty control signifies the pHIS plasmid transformed with the motif and an empty pGBT9 plasmid.

4.3.7 Geographical distribution and historical heading data suggest role of TraesCS3B02G470000 intronic deletion

Since a key difference between spring and winter wheat is their responsiveness to temperature, it was informative to investigate whether the intronic deletions in *TraesCS3B02G470000* conveyed adaptability to a specific climate. Using the Wat- seq haplotype map, lines containing the intronic deletions had been assigned to Cluster 56 out of 67 clusters total; therefore, the average daily temperature for the site in which each line originated was plotted. This analysis was then divided into seasons to identify whether the intronic deletions were associated with a particular time of year. The results revealed that lines containing the deletions (represented by Cluster 56) were endemic to a wide range of climatic regions (**Figure 4.11**).

To determine whether this was significantly different to lines without the deletions, the temperature range of Cluster 56 was compared to the other clusters for each season. Since certain clusters only contained several lines and would be unsuitable for statistical analysis, comparisons were made between clusters with sample size similar to Cluster 56. This revealed that there were no statistically significant difference in winter temperature of Cluster 56 compared to the other clusters. Greatest difference in temperature was observed for Cluster 56 versus Cluster 01, with statistically significant differences in average daily spring, summer, and autumn temperatures. Cluster 56 was associated with lower spring, summer, and winter average temperatures. As well as this, Cluster 56 versus Cluster 09 had statistically significant differences in average daily temperatures for summer and autumn, where Cluster 56 also had lower temperature compared to Cluster 09. To further explore whether presence of these intronic deletions was linked to temperature response, the geographical location of lines containing the intron deletion were mapped to see if a link existed between the deletions and tolerance or sensitivity to local temperature.

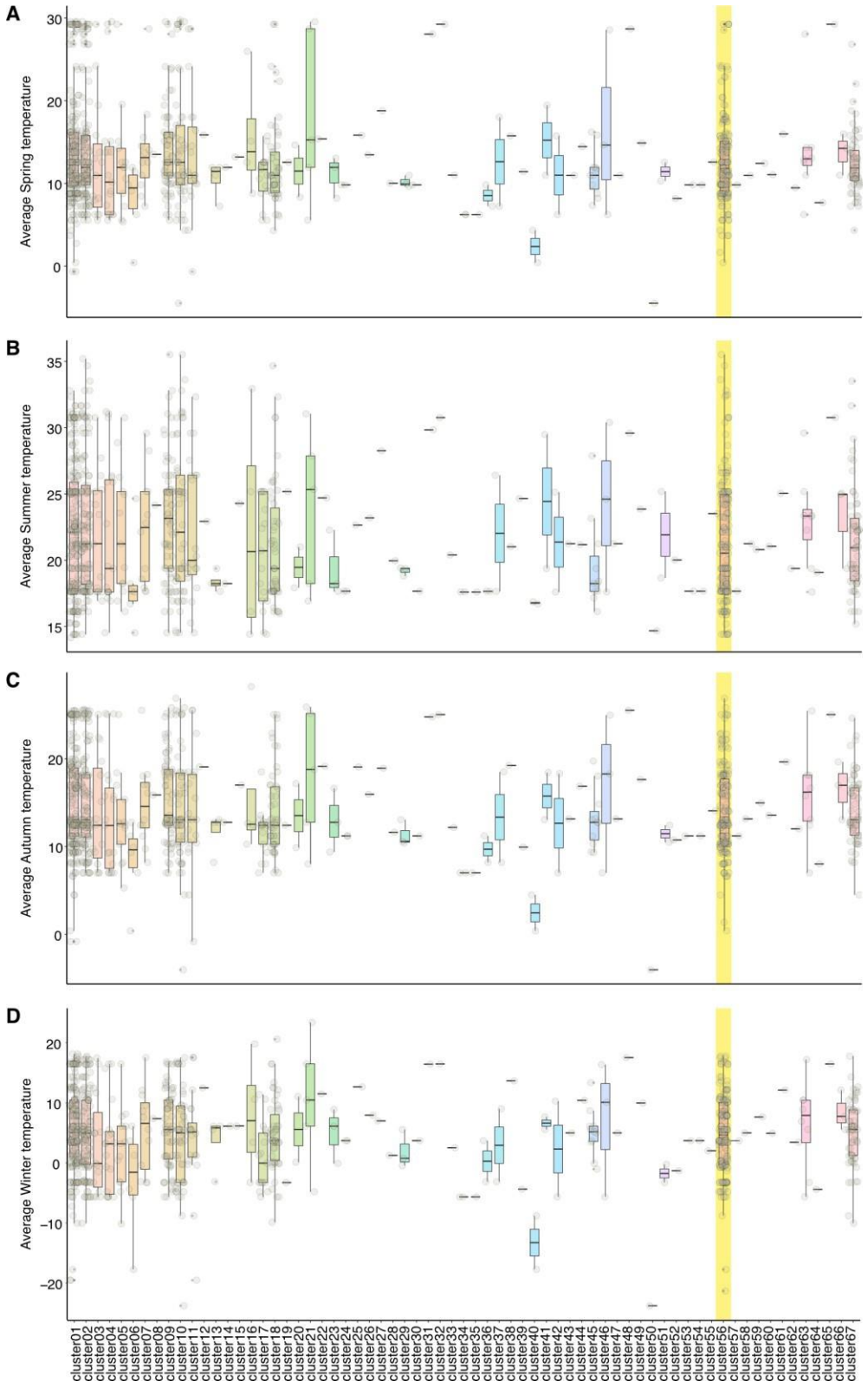


Figure 4.11: Average seasonal temperature for geographical locations Watkins collection clusters. Seasons are divided into (**A**) spring, (**B**) summer, (**C**) autumn, and (**D**) winter. Cluster 56 (highlighted yellow) contains the lines with the *TraesCS3B02G470000* intronic deletions.

However, the geographical origin of these lines didn't indicate an obvious connection (**Figure 4.12**). This analysis did identify a number of lines that stood out due to their non-intuitive habit and location. For example, winter line WATDE0981 from Algeria, and winter line WATDE0848 from Tunisia. Based on Berkely Earth data, the average autumn and winter temperatures for the Algeria site is 17.9 °C and 9.6 °C. For the site in Tunisia that WATDE0848 was sourced, the average autumn and winter temperatures are 20.9 °C and 12.6 °C. Conversely, the spring wheat line WATDE0111 was native to Finland. The average spring temperature for this site is 2.7 °C; furthermore, average temperature for this site in June and July was 16-17 °C – considerably lower than the temperatures associated with flowering. This might suggest that these intronic deletions interfere with temperature responsiveness in these lines, potentially influencing the flowering response.

To investigate whether these intronic deletions within *TraesCS3B02G470000* are linked with a flowering phenotype, flowering time of lines containing both WT intron (i.e., any wheat cultivar with an intact first intron for *TraesCS3B02G470000*) and the deletions version of the intron were used. Using historical phenotypic data from field trials undertaken and provided by the John Innes Centre, flowering of lines during the years 2006, 2010, and 2014 was examined. The 2006 field trial was a spring drilling for both spring and winter lines, whereas the 2010 and 2014 field trials were winter drillings. Therefore, spring and winter lines drilled in spring developed throughout spring/summer and flowered that summer, whereas winter drilled lines grew slowly over winter and flowered the following summer. This likely explains the statistically significant difference in flowering time of winter versus spring lines with these intronic deletions, with spring lines flowering earlier (**Figure 4.13A**).

For spring lines, the difference between WT and deletion intron across all 3 years was significant ($p = 0.005$, 0.0001, and 4.73E-07) (**Figure 4.13C**). This indicated that spring lines with the intron deletions flowered later than those with the WT version of the intron. Conversely, there was no significant difference between the flowering time of winter lines containing the intron deletions and those without (**Figure 4.13B**). This was consistent across all 3 years. This potentially indicates that the region of the first intron of *TraesCS3B02G470000* containing these deletions plays a more important role in gene regulation in spring wheat compared to winter.

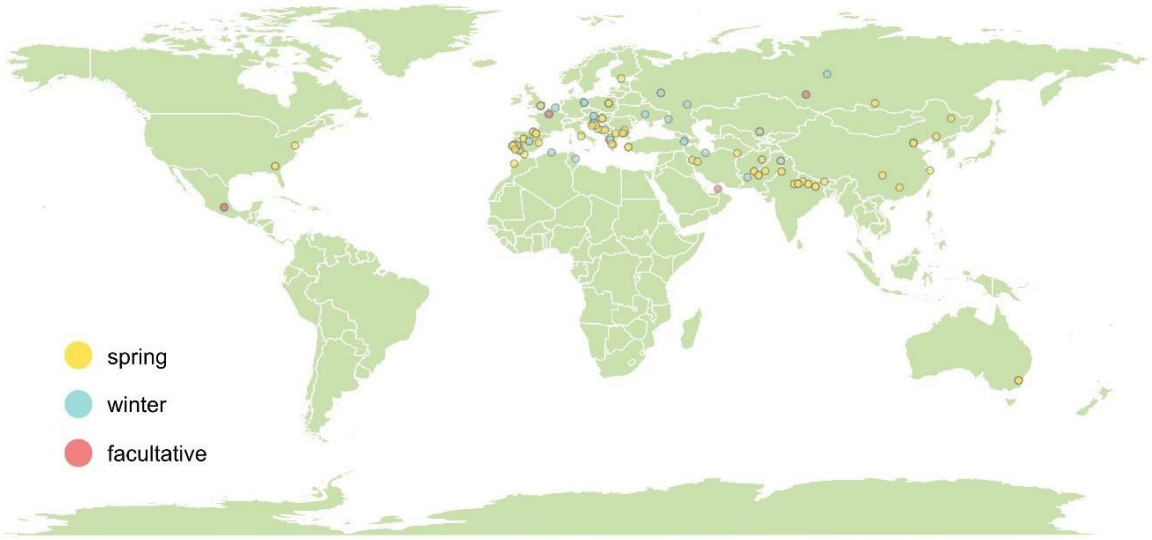


Figure 4.12: Global distribution of wheat lines containing deletions within the first intron of *TraesCS3B02G470000*. Circles represent the source location for each line based on latitude and longitude. Spring habit = yellow; winter habit = blue; facultative habit = pink.

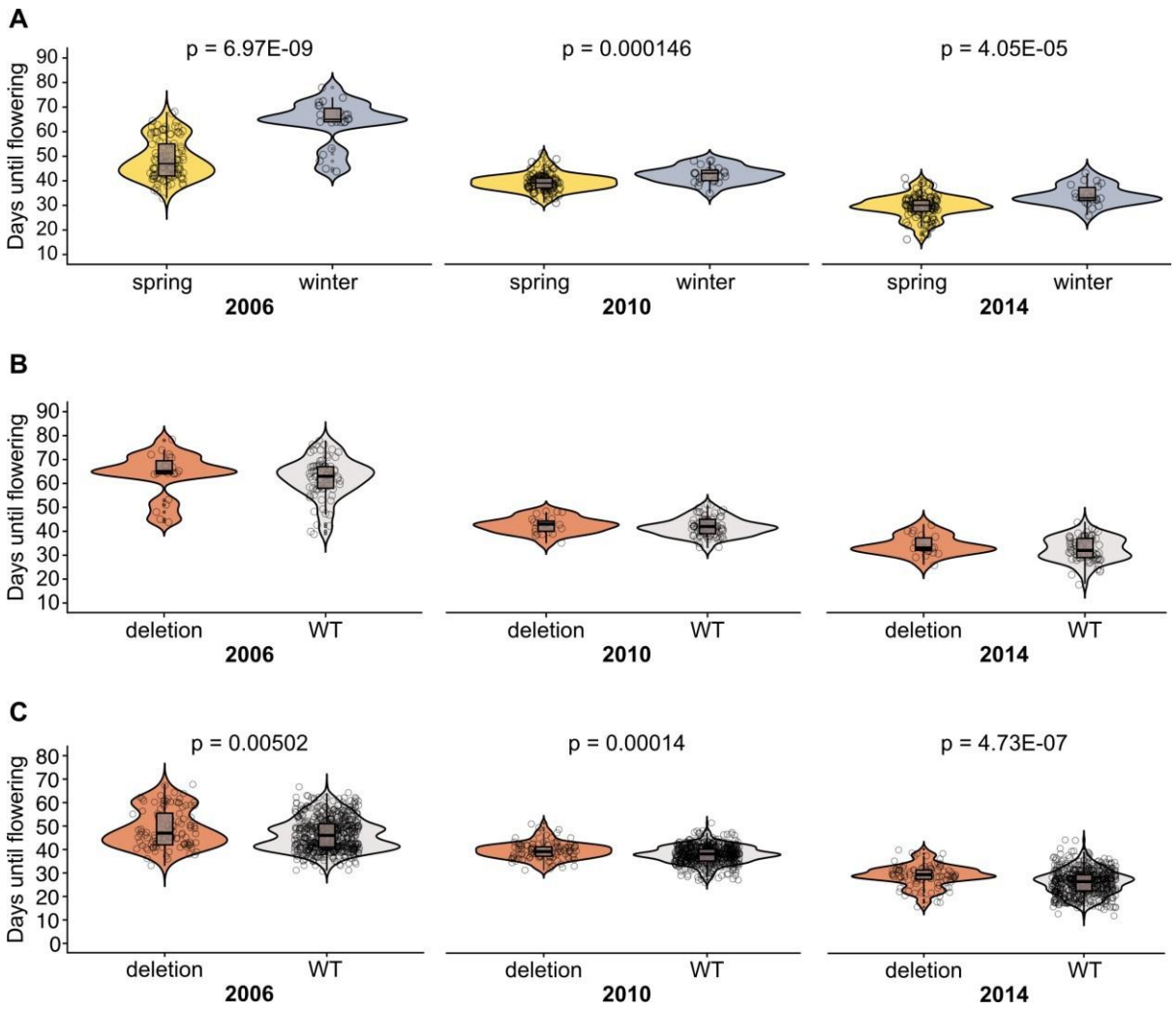


Figure 4.13: Flowering of Watkins lines across years 2006, 2010, and 2014. (**A**) Days until flowering for spring (yellow; n = 119) and winter (blue; n = 58) cultivars with deletions in the first intron of *TraesCS3B02G470000*. (**B**) Days until flowering for winter cultivars with the deletion (n = 58) and without the deletion (WT; n = 242) in the first intron of *TraesCS3B02G470000*. (**C**) Days until flowering for spring cultivars with the deletion (n = 119) and without the deletion (WT; n = 607) in the first intron of *TraesCS3B02G470000*. Plants were grown at the John Innes Centre field site; p- values from Mann-Whitney U test.

To further explore whether these intronic deletions were more important in spring wheat, *TraesCS3B02G470000* expression was analysed. No apparent trend in level of expression was observed, with lines containing the deletions showing a uniform distribution of expression (**Figure 4.14**). Approximately 33% of deletion lines and 50% of WT lines had expression higher than the median level of expression (3.86 fpkm). The highest level of expression (16.34 fpkm) was recorded in a WT line (**Figure 4.14**). The gene wasn't expressed in 4 lines, two of which were WT and two which contained the intron deletions. When comparing the expression levels of winter and spring lines containing the intron deletion from the YoGI expression dataset, a statistically significant ($p < 0.01$) difference in expression was revealed (**Figure 4.15**). This indicated a higher level of expression in winter lines compared to spring lines with the intron deletions; the highest level of expression was 28 fpkm (winter) and the lowest was 0 fpkm (spring) (**Figure 4.15**). To explore this further, expression level across all winter lines in the YoGI panel was investigated. This revealed an overall statistically significant ($p < 0.01$) trend towards higher level of expression of *TraesCS3B02G470000* in winter lines containing both deletions within the first intron (**Figure 4.16**). 28 of the 32 lines containing the deletions were above the median expression of 7.6 fpkm (**Figure 4.16**). Notably, previously mentioned Tunisian winter line WATDE0848 had the highest level of expression of all winter lines (28 fpkm) (**Figure 4.16**).

As winter lines containing the intron deletions exhibited a higher level of *TraesCS3B02G470000* gene expression, this was pursued further through conducting a detailed developmental study involving expression analysis and phenotypic analysis under conditions which replicated winter. As a control, a line containing the wild type version of the intron (i.e., no intronic deletions) was also grown under these conditions. For spring lines, no significant difference was observed in *TraesCS3B02G470000* expression for lines within and without the intron deletions at the 2- and 4-leaf stage of development. However, it was observed that spring wheat lines containing the intronic deletions exhibited slower apical development compared to lines with the wild type version of the intron (**Figure 4.17**). Moreover, there was a tenfold higher level of expression of the gene under 10 °C compared to 16 °C for the deletion lines (**Figure 4.17**).

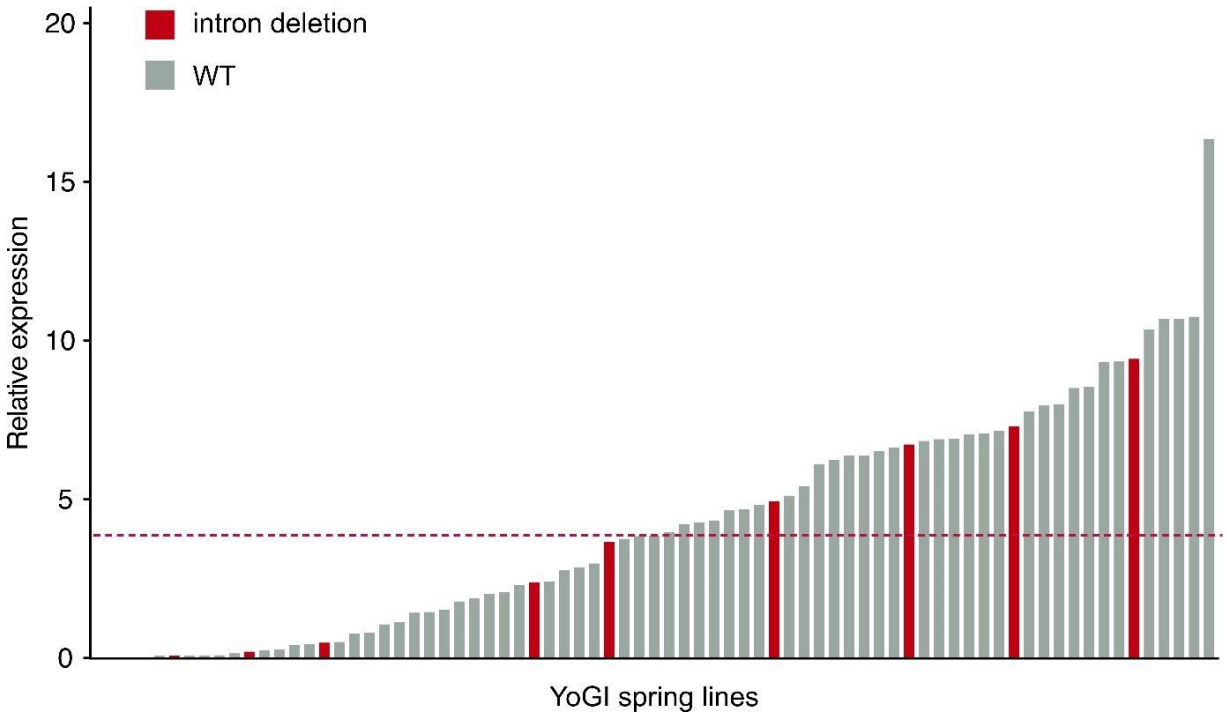


Figure 4.14: Expression of *TraesCS3B02G470000* in YoGI spring wheat lines. Lines with the WT version of the intron (n = 64) are grey and lines with the intron deletions (n = 11) are red. Red dotted line represents the median expression (3.86 fpkm) across all 75 lines.

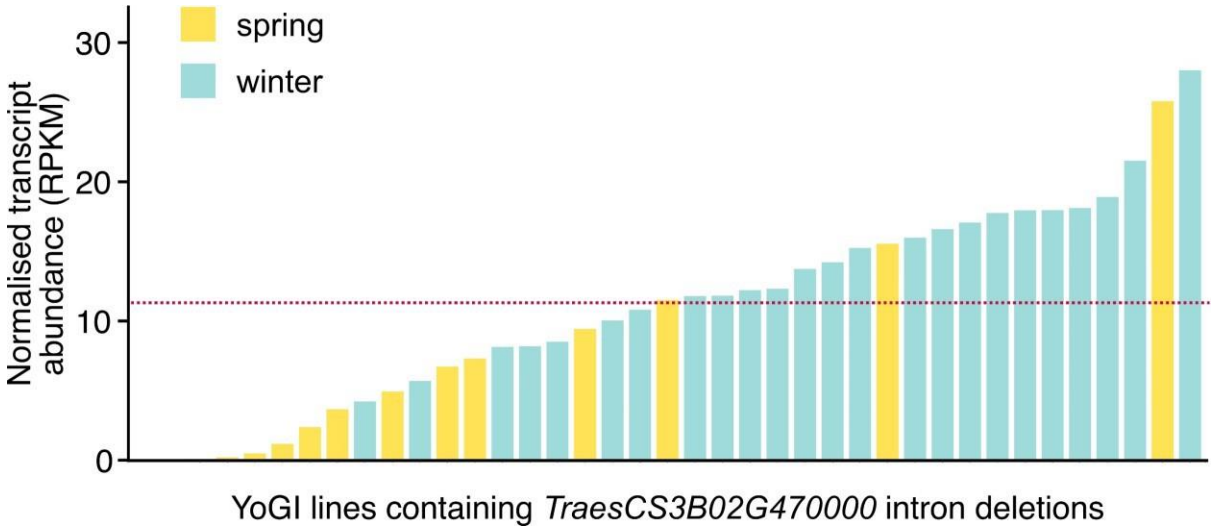


Figure 4.15: Expression of *TraesCS3B02G470000* in YoGI lines containing the first intron deletions. Spring (yellow; n = 15) and winter (blue; n = 24) habit expression is plotted. Red dotted line represents the median expression across all 39 lines.

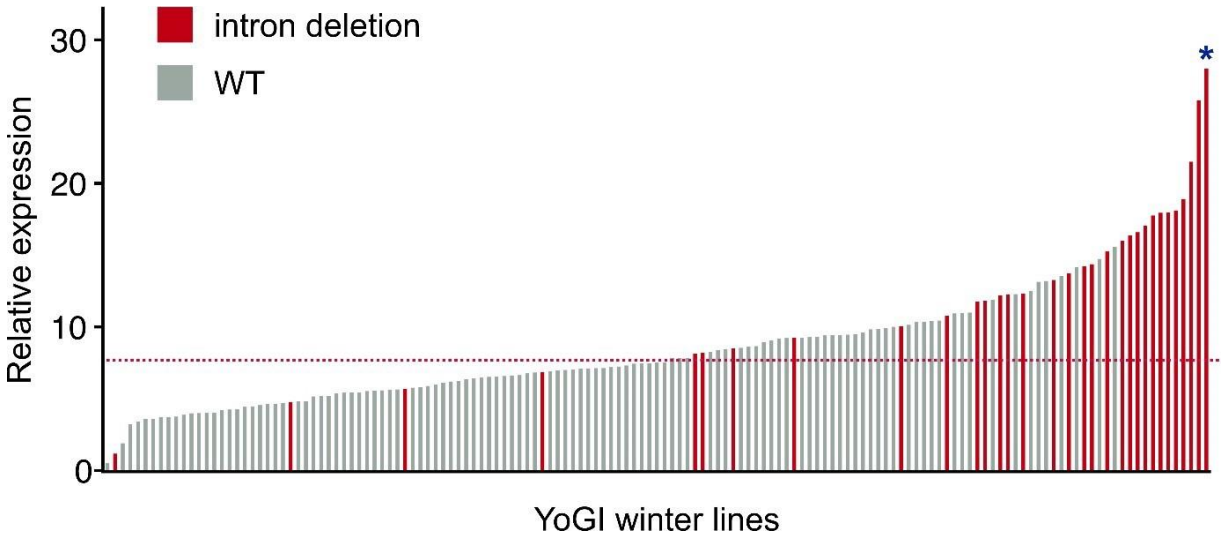


Figure 4.16: Expression of *TraesCS3B02G470000* in winter habit YoGI lines. The expression of lines containing the first intron deletions (red; n = 32) and WT lines without the deletions (grey; n = 113) is plotted. Red dotted line represents the median expression across all 145 lines. Blue asterisk represents line WATDE0848. The p-value is = 0.00001, Student's t-test.

In winter lines, a similar trend in apical development was observed. In lines containing the intron deletions, apices developed at a slower rate when compared to the wild type line; this result was quantified by the size of the apex (**Figure 4.18**). Across the developmental time series, leaf samples were also taken at each leaf emergence stage. Furthermore, higher expression of *TraesCS3B02G470000* was observed in the deletion line compared to the wild type line (**Figure 4.18**). However, this difference was minimal and didn't clarify whether the intronic deletion resulted in higher expression under these conditions.

4.3.8 Transgenic *TraesCS3B02G470000* may shed light on gene function

The population studies and developmental time course indicated a role for *TraesCS3B02G470000* in development. This gave me confidence that *TraesCS3B02G470000* was involved in a biologically relevant temperature response in wheat. To understand the function of the gene, it was informative to study *TraesCS3B02G470000* under biologically controlled conditions where the genetic background was not a confounding factor. To do this, transgenic *TraesCS3B02G470000* lines were developed. Due to the large size of the gene (31,262 bp) and in particular the first intron (29,407 bp), a construct was developed which only contained the coding region (504 bp), the 2.5 kb region upstream containing the native promoter, and HA-tag (160 bp) to facilitate subsequent protein analysis. Therefore, the total construct size including the 2,710 bp vector backbone was 5,874 bp. Due to the challenges of transformation, the construct was transformed into the spring cv. Cadenza – this additionally enabled the opportunity to investigate gene function in a spring background. The outcome of transformation was increased transgenic copies of the gene in the initial transformants (T0, received from NIAB). As a control, Cadenza plants were transformed with an empty plasmid, but still contained their original/native *TraesCS3B02G470000*. Cadenza is a spring wheat that doesn't have a vernalization requirement – therefore, the initial T0 transformants were grown under a long day photoperiod (16hr light/8hr dark) at 20°C.

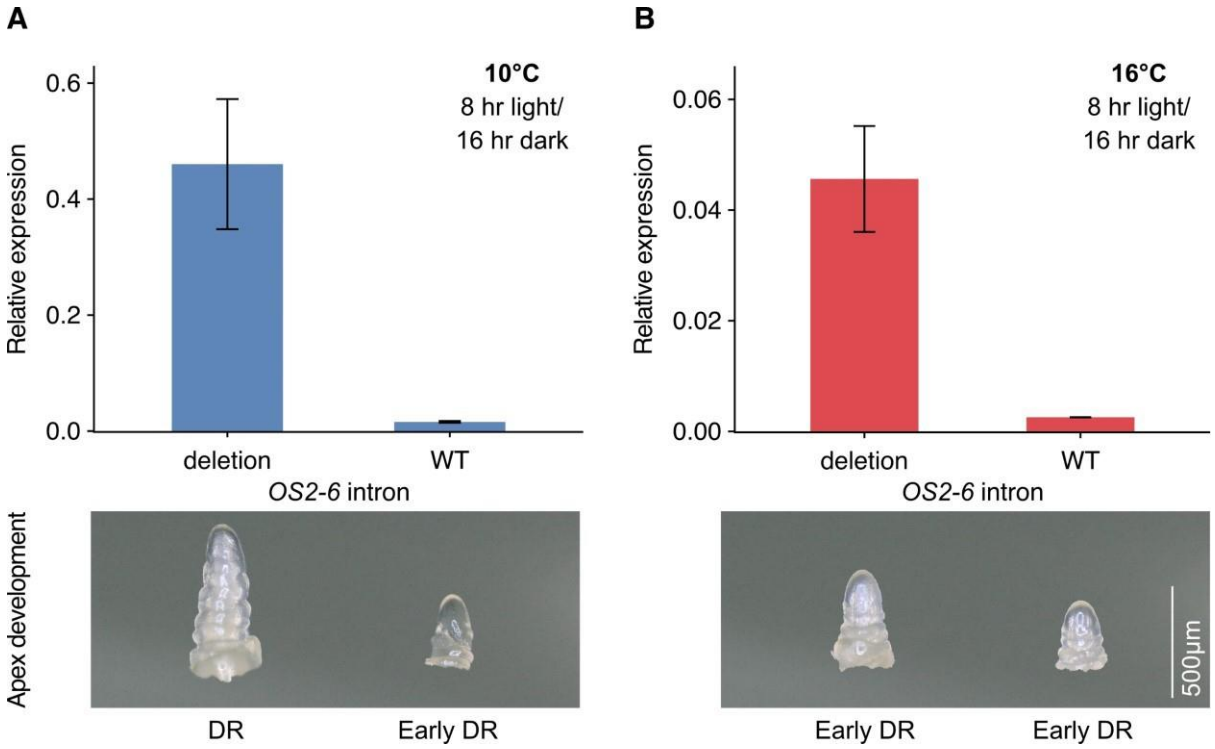


Figure 4.17: Expression of *TraesCS3B02G470000* and apex development in spring lines containing the different versions of the first intron. Expression and apex development under (**A**) 10 °C and (**B**) 16 °C short day (8 hr light/16 hr dark) conditions at the 4-leaf stage of development. Error bars represent standard error of the mean for 3 biological replicates. Below, apex stage of development; DR = double ridge apex.

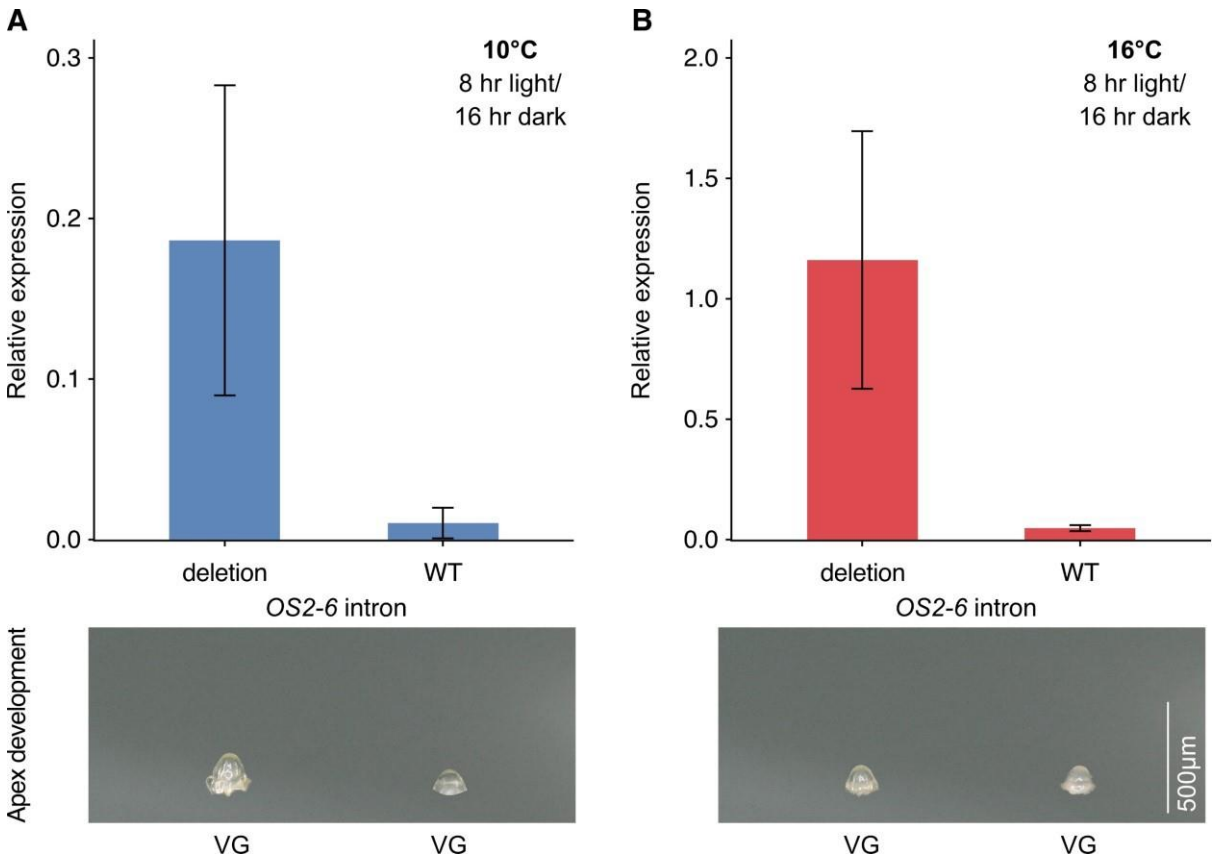


Figure 4.18: Expression of *TraesCS3B02G470000* and apex development in spring lines containing the different versions of the first intron. Expression and apex development under (**A**) 10 °C and (**B**) 16 °C short day (8 hr light/16 hr dark) conditions at the 4-leaf stage of development. Error bars represent standard error of the mean for 3 biological replicates. Below, apex stage of development; VG = vegetative apex.

Transgenic Cadenza plants that had higher expression of *TraesCS3B02G470000* were crossed into the winter cv. Hereward and Charger with the intention of studying the transgene in a winter background. Selection of the transgene was made based on the lack of introns. This generated 11 individual Charger plants containing the *TraesCS3B02G470000* transgene. Hereward x transgenic crosses were also generated, with confirmation of the presence of the transgene in progress.

When analysing the flowering response in wheat plants that had been transformed with the *TraesCS3B02G470000* transgene, no significant difference was observed between the number of days until flowering in lines transformed with the transgene and the control lines transformed with an empty construct. Furthermore, there was no significant difference in time until flowering in transgenic lines with different gene copy number (**Figure 4.19**). The transgenic line “T-7” took the longest to flower by individual plant (88 days), whereas the transgenic line “T-2” flowered the fastest, with the first plant flowering after 71 days. Other plants from this line flowered after 73 and 74 days; however, only 3 of the 7 plants successfully survived until flowering (**Figure 4.19**).

Although no substantial difference was detected in flowering time, unusual ear phenotypes were observed in the transgenic *TraesCS3B02G470000* lines compared to the control line. The control line had ears that were representative of the cv. Cadenza, the background within which the lines/transgenes were transformed. In contrast, transgenic *TraesCS3B02G470000* lines had a range of ear phenotypes (**Figure 4.20**). This included multiple lines exhibiting lack of floret dieback, resulting in distinct elongated spikelets with higher floret number (**Figure 4.20**). Also observed was necrosis of spikelets or spikelets drastically reduced in size formed on the tip of the ear (**Figure 4.20**).

The observed abnormal ears didn’t result in a recognizable impact on spikelet number. Compared to the control line, transgenic lines “T-3”, “T-4”, “T-5”, “T-6”, “T-7”, and “T-8” had no significant difference in final spikelet number (**Figure 4.21**). Conversely, line “T-2” did have a significant difference in spikelet number compared to the control line ($p = 0.0012$); however, as mentioned earlier, only 3 individuals in this line flowered, suggesting stress affected/influenced the decreased spikelet number per ear (**Figure 4.21**).

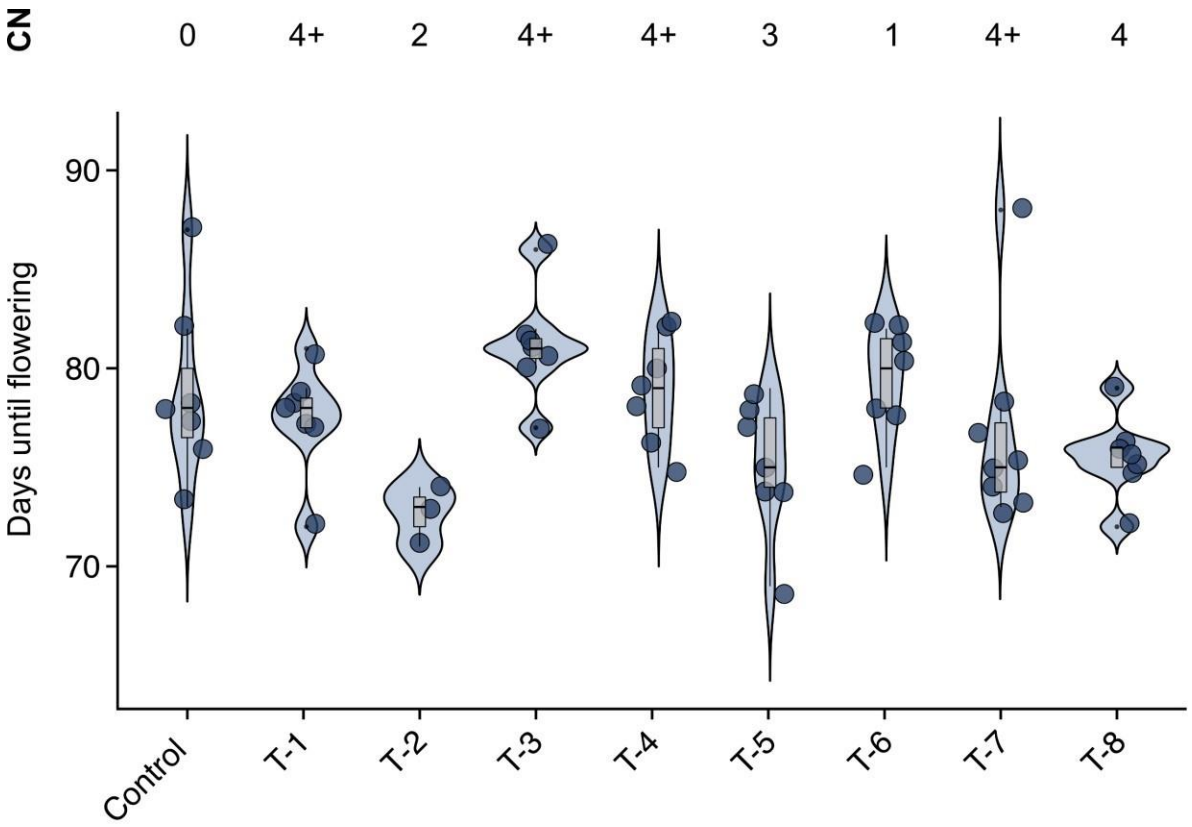


Figure 4.19: Flowering of *TraesCS3B02G470000* transgenic wheat lines. Flowering refers to the number of days from germination until half-ear emergence. T-* = lines transformed with *TraesCS3B02G470000* transgene with its native promoter; n ≤ 8. CN = *TraesCS3B02G470000* copy number.

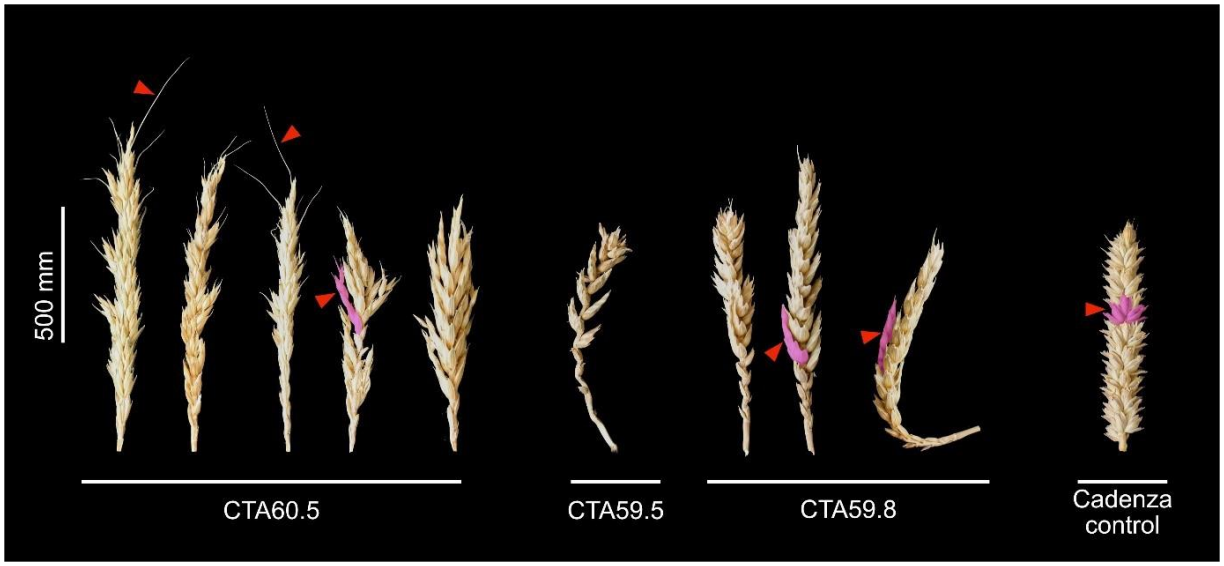


Figure 4.20: Range of ear phenotypes in Cadenza wheat transformed with *TraesCS3B02G470000* transgene. Areas highlighted in pink show the variety in spikelet appearance; red arrows indicate features of interest, including the presence of awns in transgenic line CTA60.5.

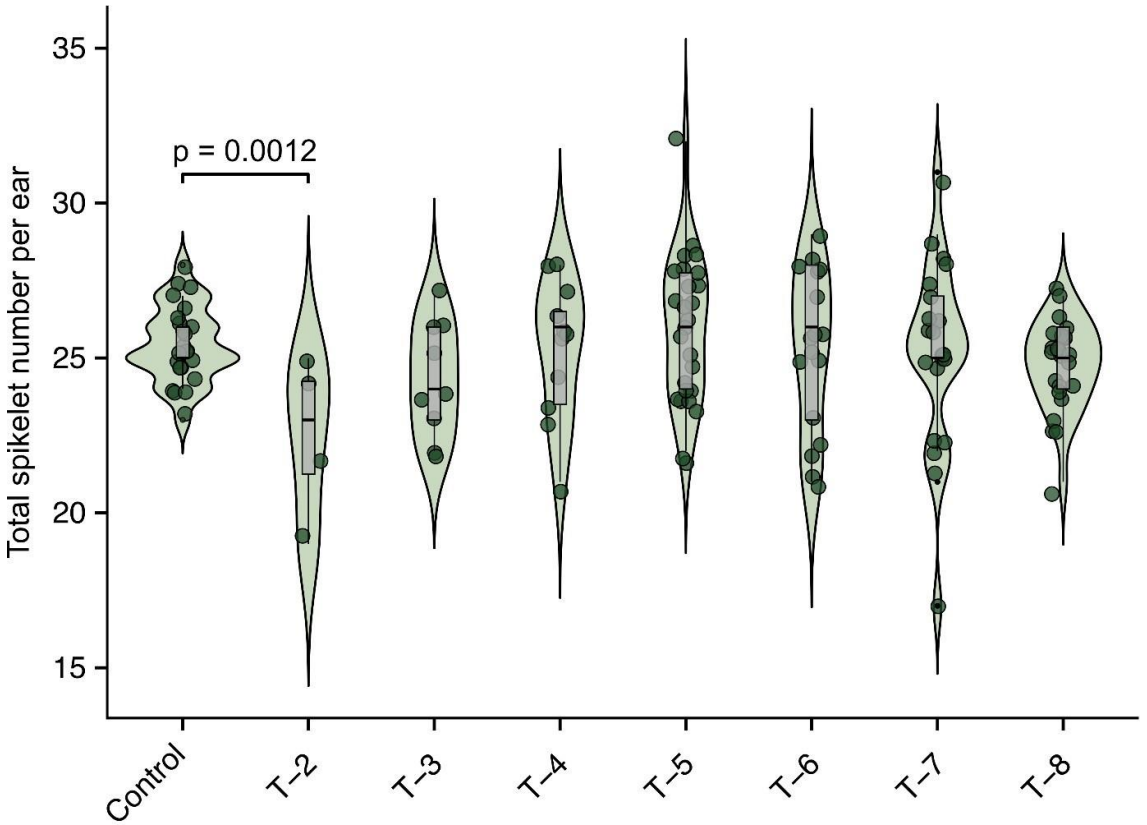


Figure 4.21: Total spikelet number of *TraesCS3B02G470000* transgenic wheat lines. T-* = lines transformed with *TraesCS3B02G470000* transgene with its native promoter.

4.4 Discussion

In the previous chapter, we introduced 15 potential *ODDSOC* candidate genes in bread wheat. Here, we aimed to explore this group of genes further to elucidate their role in vernalization of winter wheat. Firstly, expression of all 15 genes was examined in the winter bread wheat cultivar Hereward. This cultivar has a high vernalization requirement and is commonly grown in the UK; for this reason, Hereward is able to develop slowly throughout the UK winter season and is expected to require multiple weeks of low temperature before vernalization is completed. Therefore, examining expression of *ODDSOC* genes in this cultivar as it grows as part of a winter field trial can be informative in understanding the role of these genes in vernalization. As the flowering response in Hereward indicated that vernalization was saturated after 10 weeks of growth in the field, we would anticipate that, as putative floral repressors, *OS2* genes would be down regulated by this point. Although a subset of these genes were down regulated by 10 weeks and continued to be suppressed beyond this point, not all putative *OS2* genes followed this pattern.

Strikingly, when plants remained in the field beyond the 10 week saturation of vernalization point, flowering was steadily and increasing delayed when plants were unearthed and transferred into 22 °C long day glasshouse conditions. This is reminiscent of the repressive epigenetic modifications being gradually removed from the loci that are silenced upon completion of vernalization, effectively resetting the memory of winter and the need for vernalization (Sheldon et al., 2008). This is an integral process that prevents the “memory of winter” from being passed to progeny of vernalized plants, therefore preventing premature flowering (Choi et al., 2009; Sheldon et al., 2008). This process of high temperature resetting the memory for vernalization is known as de-vernalization (Purvis & Gregory, 1945). De-vernalization has been reported to occur at temperatures higher than 25 °C in a range of species such as Arabidopsis (Bouché et al., 2015; Shindo et al., 2006), root chicory (Périlleux et al., 2013), lettuce (Marks & Prince, 1979), onion (Khokhar, 2008), barley (Monteagudo et al., 2019), and bread wheat (Dixon et al., 2019). We would anticipate that in response to the warmer temperatures post-10 weeks in the field, floral repressors would be activated, as has previously been observed (Dixon et al., 2019; Monteagudo et al., 2019).

Since the *ZCCT* genes remained switched off beyond 10 weeks, this would suggest that if Hereward plants were undergoing de-vernalization, the repressive histone modifications were not removed from the *VRN2* locus. Conversely, *ODDSOC* genes *TraesCS7A02G260900*, *TraesCS7B02G158900*, and *TraesCS7D02G261900* continued to be up regulated beyond 10 weeks; this could indicate that these are floral repressors that aren't successfully repressed by floral activators within the first 10 weeks. However, another possibly more likely explanation is that these gene are functioning as floral activators; this aligns with the pattern of up regulation of *VRN-A1* across this time course (**Figure 4.3**). Interestingly, hypothesised floral activator *TraesCS3A02G432900* was not re-up regulated beyond 10 weeks, unlike *VRN-A1* which was dramatically up regulated. This could suggest that the involvement of *TraesCS3A02G432900* in vernalization and floral initiation is most significant in the early stages of development. Moreover, it is surprising that the steady reduction and increase in days until flowering doesn't correspond with the expression pattern of down regulation and subsequent up regulation of a floral repressor. This supports the hypothesis that the cereal vernalization pathway consists of elements beyond *VRN1* and *VRN2*, including genes that could function as floral repressors at different stages of development.

Since the expression of all *OS2* genes didn't behave like classical floral repressors throughout vernalization in the field, I wanted to identify which of these genes were the most prominent floral repressors. Therefore, I attempted to identify which of these genes were the most conserved using the Watkins collection haplotype maps. Although the coding regions of the *OS2* genes were relatively equally conserved within the Watkins collection, conserved large and small deletions were discovered within the first intron of *TraesCS3B02G470000*. This indicates a possible regulatory function for this intron and an important function for this gene as these deletions were conserved in close to 400 lines. The intronic deletions were most prevalent in facultative wheat, with 52% of all facultative wheat lines in the Watkins population containing the deletions. Wheat with a facultative growth habit is cold tolerant and has a low to non-existent vernalization requirement, enabling these varieties to be grown either in winter or spring plantings (von Zitzewitz et al., 2005). Therefore, the prevalence of these deletions in facultative wheat could indicate that absence of

this intronic region reduces vernalization requirement, similar to how a deletion within the first intron of *VRN1* results in spring growth habit.

Moreover, the presence of multiple predicted CArG box sites in both intronic deletions could signify that this is a region that MIKC-type MADS-Box proteins, including TraesCS3B02G470000, could bind. In addition, binding sites associated with flowering-related genes such as *FLC*, *SOC1*, and various *AGL* genes identified in these deleted intronic regions further supported the potential regulatory role of this *TraesCS3B02G470000* intron. It should be noted that predicted binding sites were selected from an Arabidopsis motif database as no comprehensive bread wheat motif database is currently available; unlike CArG-box binding sites, where the $CC(A/T)_6GG$ motif is a recognised binding site across species (de Folter & Angenent, 2006), it is unclear whether an Arabidopsis *FLC* predicted motif would be a functional binding site in bread wheat. Therefore, it's possible that these identified motifs aren't functional binding sites in bread wheat. For this reason, these identified motifs require additional verification, such as chromatin immunoprecipitation sequencing (ChIP-seq) (Park, 2009) or surface plasmon resonance (SPR) (Nguyen et al., 2015).

In this study, direct interaction between the putative DNA binding site and protein of interest was examined using a yeast one-hybrid assay. This identified an interaction between TraesCS3B02G470000 and the SPL-associated binding site present in the intronic deletions. Interestingly, a variation of the SPL-associated binding site was also identified within the first intron of *TraesCS3A02G432900* that overlapped with the lncRNA antisense transcript for this gene. Since TraesCS3A02G432900 and TraesCS3B02G470000 were shown to interact in section **3.3.7**, it's possible that TraesCS3A02G432900 is able to bind to the SPL-associated binding site in both the *TraesCS3B02G470000* intron and its own intron. We would assume that *TraesCS3A02G432900*, as a potential floral activator, could be binding to the *TraesCS3B02G470000* intron to repress expression of this gene. Conversely, expression analysis from section **3.3.10** would suggest that the lncRNA up regulates *TraesCS3A02G432900*, therefore it's possible that TraesCS3A02G432900 could bind the intronic SPL motif as part of this regulation. However, further validation is required to investigate this possible interaction.

One possible explanation for the function of this intron in regulation of *TraesCS3B02G470000* is involvement in temperature responsiveness. In section **3.3.9**, it was observed that *TraesCS3B02G470000* expression in winter cv. Buster (no intronic deletions) was higher under 16 °C compared to 10 °C short day conditions. Speculatively, this indicates a repressive role for this gene in winter wheat during the end of autumn when daylength is decreasing but temperatures are still relatively warm and vernalization is yet to be completed. However, wheat lines with the intronic deletions were geographically associated with regions with lower temperatures during autumn, winter, and spring. Moreover, considering that expression of *TraesCS3B02G470000* was higher in winter deletion lines compared to spring deletion lines, it's possible that regulation via the intron is more significant in winter wheat. Supporting the possible suggestion of floral activator TraesCS3A02G432900 binding the intron to repress *TraesCS3B02G470000* activity, winter lines with the intronic deletions had significantly higher expression. This indicates a connection between this region of the intron being involved in down regulation of *TraesCS3B02G470000*.

Conversely, the role of this intron in regulation of *TraesCS3B02G470000* in spring wheat is more ambiguous. Based on the Watkins flowering data, spring lines with the deletions flowered later. However, the lack of significant difference in expression between spring lines with and without the intronic deletions across the YoGI panel corresponded with the lack of difference in expression level of the lines.Furthermore, transgenic spring cv. Cadenza transformed with *TraesCS3B02G470000* still successfully flowered, indicating that multiple copies of this gene doesn't suppress flowering. This explanation was further supported by the lack of significant difference in the rate of flowering in lines transformed with *TraesCS3B02G470000* and the control. Since the *TraesCS3B02G470000* transgene only contained the CDS, we can assume that these transgenes could not be regulated by their intron as it was absent in the transgenic plants. Although *TraesCS3B02G470000* transgenic Cadenza didn't exhibited a different rate of flowering, there did appear to be an impact on ear formation, indicating a role in development of reproductive structures. This is comparable to MADS-Box floral repressor *VRT2*, where constitutive expression of *VRT2* in tetraploid wheat Kronos has been shown to impact spike development (Li et al., 2021), whereas high copy number of *VRT-A2* in transgenic hexaploid

wheat corresponds with increased spike, glume, lemma, and grain length (Adamski et al., 2021). Moreover, sequence rearrangement within the first intron of *VRT-A2* was linked to the elongated glume phenotype (Adamski et al., 2021). For this reason, further investigation into the effects of the *TraesCS3B02G470000* transgene in a winter cultivar, such as the transgenic Charger crosses generated in this chapter (section **4.3.2**), would be informative. Similarly, flowering response and expression analysis under different conditions needs to be examined, as well as expression at different stages of development.

Chapter 5: General Discussion

5.1 Summary of findings

5.2 Speed vernalization and vernalization under warmer temperatures adds to our understanding of the vernalization pathway in cereals

With climates anticipated to become more unpredictable in the future, safeguarding crop production from a myriad of biotic and abiotic stresses is a major concern. For winter cultivars that have a moderate to high vernalization requirement, the prospect of warmer temperatures throughout autumn and winter was predicted to impact the vernalization response. In chapter 2 section **2.3.1.1**, winter wheat cultivars with a high vernalization requirement (cultivars Charger and Hereward) were successfully vernalized under higher ambient temperatures of 10 °C-14 °C. This indicates that these cultivars are more robust to warmer winter temperatures than would have been initially thought. However, when cv. Hereward was left in the field beyond saturation of the vernalization requirement, this delayed the rate of flowering. Therefore, vernalization can proceed under a warmer temperature but variables post-vernalization can initiative resetting of repressive epigenetic modifications. For this reason, we would anticipate that the current sowing window for winter cultivars might shift in the future.

Speed vernalization reveals that vernalization can proceed under a combination of warmer temperature, extended photoperiod, and – perhaps most surprisingly – exposure to direct high light intensity from germination. Crucially, the aim of speed vernalization is not to fully replace the standard vernalization procedure; as mentioned in chapter 1, the high-yielding trait of winter wheat compared to spring wheat is partly attributed to the longer growth cycle, with the months of slow growth over winter and spring facilitating grain development. Instead, we anticipate that speed vernalization would be most effective when integrated into crossing programmes, comparable to single seed descent (SSD) where the aim is to generate F_n for trait introgression (Muehlbauer et al., 1981).

Here, the speed vernalization protocol introduced in chapter 2 has the potential to substantially contribute to plant breeding programmes. This extends beyond wheat, with an application in the generation of any crop that has a vernalization requirement.

Arguably, the greatest benefit that this offers is a reduction in generation time, enabling more generation cycles per year. This has already shown to be effective in reducing the vernalization requirement in certain winter barley cultivars (Cha et al., 2022). Typically, winter barley is vernalized at 2 °C-6 °C for >7 weeks (Greenup et al., 2011; Monteagudo et al., 2019; Sasani et al., 2009; von Zitzewitz et al., 2005). Under these standard vernalization conditions and typical long day conditions post-vernalization, a maximum of 2 generations of winter barley can be produced in a year (Watson et al., 2018). Using speed vernalization, up to 5 generations of winter barley can be achieved within 12 months (Cha et al., 2022).

We would anticipate that speed vernalization could offer a similar reduction in generation time for other winter crops such as rye and canola (*B. napus*). Rye (*Secale cereale L*.) is predominantly produced in northern Europe and is predicted to have strong adaptability to the changing climate in this part of the world (Hackauf et al., 2022). In rye breeding programmes, hybrid breeding is the main method of improving grain yield (Hackauf et al., 2022; Miedaner et al., 2022; Miedaner & Laidig, 2019). Winter rye generally requires 7-9 weeks vernalization at 2 °C – 4 °C (Erath et al., 2017; Marcińska et al., 2018; Zieliński et al., 2020).

Outside of cereals, *B. napus* (canola) is generally vernalized at 4 °C-5 °C for 8 weeks (Matar et al., 2021; Schiessl et al., 2019). As with barley, this vernalization step adds 8 weeks onto the generation process; considering that a natural neutral to long day photoperiod enables the production of 2 generations of spring canola (Watson et al., 2018), winter canola crops can be limited to a single generation within 12 months. Similarly, *Brassica oleracea* sub-species such as cabbage and broccoli, winter cultivars are vernalized at 5 °C for 8 weeks (Lin et al., 2005). Therefore, we would anticipate a similar generation time as winter canola. To maximise the efficacy of the speed vernalization reducing generation time, we recommend following speed vernalization with speed breeding. In spring cultivars, speed breeding enables the production of up to 6 generations of barley and 5 generations of wheat per cycle, compared to 2-3 generations under a natural variable photoperiod of 10-16 hours (Cha et al., 2022; Watson et al., 2018). However, up until this point, winter cultivars have not been able to experience the full benefits of reduction in generation time offered by speed breeding. This is largely due to the standard practice of up to

10 weeks vernalization, working against the reduction in generation time offered by speed breeding.

In addition to contributing to plant breeding programmes, speed vernalization can also broaden current understanding of the vernalization pathway in cereals. Considering that well-characterised vernalization genes *VRN1*, *VRN2*, and *VRT2* don't account for the vernalization responses that were observed (**Figure 3.20**), this suggests that additional overlooked genes are involved in this response or even an entirely discrete pathway. Since the SGV speed vernalization protocol incorporates exposure to high light intensity for a 22 hr photoperiod from germination, this raises the possibility of light-responsive or circadian elements such as *PHYC* or *ELF3* playing a key role in this process. Additionally, several of the *ODDSOC* genes characterised in chapters 3 and 4 were also expressed under speed vernalization conditions (**Figure 3.20**). Expression level was higher than both *VRN1* and *VRN2*, therefore indicating involvement of these *ODDSOC* genes in the speed vernalization pathway.

5.3 *ODDSOC* genes in bread wheat

The Arabidopsis gene *FLC* has been paramount to understanding vernalization. Despite having a central role in Brassicaceae vernalization, the identification of *FLC* in cereals has remained elusive. Recently, *FLC*-like genes have been identified in cereals, with a link established between *FLC*-like genes and *OS1* and *OS2* in *B. distachyon* (Ruelens et al., 2013).

In this thesis, the function of the *ODDSOC* genes with respect to temperature in bread wheat was investigated. This indicated that, excluding *TraesCS3A02G432900*, the 15 *ODDSOC* genes could all be sorted into homoeologous groups that shared a high level of similarity. An effort has been made to move away from conclusively classifying these genes as either *OS1* or *OS2*; in previous studies, the identification of *OS2* genes in different cereals has resulted in different naming classifications (e.g., *AGL33*, *OsMADS*, *FLC*-like). However, results of the expression analysis in chapter 3 revealed a range of responses. For this reason, I chose to examine the group of 15 genes as a whole to understand their role in temperature responsiveness in bread wheat.

Here, I report that *TraesCS3A02G432900* is active under lower temperatures and is a strong candidate for a floral activator, promoting flowering in a manner similar to *VRN1*. Furthermore, several lncRNA antisense transcripts associated with *TraesCS3A02G432900* were identified and are likely involved in the regulation of this gene. There was also strong evidence that the large first intron also occupies a regulatory role, especially since it is highly conserved in 12 of the 15 genes investigated in this thesis. For the gene *TraesCS3B02G470000*, regulation via the first intron is also likely; conserved deletions within the intron are associated with higher expression of this gene in winter cultivars. Paired with geographical location, this suggests a role for *TraesCS3B02G470000* in climate adaptability. Moreover, expression level of this gene is higher under warm short day conditions – therefore, we could speculate that *TraesCS3B02G470000* represses premature flowering during unfavourable conditions, ultimately functioning as a floral repressor similar to *VRN2*. Since TraesCS3A02G432900 and TraesCS3B02G470000 interacted within a yeast one-hybrid, this also supports the possibility that *TraesCS3A02G432900* is involved in repression of *TraesCS3B02G470000*, reminiscent to how *VRN1* downregulates *VRN2* during vernalization.

The interplay between *VRN1* and *VRN2* underpins the flowering response in cereals, with allelic variation at these loci driving growth habit. However, this doesn't account for the variation in vernalization response reported in cultivars with the same alleles for these loci (Rizza et al., 2016). This introduces the possibility of temperature responsive *ODDSOC* genes also being involved in determining variation in vernalization requirement. In this thesis, TraesCS3B02G470000 was shown to interact with VRN-A1, but not with a version of the VRN-A1 protein associated with higher vernalization requirement. This preferential binding further supports the potential involvement of *TraesCS3B02G470000* in variation of vernalization requirement.

This raises the possibility of other *ODDSOC* genes also occupying a role in vernalization. Here, not all 15 genes were investigated to the same level of depth, as this was beyond the scope of the PhD; however, genes including *TraesCS3A02G434400*, *TraesCS3D02G427700*, and *TraesCS3D02G428000* had intriguing expression profiles under different temperatures and photoperiods. This

warrants further research into the function of these genes, particularly in relation to *TraesCS3A02G432900* and *VRN-A1*. In addition, the winter wheat crosses containing the *TraesCS3B02G470000* transgene will enable future examination of this transgene in a winter background, furthering our understanding of the function of *TraesCS3B02G470000* in winter cultivars. To complement this, recent advances in wheat transformation protocols now enables the development of transgenic winter wheat (Hahn et al., 2021), introducing the opportunity to study other *ODDSOC* genes in a transgenic winter wheat system.

References

Adams, S. R., & Langton, F. A. (2005). Photoperiod and Plant Growth: a review. *The journal of horticultural science & biotechnology*, *80*(1), 2-10. https://doi.org/10.1080/14620316.2005.11511882

Adamski, N. M., Simmonds, J., Brinton, J. F., Backhaus, A. E., Chen, Y., Smedley, M., Hayta, S., Florio, T., Crane, P., Scott, P., Pieri, A., Hall, O., Barclay, J. E., Clayton, M., Doonan, J. H., Nibau, C., & Uauy, C. (2021). Ectopic expression of Triticum polonicum VRT-A2 underlies elongated glumes and grains in hexaploid wheat in a dosage-dependent manner. *The Plant Cell*, *33*(7), 2296-2319. https://doi.org/10.1093/plcell/koab119

Aerts, N., de Bruijn, S., van Mourik, H., Angenent, G. C., & van Dijk, A. D. J. (2018). Comparative analysis of binding patterns of MADS-domain proteins in Arabidopsis thaliana. *BMC plant biology*, *18*(1), 131. https://doi.org/10.1186/s12870-018-1348-8

Alabdullah, A. K., Borrill, P., Martin, A. C., Ramirez-Gonzalez, R. H., Hassani-Pak, K., Uauy, C., Shaw, P., & Moore, G. (2019). A Co-Expression Network in Hexaploid Wheat Reveals Mostly Balanced Expression and Lack of Significant Gene Loss of Homeologous Meiotic Genes Upon Polyploidization [Original Research]. *Frontiers in plant science*, *10*. https://doi.org/10.3389/fpls.2019.01325

Alexandre, C. M., & Hennig, L. (2008). FLC or not FLC: the other side of vernalization. *Journal of Experimental Botany*, *59*(6), 1127-1135. https://doi.org/10.1093/jxb/ern070

Alvarez-Buylla, E. R., Pelaz, S., Liljegren, S. J., Gold, S. E., Burgeff, C., Ditta, G. S., Ribas de Pouplana, L., Martínez-Castilla, L., & Yanofsky, M. F. (2000). An ancestral MADS-box gene duplication occurred before the divergence of plants and animals. *Proceedings of the National Academy of Sciences*,

97(10), 5328-5333. https://doi.org/10.1073/pnas.97.10.5328

Amano, T., Smithers, R., Sparks, T., & Sutherland, W. (2010). A 250-year index of first flowering dates and its response to temperature changes. *Proceedings. Biological sciences / The Royal Society*, *277*, 2451-2457. https://doi.org/10.1098/rspb.2010.0291

Amasino, R. M., & Sung, S. (2004). Vernalization in Arabidopsis thaliana is mediated by the PHD finger protein VIN3. *Nature*, *427*(6970), 159-164. https://doi.org/10.1038/nature02195

Amberg, D. C., Burke, D. J., & Strathern, J. N. (2006). “Quick and Dirty” Plasmid Transformation of Yeast Colonies. *Cold Spring Harbor protocols*, *2006*(1), pdb.prot4146. https://doi.org/10.1101/pdb.prot4146

Andrés, F., & Coupland, G. (2012). The genetic basis of flowering responses to seasonal cues. *Nature Reviews Genetics*, *13*(9), 627-639. https://doi.org/10.1038/nrg3291

Appels, R., Eversole, K., Stein, N., Feuillet, C., Keller, B., Rogers, J., Pozniak, C. J., Choulet, F., Distelfeld, A., Poland, J., Ronen, G., Sharpe, A. G., Barad, O., Baruch, K., Keeble-Gagnère, G., Mascher, M., Ben-Zvi, G., Josselin, A.-A., Himmelbach, A., . . . Wang, L. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*, *361*(6403), eaar7191. https://doi.org/doi:10.1126/science.aar7191

Båga, M., Chodaparambil, S. V., Limin, A. E., Pecar, M., Fowler, D. B., & Chibbar, R. N. (2007). Identification of quantitative trait loci and associated candidate genes for low-temperature tolerance in cold-hardy winter wheat. *Functional & Integrative Genomics*, *7*(1), 53-68. https://doi.org/10.1007/s10142-006-0030-7

Balla, K., Rakszegi, M., Li, Z., Bekes, F., Bencze, S., & Veisz, O. (2011). Quality of winter wheat in relation to heat and drought shock after anthesis. *Czech Journal of Food Sciences*, *29*(2), 117-128.

https://cjfs.agriculturejournals.cz/artkey/cjf-201102-0004.php http://dx.doi.org/10.17221/227/2010-CJFS

Barratt, L. J., He, Z., Fellgett, A., Wang, L., Mason, S. M., Bancroft, I., & Harper, A. L. (2023). Co-expression network analysis of diverse wheat landraces reveals markers of early thermotolerance and a candidate master regulator of thermotolerance genes. *The Plant journal : for cell and molecular biology*. https://doi.org/10.1111/tpj.16248

Barratt, L. J., Reynolds, I. J., Franco Ortega, S., & Harper, A. L. (2023). Transcriptomic and co-expression network analyses on diverse wheat landraces identifies candidate master regulators of the response to early drought [Original Research]. *Frontiers in plant science*, *14*. https://doi.org/10.3389/fpls.2023.1212559

Beales, J., Turner, A., Griffiths, S., Snape, J. W., & Laurie, D. A. (2007). Pseudo-Response Regulator is misexpressed in the photoperiod insensitive Ppd-D1a mutant of wheat (Triticum aestivum L.). *Theoretical and applied genetics*, *115*(5), 721-733. https://doi.org/10.1007/s00122-007-0603-4

Becker, A., & Theißen, G. (2003). The major clades of MADS-box genes and their role in the development and evolution of flowering plants. *Molecular Phylogenetics and Evolution*, *29*(3), 464-489. https://doi.org/https://doi.org/10.1016/S1055-7903(03)00207-0

Bennett, T., & Dixon, L. E. (2021). Asymmetric expansions of FT and TFL1 lineages characterize differential evolution of the EuPEBP family in the major angiosperm lineages. *BMC Biology*, *19*(1), 181. https://doi.org/10.1186/s12915-021-01128-8

Bieniawska, Z., Espinoza, C., Schlereth, A., Sulpice, R., Hincha, D. K., & Hannah, M. A. (2008). Disruption of the Arabidopsis Circadian Clock Is Responsible for Extensive Variation in the Cold-Responsive Transcriptome *Plant Physiology*, *147*(1), 263-279. https://doi.org/10.1104/pp.108.118059

Bin Rahman, A. N. M. R., & Zhang, J. (2018). Preferential Geographic Distribution Pattern of Abiotic Stress Tolerant Rice. *Rice*, *11*(1), 10. https://doi.org/10.1186/s12284-018-0202-9

Borlaug, N. E. (1983). Contributions of conventional plant breeding to food production. *Science (American Association for the Advancement of Science)*, *219*(4585), 689-693. https://doi.org/10.1126/science.219.4585.689

Borrill, P., Ramirez-Gonzalez, R., & Uauy, C. (2016). expVIP: a Customizable RNA-seq Data Analysis and Visualization Platform *Plant Physiology*, *170*(4), 2172-2186. https://doi.org/10.1104/pp.15.01667

Boss, P. K., Bastow, R. M., Mylne, J. S., & Dean, C. (2004). Multiple pathways in the decision to flower: enabling, promoting, and resetting. *Plant Cell*, *16 Suppl*(Suppl), S18-31. https://doi.org/10.1105/tpc.015958

Bouché, F., Detry, N., & Périlleux, C. (2015). Heat can erase epigenetic marks of vernalization in Arabidopsis. *Plant Signaling & Behaviour*, *10*(3), e990799. https://doi.org/10.4161/15592324.2014.990799

Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*, *34*(5), 525-527. https://doi.org/10.1038/nbt.3519

Brown, T. A., Jones, M. K., Powell, W., & Allaby, R. G. (2009). The complex origins of domesticated crops in the Fertile Crescent. *Trends in ecology & evolution (Amsterdam)*, *24*(2), 103-109. https://doi.org/10.1016/j.tree.2008.09.008

Bugbee, B., & Koerner, G. (1997). Yield comparisons and unique characteristics of the dwarf wheat cultivar 'USU-Apogee'. *Advances in Space Research*, *20*(10), 1891-1894.

Büntgen, U., Piermattei, A., Krusic, P. J., Esper, J., Sparks, T., & Crivellaro, A. (2022). Plants in the UK flower a month earlier under recent warming. *Proceedings of the Royal Society B: Biological Sciences*, *289*(1968),

20212456. https://doi.org/10.1098/rspb.2021.2456

Castro-Mondragon, J. A., Riudavets-Puig, R., Rauluseviciute, I., Berhanu Lemma, R., Turchi, L., Blanc-Mathieu, R., Lucas, J., Boddie, P., Khan, A., Manosalva Pérez, N., Fornes, O., Leung, Tiffany Y., Aguirre, A., Hammal, F., Schmelter, D., Baranasic, D., Ballester, B., Sandelin, A., Lenhard, B., . . . Mathelier, A. (2022). JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. *Nucleic acids research*, *50*(D1), D165-D173. https://doi.org/10.1093/nar/gkab1113

Cha, J.-K., O'Connor, K., Alahmad, S., Lee, J.-H., Dinglasan, E., Park, H., Lee, S.-M., Hirsz, D., Kwon, S.-W., Kwon, Y., Kim, K.-M., Ko, J.-M., Hickey, L. T., Shin, D., & Dixon, L. E. (2022). Speed vernalization to accelerate generation advance in winter cereal crops. *Molecular plant*, *15*(8), 1300-1309. https://doi.org/https://doi.org/10.1016/j.molp.2022.06.012

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. *PLOS Genetics*, *8*(12), e1003134. https://doi.org/10.1371/journal.pgen.1003134

Chen, A., Li, C., Hu, W., Lau, M. Y., Lin, H., Rockwell, N. C., Martin, S. S., Jernstedt, J. A., Lagarias, J. C., & 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*(28), 10037-10044. https://doi.org/10.1073/pnas.1409795111

Chiang, G. C. K., Barua, D., Kramer, E. M., Amasino, R. M., & Donohue, K. (2009). Major flowering time gene, FLOWERING LOCUS C, regulates seed germination in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences*, *106*(28), 11661-11666. https://doi.org/10.1073/pnas.0901367106

Cho, K., Falloon, P., Gornall, J., Betts, R., & Clark, R. (2012). Winter wheat yields in the UK: uncertainties in climate and management impacts. *Climate research*,

54(1), 49-68. https://doi.org/10.3354/cr01085

Choi, J., Hyun, Y., Kang, M. J., In Yun, H., Yun, J. Y., Lister, C., Dean, C., Amasino, R. M., Noh, B., & Noh, Y. S. (2009). Resetting and regulation of FLOWERING LOCUS C expression during Arabidopsis reproductive development. *The Plant Journal*, *57*(5), 918-931.

Chouard, P. (1960). Vernalization and its Relations to Dormancy. *Annual review of plant physiology*, *11*(1), 191-238. https://doi.org/10.1146/annurev.pp.11.060160.001203

Chow, C.-N., Lee, T.-Y., Hung, Y.-C., Li, G.-Z., Tseng, K.-C., Liu, Y.-H., Kuo, P.-L., Zheng, H.-Q., & Chang, W.-C. (2018). PlantPAN3.0: a new and updated resource for reconstructing transcriptional regulatory networks from ChIP-seq experiments in plants. *Nucleic acids research*, *47*(D1), D1155-D1163. https://doi.org/10.1093/nar/gky1081

Cockram, J., Chiapparino, E., Taylor, S. A., Stamati, K., Donini, P., Laurie, D. A., & O'Sullivan D, M. (2007). Haplotype analysis of vernalization loci in European barley germplasm reveals novel VRN-H1 alleles and a predominant winter VRN-H1/VRN-H2 multi-locus haplotype. *Theoretical and applied genetics*, *115*(7), 993-1001. https://doi.org/10.1007/s00122-007-0626-x

Coen, E. S., & Meyerowitz, E. M. (1991). The war of the whorls: genetic interactions controlling flower development. *Nature*, *353*(6339), 31-37. https://doi.org/10.1038/353031a0

Corbesier, L., Vincent, C., Jang, S., Fornara, F., Fan, Q., Searle, I., Giakountis, A., Farrona, S., Gissot, L., Turnbull, C., & Coupland, G. (2007). FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. *Science*, *316*(5827), 1030-1033. https://doi.org/10.1126/science.1141752

Csorba, T., Questa, J. I., Sun, Q., & Dean, C. (2014). Antisense COOLAIR mediates

the coordinated switching of chromatin states at FLC during vernalization. *Proceedings of the National Academy of Sciences - PNAS*, *111*(45), 16160-16165. https://doi.org/10.1073/pnas.1419030111

Danyluk, J., Kane, N. A., Breton, G., Limin, A. E., Fowler, D. B., & Sarhan, F. (2003). TaVRT-1, a Putative Transcription Factor Associated with Vegetative to Reproductive Transition in Cereals. *Plant physiology (Bethesda)*, *132*(4), 1849-1860. https://doi.org/10.1104/pp.103.023523

De Bodt, S., Raes, J., Florquin, K., Rombauts, S., Rouzé, P., Theissen, G., & Van de Peer, Y. (2003). Genomewide structural annotation and evolutionary analysis of the type I MADS-box genes in plants. *Journal of Molecular Evolution*, *56*(5), 573-586. https://doi.org/10.1007/s00239-002-2426-x

de Folter, S., & Angenent, G. C. (2006). trans meets cis in MADS science. *Trends in Plant Science*, *11*(5), 224-231. https://doi.org/10.1016/j.tplants.2006.03.008

De Lucia, F., Crevillen, P., Jones, A. M., Greb, T., & Dean, C. (2008). A PHD-polycomb repressive complex 2 triggers the epigenetic silencing of FLC during vernalization. *Proceedings of the National Academy of Sciences USA*, *105*(44), 16831-16836. https://doi.org/10.1073/pnas.0808687105

Deng, W., Casao, M. C., Wang, P., Sato, K., Hayes, P. M., Finnegan, E. J., & Trevaskis, B. (2015). Direct links between the vernalization response and other key traits of cereal crops. *Nature communications*, *6*(1), 5882-5882. https://doi.org/10.1038/ncomms6882

Dhillon, T., Pearce, S. P., Stockinger, E. J., Distelfeld, A., Li, C., Knox, A. K., Vashegyi, I., VÁgújfalvi, A., Galiba, G., & Dubcovsky, J. (2010). Regulation of Freezing Tolerance and Flowering in Temperate Cereals: The VRN-1 Connection *Plant Physiology*, *153*(4), 1846-1858. https://doi.org/10.1104/pp.110.159079

Díaz, A., Zikhali, M., Turner, A. S., Isaac, P., & Laurie, D. A. (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*(3), e33234-e33234. https://doi.org/10.1371/journal.pone.0033234

Distelfeld, A., Li, C., & Dubcovsky, J. (2009). Regulation of flowering in temperate cereals. *Current Opinion in Plant Biology*, *12*(2), 178-184. https://doi.org/10.1016/j.pbi.2008.12.010

Distelfeld, A., Tranquilli, G., Li, C., Yan, L., & Dubcovsky, J. (2009). Genetic and Molecular Characterization of the VRN2 Loci in Tetraploid Wheat. *Plant physiology (Bethesda)*, *149*(1), 245-257. https://doi.org/10.1104/pp.108.129353

Dixon, L. E., Karsai, I., Kiss, T., Adamski, N. M., Liu, Z., Ding, Y., Allard, V., Boden, S. A., & Griffiths, S. (2019). VERNALIZATION1 controls developmental responses of winter wheat under high ambient temperatures. *Development (Cambridge)*, *146*(3), dev172684. https://doi.org/10.1242/dev.172684

Dorca-Fornell, C., Gregis, V., Grandi, V., Coupland, G., Colombo, L., & Kater, M. M. (2011). The Arabidopsis SOC1-like genes AGL42, AGL71 and AGL72 promote flowering in the shoot apical and axillary meristems. *The Plant Journal*, *67*(6), 1006-1017. https://doi.org/10.1111/j.1365-313X.2011.04653.x

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*, e06620. https://doi.org/10.7554/eLife.06620

Dupont, F. M., & Altenbach, S. B. (2003). Molecular and biochemical impacts of environmental factors on wheat grain development and protein synthesis. *Journal of Cereal Science*, *38*(2), 133-146.

https://doi.org/https://doi.org/10.1016/S0733-5210(03)00030-4

Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC bioinformatics*, *5*(1), 113-113. https://doi.org/10.1186/1471-2105-5-113

Egea-Cortines, M., Saedler, H., & Sommer, H. (1999). Ternary complex formation between the MADS-box proteins SQUAMOSA, DEFICIENS and GLOBOSA is involved in the control of floral architecture in Antirrhinum majus. *The EMBO journal*, *18*(19), 5370-5379.

Erath, W., Bauer, E., Fowler, D. B., Gordillo, A., Korzun, V., Ponomareva, M., Schmidt, M., Schmiedchen, B., Wilde, P., & Schön, C.-C. (2017). Exploring new alleles for frost tolerance in winter rye. *Theoretical and applied genetics*, *130*, 2151-2164.

FAO. (2022). *Crop Prospects and Food Situation – Quarterly Global Report No. 4*.

FAO. (2023). *Crop Prospects and Food Situation – Quarterly Global Report No. 2*.

FAO, I., UNICEF, WFP, WHO. (2023). *The State of Food Security and Nutrition in the World 2023. Urbanization, agrifood systems transformation and healthy diets across the rural–urban continuum.*

Farooq, M., Bramley, H., Palta, J. A., & Siddique, K. H. M. (2011). Heat Stress in Wheat during Reproductive and Grain-Filling Phases. *Critical Reviews in Plant Sciences*, *30*(6), 491-507. https://doi.org/10.1080/07352689.2011.615687

Feldman, M., & Levy, A. A. (2012). Genome Evolution Due to Allopolyploidization in Wheat. *Genetics*, *192*(3), 763-774. https://doi.org/10.1534/genetics.112.146316

Feldman, M., Mello-Sampayo, T., & Avivi, L. (1972). Somatic association of homoeologous chromosomes in Triticum aestivum. *Chromosoma*, *37*(2).

https://doi.org/10.1007/BF00284940

Feng, P., Guo, H., Chi, W., Chai, X., Sun, X., Xu, X., Ma, J., Rochaix, J. D., Leister, D., Wang, H., Lu, C., & Zhang, L. (2016). Chloroplast retrograde signal regulates flowering. *Proceedings of the National Academy of Sciences USA*, *113*(38), 10708-10713. https://doi.org/10.1073/pnas.1521599113

Folter, S. d., & Angenent, G. C. (2006). trans meets cis in MADS science. *Trends in Plant Science*, *11*(5), 224-231. https://doi.org/https://doi.org/10.1016/j.tplants.2006.03.008

Fu, D., Szucs, P., Yan, L., Helguera, M., Skinner, J. S., von Zitzewitz, J., Hayes, P. M., & Dubcovsky, J. (2005). Large deletions within the first intron in VRN-1 are associated with spring growth habit in barley and wheat. *Molecular Genetics and Genomics : MGG*, *273*(1), 54-65. https://doi.org/10.1007/s00438-004-1095-4

Gabler, F., Nam, S.-Z., Till, S., Mirdita, M., Steinegger, M., Söding, J., Lupas, A. N., & Alva, V. (2020). Protein Sequence Analysis Using the MPI Bioinformatics Toolkit. *Current Protocols in Bioinformatics*, *72*(1), e108. https://doi.org/https://doi.org/10.1002/cpbi.108

Gallegos, J. E., & Rose, A. B. (2015). The enduring mystery of intron-mediated enhancement. *Plant Science*, *237*, 8-15. https://doi.org/https://doi.org/10.1016/j.plantsci.2015.04.017

Gallegos, J. E., & Rose, A. B. (2017). Intron DNA Sequences Can Be More Important Than the Proximal Promoter in Determining the Site of Transcript Initiation. *The Plant Cell*, *29*(4), 843-853. https://doi.org/10.1105/tpc.17.00020

Gan, E.-S., Xu, Y., Wong, J.-Y., Goh, J. G., Sun, B., Wee, W.-Y., Huang, J., & Ito, T. (2014). Jumonji demethylases moderate precocious flowering at elevated temperature via regulation of FLC in Arabidopsis. *Nature communications*, *5*(1), 5098-5098. https://doi.org/10.1038/ncomms6098

Gauley, A., & Boden, S. A. (2021). Stepwise increases in FT1 expression regulate seasonal progression of flowering in wheat (Triticum aestivum). *New Phytologist*, *229*(2), 1163-1176. https://doi.org/10.1111/nph.16910

Ge, S. X., Jung, D., & Yao, R. (2020). ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics*, *36*(8), 2628-2629. https://doi.org/10.1093/bioinformatics/btz931

Ghosh, S., Watson, A., Gonzalez-Navarro, O. E., Ramirez-Gonzalez, R. H., Yanes, L., Mendoza-Suárez, M., Simmonds, J., Wells, R., Rayner, T., Green, P., Hafeez, A., Hayta, S., Melton, R. E., Steed, A., Sarkar, A., Carter, J., Perkins, L., Lord, J., Tester, M., . . . Hickey, L. T. (2018). Speed breeding in growth chambers and glasshouses for crop breeding and model plant research. *Nature Protocols*, *13*(12), 2944-2963. https://doi.org/10.1038/s41596-018-0072-z

Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., Pretty, J., Robinson, S., Thomas, S. M., & Toulmin, C. (2010). Food Security: The Challenge of Feeding 9 Billion People. *Science (American Association for the Advancement of Science)*, *327*(5967), 812-818. https://doi.org/10.1126/science.1185383

Gramzow, L., & Theißen, G. (2010). A hitchhiker's guide to the MADS world of plants. *Genome Biology*, *11*(6), 214. https://doi.org/10.1186/gb-2010-11-6-214

Gramzow, L., & Theißen, G. (2013). Phylogenomics of MADS-Box Genes in Plants — Two Opposing Life Styles in One Gene Family. *Biology*, *2*(3), 1150-1164.

Greenup, A. G., Sasani, S., Oliver, S. N., Talbot, M. J., Dennis, E. S., Hemming, M. N., & Trevaskis, B. (2010). ODDSOC2 Is a MADS Box Floral Repressor That Is Down-Regulated by Vernalization in Temperate Cereals. *Plant physiology (Bethesda)*, *153*(3), 1062-1073. https://doi.org/10.1104/pp.109.152488

Greenup, A. G., Sasani, S., Oliver, S. N., Walford, S. A., Millar, A. A., & Trevaskis, B.

(2011). Transcriptome analysis of the vernalization response in barley (Hordeum vulgare) seedlings. *PloS one*, *6*(3), e17900-e17900. https://doi.org/10.1371/journal.pone.0017900

Gupta, S., Stamatoyannopoulos, J. A., Bailey, T. L., & Noble, W. S. (2007). Quantifying similarity between motifs. *Genome Biology*, *8*(2), R24. https://doi.org/10.1186/gb-2007-8-2-r24

Hackauf, B., Siekmann, D., & Fromme, F. J. (2022). Improving Yield and Yield Stability in Winter Rye by Hybrid Breeding. *Plants*, *11*(19).

Hahn, F., Sanjurjo Loures, L., Sparks, C. A., Kanyuka, K., & Nekrasov, V. (2021). Efficient CRISPR/Cas-Mediated Targeted Mutagenesis in Spring and Winter Wheat Varieties. *Plants (Basel)*, *10*(7). https://doi.org/10.3390/plants10071481

Halliwell, J., Borrill, P., Gordon, A., Kowalczyk, R., Pagano, M. L., Saccomanno, B., Bentley, A. R., Uauy, C., & Cockram, J. (2016). Systematic Investigation of FLOWERING LOCUS T-Like Poaceae Gene Families Identifies the Short-Day Expressed Flowering Pathway Gene, TaFT3 in Wheat (Triticum aestivum L.) [Original Research]. *Frontiers in plant science*, *7*. https://doi.org/10.3389/fpls.2016.00857

Harmer, S. L., Hogenesch, J. B., Straume, M., Chang, H.-S., Han, B., Zhu, T., Wang, X., Kreps, J. A., & Kay, S. A. (2000). Orchestrated Transcription of Key Pathways in Arabidopsis by the Circadian Clock. *Science (American Association for the Advancement of Science)*, *290*(5499), 2110-2113. https://doi.org/10.1126/science.290.5499.2110

Harmer, S. L., & Kay, S. A. (2005). Positive and Negative Factors Confer Phase-Specific Circadian Regulation of Transcription in Arabidopsis. *The Plant Cell*, *17*(7), 1926-1940. https://doi.org/10.1105/tpc.105.033035

Hedden, P. (2003). The genes of the Green Revolution. *Trends in Genetics*, *19*(1), 5-

9. https://doi.org/https://doi.org/10.1016/S0168-9525(02)00009-4

Helliwell, C. A., Wood, C. C., Robertson, M., James Peacock, W., & Dennis, E. S. (2006). The Arabidopsis FLC protein interacts directly in vivo with SOC1 and FT chromatin and is part of a high-molecular-weight protein complex. *The Plant journal : for cell and molecular biology*, *46*(2), 183-192. https://doi.org/10.1111/j.1365-313X.2006.02686.x

Hemming, M. N., Walford, S. A., Fieg, S., Dennis, E. S., & Trevaskis, B. (2012). Identification of High-Temperature-Responsive Genes in Cereals. *Plant physiology (Bethesda)*, *158*(3), 1439-1450. https://doi.org/10.1104/pp.111.192013

Heo, J. B., & Sung, S. (2011). Vernalization-Mediated Epigenetic Silencing by a Long Intronic Noncoding RNA. *Science (American Association for the Advancement of Science)*, *331*(6013), 76-79. https://doi.org/10.1126/science.1197349

Hepworth, J., Antoniou-Kourounioti, R. L., Berggren, K., Selga, C., Tudor, E. H., Yates, B., Cox, D., Collier Harris, B. R., Irwin, J. A., Howard, M., Säll, T., Holm, S., & Dean, C. (2020). Natural variation in autumn expression is the major adaptive determinant distinguishing Arabidopsis FLC haplotypes. *Elife*, *9*. https://doi.org/10.7554/eLife.57671

Hepworth, J., Antoniou-Kourounioti, R. L., Bloomer, R. H., Selga, C., Berggren, K., Cox, D., Collier Harris, B. R., Irwin, J. A., Holm, S., Säll, T., Howard, M., & Dean, C. (2018). Absence of warmth permits epigenetic memory of winter in Arabidopsis. *Nature communications*, *9*(1), 639-638. https://doi.org/10.1038/s41467-018-03065-7

Hurkman, W. J., McCue, K. F., Altenbach, S. B., Korn, A., Tanaka, C. K., Kothari, K. M., Johnson, E. L., Bechtel, D. B., Wilson, J. D., Anderson, O. D., & DuPont, F. M. (2003). Effect of temperature on expression of genes encoding enzymes for starch biosynthesis in developing wheat endosperm. *Plant Science*,

164(5), 873-881. https://doi.org/https://doi.org/10.1016/S0168-9452(03)00076-1

Hyun, Y., Richter, R., Vincent, C., Martinez-Gallegos, R., Porri, A., & Coupland, G. (2016). Multi-layered regulation of SPL15 and cooperation with SOC1 integrate endogenous flowering pathways at the Arabidopsis shoot meristem. *Developmental cell*, *37*(3), 254-266.

Hyun, Y., Vincent, C., Tilmes, V., Bergonzi, S., Kiefer, C., Richter, R., Martinez-Gallegos, R., Severing, E., & Coupland, G. (2019). A regulatory circuit conferring varied flowering response to cold in annual and perennial plants. *Science*, *363*(6425), 409-412. https://doi.org/10.1126/science.aau8197

Jiao, F., Pahwa, K., Manning, M., Dochy, N., & Geuten, K. (2019). Cold Induced Antisense Transcription of FLOWERING LOCUS C in Distant Grasses. *Frontiers in plant science*, *10*, 72-72. https://doi.org/10.3389/fpls.2019.00072

Jones, L., Gorst, A., Elliott, J., Fitch, A., Illman, H., Evans, C., Thackeray, S., Spears, B., Gunn, I., & Carvalho, L. (2020). Climate driven threshold effects in the natural environment. *Report to the UK Climate Change Committee*.

Jung, J.-H., Park, J.-H., Lee, S., To, T. K., Kim, J.-M., Seki, M., & Park, C.-M. (2013). The Cold Signaling Attenuator HIGH EXPRESSION OF OSMOTICALLY RESPONSIVE GENE1 Activates FLOWERING LOCUS C Transcription via Chromatin Remodeling under Short-Term Cold Stress in Arabidopsis. *The Plant Cell*, *25*(11), 4378-4390. https://doi.org/10.1105/tpc.113.118364

Kane, N., Danyluk, J., Guylaine, T., Ouellet, F., Laliberté, J.-F., Allen, E. L., Fowler, D. B., & Sarhan, F. (2005). TaVRT-2, a Member of the StMADS-11 Clade of Flowering Repressors, Is Regulated by Vernalization and Photoperiod in Wheat. *Plant Physiology*, *138*(4), 2354-2363. http://www.jstor.org/stable/4630028

Kapazoglou, A., Engineer, C., Drosou, V., Kalloniati, C., Tani, E., Tsaballa, A., Kouri,

E. D., Ganopoulos, I., Flemetakis, E., & Tsaftaris, A. S. (2012). The study of two barley Type I-like MADS-boxgenes as potential targets of epigenetic regulation during seed development. *BMC plant biology*, *12*(1). https://doi.org/10.1186/1471-2229-12-166

Kaufmann, K., Melzer, R., & Theissen, G. (2005). MIKC-type MADS-domain proteins: structural modularity, protein interactions and network evolution in land plants. *Gene*, *347*(2), 183-198. https://doi.org/10.1016/j.gene.2004.12.014

Kendon, M., Sexton, D., & McCarthy, M. (2020). A temperature of 20°C in the UK winter: a sign of the future? *Weather*, *75*(10), 318-324. https://doi.org/10.1002/wea.3811

Kennedy, A., & Geuten, K. (2020). The Role of FLOWERING LOCUS C Relatives in Cereals. *Frontiers in plant science*, *11*, 617340-617340. https://doi.org/10.3389/fpls.2020.617340

Khokhar, K. M. (2008). Effect of temperature and photoperiod on the incidence of bulbing and bolting in seedlings of onion cultivars of diverse origin. *The Journal of Horticultural Science and Biotechnology*, *83*(4), 488-496. https://doi.org/10.1080/14620316.2008.11512412

Kim, D.-H., & Sung, S. (2017). Vernalization-Triggered Intragenic Chromatin Loop Formation by Long Noncoding RNAs. *Developmental cell*, *40*(3), 302-312.e304. https://doi.org/10.1016/j.devcel.2016.12.021

Kim, D.-H., Xi, Y., & Sung, S. (2017). Modular function of long noncoding RNA, COLDAIR, in the vernalization response. *PLOS Genetics*, *13*(7), e1006939-e1006939. https://doi.org/10.1371/journal.pgen.1006939

Kippes, N., Chen, A., Zhang, X., Lukaszewski, A. J., & Dubcovsky, J. (2016). Development and characterization of a spring hexaploid wheat line with no functional VRN2 genes. *Theoretical and applied genetics*, *129*(7), 1417-1428.

https://doi.org/10.1007/s00122-016-2713-3

Kippes, N., Debernardi, J. M., Vasquez-Gross, H. A., Akpinar, B. A., Budak, H., Kato, K., Chao, S., Akhunov, E., & Dubcovsky, J. (2015). Identification of the VERNALIZATION 4 gene reveals the origin of spring growth habit in ancient wheats from South Asia. *Proceedings of the National Academy of Sciences*, *112*(39), E5401-E5410. https://doi.org/10.1073/pnas.1514883112

Kippes, N., Zhu, J., Chen, A., Vanzetti, L., Lukaszewski, A., Nishida, H., Kato, K., Dvorak, J., & Dubcovsky, J. (2014). Fine mapping and epistatic interactions of the vernalization gene VRN-D4 in hexaploid wheat. *Molecular Genetics and Genomics*, *289*(1), 47-62. https://doi.org/10.1007/s00438-013-0788-y

Kirby, E. J. M., & Appleyard, M. (1987). Development and structure of the wheat plant. In F. G. H. Lupton (Ed.), *Wheat Breeding: Its scientific basis* (pp. 287-311). Springer Netherlands. https://doi.org/10.1007/978-94-009-3131-2_10

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*(5446), 1960-1962. https://doi.org/10.1126/science.286.5446.1960

Kobayashi, Y., & Weigel, D. (2007). Move on up, it's time for change--mobile signals controlling photoperiod-dependent flowering. *Genes & development*, *21*(19), 2371-2384. https://doi.org/10.1101/gad.1589007

Kumar, L., & Futschik, M. (2007). Mfuzz: a software package for soft clustering of microarray data. *Bioinformation*, *2*(1), 5-7. https://doi.org/10.6026/97320630002005

Labuschagne, M. T., Elago, O., & Koen, E. (2009). The influence of temperature extremes on some quality and starch characteristics in bread, biscuit and durum wheat. *Journal of Cereal Science*, *49*(2), 184-189. https://doi.org/https://doi.org/10.1016/j.jcs.2008.09.001

Laudencia-Chingcuanco, D., Ganeshan, S., You, F., Fowler, B., Chibbar, R., &

Anderson, O. (2011). Genome-wide gene expression analysis supports a developmental model of low temperature tolerance gene regulation in wheat (Triticum aestivum L.). *BMC Genomics*, *12*(1), 299. https://doi.org/10.1186/1471-2164-12-299

Le Gouis, J., Oury, F.-X., & Charmet, G. (2020). How changes in climate and agricultural practices influenced wheat production in Western Europe. *Journal of Cereal Science*, *93*, 102960. https://doi.org/https://doi.org/10.1016/j.jcs.2020.102960

Leach, L. J., Belfield, E. J., Jiang, C., Brown, C., Mithani, A., & Harberd, N. P. (2014). Patterns of homoeologous gene expression shown by RNA sequencing in hexaploid bread wheat. *BMC Genomics*, *15*(1), 276. https://doi.org/10.1186/1471-2164-15-276

Lee, H., Suh, S. S., Park, E., Cho, E., Ahn, J. H., Kim, S. G., Lee, J. S., Kwon, Y. M., & Lee, I. (2000). The AGAMOUS-LIKE 20 MADS domain protein integrates floral inductive pathways in Arabidopsis. *Genes & development*, *14*(18), 2366-2376. https://doi.org/10.1101/gad.813600

Lee, J., & Lee, I. (2010). Regulation and function of SOC1, a flowering pathway integrator. *Journal of Experimental Botany*, *61*(9), 2247-2254. https://doi.org/10.1093/jxb/erq098

Lee, J. H., Yoo, S. J., Park, S. H., Hwang, I., Lee, J. S., & Ahn, J. H. (2007). Role of SVP in the control of flowering time by ambient temperature in Arabidopsis. *Genes & development*, *21*(4), 397-402. https://doi.org/10.1101/gad.1518407

Levy, A. A., & Feldman, M. (2022). Evolution and origin of bread wheat. *Plant Cell*, *34*(7), 2549-2567. https://doi.org/10.1093/plcell/koac130

Li, C., & Dubcovsky, J. (2008). Wheat FT protein regulates VRN1 transcription through interactions with FDL2. *The Plant Journal*, *55*(4), 543-554. https://doi.org/10.1111/j.1365-313X.2008.03526.x

Li, C., Lin, H., & Dubcovsky, J. (2015). Factorial combinations of protein interactions generate a multiplicity of florigen activation complexes in wheat and barley. *The Plant Journal*, *84*(1), 70-82. https://doi.org/10.1111/tpj.12960

Li, K., Debernardi, J. M., Li, C., Lin, H., Zhang, C., Jernstedt, J., Korff, M. v., Zhong, J., & Dubcovsky, J. (2021). Interactions between SQUAMOSA and SHORT VEGETATIVE PHASE MADS-box proteins regulate meristem transitions during wheat spike development. *The Plant Cell*, *33*(12), 3621-3644. https://doi.org/10.1093/plcell/koab243

Li, Q., Peng, A., Yang, J., Zheng, S., Li, Z., Mu, Y., Chen, L., Si, J., Ren, X., & Song, H. (2022). A 215-bp indel at intron I of BoFLC2 affects flowering time in Brassica oleracea var. capitata during vernalization. *Theoretical and applied genetics*, *135*(8), 2785-2797. https://doi.org/10.1007/s00122-022-04149-1

Lin, S. I., Wang, J. G., Poon, S. Y., Su, C. L., Wang, S. S., & Chiou, T. J. (2005). Differential regulation of FLOWERING LOCUS C expression by vernalization in cabbage and Arabidopsis. *Plant Physiology*, *137*(3), 1037-1048. https://doi.org/10.1104/pp.104.058974

Liu, C., Chen, H., Er, H. L., Soo, H. M., Kumar, P. P., Han, J. H., Liou, Y. C., & Yu, H. (2008). Direct interaction of AGL24 and SOC1 integrates flowering signals in Arabidopsis. *Development*, *135*(8), 1481-1491. https://doi.org/10.1242/dev.020255

Liu, F., Marquardt, S., Lister, C., Swiezewski, S., & Dean, C. (2010). Targeted 3' Processing of Antisense Transcripts Triggers Arabidopsis FLC Chromatin Silencing. *Science (American Association for the Advancement of Science)*, *327*(5961), 94-97. https://doi.org/10.1126/science.1180278

Liu, J., Chen, Z., Wang, Z., Zhang, Z., Xie, X., Wang, Z., Chai, L., Song, L., Cheng, X., Feng, M., Wang, X., Liu, Y., Hu, Z., Xing, J., Su, Z., Peng, H., Xin, M., Yao, Y., Guo, W., . . . Ni, Z. (2021). Ectopic expression of VRT-A2 underlies the origin of Triticum polonicum and Triticum petropavlovskyi with long outer

glumes and grains. *Molecular plant*, *14*(9), 1472-1488. https://doi.org/https://doi.org/10.1016/j.molp.2021.05.021

Lu, X., O'Neill, C. M., Warner, S., Xiong, Q., Chen, X., Wells, R., & Penfield, S. (2022). Winter warming post floral initiation delays flowering via bud dormancy activation and affects yield in a winter annual crop. *Proceedings of the National Academy of Sciences - PNAS*, *119*(39), 1-e2204355119. https://doi.org/10.1073/pnas.2204355119

Luo, M.-C., Gu, Y. Q., Puiu, D., Wang, H., Twardziok, S. O., Deal, K. R., Huo, N., Zhu, T., Wang, L., Wang, Y., McGuire, P. E., Liu, S., Long, H., Ramasamy, R. K., Rodriguez, J. C., Van, S. L., Yuan, L., Wang, Z., Xia, Z., . . . Dvořák, J. (2017). Genome sequence of the progenitor of the wheat D genome Aegilops tauschii. *Nature (London)*, *551*(7681), 498-502. https://doi.org/10.1038/nature24486

Ma, S., Wang, M., Wu, J., Guo, W., Chen, Y., Li, G., Wang, Y., Shi, W., Xia, G., Fu, D., Kang, Z., & Ni, F. (2021). WheatOmics: A platform combining multiple omics data to accelerate functional genomics studies in wheat. *Molecular plant*, *14*(12), 1965-1968. https://doi.org/10.1016/j.molp.2021.10.006

Magome, H., Yamaguchi, S., Hanada, A., Kamiya, Y., & Oda, K. (2004). dwarf and delayed-flowering 1, a novel Arabidopsis mutant deficient in gibberellin biosynthesis because of overexpression of a putative AP2 transcription factor. *The Plant Journal*, *37*(5), 720-729. https://doi.org/https://doi.org/10.1111/j.1365-313X.2003.01998.x

Marcińska, I., Czyczyło-Mysza, I., Skrzypek, E., Warchoł, M., Zieliński, K., & Dubas, E. (2018). Obtaining of winter rye (Secale cereale L. ssp. cereale) haploid embryos through hybridization with maize (Zea mays L.). *Cereal Research Communications*, *46*(3), 521-532.

Marcussen, T., Sandve, S. R., Heier, L., Spannagl, M., Pfeifer, M., Jakobsen, K. S., Wulff, B. B. H., Steuernagel, B., Mayer, K. F. X., & Olsen, O.-A. (2014).

Ancient hybridizations among the ancestral genomes of bread wheat. *Science (American Association for the Advancement of Science)*, *345*(6194), 1250092-1250092. https://doi.org/10.1126/science.1250092

Marks, M. K., & Prince, S. D. (1979). INDUCTION OF FLOWERING IN WILD LETTUCE (LACTUCA SERRIOLA L.). II. DEVERNALIZATION. *New Phytologist*, *82*(2), 357-363. https://doi.org/https://doi.org/10.1111/j.1469-8137.1979.tb02661.x

Marquardt, S., Raitskin, O., Wu, Z., Liu, F., Sun, Q., & Dean, C. (2014). Functional Consequences of Splicing of the Antisense Transcript COOLAIR on FLC Transcription. *Molecular cell*, *54*(1), 156-165. https://doi.org/10.1016/j.molcel.2014.03.026

Masiero, S., Imbriano, C., Ravasio, F., Favaro, R., Pelucchi, N., Gorla, M. S., Mantovani, R., Colombo, L., & Kater, M. M. (2002). Ternary complex formation between MADS-box transcription factors and the histone fold protein NF-YB. *Journal of Biological Chemistry*, *277*(29), 26429-26435.

Matar, S., Kumar, A., Holtgräwe, D., Weisshaar, B., & Melzer, S. (2021). The transition to flowering in winter rapeseed during vernalization. *Plant, cell & environment*, *44*(2), 506-518. https://doi.org/https://doi.org/10.1111/pce.13946

Mateos, J. L., Tilmes, V., Madrigal, P., Severing, E., Richter, R., Rijkenberg, C. W. M., Krajewski, P., & Coupland, G. (2017). Divergence of regulatory networks governed by the orthologous transcription factors FLC and PEP1 in Brassicaceae species. *Proceedings of the National Academy of Sciences USA*, *114*(51), E11037-e11046. https://doi.org/10.1073/pnas.1618075114

Mathews, S., & Sharrock, R. (1997). Phytochrome gene diversity. *Plant, cell & environment*, *20*(6), 666-671.

McKeown, M., Schubert, M., Marcussen, T., Fjellheim, S., & Preston, J. C. (2016). Evidence for an Early Origin of Vernalization Responsiveness in Temperate

Pooideae Grasses *Plant Physiology*, *172*(1), 416-426. https://doi.org/10.1104/pp.16.01023

Messenguy, F., & Dubois, E. (2003). Role of MADS box proteins and their cofactors in combinatorial control of gene expression and cell development. *Gene*, *316*, 1-21. https://doi.org/https://doi.org/10.1016/S0378-1119(03)00747-9

Michael, T. P., & McClung, C. R. (2002). Phase-specific circadian clock regulatory elements in Arabidopsis. *Plant Physiology*, *130*(2), 627-638. https://doi.org/10.1104/pp.004929

Michaels, S. D., & Amasino, R. M. (1999). FLOWERING LOCUS C Encodes a Novel MADS Domain Protein That Acts as a Repressor of Flowering. *The Plant Cell*, *11*(5), 949-956. https://doi.org/10.1105/tpc.11.5.949

Michaels, S. D., & Amasino, R. M. (2000). Memories of winter: vernalization and the competence to flower. *Plant, cell & environment*, *23*(11), 1145-1153. https://doi.org/https://doi.org/10.1046/j.1365-3040.2000.00643.x

Miedaner, T., Korzun, V., & Wilde, P. (2022). Effective Pollen-Fertility Restoration Is the Basis of Hybrid Rye Production and Ergot Mitigation. *Plants*, *11*(9).

Miedaner, T., & Laidig, F. (2019). Hybrid Breeding in Rye (Secale cereale L.). In J. M. Al-Khayri, S. M. Jain, & D. V. Johnson (Eds.), *Advances in Plant Breeding Strategies: Cereals: Volume 5* (pp. 343-372). Springer International Publishing. https://doi.org/10.1007/978-3-030-23108-8_9

Mikkelsen, M. D., & Thomashow, M. F. (2009). A role for circadian evening elements in cold-regulated gene expression in Arabidopsis. *The Plant Journal*, *60*(2), 328-339. https://doi.org/https://doi.org/10.1111/j.1365-313X.2009.03957.x

Monteagudo, A., Igartua, E., Contreras-Moreira, B., Gracia, M. P., Ramos, J., Karsai, I., & Casas, A. M. (2019). Fine-tuning of the flowering time control in winter barley: the importance of HvOS2 and HvVRN2 in non-inductive conditions. *BMC plant biology*, *19*(1), 113-113. https://doi.org/10.1186/s12870-019-1727-

9

Muehlbauer, F. J., Burnell, D. G., Bogyo, T. P., & Bogyo, M. T. (1981). Simulated Comparisons of Single Seed Descent and Bulk Population Breeding Methods1. *Crop Science*, *21*(4), cropsci1981.0011183X002100040025x. https://doi.org/https://doi.org/10.2135/cropsci1981.0011183X002100040025x

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*(5), 686-698. https://doi.org/10.1093/pcp/pci086

Nguyen, H. H., Park, J., Kang, S., & Kim, M. (2015). Surface plasmon resonance: a versatile technique for biosensor applications. *Sensors (Basel)*, *15*(5), 10481-10510. https://doi.org/10.3390/s150510481

Nishida, H., Yoshida, T., Kawakami, K., Fujita, M., Long, B., Akashi, Y., Laurie, D. A., & Kato, K. (2013). Structural variation in the 5′ upstream region of photoperiod-insensitive alleles Ppd-A1a and Ppd-B1a identified in hexaploid wheat (Triticum aestivum L.), and their effect on heading time. *Molecular Breeding*, *31*(1), 27-37. https://doi.org/10.1007/s11032-012-9765-0

Nurk, S., Koren, S., Rhie, A., Rautiainen, M., Bzikadze, A. V., Mikheenko, A., Vollger, M. R., Altemose, N., Uralsky, L., Gershman, A., Aganezov, S., Hoyt, S. J., Diekhans, M., Logsdon, G. A., Alonge, M., Antonarakis, S. E., Borchers, M., Bouffard, G. G., Brooks, S. Y., . . . Phillippy, A. M. (2022). The complete sequence of a human genome. *Science*, *376*(6588), 44-53. https://doi.org/10.1126/science.abj6987

Nusinow, D. A., Helfer, A., Hamilton, E. E., King, J. J., Imaizumi, T., Schultz, T. F., Farré, E. M., & Kay, S. A. (2011). The ELF4–ELF3–LUX complex links the circadian clock to diurnal control of hypocotyl growth. *Nature*, *475*(7356), 398-402. https://doi.org/10.1038/nature10182

O'Neill, C. M., Lu, X., Calderwood, A., Tudor, E. H., 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*(24), 4300-4306.e4302. https://doi.org/10.1016/j.cub.2019.10.051

OECD, Food, & Nations, A. O. o. t. U. (2022). *OECD-FAO Agricultural Outlook 2022-2031*. https://doi.org/doi:https://doi.org/10.1787/f1b0b29c-en

Oliver, S. N., Finnegan, E. J., Dennis, E. S., Peacock, W. J., & Trevaskis, B. (2009). Vernalization-induced flowering in cereals is associated with changes in histone methylation at the VERNALIZATION1 gene. *Proceedings of the National Academy of Sciences - PNAS*, *106*(20), 8386-8391. https://doi.org/10.1073/pnas.0903566106

Palosuo, T., Kersebaum, K. C., Angulo, C., Hlavinka, P., Moriondo, M., Olesen, J. E., Patil, R. H., Ruget, F., Rumbaur, C., & Takáč, J. (2011). Simulation of winter wheat yield and its variability in different climates of Europe: A comparison of eight crop growth models. *European Journal of Agronomy*, *35*(3), 103-114.

Para, A., Farré, E. M., Imaizumi, T., Pruneda-Paz, J. L., Harmon, F. G., & Kay, S. A. (2007). PRR3 Is a vascular regulator of TOC1 stability in the Arabidopsis circadian clock. *The Plant Cell*, *19*(11), 3462-3473. https://doi.org/10.1105/tpc.107.054775

Pařenicová, L., de Folter, S., Kieffer, M., Horner, D. S., Favalli, C., Busscher, J., Cook, H. E., Ingram, R. M., Kater, M. M., Davies, B., Angenent, G. C., & Colombo, L. (2003). Molecular and Phylogenetic Analyses of the Complete MADS-Box Transcription Factor Family in Arabidopsis : New Openings to the MADS World[W]. *The Plant Cell*, *15*(7), 1538-1551. https://doi.org/10.1105/tpc.011544

Park, P. J. (2009). ChIP–seq: advantages and challenges of a maturing technology. *Nature Reviews Genetics*, *10*(10), 669-680. https://doi.org/10.1038/nrg2641

Paten, B., Earl, D., Nguyen, N., Diekhans, M., Zerbino, D., & Haussler, D. (2011). Cactus: Algorithms for genome multiple sequence alignment. *Genome research*, *21*(9), 1512-1528.

Paterson, A. H., Brubaker, C. L., & Wendel, J. F. (1993). A rapid method for extraction of cotton (Gossypium spp.) genomic DNA suitable for RFLP or PCR analysis. *Plant Molecular Biology Reporter*, *11*(2), 122-127. https://doi.org/10.1007/BF02670470

Pearce, S., Kippes, N., Chen, A., Debernardi, J. M., & Dubcovsky, J. (2016). RNA-seq studies using wheat PHYTOCHROME B and PHYTOCHROME C mutants reveal shared and specific functions in the regulation of flowering and shade-avoidance pathways. *BMC plant biology*, *16*(1), 141. https://doi.org/10.1186/s12870-016-0831-3

Périlleux, C., Pieltain, A., Jacquemin, G., Bouché, F., Detry, N., D'Aloia, M., Thiry, L., Aljochim, P., Delansnay, M., Mathieu, A.-S., Lutts, S., & Tocquin, P. (2013). A root chicory MADS box sequence and the Arabidopsis flowering repressor FLC share common features that suggest conserved function in vernalization and de-vernalization responses. *The Plant Journal*, *75*(3), 390-402. https://doi.org/https://doi.org/10.1111/tpj.12208

Pieper, R., Tomé, F., Pankin, A., & von Korff, M. (2021). FLOWERING LOCUS T4 delays flowering and decreases floret fertility in barley. *Journal of Experimental Botany*, *72*(1), 107-121. https://doi.org/10.1093/jxb/eraa466

Pimentel, H., Bray, N. L., Puente, S., Melsted, P., & Pachter, L. (2017). Differential analysis of RNA-seq incorporating quantification uncertainty. *Nature Methods*, *14*(7), 687-690. https://doi.org/10.1038/nmeth.4324

Purvis, O. N., & Gregory, F. G. (1945). Devernalization by High Temperature. *Nature*, *155*(3926), 113-114. https://doi.org/10.1038/155113a0

Qüesta, J. I., Song, J., Geraldo, N., An, H., & Dean, C. (2016). Arabidopsis

transcriptional repressor VAL1 triggers Polycomb silencing at FLC during vernalization. *Science (American Association for the Advancement of Science)*, *353*(6298), 485-488. https://doi.org/10.1126/science.aaf7354

Raza, Q., Riaz, A., Atif, R. M., Hussain, B., Rana, I. A., Ali, Z., Budak, H., & Alaraidh, I. A. (2021). Genome-Wide Diversity of MADS-Box Genes in Bread Wheat is Associated with its Rapid Global Adaptability. *Frontiers in Genetics*, *12*, 818880-818880. https://doi.org/10.3389/fgene.2021.818880

Riechmann, J. L., Heard, J., Martin, G., Reuber, L., Jiang, C. Z., Keddie, J., Adam, L., Pineda, O., Ratcliffe, O. J., Samaha, R. R., Creelman, R., Pilgrim, M., Broun, P., Zhang, J. Z., Ghandehari, D., Sherman, B. K., & L. Yu, G. (2000). Arabidopsis Transcription Factors: Genome-Wide Comparative Analysis Among Eukaryotes. *Science*, *290*(5499), 2105-2110. https://doi.org/10.1126/science.290.5499.2105

Riechmann, J. L., Krizek, B. A., & Meyerowitz, E. M. (1996). Dimerization specificity of Arabidopsis MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA, and AGAMOUS. *Proceedings of the National Academy of Sciences USA*, *93*(10), 4793-4798. https://doi.org/10.1073/pnas.93.10.4793

Riechmann, J. L., & Meyerowitz, E. M. (1997). MADS domain proteins in plant development. *Journal of Biological Chemistry*, *378*(10), 1079-1101.

Rizza, F., Karsai, I., Morcia, C., Badeck, F.-W., Terzi, V., Pagani, D., Kiss, T., & Stanca, A. M. (2016). Association between the allele compositions of major plant developmental genes and frost tolerance in barley (Hordeum vulgare L.) germplasm of different origin. *Molecular Breeding*, *36*(11), 156. https://doi.org/10.1007/s11032-016-0571-y

Rosa, S., Duncan, S., & Dean, C. (2016). Mutually exclusive sense–antisense transcription at FLC facilitates environmentally induced gene repression. *Nature communications*, *7*(1), 13031. https://doi.org/10.1038/ncomms13031

Rose, A. B. (2008). Intron-Mediated Regulation of Gene Expression. In A. S. N. Reddy & M. Golovkin (Eds.), *Nuclear pre-mRNA Processing in Plants* (pp. 277-290). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-76776-3_15

Ruelens, P., Maagd, d. R. A., Proost, S., Theissen, G., Geuten, K., & Kaufmann, K. (2013). FLOWERING LOCUS C in monocots and the tandem origin of angiosperm specific MADS-box genes. *Nature communications*, *4*(1), 2280-2280. https://doi.org/10.1038/ncomms3280

Sasani, S., Hemming, M. N., Oliver, S. N., Greenup, A., Tavakkol-Afshari, R., Mahfoozi, S., Poustini, K., Sharifi, H.-R., Dennis, E. S., Peacock, W. J., & Trevaskis, B. (2009). Influence of vernalization and daylength on expression of flowering-time genes in the shoot apex and leaves of barley (Hordeum vulgare). *Journal of Experimental Botany*, *60*(7), 2169-2178. https://doi.org/10.1093/jxb/erp098

Schiessl, S. V., Quezada-Martinez, D., Tebartz, E., Snowdon, R. J., & Qian, L. (2019). The vernalisation regulator FLOWERING LOCUS C is differentially expressed in biennial and annual Brassica napus. *Scientific Reports*, *9*(1), 14911. https://doi.org/10.1038/s41598-019-51212-x

Schilling, S., Kennedy, A., Pan, S., Jermiin, L. S., & Melzer, R. (2020). Genome-wide analysis of MIKC-type MADS-box genes in wheat: pervasive duplications, functional conservation and putative neofunctionalization. *New Phytologist*, *225*(1), 511-529. https://doi.org/10.1111/nph.16122

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, *9*(7), 671-675. https://doi.org/10.1038/nmeth.2089

Searle, I., He, Y., Turck, F., Vincent, C., Fornara, F., Kröber, S., Amasino, R. A., & Coupland, G. (2006). The transcription factor FLC confers a flowering response to vernalization by repressing meristem competence and systemic

signaling in Arabidopsis. *Genes & development*, *20*(7), 898-912. https://doi.org/10.1101/gad.373506

Seo, P. J., Park, M.-J., Lim, M.-H., Kim, S.-G., Lee, M., Baldwin, I. T., & Park, C.-M. (2012). A Self-Regulatory Circuit of CIRCADIAN CLOCK-ASSOCIATED1 Underlies the Circadian Clock Regulation of Temperature Responses in Arabidopsis *The Plant Cell*, *24*(6), 2427-2442. https://doi.org/10.1105/tpc.112.098723

Sharma, N., Ruelens, P., D'Hauw, M., Maggen, T., Dochy, N., Torfs, S., Kaufmann, K., Rohde, A., & Geuten, K. (2017). A Flowering Locus C Homolog Is a Vernalization-Regulated Repressor in Brachypodium and Is Cold Regulated in Wheat. *Plant physiology (Bethesda)*, *173*(2), 1301-1315. https://doi.org/10.1104/pp.16.01161

Shaul, O. (2017). How introns enhance gene expression. *The International Journal of Biochemistry & Cell Biology*, *91*, 145-155. https://doi.org/https://doi.org/10.1016/j.biocel.2017.06.016

Shaw, L. M., 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*(1), 193-204. https://doi.org/10.1093/jxb/ery350

Shaw, L. M., Turner, A. S., & Laurie, D. A. (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*(1), 71-84. https://doi.org/https://doi.org/10.1111/j.1365-313X.2012.04971.x

Sheehan, H., & Bentley, A. (2021). Changing times: Opportunities for altering winter wheat phenology. *Plants, people, planet*, *3*(2), 113-123. https://doi.org/10.1002/ppp3.10163

Sheldon, C. C., Burn, J. E., Perez, P. P., Metzger, J., Edwards, J. A., Peacock, W. J.,

& Dennis, E. S. (1999). FLF MADS box gene: a repressor of flowering in Arabidopsis regulated by vernalization and methylation. *The Plant Cell*, *11*(3), 445-458. https://doi.org/10.1105/tpc.11.3.445

Sheldon, C. C., Hills, M. J., Lister, C., Dean, C., Dennis, E. S., & Peacock, W. J. (2008). Resetting of FLOWERING LOCUS C expression after epigenetic repression by vernalization. *Proceedings of the National Academy of Sciences USA*, *105*(6), 2214-2219. https://doi.org/10.1073/pnas.0711453105

Sheldon, C. C., Rouse, D. T., Finnegan, E. J., Peacock, W. J., & Dennis, E. S. (2000). The Molecular Basis of Vernalization: The Central Role of FLOWERING LOCUS C (FLC). *Proceedings of the National Academy of Sciences - PNAS*, *97*(7), 3753-3758. https://doi.org/10.1073/pnas.060023597

Shimada, S., Ogawa, T., Kitagawa, S., Suzuki, T., Ikari, C., Shitsukawa, N., Abe, T., Kawahigashi, H., Kikuchi, R., Handa, H., & Murai, K. (2009). A genetic network of flowering-time genes in wheat leaves, in which an APETALA1/FRUITFULL-like gene, VRN1, is upstream of FLOWERING LOCUS T. *The Plant Journal*, *58*(4), 668-681. https://doi.org/https://doi.org/10.1111/j.1365-313X.2009.03806.x

Shindo, C., Lister, C., Crevillen, P., Nordborg, M., & Dean, C. (2006). Variation in the epigenetic silencing of FLC contributes to natural variation in Arabidopsis vernalization response. *Genes & development*, *20*(22), 3079-3083. https://doi.org/10.1101/gad.405306

Shore, P., & Sharrocks, A. D. (1995). The MADS-box family of transcription factors. *European Journal of Biochemistry*, *229*(1), 1-13. https://doi.org/10.1111/j.1432-1033.1995.tb20430.x

Shrestha, R., Gómez-Ariza, J., Brambilla, V., & Fornara, F. (2014). Molecular control of seasonal flowering in rice, arabidopsis and temperate cereals. *Annals of Botany*, *114*(7), 1445-1458. https://doi.org/10.1093/aob/mcu032

Stanley, S., Antoniou, V., Askquith-Ellis, A., Ball, L. A., Bennett, E. S., Blake, J. R., Boorman, D. B., Brooks, M., Clarke, M., Cooper, H. M., Cowan, N., Cumming, A., Evans, J. G., Farrand, P., Fry, M., Harvey, D., Houghton-Carr, H., Howson, T., Khamis, D., . . . Winterbourn, J. B. (2023). *Daily and sub-daily hydrometeorological and soil data (2013-2022) [COSMOS-UK]* NERC EDS Environmental Information Data Centre. https://doi.org/10.5285/5060cc27-0b5b-471b-86eb-71f96da0c80f

Stetter, M. G., Zeitler, L., Steinhaus, A., Kroener, K., Biljecki, M., & Schmid, K. J. (2016). Crossing Methods and Cultivation Conditions for Rapid Production of Segregating Populations in Three Grain Amaranth Species [Original Research]. *Frontiers in plant science*, *7*. https://doi.org/10.3389/fpls.2016.00816

Swiezewski, S., Liu, F., Magusin, A., & Dean, C. (2009). Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycomb target. *Nature (London)*, *462*(7274), 799-802. https://doi.org/10.1038/nature08618

Sysoeva, M. I., Markovskaya, E. F., & Shibaeva, T. G. (2010). Plants under continuous light: a review. *Plant stress*, *4*(1), 5-17.

Szűcs, P., Skinner, J. S., Karsai, I., Cuesta-Marcos, A., Haggard, K. G., Corey, A. E., Chen, T. H., & Hayes, P. M. (2007). Validation of the VRN-H2/VRN-H1 epistatic model in barley reveals that intron length variation in VRN-H1 may account for a continuum of vernalization sensitivity. *Molecular Genetics and Genomics*, *277*(3), 249-261. https://doi.org/10.1007/s00438-006-0195-8

Tanaka, C., Itoh, T., Iwasaki, Y., Mizuno, N., Nasuda, S., & Murai, K. (2018). Direct interaction between VRN1 protein and the promoter region of the wheat FT gene. *Genes & Genetic Systems*, *93*(1), 25-29. https://doi.org/10.1266/ggs.17-00041

Thaler, S., Eitzinger, J., Trnka, M., & Dubrovsky, M. (2012). Impacts of climate change and alternative adaptation options on winter wheat yield and water

productivity in a dry climate in Central Europe. *The Journal of Agricultural Science*, *150*(5), 537-555. https://doi.org/10.1017/S0021859612000093

Theißen, G., Becker, A., Di Rosa, A., Kanno, A., Kim, J. T., Münster, T., Winter, K. U., & Saedler, H. (2000). A short history of MADS-box genes in plants. *Plant Molecular Biology*, *42*(1), 115-149.

Trevaskis, B., Bagnall, D. J., Ellis, M. H., Peacock, W. J., & Dennis, E. S. (2003). MADS Box Genes Control Vernalization-Induced Flowering in Cereals. *Proceedings of the National Academy of Sciences - PNAS*, *100*(22), 13099-13104. https://doi.org/10.1073/pnas.1635053100

Trevaskis, B., Hemming, M. N., Dennis, E. S., & Peacock, W. J. (2007). The molecular basis of vernalization-induced flowering in cereals. *Trends in Plant Science*, *12*(8), 352-357. https://doi.org/10.1016/j.tplants.2007.06.010

Turner, A., Beales, J., Faure, S., Dunford, R. P., & Laurie, D. A. (2005). The pseudo-response regulator Ppd-H1 provides adaptation to photoperiod in barley. *Science*, *310*(5750), 1031-1034. https://doi.org/10.1126/science.1117619

van Dijk, A. D., Morabito, G., Fiers, M., van Ham, R. C., Angenent, G. C., & Immink, R. G. (2010). Sequence motifs in MADS transcription factors responsible for specificity and diversification of protein-protein interaction. *PLoS computational biology*, *6*(11), e1001017.

Van Bel, M., Silvestri, F., Weitz, E. M., Kreft, L., Botzki, A., Coppens, F., & Vandepoele, K. (2022). PLAZA 5.0: extending the scope and power of comparative and functional genomics in plants. *Nucleic acids research*, *50*(D1), D1468-D1474. https://doi.org/10.1093/nar/gkab1024

Velez-Ramirez, A. I., van Ieperen, W., Vreugdenhil, D., & Millenaar, F. F. (2011). Plants under continuous light. *Trends in Plant Science*, *16*(6), 310-318. https://doi.org/10.1016/j.tplants.2011.02.003

von Zitzewitz, J., Szűcs, P., Dubcovsky, J., Yan, L., Francia, E., Pecchioni, N.,

Casas, A., Chen, T. H., Hayes, P. M., & Skinner, J. S. (2005). Molecular and structural characterization of barley vernalization genes. *Plant Molecular Biology*, *59*(3), 449-467. https://doi.org/10.1007/s11103-005-0351-2

Vyse, K., Schaarschmidt, S., Erban, A., Kopka, J., & Zuther, E. (2022). Specific CBF transcription factors and cold-responsive genes fine-tune the early triggering response after acquisition of cold priming and memory. *Physiologia Plantarum*, *174*(4), e13740. https://doi.org/https://doi.org/10.1111/ppl.13740

Watson, A., Ghosh, S., Williams, M. J., Cuddy, W. S., Simmonds, J., Rey, M.-D., Asyraf Md Hatta, M., Hinchliffe, A., Steed, A., Reynolds, D., Adamski, N. M., Breakspear, A., Korolev, A., Rayner, T., Dixon, L. E., Riaz, A., Martin, W., Ryan, M., Edwards, D., . . . Hickey, L. T. (2018). Speed breeding is a powerful tool to accelerate crop research and breeding. *Nature Plants*, *4*(1), 23-29. https://doi.org/10.1038/s41477-017-0083-8

Webber, H., Ewert, F., Olesen, J. E., Müller, C., Fronzek, S., Ruane, A. C., Bourgault, M., Martre, P., Ababaei, B., Bindi, M., Ferrise, R., Finger, R., Fodor, N., Gabaldón-Leal, C., Gaiser, T., Jabloun, M., Kersebaum, K.-C., Lizaso, J. I., Lorite, I. J., . . . Wallach, D. (2018). Diverging importance of drought stress for maize and winter wheat in Europe. *Nature communications*, *9*(1), 4249. https://doi.org/10.1038/s41467-018-06525-2

Wilhelm, E. P., Turner, A. S., & Laurie, D. A. (2009). Photoperiod insensitive Ppd-A1a mutations in tetraploid wheat (Triticum durum Desf.). *Theoretical and applied genetics*, *118*(2), 285-294. https://doi.org/10.1007/s00122-008-0898-9

Winfield, M. O., Allen, A. M., Wilkinson, P. A., Burridge, A. J., Barker, G. L. A., Coghill, J., Waterfall, C., Wingen, L. U., Griffiths, S., & Edwards, K. J. (2018). High-density genotyping of the A.E. Watkins Collection of hexaploid landraces identifies a large molecular diversity compared to elite bread wheat. *Plant Biotechnology Journal*, *16*(1), 165-175. https://doi.org/10.1111/pbi.12757

Winfield, M. O., Lu, C., Wilson, I. D., Coghill, J. A., & Edwards, K. J. (2009). Cold-

and light-induced changes in the transcriptome of wheat leading to phase transition from vegetative to reproductive growth. *BMC plant biology*, *9*(1), 55-55. https://doi.org/10.1186/1471-2229-9-55

Wingen, L. U., Orford, S., Goram, R., Leverington-Waite, M., Bilham, L., Patsiou, T. S., Ambrose, M., Dicks, J., & Griffiths, S. (2014). Establishing the A. E. Watkins landrace cultivar collection as a resource for systematic gene discovery in bread wheat. *Theoretical and applied genetics*, *127*(8), 1831-1842. https://doi.org/10.1007/s00122-014-2344-5

Wingen, L. U., West, C., Leverington-Waite, M., Collier, S., Orford, S., Goram, R., Yang, C. Y., King, J., Allen, A. M., Burridge, A., Edwards, K. J., & Griffiths, S. (2017). Wheat Landrace Genome Diversity. *Genetics*, *205*(4), 1657-1676. https://doi.org/10.1534/genetics.116.194688

Xiao, J., Xu, S., Li, C., Xu, Y., Xing, L., Niu, Y., Huan, Q., Tang, Y., Zhao, C., Wagner, D., Gao, C., & Chong, K. (2014). O-GlcNAc-mediated interaction between VER2 and TaGRP2 elicits TaVRN1 mRNA accumulation during vernalization in winter wheat. *Nature communications*, *5*(1), 4572. https://doi.org/10.1038/ncomms5572

Xie, L., Zhang, Y., Wang, K., Luo, X., Xu, D., Tian, X., Li, L., Ye, X., Xia, X., Li, W., Yan, L., & Cao, S. (2021). TaVrt2, an SVP-like gene, cooperates with TaVrn1 to regulate vernalization-induced flowering in wheat. *New Phytologist*, *231*(2), 834-848. https://doi.org/https://doi.org/10.1111/nph.16339

Xu, S., & Chong, K. (2018). Remembering winter through vernalisation. *Nature Plants*, *4*(12), 997-1009. https://doi.org/10.1038/s41477-018-0301-z

Xu, S., Xiao, J., Yin, F., Guo, X., Xing, L., Xu, Y., & Chong, K. (2019). The Protein Modifications of O-GlcNAcylation and Phosphorylation Mediate Vernalization Response for Flowering in Winter Wheat. *Plant Physiology*, *180*(3), 1436-1449. https://doi.org/10.1104/pp.19.00081

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*(51), 19581-19586. https://doi.org/10.1073/pnas.0607142103

Yan, L., Loukoianov, A., Blechl, A., Tranquilli, G., Ramakrishna, W., Sanmiguel, P., Bennetzen, J. L., Echenique, V., & Dubcovsky, J. (2004). The Wheat VRN2 Gene Is a Flowering Repressor Down-Regulated by Vernalization. *Science (American Association for the Advancement of Science)*, *303*(5664), 1640-1644. https://doi.org/10.1126/science.1094305

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 - PNAS*, *100*(10), 6263-6268. https://doi.org/10.1073/pnas.0937399100

Yang, H., Berry, S., Olsson, T. S. G., Hartley, M., Howard, M., & Dean, C. (2017). Distinct phases of Polycomb silencing to hold epigenetic memory of cold in Arabidopsis. *Science*, *357*(6356), 1142-1145. https://doi.org/10.1126/science.aan1121

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. L. (2012). Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC bioinformatics*, *13*, 134. https://doi.org/10.1186/1471-2105-13-134

Zadoks, J. C., Chang, T. T., & Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Research*, *14*(6), 415-421. https://doi.org/https://doi.org/10.1111/j.1365-3180.1974.tb01084.x

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 : MGG*, *276*(4), 334-350.

https://doi.org/10.1007/s00438-006-0147-3

Zhao, Y., Zhu, P., Hepworth, J., Bloomer, R., Antoniou-Kourounioti, R. L., Doughty, J., Heckmann, A., Xu, C., Yang, H., & Dean, C. (2021). Natural temperature fluctuations promote COOLAIR regulation of FLC. *Genes & development*, *35*(11-12), 888-898. https://doi.org/10.1101/gad.348362.121

Zieliński, K., Krzewska, M., Żur, I., Juzoń, K., Kopeć, P., Nowicka, A., Moravčiková, J., Skrzypek, E., & Dubas, E. (2020). The effect of glutathione and mannitol on androgenesis in anther and isolated microspore cultures of rye (Secale cereale L.). *Plant Cell, Tissue and Organ Culture (PCTOC)*, *140*(3), 577-592. https://doi.org/10.1007/s11240-019-01754-9

Zikhali, M., Leverington-Waite, M., Fish, L., Simmonds, J., Orford, S., Wingen, L. U., Goram, R., Gosman, N., Bentley, A., & Griffiths, S. (2014). Validation of a 1DL earliness per se (eps) flowering QTL in bread wheat (Triticum aestivum). *Molecular Breeding*, *34*, 1023-1033. https://doi.org/10.1007/s11032-014-0094-3