SELECTING DAYS TO FLOWERING IN LENTIL (*Lens culinaris* Medik.) FOR A NORTHERN TEMPERATE CLIMATE

A Thesis Submitted to the College of Graduate and Postdoctoral Studies

In Partial Fulfillment of the Requirements

For the Degree of Master of Science

In the Department of Plant Sciences

University of Saskatchewan

By Taryn Heidecker

PERMISSION TO USE

In presenting this thesis/dissertation in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis. Requests for permission to copy or to make other use of the material in this thesis, whole or part, should be addressed to:

Head of the Department of Plant Sciences
University of Saskatchewan
4D36 Agriculture Building, 51 Campus Drive
Saskatoon, Saskatchewan, S7N 5A8
Canada

AND/OR

Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan, S7N 5C9
Canada

ABSTRACT

Lentil varieties in western Canada are continuously improved to stay relevant as a crop. This improvement requires access to a diverse pool of genetics of which, in lentil, is difficult to use due to a lack of adaptation to our environment. Genetic markers for days to flowering (DTF) have been identified but have not yet been tested in western Canadian lentil breeding. Phenotyping DTF in Saskatchewan field experiments and mapping, using markers associated with specific loci in the lentil genome, was used to identify loci relevant to western Canadian conditions and applicable to diverse lentil germplasm. A bi-parental RIL population (LR-11) created using a Canadian line (CDC Milestone) crossed with a Bangladeshi line (ILL 8006) was phenotyped for multiple phenological traits in Saskatchewan field experiments over four site-years. A linkage map consisting of six linkage groups (LGs) was constructed using 11, 558 single nucleotide polymorphism (SNP) markers. Four quantitative trait loci (QTL) for DTF (q.DTF) were identified in more than one site-year. Of these, the two which explained the largest amount of the observed phenotypic variability each contained members of the *Flowering Locus T* (*FT*) gene family as annotated in the lentil reference genome. An exploratory expression study of these *FT* genes provided additional support that q.DTF.6-1 and q.DTF.6-2 may represent variation at *LcFTb2* and *LcFTa1*, respectively. A diversity panel, AGILE-LDP, was screened with a marker representative of each of these two q.DTF and both markers accounted for differences in DTF in Saskatchewan. The inheritance patterns of the two markers, and the implied roles of *LcFTb2* and *LcFTa1*, based on research in related legumes, provided support that the markers are accounting for at least some of the relevant variation in DTF in this panel. These q.DTF could be used to improve selection of preferred allele combinations across diverse material. This study emphasized the need for further investigation into the underlying genes and molecular pathways implicated by q.DTF in any study, and that, on their own, focusing on any one locus leaves many unanswered questions and reduces marker adoption by breeders. In addition to DTF, days to emergence, vegetative period, and reproductive period were identified as having the potential for identifying additional discrete, and phenologically relevant genomic regions that should be followed up in future studies.

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor Dr. Kirstin Bett for her patience and guidance throughout this project. I would also like to thank my committee members Drs. B. Ta'ran, T. Sharbel and C. Pozniak for their comments and advice. I would also like to thank the external examiner, Dr. Nickolas Larkan, for his time.

A special thanks belongs to Sandesh Neupane, Karsten Nielsen, Teketel Haile and Derek Wright, for their help and the camaraderie that came with data collection and analysis. Akiko Tomita and Robert Stonehouse also deserve my thanks for their mentorship and all the work they do in the lab. For the care and attention of the field trials, I'd like to thank Brent Barlow and the Crop Science Field Lab crew at USask for all their hard work. I am also thankful to acknowledge the KnowPulse team at USask for their expertise and support regarding phenotypic and genotypic data management in this project.

The assistance and kindness of the research collaborators at the University of Tasmania was also very welcome. Specifically, I'd like to recognise Drs. Jim Weller, Raul Ortega, Steven Ridge, and Jackie Vander Schoor and their families.

Research in my thesis was supported by the AGILE (Application of Genomics to Innovation in the Lentil Economy) project. I am also grateful to the Robert P. Knowles scholarship fund and the Canada 150 Travel Award which provided me financial support.

Finally, the love and patience of my friends and family, specifically my wife Nicole, has been very much appreciated and I know I have been extremely lucky.

Thank you all

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AGILE	Application of Genomics to Innovation in the Lentil Economy
AGILE-LDP	The lentil diversity panel developed at USask
ANOVA	Analysis of Variance
bp	Nucleotide base pairs
CO	CONSTANS
CDC	Crop Development Centre (located at USask)
CDF	CYCLING DOF FACTOR
°C	Degree Celsius
DNA	Deoxyribonucleic acid
DTE	Days to emergence
DTF	Days to flowering
DTH	Days to Harvest
DTM	Days to Maturity
DTSP	Days to Swollen Pod
ELF	EARLY FLOWERING LOCUS
FKF	FLAVIN-BINDING, KELCH DOMAIN, F BOX
FLC	FLOWERING LOCUS C
FT	FLOWERING LOCUS T
FUL	FRUITFUL

Gb	Giga base pair
GDD	Growing Degree Days
GI	GIGANTEA
GWAS	Genome-Wide Association Studies
h	Hour
HR	HIGH RESPONSE TO PHOTOPERIOD
KASP	Kompetitive Allele Specific PCR
LG	Linkage group
lncRNA	Long non-coding RNA
LR-11	Lentil RIL population 11
LUX	LUX ARRHYTHMO
NFD	Node of flower development
PCR	Polymerase chain reaction
PHYA	PHYTOCHROME A
PRR	PSEUDO-RESPONSE REGULATOR
PTT	Photothermal Time
qPCR	Quantitative PCR
QTL	Quantitative trait loci
RepP	Reproductive Period (DTM-DTF)
RIL	Recombinant Inbred Line
SCAR	Sequence Characterized Amplified Region
SK	Saskatchewan, Canada
Sn	STERILE NODES
ssp.	Subspecies
SSR	Simple sequence repeat
SOC	SUPPRESSOR OF OVEREXPRESSION OF CO
USask	University of Saskatchewan
VegP	Vegetative Period (DTF-DTE)
YLD	Gross Plot Yield

LIST of APPENDICES

Chapter 1. Introduction

As a crop, lentil (*Lens culinaris* Medik.) faces stiff competition for a place in a western Canadian grower's rotation. Profitable crop alternatives, increasingly competitive weeds, and diseases, as well as potentially severe environmental conditions influence lentil acreage. Breeding of elite cultivars able to meet grower expectations requires breeders have access diverse lentil material and the ability to make efficient selections during development.

A major constraint to the use of exotic material in breeding material for western Canada is that lentil grown in other parts of the world responds differently to some key local environmental conditions (Wright et al. 2020). Adaptations affecting crop phenology are strong influences of seed yield in western Canadian grown lentil (Tullu et al. 2008). Consequently, the amount of time it takes plants to flower, often investigated as days to flower after planting (DTF), is a common indicator of fitness in an environment (Tullu et al. 2008; Fedoruk et al. 2013).

Genetic markers for flowering are currently not used in the Canadian breeding program due to being poorly adaptable across diverse material (Tullu et al. 2008). In addition, in studies conducted under short day conditions (daylength of 12 hours and less), markers of large effect have been associated with genetic variation not relevant to western Canadian field conditions (Sarker et al. 1999b; Fratini et al. 2007; Saha et al. 2013). The large number of genetic markers now being generated using 'next generation' sequencing techniques, and the continued improvement of the lentil genome, are contributing to the search for loci relevant to DTF in western Canada. Field experiments conducted in Saskatchewan have now identified loci which are in or near putative genes relevant to DTF under these long day conditions (where daylength is greater than 16 hours) (Haile et al. 2021). Employing similarly dense mapping methods, indoor experiments have also identified loci of suspected importance to long day conditions (Rajandran et al. 2021; Yuan et al. 2021).The complexity of the pathway of signals influencing DTF suggests there are more loci to identify in additional genetic backgrounds. Genome wide association studies (GWAS) using diverse panels of lines may be used for this but run the risk of filtering out important contributors when found in relatively few individuals (Upadhyaya et al. 2015), which is possible for a trait like DTF, where the phenological response is often linked to climatic region-specific adaptations. Because of this, studies which have identified markers related to DTF in lentil have typically used bi-parental populations. Testing DTF markers,

initially identified in bi-parental populations, on diverse lines would help determine which ones are useful and support marker use as indicators of fitness for breeders considering exotic material. Verified markers could then be used in screening germplasm arising from a cross using exotic material to remove material with undesirable alleles and create more space and time for those carrying ideal alleles (Tullu et al. 2013). This is particularly important for development-related traits like DTF which are strongly influenced by environmental pressures, and phenotypes collected during early generations grown in controlled environments, may not be readily comparable to those in a field setting (preliminary study in Appendix 1).

1.1 Hypotheses

1. Variability in days to flowering (DTF) in a Canadian by South Asian lentil population is controlled by genetic factors at relatively few regions of the genome containing known flowering time genes.

2. Genetic markers corresponding to flowering QTLs in bi-parental populations will account for DTF in Saskatchewan (SK) in a population of diverse individuals. Markers identified in SK field trials will be more useful in long day experiments than those found in a short-day greenhouse (GH) experiment.

3. Markers associated with different candidate flowering genes will independently account for variation in DTF among diverse lentil lines.

1.2 Objectives

1. Phenotypically and genotypically characterize the bi-parental RIL population LR-11 (ILL 8006 x CDC Milestone) for development-related traits and identify QTLs controlling days to flowering and other phenology traits.

2. Test markers associated with candidate flowering genes to determine if they explain differences observed for days to flowering among diverse germplasm in both a diversity panel (AGILE-LDP) and a bi-parental population (LR-11).

Chapter 2. Literature Review

2.1 Cultivated Lentil Diversity; a product of selection for climatic region

2.1.1 Lentil domestication

Cultivated lentil (*Lens culinaris* Medik.) is a diploid (2n=2x=14) and self-pollinating pulse crop. There are six other species in this genus, including: *L. orientalis, L. odemensis, L. tomentosus, L. lamottei, L. ervoides, and L. nigricans* (Wong et al. 2015). The *Lens* species originated in the Mediterranean and Middle Eastern 'Fertile Crescent', with the cultivated type *L .culinaris* likely selected from its wild cousin *L. orientalis* there over 9000 years ago (Zohary 1972; Ladizinsky 1979; Cubero et al. 2009). Genetic studies have refined the location of initial domestication, and identify what today is Syria and Turkey as the place of origin (Ferguson et al. 1998; Sonnante et al. 2009), with some suggesting it may have been limited to an even smaller area in Southern Turkey (Alo et al. 2011).

While initial domestication mutations would have been for reduced seed dehiscence and dormancy, along with a larger seed size (Sonnante et al. 2009), diversity was further reduced by the requirement for new phenological responses as domesticated lentil reached new areas (Sarker and Erskine 2006; Khazaei et al. 2016). Today, wild species of lentil are potential sources of abiotic stress resistance, specifically to factors such as drought, salinity, alkalinity, and winter hardiness (Erskine et al. 1994b; Kahraman et al. 2004; Kumar et al. 2016; Singh et al. 2016) as well as resistance to biological stressors especially fungal diseases like Anthracnose (*Colletotrichum lenti*s), Stemphylium blight (*Stemphylium botryosum*), and Ascochyta blight (*Ascochyta lentis*) (Erskine et al. 1994b; Fiala et al. 2009; Tullu et al. 2011; Podder et al. 2013; Gela et al. 2021b). Resistance genes for Anthracnose, not found in *L. culinaris*, have been introgressed from wild *Lens*, but these crosses are accompanied by the need for many generations to select and fix ideal agronomic traits, one of which in western Canada is flowering habit (Tullu et al. 2013; Gela et al. 2021a).

2.1.2 Adaptations for flowering time allowed lentil expansion, reduce diversity

When considering the cultivated *L. culinaris*, landraces are much more genetically diverse than lines used in breeding programs (Erskine et al. 1989; Ferguson et al. 1998; Alo et al. 2011; Lombardi et al. 2014). The landraces and cultivars grown in Syria, Turkey and Greece are particularly diverse, likely due to their geographical proximity to many wild species (Erskine et

al. 1989; Singh et al. 2014). One working hypotheses for the diversity in these regions is that there were originally multiple domestication events or recombinations which occurred between early domesticated and wild species (Abbo et al. 2009; Fuller et al. 2011). Research into differences in cultivated lentil initially associated consumer preferences for seed-morphology with group diversity, however, studies from as early as the 1990s indicated that this grouping system does not account for genetic or geographical variation (Ferguson et al. 1998; Sonnante et al. 2009; Khazaei et al. 2016). Instead, a more important factor contributing to current worldwide lentil diversity has been the differences in climatic pressures, namely differences in temperature and photoperiod, in the areas lentil has spread to and the associated phenological requirements (Erskine et al. 1989, 1994b; Ferguson et al. 1998; Lombardi et al. 2014; Wright et al. 2020).

2.1.3 Contrasting flowering adaptations provide success in different climatic regions

In the late 1980s, researchers began to identify the different phenological responses lentils can have to different climatic conditions, primarily those caused by differences in daylength and temperature (Summerfield et al. 1985; Roberts et al. 1988; Erskine et al. 1990a, 1994a). They concluded that the time it takes from seeding and emergence to flowering is the most important constraint defining the region a lentil is adapted to (Erskine et al. 1989). The result of this is that the adaptations required for success in one climatic region (or 'macro environment') are not the same as in another, and fruitful material from one climatic region may be unfit for another. Cultivated lentil is now genetically distinct when comparing material from different climatic regions, largely due to the isolation caused by differences in flowering time (Erskine et al. 1998; Khazaei et al. 2016).

Across the major production areas, today lentil is generally grouped based on fitness for one of three general climatic regions: 'northern Temperate', where material is spring sown and experiences long days; 'Mediterranean', which is fall sown and responds to increasing temperatures following a period of low temperature, and increasing daylength; or 'South Asian', where lentil is also fall sown but is adapted to flower and mature quickly before the temperatures are too high (Khazaei et al. 2016; Wright et al. 2020).

Mediterranean-adapted lentil is the most diverse group, and these types include those grown where cultivation of lentil is suspected to have begun (Lombardi et al. 2014; Khazaei et al. 2016). When lentil cropping expanded from West Asia into the Indo-Gangetic Plains (northern

India), selected material was generally faster maturing, and accessions that succeed there today have comparably low photoperiod sensitivity (Erskine et al. 1994a). South Asian lentil of this climatic region flower and mature quickly under short days, to avoid exposure to the heat and drought of summer. The loss of genetic diversity in lentil has been best described in the *pilosae* ecotype, which is able to reach seed set under these climatic stresses but is susceptible to biotic and other abiotic stresses due to the small genetic pool adapted to these South Asian conditions (Khazaei et al. 2016; Wright et al. 2020). Improvement of lentil in this climatic region has focused on identifying exotic material to use in breeding programs, as well as mutation breeding (Erskine et al. 1998), but without effective tools for selection, it has been costly and this material still typically only yields 60 % of the global average (Dikshit et al. 2015).

Temperate-adapted lentil germplasm generally takes the longest to flower when compared to material adapted to other climatic regions. These lentils are typically grown at higher latitudes than material in the Fertile Crescent, are less sensitive to photoperiod and have higher temperature sensitivity when grown under short day conditions (days of 12 hours and less) (Erskine et al. 1990a, 1994a). In contrast to how material was able to colonize northern India, which was through a increased responsiveness to photoperiod and temperature, germplasm grown in Temperate conditions succeed by responding positively to increasing temperature and daylength conditions to avoid being prematurely driven to flowering (Wright et al. 2020). Conflicting environmental requirements have reduced the direct usefulness of lentil from other climatic regions in Canadian breeding programs as the production practice of spring sowing, which results in comparatively long days during early growth, is unique to material with Temperate fitness (Tullu et al. 2001, 2011; Khazaei et al. 2016).

2.2 Improving Canadian lentil requires focus on flowering habit

'Days to flower' (DTF), is used as a selection tool in lentil breeding due to its high correlation to harvest date and yield while possessing higher heritability (Tullu et al. 2008). DTF is closely linked to where lentils are grown and, in Canadian breeding programs, material which comes from other climatic regions has been avoided due to the investment required following a cross. Under the current practice of phenotypic selection for DTF, Canadian breeders find it more efficient to backcross to adapted material following a 'wider' cross. An example of this is the use of ILL 5588, a line present in the pedigree of many Canadian lines as a source of Ascochyta

blight resistance. ILL 5588 is a line considered ‘adapted enough’ to Canadian conditions that only one additional round of backcrossing to adapted lentil was sufficient prior to selfing and line selection (Vandenberg et al. 2001; Tar’an et al. 2003).

The simple inheritance of some agronomic traits, combined with markers for disease (Tar’an et al. 2003), have allowed effective improvement in Canadian lentil breeding, but there is an increasing need for efficient access to broader genetic diversity. Having molecular markers for DTF would aid in evaluation and early selection of material from crosses involving lentil outside the Temperate climatic group (Tullu et al. 2013; Dikshit et al. 2015; Wright et al. 2020).

2.3 Understanding flowering as a tool for selection in diverse lentils

2.3.1 Flowering time as a measure for phenological response in lentil

Phenology encompasses many components of plant development, including the time required to emerge, flower, pod, and mature, among other stages, many of which have been described in the context of assessing lentil development (Erskine et al. 1990b). Days to flowering (DTF), referenced multiple ways including ‘earliness’, ‘flowering’ and sometimes ‘flowering habit’, has typically been used by breeders as a measure due to its correlation with maturity and fitness (Erskine et al. 1994b; Tullu et al. 2008). Evaluating diverse lentil with adaptations for different climatic regions has generally relied on understanding flowering responses to temperature and photoperiod (Erskine et al. 1994b; Sarker and Erskine 2006).

With high throughput phenotyping and genotyping now commonplace, studying plant development as a phenological response has been done across many relevant species. Phenological response has been used both for the characterization of diverse lines (Tullu et al. 2001; Berger et al. 2011; Ghanem et al. 2015; Wright et al. 2020) and for the identification of conserved flowering time genes (Hecht et al. 2005; Weller et al. 2012; Burgarella et al. 2016; Ortega et al. 2019) across legume species.

2.3.2 Lentil flowering occurs in response to photoperiod and temperature signals

Lentil is defined as a quantitative long day plant, with flowering being independently influenced by photoperiod and temperature (Summerfield et al. 1985). Sensitivities to vernalisation and light quality also contribute to the variation for flowering in lentil (Roberts et al. 1988; Yuan et al. 2017, 2021). Early studies found that characterizing material by explaining days to flower as a

phenological response to photoperiod and temperature improved defining the differences in germplasm (Summerfield et al. 1985). Calculation of genotype-specific values for critical photoperiod were also found to be a useful phenotype for identifying lentils more easily adapted to new latitudes (Roberts et al. 1986; Wright et al. 2020). Recent work using a large lentil diversity panel (AGILE-LDP) has shown there are limitations to these early models, likely due to the large variance in photoperiod sensitivities among diverse lentil genotypes (Wright et al. 2020). Classifying early plant development stages based on only photoperiod to refine flowering prediction in lentil has been proposed (Roberts et al. 1988), and worked when using more complex equations in chickpea (Daba et al. 2016).

Daba et al. (2016) found linear mixed models for photothermal time were useful for interpreting the effects of temperature and photoperiod on transition to flowering in chickpea. Beneficial to these types of equations, lentil has been shown to have a linear and positive response to temperature, and research specific to the Indian sub-continent has defined some of the upper limits lentil can yield in (Erskine et al. 1990a; Sita et al. 2018). The relationship of lentil flowering to temperature and photoperiod will be refined by considering the genetic pathways that underly the observed responses (Wright et al. 2020). Continuing to investigate these environmental pressures individually across multiple environments identifies loci influencing flowering in distinct genetic pathways (Neupane et al. unpublished).

2.4 Synteny in legumes allows a better understanding of lentil flowering

2.4.1 Genetics of flowering in long day, cool season legumes

The *Lens* (Miller) genus is a member of the *Fabaceae* family which includes several other agronomically important genera (Cronk et al. 2006). This family is made up of some common cool-season crops including the common pea (*Pisum sativum*), faba bean (*Vicia faba*) and Chickpea (*Cicer arietinum*), as well as the model legume *Medicago truncatula*.

Understanding flowering genes in legumes is often first done in reference to the model plant *Arabidopsis*, in which the consensus is that transition from vegetative to reproductive growth is signaled by over 100 genes through a number of pathways, which interact epistatically (Ehrenreich et al. 2009; Fornara et al. 2010; Srikanth and Schmid 2011). The Flor-ID database currently has 306 *Arabidopsis* genes linked to at least one flowering pathway (Bouché et al.

2016). Flowering genes have been shown to possess different roles based on species (Jung and Müller 2009), and within legumes there is a large diversity in the occurrence and variation of certain flowering-related gene families (Hecht et al. 2005; Kim et al. 2013; Weller and Ortega 2015). Legume species in which flowering-related genes have been most extensively studied include the long day responsive pea and *Medicago truncatula*, as well as the short day induced soybean (*Glycine max*) (Kim et al. 2013; Weller and Ortega 2015).

Flowering-related genes are often grouped based on the environmental factors they respond to; the pathways of highest significance are well conserved among long day legumes. Of these, the genes involved in the photoperiod and temperature pathways are most likely to harbor variation useful for lentil breeding programs (Roberts et al. 1988; Srikanth and Schmid 2011). The genes which act as Floral Pathway Integrators (FPIs), integrating signals from multiple pathways, are also highly conserved across plant species, and are preferred candidates for selection in crop improvement (Jung et al. 2017; van Dijk and Molenaar 2017)

2.4.2 Floral Pathway Integrators

Multiple orthologs of *Flowering Locus T* (*FT*), an important floral pathway integrator (FPI) sometimes called 'florigen', have been identified in both short day and long day responsive legumes (Hecht et al. 2011; Weller and Ortega 2015; Liu et al. 2018). *FT* and *FT*-like genes are found across numerous plant species and encode for small (~175 amino acid), phosphatidylethanolamine-binding (PEBP) proteins, which act as mobile signals often linked to flowering pathways (Putterill and Varkonyi-Gasic 2016). Studies in legumes have identified unique regulators of *FT* homologs such as the *E1* locus, which is important in soybean (Zhai et al. 2014) and may also regulate *FT* homologs in long day/temperate legumes (Zhang et al. 2016). In addition, the role of some central *FT* regulators for *Arabidopsis* are altered or yet to be implicated in temperate legumes, as seems the case for *CONSTANS* (*CO*) (Wong et al. 2014; Weller and Ortega 2015; Ridge et al. 2016), and *FLOWERING LOCUS C (FLC)* (Kim et al. 2013; Jaudal et al. 2020).

The role of legume *FT* homologs themselves is expanded beyond that reported in the model *Arabidopsis*. In temperate legumes there are five conserved homologs of the *Arabidopsis FT* gene: *FTa1*, *FTa2*, *FTb1*, *FTb2*, and *FTc* (Hecht et al. 2011; Weller and Ortega 2015). *FTb* gene expression most closely follows that of the 'florigen' action of *FT* in *Arabidopsis.* In long day

responsive legumes, *FTb* is expressed in the leaves diurnally, with levels peaking ~4 and 16 hrs after sunrise and can be used as an early signal for the transition to flowering following inducive photoperiods (Hecht et al. 2011; Weller and Ortega 2015; Thomson et al. 2019).

FTa1 has been found most often in integrating signals related to photoperiod, temperature, and light quality (Hecht et al. 2011; Laurie et al. 2011; Yuan et al. 2021). In contrast to *FTb*, *FTa1* is expressed at a constant rate during the day, however increased expression of *FTa1* is linked to higher *FTb* expression levels (Thomson et al. 2019). Similar to *FTb*, it is first expressed in the leaves, but it is also suspected to produce a transcript which travels to the shoot apical meristem (SAM) (Hecht et al. 2011; Thomson et al. 2019). In the SAM, the expression of two more FPIs: *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* (*SOC1*) and *APETALA1* (*AP1*), follow an increase of *FTc* expression (Hecht et al. 2011; Fudge et al. 2018). At this time, however, the role of *FTc* remains to be systematically explored.

It is surmised that the increased diversity of the *FT* family in temperate legumes, relative to *Arabidopsis*, has resulted in their expanded importance in integrating signals, including cross-regulation between *FT* homologs themselves (Figure 1.1) (Hecht et al. 2011; Jaudal et al. 2013; Liew et al. 2014). Intergenic regions have been implicated in regulating *FTa1* in both model and crop legumes (Jaudal et al. 2013; Nelson et al. 2017; Ortega et al. 2019). In lentil, an intergenic region between *FTa1* and *FTa2* is linked to a flowering QTL, with accessions that are missing some of this region tending to flower earlier, presumably due to a loss of flowering suppression (Rajandran et al. 2021). A consequence of this complex intergenic and cross-regulation of the *FT* family in legumes is that the role of many genes suspected of influencing flowering in response to photoperiod and temperature pressures are often measured based on their adjustment of the expression of legume *FT* homologs (Hecht et al. 2011; Liew et al. 2014; Weller and Ortega 2015).

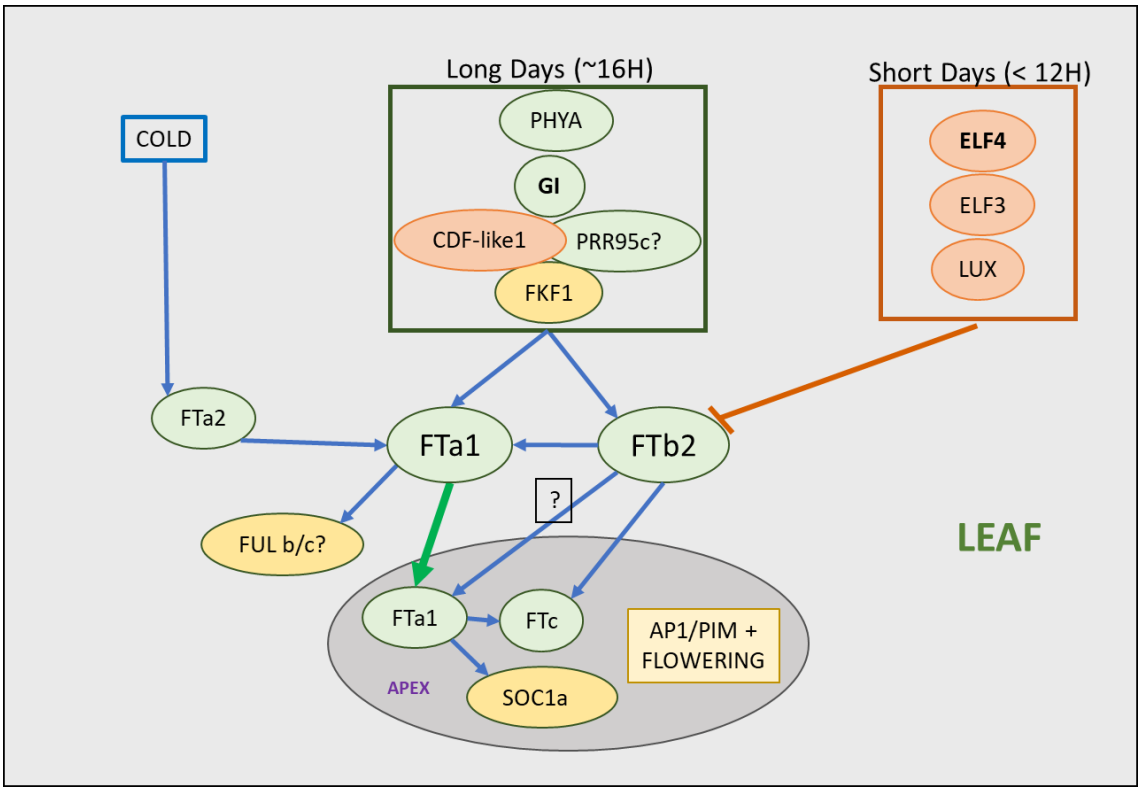

Figure 2.1 Conserved flowering pathway for temperate legumes based on pea and *Medicago truncatula.* Genes placed in light grey background have shown expression in leaf tissue, with darker background indicated apex tissue expression. Photoperiod sensitive genes are grouped by photoperiod which increases level of expression (long versus short day). Genes with expression that induces flowering are on a green background, while those which suppress are in orange. Yellow indicates genes which have not been characterized in multiple legume species. Blue arrows indicate a positive correlation in expression, orange indicate a negative correlation (direction indicates order of induction). The green arrow indicates movement of transcribed protein from leaves to apex. [?] indicates pathway which have only been speculated in legumes. Summary based on: Hecht et al. (2011), Weller and Ortega (2015), Ridge et al. (2016), Ridge et al. (2017), Fudge et al. (2018), and Rajandran et al. (2021). Gene names are those used for homologs in Medicago with phylogeny analyses (Hecht et el. 2011; unpublished) allowing this naming convention to be used in the current lentil genome annotation (Lcu.RBY2).

2.4.3 Photoperiod Pathway

In *Arabidopsis*, *CONSTANS* (*CO*) transcripts accumulate during the day and are highest near the end of the day so that, when the days are long enough, a signalling complex formed with *GIGANTEA (GI)* and *FT* induces flowering (Turck et al. 2008). A number of phytochromes are also implicated in adjusting this process (Jung and Müller 2009). It appears that, in legumes, the role of *GI* remains conserved but *CO* is not, with *GI* expressed in response to inducive

daylengths, followed by *FT* homolog expression and flowering (Hecht et al. 2007; Weller and Ortega 2015). A *CYCLING DOF FACTOR* (*CDF*) gene homolog (*CDF-like1*) has also been found to regulate *FTb2* expression in pea, possibly with the help of a *Pseudo-Response Regulator* gene homolog (*PRR95c*) (Ridge et al. 2016). In temperate legumes, blue light photoreceptor *FLAVIN BINDING KELCH REPEAT F-BOX1* (*FKF1*) gene homologs, are also implicated in the induction of flowering under long day conditions (where daylength is greater than 12 hours), and are suspected of forming a complex with *GI* to influence *FTb2* expression directly, or through a yet-unproven mechanism (Wong et al. 2014; Ridge et al. 2016).

Other genes that respond to photoperiod and have proven expression in cool-season legumes include *Phytochrome A* (*PHYA*), and a number of circadian clock homologs, including members of the *EARLY FLOWERING* (*ELF*) family (Kong et al. 2010; Jaudal et al. 2020). In cool-season legumes, short days increase the expression of *ELF3* and *ELF4*, while *LUX ARRHYTHMO* (*LUX*) acts with these members of the *ELF* family in the suppression of *FT* homologs, and thus flowering (Hecht et al. 2011; Liew et al. 2014). In addition, work in lentil suggests expression of an *ELF4* homolog is adjusted in response to light quality (Yuan et al. 2021). Across cropped plant species, mutants with reduced expression of these genes can be useful, with reduced flowering suppression allowing faster development under short photoperiods (Laurie et al. 2009; Ridge et al. 2017). In lentil, the early flowering phenotype credited with the crop's expansion into northern India, denoted by an early allele at the *Sn* flowering locus, has been shown to be a result of a mutation at *ELF3* (Sarker et al. 1999b; Weller et al. 2012).

2.4.4 Temperature Pathway

In models for flowering in legumes, including those in lentil, the effect of temperature is generally considered a linear factor, acting as an amplifier of photoperiodic influences (Summerfield et al. 1985; Cober et al. 2014; Wright et al. 2020). Temperate legume crop vernalization responses to a cold period are contested, and in *Medicago*, responses are small and often masked by photoperiod interactions (Roberts et al. 1986; Jaudal et al. 2013; Putterill et al. 2013). Perhaps due to this complex relationship between temperature and plant development, the genetic pathway regarding ambient temperature and flowering has not received as much attention as the photoperiod pathway. In *Arabidopsis*, splice variants of *FLOWERING LOCUS M* (*FLM*) genes as well as the degradation of *SHORT VEGETATIVE STAGE* (*SVP)* transcripts, have been

found linked to temperature-dependant responses that adjust flowering (Lee et al. 2007; Lutz et al. 2015). While homologs for these genes are present in legumes, the *FLOWERING LOCUS C (FLC)* gene, integral to the integration of *FLM* and *SVP* in *Arabidopsis* through regulation of *FT*, is absent (Hecht et al. 2005; Kim et al. 2013). In *Medicago*, a homolog of *VERNALISATION2-LIKE* (VRN2) was found to contribute to flowering in response to temperature by regulating *FTa1* (Jaudal et al. 2016). Additionally, in long day-responsive legumes, intergenic regions have been found to regulate *FTa1* in multiple species, and *FTa1* expression is thought to be supressed by these regions until overridden by both temperature and photoperiod requirements (Jaudal et al. 2013; Nelson et al. 2017; Ortega et al. 2019; Rajandran et al. 2021).

2.4.5 Leveraging legume synteny to understand flowering in lentil

Shared synteny of the molecular pathways among temperate legumes makes leveraging information from other species effective for lentil investigations (Weller et al. 2012; Ridge et al. 2017; Rajandran et al. 2021). In regards to flowering time, pea has been the most extensively studied with over 20 loci known to control flowering (Weller and Ortega 2015). The model legume, *Medicago truncatula*, is also a valuable resource with access to a high-quality genome useful for tracing flowering time homologs between legumes and back to studies in *Arabidopsis* (Hecht et al. 2005; Kim et al. 2013). Currently, the best consensus useful for initial searches regarding flowering time gene homologs in lentil is the 'Curated Genes' track on the KnowPulse JBrowse (https://knowpulse.usask.ca/jbrowse/Lens-culinaris/2), with annotated lentil flowering homologs identified using flowering genes from *Arabidopsis*, *Medicago*, chickpea, and soybean.

Leveraging flowering gene resources has multiple applications and in chickpea, studies focusing on variation only around known flowering time genes from other legumes were successful, and identified flowering time QTL in diverse sets of accessions linked directly to candidate genes (Kujur et al. 2015a, 2015b; Upadhyaya et al. 2015). Conservation among crop legumes allowed for the identification of the gene responsible for the lentil flowering locus *Sn*, which was resolved through comparison to a pea flowering gene homolog, *HIGH RESPONSE TO PHOTOPERIOD* (*HR*) (Weller et al. 2012). Citing studies in related legumes, Yuan et al. (2021) were led to investigate a lentil *ELF4* homolog, *LcELF4a*, and found expression was altered in their light quality experiment.

It is likely that novel genetic pathways exist in each crop legume, relative to those in the most well studied temperate legumes *Medicago* and pea (Welch et al. 2005; Srikanth and Schmid 2011). This, coupled with the level of genetic resources available across legume species, including lentil, makes the verification of candidate flowering gene homologs following marker studies worthwhile.

2.5 Molecular mapping for candidate gene identification

2.5.1 Genetic Markers for Flowering in Lentil

Flowering time in lentil was first suggested to be under the control of a single major gene, *STERILE NODES (Sn)*, with a few minor loci applying additional regional adaptations (Sarker et al. 1999b). The early-recessive allele, *sn* was initially identified in crosses between the early-to-flower *pilosae* ecotype of South Asian and the South American cultivar, Precoz (Sarker et al. 1999b). However, diversity analyses using more broadly gathered lentil accessions have since indicated that multiple loci are likely important to the variation in flowering (Alo et al. 2011; Khazaei et al. 2016). Fratini et al. (2007) and Saha et al. (2013) each identified 3 QTLs linked to flowering (q.DTF) in different populations, and Kahriman et al. (2014) and Jha et al. (2017) each identified a single q.DTF with a large effect in their populations. Rajandran et al. (2021) suggest the *Sn* locus is likely segregating in all these populations, however the anonymous markers used are not conserved across all the populations and limit comparisons of QTLs among these studies. As the *Sn* locus has since been identified as a homolog of a gene responsible for sensing short photoperiods, *ELF3*, with mutants exhibiting reduced photoperiod sensitivity (Weller et al. 2012), studies on lentil flowering conducted in long day conditions (e.g., Temperate regions) have likely identified q.DTF not linked to *Sn.*

The multiple loci identified by Tullu et al. (2008) and Fedoruk et al. (2013) are, therefore, likely under the control of other genic regions. In support of this, field trials using a panel of diverse lentil, conducted in three major climatic regions, found DTF loci identified in short day trials did not play a significant role on DTF in Temperate trials of the same material (Neupane 2019). A single, stable q.DTF identified by Fedoruk et al. (2013) on LcChr1 was the first linked to a gene-based map. Using transcript data and an early version of the lentil genome assembly, pieces of two flowering gene homologs on LcChr1 were proposed as candidates for the variation captured by this QTL (Sudheesh et al. 2016). From another Saskatchewan field trial with a separate

recombinant inbred line (RIL) population, two q.DTF were identified on LcChr6 and, with the use of the most current lentil genome assembly (Lcu.2RBY), lentil homologs of *Flowering Locus T* (*FT*) were identified as likely candidate genes (Haile et al. 2021).

In greenhouse experiments conducted on two populations derived from a single cross of an early-to-flower, photoperiod insensitive line (ILL 2601) and a frequently used breeding line (ILL 5588), two flowering QTL were identified under both long day and short day conditions (Rajandran et al. 2021). The QTL of larger effect encompasses the same cluster of *FT* homologs, *LcFTa1*, *LcFTa2* and *LcFTc*, identified as likely candidates for one q.DTF identified in Saskatchewan field experiments (Haile et al. 2021). In addition, when examining contrasting light quality in an indoor experiment, Yuan et al. (2021) found a q.DTF covering this same region of *FT* homologs. Gene expression studies on these lentil *FT* homologs further implied their importance, especially *LcFTa1*, as genes to consider for selection tools (Rajandran et al. 2021; Yuan et al. 2021). The second q.DTF identified by Haile et al. (2021) was linked to a separate *FT* cluster, which included two *FTb* homologs as likely candidates, with further study into this region, including expression analysis, yet to be done.

2.5.2 SNP Markers for Linkage Mapping

Single Nucleotide Polymorphisms (SNPs) are becoming widely applied to crop species and are the most polymorphic of genetic markers (Rafalski 2002). Linkage mapping studies in lentil have used other types of markers, however none have widely facilitated marker assisted selection (MAS) in breeding, due to low coverage and being poorly conserved across populations (Tullu et al. 2008; Saha et al. 2010; Kahriman et al. 2014).

Genotype-By-Sequencing (GBS) has been employed in lentil to discover hundreds of thousands of SNP markers (Fedoruk et al. 2013; Sharpe et al. 2013; Kaur et al. 2014; Wong et al. 2015; Khazaei et al. 2016; Aldemir et al. 2017; Gela et al. 2021c). In genomes with high levels of intergenic repetition, like lentil (Ramsay et al. 2021), care has to be taken to ensure marker quality with extra steps required to ensure markers are widely relevant. The development of a marker panel for lentil by Sharpe et al. (2013) is one example of this. Exome capture, as a method of reduced representation sequencing is regarded as a highly informative alternative, especially in plants with large genomes such as *Lens* (Hirsch et al. 2014; Bohra and Singh 2015). There are a number of bioinformatically focused tools developed for this type of marker data

(DePristo et al. 2011; Ji and Chen 2016) making analyses more straight-forward and the most recent and dense linkage maps for lentil have been made using markers identified using some form of genome sequencing followed by a complexity-reduction step (Haile et al. 2021; Rajandran et al. 2021; Yuan et al. 2021).

2.5.3 Exome Capture Assay SNPs

The exome capture assay is designed to sequence coding regions (i.e., genes), to reduce genome-wide representation but retain useful and evenly distributed variation (Hirsch et al. 2014). The DNA is fragmented using sonication, and wet-lab steps select G-C rich regions prior to sequencing (SciGenom 2012). The sequenced reads are also filtered to select away from non-coding regions by eliminating highly repetitive sequences (Hirsch et al. 2014). Reads from the exome capture sequencing are aligned to a genome assembly, and has allowed SNP and Insertion/Deletion mutations (InDel) variation to be identified in lentil (Ogutcen et al. 2018).

Variation between accessions identified using an exome capture assay can be more useful than GBS markers, and reducing variation to only within coding regions has allowed for a direct flow into finer mapping and prediction of protein changes in multiple crops (Hirsch et al. 2014; Watson et al. 2015; Kassa et al. 2017). A risk with a trait such as days to flowering, however, is that valuable intergenic variation may be filtered away before analysis, as there may be mutations in conserved non-coding regions which would likely not be captured (Salvi et al. 2007). However, the large number of SNPs identified within coding regions is a source of opportunity for a flowering time analysis and within cultivated accessions there will be sufficient linkage disequilibrium to have SNPs associate with these kinds of variation. In a genome wide association study (GWAS) using exome-capture SNPs in switchgrass, Grabowski et al. (2017) identified multiple flowering time loci, one of which was linked to an *FT* homolog. Even more promising, exome capture sequencing was used to identify the functional mutation at a flowering gene in barely (Stein et al. 2014).

2.5.4 The FTa1-FTa2 Marker

The FTa1-FTa2 marker was developed as part of Dr. Rajandran's Ph.D. research conducted at the University of Tasmania (UTAS) under Dr. Jim Weller (Rajandran et al. 2021). F2-derived populations, made by crossing ILL 2601 with ILL 5588, were grown under short day (<12HR) greenhouse conditions. The populations segregated for days to flower (DTF) and a strong QTL

was identified on LG4 (corresponding to LcChr6) when using markers identified through DArT (Diversity Array Technology) sequencing (Sansaloni et al. 2011). An InDel was identified at the QTL and a Sequence Characterized Amplified Region (SCAR) marker was developed (called FTa1-FTa2 in this study). The wild-type allele produces a PCR product of about 450 bases. An alternate allele, the deletion, results in a product of ~200 bases. Under short day greenhouse conditions, the deletion (DEL) and heterozygous genotypes flowered earlier than the wild-type (WT) lines in their population. At F2, segregation for FTa1-FTa2 marker was consistent with a single gene with DEL dominant to WT (Rajandran et al. 2021).

Prologue to Chapter 3

In a preliminary study (Appendix 1), a diverse set of lentil lines were shown to have a large range in days to flowering (DTF) in greenhouse and field experiments in Saskatchewan. A previously developed biparental mapping population (LR-11), derived from parents with DTF at either end of the range observed among the panel of diverse lines, was available at University of Saskatchewan Crop Development Centre (CDC). Use of this population was expected to identify major loci controlling flowering time variation with the goal of developing molecular tools useful for screening diverse lentil lines.

Single Nucleotide Polymorphism (SNP) marker data from an exome capture assay of this population, aligned to the lentil genome v1.2 (LcV1.2; Ramsay et al. 2016), were given to me to develop a dense marker map for quantitative trait loci (QTL) analysis.

Chapter 3. Genetics of time to flowering in Saskatchewan for a south Asian (ILL 8006) by adapted (CDC Milestone) population

3.1 Introduction

Lentil is an important crop in western Canada and, in the last decade, annual production in Saskatchewan has rarely been matched by any other individual country (Canadian Grain Commission 2021; FAOSTAT 2021). Breeding has focused on improving agronomic and end-use/quality traits using the existing pool of adapted genetic material, with the occasional introgression using unadapted material (Tullu et al. 2013). Even though there are many desirable traits available in more diverse germplasm, crosses with diverse material has generally been avoided by Canadian breeders due to detrimental traits which are difficult to overcome (Khazaei et al. 2016). The largest constraint to using germplasm from different parts of the world is adaptations for flowering time (Erskine et al. 1989). In characterizing diverse lentil accessions, how plants reach development stages in response to multiple environmental pressures is tied to genetic background (Wright et al. 2020). Because of this, predicting days to flowering (DTF) in a single environment can be difficult when comparing plants with different backgrounds. Prediction has been shown to be improved however, by considering major flowering time homologs identified in other crop species (Cockram et al. 2007; Wenden et al. 2009; Bhakta et al. 2017).

A constraint of many previous marker studies in lentil has been marker anonymity and lack of marker-associations outside of the specific population in which they were identified (Saha et al. 2013). Maps employing markers readily comparable to gene-based maps have a higher likelihood of being broadly useful in diverse material and make maker-assisted selection (MAS) more reliable. This is becoming realized in lentil and some studies have implemented these types of markers for traits related to seed quality (Fedoruk et al. 2013; Verma et al. 2015; Jha et al. 2017) and micronutrient accumulation (Ateş et al. 2016, 2018; Aldemir et al. 2017).

A large factor contributing to the under-utilization of markers for screening DTF by breeding programs is that only a few mapping studies in lentil had used markers which were easily linked back to a gene-based map (Fedoruk et al. 2013; Kahriman et al. 2014; Haile et al. 2021). The use of gene-linked markers also lends itself to the leveraging of information from related species (Weller et al. 2012). Flowering genes are well conserved across legume species, making the extensive knowledge of molecular flowering pathways in pea (*Pisum sativum*) and *Medicago truncatula* and, to a lesser extent, soybean (*Glycine max*) valuable resources to lentil breeders (Kim et al. 2013; Weller and Ortega 2015).

The Recombinant Inbred Line (RIL) population LR-11 was employed in this study to identify regions of the genome contributing to variation in DTF under Saskatchewan field conditions. Heritability estimates for DTF using similar populations are often moderate, and DTF is generally considered a quantitative trait. Trials in Saskatchewan investigating DTF using crosses narrower than that of the LR-11 have yielded heritability estimates of 30-53% (Tullu et al. 2008; Fedoruk et al. 2013; Haile et al. 2021). Careful phenotyping, and the wide cross used to create LR-11, between a Canadian cultivar, CDC Milestone, and one developed for South Asia, ILL 8006, should result in a high heritability estimate for DTF. Using multiple phenology-related traits in conjunction with DTF will help provide context to any identified QTLs. Additionally, by having many informative, gene-based markers, this study is well positioned to identify practical QTLs readily testable on diverse material such as the AGILE-LDP (Chapter 4).

3.2 Materials and Methods

3.2.1 Plant Material

The LR-11 recombinant inbred line (RIL) population was created at the University of Saskatchewan (USask) by crossing the South Asian line ILL 8006 with the temperate variety

CDC Milestone and selfing through single seed descent from F2 to F8 before bulking individual lines. The population available for use in this study consisted of 120 RILs.

ILL 8006, also called BariMasur-4, was developed in Bangladesh from a cross between ILL 5888 and ILL 5782. ILL 5888 was a selection from a Bangladeshi landrace and, interestingly, is one of the first lines to flower in a diversity panel (AGILE-LDP; Appendix 1) in both South Asian trials, as well as in Saskatchewan trials (Wright et al. 2020). ILL 8006 is small seeded, with orangey-red cotyledons and a dotted gray seed coat (Sarker et al. 1999a; Kumar 2007). Neither ILL 5888 nor ILL 8006 are adapted to the northern temperate climate of Saskatchewan summers, and both suffer due to photoperiod insensitivity, and they flower very early under long days, an undesirable adaptation for Canadian environments.

The Saskatchewan-adapted parent, CDC Milestone, was developed by selecting a line from the cross of Eston x C8L27Y. C8L27Y was a breeding line used at USask which arose from a selection of the F5 generation in an F2-derived population created by crossing Eston and ILL 5588. ILL 5588 has been used widely in Saskatchewan breeding lines as a source of Ascochyta blight resistance. CDC Milestone is small seeded, with yellow cotyledons and a pale green seed coat with faint mottling (Vandenberg et al. 2001).

3.2.2 Phenotyping

The study was conducted at two sites in North-central Saskatchewan: Sutherland (GPS: 52.17, -106.51) and Rosthern (GPS: 52.68, -106.29). In 2017, the Sutherland site was seeded on May 4th and Rosthern was seeded on May 19th. In 2018, Sutherland was seeded May 9th and Rosthern was seeded May 11th.

The 120 RIL lines were sown in a randomized complete block design with three replications at each site. Each 1m^2 microplot was seeded with 120 seeds, provided there was enough seed from previous harvests to allow it. In 2017, LR-11-133 had only enough seed for two replications (Sutherland Rep 1 and 2) and was seeded at 60 seeds per plot. In 2018, both parents were included in the trial design, but in 2017 only CDC Milestone was included due to a lack of ILL 8006 seed.

The plots were visited every 1-3 days and phenotypes called on the day 10 % of the plot emerged (DTE), flowered (DTF), and matured (DTM). Vegetative Period (VegP) and Reproductive

Period (RepP) were calculated by taking the difference between DTF and DTE, and DTM and DTF, respectively. Node of Flower Development (NFD) was recorded at 10 % flowering by counting from the ground to the lowest node on the main stem with at least one partially open flower. Total plot seed yield (YLD) was collected on lightly cleaned samples following harvest and threshing. Days to 10 % of pods were swollen (DTSP) was recorded only in 2018. Raw phenotypic data used in this study can be found at https://knowpulse.usask.ca/Heidecker-Thesis-Association-Data; Supplementary 1.

3.2.3 Statistical Analysis

The years and locations of the field trials were combined and used as site-years. All statistical analyses were done using the software R v3.4.4. (R Development Core Team 2018). A linear mixed-model was fitted treating the replications (blocks) and site-years as random factors, while the genotypes were considered fixed. The "lmer" function from the lme4 package-v1.1-15 (Bates et al. 2015) was used to fit a linear mixed-model, with significance and variance components tested using the R package lmerTest v3.0-1 (Kuznetsova et al. 2017).

3.2.4 LR-11 Linkage Map Construction

All 120 lines and parents of the LR-11 population were genotyped using a custom lentil exome capture assay (Ogutcen et al. 2018). Using an in-house pipeline, reads were aligned to the lentil genome v1.2 (LcV1.2; Ramsay et al. 2016) to identify Single Nucleotide Polymorphisms (SNPs) which were then made available for this study.

Exome capture-SNP markers were filtered for those with less than 10% missing or heterozygous calls. Afterward, only markers which were also present in the AGILE-LDP genotyping results (Haile et al. 2020) were retained. This was done as another marker quality check, and ensures identified markers were not variants unique to the bi-parental population, and therefore more tractable for future studies. Markers mapping to unanchored contigs in the current assembly were included without the requirement of being found in the AGILE-LDP, provided only 1 line at most scored heterozygous, and less than 10 % of lines were missing an allele call. A total of 11,560 markers were used for linkage mapping.

SimpleMap (Jighly et al. 2015) was used to combine markers into bins of co-segregating markers prior to mapping. SimpleMap retains the repulsion information within each bin which allows the markers to be re-introduced for analysis following mapping. In the input file, markers were

ordered based on the number of missing and heterozygous lines, with markers mapping to contigs placed at the bottom of the file, so that the markers mapped to chromosomes with the largest number of informative scores would be chosen to represent the bin first. Repulsion was set to 0 so only markers scoring the same across all 120 lines, or those only different due to missing data, would be combined into a single bin. This resulted in 2433 marker-bins being used for linkage mapping.

MapDisto2.0 (Heffelfinger et al. 2017) was used to create the linkage map. A LOD of 8 resulted in 6 Linkage Groups, with only 1 marker-bin with 2 markers being unlinked and thus removed. Double recombinants were identified and replaced with missing data after each round of ordering/rippling. The final map contained 1,448 uniquely mapping marker-bins, representing a total of 11,558 markers. The order of these 1448 marker-bins can be found at https://knowpulse.usask.ca/Heidecker-Thesis-Association-Data; Supplementary 2.

3.2.5 QTL Analysis

Quantitative Trait Loci (QTL) analysis was conducted in QTL IciMapping (Meng et al. 2015). Interval mapping (IM) and Inclusive Composite Interval Mapping (ICIM) were used with the individual phenotypes for each site-year, as well as using the combined LSMeans across all site-years. Mapping parameters were set to 0.50 centimorgans. One thousand permutation tests were run to determine the LOD threshold value of 3.5 which was used to declare significant QTLs.

3.2.6 DTF Candidate Gene Identification

The positions in LcV1.2 of markers at QTL peaks were compared to known flowering-time gene homologs using the Curated Genes track of the lentil JBrowse on KnowPulse (https://knowpulse.usask.ca/jbrowse/Lens-culinaris/1).

3.2.7 *FT* homolog expression in a selection of lines segregating for main flowering QTLs

Plant tissue of the parents and nine RILs segregating for the three most significant DTF QTLs (q.DTF) was harvested from block 1 at Sutherland on June 8th, 2018, over four weeks after sowing (Table 3.1). Plants were sampled at the 6-8 node stage as this is close to the timepoint found to be the most useful for identifying flowering gene expression in lentils in controlled experiments (Ortega 2018, pers. comm.). Samples were collected multiple hours after sunrise, with the final plant sampled within twenty minutes of the first, and before 11:00AM. The top-most, fully open leaf was taken from two plants and combined for each biological replication,

with two biological replications harvested for each line tested. RNA was extracted using the Qiagen RNEasy Plant MiniKit (QIAGEN, Hilden, Germany), following the standard protocol for plant tissue. An Agilent 2100 Bioanalyzer (Agilent, Santa Clara, California, USA) was used to check for RNA quality and quantity.

Lines were screened for gene expression using primers for the lentil *Flowering Locus T* (*FT*) flowering gene homologs, *LcFTa1* and *LcFTb2* as well as the housekeeping gene *LcTIF* (Rajandran et al. 2021). Reverse Transcription PCR required 1 µg of RNA, 4 µL of 5x buffer, 1 µL each of dNTPs and oligio dTs, 0.5 µL Invitrogen SuperScript™ IV Reverse Transcriptase (ThermoFischer Scientific, Wilmington, Delaware, USA), with dDH_2O to make a final volume of 20 µL. The PCR protocol was a single step, with an annealing temperature of 42°C for 30 min and a cleavage step of 85°C for 5 min. The 20 µL samples of concentrated cDNA were diluted to a final volume of 100µL.

Table 3.1. LR-11 RILs chosen for *FT* homolog expression. Lines were chosen as representatives based on allelic state of 3 identified QTL for DTF (q.DTF). Tissue was taken from block 1 of Sutherland 18 (S18), 30 days after sowing. Days to emergence (**DTE**) and flower (**DTF**) for the plot are reported for reference.

Entry	q.DTF.6-1/2/3*	DTE	DTF
CDC Milestone	A/A/A	13	46
LR-11-153	A/A/A	13	46
LR-11-24	A/B/A	13	40
LR-11-94	A/A/B	12	39
LR-11-47	A/B/B	15	43
LR-11-154	B/A/A	15	43
LR-11-58	B/B/A	16	43
LR-11-122	B/A/B	12	42
LR-11-20	B/B/B	15	42
LR-11-15	B/B/B	14	42
ILL 8006	B/B/B	19	43
*allele at q.DTF: **A** = CDC Milestone, **B** = ILL 8006			

FastSYBR® Green (ThermoFischer Scientific, Wilmington, Delaware, USA) was used for fluorescence detection and sample setup followed the standard protocol. The qPCR program was run on a CFX 384 BioRad (ThermoFischer Scientific, Wilmington, Delaware, USA): 40 rounds of 95°C for 15s, 95°C for 3s, and 60°C for 30s with plate fluorescence quantified every round. This procedure was followed by a melt curve analysis (as a quality check) in which temperature

was increased from 65°C to 95°C, with fluorescence quantified every 5s as temperature increased by 0.5°C. Three technical replicates were run for each biological sample, generating six datapoints for each line. Expression of the flowering homologs (*LcFTa1*, *LcFTb2*) were evaluated by comparing to the housekeeping gene *LcTIF* (Appendix 2).

3.3 Results

3.3.1 Phenotyping

Distributions of each phenology-related trait were similar across site-years. The distribution for days to emergence (DTE) was the tightest, especially in Rosthern 2017 (R17) where the last plots emerged only four days after the first plots (Figure 3.1). Mean correlations for the individual traits across site-years ranged from r=0.31 to r=0.84 (Table 3.2).

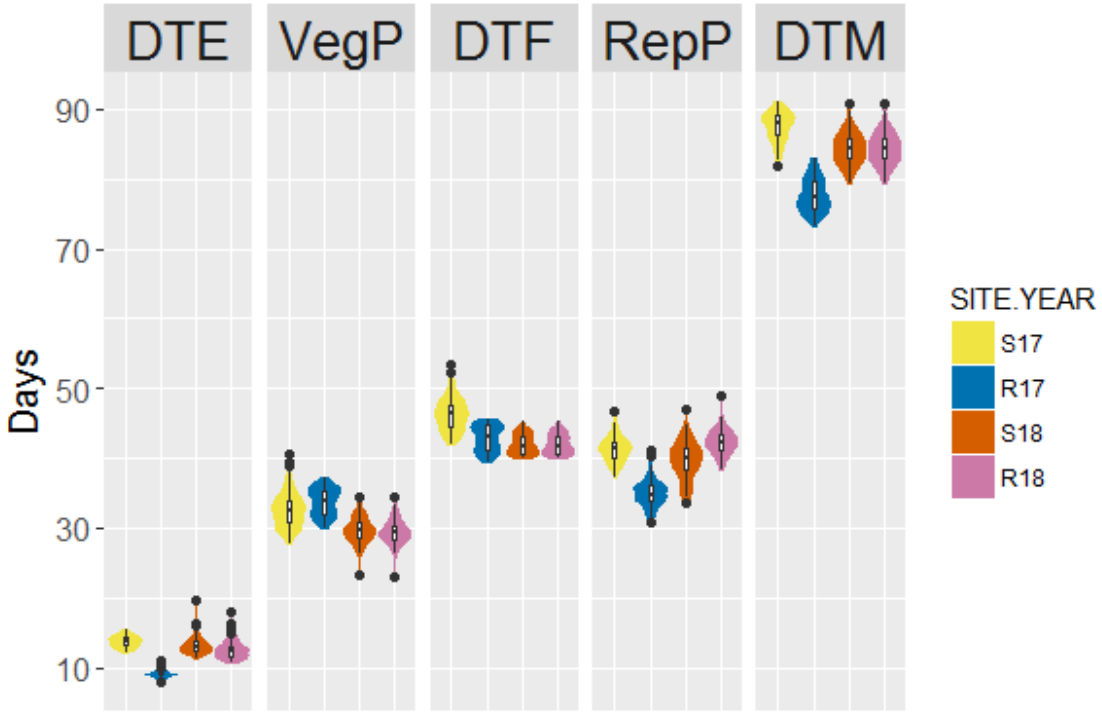

Figure 3.1. Distribution of phenotypic data for phenology-related traits in a recombinant inbred population (LR-11: ILL 8006 x CDC Milestone) across 4 site-years; days to: emergence (DTE), flowering after emergence (VegP), flower (DTF), maturity after flowering (RepP) and days to maturity (DTM). Site-years were Sutherland 2017 (S17), Rosthern 2017 (R17), Sutherland 2018 (S18), and Rosthern 2018 (R18).

Heritability estimates for most traits were moderate to high, with most estimates over 0.50 (Table 3.3). Exceptions to this were RepP (0.29) and NFD (0.14). VegP had the highest heritability of

all traits (0.81) and was higher than either of the phenotypic values from which it was derived - DTE and DTF.

Table 3.2. Correlation matrix for phenotypic traits measured on a recombinant inbred population (LR-11: ILL 8006 x CDC Milestone) across 4 site-years. Significance was determined using Pearson's correlation coefficient.

	Site-Year	DTE	VegP	DTF	RepP	DTM	YLD	DTSP[1]
DTE	S17	**0.42**	-0.33***	-0.02^{NS}	-0.01^{NS}	-0.03^{NS}	-0.38***	
	R17		-0.33***	-0.08^{NS}	-0.08^{NS}	-0.12^{NS}	-0.23*	
	S18		-0.65***	-0.03^{NS}	-0.05^{NS}	-0.05^{NS}	-0.25**	0.01^{NS}
	R18		-0.58***	0.25**	0.11^{NS}	0.24**	-0.18*	-0.09^{NS}
VegP	S17		**0.84**	0.95***	-0.49***	0.68***	0.54***	
	R17			0.97***	-0.17^{NS}	0.63***	0.50***	
	S18			0.78***	0.36***	0.57***	0.48***	0.61***
	R18			0.64***	-0.13^{NS}	0.34***	0.44***	0.56***
DTF	S17			**0.79**	-0.52***	0.71***	0.45***	
	R17				-0.20*	0.63***	0.46***	
	S18				0.43***	0.72***	0.43***	0.81***
	R18				-0.06^{NS}	0.58***	0.35***	0.56***
RepP	S17				**0.31**	0.24**	0.12^{NS}	
	R17					0.63***	0.29**	
	S18					0.94***	0.43***	0.56***
	R18					0.72***	0.25**	0.06^{NS}
DTM	S17					**0.64**	0.61***	
	R17						0.60***	
	S18						0.50***	0.75***
	R18						0.42***	0.43***
YLD	S17						**0.53**	
	R17							
	S18							0.31***
	R18							0.32***
DTSP[1]								**0.57**

Traits reported**:** days to emergence (DTE), flowering (DTF), swollen pod (DTSP) maturity (DTM), as well as Vegetative Period (VegP), Reproductive Period (RepP), and yield (YLD).
Bolded values are the mean correlation of trait to itself across all site-years tested.
*****p < 0.001 , **p<0.01, *p<0.05, ^{NS}p > 0.05** ; [1]2018 only

Table 3.3. Estimates of variance components from "lmer" modeling and broad-sense heritability for phenology-related traits in LR-11 across four site-years. A random intercept model was used.

	DTE	VegP	DTF	RepP	DTM	NFD	YLD	DTSP[1]
Genotype	0.446	3.171	2.301	1.123	3.869	0.064	2619.936	1.764
Genotype*Site-Year	0.075	0.358	0.418	1.349	1.019	0.109	1459.036	0.318
Environment	1.245	1.202	1.149	4.260	4.282	0.813	3154.606	3.040
Heritability (H^2)	0.48	0.81	0.74	0.29	0.61	0.14	0.51	0.57

Traits reported: days to emergence (DTE), flowering (DTF), swollen pod (DTSP) maturity (DTM), as well as Vegetative Period (VegP), Reproductive Period (RepP), Node of Flower Development (NFD), and yield (YLD).
[1]2018 only

The linear mixed model results indicate no significant interaction between Genotype and Site-Year for DTE (Table 3.4). The interaction was significant for all the other traits, however, so it was necessary for site-years to be kept separate in later analyses. This was particularly important for QTL mapping due to the different levels of influence the environment had across the site-years.

Table 3.4. Results summary for linear mixed models fit using "lmer" for phenology-related traits in LR-11 across four site-years. Genotype (Entry) was used as fixed effect while interactions with Site-Year (Genotype*Site-Year) and Site-Year/Replication were included as random effects.

	Significance Test	n	DTE	VegP	DTF	RepP	DTM	NFD	YLD	DTSP[1]
Genotype	*F* [A]	122	4.31*	16.99*	12.46*	2.6*	7.31*	1.69*	5.14*	3.65*
Genotype* Site-Year	*LRT* [B]	485	2.64^{NS}	57.77*	78.36*	63.75*	40.89*	15.28*	1459.04*	3.81*
Environment		4	1.24*	22.39*	12.2*	18.88*	17.98*	11.36*	3154.61*	0.1*

Traits reported: days to emergence (DTE), flowering (DTF), swollen pod (DTSP) maturity (DTM), as well as Vegetative Period (VegP), Reproductive Period (RepP), Node of Flower Development (NFD), and yield (YLD).
[1]2018 only; [A]F-value and test of Fixed Effect [B]Likelihood Ratio Test of Chi-square values -- Random Effects
***p < 0.001, [NS]p > 0.05**

3.3.2 LR-11 Linkage Map Construction

A large group of co-segregating markers associated with reference genome LcV1.2 chromosomes 2 (LcChr2) and 3 (LcChr 3) were pseudolinked so only 6 linkage groups (LGs) rather than the expected 7 were generated. Other than this region on linkage group 2 (LG2), markers generally grouped into LGs which match the LcV1.2 pseudomolecules (chromosomes). The complete linkage map was 1604.3 cM long with an average of 1.1 cM between uniquely

mapping bins. The marker-bin density by LG ranged from 1.33 cM per uniquely mapping bin, on LG1, to 0.82 cM per unique marker-bin, on LG7 (Table 3.5).

Table 3.5. Linkage map summary for the recombinant inbred population LR-11 (ILL 8006 x CDC Milestone). Data is presented both for the entire map **(LR-11)** but is also presented based on each linkage group (**LG**). There were 6 linkage groups, which corresponded to chromosomes of the reference genome assembly (LcV1.2), with the exception being LG2, which was made up of markers from both chromosome 2 and 3 (LcChr2 and LcChr3). For referring to LcV1.2, LG1=LcChr1, LG2 = LcChr2/3, LG3=LcChr4, LG4=LcChr5, LG5=LcChr6 and LG6=LcChr7

LG	No.Markers	No.Unique	cM	Max.cM	Med.cM	Max. Dup	Med. Dup
LR-11	**11558**	**1448**	**1604.26**	**40.69**	**0.50**	**389**	**3**
LG1	1195	173	230.28	25.99	0.52	228	3
LG2	4043	480	560.80	31.08	0.50	156	2
LG3	1887	230	261.01	14.49	0.50	82	3
LG4	1135	149	156.28	17.32	0.51	173	2
LG5	1562	254	261.90	40.69	0.49	87	3
LG6	1736	162	133.99	5.34	0.49	389	3

Results reported: number of markers (**No.Markers**), number of unique marker-bins (**No.Unique**), linkage group size in centimorgans (**cM**), maximum distance between two linked markers in cM (**Max.cM**), median distance between two linked markers in cM (**Med.cm**), maximum number of markers mapping to a single bin (**Max.Dup**), and median number of markers mapping to a single bin (**Med.Dup**).

3.3.3 QTL Identification

Phenology-related QTLs were found on 4 of the 6 linkage groups (Table 3.6). Many of the confidence intervals for different traits overlapped when considering the IM interval, which allowed them to be placed into QTL Groups. In total, there were 13 phenology-related QTL groups: 2 on LG1, 5 on LG2, 5 on LG5 and 1 on LG6. The only QTLs which were found in all site years fell into the QTL group LcChr6-A1, located at the top of LG5. The QTL located in this region were for VegP, DTF, and DTM and, respectively, these QTL explained a minimum of 33.0%, 16.5% and 10.3% of the phenotypic variation on a site-year basis. This region explained 52.7% (VegP), 51.9% (DTF) and 36.8% (DTM) of the phenotypic variation for these same traits when mapping the LSMeans generated using all 4 site years together indicating this locus has a strong effect on development of the LR-11 RILs under Saskatchewan conditions. Figure 3.2 highlights the relative stability of these three main loci on LG5 and plots the LOD scores using the LSMeans for DTF, VegP, DTM and RepP.

QTLs for DTE (q.DTE) were present on LG2 and LG6. Due to the limitations of LG2 caused by the markers of LcChr2 and LcChr3 co-segregating, the confidence intervals for q.DTE on LG2

are large (as much as 30 centimorgans), even though they explain 20 - 31% of the variation in any given site-year. The late allele for DTE is being provided by ILL 8006 and not CDC Milestone, unlike the rest of the phenology-related traits where the late allele is provided by CDC Milestone (Table 3.6).

Gross plot yield (YLD) QTLs were only identified when mapping using the LSMeans and were always located within the confidence intervals of QTLs for phenology-related traits with major effects.

Table 3.6. QTL summary for phenology-related traits for the recombinant inbred population LR-11 (ILL 8006 x CDC Milestone). Site-Year indicates in which site-year the Quantitative Trait Loci (QTL) reached significance. **Confidence Interval** is the map position (in centimorgans (**cM**)) on the linkage group. If a QTL for a trait was significant in multiple Site-Years where confidence intervals between site years partially overlapped, the Confidence Interval reported here includes the entire region. **LOD** is the peak logarithm of the odds score for the QTL. Percent Variation Explained **(PVE)** is the amount of the variation explained by the locus at the peak of the QTL, estimated by ICIM. **Source** indicates the parental source of the later or allele of higher numerical value (MLN = CDC Milestone, BM4 = ILL 8006). Days to Flowering (DTF) QTL (q.DTF) have been bolded.

LG	QTL Group	Trait	QTL Name	Site-Year	Confidence Interval (cM)	LOD	PVE (%)	Source
1	LcChr1-A	DTSP	qDTSP.1-1	S18	62.5-63.5	5.1	8.8	MLN
		DTM	qDTM.1-1[a]	R17, R18, S18	64.5-76.5	3.6-5.4	7.0-13.7	MLN
		RepP	qRepP.1-1[a]	S18	52.5-54.5	5.7	12.6	MLN
	LcChr1-B	VegP	qVegP.1-1	S17	205.5-215.5	6.7	6.8	MLN
		DTF	**qDTF.1-1[a]**	**S17, S18**	**203.5-222.5**	**4.2-5.1**	**6.0-6.5**	**MLN**
2	LcChr2/3-A	VegP	qVegP.2/3-1	R18	67.5-68.5	3.8	7.5	MLN
	LcChr2/3-B1	**DTF**	**qDTF.2/3-1**	**S17**	**176.5-183.5**	**3.5**	**3.7**	**BM4**
	LcChr2/3-B2	**DTF**	**qDTF.2/3-2**	**R18**	**209.5-210.5**	**6.7**	**16.7**	**BM4**
	LcChr2/3-C1	DTE	qDTE.2/3-1[a]	R18	260.5-260.5	5.3	20	BM4
		VegP	qVegP.2/3-2	S17, S18	271.5-273	5.2-10.7	8.3-10.3	MLN
		DTF	**qDTF.2/3-3**	**S17**	**266.5-268.5**	**3.7**	**4**	**MLN**
	LcChr2/3-C2	DTE	qDTE.2/3-2	R17, S17, R18	328.5-358.5	7.6-9.4	25.0-31.1	BM4
5	LcChr6-A1	VegP	qVegP.6-1[a]	All Site Years	2.5-5.5	13.8-32.8	33.0-59.3	MLN
		DTF	**qDTF.6-1[a]**	**All Site Years**	**2.5-6.5**	**6.6-27.9**	**16.5-61.3**	**MLN**
		DTSP	qDTSP.6-1	S18	4.5-7.5	14.7	30.1	MLN
		DTM	qDTM.6-1[a]	All Site Years	2.5-7.5	4.2-13.4	10.3-31.9	MLN
		NFD	qNFD.6-1	R17	2.5-6.5	7.3	20.9	MLN
	LcChr6-A2	DTSP	qDTSP.6-2	R18	19.5-20.5	3.6	12.3	MLN
		RepP	qRepP.6-1[a]	S17, R18	20.5-22, 42.5-43.5	4.7	16.44	MLN
		DTM	qDTM.6-2[a]	S17, R17, R18	20.5-22, 42.5-44.5	5.7-9.1	14.6-17.9	MLN
	LcChr6-B	VegP	qVegP.6-2	S17, S18	86.5-89.5	4.1-5.3	4.5-6.3	MLN
		DTF	**qDTF.6-2[a]**	**R17, S17, S18**	**87.5-89.5**	**3.7-7.1**	**4.8-10.4**	**MLN**
	LcChr6-C	VegP	qVegP.6-3	R17	110.5-111.5	4.1	5.2	MLN
		NFD	qNFD.6-2[a]	R17	103-105	6.0	16.7	MLN
	LcChr6-D	VegP	qVegP.6-4[a]	S17	259.5-261	4.5	3.8	MLN
		DTF	**qDTF.6-3[a]**	**S17, S18**	**258.5-259.5**	**5.0-10.0**	**5.5-14.9**	**MLN**
		DTSP	qDTSP.6-3	S18	260.5-261	7.7	13.9	MLN
6	LcChr7-A	DTE	qDTE.7-1[a]	S17	24.5-28.5	3.6	10.6	BM4

Traits with reported QTLs: days to emergence (DTE), flowering (DTF), swollen pod (DTSP) maturity (DTM), as well as Vegetative Period (VegP), Reproductive Period (RepP), Node of Flower Development (NFD)
[a]QTL Confidence Interval shared by trait LSMeans generated QTL

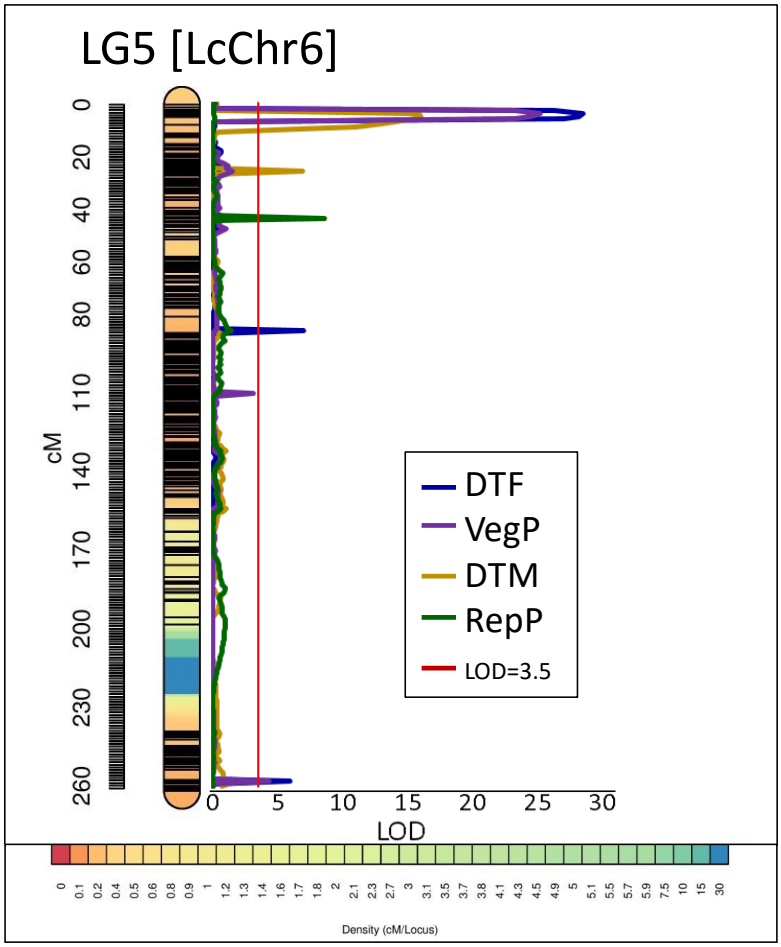

Figure 3.2. Linkage group 5 (LG5) and QTL for days to flower and other development traits. LG5 is shaded based on marker density in cM/Locus. To the right of the linkage group, LOD score is plotted at each locus based on the LSMeans (four site years) of days to flower (DTF), vegetative period (VegP), days to maturity (DTM) and reproductive period (RepP). A LOD threshold, shown in red, of 3.5 was determined using a 1000-permutation test.

3.3.4 Candidate Gene Identification

Markers within the major q.DTF, q.DTF.6-1 correspond to the top of LcChr6 (LcV1.2), with only a few markers within the confidence interval mapping outside of LcChr6:0-3Mbp region. In this region, there is one curated flowering gene homolog, *LcFTb1*/*LcFTb2* (http://knowpulse.usask.ca/portal/jbrowse/Lentil; Lc27932a), identified as a candidate gene for consideration in another field study conducted in Saskatchewan (Haile et al. 2021).

The two additional QTL groups which each harbor a QTL for DTF on LG5: q.DTF.6-2 and q.DTF.6-3, each explain over ten percent of the phenotypic variation in at least one site-year (Table 3.6). q.DTF.6-2 contains markers mapping to a 24 Mb region on LcChr6 (from 138 Mbp to 162 Mbp), and includes a cluster of florigen genes, *LcFTa1*, *LcFTa2* and *LcFTc*, that were recently implicated in signalling the change from vegetative to flowering in lentil (Rajandran et al. 2021; Yuan et al. 2021). DTF.6-3 includes markers mapping between both 107-109 Mbp, and 205-208 Mbp on LcChr6 and contains four known flowering gene homologs: *LcLFY*, *LcDCL1*, *LcATX3b* and *LcPRR95c*.

3.3.5 Expression of candidate genes for DTF in LR-11

In a selection of lines segregating at the three most significant q.DTF, outlined in Table 3.1, higher expression of *LcFTb2* was generally found in lines with the ILL 8006 allele at q.DTF.6-1 (Figure 3.3). *LcFTb2* expression in CDC Milestone and LR-11-153 (CDC Milestone allele at all loci tested) was significantly lower than expression in all other lines sampled. The three RILs with the ILL 8006 allele at LcChr6-A1(LR-11-20, LR-11-15, and LR-11-58) expressed *LcFTb2* at a level statistically equivalent to ILL 8006, which had the highest expression of *LcFTb2*.

Possession of the CDC Milestone allele at q.DTF.6-1 was not associated with expression of *LcFTa1*, the gene found within the confidence interval of q.DTF.6-2, supporting the assertion that, in the selected lines at least, these two loci are in fact two separate genetic factors. Instead, the marker scores underlying q.DTF.6-2 (Table 3.1) are better for grouping material when considering *LcFTa1* expression. Apart from LR-11-94, all lines which possessed the CDC Milestone allele at q.DTF.6-2 expressed *LcFTa1* at the same level as CDC Milestone. Possession of either parental allele at q.DTF.6-3 did not fit an observable trend in *LcFTb2* or *LcFTa1* expression.

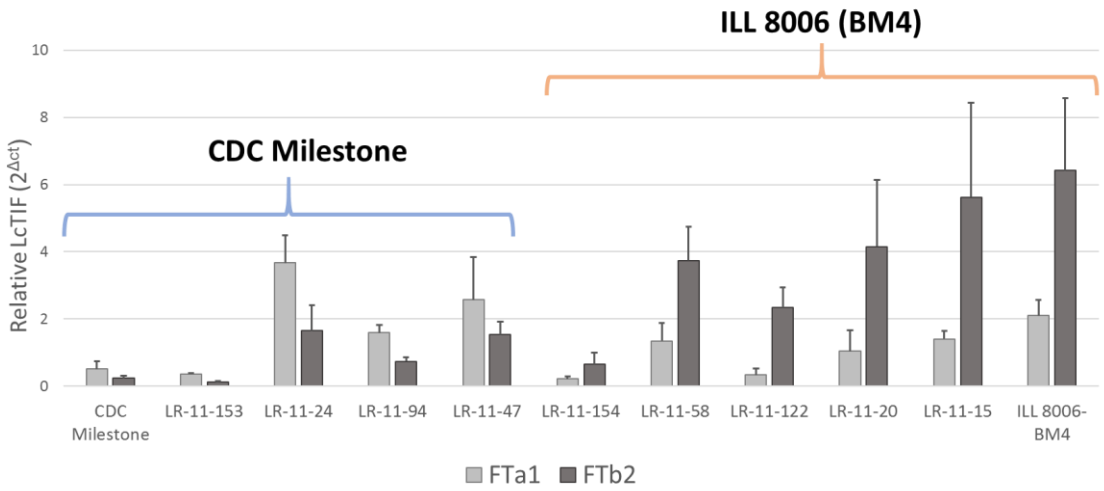

Figure 3.3. RNA expression of *LcFTa1* and *LcFTb2* from leaf tissue in a subset of LR-11 lines. Leaf tissue was collected from replication 1 of the Sutherland 2018 experiment. Lines are ordered based on similarity to parental lines based on scores at three q.DTF on LG5 (LcChr6) and grouped based on marker variant at q.DTF.6-1 (QTL group LcChr6-A1). Additional information related to this figure can be found in Appendix 2.

3.4 Discussion

3.4.1 LR-11 grown in Saskatchewan field trials generated valuable phenotypic data for phenology-related traits

The LR-11 population was used because it was made by crossing an adapted-to-western Canada line and an unadapted South Asian line, and the different development responses these parents show in the western Canadian environment make it a good population for dissecting the requirements for flowering time in this region. Other studies, similarly using populations made using parents adapted to two climatic regions, have identified important loci responsible for controlling plant development, with a focus on flowering, and linked to success an environment (Sarker et al. 1999b; Haile et al. 2021).

Heritability (H^2) estimates, calculated using the variance components (Table 3.3), for most components of plant development in this study were moderate to high, comparable to other long day field studies in lentil (Tullu et al. 2008; Haile et al. 2021), indicating they could be used to search for quantitative trait loci (QTL). The relationships among these phenological traits, which were stable across site-years in this experiment, were expected to provide context when investigating the underlying genomic regions. While most traits were significantly correlated with one another, even across site-years (Table 3.2), days to emergence (DTE) was only significantly correlated with flowering (DTF) and maturity (DTM) in one site-year and was

never correlated with reproductive period (RepP) or days to swollen pod (DTSP). Vegetative period (VegP) and RepP were calculated and investigated as phenological traits because previous studies in lentil have found them useful for parsing out plant development, especially when looking at the environmental factors which help drive flowering and maturity (Roberts et al. 1988; Wright et al. 2020). Gross plot seed yield (YLD) was significantly correlated with all phenology-related traits in almost all site-years, consistent with other studies which recommend using these traits useful for selecting yield (Tullu et al. 2011; Fedoruk et al. 2013). Longer VegP and DTM were most highly associated with higher YLD, but early emergence was also associated with higher yields across multiple site-years.

In contrast, RepP, along with Node of Flower Determination (NFD), had low heritability in this study, making them poor candidates for QTL identification, with any QTL identified likely of little use to breeders. This is supported by Wright et al. (2020), where RepP showed little variability across a large number of lentil accessions grown in multiple environments. Instead, it is the development prior to flowering which is important for lentil fitness.

3.4.2 Exome-capture derived SNP markers are useful for linkage and QTL mapping of phenological traits, with additional value derived from access to the lentil genome.

With 1448 uniquely mapping markers across 1604.3 cM, this is one of the denser lentil linkage maps for a bi-parental population, with a similar total map size comparable to other SNP-based marker maps using two cultivated parents (Ateş et al. 2016, 2018; Haile et al. 2021). Linkage group 5 (LG5) has a large (40.7 cM) gap partially due to filtering criteria that markers also be found polymorphic in the AGILE-lentil diversity panel (AGILE-LDP). This trade-off was necessary to ensure that markers used in this study had a high chance of being useful in diverse lentil germplasm, and readily available to a breeding program. Even with densely covered maps, conservation of markers outside of mapping populations is often cited as an impediment to their utility (Tullu et al. 2011; Kumar et al. 2021).

Multiple variations of marker clean-up and linkage map construction parameters (not included) always resulted in 6 linkage groups, rather than the 7 that were expected based on the lentil genome and previous linkage mapping studies. The linkage groups each correspond to a single lentil genome pseudomolecule (LcChr), except for LG2, where markers mapping to LcChr2 and LcChr3 were inseparable. Other linkage maps in lentil have had similar problems, and it is

suspected that a translocation in one parent relative to the another is the cause (Ramsay et al. 2021; Yuan et al. 2021). It is common to have linkage map issues due to suspected translocations in populations where cultivated lentil is crossed to a wild parent (Ramsay et al. 2021; Yuan et al. 2021), and an analysis of marker diversity by Sharpe et al. (2013) found ILL 8006 was almost as divergent from Canadian lentils as wild *Lens* genotypes. Due to the reduced recombination associated with suspected translocations, QTL found in these regions cover large confidence intervals in the linkage map and end up being associated with many markers in more than one genomic region, making candidate gene searches fruitless.

The most important q.DTF , q.DTF.6-1, was co-located with QTL for other phenology-related traits on LG5, which corresponds to LcChr6 of the LcV1.2 genome. The peak consisted almost entirely of markers located within 1-3 Mbp on LcChr6. Within this region are two nearly identical lentil homologs of *FTb*: *LcFTb1* and *LcFTb2*, a gene that integrates responses from multiple pathways in response to long photoperiods in pea (Hecht et al. 2011). Under Saskatchewan field conditions, where lentil is exposed to long days (daylength greater than 12 hours) as soon as it emerges, a loss of flowering regulation control regarding photoperiod is expected to hasten flowering. A separate lentil population grown in Saskatchewan had a q.DTF in this same region, with the un-adapted parent also providing early flowering under the long days (Haile et al. 2021). Early to flower mutants in pea show a reduced long day requirement and express *FTb2* transcripts at a higher level (Hecht et al. 2011; Ridge et al. 2016), and in the expression analysis on a subset of LR-11 lines, the lentil *LcFTb2* gene was found to be expressed at a higher level in the RILs which possessed the ILL 8006 allele at q.DTF.6-1 than those with the CDC Milestone allele (Figure 3.3).

The secondary q.DTF on LG5, q.DTF.6-2 and q.DTF.6-3, each overlapped genomic regions previously identified in an indoor study conducted by Rajandran et al. (2021). They found the loss of an intergenic region near *LcFTa1* and *LcFTa2*, two genes found within q.DTF.6-2, was linked to differences in the expression of these genes, and consequently DTF. In another bi-parental population, Yuan et al. (2021) also found a q.DTF, which overlaps q.DTF.6-2, containing *LcFTa1*, *LcFTa2* and *FTc.* They found expression of *LcFTa1* was altered under different light quality regimes (red to far red ratio) leading to changes in DTF depending on the allelic state and light quality regime. Investigating *LcFTa1* expression in a selection of LR-11

lines suggests the state of the allele at q.DTF-2 may be important to the final DTF phenotype (Table 3.6). *FTa1* and *FTa2* along with *FTc* are implicated in flowering time across long day legumes, and are suspected to integrate signals from both photoperiod and temperature (Weller and Ortega 2015; Ortega et al. 2019; Thomson et al. 2019), making them good candidates for further investigation into what is happening at q.DTF.6-2. Homologs of four flowering genes are found within q.DTF.6-3: *LcLFY* , *LcDCL1*, *LcATX3b* and *LcPRR95c*, although thus far, they are relatively under-studied in legumes grown under long day field conditions. The markers identified within q.DTF.6-3 map to two regions on the LcV1.2 assembly, 107-109 Mbp and 205-208 Mbp, but in the *Medicago truncatula* assembly, these four gene homologs are all found close to one another. This suggests a possible mistake in the LcV1.2 reference assembly, or perhaps there is structural variation between the parents on LcChr6. The latter explanation would support the large 40.7cM gap found in LG5. In studying the genes associated with their q.DTF, Rajandran et al. (2016) identified a suspected frame shift mutation at *LcPRR95c* in their early-to-flower parent that has yet to be further explored. Additionally, a pea homolog of *PRR95c* was noted as a candidate of interest following a gene expression study in a long day photoperiod (Ridge et al. 2016), suggesting it may be a plausible candidate for q.DTF.6-3.

q.DTF.1-1 was only significant at Sutherland, nevertheless it consistently explained at least 6% of phenotypic variation in each site-year. While q.DTF.1-1 is more site-year specific than the QTL on LcChr6, this QTL may warrant further investigation as markers there map close to a q.DTF in another lentil RIL population grown under Saskatchewan field conditions (Fedoruk et al. 2013) which was linked to expression of a flowering gene homolog (likely *GIGANTEA* (*GI*)) in a separate population (Sudheesh et al. 2016). *GI* is a good candidate to start with for this q.DTF as *GI* has been found to help regulate *FT* genes in pea by integrating photoperiod signals (Hecht et al. 2007), and may contribute to different sensitivities to daylength which are suspected to exist in this population.

3.4.3 Days to Flower (DTF) in LR-11 may be the result of two growth phases under distinct genetic control

In a controlled environment study, four distinct phases of lentil plant development were identified (Summerfield et al. 1985) pre-emergence (equivalent to DTE) and pre-inductive, inductive, and post-inductive phases (collectively equivalent to VegP). Understanding lentil

development prior to flowering as a sum of these phases is useful for acknowledging distinct genetic controls. Regardless, in evaluating diverse material in both greenhouse and field experiments, the entire period from planting to flowering has also been effectively explained as a single phase (days to flowering; DTF) when accounting for genetic sensitivities to temperature and photoperiod (Roberts et al. 1988; Erskine et al. 1994a; Wright et al. 2020). Relatively low variability, likely compounded by strong environmental cues, in days to emergence (DTE) relative to the vegetative period (VegP) has contributed to it being overlooked, and DTF has been sufficient for identifying QTL for developmental differences in other Saskatchewan field trials involving bi-parental crosses (Tullu et al. 2008; Fedoruk et al. 2013; Haile et al. 2021).

In this study, while a statistically significant environmental term of the linear mixed model was identified for DTE (Table 3.4), an insignificant Genotype*Site-Year interaction was also returned, which suggests the environment had a generally stable effect across genotypes and was not causing differences in rank in DTE. This supported the use of DTE as an independent trait for analysis. Using DTE independent of VegP adds a facet to explaining the observed differences in DTF between the parents of LR-11 (Figure 3.4).

The early flowering parent, ILL 8006, emerged seven days later, on average, than CDC Milestone at both Sutherland and Rosthern in 2018 but ILL 8006 always reached flowering slightly earlier by having a severely reduced VegP: about 11 days shorter at both 2018 sites. When considering DTE and VegP together (DTF), there is no transgressive segregation as CDC Milestone was one of the earliest lines to emerge and had the longest VegP, while ILL 8006 was always the last to emerge yet had the shortest VegP. DTE and VegP for LR-11 individual RILs were negatively correlated within site-years ($r = -0.33$ to -0.65, $p<0.001$) and followed the trend established by the parents, with later emerging lines tending to be those with a shorter VegP. In Saskatchewan, material with growth habit matching CDC Milestone, i.e., earlier emerging, and later flowering, is preferred. Lines from this cross that combine a habit of early emergence (CDC Milestone-like) but also possess alleles for hastened flowering (ILL 8006-like) may be useful in South Asian climates by maximizing VegP in a short season. In this event, markers for both DTE and VegP/DTF would be useful.

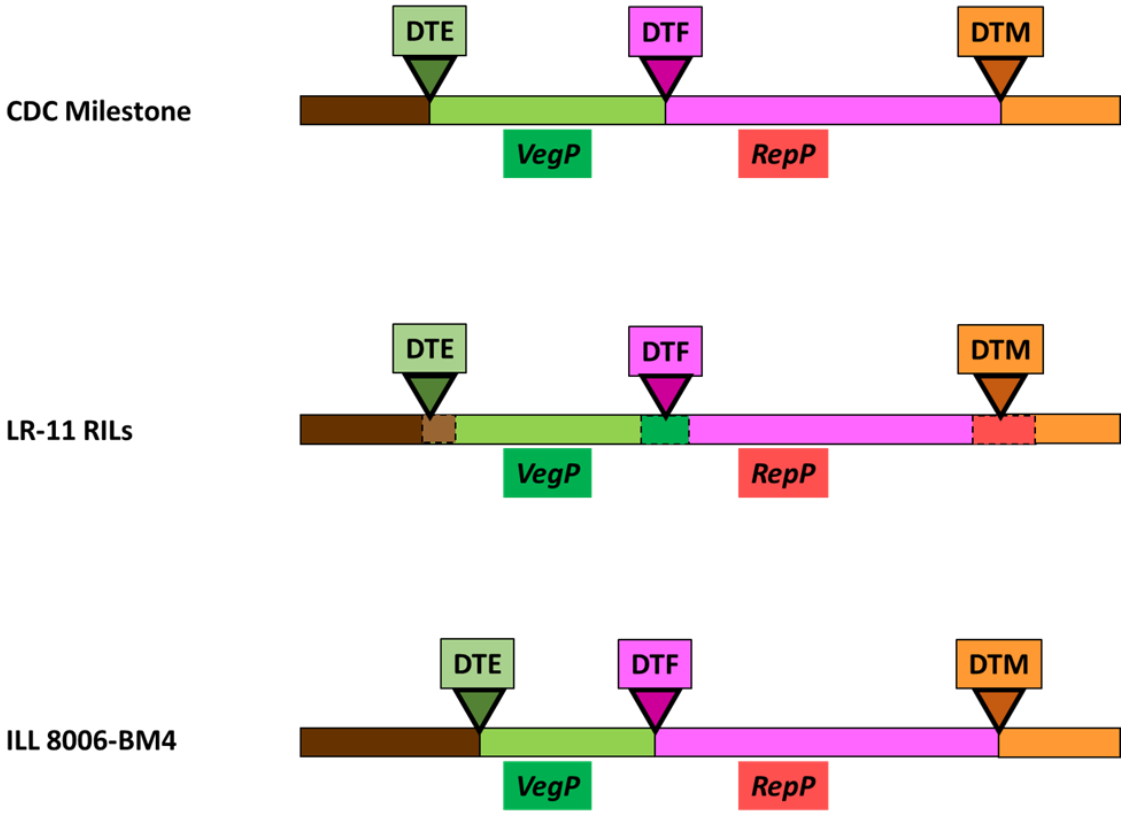

Figure 3.4. Representation of phenological progression for LR-11 parents and RILs. Important stages noted along the route to harvest from planting include days to: emergence (DTE), flower (DTF) and pod maturity (DTM). The time from emergence to flowering (Vegetative period, *VegP*) and from flowering to maturity (Reproductive period, *RepP*) have also been considered important indices along plant development in lentil. Arrows indicate the event taking place, with the hatched segments on **LR-11 RILs** timeline used to indicate there is phenotypic variation for that trait between members of the population.

The lack of correlation between DTE and DTF, yet consistent correlation of DTE with VegP could be attributed to unaccounted environmental variation (e.g., sowing depth, soil temperature, and soil moisture) as DTE was shown to have only moderate heritability, at 0.48, compared to the higher heritability of VegP, (0.81) and DTF (0.74). DTE and VegP likely do share some genetic control, and QTL for DTE, VegP and DTF were found co-located in the region LcChr2/3-C1. Genes selected as DTE candidates in legumes include those that are induced by growth regulators and temperature changes and would play a role in development beyond emergence (Dias et al. 2011; Weller and Ortega 2015).

Of the seven q.DTF which were identified (Table 3.6), all but two (q.DTF.2/3-1 and q.DTF.2/3-2) had a confidence interval which overlapped with a q.VegP and/or q.DTE. This contributes doubt on the importance of these two QTL and they were each identified in only a single site-

year. In comparison, a q.VegP for at least one site year, was co-located at all the q.DTF with larger effects, i.e., those found on LG5.

In this population, the most important contributors to DTE and VegP are loci which are distinct from one another. No q.DTE were located on LG5 and, while the early alleles for VegP on LG5 come from ILL 8006, the early allele at q.DTE on LG2 was contributed by CDC Milestone. A stable q.DTE, q.DTE2/3-2, located on LG2 explains much of the observed variation (25-30% depending on the site-year), yet no q.DTF were identified within the confidence interval. The identified q.DTE should be further investigated as a potential tool for dissecting the differences in development that contribute to DTF. In addition to the large confidence intervals for q.DTE on LG2, the implication that there may be translocations in this population relative to the reference line CDC Redberry, adds further complications to candidate gene searches, and markers map to wide swaths of both LcChr2 and LcChr3. Ramsay et al. (2021) found that this issue is tied to reduced recombination in these areas, increasing linkage drag and makes selections within the areas difficult. Markers identified in this study would require extensive evaluation on other material without this issue, to refine the implicated region to one where gene searches would be useful.

3.4.4 *LcFTb2* is a candidate gene for q.DTF.6-1 in need of further investigation

FTb2 is a floral promoter, typically implicated under long day conditions in both pea and *Medicago*, and is believed to integrate signals mostly from the photoperiod pathway (Hecht et al. 2011; Weller and Ortega 2015; Thomson et al. 2019). In this study, the lentil homolog of this gene was found within q.DTF.6-1 and, is a strong candidate for explaining some of the differences in phenology-related traits observed under the temperate Saskatchewan field conditions. It is surmised that RILs possessing the ILL 8006 allele at *LcFTb2* could be earlier flowering due to a change which was allowing *LcFTb2* to be expressed to a higher degree sooner than lines with the wild-type (CDC Milestone) allele. In long day, indoor studies in lentil and related legumes, *FTb2* expression increases around two weeks before flowering, before the other main floral pathway integrator *FTa1* (Hecht et al. 2011; Ortega et al. 2019; Thomson et al. 2019; Rajandran et al. 2021). In a subset of LR-11 individuals which included the parents and 9 RILs segregating independently for the 3 q.DTF on LG5, CDC Milestone and LR-11-153, two lines which had the wild-type allele at all three loci, had statistically lower levels of *LcFTb2* expressed

than all lines possessing the ILL 8006 allele at q.DTF-6-1 (Figure 3.3, Appendix 2). All lines with the ILL 8006 allele at q.DTF-6-1 expressed *LcFTb2* at levels numerically higher than all lines possessing the CDC Milestone allele. While these results are not entirely definitive, there is enough of an observable trend to warrant further investigation. One possible explanation for the increased *LcFTb2* expression (relative to CDC Milestone) in LR-11-24 and LR-11-47, two lines which possess the wild-type allele at q.DTF.6-1 but ILL 8006-allele at q.DTF.6-2, may be due to the observed increased expression of *LcFTa1* (Figure 3.3, Appendix 2). These two *FT* orthologs were first shown to act on one another in pea (Hecht et al. 2011). Since then, *FTb2* and *FTa1* have been shown to work in tandem in the integration of molecular signals and, along with *FTc*, are required for flowering to occur in legumes (Laurie et al. 2011; Weller and Ortega 2015; Ortega et al. 2019) The role of a candidate gene at q.DTF.6-3, the *PRR95c* homolog, is also suspected to influence flowering in conjunction with *FTb2* expression (Ridge et al. 2016; Rajandran et al. 2021), however allelic state q.DTF.6-3 was not clearly linked to expression of either *FT* homolog here. Like many genes related to flowering, tissue sampling at the correct time is important for capturing *LcFTb2* expression (Hecht et al. 2011; Ortega 2018), and future studies may find value in sampling tissue at multiple timepoints.

3.5 Conclusions

The RIL population LR-11 possesses sufficient diversity for identifying useful markers for potential use by breeders to screen potential germplasm, and accurate phenotyping coupled with using multiple phenology-related traits instead of just DTF, helped to identify markers, and in some cases candidate genes, associated with control of flowering time under Saskatchewan growing conditions.

A major q.DTF was identified at the top of LG5, associated with a candidate gene *LcFTb2*. The genetic variation associated with this locus may be important to selecting material from other climatic regions that may also be readily adapted to Saskatchewan conditions. The two other main q.DTF also contained homologs for flowering genes important to crop legumes, namely *LcFTa1*, *LcFTa2*, *LcFTc* in q.DTF.6-2, and *LcPRR95c* in q.DTF.6-3 (Weller and Ortega 2015; Ridge et al. 2016; Ortega et al. 2019). Use of markers for all three q.DTF may help identify pathways for discrete adjustments to phenology-related traits.

Even though DTF is usually sufficient for identifying the most important flowering-related QTL, more genetic contributors can be identified by investigating DTF as the sum of sowing to emergence (DTE) and emergence to flowering (VegP). DTE was found to be under genetic control separate to that of VegP. A significant q.DTE was found on LG2, however a large confidence interval and possible structural variation between the parents near this site inhibited the identification of any candidate genes.

The most significant outcome of this study was the identification of q.DTF.6-1. Further investigation into the candidate flowering gene, *LcFTb2*, may help determine the genetic differences that are causing adjustments to DTF under warm, long day growing conditions. Markers from this study are readily testable on additional diverse material because of the filtering criteria that they also be found among the members of the AGILE-LDP (a panel first reported in Haile et al. (2020)). In the subsequent chapter, the AGILE-LDP is screened with a marker from the peak of q.DTF.6-1 to determine whether the genetic control seen here is consistent across diverse material, or just unique to this population.

Prologue to Chapter 4

In the following chapter, two markers were tested for their ability to account for days to flowering (DTF) under long day experiments involving the lentil diversity panel, AGILE-LDP, and a bi-parental RIL, LR-11. The AGILE-LDP phenology data from the greenhouse were collected by one or more of Derek Wright, Sandesh Neupane and myself. The four site-years of AGILE-LDP field experiments were phenotyped by Derek Wright, Sandesh Neupane and myself and appear as a subset of the dataset in (Wright et al. 2020). The phenology data for LR-11 are described in chapter 3 of this thesis.

The FTa1-FTa2 marker was identified in an F2-F3 population, L25, grown under greenhouse conditions, and reported by Rajandran et al. (2021). This marker was run on the AGILE-LDP as part of my undergraduate thesis project, and lines were found to score either wild-type (n=250) or deletion (n=74). I included this data from my undergraduate thesis as it helps with the discussion of how the markers can be used outside their original populations. The other marker, LcChr6-A1, arose from the work described in Chapter 3 of this thesis.

Chapter 4. Utility of available markers for predicting DTF in long day experiments

4.1 Introduction

Considering diverse lentil from the major production regions, individuals can be categorized genetically based on fitness to one of three broad climatic regions: ‘northern Temperate’, ‘Mediterranean’, and ‘South Asian’ (Khazaei et al. 2016; Wright et al. 2020). The distinct environmental cues lentils are exposed to in these environments make adaptations for development related traits, specifically days to flowering (DTF), different between climatic regions (Sarker et al. 1999b). Predicting how lentil from one region will respond from one environment to the next is difficult, as mechanisms behind the phenological response may be divergent (Ferguson et al. 1998; Lombardi et al. 2014; Khazaei et al. 2016; Wright et al. 2020).

Having molecular markers and understanding how they relate to the phenological responses causing differences in development would improve selection of material to use in Saskatchewan. In theory, markers for days to flowering genes and surrounding regions could be used to reduce linkage drag of undesirable alleles (Sharma et al. 2013) and improve evaluation by providing a

selection tool for screening prior to large scale field testing (Tullu et al. 2013; Dikshit et al. 2015). Markers specifically for time-to-flowering have been identified in lentil, however none are being used in the USask breeding program, largely due to them having been identified in low marker density linkage maps with variants unique to the population (Tullu et al. 2008; Fedoruk et al. 2013) or markers being identified in other climatic regions (Ateş et al. 2016; Aldemir et al. 2017).

Recently, Rajandran et al. (2021) identified a candidate genomic region which explains variation for days to flowering under short days in specific lentil populations. This marker may prove valuable to Saskatchewan breeders, and the cross used for mapping involved ILL 5588, a line used in multiple CDC varieties for its disease resistance (Tar'an et al. 2003), and an early flowering landrace prevalent in South Asia (ILL 2601). The flowering gene this marker is expected to account for, *LcFTa1*, has since been implicated in two more studies, one of which was a Saskatchewan field experiment (Haile et al. 2021) and the other an indoor experiment under long days (Yuan et al. 2021). It also falls in the same region as q.DTF.6-2 from chapter 3. Evaluating the bi-allelic marker (FTa1-FTa2) associated with this gene on lentil genotypes with diverse backgrounds, grown under long day field conditions, will help determine if it will be useful to the USask breeding program.

As reported in Chapter 3, a major QTL for flowering time was identified in the recombinant inbred line (RIL) population, LR-11, when grown in Saskatchewan field conditions. This marker, LcChr6-A1, was shown to account for a large amount of the variability in DTF in LR-11 but has not been tested to measure its ability to account for DTF widely across accessions of different backgrounds. The expectation is that, as they have both been implicated in SK field studies, LcChr6-A1 and FTa1-FTa2 will each account for loci relevant to DTF in SK field conditions due to independent flowering genes. This study sets out to achieve this by evaluating these two markers on the diverse group of accessions found in the AGILE-LDP.

4.2 Materials and Methods

4.2.1 Phenotypic Data

The days to flowering (DTF) data used in this study were taken from the greenhouse (GH) and field datasets for the AGILE-LDP (https://knowpulse.usask.ca/Heidecker-Thesis-Association-

Data; Supplementary 3) as well as the LR-11 field phenotypes (Chapter 3; https://knowpulse.usask.ca/Heidecker-Thesis-Association-Data; Supplementary 1).

4.2.2 Genotyping

DNA extractions of the AGILE-LDP accessions used leaf tissue from two-week-old plants grown in the Agriculture Greenhouse at USask using a Qiagen Plant DNeasy MiniKit (QIAGEN, Hilden, Germany). DNA extraction of the LR-11 population is described in section 3.2.4.

4.2.2.1 Screening with FTa1-FTa2 marker for DTF

Rajandran et al. (2021) identified a QTL of strong effect for days to flowering in short day (daylength set to 12 hours) greenhouse conditions, which maps to a region containing multiple *Flowering Locus T* (*FT*) gene homologs in the lentil reference genome, LcV1.2. They identified an intergenic deletion between two of these homologs, *LcFTa1* and *LcFTa2*, with lines possessing the deletion allele expressing higher levels of *LcFTa1* and flowering earlier. A Sequence Characterized Amplified Region (SCAR) marker for this deletion (FTa1-FTa2) was provided by Dr. Weller (University of Tasmania) to be used in this study. The marker mix added to each sample included two primer forward (F1:TGGGCTTGATACTTTGTACTCC and F2: TCTACACACTTTGCTGGTTTTG) and one common reverse primer (R: CCATCACAATTCAAAGCAATG). A master mix was made up for the PCR reactions so that that 2.5 µL of 10x Buffer, 1.0 µL 10 mM dNTPs, 1.0 µL of the working primer, 0.6 µL of 50 mM MgCl2 and 0.25 µL Genscript® (5 unit/µL) Taq Polymerase and 18.5 µL of Ultradistilled water would be added to each well on a 96-well plate. This working primer was added to 1.25 µL of the 75 ng/µL DNA sample. The PCR program began with 5 min of annealing at 94°C, followed by 40 cycles of 94°C for 45s , 58°C for 45s, and 72°C for 45s. Once the 40 cycles had completed the temperature was held at 72°C for 5 min before being cooled to 8°C, where it was kept until taken off the PCR machine.

There are four potential outcomes to this PCR reaction which can be observed when run on a gel: a larger strand of amplified DNA (~450bp) (wild-type), a smaller amplified strand (~200bp) (deletion), heterozygous, or fail. Both AGILE-LDP and LR-11 were screened, with individuals receiving scores of either wild type (WT) or deletion (DEL).

4.2.2.2 Exome Capture Assay on AGILE-LDP

Exome capture library preparation and variant detection in lentil is described in Ogutcen et al. (2018). The 324 accessions of the AGILE-LDP were genotyped using this array and the genetic population structure of the set was determined by Ogutcen et al. (unpublished). For the purposes of this experiment, each accession was placed in one of nine groups, three of which are admixtures. Groups were named based on where the majority of the accessions in that group originated. The 324 accessions were grouped as: Americas/Europe (Am/Eu) 1 [n=32], Am/Eu 2 [n=67], Am/Eu Mixed [n=39], Africa [n=21], Asia 1[n=35], Asia 2[n=41], South Asia [n=32], Asia Mixed [n=14] and Mixed [n=43] (https://knowpulse.usask.ca/Heidecker-Thesis-Association-Data; Supplementary 3).

A marker present in the exome capture arrays of both AGILE-LDP and LR-11 (LcChr6:2410977) was chosen to represent the peak of the major QTL for flowering identified in LR-11 (LcChr6-A1; see section 3.3.3). This SNP marker, located at the peak of q.DTF.6-1, was chosen because it possessed the least amount of missing data across the two populations. This marker is expected to account for variation in DTF due to its proximity to a lentil *FTb* homolog. Based on their effect in LR-11, the supposed early allele is the 'A' variant, found in ILL 8006, while the 'G' variant, found in the LR-11 parent CDC Milestone, is the suspected late allele.

4.2.3 Statistical analysis

For marker FTa1-FTa2 in the AGILE-LDP, naïve marker testing was done using a Wilcox test rather than t-tests, as the variation for DTF in the two marker groups was found to be unequal.

Mixed linear model testing of both putative DTF markers was done using "lmer" from the *lme4* package (Bates et al. 2015) where the marker was fit as a fixed effect and the experimental parameters, Entry, Site-Year and Block, were included as random terms. The ability of the marker data to improve prediction of days to flowering was evaluated based on the model fitness relative to a null model, with differences reported using Akaike's Information Criterion (AIC).

4.3 Results

4.3.1 Evaluation of the previously identified marker, FTa1-FTa2 in diverse populations

Of the 324 AGILE-LDP accessions, 74 possess the deletion (DEL) and 250 the wild type (WT) allele when screened with the FTa1-FTa2 marker. The mean days to flowering (DTF) for the DEL group is shorter than the lines possessing the WT allele for all growing locations (Table

4.1). Wilcox tests indicate that the means of the two groups are significantly different in each experiment, suggesting the marker accounts for at least some of the observed differences in mean DTF between the two marker groups (Table 4.1A).

Table 4.1. Mean and Standard Deviation (StDev) for days to flowering (DTF) in the AGILE-LDP, in greenhouse and field experiments, grouped by marker score at FTa1-FTa2. Wilcox tests, used to determine differences between allele groups within each experiment/site-year, all rejected the null hypothesis (H_0) that there are no differences in mean between the two allelic states.

	Greenhouse						Field							
AGILE-LDP	Fall 2015		Winter 2016		Fall 2016		Sutherland 2016		Sutherland 2017		Rosthern 2016		Rosthern 2017	
FTa1-FTa2	DEL	WT	DEL	WT	DEL	WT	DEL	WT	DEL	WT	DEL	WT	DEL	WT
Mean	34.82	39.84	37.59	42.11	34.41	37.58	47.47	50.07	46.89	50.34	45.06	46.96	42.74	45.24
StDev	5.17	7.07	4.46	6.38	3.97	6.44	2.17	3.74	3.20	4.69	1.44	2.72	2.62	3.47
Wilcox*	p=1.471e-07		p= 5.379e-09		p=5.349e-05		p=5.33e-08		p=2.993e-08		p=2.524e-08		p= 3.068e-08	

*Wilcox Test, H_0=no difference in mean between the two groups

Mixed linear modeling, comparing FTa1-FTa2 scores to a null model (a base model with no additional grouping), results in a reduced Akaike Information Criterion (AIC) for both the greenhouse and field experiments (Table 4.2). This supports the Wilcox test results, indicating that the FTa1-FTa2 marker is accounting for at least some of the observed differences in DTF in the AGILE-LDP.

For LR-11, ILL 8006 and 55 RILs carry the DEL allele while CDC Milestone and the other 65 RILs are WT for the FTa1-FTa2 marker. Using FTa1-FTa2 scores in LR-11 also results in an improved model for DTF compared to a null, suggesting this marker is useful in explaining differences in DTF in the bi-parental RIL population (Table 4.2). Plotting the distributions for DTF based on the FTa1-FTa2 alleles shows plenty of overlap between the two allelic states, indicating much of the variation is left unexplained by only using FTa1-FTa2, however (Figure 4.1B).

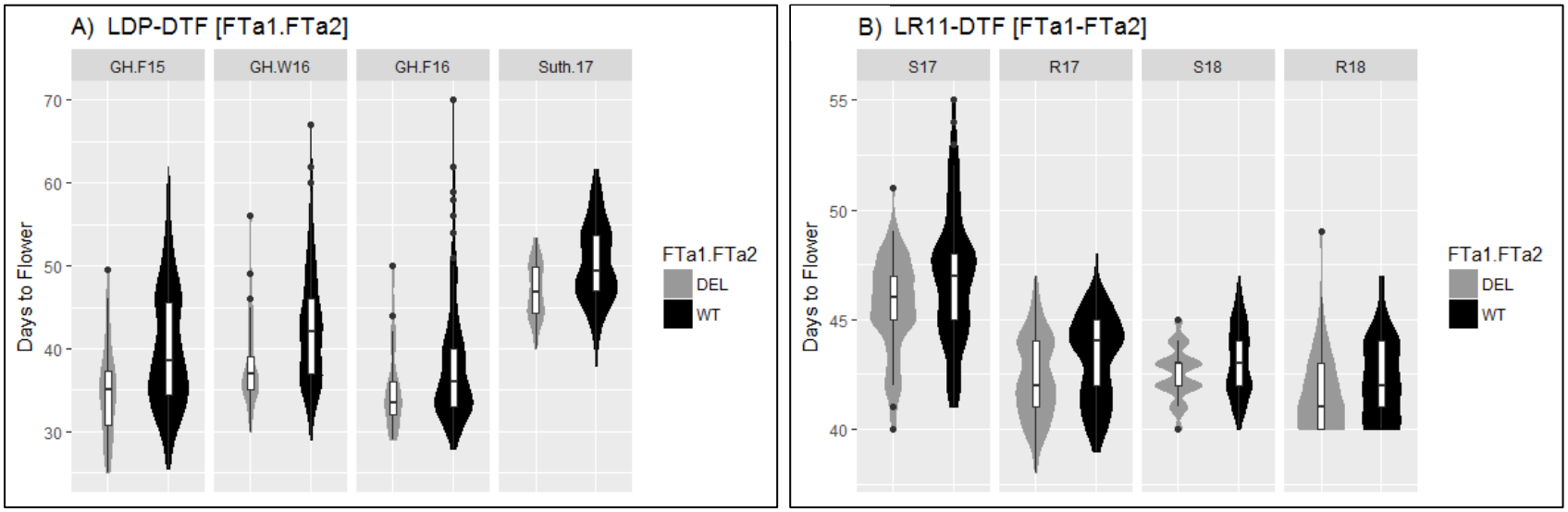

Figure 4.1. Distribution of days to flowering (DTF) in the AGILE-LDP and LR-11 based on allelic state of the FTa1-FTa2 marker. A) AGILE-LDP accessions possessed the deletion (DEL, n=74) or wild-type (WT, n=250) allele. Days to flowering (DTF) is plotted for the greenhouse (GH) in three experiments: Fall 2015 (F15), Winter 2016 (W16) and Fall 2016 (F16) and in the field Sutherland 2017 (Suth.17). **B)** At the FTa1-FTa2 marker, LR-11 individuals scored as deletion (DEL; ILL 8006; n=55), or wild-type (WT; CDC Milestone; n=65). Shown here are the four site-years of field experiment: Sutherland 2017 (S17) and 2018 (S18) and Rosthern 2017 (R17) and 2018 (R18). The mean DTF of the two allelic states in each experiment is marked with the line in the center of the interior boxplot, with the distributions shown in the shaded violin plots.

4.3.2 Evaluation of LcChr6-A1 marker in diverse populations

In the LR-11 population, 47 RILs possess the ILL 8006 variant ('A' SNP) and 71 possess the CDC Milestone variant ('G' SNP) at the marker used to represent the QTL group LChr6-A1. Two lines had a result of 'missing data' (https://knowpulse.usask.ca/Heidecker-Thesis-Association-Data; Supplementary 1). In the AGILE-LDP, 91 accessions possess the 'A' allele, 229 the 'G' allele, and 4 score heterozygous (H) (Figure 4.2A; https://knowpulse.usask.ca/Heidecker-Thesis-Association-Data; Supplementary 3).

Models which include the candidate markers FTa1-FTa2 or LcChr6-A1 show a reduction in AIC compared to the null in all experiments (Table 4.2), suggesting both markers are accounting for differences in DTF in both LR-11 and the AGILE-LDP. By using both FTa1-FTa2 and LcChr6-A1 in a single model, AIC was reduced further than either marker individually. This suggests the best model is the one which includes both markers and supports that the FTa1-FTa2 and LcChr6-A1 are accounting for discrete variation.

Table 4.2. AIC values from model testing using the markers FTa1-FTa2 and LcChr6-A1 for days to flowering (DTF) versus a null to compare marker fitness in the bi-parental RIL, LR-11 and the diversity panel, AGILE-LDP. Significance testing, determined by Chi-square test, indicates the models using FTa1-FTa2 and LcChr6-A1 separately, and together, were superior to a null model (no marker), suggesting the markers are accounting for differences in DTF. While minimal, the consistent reduction in AIC in models for DTF which includes both markers, over models using markers individually, suggests it may be the best model. Appendix 3 shows the estimated effect of the allelic state FTa1-FTa2 and LcChr6-A1 for each model.

	LR-11		**AGILE-LDP**		
	df	Field	df	Greenhouse	Field
Null	6	5055.9	4	5831.9	5393.9
FTa1-FTa2	7	5048.7**	5	5803.0***	5359.4***
LcChr6-A1	7	4977.0***	6	5789.4***	5339.4***
Both Loci	8	4962.0***	7	5770.7***	5318.4***
***p<0.001, **p<0.01					

4.3.3 Marker inheritance of FTa1-FTa2 and LcChr6-A1

Thirty-nine LR-11 RILs possess the CDC Milestone parent allele at both FTa1-FTa2 and LcChr6-A1 (wild-type and 'G' allele (**WT.G**)) and 22 RILs score the same as ILL 8006 at both loci (deletion and 'A' allele (**DEL.A**)). 32 RILs match ILL 8006 at FTa1-FTa2 but CDC

Milestone at LcChr6-A1 (**DEL.G**) and 25 RILs score the same as CDC Milestone at FTa1-FTa2 and ILL 8006 at LcChr6-A1 (**WT.A**) (Figure 4.2B).

In the AGILE-LDP, 190 accessions have the late allele at both loci (WT.G). There are 4 accessions which were heterozygous ([H]) at the LcChr6-A1 marker and all of these were wild-type (WT) for FTa1-FTa2. The remaining 56 accessions that score WT for FTa1-FTa2 possess the ‘A’ allele at LcChr6-A1 (WT.A). The accessions with the deletion (DEL) allele are split into groups with the ‘G’allele at LcChr6-A1 (DEL.G, n=39) or the 'A’ allele (DEL.A, n=35) (Figure 4.2D)

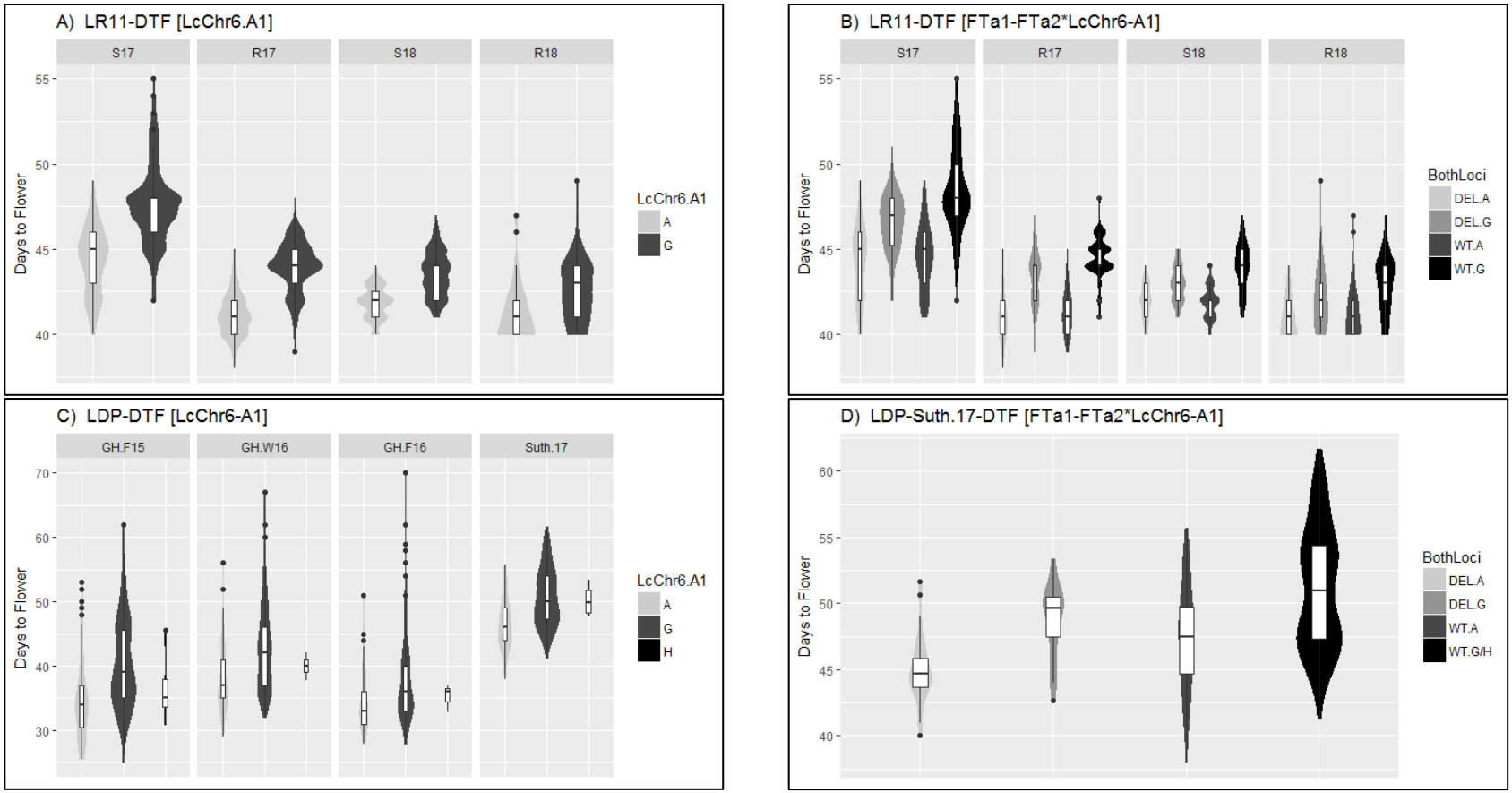

Figure 4.2. Distribution of days to flower (DTF) in the LR-11 and AGILE-LDP populations grouped by allelic state in field and greenhouse experiments. LR-11 four field site-years: Sutherland 2017 (S17) and 2018 (S18) and Rosthern 2017 (R17) and 2018 (R18). **A)** LR-11 individuals grouped by the LcChr6-A1 marker as 'A'-SNP (ILL 8006; n=47), or 'G'-SNP (CDC Milestone; n=71) and two excluded for missing scores. **B)** Grouping LR-11 by both FTa1-FTa2 and LcChr6-A1 results in a DEL.A group (ILL 8006; n=22), a DEL.G group (n=32), a WT.A group (n=25) and at WT.G group (CDC Milestone; n=39) with two lines excluded due to missing scores. **C)** The AGILE-LDP experiments shown here include the three greenhouse experiments Fall 2015 (F15), Winter 2016 (W16) and Fall 2016 (F16) and the field experiment Sutherland 2017 (Suth.17) and accessions grouped based on LcChr6-A1: as 'A'-SNP (n=91), 'G'-SNP (n=229) or heterozygous (H; n=4). **D)** Plotting only the Sutherland 2017 field experiment (Suth.17), the AGILE-LDP accessions are grouped by scores at the FTa1-FTa2 marker as well as at LcChr6-A1 resulting in four possibilities based on possession of DEL.A (n=35), DEL.G (n=39), WT.A (n=56) or WT.G/H (n=194).

4.3.4 Differences in marker inheritance in the AGILE-LDP refined using population structure

Combining the scores of both FTa1-FTa2 and LcChr6-A1, along with the population structure data that were provided, allowed additional investigation into control of DTF in the AGILE-LDP. The results above (4.3.2 and 4.3.3) lead to the designation of the 'A' allele at LcChr6-A1 and the deletion (DEL) at FTa1-FTa2 as the 'early' alleles and the 'G' allele at LcChr6-A1 and the wild-type (WT) at FTa1-FTa2 as the 'late' alleles. Carrying the early allele at one locus does not predict the allele state at the other, i.e., they are independent. Just over one third of the AGILE-LDP population (85 accessions) carry an early allele at one locus and a late allele at the other. A population structure analysis of the AGLIE-LDP separates the panel into nine groups, largely based on geography (Ogutcen et al. unpublished). The allelic state FTa1-FTa2 and LcChr6-A1 within the structure groups (Figure 4.3) suggest these two markers may be separately accounting for climatic region-specific adaptations.

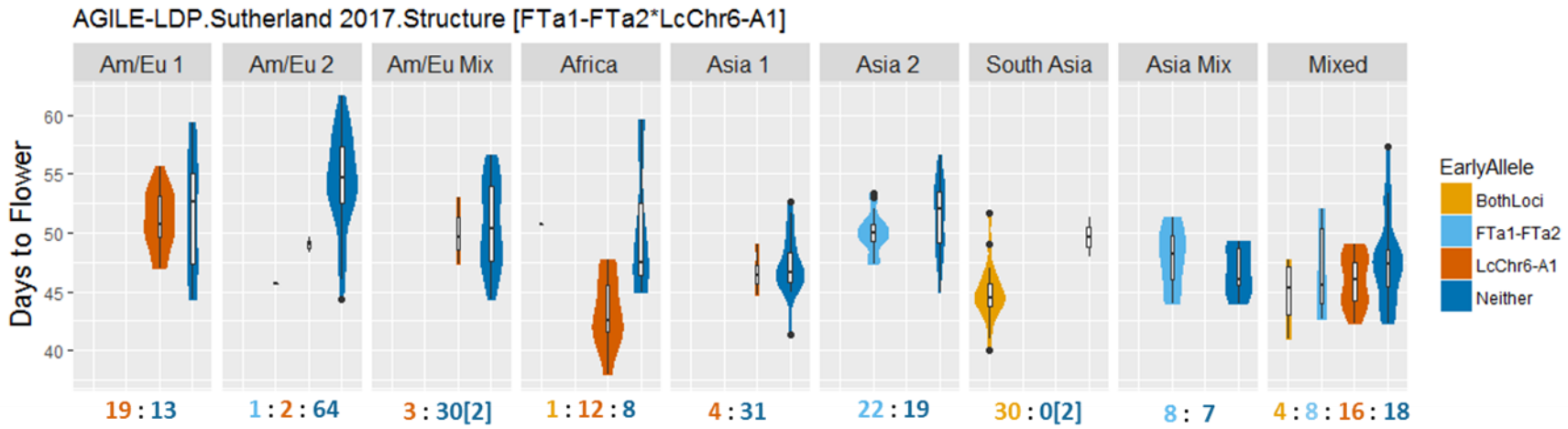

Figure 4.3. Days to flowering (DTF) in Sutherland 2017 of the AGILE-LDP when grouped according to population structure and allelic state at FTa1-FTa2 and LcChr6-A1. Population structure was used to group accessions based on genetic relatedness, with labels referring to where the majority of the accessions in that group are grown today: America/Europe (Am/Eu) 1 (n=32), Am/Eu 2 (n=67), Am/Eu Mixed (n=35), Africa (n=21), Asia 1 (n=35), Asia 2 (n=41), South Asia (n=32), Asia Mixed (n=15) and Mixed (n=46). Accessions are then grouped based on allelic state at both markers with early at both loci labeled **BothLoci**. Accessions early for the FTa1-FTa2 locus but scoring late for LcChr6-A1 are labeled (**FTa1-FTa2**). Accessions with the late FTa1-FTa2 allele were then split based on early for LcChr6-A1 (**LcChr6-A1**) and late/heterozygous [H] (**Neither**). The numbers at the bottom of each population structure group indicate how accessions within each are split based on allelic state.

Group America/Europe 2 is populated by accessions mostly grown in northern Temperate areas and these accessions are generally the last to flower in the Saskatchewan trials, with mean DTF in S17 greater than 54 days, and almost every individual carries the late allele at both loci. Many accessions in Asia 1 also score late for both FTa1-FTa2 and LcChr6-A1, but the mean DTF in Saskatchewan (SK) field conditions is consistently earlier, with a difference of over 7 days, over America/Europe 2 accessions in the field in Sutherland 2017 (Figure 4.3; https://knowpulse.usask.ca/Heidecker-Thesis-Association-Data; Supplementary 3).

The America/Europe 1 and Africa groups both have many accessions with passport data indicating a 'Mediterranean' origin. Also included in these groups are accessions originating in Latin America and a few accessions that originate from South Asia (https://knowpulse.usask.ca/Heidecker-Thesis-Association-Data; Supplementary 3). All individuals in these groups carry the late allele at the FTa1-FTa2 locus, however, nearly half the accessions possess the early allele at LcChr6-A1 (Figure 4.3). In both groups, the average DTF is earlier for the accessions with the early LcChr6-A1 allele than those with the late allele at this locus, however, there is considerable overlap between the two allelic states (Figure 4.3).

The early allele at FTa1-FTa2, DEL, is generally limited to accessions of 'Asian' origin and is almost exclusive to the structure groups Asia 2, South Asia, and the admixture group, Asia Mixed. Among the Asia 2 and Asia Mix groups, all accessions have the late allele at the LcChr6-A1 locus. The difference at FTa1-FTa2 accounts for a minor difference in DTF across the long day conditions, with accessions possessing the DEL allele likely to flower earlier than those scoring WT in the Asia 2 group, although this is not the case in the Asia Mix group (Figure 4.3).

Thirty-five (35) accessions possess early alleles for both FTa1-FTa2 and LcChr6-A1, and most of these are in the South Asia structure group. In the long day conditions in this study, they are among the earliest to flower in the GH and field, although some lines flower closer to the average DTF of the whole panel.

4.4 Discussion

4.4.1 Allelic State at FTa1-FTa2 accounts for differences in DTF, but alone is not informative enough for selecting diverse lentil germplasm for Saskatchewan

When considering the entire AGILE-LDP, allelic state at FTa1-FTa2 accounts for a similar amount of difference in DTF in both long day greenhouse (GH) and Saskatchewan field

experiments (Table 4.1; Figure 4.1A). Grouping by FTa1-FTa2 allele also explains differences in DTF in LR-11 (Figure 4.1B; Table 4.2). The major DTF QTL that Haile et al. (2021) found in their Saskatchewan field study matches the genomic region the FTa1-FTa2 marker represents and accounted for significant differences in mean DTF (1.8 days) in their bi-parental population. The distribution of DTF between the marker groups in these field studies are all much narrower, with much more overlap between the two allelic states, than reported in Rajandran et al. (2021). In their case, a bi-parental F2 population grown in greenhouse experiments, however, they found a larger difference for DTF between allelic states during short days (12-hour daylength) compared to long day experiments (16-hour daylength). The increased daylength and larger variability in ambient temperatures prior to flowering in Saskatchewan field studies relative to indoor ones, likely induces strong responses from additional loci which are reducing the differences between states at FTa1-FTa2, and could be used to refine screening of diverse material like the AGILE-LDP (Fedoruk et al. 2013; Wright et al. 2020; Haile et al. 2021).

4.4.2 Allelic state at LcChr6-A1 accounts for genetic control distinct to FTa1-FTa2, and they can be used together to improve DTF predictions

Even though both FTa1-FTa2 and LcChr6-A1 are located on LcChr6, it is unlikely the two markers are accounting for the same source of genetic control for DTF, and LcChr6-A1 appears to account for a slightly larger amount of the observed differences in the long day trials reported here (Table 4.2), especially in the LR-11 population where it was the most significant QTL. In the genome assembly LcV1.2, the putative genes these markers represent are located over 100 Mbp apart and the linkage map for LR-11 places markers at *LcFTa1* and *LcFTa2,* in QTL group LcChr6-B which can be represented by FTa1-FTa2, about 70 cM away and in a separate QTL group than the LcChr6-A1 marker (Table 3.6). Only around half of the LR-11 population, 57 RILs, score the same as a parent at both FTa1-FTa2 and LcChr6-A1, further supporting that the variability in DTF the two markers are accounting for is coming from genetically distinct sources. The two discrete DTF QTL identified in the Saskatchewan field experiments by Haile et al. (2021) also match the regions these two markers represent. Further, expression studies conducted by Rajandran et al. (2021) under short days indicate that increased expression of *LcFTa1* was linked to the earlier flowering associated with the FTa1-FTa2 marker but saw little to no expression of *LcFTb2.* In contrast, preliminary expression analysis on selected lines from LR-11 (section 3.3.5, Appendix 2), indicate the LcChr6-A1 marker is linked to *LcFTb2*

expression, with lines with the early allele at LcChr6-A1 expressing higher levels than those with the late allele. In models created to test the two markers' ability to predict of DTF, a model including both markers was superior to using only one, for both LR-11 and AGILE-LDP experiments (Table 4.2). Further, while there was significant overlap between the four potential allelic states, when grouping with both FTa1-FTa2 and LcChr6-A1 (Figure 4.2), there appeared to be a trend indicating a compounding effect of the two loci.4.4.3 Phenological pressures associated with origin contribute to allelic state of flowering markers in the AGILE-LDP

The genetic population structure groups used for the AGILE-LDP could be broadly aligned to the three climatic groups previously suggested by Khazaei et al. (2016), with America/Europe 2 closest to 'northern Temperate', America/Europe 1, Asia 1, and Africa generally matching 'Mediterranean', and lines of Asia 2 and South Asia structure groups being 'South Asian'. Generally, northern Temperate lentil is spring sown and experiences long days; Mediterranean is fall sown and responds to increasing temperature and daylength, following a period of low temperature and shorter days; while South Asian lentil is also fall sown but flowers quickly prior to restrictive heat and reduced water availability (Wright et al. 2020).

In the long day conditions of this experiment, accessions which score late for both FTa1-FTa2 and LcChr6-A1 had the latest mean DTF of the four allelic states (Figure 4.2D), and this was mostly the case for each group the AGILE-LDP was split into based on population structure. The Saskatchewan-bred accessions in the AGILE-LDP possessed the late allele at both marker loci, and most fell in America/Europe 2. In the cases where exotic material scores late at both markers, access to the origins of these accessions can then be used to provide additional information regarding the potential genetic variability useful to Saskatchewan breeders. For example, accessions in Asia 1 may be a source of useful diversity, readily adaptable to Temperate conditions as most which grouped into Asia 1 do not possess the early allele at these two DTF loci. However, the relative earliness to flower of accessions in the Asia 1 group may be linked to other uncaptured loci that could further influence the actual DTF. Further investigation, comparing the more temperate-adapted accessions in America/Europe 2, to those in Asia 1 may be useful in identifying additional loci useful for screening material in Saskatchewan.

Most of the accessions with only the early allele at LcChr6-A1 fell into either America/Europe 1 or the Africa structure group, and are accessions expected to be adapted to a Mediterranean

climate. The two alleles at LcChr6-A1 are present in a similar number of lines in America/Europe 1 and Africa indicating that, rather than being necessary for adaptation, the early allele may instead provide fitness benefits under certain circumstances these lentils can be exposed to. In this study, where they were grown under 'temperate' conditions, the early allele at LcChr6-A1 was only associated with a slight reduction in DTF relative the late allele, even after sorting into the individual structure groups (Figure 4.3). The early allele of LcChr6-A1, with a potential link to *LcFTb2* expression (Figure 3.3), may be conferring earliness by reducing the photoperiod requirement prior to flowering (Hecht et al. 2011). This may provide an advantage in Mediterranean conditions where, prior to flowering, low photoperiods and low temperatures are the norm. However, in long day-GH and Saskatchewan field trials, the difference based on photoperiod sensitivity would be less obvious, as the requirements would be quickly fulfilled in accessions with either allele variant. Similar to how the early allele at LcChr6-A1 is not found in Temperate material, the early allele at FTa1-FTa2 (DEL) may be associated with a negative phenotype in the Mediterranean regions as is found in only a single Africa accession. As these lentils are typically planted prior to a cold period and need to wait to flower until daylength increases, which consequently is tied with rising temperatures in these areas, loss of regulation at *LcFTa1* may induce flowering prematurely. In *Medicago*, deletions in the intergenic region near *FTa1* result in a loss of flowering suppression prior to a cold period (Jaudal et al. 2013), an effect undesirable for Mediterranean grown lentil (Rajandran et al. 2021). It is possible that the difference in flowering linked to the LcChr6-A1 marker, caused by variation at the candidate gene *LcFTb2*, is adequate for providing earlier DTF in the lines which do possess the early allele, for Mediterranean climatic regions while the response to cold would be retained in all lines.

The early allele at FTa1-FTa2 is almost exclusive to the Asia 2 and South Asia structure groups, accessions from South Asian climatic regions. In contrast to lentil grown in Mediterranean regions, flowering suppression during a cold period would not be needed. The decrease in DTF associated with the loss of this regulatory region may have a significant enough of an effect in germplasm of Asia 2 structure group, and there was a difference DTF even in the long day experiments a between allelic states at FTa1-FTa2 (Figure 4.3). Germplasm in the South Asia structure group, however, originated from a climatic zone with harsh environmental constraints and has experienced a genetic bottleneck due to the extreme early flowering requirement (Erskine et al. 1998; Dikshit et al. 2015). This group is the only one where early alleles at both

markers exist, and perhaps both are required to achieve adequate earliness and have been retained regardless of any possible agronomic penalties the stacking of these early alleles may cause. Material from this area is also the only bearer of another early flowering variant at another flowering locus, *Sn*, with earliness conferred by variation in the regulatory region of an *Early Flowering Locus 3* (*ELF3*) homolog (Weller et al. 2012).

4.5 Conclusions

Screening a diverse group of material with the markers FTa1-FTa2 and LcChr6-A1 provided support that the two markers are accounting for a portion of the differences in DTF through the allelic status of two genic regions. This is a useful first step in evaluating their potential use in a breeding program, however the biological significance of these markers on germplasm grown in Saskatchewan remains be further explored.

The ability of these alleles to influence agronomic traits, such as yield, branching and plant height in long day conditions may be of value, and would contribute to understanding how variation at the underlying genes may be exploited by breeders. In particular, the potential agronomic effects the early allele at LcChr6-A1 may be useful to understand, as the differences in DTF under long day conditions explained by this marker on diverse accessions were minor, yet all Saskatchewan material possessed the ‘late’ allele. If a non-DTF factor can be identified, screening with LcChr6-A1 may be useful to avoid potential undesirable effects when accessing exotic material, with the added benefit of slightly accounting for DTF under long day conditions.

The inheritance of FTa1-FTa2 and LcChr6-A1 within the structure groups of the AGILE-LDP provides a possible explanation for why their effects on DTF are more muted under long day field studies than expected. Instead, the gene variants may be providing a signalling response not as relevant in Saskatchewan field experiments yet useful in other climatic regions. Investigating the relationship of FTa1-FTa2 and LcChr6-A1 on DTF in other environments could be useful to further understanding the mechanisms through which the genes they are expected to represent are integrating signals for climatic pressures, including photoperiod and temperature.

Chapter 5. General Discussion

5.1 General Discussion

When developing improved cultivars, breeders are constrained by the diversity in the material available to them, and for lentil only a portion of the worldwide diversity is found in lines grown in Saskatchewan (Lombardi et al. 2014; Khazaei et al. 2016). However, when breeders attempt to use lentil from other parts of the world, significant time and resources are spent selecting adaptations related to development in the temperate climate of western Canada, where days are long and daily temperature fluctuations can be large (Sarker et al. 1999b; Wright et al. 2020). The unique combinations of photoperiod and temperature lentil experiences in different climatic zones are so important to development that responses to these pressures can be used to genotypically group diverse lentil (Khazaei et al. 2016; Wright et al. 2020). Even when comparing experiments from two closely related environments like long day greenhouse (GH) to Saskatchewan field experiments (preliminary study in Appendix 1), simple adjustments for photoperiod and temperature are not enough on their own to account for the range of sensitivities diverse material can possess. Having a set of genetic markers for development-related fitness, typically measured as ‘days to flowering’ (DTF), would improve screening material from other climatic regions, and broaden the diversity available to breeders. The few genetic loci which have been identified as markers for fitness to an environment are attributed to photoperiod and temperature responses, further enforcing genes in this vein as candidates for selection (Sarker et al. 1999b; Weller et al. 2012; Haile et al. 2021; Rajandran et al. 2021).

The first hypothesis set to be tested in this thesis was that variability in DTF in a in a Canadian by South Asian lentil population is controlled by genetic factors at relatively few regions of the genome containing known flowering time genes, and this was shown to be the case. In Chapter 3, a recombinant inbred line (RIL) population called LR-11, previously created by crossing a variety developed for western Canada (CDC Milestone) with a variety adapted for the South Asian climatic region (ILL 8006) was grown in a Saskatchewan field experiment involving two sites and two years (four site-years). Using single nucleotide polymorphism (SNP) markers generated from an exome capture, the objective was to identify quantitative trait loci (QTL) for development, with a focus on DTF, useful to western Canada. Under these conditions, seven QTL for DTF (q.DTF) were identified on three linkage groups (LGs) (Table 3.6). Four of the

q.DTF were considered of higher value because they were implicated in more than one site-year. Additionally, the two constituents of DTF, days from sowing to emergence (DTE) and days from emergence to flowering (VegP) were investigated. A q.VegP shared confidence intervals with each the four q.DTF found in multiple site-years and any q.VegP found in more than one site-year co-located with a q.DTF. This suggests the adjustment made by DTE on DTF are minor, and other Saskatchewan field studies in lentil have also found DTF is sufficient for identifying loci which account for variation (Tullu et al. 2008; Fedoruk et al. 2013; Haile et al. 2021). What was less expected was that three q.DTE were found on two LGs, with the late allele for DTE provided by ILL 8006, compared to CDC Milestone being the source of the late allele at all q.DTF. One q.DTE was found in multiple site years suggesting DTE can be used to identify development related loci under genetic control distinct from those identified using only DTF. In lentil, no q.DTE have yet been reported in any other Saskatchewan field study, however the indoor lentil flowering study conducted by Rajandran et al. (2021) identified two q.DTE which co-located with multiple branching and height related QTL and were also at positions distinct from their q.DTF. Early emergence was linked to longer main stem internodes leading them to suggest the variation at their q.DTE may be due to the influence of gene(s) related to growth rate of main stem internodes.

The other development related traits recorded in this study, node of flower development (NFD), days to swollen pod (DTSP), days to maturity (DTM), reproductive period (RepP) and seed yield (YLD) were almost always correlated with DTF, with the exception being one site year of each RepP and NFD (Table 3.2). As RepP is the time from DTF to DTM, it was surmised that plant development unrelated to DTF, and therefore under unique genetic control, could be reducing correlations between the two. Two q.RepP were identified, with neither sharing confidence intervals with any q.DTF found. However, RepP showed low heritability (Table 3.3) and the q.RepP were considered of little use. Additional doubt was cast on the value of these q.RepP as only one was significant in more than one site-year, and the one that was identified in more than one site-year fell between two q.DTF of large effect. However, in the Saskatchewan field study conducted by Haile et al. (2021), they reported a higher level of heritability for RepP and identified a q.RepP stable across all site-years of their experiment that did not overlap with q.DTF. Additionally in their study, q.DTSP and q.DTM shared confidence intervals with q.RepP and not q.DTF suggesting that, even though RepP appeared to be a poor source of variation in

plant development in this study, RepP may identify development-related loci that are overlooked when only using DTF. Node of Floral Initiation (NFI) as well as NFD have been widely used in pea studies to find markers associated with flowering-related genes (Weller et al. 1997, 2012; Hecht et al. 2011; Liew et al. 2014). In lentil grown indoors, Rajandran et al. (2021) found NFD was highly correlated with DTF and suggested variation in NFD was influenced by photoperiod sensitivities. The long daylengths in the field study of chapter 3, however, would mask differences in photoperiod sensitivity and may be an explanation for the very low heritability of NFD and its poor correlation to DTF. These results highlight that, for this study at least, q.DTF are capturing the most important variation for development and the associated markers are worth further investigation to understand how they can be applied in screening diverse lentil. These additional traits, namely DTE, VegP, and RepP, which are suspected to be phenologically relevant might be further dissected in a population with more diversity than a bi-parental population, and genome wide association studies is a tool expected to provide further insight when used with diverse accessions, such as those of the AGILE-LDP (Neupane et al. unpublished).

The four q.DTF which were identified in more than one site-year were q.DTF.1-1 on LG1 and q.DTF.6-1, q.DTF.6-2 and q.DTF.6-3 located on LG5. The q.DTF.1-1 explained less variation for DTF than any of the three q.DTF on LG5, but markers under q.DTF.1-1 mapped to areas of the lentil genome which were covered by the most important q.DTF identified in a different Saskatchewan field study (Fedoruk et al. 2013). The related q.DTF in Fedoruk et al. (2013) was identified using a RIL population created using two parents both adapted to Saskatchewan. In contrast, the parents of the RIL population used in this study, LR-11, were from two different climatic regions, providing much more variability in DTF. q.DTF.1.1 may be accounting for a region useful for making smaller adjustments, such as those for development within lentil already adapted for Saskatchewan field conditions. Future work might start with the lentil homolog of the DTF related gene *GIGANTEA* (*GI*) (Hecht et al. 2007), which is found annotated on LcV1.2 among markers which are associated with q.DTF.1-1, and Sudheesh et al. (2016) identified differentially expressed transcripts in lentil tissues which matched pieces of the *Medicago GI* homolog. The AGILE-LDP, with accessions representing all the major climatic regions, and the large number of markers provided by the exome capture, is well positioned for further investigation into the markers under this q.DTF.

The second hypothesis tested in this thesis was that genetic markers corresponding to flowering QTLs in bi-parental populations will account for DTF in Saskatchewan (SK) in a population of diverse individuals, with markers identified in SK field trials being more useful in long day experiments than those found in a short-day greenhouse (GH) experiment. Additionally, the third hypothesis suggested that markers associated with different candidate flowering genes will independently account for variation in DTF among diverse lentil lines. While both markers (initially identified in bi-parental populations) which were tested here were found to account for DTF in a diverse panel of lentil accessions (AGILE-LDP), it was not shown that the marker identified in the long day experiments (LcChr6-A1) is more useful than the one first found under short days (FTa1-FTa2). The inheritance of the allelic state at these two markers in the AGILE-LDP indicates the genes underlying the two markers are different, however additional work is needed to confirm this and identify the causal mutations.

The two q.DTF of largest effect in this study, q.DTF.6-1 and q.DTF.6-2, were found to each contain homologs of the *Flowering Locus T* (*FT*) gene. While only a single *FT* is present in the model plant *Arabidopsis*, in cool season legumes the role of *FT* is expanded, and *FT* genes are fit into three distinct clades *FTa*, *FTb*, and *FTc* (Hecht et al. 2011). These genes work together in the integration of signals for both photoperiod and temperature (Kong et al. 2010; Hecht et al. 2011; Laurie et al. 2011) with variability in the regulation of these genes linked to adaptations for time to flowering across cool season legumes (Weller & Ortega, 2015). The number of *FT* genes within clades is conserved across *Medicago*, pea, and lentil and there are three *FTa* genes (*FTa1*, *FTa2* and *FTa3*), two *FTb* (*FTb1* and *FTb2*), and one *FTc* (Hecht, unpublished). The roles of these genes are also conserved, and *FTb* genes (mainly *FTb2*) have been found responsible for integrating photoperiod signals (Hecht et al. 2011; Laurie et al. 2011). In addition to integrating photoperiod, *FTa1*, *FTa2* and *FTc* (mainly *FTa1*) are implicated in responses to light quality, temperature and vernalization pressures (Thomson et al. 2019; Jaudal et al. 2020; Rajandran et al. 2021; Yuan et al. 2021). In support of this, markers which fall under q.DTF.6-1 and q.DTF.6-2 in the LR-11 analysis account for some of the variability in the diverse accessions of the AGILE-LDP, even under Saskatchewan field conditions (Figure 4.2), where differences at these floral pathway integrators would be muted.

q.DTF6-1 accounted for the largest differences in DTF in LR-11 and the associated markers mapped to the region of the lentil genome where *LcFTb1* and *LcFTb2* are located. As *FTb* genes have been shown to respond only to photoperiod, it was understandable that the two parents, one adapted for growth under long days and the other for relatively short days, would possess variation in regulating these genes. In the long days experienced in Saskatchewan field conditions, where there would be strong induction of *FTb* genes regardless of allelic variant (Ridge et al. 2017), the responses observed in this study may be muted compared to what would be seen under short days. However, in support of the utility of this region as a selection tool for lentil in western Canada, a q.DTF containing these genes was separately found under Saskatchewan field conditions, in another population with an unadapted parent unrelated to the ones used in this study (Haile et al. 2021). A marker which was located at the peak of q.DTF.6-1, referred to as LcChr6-A1 in Chapter 4, was tested by screening the diverse set of lentils in the AGILE-LDP first used in Wright et al. (2020). In long day greenhouse experiments and Saskatchewan field experiments, LcChr6-A1 was shown to account for some of the variation in DTF (Table 4.2; Figure 4.2). Additionally, an exploratory expression study using a selection of lines from the LR-11 population found *LcFTb2* expression was higher in the lines with ILL 8006 parental allele at q.DTF.6-1 than the adapted parent CDC Milestone (Figure 3.3; Appendix 2), highlighting this gene as a prime candidate for future work while investigating what this marker is representing. Taken together, these results support the inclusion of a marker for this locus in a panel of markers for screening DTF in diverse lentil but, as *FTb2* has only been found to integrate photoperiod signals and the climatic regions lentil it is grown in differ in more than just daylength, additional loci are bound to play important roles.

The lentil homologs of *FTa1*, *FTa2* and *FTc* are all found close to one another in the lentil genome (Hecht et al. 2011; Ramsay et al. 2021), and markers associated with q.DTF.6-2 encompass a wide area which includes them all. Flowering loci which include these genes have been identified in both the model, *Medicago*, and crop legumes, including lentil, pea and chickpea (Hecht et al. 2011; Putterill et al. 2013; Ortega et al. 2019; Haile et al. 2021; Rajandran et al. 2021; Yuan et al. 2021). Promoted expression of these three genes corresponds to flowering, with *FTa1* acting as a primary integrator of photoperiod and vernalization signals, and *FTc* suspected to act as an integrator of signals from other *FT* genes, which are primarily leaf tissue based, in shoot apex tissues (Hecht et al. 2011; Laurie et al. 2011). The role of *FTa2* is the

most poorly conserved of the three *FT* genes that fall under q.DTF.6-2 and, where implicated, it plays a similar but weaker role to *FTa1* (Hecht et al. 2011; Ortega et al. 2019; Rajandran et al. 2021). In *Medicago*, chickpea and lentil, variation in the intergenic regions of this cluster of *FT* genes has been associated with adjusted time to flowering through differences in *FTa1* expression (Jaudal et al. 2013; Ortega et al. 2019; Rajandran et al. 2021). A bi-allelic marker, called FTa1-FTa2, associated with the intergenic deletion identified in an early flowering lentil line by Rajandran et al. (2021), segregated in LR-11, falling within q.DTF.6-2, suggesting it may be the source of genetic difference represented by the q.DTF. Additionally, the LR-11 parent, CDC Milestone, had lower levels of *FTa1* expressed than the selection of LR-11 RILs with the ILL 8006 allele at q.DTF.6-2 (Figure 3.3, Appendix 2). Screening the AGILE-LDP with the FTa1-FTa2 marker found the two allelic groups did show differences for DTF in the long day greenhouse experiments and Saskatchewan field and accounted for differences to a degree comparable to grouping using LcChr6-A1 (Table 4.2, Figure 4.1). In contrast to studies which identify the allelic state near *FTa1* as a strong influencer of DTF (Rajandran et al. 2021; Yuan et al. 2021), the role of an early *FTa1* allele would be reduced under the long day field conditions in this study, where plants experience cool nights and plenty of long days with high levels of solar radiation (Wright et al. 2020; Figure A.1.1).

The inheritance of the allelic state for LcChr6-A1 and FTa1-FTa2 in the AGILE-LDP, after accounting for population structure, further support *FTb2* and *FTa1*, respectively, as candidate genes for these markers. South Asian germplasm, which experiences short days and has to reach flowering with the smallest amount of photothermal units to avoid severe yield penalties (Dikshit et al. 2015; Sita et al. 2018; Wright et al. 2020) were the only lines with the early allele at both loci (Figure 4.3). Further, most of the lines which contained the early allele for FTa1-FTa2 but not at LcChr6-A1 were those suspected to be used on the edge of the South Asian climatic region, and reduced flowering suppression without a cold period (i.e., vernalization) may be beneficial (Rajandran et al. 2021). In contrast, lentil grown in the Mediterranean climatic region does experience a cold period prior to flowering(Wright et al. 2020), and retaining this response may be of value. Lines in the AGILE-LDP of the Mediterranean background are unlikely to possess the early allele at FTa1-FTa2, but nearly half have the early LcChr6-A1 allele. If LcChr6-A1 is denoting *LcFTb2*, allelic variation could be retained in some lines as a discrete

adjustment in photoperiod sensitivity for development under the short days, while allowing the integration of temperature signals to remain intact.

The q.DTF.6-3 was found in both years at the Sutherland site but was never significant at Rosthern. It covered a narrow region on the linkage map, but markers at this q.DTF mapped to two discrete regions of the genome assembly ~100Mbp apart. The four candidate gene homologs in this region: *LcLFY, LcDCL1, LcATX3b* and *LcPRR95c,* while conserved across a number of legumes, have yet to be extensively implicated in expression studies (Kim et al. 2013; Weller and Ortega 2015). *LcPRR95c* is currently the strongest candidate, as it was also included in a q.DTF found by Rajandran et al. (2021). *Pseudo Response Regulator* (*PRR*) genes implicated in photoperiod signalling pathways in legumes (Liew et al. 2014; Weller and Ortega 2015) and, in *FTb2* pea mutants, Ridge et al. (2016) found the *PRR95c* homolog also showed adjusted expression.

5.2 Conclusions and Future Work

The goal of this study was to identify genetic markers for days to flowering (DTF) which could be used as tools by which lentil breeders in western Canada could more easily consider exotic material. In Chapter 3, a Saskatchewan field experiment using a bi-parental population with a Canadian line and a South Asian line, identified four q.DTF as places likely to possess useful markers. Further, two of these, q.DTF.6-1 and q.DTF.6-2, contain genes implicated across multiple cool season legume studies in the integration of environmental signals related to flowering and development, members of the *FT* family. In Chapter 4, markers representing the variation at these two q.DTF were then tested on diverse lentil accessions (AGILE-LDP) and implicated in independently accounting for DTF under Saskatchewan field studies, however as much variation in DTF still exists after using the two markers, the genetic pathways underlying these markers need to be further investigated.

Future work investigating the flowering time pathway suggested here will improve our understanding of how these loci are influencing this important adaptation trait. Increasingly, markers readily comparable to the lentil genome are being used in flowering studies and will allow the identification of conserved variation. Markers in q.DTF.6-1 are in the same area as markers associated with a q.DTF in both Haile et al. (2021) and Polanco et al. (2019). Similarly, the region covered by q.DTF.6-2 is expected to be the same genes as q.DTF in multiple recent

lentil studies (Haile et al. 2021; Rajandran et al. 2021; Yuan et al. 2021). The colocation of QTL across studies at these two loci provide a strong suggestion there is relevant genetic variability affecting DTF. Identifying the genetic variation these q.DTF are accounting for could; A) confirm the flowering genes implicated and B) identify a marker directly accounting for the source of the variability, both of which would improve the likelihood of adoption into breeding programs. As *FT* genes are downstream in signalling pathways, and their core roles are conserved across cool season legumes species (Hecht et al. 2011; Weller and Ortega 2015; Ortega et al. 2019), they are good candidates for crop improvement. However, mutations causing large differences in flowering habit in these genes are likely not conserved and instead, variability which influences how they interpret upstream signals is associated with small sensitivity adjustments (Ortega et al. 2019). Coupled with the implication that the source of variation is likely located in intergenic regions, identifying causal variants which are widely applicable may be difficult. In the South Asian climatic region, variation at a single locus *Sn*, which denotes an *Early Flowering 3* (*ELF3*) homolog, fits the main selection criteria for determining ideal flowering habit (Sarker et al. 1999b; Weller et al. 2012). In contrast to this, the variation in flowering after accounting for the allelic states of the two markers in Chapter 4 shows there are likely multiple loci yet to identify which will be relevant to DTF under western Canadian field conditions. Thus, an additional consideration to breeding for western Canada is that, even though lentil development occurs under photoperiod and temperature conditions contrasting those experienced by germplasm from sources of lentil diversity (i.e. the Mediterranean climatic region), the probability of identifying a single locus of large effect, comparable to *Sn* for the South Asian climatic region, for western Canada is low because the associated signals which induce flowering genes are relatively inductive (Wright et al. 2020). Unlike in other climatic regions, breeding for western Canada involves selecting for reduced responses to photoperiod and temperature signals. Integrators like *FT* make good candidates for this, and this highlights the potential of q.DTF.6-1 and q.DTF.6-2.

In addition to strictly selecting for DTF, knowing which genes are behind observed variation is valuable because they often have pleiotropic effects and, in lentil and its relatives, *FT* genes have been found associated with additional growth habit traits relevant to breeders such as branching, as well as root and general biomass accumulation (Danilevskaya et al. 2011; Pin and Nilsson 2012; Ortega et al. 2019; Rajandran et al. 2021). In lentil studies, DTF is often investigated in

conjunction with agronomic traits, with parents of crosses also segregating for traits like plant structure and size (Fratini et al. 2007; Tullu et al. 2008) and seed related traits (Fedoruk et al. 2013; Verma et al. 2015; Jha et al. 2017; Polanco et al. 2019) and these traits may provide additional clues when searching for the underlying genetic factors for DTF in the future.

Another outcome of this study was the indication that there were stages along plant development poorly correlated to DTF but with enough heritability under Saskatchewan field conditions for QTL identification. QTL for Reproductive Period (RepP) and days to emergence (DTE), at positions separate from q.DTF, may be areas of interest for study in other populations. Dense linkage maps and accurate phenotyping can identify useful loci for these additional development-related traits in lentil (Haile et al. 2021; Rajandran et al. 2021). The q.DTE from this study identifies a large confidence interval on LG2 even though it explains a large amount of the phenotypic variation, because of poor separation of markers from both LcChr2 and LcChr3 due to a suspected translocation near where the q.DTE is expected to be located. Future investigation into markers underlying q.DTE may be improved by recording growth habit-related phenotypes like what was done in Rajandran et al. (2021). Additionally, markers studies involving QTLs expected to fall in the region covered by LG2 (LcChr2 and LcChr3) may be more effective in another population where LcChr2 and LcChr3 each map to individual LGs.

References

Abbo, S., Saranga, Y., Peleg, Z., Kerem, Z., Lev-yadun, S., and Gopher, A. 2009. Reconsidering domestication of legumes versus cereals in the ancient near east. Q. Rev. Biol. 84: 29–50.

Aldemir, S., Ateş, D., Yilmaz Temel, H., Yağmur, B., Alsaleh, A., Kahriman, A., Özkan, H., Vandenberg, A., and Tanyolac, M.B. 2017. QTLs for iron concentration in seeds of the cultivated lentil (*Lens culinaris* Medic.) via genotyping by sequencing. Turkish J. Agric. For. 41: 243–255. doi:10.3906/tar-1610-33.

Alo, F., Furman, B.J., Akhunov, E., Dvorak, J., and Gepts, P. 2011. Leveraging genomic resources of model species for the assessment of diversity and phylogeny in wild and domesticated lentil. J. Hered. 102: 315–329. doi:10.1093/jhered/esr015.

Ateş, D., Aldemir, S., Alsaleh, A., Erdogmus, S., Nemli, S., Kahriman, A., Ozkan, H., Vandenberg, A., and Tanyolac, B. 2018. A consensus linkage map of lentil based on DArT markers from three RIL mapping populations. PLoS One 13: 1–13. doi:10.1371/journal.pone.0191375.

Ateş, D., Sever, T., Aldemir, S., Yagmur, B., Temel, H.Y., Kaya, H.B., Alsaleh, A., Kahraman, A., Ozkan, H., Vandenberg, A., and Tanyolac, B. 2016. Identification QTLs controlling genes for Se uptake in Lentil seeds. PLoS One 11. doi:10.1371/journal.pone.0149210.

Berger, J.D., Milroy, S.P., Turner, N.C., Siddique, H.M., Imtiaz, M., and Malhotra, R. 2011. Chickpea evolution has selected for contrasting phenological mechanisms among different habitats. Euphytica 180: 1–15. doi:10.1007/s10681-011-0391-4.

Bhakta, M.S., Gezan, S.A., Clavijo Michelangeli, J.A., Carvalho, M., Zhang, L., Jones, J.W., Boote, K.J., Correll, M.J., Beaver, J., Osorno, J.M., Colbert, R., Rao, I., Beebe, S., Gonzalez, A., Ricaurte, J., and Vallejos, C.E. 2017. A Predictive Model for Time-to-Flowering in the Common Bean Based on QTL and Environmental Variables. G3 Genes|Genomes|Genetics 7: 3901–3912. doi:10.1534/g3.117.300229.

Bohra, A., and Singh, N.P. 2015. Whole genome sequences in pulse crops: a global community resource to expedite translational genomics and knowledge-based crop improvement. Biotechnol. Lett. 37: 1529–1539. Springer Netherlands. doi:10.1007/s10529-015-1836-y.

Bouché, F., Lobet, G., Tocquin, P., and Périlleux, C. 2016. FLOR-ID: An interactive database of flowering-time gene networks in *Arabidopsis thaliana*. Nucleic Acids Res. 44: D1167–D1171. doi:10.1093/nar/gkv1054.

Burgarella, C., Chantret, N., Gay, L., Prospéri, J.-M., Bonhomme, M., Tiffin, P., Young, N.D., and Ronfort, J. 2016. Adaptation to climate through flowering phenology: a case study in *Medicago truncatula*. Mol. Ecol. 25: 3397–3415. doi:10.1111/mec.13683.

Canadian Grain Commission. 2021. Scientific analysis of western Canadian lentil quality. Statistics Canada. Available: https://www.grainscanada.gc.ca/en/grain-research/export-quality/pulses/lentils.

Cober, E.R., Curtis, D.F., Stewart, D.W., and Morrison, M.J. 2014. Quantifying the Effects of Photoperiod, Temperature and Daily Irradiance on Flowering Time of Soybean Isolines. Plants 3: 476–497. doi:10.3390/plants3040476.

Cockram, J., Jones, H., Leigh, F.J., O'Sullivan, D., Powell, W., Laurie, D.A., and Greenland, A.J. 2007. Control of flowering time in temperate cereals: Genes, domestication, and sustainable productivity. J. Exp. Bot. 58: 1231–1244. doi:10.1093/jxb/erm042.

Cronk, Q., Ojeda, I., and Pennington, R.T. 2006. Legume comparative genomics: Progress in phylogenetics and phylogenomics. Curr. Opin. Plant Biol. 9: 99–103. doi:10.1016/j.pbi.2006.01.011.

Cubero, J. I., Perez de la Vega, M. and Fratini, R. 2009. Origin, phylogeny, domestication and spread. p. 13-33 in W. Erskine, F. J. Muehlbauer, A. Sarker, and B. Sharma, eds. The lentil: Botany, Production and Uses. CABI, Wallingford, UK.

Daba, K., Tar'an, B., and Warkentin, T.D. 2016. Flowering response of diverse chickpea (*Cicer arietinum* L.) accessions to photoperiod. Genet. Resour. Crop Evol. 63: 1161–1172. Springer Netherlands. doi:10.1007/s10722-015-0308-5.

Danilevskaya, O.N., Meng, X., McGonigle, B., and Muszynski, M.G. 2011. Beyond flowering time: Pleiotropic function of the maize flowering hormone *Florigen*. Plant Signal. Behav. 6: 1267–1270. doi:10.4161/psb.6.9.16423.

DePristo, M.A., Banks, E., Poplin, R., Garimella, K. V, Maguire, J.R., Hartl, C., Philippakis, A.A., Angel, G. del, Rivas, M.A., Hanna, M., McKenna, A., Fennell, T.J., Kernytsky, A.M., Sivachenko, A.Y., and Daly, M.J. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. 43: 491–501. doi:doi:10.1038/ng.806.

Dias, P.M.B., Brunel-Muguet, S., Dürr, C., Huguet, T., Demilly, D., Wagner, M.-H., and Teulat-Merah, B. 2011. QTL analysis of seed germination and pre-emergence growth at extreme temperatures in *Medicago truncatula*. Theor. Appl. Genet. 122: 429–444. doi:10.1007/s00122-010-1458-7.

van Dijk, A.D.J., and Molenaar, J. 2017. Floral pathway integrator gene expression mediates gradual transmission of environmental and endogenous cues to flowering time. PeerJ 5: e3197. doi:10.7717/peerj.3197.

Dikshit, H.K., Singh, A., Singh, D., Aski, M.S., Prakash, P., Jain, N., Meena, S., Kumar, S., and Sarker, A. 2015. Genetic diversity in *Lens* species revealed by EST and genomic simple sequence repeat analysis. PLoS One 10: e0138101. doi:10.1371/journal.pone.0138101.

Ehrenreich, I.M., Hanzawa, Y., Chou, L., Roe, J.L., Kover, P.X., and Purugganan, M.D. 2009. Candidate gene association mapping of *Arabidopsis* flowering time. Genetics 183: 325–335. doi:10.1534/genetics.109.105189.

Erskine, W., Adham, Y., and Holly, L. 1989. Geographic distribution of variation in quantitative traits in a wold lentil collection. Euphytica 43: 97–103. doi:10.1007/BF00037901.

Erskine, W., Chandra, S., Chaudhry, M., Malik, I.A., Sarker, A., Sharma, B., Tufail, M., and Tyagi, M.C. 1998. A bottleneck in lentil: widening its genetic base in South Asia. Euphytica 101: 207–211.

Erskine, W., Ellis, R.H., Summerfield, R.J., Roberts, E.H., and Hussain, A. 1990a. Characterization of responses to temperature and photoperiod for time to flowering in a world lentil collection. Theor. Appl. Genet. 80: 193–199. doi:10.1007/BF00224386.

Erskine, W., Hussain, A., Tahir, M., Bahksh, A., Ellis, R.H., Summerfield, R.J., and Roberts, E.H. 1994a. Field evaluation of a model of photothermal flowering responses in a world lentil collection. Theor. Appl. Genet. 88: 423–428. doi:10.1007/BF00223655.

Erskine, W., Muehlbauer, F.J., and Short, R.W. 1990b. Stages of Development in Lentil. Exp. Agric. 26: 297. doi:10.1017/S0014479700018457.

Erskine, W., Tufail, M., Russell, A., Tyagi, M.C., Rahman, M.M., and Saxena, M.S. 1994b. Current and future strategies in breeding lentil for resistance to biotic and abiotic stresses. doi:10.1007/BF00027189.

FAOSTAT. 2021. Food and Agriculture Organization of the United Nations (FAO) Statistical Division. Available: http://www.fao.org/faostat/en/#data.

Fedoruk, M.J., Vandenberg, A., and Bett, K.E. 2013. Quantitative trait loci analysis of seed quality characteristics in lentil using single nucleotide polymorphism markers. Plant Genome 6: 1–10. doi:10.3835/plantgenome2013.05.0012.

Ferguson, M.E., Robertson, L.D., Ford-Lloyd, B. V., Newbury, H.J., and Maxted, N. 1998. Contrasting genetic variation amongst lentil landraces from different geographical origins. Euphytica 102: 265–273.

Fiala, J. V, Tullu, A., Banniza, S., Séguin-Swartz, G., and Vandenberg, A. 2009. Interspecies Transfer of Resistance to Anthracnose in Lentil (*Lens culinaris* Medic.). Crop Sci. 49: 825–830. doi:10.2135/cropsci2008.05.0260.

Fornara, F., de Montaigu, A., and Coupland, G. 2010. SnapShot: Control of flowering in *Arabidopsis*. Cell 141: 3–5. doi:10.1016/j.cell.2010.04.024.

Fratini, R., Durán, Y., García, P., and Pérez De La Vega, M. 2007. Identification of quantitative trait loci (QTL) for plant structure, growth habit and yield in lentil. Spanish J. Agric. Res. 5: 348–356.

Fudge, J.B., Lee, R.H., Laurie, R.E., Mysore, K.S., Wen, J., Weller, J.L., and Macknight, R.C. 2018. *Medicago truncatula SOC1* genes are up-regulated by environmental cues that promote flowering. Front. Plant Sci. 9: 496. doi:10.3389/fpls.2018.00496.

Fuller, D.Q., Willcox, G., and Allaby, R.G. 2011. Cultivation and domestication had multiple origins: Arguments against the core area hypothesis for the origins of agriculture in the Near East. World Archaeol. 43: 628–652. doi:10.1080/00438243.2011.624747.

Gela, T.S., Adobor, S., Khazaei, H., and Vandenberg, A. 2021a. An advanced lentil backcross population developed from a cross between *Lens culinaris* × *L. ervoides* for future disease resistance and genomic studies. 19: 167–173. doi:10.1017/S1479262121000216.

Gela, T.S., Banniza, S., and Vandenberg, A. 2021b. Lack of effective resistance to the virulent race of *Colletotrichum lentis* in *Lens culinaris* Medikus subsp. *culinaris*. 18: 81–87. doi:10.1017/S1479262120000027.

Gela, T.S., Koh, C.S., Caron, C.T., Chen, L., Vandenberg, A., and Bett, K.E. 2021c. QTL mapping of lentil anthracnose (*Colletotrichum lentis*) resistance from *Lens ervoides* accession IG 72815 in an interspecific RIL population. Euphytica 217. doi:10.1007/s10681-021-02804-0.

Ghanem, M.E., Marrou, H., Soltani, A., Kumar, S., and Sinclair, T.R. 2015. Lentil variation in phenology and yield evaluated with a model. Agron. J. 107. doi:10.2134/agronj15.0061.

Grabowski, P.P., Evans, J., Daum, C., Deshpande, S., Barry, K.W., Kennedy, M., Ramstein, G., Kaeppler, S.M., Buell, C.R., Jiang, Y., and Casler, M.D. 2017. Genome-wide associations with flowering time in switchgrass using exome-capture sequencing data. New Phytol. 213: 154–169. doi:10.1111/nph.14101.

Haile, T.A., Heidecker, T., Wright, D.M., Neupane, S., Ramsay, L., Vandenberg, A., and Bett, K.E. 2020. Genomic selection for lentil breeding : Empirical evidence. Plant Genome 13: e20002. doi:10.1002/tpg2.20002.

Haile, T.A., Stonehouse, R., Weller, J.L., and Bett, K.E. 2021. Genetic basis for lentil adaptation to summer cropping in northern temperate environments. Plant Genome: e20144. doi:10.1002/tpg2.20144.

Hecht, V.F.G., Foucher, F., Ferrándiz, C., Macknight, R.C., Navarro, C., Morin, J., Vardy, M.E., Ellis, N., Beltrán, J.P., Rameau, C., and Weller, J.L. 2005. Conservation of *Arabidopsis* flowering genes in model legumes. Plant Physiol. 137: 1420–1434. doi:10.1104/pp.104.057018.

Hecht, V.F.G., Knowles, C.L., Vander Schoor, J.K., Liew, L.C., Jones, S.E., Lambert, M.J.M., and Weller, J.L. 2007. Pea *LATE BLOOMER1* is a *GIGANTEA* ortholog with roles in photoperiodic flowering, deetiolation, and transcriptional regulation of circadian clock gene homologs. Plant Physiol. 144: 648–61. doi:10.1104/pp.107.096818.

Hecht, V.F.G., Laurie, R.E., Vander Schoor, J.K., Ridge, S., Knowles, C.L., Liew, L.C., Sussmilch, F.C., Murfet, I.C., Macknight, R.C., and Weller, J.L. 2011. The Pea GIGAS Gene Is a *FLOWERING LOCUS T* Homolog Necessary for Graft-Transmissible Specification of Flowering but Not for Responsiveness to Photoperiod. Plant Cell 23: 147–161. doi:10.1105/tpc.110.081042.

Heffelfinger, C., Fragoso, C.A., and Lorieux, M. 2017. Constructing linkage maps in the genomics era with MapDisto 2.0. Bioinformatics 33: 2224–2225. doi:10.1093/bioinformatics/btx177.

Hirsch, C.D., Evans, J., Buell, C.R., and Hirsch, C.N. 2014. Reduced representation approaches to interrogate genome diversity in large repetitive plant genomes. Briefings Funct. Genomics Proteomics 13: 257–267. doi:10.1093/bfgp/elt051.

Jaudal, M., Wen, J., Mysore, K.S., and Putterill, J. 2020. *Medicago PHYA* promotes flowering, primary stem elongation and expression of flowering time genes in long days. BMC Plant Biol. 20: 329. BMC Plant Biology. doi:10.1186/s12870-020-02540-y.

Jaudal, M., Yeoh, C.C., Zhang, L., Stockum, C., Mysore, K.S., Ratet, P., and Putterill, J. 2013. Retroelement insertions at the *Medicago FTa1* locus in spring mutants eliminate vernalisation but not long-day requirements for early flowering. Plant J. 76: 580–591. doi:10.1111/tpj.12315.

Jaudal, M., Zhang, L., Che, C., Hurley, D.G., Thomson, G., Wen, J., Mysore, K.S., and Putterill, J. 2016. *MtVRN2* is a Polycomb *VRN2-like* gene which represses the transition to flowering in the model legume *Medicago truncatula*. Plant J. 86: 145–160. doi:10.1111/tpj.13156.

Jha, R., Bohra, A., Chand, U., and Maneet, J. 2017. Analysis of an intraspecific RIL population uncovers genomic segments harbouring multiple QTL for seed relevant traits in lentil (*Lens culinaris* L.). Physiol. Mol. Biol. Plants 23: 675–684. Springer India. doi:10.1007/s12298-017-0438-7.

Ji, T., and Chen, J. 2016. Statistical models for DNA copy number variation detection using read-depth data from next generation sequencing experiments. Aust. N. Z. J. Stat 58: 473–491. doi:10.1111/anzs.12175.

Jighly, A., Joukhadar, R., and Alagu, M. 2015. SimpleMap : A pipeline to streamline high density linkage map construction. Plant Genome 8: 1–24. doi:10.3835/plantgenome2014.09.0056.

Jung, C., and Müller, A.E. 2009. Flowering time control and applications in plant breeding. Trends Plant Sci. 14: 563–573. doi:10.1016/j.tplants.2009.07.005.

Jung, C., Pillen, K., Staiger, D., Coupland, G., and Korff, M. Von 2017. Editorial : Recent Advances in Flowering Time Control. Front. Plant Sci. 7: 2011. doi:10.3389/fpls.2016.02011.

Kahraman, A., Kusmenoglu, I., Aydin, N., AYDOĞAN, A., Erskine, W., and Muehlbauer, F.J. 2004. Genetics of Winter Hardiness in 10 Lentil Recombinant Inbred Line Populations. Crop Sci. 44: 5–12. doi:10.2135/cropsci2004.5000.

Kahriman, A., YILMAZ TEMEL, H., AYDOĞAN, A., and Tanyolac, B. 2014. Major quantitative trait loci for flowering time in lentil. Turkish J. Agric. For. 38: 588–595. doi:10.3906/tar-1408-16.

Kassa, M.T., You, F.M., Hiebert, C.W., Pozniak, C.J., Fobert, P.R., Sharpe, A.G., Menzies, J.G., Humphreys, D.G., Rezac Harrison, N., Fellers, J.P., McCallum, B.D., and McCartney, C.A. 2017. Highly predictive SNP markers for efficient selection of the wheat leaf rust resistance gene *Lr16*. BMC Plant Biol. 17: 45. doi:10.1186/s12870-017-0993-7.

Kaur, S., Cogan, N.O.I., Stephens, A., Noy, D., Butsch, M., Forster, J.W., and Materne, M. 2014. EST-SNP discovery and dense genetic mapping in lentil (*Lens culinaris* Medik.) enable candidate gene selection for boron tolerance. Theor. Appl. Genet. 127: 703–713. doi:10.1007/s00122-013-2252-0.

Khazaei, H., Caron, C.T., Fedoruk, M.J., Diapari, M., Vandenberg, A., Coyne, C.J., Mcmullen, M.D., and Bett, K.E. 2016. Genetic diversity of cultivated lentil (*Lens culinaris* Medik.) and its relation to the world's agro-ecological zones. Front. Plant Sci 7: 1093. doi:10.3389/fpls.2016.01093.

Kim, M.Y., Kang, Y.J., Lee, T., and Lee, S.-H. 2013. Divergence of Flowering-Related Genes in Three Legume Species. Plant Genome 6. doi:10.3835/plantgenome2013.03.0008.

Kong, F., Liu, B., Xia, Z., Sato, S., Kim, B.M., Watanabe, S., Yamada, T., Tabata, S., Kanazawa, A., Harada, K., and Abe, J. 2010. Two Coordinately Regulated Homologs of *FLOWERING LOCUS T* Are Involved in the Control of Photoperiodic Flowering in Soybean. Plant Physiol. 154: 1220–1231. doi:10.1104/pp.110.160796.

Kujur, A., Bajaj, D., Upadhyaya, H.D., Das, S., Ranjan, R., Shree, T., Saxena, M.S., Badoni, S., Kumar, V., Tripathi, S., Gowda, C.L.L., Sharma, S., Singh, S., Tyagi, A.K., and Parida, S.K. 2015a. Employing genome-wide SNP discovery and genotyping strategy to extrapolate the natural allelic diversity and domestication patterns in chickpea. Front. Plant Sci. 6: 162. doi:10.3389/fpls.2015.00162.

Kujur, A., Upadhyaya, H.D., Shree, T., Bajaj, D., Das, S., Saxena, M.S., Badoni, S., Kumar, V., Tripathi, S., Gowda, C.L.L., Sharma, S., Singh, S., Tyagi, A.K., and Parida, S.K. 2015b. Ultra-high density intra-specific genetic linkage maps accelerate identification of functionally relevant molecular tags governing important agronomic traits in chickpea. Sci. Rep. 5: 1–13. doi:10.1038/srep09468.

Kumar, J., Gupta, D., Baum, M., Varshney, R.K., and Kumar, S. 2021. Genomics-assisted lentil breeding : Current status and future strategies. Legum. Sci. 3: e71. doi:10.1002/leg3.71.

Kumar, J., Gupta, S., Gupta, P., Dubey, S., Tomar, R.S.S., and Kumar, S. 2016. Breeding strategies to improve lentil for diverse agro-ecological environments. Indian J. Genet. Plant Breed. 76: 530. doi:10.5958/0975-6906.2016.00071.7.

Kumar, P. 2007. Genetics Of Resistance To Stemphylium Leaf Blight Of Lentil (*Lens Culinaris*) In The Cross Barimasur-4 × CDC Milestone. MSc. Thesis. University of Saskatchewan.

Kuznetsova, A., Brockhoff, P.B., and Christensen, R.H.B. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. J. Stat. Softw. 82. doi:10.18637/jss.v082.i13.

Ladizinsky, G. 1979. The origin of lentil and its wild genepool. Euphytica 28: 179–187. doi:10.1007/BF00029189.

Laurie, R.E., Diwadkar, P., Jaudal, M., Zhang, L., Hecht, V.F.G., Wen, J., Tadege, M., Mysore, K.S., Putterill, J., Weller, J.L., and Macknight, R.C. 2011. The *Medicago FLOWERING LOCUS T* Homolog, *MtFTa1*, Is a Key Regulator of Flowering Time. Plant Physiol. 156: 2207–2224. doi:10.1104/pp.111.180182.

Laurie, R.E., Knowles, C.L., Weller, J.L., Hecht, V.F.G., Vander Schoor, J.K., Liew, L.C., and Macknight, R.C. 2009. *DIE NEUTRALIS* and *LATE BLOOMER 1* Contribute to Regulation of the Pea Circadian Clock. Plant Cell 21: 3198–3211. doi:10.1105/tpc.109.067223.

Lee, J.H., Yoo, S.J., Park, S.H., Hwang, I., Lee, J.S., and Ahn, J.H. 2007. Role of *SVP* in the control of flowering time by ambient temperature in *Arabidopsis*. Genes Dev. 21: 397–402. doi:10.1101/gad.1518407.GENES.

Liew, L.C., Hecht, V.F.G., Sussmilch, F.C., and Weller, J.L. 2014. The Pea Photoperiod Response Gene *STERILE NODES* Is an Ortholog of *LUX ARRHYTHMO*. Plant Physiol. 165: 648–657. doi:10.1104/pp.114.237008.

Liu, W., Jiang, B., Ma, L., Zhang, S., Zhai, H., Xu, X., Hou, W., Xia, Z., Wu, C., Sun, S., Wu, T., Chen, L., and Han, T. 2018. Functional diversification of *Flowering Locus T* homologs in soybean: *GmFT1a* and *GmFT2a/5a* have opposite roles in controlling flowering and maturation. New Phytol. 217. doi:10.1111/nph.14884.

Lombardi, M., Materne, M., Cogan, N.O.I., Rodda, M.S., Daetwyler, H.D., Slater, A.T., Forster, J.W., and Kaur, S. 2014. Assessment of genetic variation within a global collection of lentil (*Lens culinaris* Medik.) cultivars and landraces using SNP markers. BMC Genet. 15. doi:10.1186/s12863-014-0150-3.

Lutz, U., Posé, D., Pfeifer, M., Gundlach, H., Hagmann, J., Wang, C., Weigel, D., Mayer, K.F.X., Schmid, M., and Schwechheimer, C. 2015. Modulation of Ambient Temperature-Dependent Flowering in *Arabidopsis thaliana* by Natural Variation of *FLOWERING LOCUS M*. PLoS Genet. 11: 1–26. doi:10.1371/journal.pgen.1005588.

Meng, L., Li, H., Zhang, L., and Wang, J. 2015. QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. Crop J. 3: 269–283. doi:10.1016/j.cj.2015.01.001.

Nelson, M.N., Ksia, M., Rychel, S., Besharat, N., Taylor, C.M., Wyrwa, K., Jost, R., Erskine, W., Cowling, W.A., Berger, J.D., Batley, J., Weller, J.L., Naganowska, B., and Wolko, B. 2017. The loss of vernalization requirement in narrow-leafed lupin is associated with a deletion in the promoter and de-repressed expression of a *Flowering Locus T* (*FT*) homologue. New Phytol. 213: 220–232. doi:10.1111/nph.14094.

Neupane, S. 2019. Flowering Time Response Of Diverse Lentil (*Lens Culinaris* Medik.) Germplasm Grown In Multiple Environments. MSc. Thesis. University of Saskatchewan.

Ogutcen, E., Ramsay, L., von Wettberg, E.B., and Bett, K.E. 2018. Capturing variation in *Lens* (*Fabaceae*): Development and utility of an exome capture array for lentil. Appl. Plant Sci. 6: e1165. doi:10.1002/aps3.1165.

Ortega, R. 2018. The role of the *FT* genes in the control of flowering in chickpea. PhD. Thesis. University of Tasmania.

Ortega, R., Hecht, V.F.G., Freeman, J.S., Rubio, J., Carrasquilla-garcia, N., Mir, R.R., Penmetsa, R.V., Vallejos, C.E., Cook, D.R., Millán, T., and Weller, J.L. 2019. Altered Expression of an *FT* Cluster Underlies a Major Locus Controlling Domestication-Related Changes to Chickpea Phenology and Growth Habit. Front. Plant Sci. 10. doi:10.3389/fpls.2019.00824.

Pin, P.A., and Nilsson, O. 2012. The multifaceted roles of *FLOWERING LOCUS T* in plant development. Plant, Cell Environ. 35: 1742–1755. doi:10.1111/j.1365-3040.2012.02558.x.

Podder, R., Banniza, S., and Vandenberg, A. 2013. Screening of wild and cultivated lentil germplasm for resistance to stemphylium blight. Plant Genet. Resour. 11: 26–35. doi:10.1017/S1479262112000329.

Polanco, C., Miera De, L.E., Gonzalez, A.I., García, P., Fratini, R., Vaquero, F., Vences, F.J., and Perez, M. 2019. Construction of a high-density interspecific (*Lens culinaris* x *L.odemensis*) genetic map based on functional markers for mapping morphological and agronomical traits, and QTLs affecting resistance to Ascochyta in lentil. PLoS One 14. doi:journal.pone.0214409.

Putterill, J., Zhang, L., Yeoh, C.C., Balcerowicz, M., Jaudal, M., and Gasic, E.V. 2013. *FT* genes and regulation of flowering in the legume *Medicago truncatula*. Funct. Plant Biol. 40: 1199–1207. doi:10.1071/FP13087.

Putterill, J., and Varkonyi-Gasic, E. 2016. *FT* and florigen long-distance flowering control in plants. Curr. Opin. Plant. Biol. 33:77-82. doi: 10.1016/j.pbi.2016.06.008.

Rafalski, A. 2002. Applications of single nucleotide polymorphisms in crop genetics. Curr. Opin. Plant Biol. 5: 94–100. doi:10.1016/S1369-5266(02)00240-6.

Rajandran, V. 2016. Genetic control of flowering time in lentil. PhD. Thesis. University of Tasmania.

Rajandran, V., Ortega, R., Vander Schoor, J.K., Butler, J.B., Freeman, J.S., Hecht, V.F.G., Erskine, W., Murfet, I.C., Bett, K.E., and Weller, J.L. 2021. Genetic analysis of early phenology in lentil identifies distinct loci controlling component traits. bioRxiv. doi:10.1101/2021.08.16.456438.

Ramsay, L., Koh, C.S., Kagale, S., Gao, D., Kaur, S., Gela, T.S., Chen, L., Cao, Z., Konkin, D.J., Toegelová, H., McGee, R.J., Cook, D.R., Penmetsa, R.V., and Vandenberg, A. 2021. Genomic rearrangements have consequences for introgression breeding as revealed by genome assemblies of wild and cultivated lentil species. bioRxiv. doi:10.1101/2021.07.23.453237.

Ridge, S., Deokar, A., Lee, R.H., Daba, K., Macknight, R.C., Weller, J.L., and Tar'an, B. 2017. The chickpea *Early Flowering 1* (*Efl1*) locus is an ortholog of *Arabidopsis ELF3*. Plant Physiol. 175: 802–815. doi:10.1104/pp.17.00082.

Ridge, S., Sussmilch, F.C., Hecht, V.F.G., Vander Schoor, J.K., Lee, R.H., Aubert, G., Burstin, J., Macknight, R.C., and Weller, J.L. 2016. Identification of *LATE BLOOMER2* as a *CYCLING DOF FACTOR* Homolog Reveals Conserved and Divergent Features of the Flowering Response to Photoperiod in Pea. Plant Cell 28: 2545–2559. doi:10.1105/tpc.15.01011.

Roberts, E.H., Summerfield, R.J., Ellis, R.H., and Stewart, K.A. 1988. Photothermal Time for Flowering in Lentils (*Lens culinaris*) and the Analysis of Potential Vernalization Responses. Ann. Bot. 61: 29–39. doi:10.2307/42758664.

Roberts, E.H., Summerfield, R.J., Muehlbauer, F.J., and Short, R.W. 1986. Flowering in lentil (*Lens culinaris* Medic.): the duration of the photoperiodic inductive phase as a function of accumulated daylength above the critical photoperiod. Ann. Bot. 58: 235–248. [Online] Available: http://aob.oxfordjournals.org/content/58/2/235.

Saha, G.C., Sarker, A., Chen, W., Vandemark, G.J., and Muehlbauer, F.J. 2013. Inheritance and Linkage Map Positions of Genes Conferring Agromorphological Traits in *Lens culinaris* Medik. Int. J. Agron. 2013. [Online] Available: http://dx.doi.org/10.1155/2013/618926.

Saha, G.C., Sarker, A., Chen, W., Vandemark, G.J., Muehlbauer, F.J., and Murat, F. 2010. Inheritance and Linkage Map Positions of Genes Conferring Resistance to Stemphylium Blight in Lentil. Crop Sci. 50: 1831–1839. doi:10.2135/cropsci2009.12.0709.

Salvi, S., Sponza, G., Morgante, M., Tomes, D., Niu, X., Fengler, K.A., Meeley, R., Ananiev, E. V, Svitashev, S., Bruggemann, E., Li, B., Hainey, C.F., Radovic, S., Zaina, G., Rafalski, A., Tingey, S. V, Miao, G.-H.H., Phillips, R.L., and Tuberosa, R. 2007. Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize. Proc. Natl. Acad. Sci. U. S. A. 104: 11376–11381. doi:0704145104 [pii]\r10.1073/pnas.0704145104.

Sansaloni, C., Petroli, C., Jaccoud, D., Carling, J., Detering, F., Grattapaglia, D., and Kilian, A. 2011. Diversity Arrays Technology (DArT) and next- generation sequencing combined : genome-wide, high throughput, highly informative genotyping for molecular breeding of Eucalyptus. 5. doi:10.1186/1753-6561-5-S7-P54.

Sarker, A., and Erskine, W. 2006. Recent progress in the ancient lentil. J. Agric. Sci. 144: 19–29. doi:10.1017/S0021859605005800.

Sarker, A., Erskine, W., Hassan, M.S., Afzal, M.A., and Murshed, A.N.M.M. 1999a. Registration of "Barimasur-4" Lentil. Crop Sci. 39: 876.

Sarker, A., Erskine, W., Sharma, B., and Tyagi, M.C. 1999b. Inheritance and linkage relationship of days to flower and morphological loci in lentil (*Lens culinaris* Medikus subsp. *culinaris*). J. Hered. 90: 270–275. doi:10.1093/jhered/90.2.270.

SciGenom 2012. Exome Sequencing and Data Analysis - (Whitepaper). SciGenom.

Sharma, S., Upadhyaya, H.D., Varshney, R.K., and Gowda, C.L.L. 2013. Pre-breeding for diversification of primary gene pool and genetic enhancement of grain legumes. Front. Plant Sci. 4. doi:10.3389/fpls.2013.00309.

Sharpe, A.G., Ramsay, L., Sanderson, L.-A., Fedoruk, M.J., Clarke, W.E., Li, R., Kagale, S., Vijayan, P., Vandenberg, A., and Bett, K.E. 2013. Ancient orphan crop joins modern era: gene-based SNP discovery and mapping in lentil. BMC Genomics 14: 192. doi:10.1186/1471-2164-14-192.

Singh, D., Singh, C.K., Tomar, R.S.S., Taunk, J., Singh, R., Maurya, S., Chaturvedi, A.K., Pal, M., Singh, R., and Dubey, S.K. 2016. Molecular Assortment of Lens Species with Different Adaptations to Drought Conditions Using SSR Markers. PLoS One 11: e0147213. doi:10.1371/journal.pone.0147213.

Singh, M., Bisht, I.S., Kumar, S., Dutta, M., Bansal, K.C., Karale, M., Sarker, A., Amri, A., Kumar, S., and Datta, S.K. 2014. Global wild annual lens collection: A potential resource for lentil genetic base broadening and yield enhancement. PLoS One 9. doi:10.1371/journal.pone.0107781.

Sita, K., Sehgal, A., Bhandari, K., Kumar, J., Singh, S., Siddique, H.M., and Nayyar, H. 2018. Impact of heat stress during seed filling on seed quality and seed yield in lentil *(Lens culinaris* Medikus) genotypes. J Sci Food Agric 98: 5134–5141. doi:10.1002/jsfa.9054.

Sonnante, G., Hammer, K., Pignone, D., Sonnante, G., Hammer, K., and Pignone, D. 2009. From the cradle of agriculture a handful of lentils: History of domestication. Rend. Lincei 20: 21–37. doi:10.1007/s12210-009-0002-7.

Springate, D.A., and Kover, P.X. 2014. Plant responses to elevated temperatures: A field study on phenological sensitivity and fitness responses to simulated climate warming. Glob. Chang. Biol. 20: 456–465. doi:10.1111/gcb.12430.

Srikanth, A., and Schmid, M. 2011. Regulation of flowering time: All roads lead to Rome. Cell. Mol. Life Sci. 68: 2013–2037. doi:10.1007/s00018-011-0673-y.

Stein, N., Davis, S.J., Kilian, B., Schneeberger, K., Sharma, R., Saini, R., Pankin, A., Himmelbach, A., Dong, X., Campoli, C., and von Korff, M. 2014. Mapping-by-Sequencing Identifies *HvPHYTOCHROME C* as a Candidate Gene for the *Early Maturity 5 Locus* modulating the circadian clock and photoperiodic flowering in barley. Genetics 198: 383–396. doi:10.1534/genetics.114.165613.

Sudheesh, S., Verma, P., Forster, J.W., Cogan, N.O.I., and Kaur, S. 2016. Generation and characterisation of a reference transcriptome for lentil (*Lens culinaris* Medik.). Int. J. Mol. Sci. 17: 1887. doi:10.3390/ijms17111887.

Summerfield, R.J., Roberts, E.H., Erskine, W., and Ellis, R.H. 1985. Effects of Temperature and Photoperiod on Flowering in Lentils (*Lens culinaris* Medic.). Ann. Bot. 56: 659–671. doi:10.1093/oxfordjournals.aob.a087055.

Tar'an, B., Buchwaldt, L., Tullu, A., Banniza, S., Warkentin, T.D., and Vandenberg, A. 2003. Using molecular markers to pyramid genes for resistance to ascochyta blight and anthracnose in lentil (*Lens culinaris* Medik). Euphytica 134: 223–230. doi:10.1023/B:EUPH.0000003913.39616.fd.

Thomson, G., Taylor, J., and Putterill, J. 2019. The transcriptomic response to a short day to long day shift in leaves of the reference legume *Medicago truncatula*. PeerJ: 7:e6626. doi:10.7717/peerj.6626.

Tullu, A., Bett, K.E., Banniza, S., Vail, S.L., and Vandenberg, A. 2013. Widening the genetic base of cultivated lentil through hybridization of *Lens culinaris* "Eston" and *L.ervoides* accession IG 72815. Can. J. Plant Sci. 93: 1037–1047. doi:10.4141/cjps2013-072.

Tullu, A., Diederichsen, A., Suvorova, G., and Vandenberg, A. 2011. Genetic and genomic resources of lentil: status, use and prospects. Plant Genet. Resour. 9: 19–29. doi:10.1017/S1479262110000353.

Tullu, A., Kusmenoglu, I., Mcphee, K.E., and Muehlbauer, F.J. 2001. Characterization of core collection of lentil germplasm for phenology, morphology, seed and straw yields. Genet. Resour. Crop Evol. 48: 143–151. doi:10.1023/A:1011254629628.

Tullu, A., Tar'an, B., Warkentin, T.D., and Vandenberg, A. 2008. Construction of an intraspecific linkage map and QTL analysis for earliness and plant height in lentil. Crop Sci. 48: 2254–2264. doi:10.2135/cropsci2007.11.0628.

Turck, F., Fornara, F., and Coupland, G. 2008. Regulation and Identity of Florigen: *FLOWERING LOCUS T* Moves Center Stage. Annu. Rev. Plant Biol. 59: 573–594. doi:10.1146/annurev.arplant.59.032607.092755.

Upadhyaya, H.D., Bajaj, D., Das, S., Saxena, M.S., Badoni, S., Kumar, V., Tripathi, S., Gowda, C.L.L., Sharma, S., Tyagi, A.K., and Parida, S.K. 2015. A genome-scale integrated approach aids in genetic dissection of complex flowering time trait in chickpea. Plant Mol. Biol. 89: 403–420. doi:10.1007/s11103-015-0377-z.

Vadez, V., Berger, J.D., Warkentin, T.D., Asseng, S., Ratnakumar, P., Rao, K.P.C., Gaur, P.M., Munier-Jolain, N., Larmure, A., Voisin, A.S., Sharma, H.C., Pande, S., Sharma, M., Krishnamurthy, L., and Zaman, M.A. 2012. Adaptation of grain legumes to climate change: A review. Agron. Sustain. Dev. 32: 31–44. doi:10.1007/s13593-011-0020-6.

Vandenberg, A., Kiehn, F.A., Vera, C., Gaudiel, R., Buchwaldt, L., Kirkland, K.J., Morrall, R.A.A., Wahab, J., and Slinkard, A.E. 2001. Cultivar description CDC Milestone lentil. Can. J. Plant Sci. 81: 113–114.

Verma, P., Goyal, R., Chahota, R.K., Sharma, T.R., Abdin, M.Z., and Bhatia, S. 2015. Construction of a genetic linkage map and identification of QTLs for seed weight and seed size traits in lentil (*Lens culinaris* Medik.). PLoS One 10: e0139666. doi:10.1371/journal.pone.0139666.

Watson, M., Hume, D., Archibald, A., Deeb, N., Warr, A., and Robert, C. 2015. Exome Sequencing: Current and Future Perspectives. G3 Genes|Genomes|Genetics 5: 1543–1550. doi:10.1534/g3.115.018564.

Welch, S.M., Dong, Z., Roe, J.L., and Das, S. 2005. Flowering time control: Gene network modelling and the link to quantitative genetics. Aust. J. Agric. Res. 56: 919–936. doi:10.1071/AR05155.

Weller, J.L., Liew, L.C., Hecht, V.F.G., Rajandran, V., Laurie, R.E., Ridge, S., Wenden, B., Vander Schoor, J.K., Jaminon, O., Blassiau, C., Dalmais, M., Rameau, C., Bendahmane, A., Macknight, R.C., and Lejeune-Hénaut, I. 2012. A conserved molecular basis for photoperiod adaptation in two temperate legumes. Proc. Natl. Acad. Sci. U. S. A. 109: 21158–21163. doi:10.1073/pnas.1207943110.

Weller, J.L., and Ortega, R. 2015. Genetic control of flowering time in legumes. Front. Plant Sci. 6: 207. doi:10.3389/fpls.2015.00207.

Weller, J.L., Reid, J.B., Taylor, S.A., and Murfet, I.C. 1997. The genetic control of flowering in pea. Trends Plant Sci. 22: 412–418. doi:10.1016/S1360-1385(97)01127-8.

Wenden, B., Dun, E.A., Hanan, J., Andrieu, B., Weller, J.L., Beveridge, C.A., and Rameau, C. 2009. Computational analysis of flowering in pea (*Pisum sativum*). New Phytol. 184: 153–167. doi:10.1111/j.1469-8137.2009.02952.x.

Wong, A.C.S., Hecht, V.F.G., Picard, K., Diwadkar, P., Laurie, R.E., Wen, J., Mysore, K.S., Macknight, R.C., and Weller, J.L. 2014. Isolation and functional analysis of *CONSTANS-LIKE* genes suggests that a central role for *CONSTANS* in flowering time control is not evolutionarily conserved in *Medicago truncatula*. Front. Plant Sci. 5: 486. doi:10.3389/fpls.2014.00486.

Wong, M.M.L., Gujaria-Verma, N., Ramsay, L., Yuan, H.Y., Caron, C.T., Diapari, M., Vandenberg, A., and Bett, K.E. 2015. Classification and characterization of species within the genus *Lens* using genotyping-by-sequencing (GBS). PLoS One 10. doi:10.1371/journal.pone.0122025.

Wright, D.M., Neupane, S., Heidecker, T., Haile, T.A., Chan, C., Coyne, C.J., McGee, R.J., Udupa, S.M., Henkrar, F., Barilli, E., Rubiales, D., Gioia, T., Logozzo, G., Marzario, S., Mehra, R., Sarker, A., Dhakal, R., Anwar, B., Sarkar, D., Vandenberg, A., and Bett, K.E. 2020. Understanding photothermal interactions will help expand production range and increase genetic diversity of lentil (*Lens culinaris* Medik.). Plants People Planet 2020. doi:10.1002/ppp3.10158.

Yuan, H.Y., Caron, C.T., Ramsay, L., Fratini, R., De, M.P., Vandenberg, A., Weller, J.L., and Bett, K.E. 2021. Genetic and gene expression analysis of flowering time regulation by light quality in lentil. Ann. Bot. 128: 481–496. doi:https://doi.org/10.1093/aob/mcab083.

Yuan, H.Y., Saha, S., Vandenberg, A., and Bett, K.E. 2017. Flowering and Growth Responses of Cultivated Lentil and Wild Lens Germplasm toward the Differences in Red to Far-Red Ratio and Photosynthetically Active Radiation. Front. Plant Sci. 8: 386. doi:10.3389/fpls.2017.00386.

Zhai, H., Lü, S., Liang, S., Wu, H., Zhang, X., Liu, B., Kong, F., Yuan, X., Li, J., and Xia, Z. 2014. *GmFT4*, a homolog of *FLOWERING LOCUS T*, is positively regulated by *E1* and functions as a flowering repressor in soybean. PLoS One 9. doi:10.1371/journal.pone.0089030.

Zhang, X., Zhai, H., Wang, Y., Tian, X., Zhang, Y., Wu, H., Lü, S., Yang, G., Li, Y., Wang, L., Hu, B., Bu, Q., and Xia, Z. 2016. Functional conservation and diversification of the soybean maturity gene *E1* and its homologs in legumes. Sci. Rep. 6: 1–14. doi:10.1038/srep29548.

Zohary, D. 1972. The Wild Progenitor and the Place of Origin of the Cultivated Lentil : *Lens culinaris*. Econ. Bot. 26: 326–332. [Online] Available: http://www.jstor.org/stable/4253374.

Appendices

Prologue to Appendix 1. Preliminary Experiment

324 lentil accessions, gathered to represent the current diversity in cultivated lentil, were evaluated for days to flowering (DTF) in long day greenhouse (GH) and Saskatchewan field experiments.

The three GH experiments were maintained by Devini De Silva and others from the Crop Development Centre (CDC). Phenotypes used here collected by Derek Wright (F15, W16, F16), Sandesh Neupane (W16, F16) and myself (F15, F16).

The four site-years of field experiment were phenotyped by Derek Wright, Sandesh Neupane and myself. The GDD calculation used as a transformation here was found to be the best for comparing across environments in Mr. Neupane's MSc. thesis (Neupane 2019). The phenotypic field data used here also appear in (Wright et al. 2020).

Appendix 1. Preliminary Experiment: Characterising diverse lentils for days to flowering in the greenhouse and field

A.1.1 Introduction

Lentil (*Lens culinaris*) currently ranks fourth among grain legumes in terms of human consumption and is valued for its high protein content and nutritional profile (Zhang 2015, FAOSTAT 2019). The single largest lentil producing country today is Canada, and a large part of this has been made possible through the breeding which has taken place at the University of Saskatchewan (USask). However, lentils originated from within the fertile crescent in what is now Syria and Turkey where growing conditions are vastly different than those in Canada (Alo et al. 2011). The genetic adaptations which enable successful lentil production in Canada are not the same those required in other parts of the world. This reduces the genetic diversity available to the USask breeding program, as it is expensive to evaluate large numbers of unadapted lines and their progeny.

Sarker et al. (1999) identified that 'days to flowering' (DTF) in lentil is the most important trait related to adaptation to an environment. Lentil breeders find DTF a useful indicator for days to maturity and fitness (i.e. potential yield) (Tullu et al. 2008). More broadly, in all legume crops, DTF is the most important adaptation made when moving germplasm among environmental

regions, as it also allows for the appropriate phenological responses required to overcome numerous abiotic stresses, including heat and drought stress (Vadez et al. 2012). Research conducted in multiple climatic regions using a diverse panel of lentils determined the best way to define time to flowering in the field across the major production regions of the world involves considerations for photoperiod and temperature, with variable interactions influenced by the genetic background of tested lentil accessions (Neupane 2019).

A limited number of studies involving greenhouse (GH) and field replications also find success in explaining DTF as a phenological response to photoperiod and temperature (Summerfield et al. 1985; Roberts et al. 1988). Like studies in other climatic regions, these conditions are not the same as experienced by lentil when grown in western Canada and should be investigated. Early generations of germplasm assessment are often done in the greenhouse, and results gathered DTF may translate poorly to Saskatchewan field conditions due to environmental constraints which are not present in the more controlled indoor environment.

This study was done to test if cues for DTF are the same in the GH and field, with the controlled environmental conditions of the GH leading to a more stable response in DTF than in the field. Also, it is surmised that accounting for differences in temperature and photoperiod will help in comparing between experiments. These are addressed by characterizing DTF in the AGILE Lentil Diversity Panel (AGILE-LDP) grown in long day GH conditions, as well as a Saskatchewan field setting. Added comparisons, involving simple transformed versions (proxies) of DTF were included to evaluate differences between the GH replications and Saskatchewan field experiments.

A.1.2 Materials and Methods

A.1.2.1 Plant Material

The AGILE Lentil Diversity Panel (AGILE-LDP) of 324 cultivated lentils representing much of cultivated lentil diversity from around the world was used for this experiment. The accession names and labeled place of origin are included in https://knowpulse.usask.ca/Heidecker-Thesis-Association-Data; Supplementary 3. The seed source for the field experiments were derived from seed bulked from 1-2 phenotypically uniform plants from the first replicate of the GH trial, which was used to increase the seed from multiple sources.

A.1.2.2 Phenotyping

a) Indoor Greenhouse Trials

The experiment was conducted in the Agriculture Greenhouses on the University of Saskatchewan campus (room C2) 3 times. There was only space for one replicate at a time: F15 was planted September 3rd and 9th, 2015, W16 was planted January 25th, 2016, and F16 was planted October 12th, 2016.

3-4 seeds were sown in an 8-inch pot (~20.5cm) and then thinned to 2 plants per pot at the seedling stage. The plants were grown under long day conditions with supplemental lighting used to achieve seventeen-hour days, with temperature set for 22/18^{o}C (day/night). Each experiment was seeded at a different date, and therefore experienced a different amount of supplemental lighting.

Potting mix, fertility and watering requirements were standard for lentil plant growth in the pulse program at the U of S. The population was checked every one to three days for emergence, tendril elongation (W16, F16), first flower, first pod (F15, W16)/first swollen pod (F16), and dry pod. One line, ILL 5480 did not successfully reach flowering in F15. In both W16 and F16 two lines did not grow successfully (ILL 4671 and ILWL 118). Experimental conditions are plotted in Figure A.1.1.

b) Outdoor Field Trials

The Saskatchewan AGILE LDP field trials took place during the summers of 2016 and 2017. The LDP population was grown in 1m^2 square microplots, replicated three times at each of two sites in North-central Saskatchewan: Sutherland (GPS: 52.17, -106.51) and Rosthern (GPS: 52.68, -106.29). The 2016 the Sutherland site was planted April 26th and Rosthern was planted May 6th. In 2017, the Sutherland site was seeded on May 4th and Rosthern was seeded on May 19th.

Notes were taken every 1-3 days on a wide range of traits. Days to emergence (DTE), and days to 10% of plants with an opened flower (DTF) were the primary focus for comparing to the greenhouse phenotypes.

A.1.2.3 Transformation of DTF

a) Indoor Greenhouse Trials

Light intensity (reported as photosynthetic active radiation (P.A.R.; µmol/m^2s) and temperature (°C) were recorded for room C2 every hour, with the raw data obtained from https://agbio.usask.ca/research/centres-and-facilities/greenhouses.php. As slight deviations from the pre-set temperature and daylength were observed, DTF transformations used recorded temperature and hours of 'day-time' as those where PAR was >50 (Figure A.1.1).

To account for differences in days to emergence and actual temperatures experienced in the GH, DTF data were transformed to Vegetative Period (VegP), Photothermal time to Flowering (DTF.PTT), and Vegetative Photothermal Time (VegP.PTT) using the equations (Eq.) A.1.1 – A.1.3, where *ft* is a plot's flowering date (DTF), *j* is the date a plot emerged (DTE), and *i* is the date the experiment was planted.

$$\mathbf{VegP} = ft - j \quad \text{[A.1.1]}$$

$$\mathbf{DTF.PTT} = \sum_{i}^{ft}[(\text{Hourly Daytime Temperature})/24] \quad \text{[A.1.2]}$$

$$\mathbf{VegP.PTT} = \sum_{j}^{ft}[(\text{Hourly Daytime Temperature})/24] \quad \text{[A.1.3]}$$

b) Outdoor Field Trials

Growing Degree Days to Flowering (GDD.DTF) is the best environmental transformation for comparing days to flowering in multiple field experiments experiencing similar photoperiods (Neupane MSc. Thesis 2019). DTF in the field were transformed to VegP and GDD.DTF for comparison to GH experiments. P.A.R. values recorded by the exterior weather station located at the USask greenhouses were used for comparison of summer conditions experienced during field experiments to the GH trials.

$$\mathbf{GDD} = \sum_{i}^{ft}[\frac{(\text{Max.Temp.}) - (\text{Min.Temp.})}{2} - 5°\text{C}] \quad \text{[A.1.4]}$$

A.1.2.4 Statistical Analysis

Statistical analyses were done in R v3.4.4. (R Development Core Team 2018). Simple ANOVA using the built in "aov" function, as well as tests using "lmer" (from *lme4* v1.1-15 (Bates et al. 2015), were done to test for variation across experiments. Variation between the GH experiments

was always significant, even after accounting for genotype, and therefore experiments were kept separate for subsequent analyses.

Outdoor field trials, Sutherland 2016 (S16), Rosthern 2016 (R16), Sutherland 2017 (S17), and Rosthern 2017 (R17) were treated as four separate site-years (experiments). ANOVAs indicated variation between replications within a site-year were insignificant, and replication means were used in analyses.

A.1.3 Results

A.1.3.1 Within-Experiment Variability and Correlations for DTF in Greenhouse and Field

The greenhouse experiments (GH) received a lower level of light intensity than the summer field trials, with the GH day-time P.A.R mean only exceeding 400 µmol/m^2s in Winter 2016 (W16), and only on seven days prior to the last accession flowering. W16 was the only GH experiment to experience a P.A.R. maximum over 650 µmol/m^2s, a typical day-time P.A.R mean during the summer field trials. Conversely, the GH experienced a more stable temperature than field experiments, with relatively fixed daytime means and minimums never below 10°C. While later season day-time mean temperatures in the field were consistently above 15°C, early season temperature could be quite variable (Figure A.1.1).

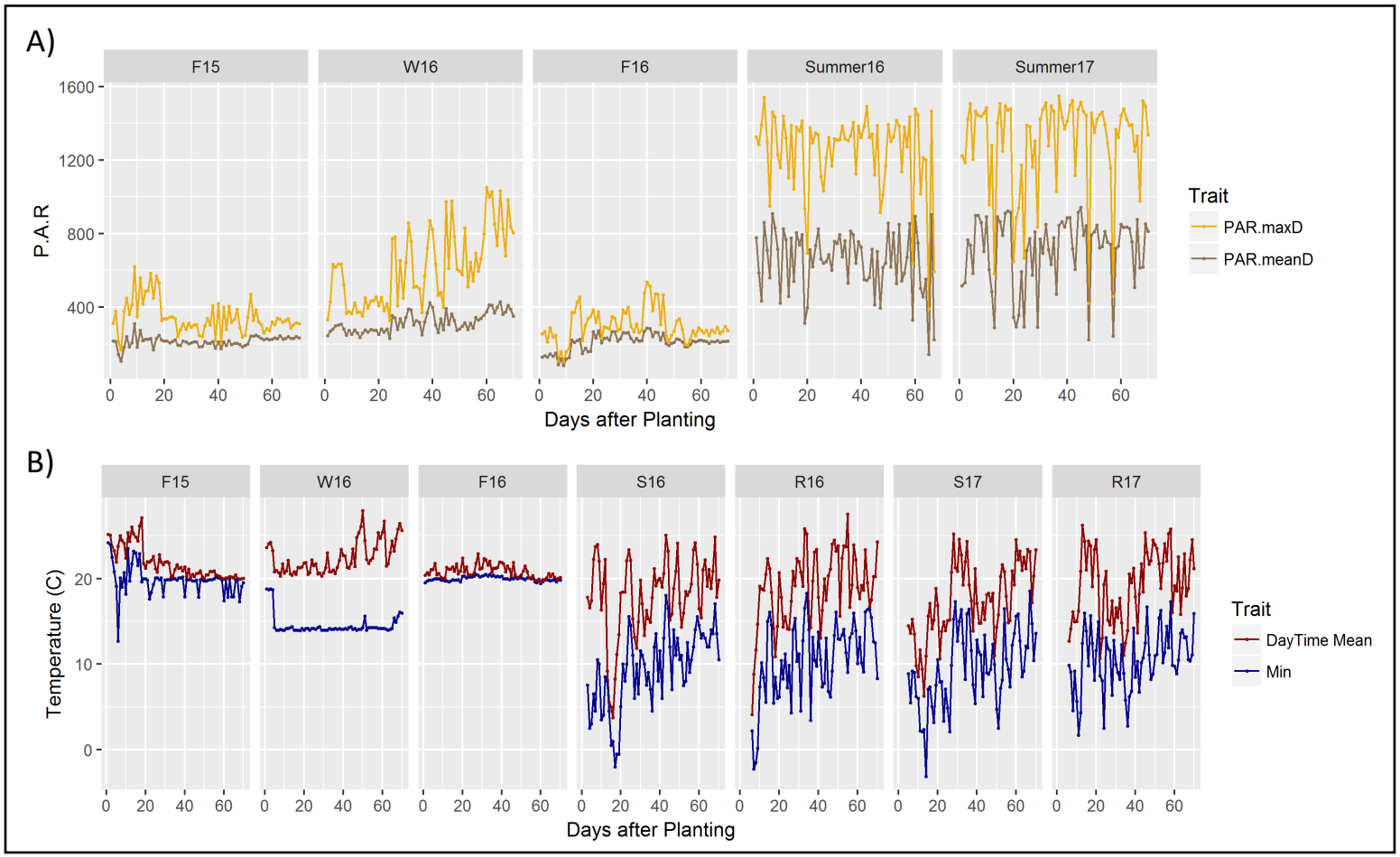

Figure A.1.1. Experimental light quality and temperature conditions from seeding to flowering. A) daily Photosynthetically Active Radiation (P.A.R; µmol/m^2s) maximum (PAR.maxD) and daytime mean (PAR.meanD) are plotted, comparing greenhouse; Fall 2015 (F15), Winter 2016 (W16), Fall 2016 (F16), and field (Summer16/17) experiments. **B)** Day-time mean temperature (°C) (DayTime Mean) as well as daily minimum temperature (Min) are also plotted for both greenhouse; F15, W16, F16, and field; Sutherland 2016/2017 (S16/17) and Rosthern 2016/2017 (R16/17) experiments.

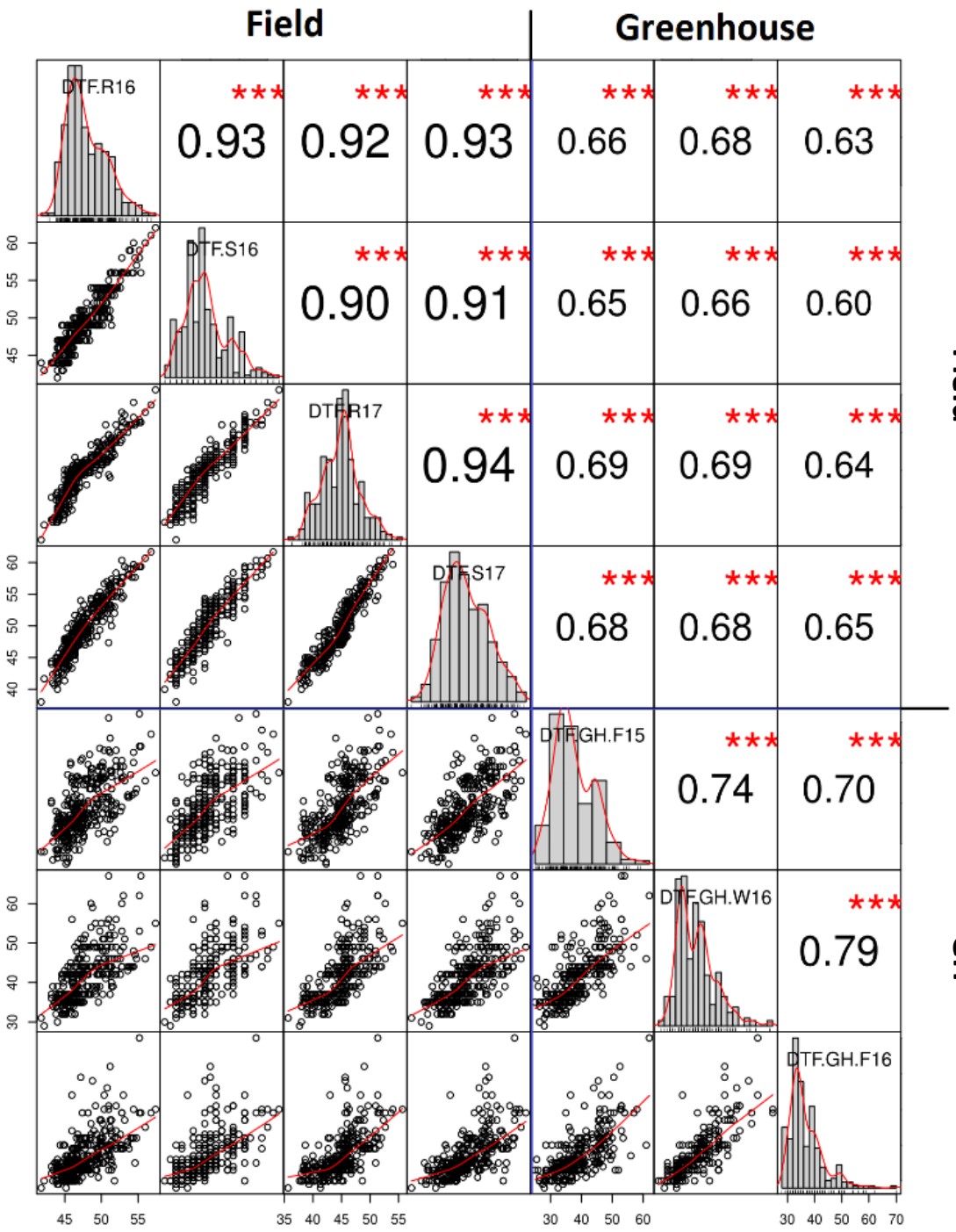

Figure A.1.2. Field and Greenhouse (GH) correlations for days to flower (DTF). Significance was determined using Pearson's correlation coefficient. While correlations between individual experiments ranged from 0.60 to 0.93, all experiments were highly correlated (*** = p< 2.2e-16).

Correlations for DTF among all experiments were significant (p< 2.2e-16) (Figure A.1.2). GH to GH correlation coefficients ranged from 0.70 to 0.79 compared to the field-to-field correlations

which were all 0.90 and greater. Comparisons of the two types of experiments resulted in correlations for DTF between 0.60 to 0.69 between the GH and field experiments.

DTF was more variable in the GH than the field (Figure A.1.3), with within-experiment variability for GH experiments ranging from 15.9 to 18.1% compared to 6.1% to 9.2% in the field experiments (Table A.1.2).

A.1.3.2 Evaluation of transformed 'days to flower'

Using Vegetative Period (VegP) to represent development in the GH experiments never reduced correlations to the field trials (DTF or proxy representatives), and generally increased correlations between GH experiments (Table A.1.2). Correlation values for both GH-DTF and GH-VegP to the field experiments were slightly improved when the field experiments were represented using Growing Degree Days (GDD) (Table A.1.1).

There was no consistent increase in correlation (r) values across GH experiments when using the Photothermal (.PTT) transformation, of DTF or VegP. Of minor note, using VegP for the field experiments consistently increased variability within-experiment and reduced correlation coefficients between field experiments (to as low as 0.71) and when compared to the GH experiments (Table A.1.1).

Table A.1.1. Correlations (r) of the field experiment Sutherland 2017 (S17) to the greenhouse (GH) experiments F15, W16, and F16 for days to flowering (DTF) and representative phenotypes. Representatives for DTF include vegetative period (VegP), Photothermal time for greenhouse experiments (PTT), and Growing Degree Days for S17 (GDD). All correlations are highly significant (p< 2.2e-16).

		DTF			VegP			DTF.PTT			VegP.PTT		
		F15	W16	F16	F15	W16	F16	F15	W16	F16	F15	W16	F16
S17	**DTF**	0.68	0.68	0.65	0.68	0.71	0.65	0.69	0.68	0.66	0.69	0.71	0.66
	VegP	0.66	0.66	0.63	0.66	0.70	0.64	0.66	0.65	0.64	0.66	0.69	0.64
	GDD	0.69	0.69	0.66	**0.69**	**0.72**	**0.66**	0.68	0.68	0.66	0.68	0.71	0.67

Table A.1.2. 'Within-experiment variability', calculated as the size of standard deviation to the mean (%) , for days to flowering (DTF) and representative phenotypes. Included are the four field experiment site-years; Sutherland 2016 (S16) and 2017 (S17) and Rosthern 2016 (R16) and 2017 (R17) as well as the three greenhouse experiments (GH); Fall 2015 (F15), Winter 2016 (W16) and Fall 2016 (F16). Representatives for DTF include vegetative period (VegP), Photothermal time for greenhouse experiments (PTT), and Growing Degree Days for field experiments (GDD).

	Experiment	Mean	Standard Deviation	'Within-Experiment Variability'
DTF	R16	47.72	2.917	6.114
	S16	49.34	3.779	7.659
	R17	44.71	3.426	7.662
	S16	49.51	4.574	9.238
	GH.F15	38.69	7.005	18.106
	GH.W16	41.29	6.547	15.856
	GH.F16	36.89	6.139	16.644
VegP	R16	23.56	4.253	18.049
	S16	23.44	6.200	26.451
	R17	35.39	3.511	9.920
	S16	36.77	4.709	12.806
	GH.F15	33.69	7.005	20.794
	GH.W16	34.432	5.731	16.645
	GH.F16	31.581	5.890	18.651
GDD/PTT	R16	559.154	44.247	7.913
	S16	591.995	51.874	8.763
	R17	496.108	48.690	9.814
	S16	482.600	45.895	9.510
	GH.F15	590.930	117.055	19.809
	GH.W16	648.459	104.215	16.071
	GH.F16	597.297	95.168	15.933

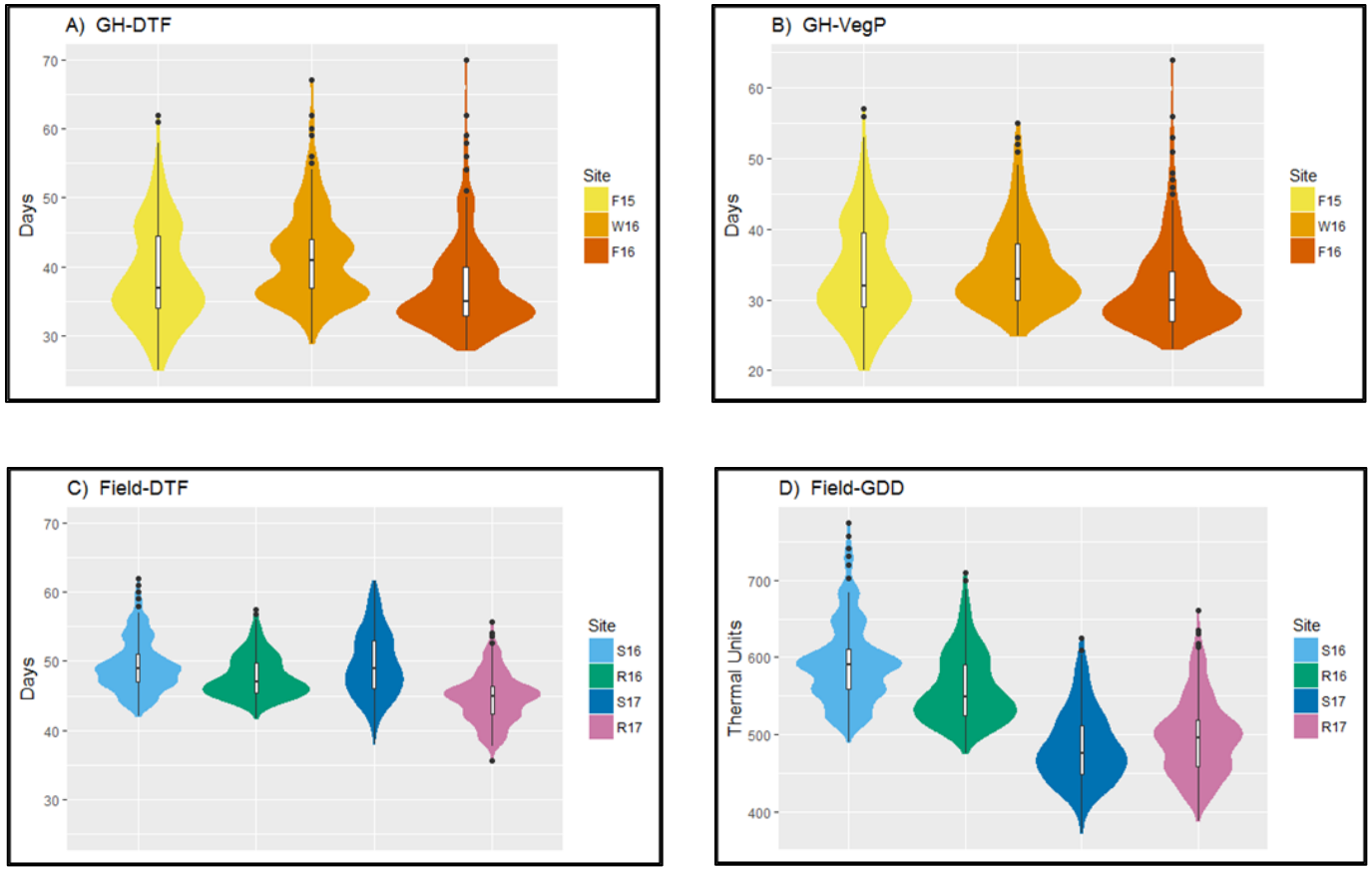

Figure A.1.3. Distribution of flowering (both DTF and best representative) among accessions of the AGILE-LDP in the GH and field. A) greenhouse (GH), days to flowering (DTF) were **B)** transformed to vegetative period (VegP). **C)** DTF in the field were **D)** transformed using GDD, with time to flowering being represented by thermal units. Experiments were conducted in the greenhouse during Fall 2015 (F15), Winter 2016 (W16), and Fall 2016 (F16). The field site-years were Sutherland 2016 (S16), Rosthern 2016 (R16), Sutherland 2017 (S17), and Rosthern 2017 (R17).

A.1.4 Discussion

A.1.4.1 Within-experiment variability reduces correlations for DTF in GH and field experiments

Correlations for days to flowering (DTF) between the greenhouse (GH) and Saskatchewan field experiments were moderate, and a coefficient of variation (r) greater than 0.69 was not obtained. This is likely due, in part, by the 'controlled environment' of the GH resulting in a larger variation for DTF than experienced in the field experiments (Figure A.1.3 A vs. C). This can be explained however, as the field experiments experienced much stronger (and stable) signals which drive flowering (light quality, daylength, and temperature (after early spring)). Lentil grown in different parts of the world responds differently to these signals due to the climatic regions they are best adapted for (Khazaei et al. 2016). Because of this, complexity of using GH data is compounded by GH experiments receiving different levels of these important signals, including varying levels of artificial light leading to differences in light quality across experiments (Figure A.1.1).

A.1.4.2 Correlations between field and GH experiments are complex due to the interaction of many factors

Previous lentil studies concur that variation about DTF, especially when considering diverse material in different environments, is accounted for best when considering temperature and photoperiod influences of the climate (Summerfield et al. 1985; Roberts et al. 1988; Erskine et al. 1994a). In comparing field trials of the AGILE-LDP in different climatic regions, accounting for temperature differences (growing degree day (GDD) proxy) was most useful for comparisons (Neupane 2019). Using the effect of both temperature and photoperiod and their interaction has effectively accounted for in some species with a photothermal model (PTT) (Springate and Kover 2014; Daba et al. 2016), however transformation of the GH data this way did not improve correlations as much as expected. Instead, removing the effect of emergence, and reporting VegP, was the most effective transformation used for DTF in the GH. Of note, it is suspected that variation in temperature within GH experiments existed, and as space within the room only allowed for one pot per accession, this might have added to differences in variation.

This experiment suggests the different environmental factors experienced in the GH compared to Saskatchewan field trials reduces the utility of GH study for predicting phenotypes in the field. Even though the photoperiods were of similar length, light quality (P.A.R.) and temperature were

much different between the two experiment types. In both lentil and related legumes, variation near flowering genes relevant to Saskatchewan field conditions has been linked to temperature responses, namely response to a cold period (Nelson et al. 2017; Ortega et al. 2019). Reports on vernalisation response in lentil are contested indicate it is overridden by photoperiod effects (Roberts et al. 1986; Weller et al. 2012; Rajandran et al. 2021). In this study, the GH experiment night-time temperatures were quite stable while field conditions experienced significant drops following sowing and prior to flowering (Figure A.1.1). Even though photoperiod requirements for all genotypes were fulfilled in both experiments (Wright et al. 2020), perhaps temperature-related variability is overridden to different degrees between the two experiment types.

The other major difference lentil experienced in the two experiment types was light quality (intensity). Similar to temperature, it is suspected that in diverse lentil there is variability for sensitivities to light intensity. Genotype specific responses to light quality, which includes photosynthetic active radiation (P.A.R), has recently been shown to influence flowering in lentil (Yuan et al. 2017, 2021). In comparing GH to field experiments in soybean, models for DTF already accounting for photoperiod and temperature were improved when a light quality parameter was included, and a threshold level was reached during the day-time of field experiments but only sometime met in the GH (Cober et al. 2014). Rate of photoperiod change was also included in the final DTF models in soybean and may be a useful way of incorporating photoperiod in predicting DTF in lentil across experiments.

Appendix 2. Expression of *LcFTa1* and *LcFTb2* relative the housekeeping gene *LcTIF* in the parents of LR-11 and nine RILs.

Sampling methods and gene expression measurement protocol can be found in chapter 3, section 3.2.7. Below in Appendix 2.A, the mean and standard deviation in expression level of *LcFTa1* is reported relative the housekeeping gene *LcTIF*. The nine RILs and parents are ordered based on the allelic state or 'score' at q.DTF.6-2 out of section 3.3.3, where 'A' refers to CDC Milestone allele and 'B' refers to the ILL 8006 allele. In Appendix 2.B, the same nine RILS and parents are ordered based on the allelic state at q.DTF.6-1 and expression of *LcFTb2*, relative *LcTIF*.

Appendix 2. A		***LcFTa1*** Relative *LcTIF* ($2^{\Delta ct}$)		
Score at q.DTF (qDTF6-1/2/3)	**Line Name**	**Mean**	**Standard Deviation**	**Significance Group**
AAA	CDC Milestone	0.507654	0.239507	c
AAA	LR-11-153	0.360735	0.03499	c
AAB	LR-11-94	1.592191	0.224662	b
BAA	LR-11-154	0.217615	0.073056	c
BAB	LR-11-122	0.33268	0.183824	c
ABA	LR-11-24	3.668327	0.81635	a
ABB	LR-11-47	2.573394	1.271646	ab
BBA	LR-11-58	1.341066	0.534475	b
BBB	LR-11-20	1.042294	0.619879	b
BBB	LR-11-15	1.405051	0.235295	b
BBB	ILL 8006	2.097759	0.47601	b
Appendix 2. B		***LcFTb2*** Relative *LcTIF* ($2^{\Delta ct}$)		
Score at q.DTF (qDTF6-1/2/3)	**Line Name**	**Mean**	**Standard Deviation**	**Significance Group**
AAA	CDC Milestone	0.239676	0.065802	e
AAA	LR-11-153	0.121885	0.025875	e
ABA	LR-11-24	1.648926	0.756718	bcd
AAB	LR-11-94	0.734239	0.127087	d
ABB	LR-11-47	1.53037	0.389438	c
BAA	LR-11-154	0.646854	0.340306	d
BBA	LR-11-58	3.74212	0.993307	ab
BAB	LR-11-122	2.338179	0.600735	bc
BBB	LR-11-20	4.152571	1.98824	ab
BBB	LR-11-15	5.622717	2.808812	ab
BBB	ILL 8006	6.419045	2.15456	a

Appendix 3. Estimated marker effect of FTa1-FTa2 and LcChr6-A1 for days to flowering (DTF) in the bi-parental RIL, LR-11 and the diversity panel, AGILE-LDP.

These estimates of marker effect come from the models compared in Table 4.2. The FTa1-FTa2 and LcChr6-A1 markers were tested using mixed linear models to see if they would be useful for predicting DTF. Models using each marker separately, and together, were tested against a null model for each experiment type. For LR-11 models, entry, entry:site year, and site year:replication were included as random terms. For the AGILE-LDP experiments, entry and Site/Site Year were included as random terms. The DTF-intercept is the estimated mean ($\bar{x}$) DTF returned by the model, followed by the standard error ($\sigma\bar{x}$). Estimated Marker Effect, reported for FTa1-FTa2 and LcChr6-A1 in the applicable models, is the adjustment in mean DTF (and standard error) between allelic states at the maker, as estimated by each model. In the AGILE-LDP, four accessions scored heterozygous at LcChr6-A1, the estimated effect of which is not reported here.

	LR-11 – Field Experiments			
			Estimated Marker Effect (Days)	
Model	df	DTF-Intercept	FTa1-FTa2	LcChr6-A1
		$\bar{x}$; $\sigma_{\bar{x}}$	**$\bar{x}$; $\sigma_{\bar{x}}$**	**$\bar{x}$; $\sigma_{\bar{x}}$**
Null	6	43.51; 0.95	--	--
FTa1-FTa2	7	43.92; 0.96	0.88; 0.27	--
LcChr6-A1	7	44.42; 0.95	--	2.26; 0.21
Both Loci	8	44.81; 0.95	0.85; 0.19	2.24; 0.19
	AGILE LDP- Field Experiments			
			Estimated Marker Effect (Days)	
Model	df	DTF-Intercept	FTa1-FTa2	LcChr6-A1
		$\bar{x}$; $\sigma_{\bar{x}}$	**$\bar{x}$; $\sigma_{\bar{x}}$**	**$\bar{x}$; $\sigma_{\bar{x}}$**
Null	4	47.83; 1.13	--	--
FTa1-FTa2	5	45.69; 1.18	2.77; 0.45	--
LcChr6-A1	6	45.56; 1.17	--	3.17; 0.41
Both Loci	7	44.26; 1.19	2.11; 0.43	2.72; 0.41
	AGILE LDP – GH Experiments			
			Estimated Marker Effect (Days)	
Model	df	DTF-Intercept	FTa1-FTa2	LcChr6-A1
		$\bar{x}$; $\sigma_{\bar{x}}$	**$\bar{x}$; $\sigma_{\bar{x}}$**	**$\bar{x}$; $\sigma_{\bar{x}}$**
Null	4	38.96; 1.32	--	--
FTa1-FTa2	5	35.73; 1.44	4.19; 0.75	--
LcChr6-A1	6	35.70; 1.40	--	4.57; 0.69
Both Loci	7	33.67; 1.47	3.31; 0.74	3.87; 0.69