THE UNIVERSITY of EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

The Genetic Basis for Natural Variation in Flowering Time in Local Populations of *Arabidopsis thaliana*

Hayley L McCulloch

Thesis Submitted for the degree of Doctor of Philosophy

University of Edinburgh

September 2011

Abstract

Factors affecting flowering time have been extensively studied for decades. Greater understanding of flowering time has wider implications in agriculture and ecology as the trait is crucial to optimising reproductive success. It is best understood in the genetic model *Arabidopsis thaliana* (Arabidopsis), in which loss and gain of function mutations have identified several pathways that regulate flowering and its response to the environment. This has been complemented by studies of natural variation in flowering. Worldwide accessions of Arabidopsis have been used to identify additional flowering regulators and to examine the evolution of these genes and their potential involvement in adaptation to different environments.

One of the most extensively studied pathways is responsible for accelerated flowering in response to an extended period of cold (vernalization). Several studies have attributed a substantial proportion of worldwide variation to the genes *FRIGIDA* (*FRI*) and *FLOWERING LOCUS C* (*FLC*), both of which are instrumental in conferring sensitivity to vernalization, though other genes have also been found.

This study examines flowering time variation locally in populations of Arabidopsis from in and around Edinburgh. It identifies substantial, genetically determined variation in flowering time and in sensitivity to photoperiod and vernalization between local accessions.

Variation in *FRI* and *FLC* sequences and in their levels of expression were detected in local accessions, but these were able to explain little of the phenotypic variation observed. Hybrids between local accessions showing extreme differences in flowering time or responses to photoperiod and vernalization were therefore used to map genes underlying their differences as quantitative trait loci (QTL). This analysis identified a locus in chromosome 5 that could account for differences in vernalization sensitivity. This region includes the *VERNALIZATION INSENSITIVE 3* (*VIN3*) gene. Sequence differences between *VIN3* alleles and their expression in response to vernalization supported the potential involvement of this gene in local flowering time variation.

Declaration

I hereby declare that I am the author of this thesis. The work presented is my own original work and I have not previously submitted this work in its entirety, or in part, for any other degree or professional qualification.

H.L McCulloch

Date:

Acknowledgements

Firstly I would like to thank my supervisor Professor Andrew Hudson, without his support and patience this thesis would not have been possible.

I would also like to thank past and present members of the Hudson group for all of their support and assistance; particularly Poay Lim and Abigail Harter who have contributed to work done in this thesis. I would also like to especially thank Jo Critchley who provided me with a steady stream of biscuits to aid my writing.

Thank you also to members of the Big Lab who have taught me techniques and been available for advice and discussion about my work, in particular members of the Goodrich group who have on several occasions been an invaluable help to me. I am also grateful for the emotional support and friendship I have had in the Big Lab and would particularly like to thank Anne and Sophie for keeping me calm when needed.

This work was assisted by Honours students whom I supervised during the course of my PhD. I am grateful for their input specifically the work done by Emma Lindsay who made a valuable contribution.

Thank you to Sibum Sung who supplied *vin3* mutant seeds for this study. I am also grateful to the BBSRC who funded this project.

Finally I would like to thank my parents and grandparents who have always encouraged me in my education. I am grateful for your support and guidance

1 Introduction ... 1

1. 1 Flowering ... 1
1.2 *Arabidopsis thaliana* ... 2
1.3. Mechanisms controlling flowering in *Arabidopsis thaliana* ... 4
1.3.1 Integration of flowering pathways ... 4
1.4 *FLC* independent pathways ... 7
1.4.1 The GA pathway ... 7
1.4.2 The photoperiod pathway ... 8
1.5. *FLC* dependent pathways ... 12
1.5.1 *FLC* ... 12
1.5.2 Up-regulation of *FLC* by *FRI* ... 13
1.5.3 The Autonomous pathway ... 14
1.5.4 The vernalization pathway ... 16
1.6. The control of flowering in other species ... 20
1.6.1 Brassicacaea ... 20
1.7. Natural Variation ... 23
1.7.1 Methods of studying the genetic basis of natural variation ... 23
1.8. Natural Variation in flowering time ... 24
1.8.1 Variation in flowering behaviour observed worldwide ... 24
1.8.2 Natural Variation at the *FRI* locus ... 24
1.8.3 Natural Variation at the *FLC* locus ... 25
1.8.4 Natural variation in other pathways ... 26
1.8.5 Adaptation of flowering behaviour ... 27
1.8.6 Local variation ... 28
1.9. Populations on which this study is based ... 28
1.10. Aims of the project ... 29

2 Materials and Methods ... 30

2.1 Plant materials and growth conditions ... 30
2.1 1 Collection sites of local Arabidopsis populations ... 30
2.1 2 Sterilising and germinating seeds ... 33
2.1 3 Vernalization treatments ... 34
2.1 4 Flowering time analysis ... 34
2.2 DNA preparation and analysis ... 35
2.2 1 DNA extraction from plants ... 35
2.2 2 PCR ... 36
2.2 3 Agarose gel electrophoresis ... 37
2.2 4 Sequencing ... 38
2.3 RNA preparation and analysis ... 38
2.3 1 RNA extractions ... 38
2.3 2 cDNA Synthesis ... 38
2.3 3 qRT-PCR ... 39
2.4 Complementation of *VIN3* alleles ... 40
2.4 1 Preparations of constructs by extension overlap PCR ... 40
2.4 2 Cloning into plasmid vectors and transformation of *E. coli* ... 43
2.4 3 Screening clones by PCR ... 44
2.4 4 *E. coli* cultures in liquid media ... 44

2.4 5 Plasmid mini-preps 44
2.5 Transformation of genes into *Arabidopsis* 44
2.5.1 Transformation into Gateway Ti-Plasmid 44
2.5.2 Transformation of *Agrobacterium tumefaciens* (*Rhizobium radiobacter*) 47
2.5.3 Transformation of *Arabidopsis* 48
3 Flowering time variation in local summer and winter annual populations 50
3.1 Introduction 50
3.2 Response to photoperiod from local winter annuals 51
3.3 Response to photoperiod from local summer annuals 55
3.4 Response to vernalization in local winter annuals 58
3.5 Response to vernalization in local summer annuals 62
3.6 Relationship between distance of collection site and difference in flowering time 62
3.7 Discussion 70
4 Sequence and expression variation in *FRIGIDA* and *FLC* 74
4.1 Introduction 74
4.2 Sequence variation in *FRIGIDA* 75
4.3 Sequence variation in *FLOWERING LOCUS C* 81
4.4 Flowering time in response to vernalization 85
4.5 Variation in expression of *FRIGIDA* and *FLOWERING LOCUS C* 88
4.6 Discussion 92
5 Production and use of a hybrid population (1E1 X10B2) to identify candidate loci involved in flowering time variation locally 95
5.1 Introduction: 95
5.2 Production of mapping population 96
5.3 Late flowering in F1 and F2 hybrids 96
5.4 Mapping candidate loci 101
5.6 Discussion 104
Chapter 6 Production and use of a hybrid population (11C1 X4D1) to identify candidate loci involved in flowering time variation locally 106
6.1 Introduction: 106
6.2 Production of a mapping population 106
6.3 Flowering time variation in F4 generation 108
6.4 Mapping candidate loci 110
6.5 *VIN3* sequence variation 115
6.6 Variation in expression level of *VIN3* and its target *FLC* 120
6.7 Discussion 124
7 Concluding Remarks 127
7.1 Future Work 130
8 References Cited 132

List of figures

Figure 1.1 Integration of major flowering pathways **5**
Figure 1.2 The photoperiod pathway **10**
Figure 1.3 Mechanisms by which *FLC* is regulated **18**
Figure 1.4 Comparison of flowering pathways between Arabidopsis and cereals. **21**

Figure 2.1 Explanation of overlapping extension PCR **42**
Figure 2.2 pENTR^{TM}/D-TOPO vector used for cloning *VIN3* constructs into *E.coli* **45**
Figure 2.3 Gateway Binary Vector **46**

Figure 3.1 Figure 3.1 Flowering times if winter annual (WA) families under short day (SD) and long day (LD) conditions **52**
Figure 3.2 Percentage effect of photoperiod on flowering time **54**
Figure 3.3 Figure 3.3 Flowering times of summer annual (SA) families under short day (SD) and long day (LD) conditions **56**
Figure 3.4 Percentage effect of photoperiod on flowering time in summer annual (SA) families **57**
Figure 3.5 Flowering times if winter annual (WA) families under vernalized and unvernalized conditions **60**
Figure 3.6 Percentage effect of vernalization on flowering time in winter annual (WA) families **61**
Figure 3.7 A tree showing genetic relatedness of winter annual (WA) populations **63**
Figure 3.8 Flowering times of summer annual (SA) families under vernalized and unvernalized conditions **64**
Figure 3.9 Percentage effect of vernalization on flowering time in summer annual (SA) families **65**
Figure 3.10 Relationship between flowering time and geographical position of winter annual collection site **68**
Figure 3.11 Relationship between flowering time and geographical position of summer annual collection site **69**

Figure 4.1 Sequence alignment of *FRI* alleles identified in local populations **77**
Figure 4.2 Sequence alignment and trace for *FRI* in local accessions and Col-0. **78**
Figure 4.3 Alignment of FRI amino acid sequences identified in local populations **79**
Figure 4.4 Sequence alignment of *FLC* alleles identified in local populations **83**
Figure 4.5 Relationship between flowering time and length of vernalization period. **87**
Figure 4.6 Relative expression of *FRI* **89**
Figure 4.7 Relative expression of *FLC* **90**

Figure 5.1 Parents and F1 plants **97**

Figure 5.2 Transverse sections of stems stained with toluidine blue to identify lignified cells **98**
Figure 5.3 Approximate positions of flowering time genes and SSLP markers **102**
Figure 5.4 Relationship between genotype and flowering time at three markers **103**

Figure 6.1 Creation of F4 mapping population **107**
Figure 6.1 Flowering times of parents and F4 generation of a hybrid mapping population **109**
Figure 6.2 Approximate positions of flowering time genes and SSLP markers **111**
Figure 6.3 relationship between genotype and flowering time under vernalized and unvernalized conditions and with respect to overall vernalization response (NGA249) **112**
Figure 6.4 relationship between genotype and flowering time under vernalized and unvernalized conditions and with respect to overall vernalization response (JV30/31) **113**
Figure 6.5 relationship between genotype and flowering time under vernalized and unvernalized conditions and with respect to overall vernalization response (JV75/76) **114**
Figure 6.6 Alignment of VIN3 amino acid sequences of 4D1 and 11C1 **116**
Figure 6.7 Sequence alignment of *VIN3* alleles from families 4D1 and 11C1 **119**
Figure 6.8 The relative expression levels of *FLC* in the parental lines 4D1 and 11C1 and the F1 generation of a hybrid mapping population **122**
Figure 6.9 The relative expression levels of *VIN3* in the parental lines 4D1 and 11C1 and the F1 generation of a hybrid mapping population **123**

List of tables

Table 2.1 Collection sites for local winter annual populations used in this study **31**
Table2.2 Collection sites for local summer annual populations used in this study **32**
Table 4.1 The local and worldwide distribution *FRI* alleles identified in this study locally **80**
Table 4.2 The local and worldwide distribution of *FLC* alleles identified in this study **84**
Table 5.1 Summary of F1 flowering times **98**

Abbreviations

bp	base pairs
BSA	Bovine Serum Albumen
cDNA	Complementary DNA
$CHCl_3$	Chloroform
CTAB	hexadacyltrimethylammonium bromide
DNA	deoyribonucleic acid
dNTPs	deoxynucleotide triphosphate
EDTA	ethylenediaminetetraacetic acid (disodium salt)
EtOH	Ethanol
FDH	Formate Dehydrogenase
g	local grativational field
Hyg	Hygromycin
Kan	Kanamycin
kb	kilobase
LB	Lysogeny broth
LD	Long Day
M	Molar
mM	millimolar
min	minutes
nM	nanomolar
MS	Murashige and Skoog
NaOAc	Sodium Acetate
PCR	Polymerase Chain Reaction
qRT-PCR	Quantitative Reverse Transcription - Polymerase Chain Reaction
RLF	Rosette Leaves to Flower
RNA	ribonucleic acid
SD	Short Day
sec	seconds
SOC	Super optimal broth
SSLP	Simple sequence length polymorphism
TE	Tris-EDTA buffer
TLF	Total Leaves to Flower
Tm	melting temperature
Tris	Tris(hydroxymethyl)aminomethane
Tween	polyethylenesorbitan monolaurate
UTR	untranslated region
UV	ultraviolet
YEP	Yeast extract peptone
β-MCE	β-Mercaptoethanol

1 Introduction

1. 1 Flowering

It is crucial for plants to regulate the timing of developmental changes in order to maximise chances of survival and successful production of offspring. One important development stage is the change from vegetative to reproductive growth, i.e., flowering. This is important to ensure that flowering occurs at a time when conditions are optimal. Ecologically flowering time has major implications as if it occurs too early there is the potential for frost to damage the developing floral buds while too late may result in damage to developing seeds. It also allows synchronisation with seasonally active pollinators.

Research into the regulation of flowering has been carried out for over 100 years (Srikanth and Schmid, 2011): early work showed that the transition to flowering was controlled by a combination of environmental factors. The importance of day-length on flowering time was recorded as early as 1920 when Wightman Garner and Henry Allard observed that shortening days induced flowering in a population of Maryland Mammoth tobacco plants (Salisbury, 1985). From this observation they went on to explore the effect that day-length had on other plant species and coined the term photoperiod, categorising the plants that they studied into three groups according to whether their flowering was promoted by long days (LD), short days (SD) and or was insensitive to photoperiod (day neutral).

In the 1930's Erwin Bünning hypothesised that this response to photoperiod was due to an internal clock that was divided into two phases: light sensitive (photophile) and dark sensitive (scotophile) (Srikanth and Schmid, 2011). This led to the development of the internal and external coincidence models by Pittendrigh (Pittendrigh, 1960; Pittendrigh, 1972). The external coincidence model proposed that oscillations of the clock required the input of light signals at specific times during the day whereas the internal coincidence model proposed that the clock had two separate rhythms and, as the seasons progressed and photoperiod changed, these rhythms would shift and overlap resulting in the activation of downstream components.

In 1934 it was reported that exposure of leaves to an inductive photoperiod was necessary for the floral transition and not exposure of the shoot apex where floral organs initiate (Knott, 1934). It was later shown that leaves from plants grown under inductive photoperiods were capable of inducing flowering when grafted to plants that had not been induced (King and Zeevaart, 1973; Lang et al., 1977), pointing towards a key component of flowering regulation being a mobile signal. This mobile signal was named florigen. Further details about the identity of florigen or the mechanism by which it regulated flowering would remain elusive until tools in molecular biology and genetics were developed.

Another environmental factor that is important in floral induction is temperature. The role of temperature in plant development has been recognised since the 18th Century. However it was not until 1918 that an extended period of cold was proposed to be necessary for flowering in some species (Gassner, 1918). The term put forward by the Russian botanist Lysenko for this process was “jarovization”, which was then translated into English as “vernalization”. Further studies showed the importance of a continuous period of cold in vernalization, for example that a short warm disruption of the cold period reduced the subsequent induction of flowering in rye plants (Purvis and Gregory, 2011).

In 1957 gibberellic acid (GA) was found to induce the onset of flowering when applied exogenously onto the vegetative rosettes of *Samolus parviflorus* and *Crepis tectorum* (Lang, 1957).

Although many factors had been identified as important for floral regulation, the mechanisms by which they controlled the onset of flowering remained unknown for many years.

1.2 *Arabidopsis thaliana*

The underlying mechanisms regulating flowering time are now best understood in the model plant *Arabidopsis thaliana*. *A. thaliana* was first proposed for plant

genetics in 1943 by Laibach (Laibach, 1943). It has many advantages for studying ecological and molecular genetic aspects of plant development. It has a relatively small genome (1.56 x 10^6 bp) and in 2000 the full genome sequence was published (The Arabidopsis Genome Initiative, 2000). *A. thaliana* is also a relatively small plant so large populations can be grown in a relatively small space and commonly used accessions have been identified that will develop from germination to seed set within ~6 weeks. It is self fertile, but can be easily cross-pollinated.

The worldwide distribution of *Arabidopsis thaliana* also makes it a useful organism for studies of ecology, adaptation and evolution. The genus *Arabidopsis* is comprised of nine species and eight subspecies; of these only *Arabidopsis thaliana* (Arabidopsis from now on) is found worldwide. It is naturally distributed throughout Eurasia and has been introduced elsewhere. Arabidopsis is found in disturbed ground in a diverse range of habitats from sea level to high altitudes, suggesting that it has adapted efficiently to a wide range of environments. It is self-compatible and predominantly self-pollinating in the wild (Platt et al., 2010), and this has probably contributed to its spread. The self compatibility of Arabidopsis is also useful experimentally as it makes the generation of recombinant inbred lines (RILs) and other mapping populations relatively simple. There are also increasingly large collections of natural and mutant lines, distributed by facilities such as ABRC and NASC, and phenotypic and genetic information is easily accessible from The Arabidopsis Information Resource (TAIR), including genome sequences of several natural accessions.

Arabidopsis is a monocarpic annual in that it dies after flowering. Its flowering behaviour is classically divided into two categories; winter annual and summer annual. Winter annual plants germinate in the late summer or autumn and overwinter as vegetative rosettes; flowering is then initiated the following spring. Plants that have a summer annual habit have the capacity to complete their entire life cycle in spring or summer without over-wintering.

1.3. Mechanisms controlling flowering in *Arabidopsis thaliana*

Studies have shown that flowering in Arabidopsis is regulated by environmental signals, such as light, temperature and stress, and pathways independent of external stimuli, such as age. Extensive mutagenesis studies identified a large numbers of genes involved in flowering and allowed their subsequent integration into four major regulatory pathways: the vernalization, autonomous, giberellic acid (GA) and photoperiod pathways. More recently the molecular mechanisms by which other factors such as ambient temperature and plant age influence flowering time have also been characterised.

1.3.1 Integration of flowering pathways

Flowering is induced by both endogenous and environmental cues; as such it is controlled by a complex network of signals. These converge on the floral integrator genes *FLOWERING LOCUS T* (*FT*), *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* (*SOC1*), *AGAMOUS-LIKE 24* (*AGL24*) and *LEAFY* (*LFY*), which promote flowering and are regulated antagonistically by different pathways (Fig.1.1).

FT is repressed by the vernalization pathway and activated by the photoperiod pathway. Over-expressing *FT* gave an early flowering phenotype whereas plants carrying mutant *ft* alleles were extremely late flowering indicating that *FT* is sufficient and necessary for floral induction (Mathieu et al., 2007). *FT* was cloned independently by two different groups and found to encode a protein similar to Raf kinase inhibitory protein (RKIP) and phosphatidylethanolamine-binding protein (PEBP), both of which regulate MAP kinase signalling. However, FT lacks key conserved residues from RKIP and PEBP and therefore it is not clear whether it acts in a similar way (Kardailsky et al., 1999; Weigel et al., 2000).

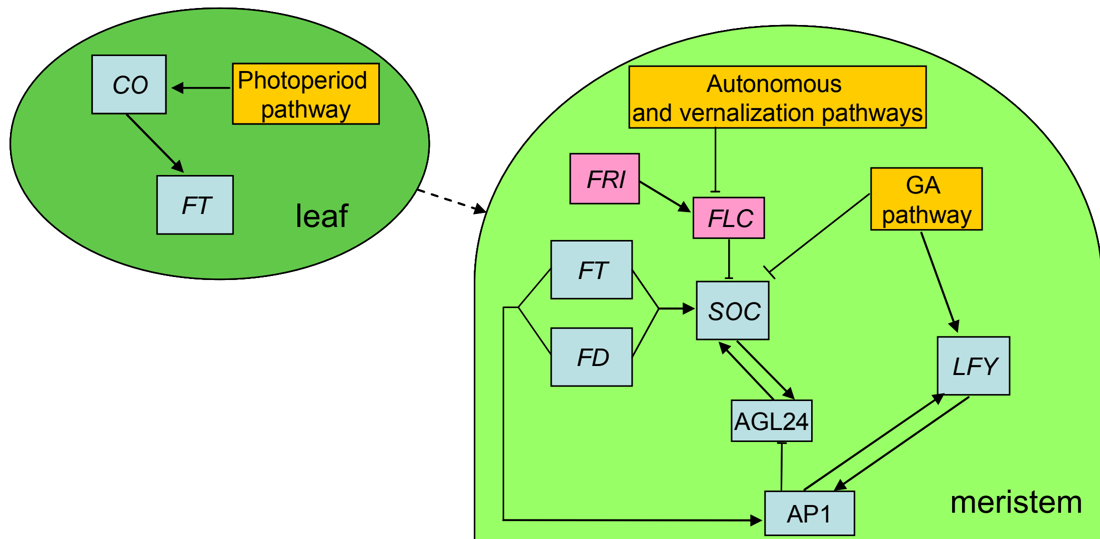


Figure 1.1 Integration of major flowering pathways Floral integrator genes are shown in blue, genes from other pathways are indicated in pink, pathways upstream of the floral integrator genes are represented in orange and those downstream are shown in yellow. Movement of FT is indicated using a dashed arrow.

FT is, however, capable of binding to the basic leucine zipper (bZIP) transcription factor FD, which is expressed in the shoot apex and is capable of activating the transcription of floral meristem identity genes such as *APETELA1* (*AP1*) in the presence of FT (Abe et al., 2005; Wigge et al., 2005), although binding of FT and FD has yet to be demonstrated in the shoot apical meristem. A similar role to FT has been proposed for its paralogue, TWIN SISTER OF FT (TSF) (Kobayashi and Weigel, 2007).

An intriguing aspect of *FT* function is that it is expressed in leaves but promotes initiation of flowering in the shoot apex (Wigge, 2011). Studies using GFP:FT fusion proteins suggested that FT travelled through the vasculature from its site of expression in leaf veins to the shoot apex (Mouradov et al., 2002), pointing to FT being the mobile flowering signal, florigen.

SOC1 also known as *AGAMOUS LIKE 20* (*AGL20*) is a MADS box transcription factor that is up-regulated by inductive photoperiod and the GA pathway and repressed by the vernalization pathway (Mouradov et al., 2002). It is expressed throughout the plant, including the shoot apex. *SOC1* expression is also promoted by members of the SQUAMOSA BINDING PROTEIN-LIKE (SPL) transcription factor family (Wang et al., 2009a; Wu et al., 2009). Since *SPL* expression is inhibited by the microRNA miR156, which declines as plants develop, this provides a link between flowering and plant age.

SOC1 functions in a heterodimer with another MADS box transcription factor, AGL24, to promote expression of *LFY* via direct binding of its promoter (Lee and Lee, 2010). Like *SOC1*, *AGL24* over-expression causes early flowering and *agl24* mutants flower late. *SOC1* and *AGL24* also promote each other's expression directly, providing a mechanism that might maintain flower induction. *LFY* was first characterised for its role in the development of the floral meristem; *lfy* mutants produce inflorescence axes, with cauline leaves and additional inflorescences in place of flowers (Weigel et al., 1992). In addition to being a direct target of SOC1-

AGL24, *LFY* might also be a direct target of the GA pathway, possibly responding to GA-activated MYB transcription factors (Eriksson et al., 2006) and, like *SOC1*, is a target in the age dependent control of flowering via miR156 regulation of *SPL*s (Yamaguchi et al., 2005).

Although both *SOC1* and *AGL24* respond to vernalization, the mechanisms appear to be different in each case. *SOC1* is repressed by a complex containing FLOWERING LOCUS C (FLC) and another MADS box transcription factor, SHORT VEGETATIVE PHASE (SVP). In contrast, *AGL24* expression is promoted by vernalization independently of *FLC* (Michaels et al., 2003a). Repression of *SOC1* by the vernalization pathway still occurs in the presence of functional CONSTANS (CO) protein, a component of the photoperiod pathway. It has been proposed that this ensures flowering is initiated in response to a balance of environmental stimuli i.e. vernalization is still needed to advance flowering in inductive long days (Lee et al., 2007; Li et al., 2008; Searle et al., 2006). The up-regulation of *SOC1* by CO occurs via FT.

1.4 *FLC* independent pathways

1.4.1 The GA pathway

GA had been known to have roles in plant development and the induction of flowering since the 1950s; however the mechanisms by which it acted were not discovered until much later. In Arabidopsis, it was originally thought that GA was only involved in flowering under SD conditions, as the GA-deficient mutant *ga1-3* appeared to flower normally under LD conditions but failed to flower under SD conditions (Wilson et al., 1992). However a LD phenotype has been observed in studies of the GA receptor gene, *GIBERELLIC INSENSITIVE DWARF 1* (*GID1*). *GID1* was initially identified in rice and Arabidopsis was subsequently shown to have three functionally redundant copies. The triple mutant in Arabidopsis was late flowering even under LD conditions (Griffiths et al., 2006; Willige et al., 2007), supporting a role for GA in promoting flowering in both SD and LD conditions. In addition, several findings suggested that GA promotes *FT* expression in parallel with

photoperiod. Firstly double mutants of *constans* (*co*), which abolished the response to long days (Section 1.4.2) and *ga1-3* flowered later than either single mutant, suggesting that photoperiod and GA have independent effects on flowering (Putterill et al., 1995). Secondly, treating wild type plants with the GA inhibitor paclobutrazol delayed flowering under LD conditions and application of exogenous GA was sufficient to restore the wild type flowering phenotype (Hisamatsu and King, 2008). Thirdly, the application of exogenous GA to the GA-deficient *ga1-3* mutant could substantially increase *FT* expression when these plants were moved from SD to LD conditions.

GA signalling also acts through *SOC1*. Under SD conditions *SOC1* expression is virtually undetectable in a *ga1-3* mutant background, but increases on exogenous application of GA (Moon et al., 2003). The exact mechanism by which *SOC1* expression is promoted by the GA pathway is unclear because the *SOC1* promoter does not contain recognised GA-responsive elements. It seems likely that at least part of the GA regulation of *SOC1* involves the floral repressor SHORT VEGETATIVE PHASE (SVP), because *SOC1* expression is increased in *svp* mutants and application of GA to wild type plants results in a decrease of *SVP* expression (Mouradov et al., 2002). The regulation of *SOC1* by GA is potentially important for flowering, because flowering in response to GA is reduced in *soc1* mutants and *SOC1* over-expression can rescue the late flowering phenotype of GA pathway mutants (Srikanth and Schmid, 2011).

GA also regulates flowering via the up-regulation of *AGL24*. This is dependent on SOC1 activity: application of GA resulted in a marked increase in *AGL24* expression in wild-type plants but not in a *soc-1* mutant background (Yu et al., 2002). Given that *SOC1* and *AGL24* positively regulate each other, GA seems likely to regulate *AGL24* via *SOC1*.

1.4.2 The photoperiod pathway

A key component of the photoperiod pathway involves regulation of *CONSTANS* (*CO*) expression by the circadian clock. The circadian clock comprises inputs that

synchronise it with day-length, oscillators to maintain circadian rhythms, and output pathways which regulate genes such as *CO*. The input pathways involve the red and far red light receptor phytochromes (PHYA to PHYE) and the blue light receptor chryptochromes (CRY1 and CRY2) (Thomas, 2006).

The oscillators of the clock comprise of three interlocked transcriptional feedback loops that converge on the Myb-related transcription factor genes *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*) and *LATE ELONGATED HYPOCOTYL* (*LHY*). In the first loop TIMING OF CAB EXPRESSION1 (TOC1), also called PSEUDO-RESPONSE REGULATOR 1 (PRR1), indirectly promotes expression of *LHY* and *CCA1* and the products of these genes then bind directly to *TOC1* and inhibit its expression. As levels of the CCA1 and LHY inhibitors subsequently decrease, *TOC1* expression can increase, producing a rhythmic variation in activities of the components of the loop. *CCA1* and *LHY* expression is promoted by light and is highest in the early morning. In the evening loop *TOC1* expression is promoted by an unknown component, which might involve the GIGANTEA (GI) protein (Harmer, 2009). In the third loop CCA1 and LHY promote expression of PRR7 and PRR9 which in turn inhibit CCA1 and LHY expression.

Posttranscriptional regulation is necessary for the correct function of the clock. One of the first clock components to be characterised was ZEITLUPE (ZTL), an F-box protein component of a Skp/Cullin/F-box (SCF) E3 ubiquitin ligase complex which targets TOC1 and PRR5 for degradation by the 26S proteasome (Fujiwara et al., 2008). ZTL been demonstrated more recently to have the capacity to sense blue light as it contains a LOV domain (a flavin mononucleotide binding region), thus suggesting a role for ZTL in synchronising the clock with light cues. ZTL has been suggested to regulate GI postranscriptionally because ZTL physically interacts with GI in a blue light-dependent manner, resulting in stabilisation of both proteins. This might prevent ZTL from degrading other proteins such as TOC1 and PRR5, resulting in a more rapid turnover of all four proteins in the dark than in the light (Harmer, 2009).

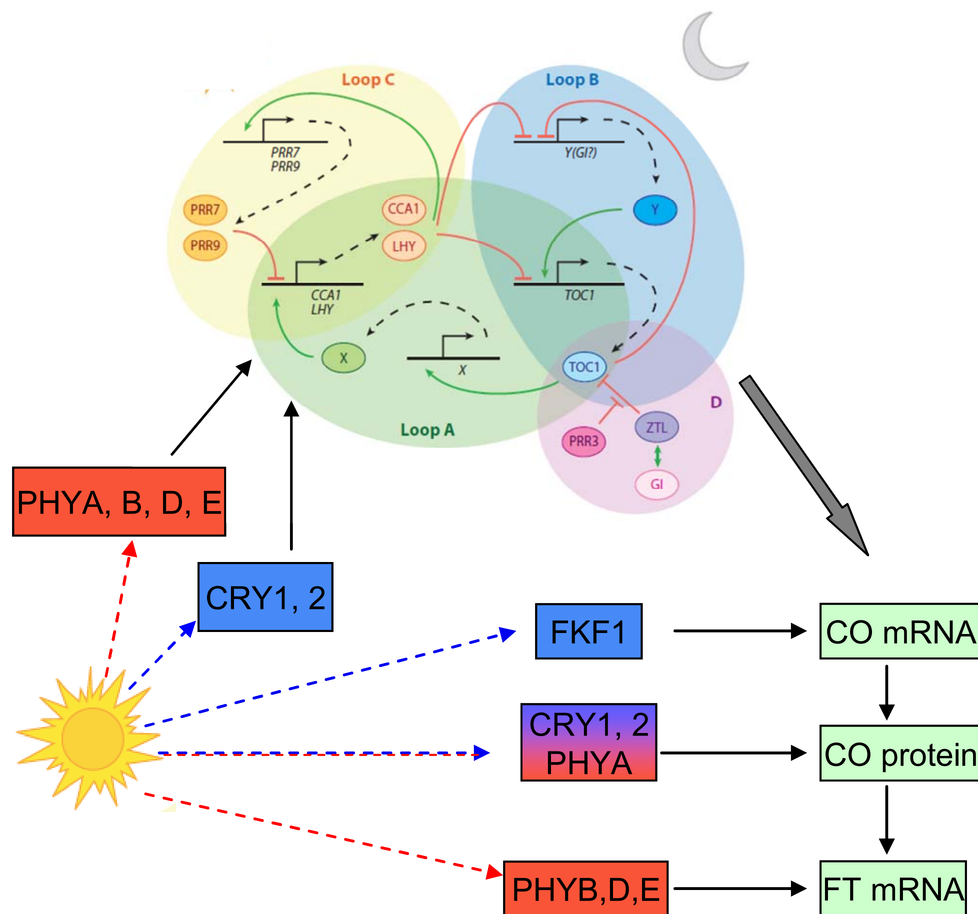


Figure 1.2 The photoperiod pathway. The mechanism by which light signals induce flowering via the integration of photoreceptors, the circadian clock and the external coincidence model. Figure modified from (Harmer, 2009)

GI also interacts with FLAVIN-BINDING, KELCH REPEAT, F-BOX 1 (FKF1) in a blue light-dependent manner. FKF1 appears to play a similar role to ZTL in the clock (Baudry et al., 2010).

Several models were proposed to explain how the circadian clock is involved in photoperiodic control of flowering, of which the external coincidence model is the most consistent with molecular genetic data. This proposes that photoperiodic responses will only be triggered when expression levels of clock-regulated components coincide with an external signal, in this case light.

CO is a key component of the photoperiod pathway and of the external coincidence model in Arabidopsis. *co* mutants are unable to respond to inductive LD and *CO* is expressed in leaves and activates expression of FT (Suarez-Lopez et al., 2001). Expression of the *CO* gene is repressed during the morning by a family of DOF transcription factors. These are targeted for degradation by the GI-FKF ubiquitin ligase, which is activated by blue light during the day (Harmer, 2009). As DOF activity decreases, *CO* transcript levels increases. The difference between SD and LD is therefore that high *CO* expression coincides with light towards the end of a LD, but occurs in darkness in SD conditions (Harmer, 2009).

Blue light promotes the stability of CO protein and therefore more CO protein accumulates in LD conditions than in SD, as a result of the co-incidence between the clock-regulation of *CO* transcription and the external regulation of CO stability by light (Valverde et al., 2004).

Blue light promotes CO stability via the CRY2 photoreceptor, while CO stability is reduced by red light perceived by PHYB. Degradation of CO appears to occur via the 26S proteasome and is dependent on the E3 ubiquitin ligase CONSTITUTIVELY PHOTOMORPHOGENIC (COP1). (Liu et al., 2008) demonstrated that COP1 contained a WD repeat domain which was able to interact directly with CO, that COP1 ubiquitinated CO, and that a CO:luciferase protein fusion was expressed at a higher level in a *cop1* mutant background compared to wild-type.

CO degradation also involves the SPA family of proteins, which also contain a WD repeat domain and are part of the COP1 complex (Laubinger et al., 2006). It is evident that SPA regulation of CO is posttranslational as *CO* mRNA remains unchanged in a *spa* mutant background whereas CO protein levels are significantly elevated (Ishikawa et al., 2006). The capacity for SPA proteins to interact with CO was demonstrated using co-immunoprecipitation studies that also showed interaction occurred via the CCT (CONSTANS, CO-LIKE, TOC1) domain towards the C-terminus of CO. It was also shown that SPA proteins interacted with the coiled-coil domain of COP1 and it was therefore likely that the degradation of CO by SPA was mediated by its interaction with COP1 (Laubinger et al., 2006).

Increased CO levels result in the up-regulation of *FT*. Both *FT* and *CO* are expressed in phloem cells and once plants have been transferred from SD to LD conditions an elevation in *FT* expression is dependent upon the presence of CO. However CO does not contain a typical DNA binding domain, so the mechanism by which this happens is unclear. One hypothesis is that CO is recruited to the *FT* promoter by the Arabidopsis HEAM ACTIVATOR COMPLEX 2 (HAP2) protein, which can dimerize with CO and is homologous to a DNA-binding protein that in yeast forms a complex with a CCT domain protein analogous to CO.

1.5. *FLC* dependent pathways

1.5.1 *FLC*

The vernalization and autonomous pathways regulate flowering at least in part through their action on the well characterised floral repressor *FLC*. *FLC* encodes a MADS box transcription factor which acts as a floral repressor via its action on the floral integration genes.

The structure of the *FLC* gene is unusual as it has an exceptionally long intron. Less than 5% of all plant introns are longer than 850 bp whereas the first intron of *FLC* is 3.5 kb - approximately half the length of the entire gene (Hanada et al., 2007). Long

introns in human genes were found to have more conserved sequences than average indicating that they may function in regulation (Sironi et al., 2005). This has been found to be the case for *FLC*, where regions identified in intron 1 are crucial for *FLC* regulation, as will be discussed in more detail in Chapter 4.

FLC belongs to a clade of MADS domain genes that includes MADS AFFECTING FLOWERING (*MAF1*; also known as *FLM*) and *MAF2* to *MAF5*. All other members of the family that have been studied have been found to have some involvement in flowering (Ratcliffe et al., 2001; Ratcliffe et al., 2003).

1.5.2 Up-regulation of *FLC* by *FRI*

Expression of *FLC* is promoted by the *FRIGIDA* (*FRI*) gene. The FRI protein contains 609 amino acids with two coiled-coil domains that appear important for function. FRI acts on *FLC* as part of a complex that consists of FRI, FRL1 and FRL2 (FRIGIDA LIKE1 and 2) and the C2H2 zinc-finger protein SUPPRESSOR OF FRIGIDA4 (SUF4). Mutations in the genes encoding any of these proteins results in reduced *FLC* expression. The complex activates transcription of FLC via direct binding to its 5' promoter region (Sheldon et al., 2002).

Other components are required for FRI-mediated up-regulation of *FLC*. Two are similar to the SWR1 and Paf1C transcriptional activator complexes identified in yeast. SWR1 in Arabidopsis is made up of PHOTOPERIOD-INDEPENDENT EARLY FLOWERING1 (PIE1), ACTIN-RELATED PROTEIN6 (ARP6) and SERRRATED LEAVES AND EARLY FLOWERING (SEF) and Paf1C consists of VIP4 and VIP5 (VERNALIZATION INDEPENDENCE4 and 5) and ELF7 and ELF8 (EARLY FLOWERING7 and 8). These complexes differ from the FRI complex in that they have a more general role in regulation of gene expression outside flowering (Choi et al., 2007; He et al., 2004; March-Diaz et al., 2007; Martin-Trillo et al., 2006; Noh and Amasino, 2003; Oh et al., 2004; Zhang and van, 2002).

FLX (FLC EXPRESSOR), a leucine zipper containing protein, has also been identified as an important component in the regulation of *FLC* by FRI. Plants carrying an *flx* mutation have strongly reduced levels of *FLC*, but other known regulators of *FLC* are unaffected (Andersson et al., 2008). This effect of *flx* mutations was suppressed by mutations in the autonomous pathway, suggesting that FLX is involved in the regulation of *FLC* by *FRI*. However, unlike FRI, FLX was also needed to promote expression of the *FLC*-like *MAF1* and *MAF2* genes (Andersson et al., 2008).

1.5.3 The Autonomous pathway

The autonomous pathway is so called because it acts to control flowering independently of environmental cues such as light and vernalization. It affects flowering via RNA processing and epigenetic regulation of the floral repressor *FLC* (Fig 1.3A). It consists of several known components, which despite being termed a pathway regulate *FLC* independently, not in a sequential series of events. The components of the autonomous pathway were identified via mutagenesis screen and includes: FCA, FY, FPA, FVE, FLD and FLK (FLOWERING LOCUS CA, Y, PA, VE, LD and LK), LUMINIDEPENDENS (LD), REF6 and PCFS4..

FVE regulate *FLC* via histone deacetylation, independently of all other components in the pathway (mutations in other components do not affect histone deacetylation at *FLC*) (Ausin et al., 2004; He et al., 2003).

FCA and FY are RNA-binding proteins that interact to determine polyadenylation site selection in antisense transcripts from *FLC* (Section 1.5.4), so that *fy* and *fca* mutants produce more transcripts that are antisense to the *FLC* promoter. This might account for increased *FLC* expression in the mutants (Hornyik et al., 2010; Quesada et al., 2003; Simpson et al., 2003). FY and FCA activity is not, however, limited to *FLC*. It has also been demonstrated through chromatin immunoprecipitation (ChIP) and genetic analysis that FLD, a histone demethylase, is required for FCA and FY to function as regulators of FLC(Liu et al., 2007). Additionally another autonomous pathway component has been shown to interact with FCA and FY and regulate the

alterative polyadenylation of FCA, this component is PCFS4, a homologue of the yeast polyadenylation factor Pcf11p (Xing et al., 2008).

FPA and FLK are also RNA binding proteins that work independently from each other and from FCA and FY (Lim et al., 2004; Schomburg et al., 2001). FPA has a similar effect on polyadenylation site selection to FCA and FY, though it functions independently of them (Hornyik et al., 2010).

LD has been shown to regulate *FLC* independently from other components of the autonomous pathway although the exact mechanism by which it does this is still unclear. It is known to encode a homeodomain protein. Although more typically associated with binding to DNA there is evidence that homeodomains can bind to RNA and this would be consistent with regulation of *FLC* by different non-coding RNAs (Dubnau and Struhl, 1996; Lee et al., 1994; Rivera-Pomar et al., 1996). LD has also been shown to bind to SUF4, which prevents the upregulation of *FLC* via the FRI complex, suggesting that it might act as a transcriptional regulator of *FLC* (Kim et al., 2006).

Components of the autonomous pathway are widely conserved in flowering plants whereas *FLC* is not; as yet orthologs of *FLC* have been identified only in the Brassicaceae. This raises the possibility that the autonomous pathway has a more general role in regulation of gene expression than just flowering. There is some evidence to support this, for example *FPA* regulates the biosynthesis of GA and *FVE* has a function in the cold acclimation pathway (Simpson, 2004) and *FPA* regulates polyadenylation site selection in other genes (Hornyik et al., 2010). These roles are not shared with all components of autonomous pathway providing further evidence that the components act in parallel rather than a linear pathway.

Inhibition by 24 nt short interfering RNAs (siRNA) is also proposed to play a role independent of the vernalization pathway in the silencing of *FLC*. An siRNA complementary to a region downstream of the sense transcript of *FLC* promotes formation of heterochromatin at *FLC*. Mutations in *DICER-LIKE1* and *3*, which are

necessary for the formation of the siRNA, lead to increased levels of *FLC* expression and later flowering in the absence of vernalization (Liu et al., 2004).

1.5.4 The vernalization pathway

Vernalization is the process by which floral initiation is promoted by an extended period of cold. Under natural conditions this would prevent plants flowering until they had over-wintered, in the more favourable conditions of spring. It has been known for many years that a vernalized plant can maintain the memory of this vernalization mitotically under warmer conditions and that an unvernalized state is reset during development of the embryo (Sheldon et al., 2008).

The vernalization pathway works in parallel with the autonomous pathway to downregulate *FLC*. However there is also evidence for vernalization responses that are independent of *FLC*. For example a member of the *MADS AFFECTING FLOWERING* (*MAF*) clade of genes; *MAF2* was shown to be involved in preventing flowering after short periods of cold. *maf2* mutants showed the same response to 10 days of cold as was induced in wild type plants after 85 days of cold (Ratcliffe et al., 2003).

Additionally ectopic expression of the *SOC1* homolog *AGAMOUS LIKE 19* (*AGL19*) has been shown to result in rapid induction of flowering. *AGL19* expression increases in response to cold independently of *VERNALIZATION INSENSITIVE 3* (*VIN3*; see below), as *agl19 vin3* double mutants have a more impaired vernalization response than either of the single mutants. In contrast to *FLC*, the response of *AGL19* to vernalization is transient and decreases once plants are returned to the warm (Schönrock et al., 2006)

Several studies have shown that vernalization regulates flowering time and *FLC* expression in a dose dependent manner; longer cold periods giving lower *FLC* mRNA levels and shorter times to flowering (Michaels and Amasino 1999, Sheldon et al 2000). However there is also evidence to suggest that it is not the absolute amount by which cold initially reduces *FLC* expression that is crucial for the effect on flowering time but the stability of that repression if the plant is returned to warm

conditions (Shindo et al., 2006). In natural Swedish accessions of Arabidopsis neither the initial unvernalized level of *FLC* nor the decrease in its RNA levels in response to cold correlated with flowering times. Instead a correlation was found with the length of vernalization required for stable repression of *FLC* so that expression did not increase again if plants were returned to the warm.

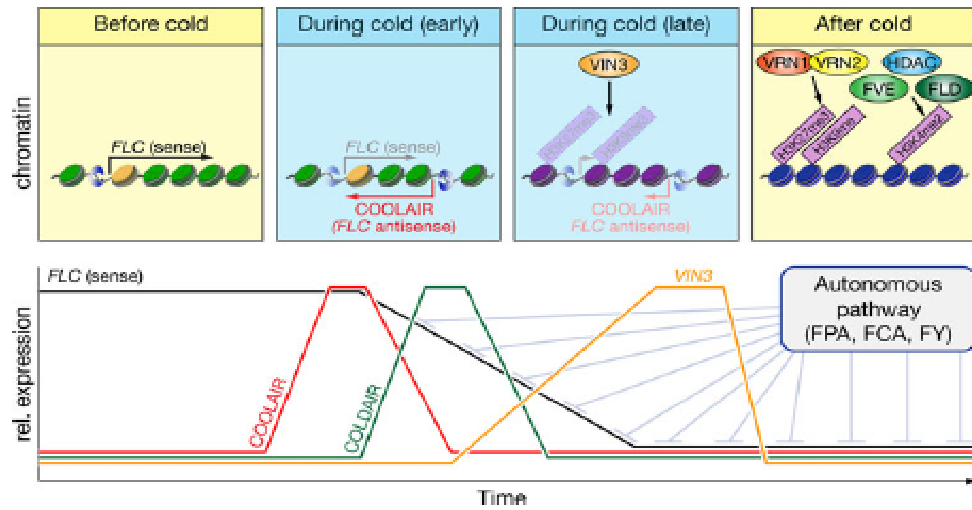
Three genes are known to be involved in vernalization pathway: *VIN3* (Sung and Amasino, 2004) and *VERNALIZATION1* and *2* (*VRN1* and *VRN2*) (Chandler et al., 1996). *VIN3* encodes a protein containing a plant homeodomain (PHD) and fibronectrin type III repeats, usually involved in protein-protein interactions. It is

induced in response to cold and required to establish cold-induced epigenetic silencing of *FLC*. *VRN1* encodes a DNA-binding protein and *VRN2* a component of the Polycomb repressive complex 2 (PRC2), which acts by modifying histones in the chromatin of target genes (Gendall et al., 2001; Levy et al., 2002). Unlike *VIN3*, neither *VRN1* nor *VRN2* are induced by vernalization. Similarly, *vrn1* and *vrn2* mutants still show the initial downregulation of *FLC* in response to cold, unlike *vin3* mutants, though this is not maintained if plants are returned to the warm (Sung and Amasino, 2004). This evidence has suggested a model in which cold-induced VIN3 initially represses *FLC* and targets VRN1 and VRN2 to *FLC* to maintain stable repression. This repression can account for the mitotic stability of vernalization.

Whilst it has been understood for some time that increased recruitment of the PRC2 complex to *FLC* during vernalization leads to an increase in the repressive histone mark, H3K27me3, the exact mechanism by which recruitment occurs is still debated. Recent studies suggest that long non-coding RNA (ncRNA) molecules transcribed from *FLC* itself play a role in repression of *FLC* and PRC2 recruitment even earlier than VIN3 .

Two non-coding RNA molecules have been identified as important for the epigenetic downregulation of *FLC* in response to vernalization, *COOLAIR* (*COLD INDUCED LONG ANTISENSE INTRAGENIC RNA*) and *COLDAIR* (*COLD ASSISTED INTRONIC NONCODING RNA*; Fig. 1.3B). It has been demonstrated that

A



B

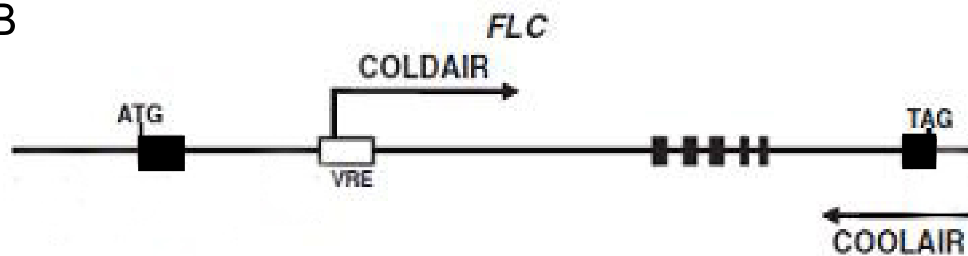


Figure 1.3 Mechanisms by which *FLC* is regulated (A) illustrates three mechanisms by which FLC is down-regulated and the timing and relative expression of the components of each mechanism. (B) shows a map of the FLC gene and the sites of transcription for the noncoding transcripts *COLDAIR* and *COOLAIR*. The figure is modified from (Heo and Sung, 2011; Srikanth and Schmid, 2011)

COLDAIR physically associates with components of the PRC2 complex and might therefore target PRC2 to *FLC* (Heo and Sung, 2011).

These transcripts are suggested to be involved in the regulation of *FLC* by cold as an earlier step than VIN3 as their up-regulation by cold can be detected 10 days earlier than *VIN3* expression and is not dependent on VIN3 activity. Additionally up-regulation of the *COOLAIR* antisense *FLC* RNA coincides with the downregulation of sense transcripts of *FLC*. This suggests that regulation via non-coding RNA molecules is an earlier step in the vernalization process than recruitment via *VIN3*. However another study showed that although the promoter, first exon and first introns of *FLC* are essential for its downregulation the promoter for *COOLAIR* was not required. This may suggest that *COOLAIR* acts redundantly with other components of the epigenetic mechanism but that it is not a requirement for downregulation of *FLC* (Helliwell et al., 2011).

Knock down lines produced using RNAi that had reduced *COLDAIR* expression were late flowering after vernalization, consistent with *COLDAIR* acting as a repressor of FLC activity. *FLC* expression was reduced in *COOLAIR* knock down lines, although repression was not maintained once plants are returned to the warm. This is consistent with the proposed role of *COLDAIR* in recruiting the PRC2 complex necessary for stable silencing of *FLC*.

A similar mechanism had been identified previously in animals, in which the non-coding RNA molecule *HOTAIR* is transcribed from the human *HOXC* homeobox gene cluster and promotes silencing of the *HOXD* cluster on a different chromosome, via PRC-mediated chromatin modification (Tsai et al., 2010). The *HOTAIR* transcript was found to be associated with PRC2 (Rinn et al., 2007). The proposed mechanism in Arabidopsis differs in that it is the target of repression itself, *FLC*, from which the non coding RNA is transcribed.

1.6. The control of flowering in other species

1.6.1 Brassicacaea

Orthologs of *FLC* have yet to be identified outside the Brassicacaea family. Within the Brassicacae, *FLC* orthologs are important for floral regulation in species other than Arabidopsis and can determine life history. For example in *Arabis alpina* the *FLC* ortholog *PERPETUAL FLOWERING 1* (*PEP1*) is crucial for establishing a perennial life history. As in Arabidopsis, *PEP1* is downregulated in response to vernalization and via methylation of H3K27 this repression is maintained. However vernalization only affected shoots that had formed prior to the start of the vernalization treatment and any additional shoots that formed during or after the vernalization process remained vegetative, therefore conferring a perennial behaviour (Wang et al., 2009b).

FRI variation has been identified in *Arabidopsis lyrata* that is responsible for flowering time variation. Two different *FRI* alleles were identified that differed by an indel corresponding to a difference of 14 amino acids. Both *FRI* proteins were believed to be functional but were associated with a difference in flowering time of approximately 15 days (Kuittinen 2008).

1.6.2 Cereals

Many studies have examined flowering regulation in cereals and have shown that environmental factors such as photoperiod and vernalization are involved. In fact the process of vernalization was first identified in grasses. Studies in barley and wheat led to the discovery of the *VRN1*, *VRN2*, *VRN3*, and *VRT2* genes involved in the vernalization response. It should be noted that despite their identical nomenclature *VRN1* and *VRN2* identified in grasses share no homology with *VRN1* and *VRN2* from Arabidopsis.

The proposed model by which grasses control flowering in response to vernalization has only some similarities with Arabidopsis. In its simplest form both models consist of activators that promote flowering in response to inductive photoperiods but are down-regulated by a vernalization-sensitive repressor. The activators of

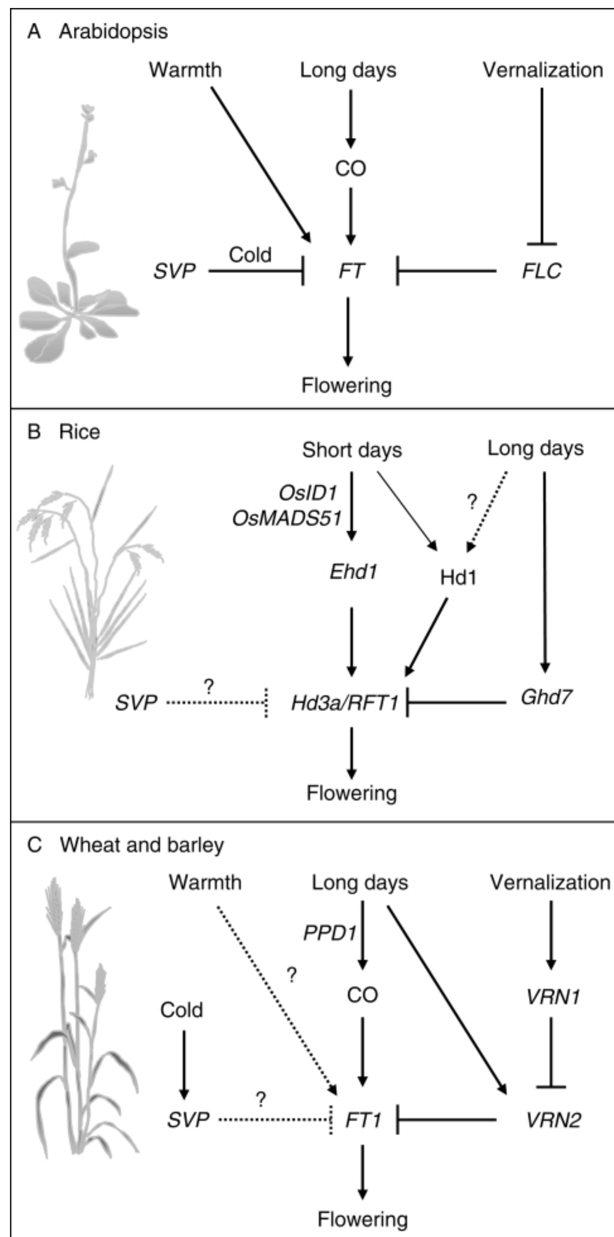


Figure 1.4 Comparison of flowering pathways between Arabidopsis and cereals. Simplified flowering pathways for (A) Arabidopsis, (B) Rice and (C) Wheat and Barley. Figure from (Greenup et al., 2009)

flowering in grasses, VRN3 and VRN1, belong to the same protein families as their Arabidopsis counterparts, FT and SOC1. However vernalization in Arabidopsis acts to repress the MADS box transcription factor gene *FLC*, whereas in grasses one model suggests that it acts on the CCT domain protein, VRN2.

One model for the regulation of flowering in cereals via vernalization is that the protein VRT2 (identified as homologous to the Arabidopsis proteins AGL24 and SVP) binds directly to the promoter of *VRN1* and recruits VRN2; together they repress the expression of *VRN1*. It has been demonstrated that SD conditions and vernalization can repress *VRN2* and therefore allow the expression of *VRN1*.

Cereal *VRN3*, the homologue of the Arabidopsis FT protein, fits into the model as a promoter of flowering in response to LD photoperiods because it up-regulates *VRN1* and is repressed by VRN2.

There is also evidence for an alternative model in which VRT2 does not function in the vernalization-regulated control of flowering time. It has been proposed that much like its Arabidopsis homologs, AGL24 and SVP, it has a role in regulation of floral meristem identity. (Trevaskis et al., 2007).

The exact mechanism by which vernalization is 'remembered' in grasses is still unclear. There is some evidence that it is similar to Arabidopsis, for example *VIN3*-like genes have been identified (*TmVIL1-3*; (Fu et al., 2006) and their transcripts shown to accumulate in response to cold and to subsequently disappear once plants are returned to the warm. Further to this a JUMONJI (JmjC)-like gene, *LpJMJC* has been identified. In Arabidopsis a JmjC protein REF6 has been shown to repress *FLC*, potentially via chromatin remodelling, although a role for JmjC proteins in vernalization has yet to be found (Noh et al., 2004).

1.7. Natural Variation

1.7.1 Methods of studying the genetic basis of natural variation

Quantitative trait locus (QTL) analysis has proved to be a useful tool in the study of natural variation in plants, the concept of QTL was first identified in cereals when it was observed that segregation of multiple loci resulted in a continuous range of phenotypes (Nordborg and Weigel, 2008). The study of natural variation using QTL analysis has identified genes not previously identified as involved in a trait of interest using other techniques such as mutagenesis screening, for example the role of the *DOG1* gene in seed dormancy (Bentsink et al., 2006). QTL analysis also has disadvantages, including the labour involved in genotyping mapping populations at many loci and the subsequent difficulty of fine mapping loci that might have relatively small effects on the phenotype to identify individual candidate loci.

With genotyping techniques, particularly those based on next generation sequencing, becoming more time and cost efficient another technique for studying the genetic basis of natural variation is becoming more commonly used. Genome wide association (GWA) mapping does not require the production of a hybrid population but instead examines polymorphisms across genomes of populations and tests for associations between genotypes and phenotypes. One clear advantage of GWA is in the study of natural variation in animal populations, where controlled crosses are either more difficult or, as in the case of humans, impossible. However a major disadvantage with this technique is that population structure can bias the results. For example GWA can identify *FRIGIDA* as a determinant of flowering time in populations from north-western Europe because most summer annuals from this regions carry frequent *fri* alleles. However in central Asia individual *fri* alleles are rarer and GWA fails to identify *FRI* as an important QTL (Shindo et al., 2005).

As a result several GWA studies have also beeen followed up with classical QTL mapping technique to ensure accurate detection of all QTL involved.

1.8. Natural Variation in flowering time

1.8.1 Variation in flowering behaviour observed worldwide

Extensive variation in flowering time has been observed wordwide, as well as variation in the response of flowering to vernalization and photoperiod. There is some evidence to suggest that this variation might be adaptive because latitudinal clines are apparent with respect to flowering time, vernalization sensitivity and the frequency of alleles of flowering time genes. However other studies reach different conclusions as the demonstration of a cline appears to be dependent upon the experimental design and the accessions chosen for the study. (Caicedo et al., 2004; Lempe et al., 2005; Shindo et al., 2005; Stinchcombe et al., 2004).

1.8.2 Natural Variation at the *FRI* locus

Several studies have shown that variation in *FRI* is a major determinant of worldwide variation in flowering time of Arabidopsis. Different estimates suggest that between 12.6% to 70% of observed flowering time variation (under controlled conditions) could be accounted for by *FRI* (Lempe et al., 2005; Scarcelli et al., 2007; Shindo et al., 2005; Werner et al., 2005). Two common *FRI* sequence variants that result in rapid cycling (i.e., early flowering without vernalization) were identified in the commonly used laboratory accessions Col-0 and L*er*. The Col-0 allele contains a 16 bp deletion at the end of the first exon which results in a premature stop codon. L*er* was found to contain a larger deletion in the promoter which removed the start codon. Both Col-0 and L*er* therefore appear to be rapid cycling because they carry *fri* loss of function mutations. These alleles were shown to be common throughout Europe (Johanson et al., 2000; Le, V et al., 2002; Shindo et al., 2005; Stinchcombe et al., 2004; Toomajian et al., 2006; Werner et al., 2005), and have been suggested to be advantageous in disturbed habitats, such as cultivated fields, leading to the suggestion that they spread with agriculture.

There is some evidence to suggest that variation at *FRI* might be adaptive because of roles in processes other than *FLC*, regulation such as a potential function in water use

efficiency [McKay et al (2003), Stinchcombe et al (2004)]. However, it is not clear whether these other effects are independent of flowering or a consequence of it.

1.8.3 Natural Variation at the *FLC* locus

Natural variation in flowering time and vernalization response has been attributed to differences in the *FLC* sequence itself, with weakly expressed or non-functional alleles correlating with rapid cycling (Lempe et al., 2005; Michaels et al., 2003b).

In a detailed study of several worldwide accessions showing different flowering responses to vernalization, flowering responses were not correlated with *FLC* expression levels before vernalization or in the cold, but to the variation in time taken for *FLC* to become stably repressed by cold, which in turn correlated with accumulation of repressive chromatin marks at *FLC* (Shindo 2006). This study also failed to detect any differences between accessions in the expression of genes in the vernalization pathway, including *VIN3*. QTL analysis to identify genes underlying the vernalization response was then done with hybrids between four accessions differing in vernalization response, and Col-0. *FRI* was identified as a major effect QTL in all populations, reflecting the presence of a null *fri* mutation in Col-0 and a chromosome region containing *FLC* in three of the hybrid populations, but not in the population with the Edinburgh-0 (Edi-0) accession as a parent. This suggests that variation at the *FLC* locus itself is responsible for its ability to be stably repressed by cold. Other QTL were identified in different populations, but none mapped to components of the vernalization pathway.

Regulation of *FLC* by siRNA (Section 1.5.4) has also been implicated in natural variation in flowering time. The L*er* accession was found to have a high abundance of 24 nt siRNAs targeting a hAT transposon in the *FLC* promoter, correlating with high levels of DNA methylation and heterochromatic formation at *hAT*. In contrast, Col-0 produced lower levels of the siRNA and the *hAT* sequence showed less methylation and heterochromatin formation (Zhai et al., 2008). This study suggests that natural differences in siRNA abundance might affect flowering responses, though the effect of siRNA on *FLC* expression and flowering was not demonstrated.

1.8.4 Natural variation in other pathways

Although differences in *FRI* or *FLC* can account for a lot of the worldwide variation in flowering time, several studies have identified roles for variation in other pathways.

Variation in flowering time in response to photoperiod has been observed in previous studies of worldwide accessions. The red light receptor gene *PHYTOCHOME C* (*PHYC*), had been shown through analysis of loss-of-function mutations to play a role in flowering time in response to photoperiod (Monte et al., 2003). A role for *PHYC* in natural variation was first identified in the photoperiod insensitive German accession Fr-2, which carries 13 nonsynomynous differences in the coding sequence relative to Col-0, one of which results in a premature stop codon. Ten of these differences are shared with commonly used laboratory accession L*er*, which has a lower response to photoperiod than Col-0. Hybrid populations of L*er* and Col-0 had detected a significant QTL for flowering time under field conditions in the autumn, when daylength is short, in the chromosomal region where *PHYC* is located (Weinig et al., 2002), It was subsequently shown that the Col-0 and L*er* type alleles of PHY-C are common amongst worldwide accessions and correlate with responsiveness to photoperiod. The less responsive L*er* type *PHYC* allele was found to be more frequent and as such it was suggested that this variation may be adaptively significant. To further support this a latitudinal cline was detected in which the more responsive Col-0 type allele was more common in northern latitudes. This was particularly evident in accessions that carried apparently functional *FRI* alleles (Balasubramanian et al., 2006; Samis et al., 2008).

Natural variation in two other photoreceptor genes *CRYPTOCHOME 2* (*CRY2*) and *PHYTOCHROME B*, has been found to account for the differences in photoperiodic response to flowering between particular accessions. *CRY2* was shown to be important for photoperiod variation observed between the Cvi-0 accession from the Cape Verde Islands and L*er*. The Cvi-0 allele of *CRY2* contains a single amino acid substitution that prevents the downregulation of *CRY2* in response to short day

conditions and as such results in early flowering under short days (El-Din El-Assal et al., 2001). However the Cvi-0 allele of *CRY2* has not been identified in other accessions and variation in *CRY2* has not proved to be significant in other QTL studies. For instance a recent QTL analysis of 18 different accessions that identified four chromosomal regions important for natural variation in flowering time found a significant QTL in the region containing *PHYC* but not *CRY2* (Salome et al., 2011).

Sequence variation in *PHYB* was also identified between natural accessions. This sequence variation was shown to affect protein function although its effect on flowering was not examined. As with *CRY2*, the novel *PHYB* alleles have proved to be rare (Filiault et al., 2008; Salome et al., 2011).
Additionally natural variation has also been observed in the flowering time gene *VIP4* and a QTL at the bottom of chromosome five this will be discussed further in Chapter 6.

1.8.5 Adaptation of flowering behaviour

Considerable genetically-determined variation in flowering time has been detected between worldwide accessions under controlled conditions. This raises the question of whether this variation could be adaptive. For adaptation to explain genetic variation in flowering, genetic differences should affect flowering time under natural conditions and lead to differences in fitness (e.g., seed yield). A latitudinal cline in responses to vernalization has been reported, with accessions from nearer the equator tending to be less responsive to vernalization (Stinchcombe et al., 2004). Although consistent with adaptation to differences in temperature, the responses of different genotypes was not tested in the field and the fitness of accessions was not compared under different conditions.

Mutants in several flowering pathways in an isogenic background, including *fri* mutants have been compared for flowering time under natural conditions. Although *fri* is associated with rapid flowering in the absence of vernalization under controlled conditions, it only allowed early flowering in plants that germinated within a particular time window in late summer and only at some field sites (Wilczek et al.,

2009). Other studies have compared the performance of worldwide accessions under common garden conditions (Brachi et al., 2010) or have carried out QTL analysis in controlled conditions that mimic temperature and daylength differences between regions from which parental accessions were collected (Brock et al., 2010) These suggest that many genes with a major effect on flowering behaviour under controlled conditions have little effect on flowering in the field.

1.8.6 Local variation

Using worldwide Arabidopsis accessions to study natural variation has several disadvantages. These include the fact that the precise collection sites might not be known and therefore local environmental conditions for many of the accessions are unknown. A second problem is that Arabidopsis shows population structure in that accessions from closer locations tend to share more genes and are therefore likely to show similar responses, even if these are not adaptive.

One solution is to study variation on a more local scale, where there is likely to be less genetic variation and less population structure, making detection of association between genes and phenotypes more feasible and where there is ready access to populations and sites for field studies, including reciprocal transplantation experiments. This, however, depends on local populations showing significant variation in the character.

1.9. Populations on which this study is based

This study uses populations of *Arabidopsis thaliana* collected from in and around Edinburgh and the Lothians. Two separate collections were made - one of plants considered to functional winter annuals and one of plants considered to be functional summer annuals; further details of how these plants were collections can be found in Chapter 2.

1.10. Aims of the project

The aims of this study are to assess variation in flowering time and flowering responses to different environmental conditions in local genotypes of Arabidopsis. This project further aims to gain an understanding of the underlying genetic variation that contributes to these differences, firstly by examining variation in the *FRI* and *FLC* genes previously shown to play a role in flowering time variation, particularly in the vernalization response, and secondly by mapping genes that contribute to natural variation in flowering time on a local scale.

2 Materials and Methods

2.1 Plant materials and growth conditions

2.1 1 Collection sites of local Arabidopsis populations

Andrew Hudson and Sarah Whittall collected populations of *Arabidopsis thaliana* from sites around Edinburgh and the Lothians (exact locations can be found in Table 2.1). In the majority of cases these were collected as functional winter annuals – i.e., as rosettes in the winter, although a few individuals were observed to have started flowering, and were therefore not overwintering as rosettes. Plants were then grown in a controlled glass house and any flowers or siliques that had begun to develop in the field were removed so that it was clear that all seeds collected later came from self fertilisation.

Plants were given a number according to the location they were collected from, a letter representing the tray they were transplanted into and then another number to identify the unique individual within that tray. For example, family 10B2 comes from a plant that was originally collected from location 10 (Loanhead, Midlothian) and was the second plant in tray B from that location.

A second collection was made by Abigail Harter (Table 2.2). This time the plants were collected as functional summer annuals, meaning that they were identified as seedlings early in the spring and then collected later in the spring or early summer once they had started to flower. It was therefore clear that these

Table 2.1 Collection sites for local winter annual populations used in this study. The identity of each population along with a description and coordinates for the collection sites, flowering habit shows whether these populations were collected as functional winter annuals (WA) or summer annuals (SA)

	Location	Position	Elevation (m)	Flowering Habit
01	University of Edinburgh, King's Buildings Campus, S. of Swann Building	N55º55.29 W003º10.266'	70	WA
02	University of Edinburgh, King's Buildings Campus, N. of Swann Building	N55º55.306' W003º10.265'	70	WA
03	University of Edinburgh, King's Buildings Campus, S. of Rutherford Building	N55º55.284' W003º10.237'	67	WA
04	University of Edinburgh, King's Buildings Campus, Forestry plots	N055º55.295' W003º10.180'	62	WA
05	University of Edinburgh, King's Buildings Campus, S of JCMB	N055º55.271' W003º10.397'	75	WA
06	University of Edinburgh, King's Buildings Campus, W. of CSEC	N055º55.278' W003º10.509'	78	WA
07	University of Edinburgh, King's Buildings Campus, W. of CSEC	N055º55.277' W003º10.749'	79	WA
08	University of Edinburgh, King's Buildings Campus, W. of CSEC	N055º55.282' W003º10.750'	79	WA
09	University of Edinburgh, King's Buildings Campus - W. of SAC	N055º55.270' W003º10.714'	79	WA
10	Loanhead	N055º52.785' W003º09.306'	148	WA
11	Hillend	N055º53.260' W003º12.643'	249	WA
12	Liberton	N055º53.260' W003º09.826'	115	WA
13	Straiton	N055º53.147' W003º09.609'	152	WA

Table 2.2 Collection sites for local summer annual populations used in this study. The identity of each population along with a description and coordinates for the collection sites, flowering habit shows whether these populations were collected as functional winter annuals (WA) or summer annuals (SA)

	Location	Position	Flowering Habit
101	Penicuik	N55° 49. 37.56’ W003° 10.2’	SA
102	Alva	N56° 9 6.1194’ W003° 54.1194’	SA
103	Dalmeny	N55° 58.0794’ W3° 22.44’	SA
104	KingHorn	N56° 53.9994’ W3° 56.28’	SA
105	Cramond	N55° 38.9994’ W3° 1.08’	SA
106	Bo’ness	N56°45.72’ W3° 54’	SA
107	Pittenweem	N56° 56.5194’ W2° 56.2794’	SA
108	Aberfeldy	N56° 31.08’ W3° 1.7994’	SA
109	Kennoway	N56° 54.3594’ W3° 6.8394’	SA
110	Inverkeithing	N56° 32.1594’ W3° 15.48’	SA
111	Torryburn	N56° 33.8394’ W3° 9.72’	SA
112	Kincardine	N56° 58.6794’ W3° 21.24’	SA
113	Dirleton	N56° 2' 24.7194 W2° 47' 17.5194	SA
114	Prestonpans	N55° 42.4794’ W2° 33.6’	SA
115	Dalkeith	N55° 30.12’ W3° 30’	SA
116	Stockbridge	N55° 34.9194’ W3° 24.12’	SA

plants had both germinated and flowered between spring and summer of the same year. As with the winter annual populations, all flowers and siliques were removed prior to growth in the glasshouse.

2.1 2 Sterilising and germinating seeds

Seeds were either germinated on soil or in Petri-dishes on 0.5 x MS medium (0.49% Murashige and Skoog, 0.3% sucrose and 0.1% Difco Agar) (Murashige and Skoog F, 1962). If they were to be germinated on soil then they were first imbibed in 0.1% agar and then kept in the dark at 4°C for approximately 72 hours in order to stratify them and therefore break dormancy. After this they were then sown (usually one or two seeds per pot) using a pipette. After germination seedlings were thinned so that only one plant per pot remained.

Seeds germinated on MS plates were first sterilised. Seeds were transferred into 1.5 ml Eppendorf tubes, 1 ml of 70% EtOH, 0.05% Tween 20 (v/v) was added and the tubes were agitated for 10 minutes. The ethanol was then poured off and 500 µl of 100% EtOH added. The seeds were then immediately transferred to a labelled sterile filter paper to dry in a laminar flow hood. After all of the alcohol had evaporated the seeds were evenly sprinkled over a sterile plate of MS medium and the plate sealed with Micropore tape. Plates were then wrapped in foil to keep them dark and stored at 4°C for approximately 72 hours, after which time they were transferred to 20°C under constant light for germination to occur. A week after germination, seedlings were transplanted into 8 cm x 8 cm x 7 cm pots of soil (Fison's F2 compost), to which had been added by volume, 33% perlite, 20% sand and ~0.03% Intercept

insecticide. Pots were placed into trays (typically 21 pots per tray). Trays were turned 180° and their positions swapped daily, to reduce any effect of position on flowering time.

2.1 3 Vernalization treatments

Plants were either vernalized as seeds or, more often, as seedlings. Seeds were vernalized by imbibing them in 0.1% agar and storing them in the dark at 4°C for 6-8 weeks. To vernalize seedlings, seeds were sterilised and germinated on MS plates, transplanted into soil one week after germination and transferred into a 5°C growth cabinet under long day conditions (16hr light, 8hr dark). For experiments to test vernalization requirements, plants remained at 5°C for different lengths of time before being moved into a 21°C LD growth cabinet.

2.1 4 Flowering time analysis

The flowering times of plants were scored as the total number of leaves produced before flowering. The leaves of each plant were counted approximately once a week and marked with a permanent pen until the first floral bud had opened. Care was taken to avoid leaves from axillary shoots in late flowering plants. Separate counts were kept for the numbers of rosette leaves and cauline leaves on the main inflorescence and these two values were added to give the total leaf number. In instances where it was no longer possible to count rosette leaf number (due to extensive growth of axillary shoots) plants were considered to be very late flowering and given a nominal leaf number that was higher than plants that had flowered, usually >150 leaves. However where possible; the date at which the first flower

opened was also recorded for extremely late flowering plants, this meant that their flowering time in days to flower could be analysed relative to all other plants in that experiment.

2.2 DNA preparation and analysis

2.2 1 DNA extraction from plants

Fresh leaf tissue was stored at -80°C in a 1.5ml Eppendorf tube with two 4 mm diameter steel ball bearings. Up to 24 tubes were then transferred to a block pre-cooled at -80°C and macerated twice in a bead mill, rotating the block 180° between macerations. Each maceration involved shaking the block at 40 Hz for 1 min. After the second maceration 0.5 ml of extraction buffer (100 mM Tris 8.0, 1.4 M NaCl, 20 mM EDTA, 2% w/v. CTAB, 0.2% (v/v) β-MCE, modified from (Doyle and Doyle JL, 1990) was added to each tube and the tubes were shaken again for 1 min in the bead mill. Tubes were then incubated in a heating block at 65°C in a fume hood for 20 minutes and allowed to cool for ~ 2 min before addition of 0.5 ml of $CHCl_3$ and mixing with a vortex. The tubes were then spun in a centrifuge at 13,000 *g* for 5 minutes and the aqueous layers transferred into clean labelled Eppendorf tubes. In order to precipitate nucleic acids 220 µl of isopropanol was added to each tube and gently mixed by rocking the tube backwards and forward. The samples were then centrifuged at 13,000 *g* for 5 minutes to pellet nucleic acids, the alcohol was poured off and most of the remaining liquid removed with a pipette. The tubes were then centrifuged again for 1 min and any remaining alcohol removed with a pipette. The pellet was dissolved in 50 µl of TE with 10 µg/ml RNase A and stored at 4°C overnight. The following day the DNA was precipitated by adding 0.1 volume of 3 M NaOAc (pH 7.4) and 2.5 volumes of EtOH and leaving it on ice for approximately

1 hr. The tubes were then spun at 21,000 *g* for 5 minutes, the alcohol removed and the tubes dried before each pellet was dissolved in 50 µl of TE.

2.2 2 PCR

The local populations were screened for previously identified disruptions in genes known to be involved in natural variation in flowering time. This was done using primers to amplify the relevant sections of the gene via PCR. A typical PCR reaction mix is given below:

2.0 µl 10X PCR Buffer (50 mM TrisCl pH 8.3, 500 µg/ml BSA, 0.5% Ficoll, 1% sucrose, 30 mM KCl, 3 mM $MgCl_2$, 1 mM tartrazine)

0.4 µl 10 mM dNTPs

0.4 µl 10 µM Forward Primer

0.4 µl 10 µM Reverse Primer

0.08 µl *Taq* DNA Polymerase (~300 U/µl)

14.7 µl Sterile distilled water

2.0 µl DNA (approx 20 ng genomic DNA)

Commercial *Taq* DNA polymerase (NEB) was used for amplifying cDNA. In these cases the reaction mix was as follows:

2.0 µl 10X Reaction Buffer

2.0 µl 10 µM Forward Primer

2.0 µl 10 µM Reverse Primer

1.6 µl 2.5 mM dNTPs

2.0 µl 25 mM $MgCl_2$

5.3 µl sterile distilled water

5.0 µl cDNA template (approx 2 ng cDNA)

PCR reaction programs varied according to the annealing temperatures of the primers (Table 2.1) and the size of the product (~40 sec extension at 72°C per kb of product). A typical reaction cycle is shown below:

95°C 3 min

95°C 10 sec

X°C 15 sec — 35 Cycles

72°C 35 sec

72°C 10 min

4°C ∞

2.2 3 Agarose gel electrophoresis

Agarose gel electrophoresis was used to visualise DNA from PCR. Gel was prepared using 1.7-3% (w/v) agarose in 0.5x TBE buffer (0.089 M Tris base, 0.089 M boric acid and 0.022 M EDTA) by heating in a microwave oven until the agarose was completed dissolved. The gel was allowed to cool to approximately 65°C before adding ethidium bromide to a concentration of 200 ng/ml (w/v). Using a pipette 25 ml of gel was transferred onto a glass slide with a comb and allowed to set. The gel was submerged in a tank of 0.5x TBE buffer, the samples loaded and electrophoresis carried out at ~3 V/cm for the length of time required to get appropriate resolution. A DNA size marker was included to estimate the sizes of PCR fragments. DNA was visualized using a GeneFlash (Syngene) UV transilluminator.

2.2 4 Sequencing

Sequencing reactions were performed by the University of Edinburgh sequencing service (The Gene Pool). Samples for sequencing consisted of: 0.6 µl DNA template (taken directly from a PCR), 1 µl of 10 µM primer and 4.4 µl of sterile distilled water.

Sequence data was subsequently aligned, edited and analysed using DNAstar software.

2.3 RNA preparation and analysis

2.3 1 RNA extractions

RNA was extracted from fresh tissue (whole seedlings) that had been stored at -80°C using an Invitrogen Pure Link RNA Mini Kit, according to the manufacturer's instructions. The RNA was eluted from the purification columns in a volume of 30 µl, and its concentration and quality assessed using a Nanodrop spectrophotometer and by agarose gel electrophoresis. The RNA was then diluted to 0.24 µg/µl in RNase free water (as provided in the kit).

2.3 2 cDNA Synthesis

cDNA was synthesised using a Promega Access RT-PCR System and an Oligo(dT) primer (100 mM). The manufacturer's instructions were used, except that the volume of each reaction was reduced to 20 µl and the temperature cycle modified to:

25°C 5 min

45°C 120 min

94°C 2 min

4°C ∞

cDNA samples were diluted to either 100 µl or 200 µl with sterile distilled water.

2.3 3 qRT-PCR

In order to quantify the level of RNA for relevant flowering time genes quantitative reverse transcription - PCR was used. For all reactions the intercalating dye SYBR green (produced by Roche) was used. Once a molecule of SYBR green has bound to double stranded DNA it fluoresces, therefore the more bound SYBR green (therefore by extension the more product present) the higher the level of fluorescence detected by the LightCycler PCR machine. There are caveats to this in that the SYBR green will not discriminate between the desired PCR product and other double-stranded DNA products such as primer dimers. For this reason melting curves were obtained for each reaction. Artefacts such as primer dimers that are shorter fragments of double stranded DNA than the product of interest, melt and lose fluorescence at a lower temperature (VanGuilder, 2008). Any reaction in which smaller artefacts were identified from melt curves were not analysed further.

For each combination of cDNA and primers, a 10x dilution series (undiluted, 1:10, 1:100, 1:1000) of cDNA was used to calculate the efficiency of the reaction and create a standard curve by which the experimental reactions were calibrated. In an ideal scenario the values obtained from this series would show ten times less signal in the 1:10 sample compared to the undiluted sample and so on to the end of the

series. This then gave an efficiency value representing the fold-change in product after each cycle of Q-PCR. Ideally the efficiency would be two, but lower values were obtained and used to estimate the relative concentration of template in the experimental samples.

In order to normalise reactions from different cDNA samples, a housekeeping gene *FDH* was amplified from each, using the primers shown in Appendix A.

A Q-PCR reaction contained the following components:

5 µl LightCycler® 480 SYBR Green I Master

5 µl cDNA Template

0.04 µl primer mix (50 µM each of Forward and Reverse)

Reactions were made up in a 96 well plate (LightCycler ® 480 Multiwell Plate) that was sealed with transparent Multiwell Sealing Foil. They were mixed thoroughly and spun in a centrifuge for 2 minutes at 3,000 *g*. The plate was then put into a LightCycler® 480 and set to the relevant program.

The abundance of template is estimated using the crossing point, which is the cycle number at which florescence is first detected by the Light Cycler.

2.4 Complementation of *VIN3* alleles

2.4 1 Preparations of constructs by extension overlap PCR

In order to analyse the effect of the *VIN3* alleles found in families 4D1 and 11C1 constructs were required that included the coding region and the promoter of *VIN3* from these two accessions. This was done with extension overlap PCR (Fig 2.1).

Primers were designed that would amplify either the promoter region or the coding region of *VIN3* with an overlap between the products in part of the 5' UTR, which did not differ between the alleles from 4D1 and 11C1. Initial PCR was carried out to amplify the promoter regions and the coding regions of the *VIN3* alleles from 4D1 and 11C1 in separate reactions. A second extension overlap PCR was then carried out to amplify both promoter and coding regions by combining the products from the first reaction in the following way:

4D1 promoter + 4D1 coding region

11C1 promoter + 11C1 coding region

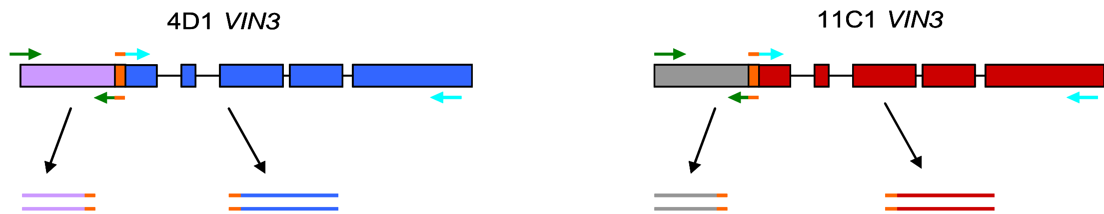
4D1 promoter + 11C1 coding region

11C1 promoter + 4D1 coding region

The overlapping region of the two products was expected to anneal and prime synthesis of a full-length *VIN3* fragment. Primers from the extreme 5' and 3' ends of *VIN3* were included in the reaction to amplify this full-length product. A proof-reading *Taq* polymerase, Phusion *Taq* from Finnzymes, was used to amplify the coding regions and for the final overlap amplification. However, this polymerase amplified non-specific bands with the promoter primers and so lab-produced *Taq* polymerase was used for the initial promoter amplification and products sequenced before use in overlap PCR.

Product generated from the overlapping extension PCR was purified using a QIAquick Gel extraction kit following the manufacturer's instructions.

1. In separate reactions amplify the promoter and coding region of the gene, creating products with an overlap in the 5’ UTR



2. Mix products, denature and anneal without primers

3. Extend product using *Taq* Polymerase

4. Amplify using primers that will anneal to 5’ and 3’ of new product

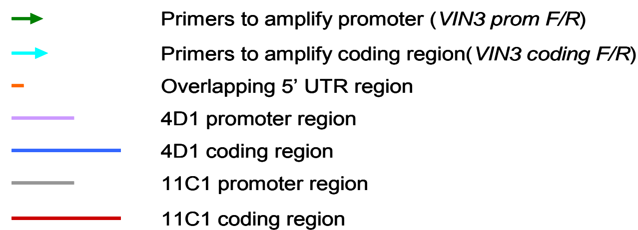


Figure2.1 Explanation of overlapping extension PCR. At each stage four separate reactions were carried out. Full details of the primers used can be found in Appendix A. For clarity the overlapping region is indicated in orange.

2.4 2 Cloning into plasmid vectors and transformation of *E. coli*

A Gateway entry vector system (Invitrogen D-TOPO cloning kit) was used for cloning the overlap PCR products into a plasmid vector (pTopo; Fig 2.2). The TOPO reaction is shown below:

4.0 µl PCR product, approx 40ng DNA (from overlap extension PCR)

1.0 µl salt solution (as included in the kit)

1.0 µl TOPO vector (pENTR D-TOPO)

This reaction was left at room temperature for 5 minutes and then put on ice.

Tubes of DH5α competent *E. coli* cells (produced according to (Hanahan, 1983) were thawed on ice; they were then added to the TOPO reaction and gently mixed before incubating on ice for 20 min. The cells were then heat shocked by putting the tubes into a 42°C water bath for 1 min and then immediately transferred to ice for 3 min. SOC medium (1 ml of 2% bacto-tryptone (w/v), 0.5% bacto-yeast extract (w/v), 8.56mM NaCl, 2.5mM KCl, 10mM $MgCl_2$ and 20mM glucose) was then added to each tube and tubes were incubated at 37°C for 60 min to allow antibiotic resistance to be expressed. Cells were then pelleted by centrifugation at 2,700 *g* for 3 min, most of the supernatant discarded and the cells resuspended in the remaining medium before spreading onto plates of LB agar medium containing 50 µg/ml kanamycin sulphate (LB-Kan). LB agar medium contains 1% agar, 1% Difco Bacto tryptone, 0.5% yeast extract, 0.5% NaCl, 0.1% D-glucose (all w/v) (Bertani, 1951). Plates were inverted and incubated overnight at 37°C.

2.4 3 Screening clones by PCR

In order to test for the presence of *VIN3* clones a few colonies from the transformation were streaked onto LB-Kan plates using sterile pipette tips. Bacterial cells remaining on the tip provided templates in PCR with primers specific to the relevant part of *VIN3*. PCR products were then visualised using gel electrophoresis to ascertain which colonies contained *VIN3* fragments.

2.4 4 *E. coli* cultures in liquid media

Colonies that tested positive for the presence of *VIN3* sequences were used to inoculate 5 ml of sterile LB-Kan liquid medium (as for LB-Kan agar, but with without agar). The cultures were incubated at 37°C with constant shaking overnight.

2.4 5 Plasmid mini-preps

Plasmid DNA was extracted from 3 ml of overnight bacterial cultures using a QIAprep® spin mini prep kit following the manufacturer's instructions. DNA was eluted from the column using 50 µl of the kit's EB buffer. Plasmid inserts were then sequenced using *VIN3*-specific primers (Appendix A).

2.5 Transformation of genes into *Arabidopsis*

2.5.1 Transformation into Gateway Ti-Plasmid

The destination vector chosen for expressing the construct was the Gateway Binary Vector pGWB1 (Nakagawa et al., 2007) In order to incorporate the *VIN3* construct into pGWB1 an LR reaction was done, this causes

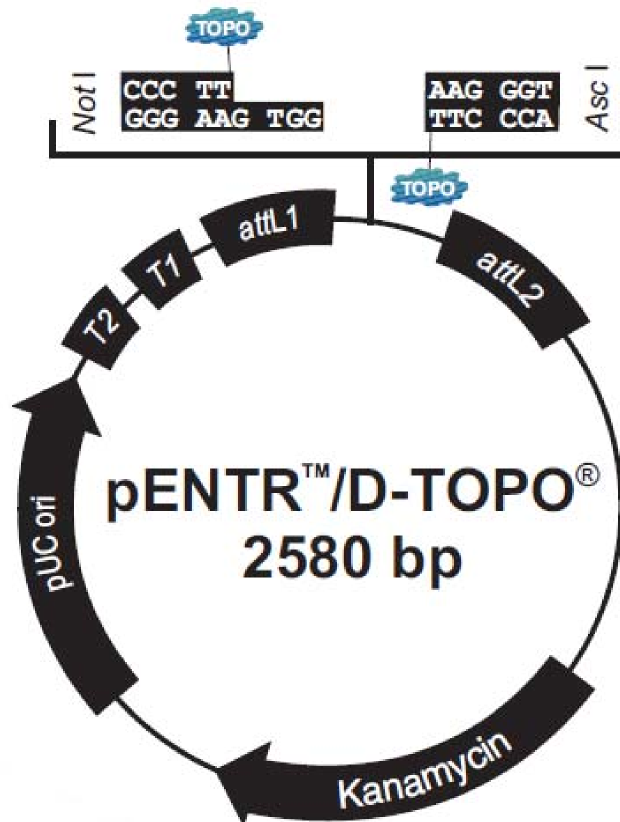


Figure 2.2 pENTR^{TM}/D-TOPO vector used for cloning *VIN3* PCR products in *E. coli.* Diagram taken from Invitrogen pENTR Directional TOPO Cloning Kits manual (version G, 6th April 2006)

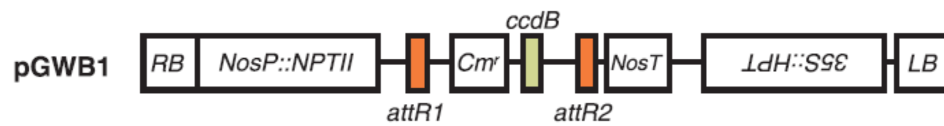


Figure 2.3 Gateway Binary Vector. Illustration of structures between right and left border in pGWB1 vector. PGWB1 is a simple vector used for complementation figure from (Nakagawa et al., 2007).

recombination to occur between the *attL* sites flanking the insertion in the entry clone (production of which is described in section 2.4ii) and the *attR* sites present in the destination vector. This LR was done according the manufacturer's instructions for Invitrogen Gateway® LR clonase^{TM} II Enzyme Mix, the only exception being that the reaction was incubated at 25°C overnight instead of the suggested 1 hour to improve efficiency of recombination.

Because the pTOPO entry clone produced in section 2.4.2 encoded Kan resistance, while pGWB1 had an additional marker, HPT, conferring resistance to hygromycin, bacteria carrying pGWB1 were selected with both Kan and Hyg. Recombination between *attL* and *attR* sites leads to loss of the lethal *ccdB* marker gene from pGWB1, allowing only cells with recombinant plasmid to grow.

2.5.2 Transformation of *Agrobacterium tumefaciens* (*Rhizobium radiobacter*)

Plasmid DNA (10 µl containing ~500ng DNA) was added to 100 µl of competent *Agrobacterium* GV101 cells and thoroughly mixed. Competent *Agrobacterium* cells were produced by the method of Weigel and Glazebrook (Weigel and Glazebrook, 2010). The cells were then put into liquid nitrogen for one minute before immediately being transferred to a 37°C incubator, until completely thawed. Then 1 ml of liquid YEP medium (1% yeast extract, 1% peptone, 0.5% NaCl, - all w/v) was added to the tube and cells incubated at 28°C for one hour. After this time the cells were collected by centrifuging for one minute at 20,800 *g*. At this point 900 µl of YEP medium was removed using a pipette and discarded; the cells were then resuspended in the remaining medium and spread on YEP plates containing Kan (50

μg/ml) and Hyg (50 μg/ml). The plates were incubated at 28°C for 2-3 days until colonies of a suitable size had grown.

As described for the entry clone, plasmid mini-preps were made from overnight cultures, followed by PCR with *VIN3* specific primers. This was done to ensure the correct insert was present in the destination vector.

2.5.3 Transformation of *Arabidopsis*

Under sterile conditions a sample from a single *Agrobacterium* colony was added to 10ml of liquid YEP medium containing 80 μg/ml of gentamycin; to select against contaminant *E. coli*, and 25 μg/ml of Kan, and 14 μg/ml of Hyg to select for the destination vector containing the relevant insert. The cultures were then put in a shaking incubator at 28°C for 18-24 hours.

Prior to inoculation of *Arabidopsis* plants, 0.02% (v/v) Silwet L77 was added to the *Agrobacterium* culture and mixed well. The culture was pipetted onto unopened floral buds. Plants were then sealed into an autoclave bag for approximately three hours. They were inoculated in the same way 5 days later.

Inoculated plants were allowed to set seed. After the seeds had dried, they were sterilised and sown onto selection plates; consisting of 0.5x MS medium, 1% agar with 25 μg/ml Kan. Seeds were stratified in the dark at 4°C for 72 hours before allowed to germinate at 20°C under constant light conditions. After germination they were allowed to grow sufficiently to be able to select transformants. Transformants

were then transferred to soil and grown under long day conditions at ~20°C, either with or without a period of vernalization at 5°C.

3 Flowering time variation in local summer and winter annual populations

3.1 Introduction

The aim of this chapter is to assess the extent of flowering time variation locally and how local genotypes vary in their response to photoperiod and vernalization. The timing of floral induction is crucial for a plant to ensure reproduction and seed set occur under favourable conditions and therefore at the optimum time of year. It has also been demonstrated in Arabidopsis and other species that flowering is plastic and can therefore respond to seasonal and environmental cues present for a given populations (Korves et al., 2007).

This chapter aims to examine how much variation is present locally in flowering behaviour and how this relates to variation found in studies of worldwide accessions. It has been reported that worldwide there is extensive variation in flowering time in response to variables such as daylength and temperature. Rapid cycling accessions that do not require vernalization have been shown to have evolved independently in several locations suggesting the adaptive importance of this trait (Johanson et al., 2000). To a lesser extent accessions that vary in photoperiod response have also been identified, supporting the idea that genetically determined variation in flowering responses might represent adaptation to the local environment.

Additionally this chapter will assess how flowering time variation relates to the location at which a population were collected as a way of exploring whether flowering time might be adapted to the local environment. There is some evidence for adaptation on a worldwide scale because several studies have found latitudinal clines in flowering time and flowering response to environmental cues such as photoperiod (Stinchcombe et al., 2004). However this is by no means conclusive as other studies have failed to find correlations and instead suggest that apparent clines are an artefact of sampling. They argue that it is more ecologically relevant to look at adaptation to specific variables

between local environments rather than using correlation with latitude, which is only a crude measure of variation between environments (Shindo et al., 2005).

3.2 Response to photoperiod from local winter annuals

Two collections of local Arabidopsis populations were made; they were either collected as functional winter annuals (Andrew Hudson and Sarah Whittall) or functional summer annuals (Abigail Harter). Functional winter annuals were collected in the winter as rosettes. A few plants had begun flowering and in these cases all flowers and siliques were removed. All plants were then grown under controlled conditions and left to self fertilize. Due to the high rate of self fertilization that occurs in Arabidopsis, a plant is likely to be homozygous. This was supported by genotyping of the field collected winter annual plants at six polymorphic microsatellite loci by Sarah Whitall. Only one plant was found to be heterozygous, at only one of the loci.

In order to gain a better understanding of how these local genotypes responded to different environments they were grown under four sets of controlled conditions. In order to assess the effect of photoperiod they were grown in both long days (LD) of 16 hours light and 8 hours dark or short days (SD) of 8 hours light and 16 hours dark. They were also either given a vernalization treatment - imbibed in 0.01% agar and left at $4^{o}C$ in the dark for a period of 6 weeks – or were sown without vernalization. An average of five plants per family were grown under each condition (a family refers to the progeny of a single parent). More plants survived under SD vernalized than SD unvernalized conditions, and therefore analysis of responses to photoperiod concentrated on plants that had undergone six weeks vernalization and were grown in either short and long days.

Figure 3.1 Flowering times of winter annual (WA) families under short day (SD) and long day (LD) conditions. Flowering times are given in mean rosette leaves until flowering (RLF) ± their standard errors. All families are shown in (A), except the extremely late flowering families from site 10, which are shown in the inset (B) for clarity.

The average flowering time for all the plants (in rosette leaves to flower) was longer in SD (55.3) compared to LD (10.2), as expected (Fig. 3.1A). However, there was considerable variation between families within each experiment, ranging from 5.2 to 83.7 leaves in LD and 31.2 to 81.5 leaves in SD.

Because members of the same family are assumed to be genetically identical, variation within families will be caused by non-genetic factors. The variance within families compared to the total variance therefore estimates how much of the observed variation in a character is due to genetic differences. This approximates to the character's broad-sense heritability, H^2 (Falconer and Mackay, 1996). The winter annual collection showed a heritability value of 0.80 under both long and short day conditions, suggesting that up to 80% of the observed variation in flowering time under either condition is genetically determined.

Although most winter annual families flowered earlier in LD than SD, families from location 10 flowered late under both conditions (as a result of their extreme late flowering they are displayed on a separate axis for clarity in Figure 3.1B). Excluding families from location 10, the average response to photoperiod resulted in 82% fewer leaves under LD conditions compared to SD and there was little variation between families for the response (Figure 3.1B). In contrast, families from location 10 showed only a 3% reduction in leaf number which a t-test suggested was not significant (p value of 0.66) and therefore that families from location 10 could be considered insensitive to photoperiod. One possible reason for their extreme lateness is a disruption in the photoperiod pathway.

The winter annual families could therefore be put into one of two photoperiod response categories: they either respond or they do not. Within these categories there is little variation.

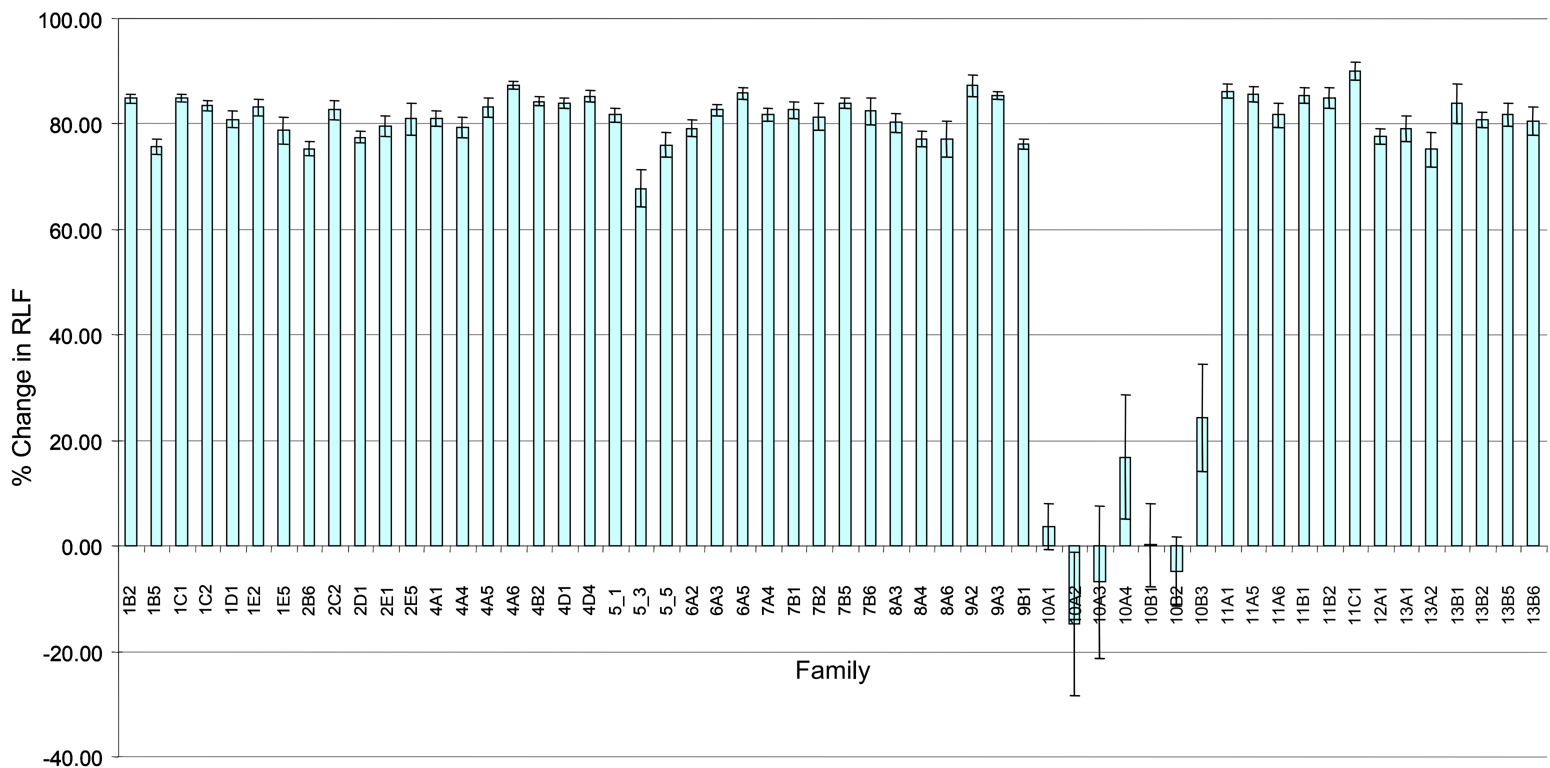


Figure 3.2 Percentage effect of photoperiod on flowering time. Percentage change in RLF of winter annual families when grown under SD compared to LD conditions. Values are family means ± their standard errors.

3.3 Response to photoperiod from local summer annuals

Functional summer annuals were initially identified as seedling in the early spring and then later in the same season were collected once flowering had begun. As with the winter annuals all developing flowers and siliques were removed and the plants were grown under controlled conditions and allowed to self fertilize.

Considerable variation in flowering time under both SD and LD conditions was also observed for the summer annual population. With one exception, families were also found to flower earlier under LD than SD conditions. Heritability values of 0.80 were calculated for flowering time variation in summer annual families under both long and short day conditions, suggesting that most of the variation observed is determined genetically.

Although family 11.31 appeared to flower earlier under SD conditions (with an average of 8.8 leaves) than LD (17 leaves; Fig. 3.3; 3.4), the value for LD involved a larger standard error (values ranged from 9 to 26 leaves). As such it is unlikely that this family flowers earlier in response to SD conditions but instead is unresponsive to photoperiod.

The summer annual families showed a wider range of responses to photoperiod than the winter annual population. Winter annual families flowered with an average of 55.3 leaves under SD conditions and 10.2 leaves under LD conditions with an overall average 82% reduction in leaf number in response to LD conditions. Whereas the summer annual families flowered with an average of 62.3 leaves under SD conditions and 21.7 leaves under LD conditions overall the summer annual families had an average 65% reduction in RLF in response to LD conditions. This variation in response between winter and summer annuals has not been reported before, previous studies have assigned the functional flowering behaviour of their collections after genetic analysis. In this instance the collections were made based on their flowering behaviour in the wild. As such this finding is novel however it should be noted that the summer annuals were

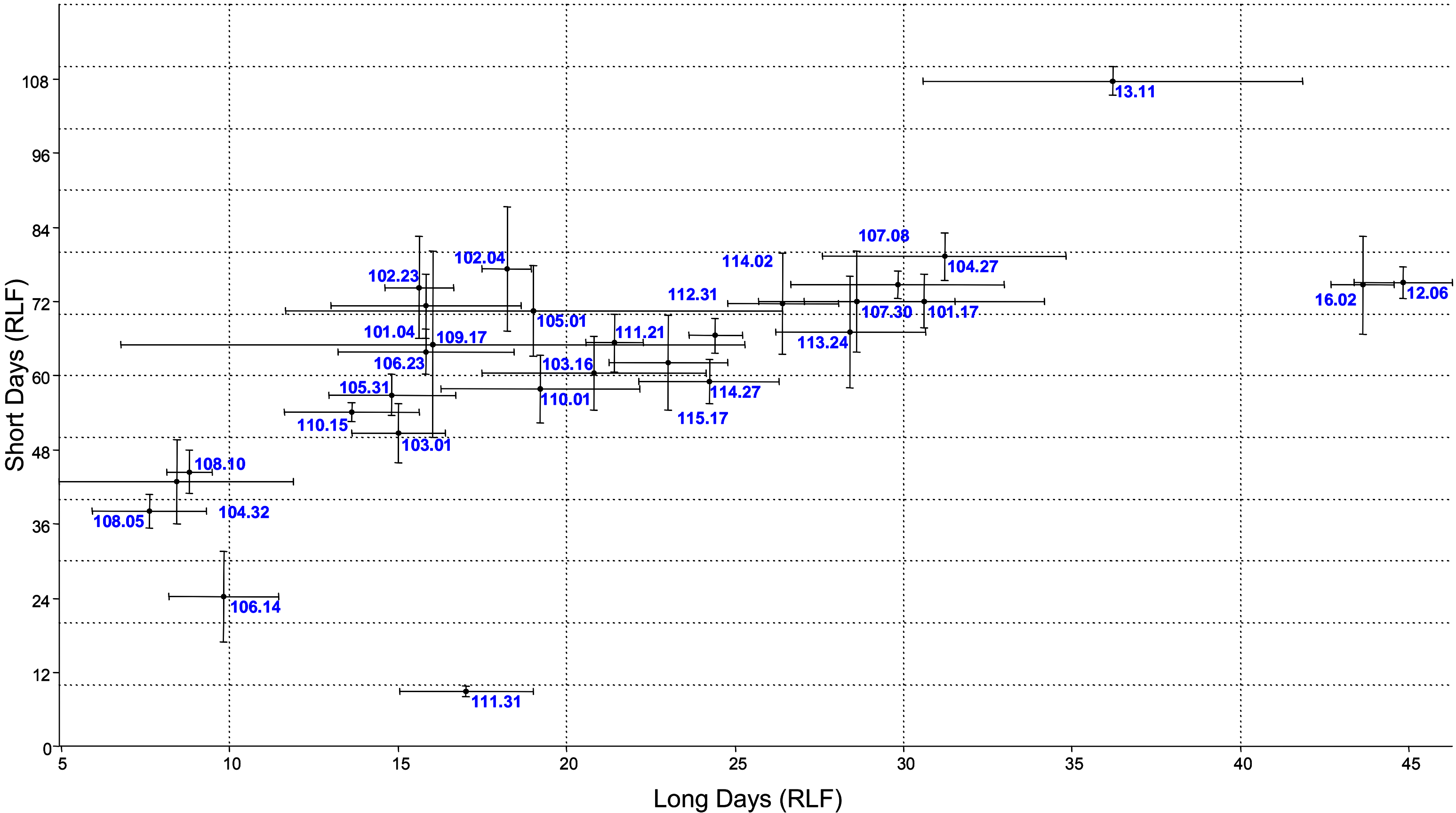


Figure 3.3 Flowering times of summer annual (SA) families under short day (SD) and long day (LD) conditions. Flowering times are shown as mean rosette leaves until flowering (RLF) with bars representing standard errors.

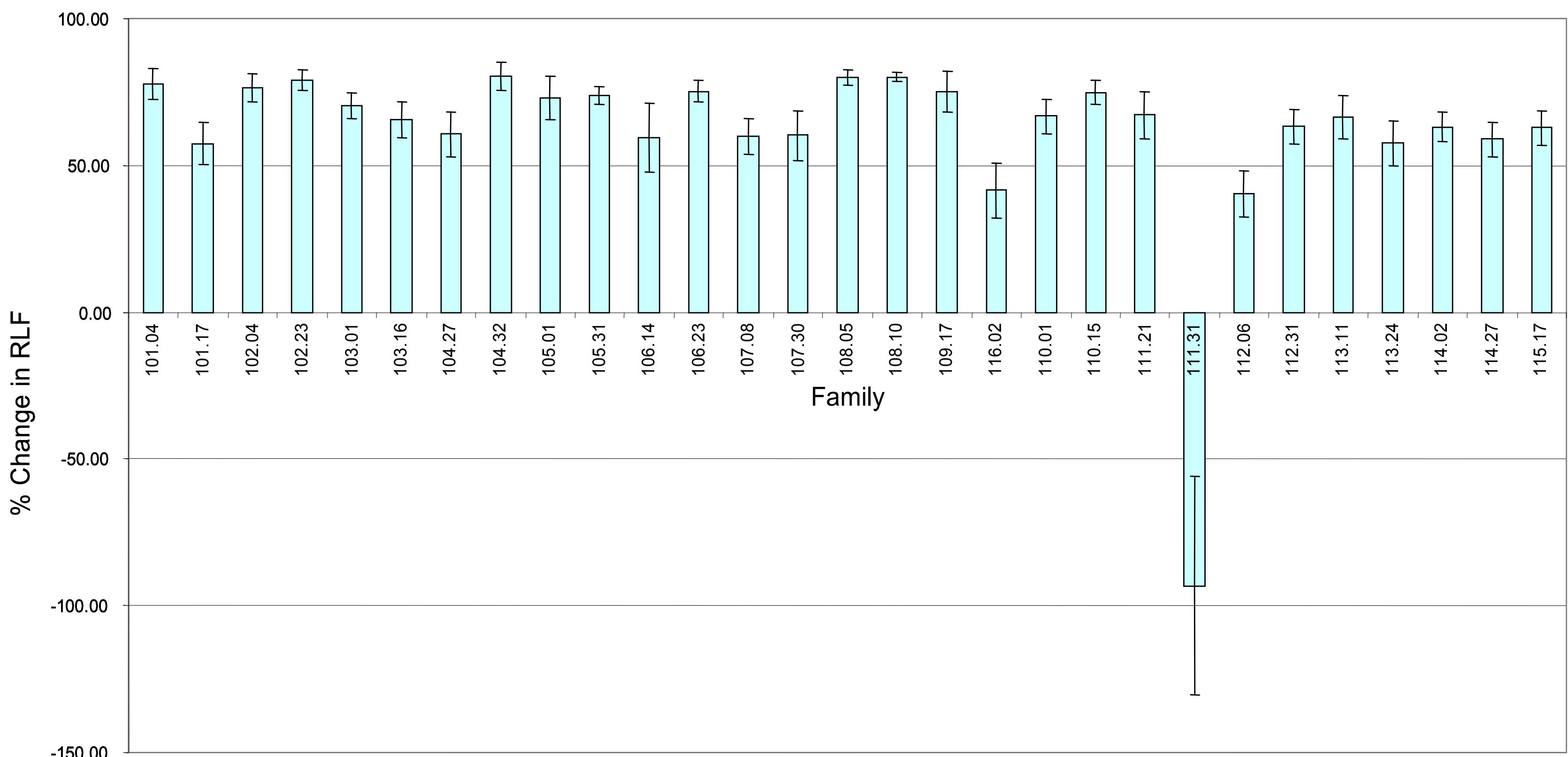


Figure 3.4 Percentage effect of photoperiod on flowering time in summer annual (SA) families. Percentage change in RLF of winter annual populations when grown under SD compared to LD conditions. Values are means from each family from ~X individuals grown under each condition, ± standard errors.

collected from a much wider area and as such this may have contributed to sampling more variation and so comparisons between the winter and summer annual populations should be viewed cautiously.

Although this suggests that summer annuals might in general flower earlier than winter annuals, the winter and summer annual populations had been collected from different locations and these locations covered a much larger geographic region in the case of the summer annuals. Therefore differences in flowering behaviour might reflect genetic differences that are related to geography or sampling (see Chapter 2). Flowering times for the two populations were also carried out independently and therefore variation between experimental conditions might also have contributed to differences in the estimates of average flowering times.

3.4 Response to vernalization in local winter annuals

In order to assess variation in vernalization response seeds were imbibed in agar and treated at $4^{o}C$ for a period of 6 weeks. They were then germinated on soil and grown under LD conditions. As before rosette leaves to flower (RLF) was used as a measure for flowering time.

Again a significant range in flowering times was observed under both conditions tested (Fig.3.5). Of the variation observed, 80% and 79% could be attributed to genetics in vernalized and unvernalized conditions, respectively.

For almost all families, vernalization shortened the time to flower, as expected. However, families varied considerably in their response to vernalization. This is most clearly illustrated as the percentage decrease in leaf number in response to vernalization (Fig.3.6); the responses ranged from zero in families 1E5, 7A4 and 8A3 to 86% in family 11C1.

It should be noted that families from location 1 show little response to vernalization (Fig. 3.5), however this is not reflected in the percentage effect data (Fig.3.6). These families flower after producing very few (<15) leaves without vernalization; as a result a small difference in leaf number appears to cause a large effect in percentage change. The standard error for the mean percentage effect values for these families is correspondingly large.

Families from the same location tended to show similar flowering responses. However, there were obvious exceptions. For example family 11A6 is significantly less vernalization sensitive than other families from location 11 (Fig. 3.6). One explanation for this is that it reflects differences in the relatedness of families from a given location. Figure 3.7 shows a neighbour-joining tree produced by Poay Lim and Andrew Hudson, showing the genetic similarities of different families estimated from amplified fragment length polymorphisms (AFLP). Population 11A6 is more distantly related to all other families from location 11 which are all quite closely related to each other. Relatedness would go some way to explain variation within locations; however this does not fit in all instances. From Figure 3.7 it is clear that families from location 8 are not as closely related to each other as are families from other locations. However families 8A4 and 8A6 are considered to be genetically very similar and yet show very different vernalization responses. This could be explained by polymorphisms in genes involved in the vernalization response within location 8.

As was seen for the effect of photoperiod, families from location 10 are extremely late flowering; they are also some of the least vernalization sensitive families with an average vernalization response of 20% reduction in RLF. It is possible that their relative insensitivity to both photoperiod and vernalization contributes to their extreme lateness.

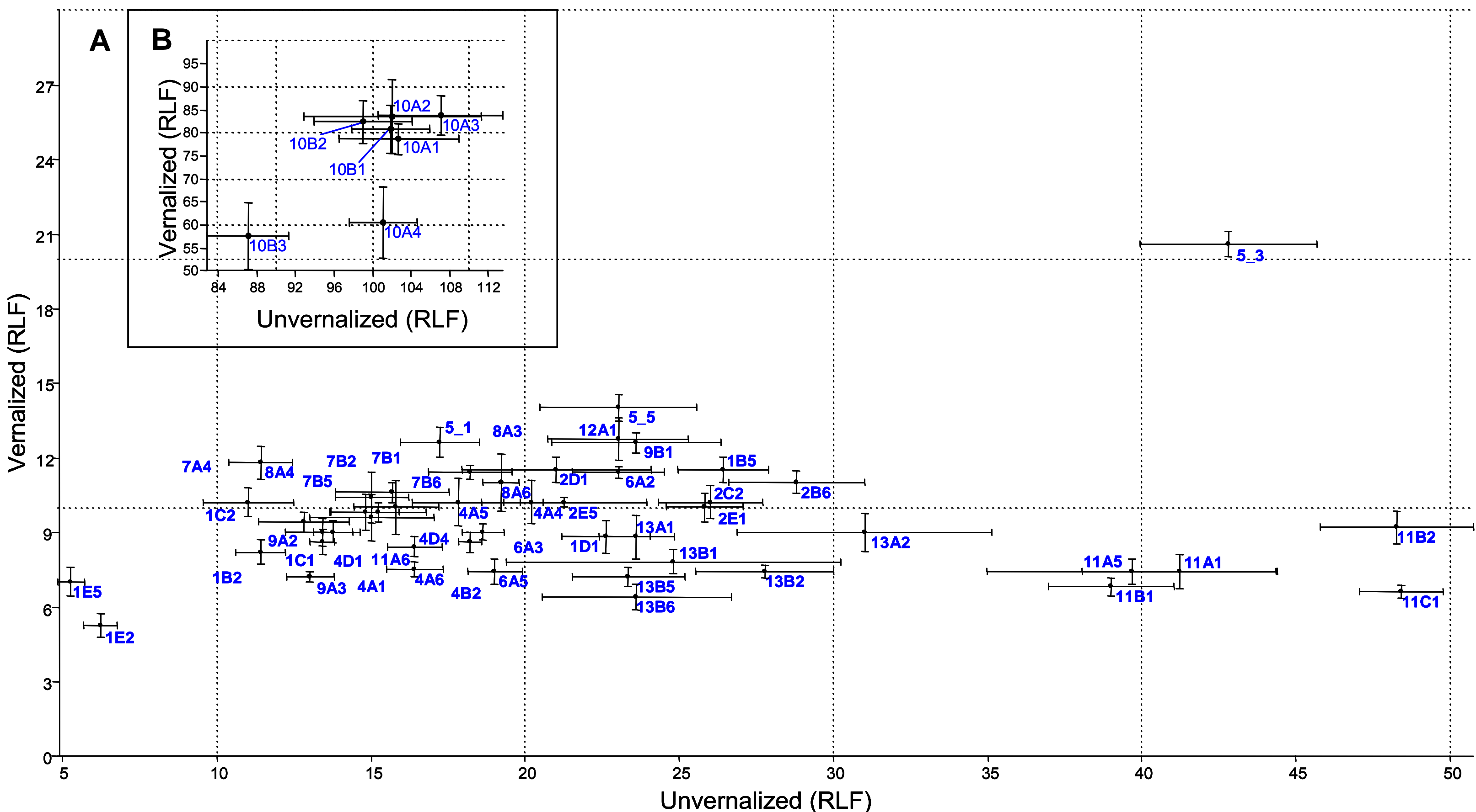


Figure 3.5 Flowering times if winter annual (WA) familes under vernalized and unvernalized conditions. Flowering times are given in rosette leaves until flowering (RLF) and all winter annual populations are shown in (A) with the exception of extremely late flowering families from 10 which are in inset (B) for clarity. Values are mean for ~X members of each family under each condition ± their standard errors.

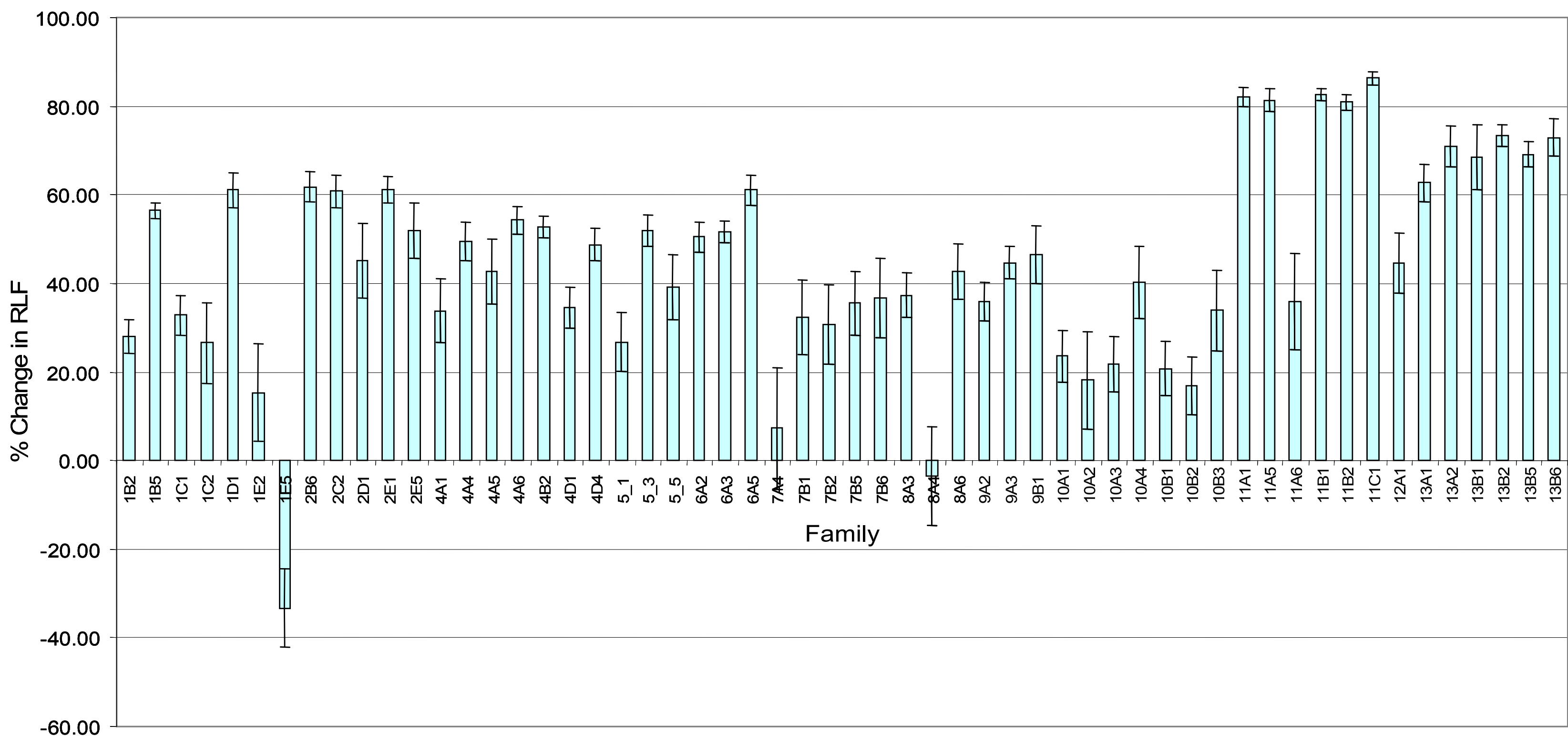


Figure 3.6 Percentage effect of vernalization on flowering time in winter annual (WA) families. Mean percentage change in RLF after 6 week vernalization are shown for ~ members under each condition ± their standard errors.

3.5 Response to vernalization in local summer annuals

Significant variation in flowering time was also observed locally for the summer annual families. As with the winter annuals they showed greater variation in their response to vernalization than to photoperiod. Heritability values were similar to the winter annuals: 81% and 80% in unvernalized and vernalized conditions, respectively.

On average the summer annuals were less responsive to vernalization than the winter annuals (38% compared to 49%). There were also more locations with families that showed no response. Of the 16 summer annual locations families from five of them appear to be insensitive to vernalization compared to three locations for the winter annuals. As discussed previously, direct comparisons cannot be drawn between the winter and summer annual populations because of variation in sampling; however it is not unexpected that more summer annuals are early flowering and vernalization insensitive than winter annuals. It is also unremarkable that overall the summer annuals do not have as strong a response to vernalization as the winter annuals; although given the expectation that summer annual populations rapid cycle without the requirement for vernalization it is perhaps surprising that there are not more early flowering, vernalization insensitive families among the summer annual population.

3.6 Relationship between distance of collection site and difference in flowering time

The difference between vernalization sensitivity of the summer and winter annual populations was consistent with adaptation to different seasons. However, considerable variation had been found within each population. One explanation was that this resulted from adaptation to environmental factors that varied between the collection sites. Because collection sites that are further apart might be expected to have more different

Figure 3.7 A tree showing genetic relatedness of winter annual (WA) populations. Families were genotyped for ~256 amplified fragment length polymorphisms (AFLP). Jaccard distances between families were used to construct a neighbour-joining tree. Note that this tree is not rooted Families are colour coded according to the location from which they were collected. The tree was produced by Poay Lim and Andrew Hudson

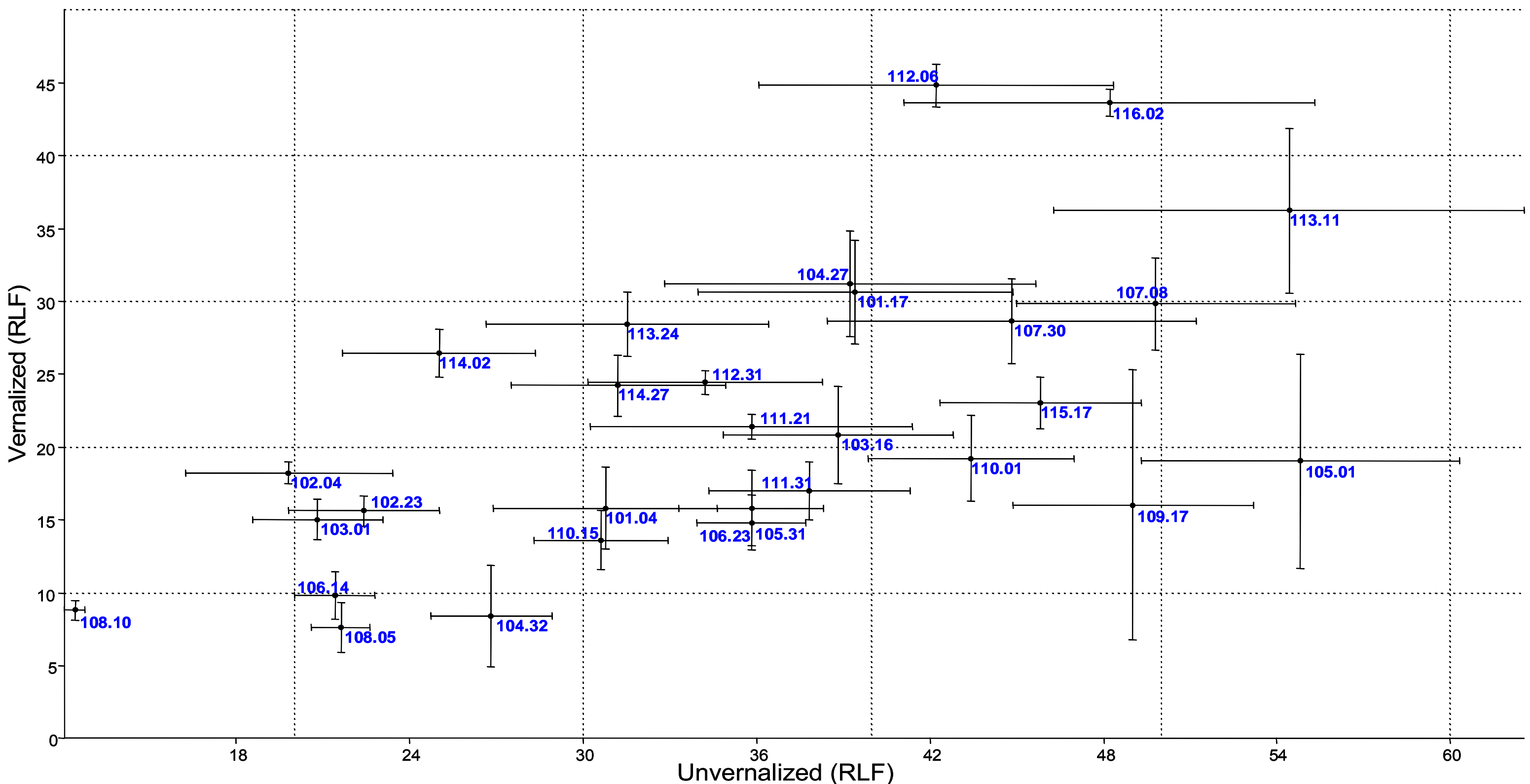


Figure 3.8 Flowering times of summer annual (SA) families under vernalized and unvernalized conditions. Flowering times are given in rosette leaves until flowering (RLF). Flowering times in vernalized conditions were recorded after a 6 week vernalization period. Values are means of ~X plants under each condition ± standard errors.

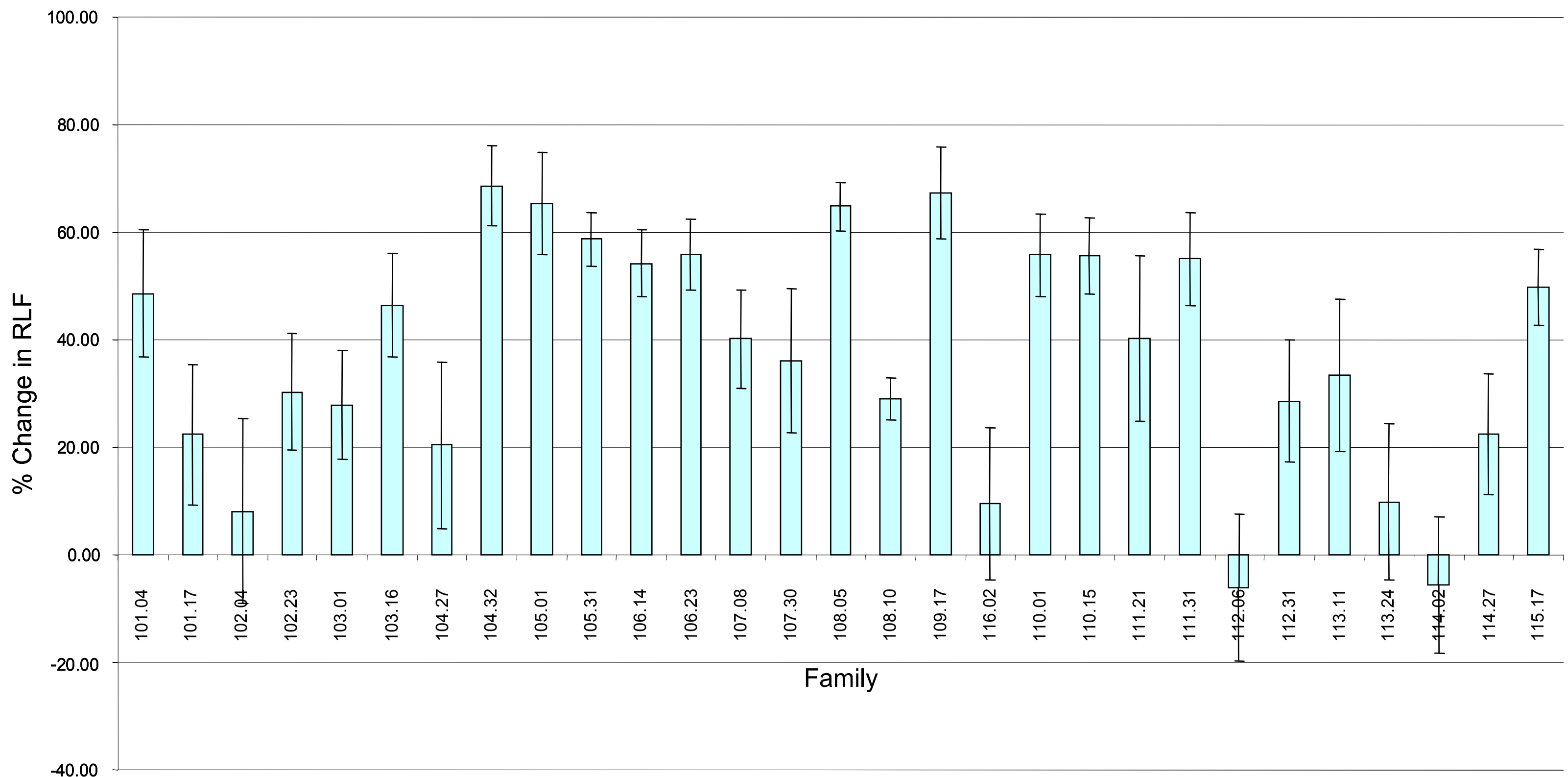


Figure 3.9 Percentage effect of vernalization on flowering time in summer annual (SA) families. Percentage change in RLF after six weeks vernalization. Values are means of ~5 plants under each condition ± standard errors.

vironments, this might result in a correlation between the geographic distance between collection sites and the differences in flowering behaviour of families from the sites.

A Mantel test was therefore used to assess if there was a correlation between geographic distance and difference in flowering behaviour. This was done for all of the families together and for the winter and summer annual populations separately. Flowering times for plants grown under three different conditions were used, SD after 6 weeks vernalization, LD after 6 weeks vernalization and LD without any vernalization.

Across all populations used in this a study there were significant but weak correlations under all environmental conditions between geographic distance between collection sites and differences in flowering time (Table. 3.1). However for summer annual populations alone there was no significant correlation for flowering in LD or SD vernalized conditions ($p = 0.06$ and 0.19, respectively) but a significant correlation under unvernalized conditions. A significant correlation was observed under all conditions for winter annual populations.

However both summer and winter annual populations contained a location that was far away from all others and contained families that exhibited extreme flowering behaviour – site 10 for the winter annuals and site 108 for the summer annuals. To test whether the significant correlations were dependent on these extreme sites, they were removed from the data and the Mantel tests were performed again.

A significant correlation was still observed for the combined populations under all conditions. However when looking at the summer and winter annual collections separately, correlations were only seen under unvernalized conditions.

A significant positive correlation implies that flowering time is more similar between populations found closer together. This is consistent with the idea that families are adapted to their local environment and therefore behave in a similar way to plants

Table 3.1 Correlations between geographic distance between collection sites and differences in flowering times Mantel tests were performed using data for summer annual and winter annual populations together or individually. A LD photoperiod was used for the vernalized and unvernalized experiments shown here. Test are shown with, or without winter annual families from site 10 or summer annual families from site 108, to test the dependence of correlations on these extreme families.

A. Winter Annual Populations

With location 10	**r**	**p (uncorr)**
Unvernalized	0.53	0
Vernalized	0.43	0
Short Day (vernalized)	0.32	0
Without Location 10	**r**	**p (uncorr)**
Unvernalized	0.43	0
Vernalized	0.03	0.3
Short Day (vernalized)	0.17	0.02

B. Summer Annual Populations

With Location 108	**r**	**p (uncorr)**
Unvernalized	0.32	0.01
Vernalized	0.20	0.06
Short Day (vernalized)	0.11	0.19
Without Location 108	**r**	**p (uncorr)**
Unvernalized	0.143	0.051
Vernalized	-0.03	0.57
Short Day (vernalized)	-0.02	0.54

C. Winter and Summer Annual Populations

All families	r	p (uncorr)
Unvernalized	0.20	0
Vernalized	0.41	0
Short Day (vernalized)	0.16	0
Without 10 and 108	**r**	**p (uncorr)**
Unvernalized	0.34	0
Vernalized	0.46	0
Short Day (vernalized)	0.23	0

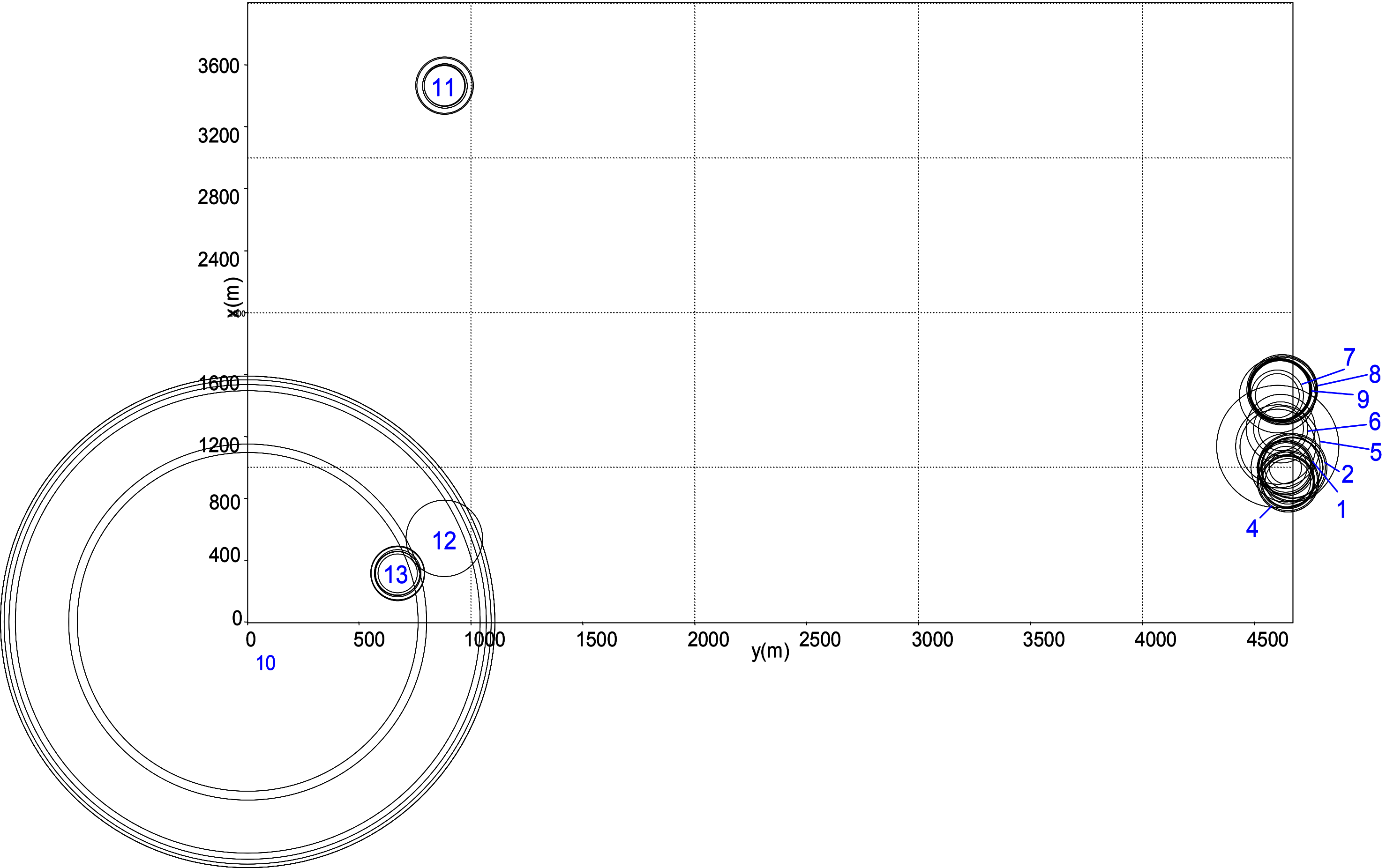


Figure 3.10 Relationship between flowering time and geographical position of winter annual collection site. The size of each bubble represents the relative flowering time of populations, the most south-westerly collection sit was selected as the origin and x(m) refers to distance in metres north of this population and y(m) refers to distance east in metres from this population.

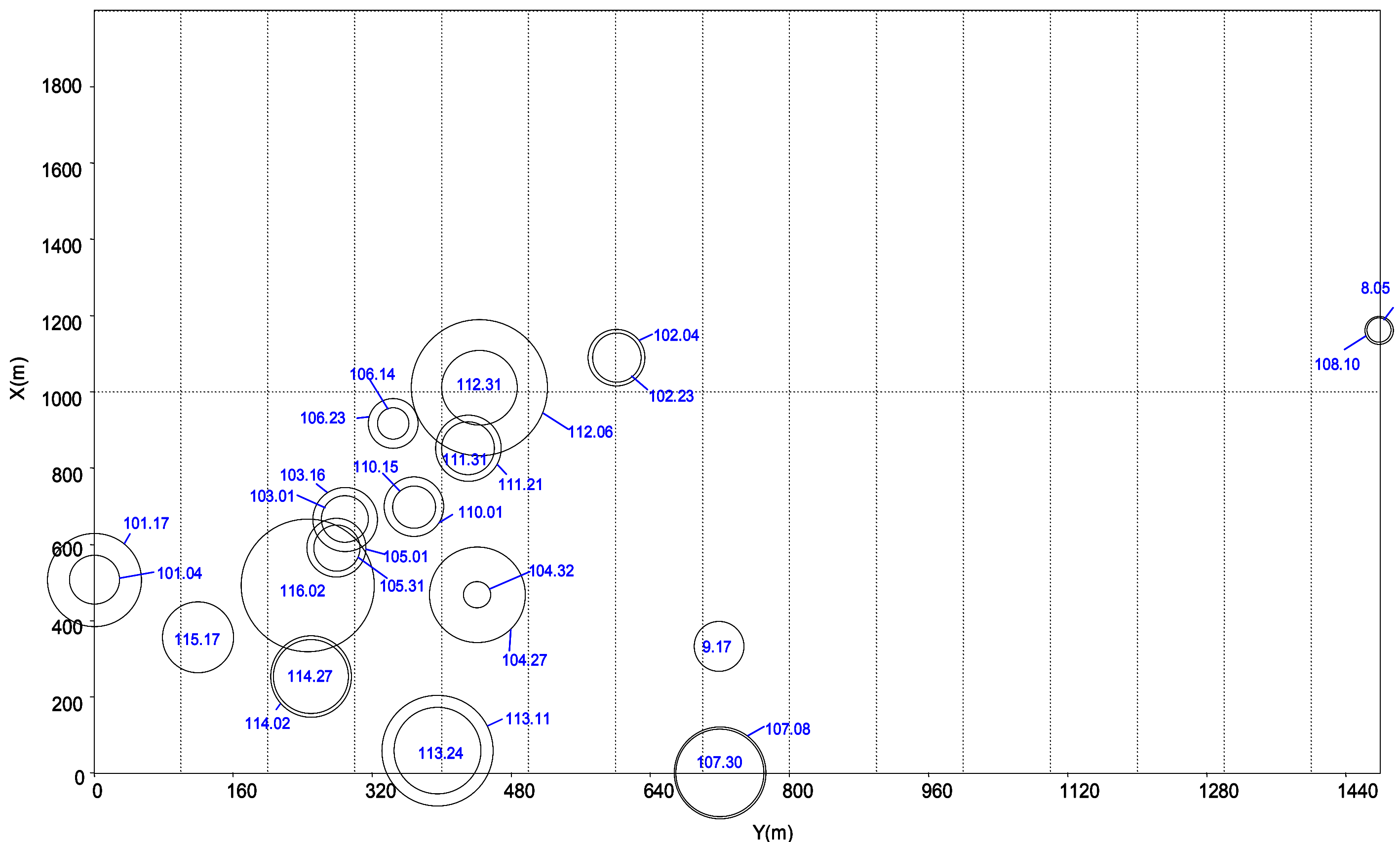


Figure 3.11 Relationship between flowering time and geographical position of summer annual collection site. The size of each bubble represents the relative flowering time of populations, the most south-westerly collection site from both winter and summer annual populations together was selected as the origin and x(m) refers to distance in metres north of this population and y(m) refers to distance east in metres from this population.

collected from a similar environment nearby. An alternative explanation is that the relationships between flowering time and geographic distance results from restricted gene flow, either through pollen dispersal or seed migration, so that plants found geographically closer to each other share more similar flowering behaviour because they are more closely related. However in the winter annual population, similarity of flowering time did not always reflect genetic relatedness of families (Fig.3.7) and so this information is insufficient to favour either adaptation or dispersal as an explanation.

3.7 Discussion

For both summer and winter annual populations there was considerable genetically determined variation in flowering time and in responses to photoperiod and vernalization.

Flowering time varied considerably between families from the earliest family 1E5 with an average of 5.25 RLF, to the latest flowering family 10A3 which flowered with an average of 107 RLF. This demonstrates how extensive the variation is over such a small geographic region as these families are both winter annuals that were collected within 5 km of each other. The summer annuals showed a smaller range, 12.4 to 54.8 RLF. From this it is evident that variation in flowering time locally is on a similar scale to variation observed worldwide. Due to variation in sampling methods conclusions cannot be drawn about the differences in flowering behaviour between the summer and winter annual populations.

Winter annual families could be categorised as either strongly responsive or not responsive to photoperiod. More continuous variation was observed for the summer annuals compared to the winter annuals, although this may reflect the wider separation of collection sites for summer annual or variation in sampling or experimental methods. Photoperiod insensitive families were found in both summer and winter annual populations although they were rare. Families from location 10 were among the latest

flowering and found to be both relatively insensitive to both photoperiod and vernalization.

As discussed in chapter 1 naturally occurring variation has been identified in components of the photoperiod pathway. *PHYC* has been identified as significant cause for natural variation in flowering time in response to day length. Based on evidence from these previous studies *PHYC* variation in local accessions might account for the photoperiod insensitivity of a few families (Balasubramanian et al., 2006; Samis et al., 2008; Weinig et al., 2002).

More variation was found in local Scottish accessions in their response to vernalization, and families showed an almost continuous range of responses. Winter annuals responded more than summer annuals on average, but both populations overlapped considerably in their responses.

The extensive local variation in response to vernalization is consistent with results obtained with worldwide accessions. It has also been demonstrated many times that the floral regulators *FRI* and *FLC* make a major contribution to this variation, as will be discussed in the next chapter. It is also worth noting that *FRI* and *FLC* do not explain all of the variation in vernalization responses between worldwide accessions (Salome et al., 2011; Strange et al., 2011). It is therefore likely that other vernalization related genes also contribute to flowering time variation in the wild. If their alleles are rare, they might have escaped identification, because they have not ended up in a QTL analysis and are unable to produce a significant signal in a genome-wide association study.

A correlation was detected between geographic distance and flowering responses of local populations. This is consistent with adaptation to factors that vary locally, although no cline in flowering behaviour, for example with altitude, was detected, this does not rule out adaptation to a factor that does not vary clinally.

For differences in flowering time to be adaptive, they must be expressed in the field. It would be interesting to explore how the flowering time variation under controlled conditions that is discussed in this chapter relates to variation in the field. Flowering time is a highly plastic trait that responds to several environmental and seasonable variables, including ambient temperature and stress. For example, evidence has been found for a correlation between flowering time and drought avoidance (McKay et al., 2003). As such, the variation observed in controlled conditions is not necessarily a reflection of flowering behaviour under natural conditions. One field study examined the flowering behaviour of mutants in pathways that regulate flowering in response to environmental signals, i.e., mutants in the photoperiod and vernalization pathways. Field sites were selected that represented the native range of European climates for Arabidopsis and flowering time was studied in plants that were germinated in the spring, summer and autumn. The results suggested that Arabidopsis genotypes cannot be discretely divided into winter or summer annuals and instead that all genotypes were capable of behaving as both – for example that both *fri* mutants and FRI^{+} genotypes could flower in the same season or over winter as rosettes (Korves et al., 2007).

It is evident that there is a similar amount of variation in flowering time locally as has been observed worldwide. This raises the possibility that correlations observed world wide could be artefacts of sampling. Several studies have reported a correlation between flowering time and the geographic origins of accessions. For example clines in *FRI* allele frequencies and vernalization response have been detected with latitude. However other studies that have looked at worldwide accessions have failed to replicate this finding. In one such study it was found that on a gross scale there was a weak correlation between flowering time and latitude of collection site, for example all accessions collected above 62°N exhibited extremely late flowering while extensive variation was found at latitude 55°N, though no significant correlation between flowering time and latitude was detected overall (Shindo et al., 2005). Given the

extensive variation in environmental conditions at any given latitude a lack of correlation with flowering time is not unexpected and that correlation between flowering time and more specific environmental variables would be ecologically more relevant.

Heritability calculations suggest that approximately 80% of the variation observed between local accessions under any of the environmental conditions was due to genetics. This made it feasible to explore which loci are responsible. As most variation had been observed in response to vernalization, this forms the main focus of the work that is to follow.

4 Sequence and expression variation in *FRIGIDA* and *FLC*

4.1 Introduction

In the previous chapter I showed that there is considerable genetically determined variation in flowering time and vernalization response among local populations of Arabidopsis. The aim of the work described in this chapter was to analyse the contribution of *FRI* and *FLC* to this phenotypic variation, by comparing *FRI* and *FLC* genotypes in local populations and testing whether they are associated with differences in gene expression and flowering behaviour.

Several studies have demonstrated that *FRI* and *FLC* make a substantial contribution to natural variation in flowering time worldwide, with some studies estimating that *FRI* is responsible for as much as 70% of the variation observed under controlled conditions (Gazzani et al., 2003; Shindo et al., 2005). In particular extensive variation has been found at the *FRI* locus, mainly within the first exon that encodes most of the FRI protein. The two commonly used laboratory strains of Arabidopsis, Col-0 and L*er*, have different *fri* loss of function alleles that have also been identified in many other natural accessions throughout Europe (Johanson et al., 2000). In the case of Col-0 a 16 bp deletion at the end of exon 1 results in a frameshift that leads to the termination of the ORF shortly after the start of the second exon. The L*er* allele of *FRI* is disrupted at the beginning of the ORF by a combined 376 bp deletion and a 31 bp insertion resulting in removal of the start codon.

L*er* carries a naturally occurring weak allele of *FLC*, which carries an insertion of a ~1.2 kb transposable element at the 3' end of intron 1. The effect of this allele is suppressed by the *fri* mutation in L*er*, however, when a functional *FRI* allele was introgressed into L*er* its *FLC* allele resulted in early flowering (Michaels et al., 2003c; Sanda and Amasino, 1996). Transposable elements inserted into intron 1 have been identified in other weak *FLC* alleles (Michaels et al., 2003c).

The importance of intron 1 in regulation of *FLC* activity was confirmed by Sheldon *et al.* (2002), who showed that a 75 bp section of the *FLC* promoter and a large

proportion of intron 1 were sufficient and necessary for the expression of *FLC* in non vernalized plants and for the repression of *FLC* in response to long periods of cold and the maintenance of repression. Intron 1 is also transcribed to produce the regulatory non-coding *COOLAIR* transcript (Heo and Sung, 2011)

In addition to examining the structure of *FRI* and *FLC* alleles in local populations, a further aim of the work in this chapter was to examine the expression of the alleles in relation to flowering behaviour. For example whether lack of *FRI* expression correlated with early flowering in the absence of vernalization and resulted in low *FLC* expression, as is seen for Col-0 and L*er*.

4.2 Sequence variation in *FRIGIDA*

To first assess the presence of known mutations in *FRI,* primers *FRI UJ26*/*UJ34* and *FRI promoter F/R* (Appendix A) were used in polymerase chain reactions (PCR) to amplify regions of *FRI* that included the deletions in Col-0 and L*er FRI* alleles. All the local accessions were tested using these primers and the products analysed using agarose gel electrophoresis. Deletions corresponding to the Col-0 or L*er* alleles were not identified in any of the local accessions.

The next step in analysing *FRI* variation was to amplify and sequence most of the first exon of *FRI* from a subset of local populations representing accessions collected in both winter and summer using the primers *FRI coding F*/*R* (Appendix A). Initially DNA from a single plant representing one family from each location was sequenced along both the forward and reverse strands. In instances where variation of interest was found a maximum of two further biological replicates were performed, for this DNA from two further individual plants from the same family was sequenced in the same way. Additionally individual plants from other families from the same locations were also sequenced.

The first exon was chosen because it is known to vary extensively between worldwide accessions and contains most of the FRI protein coding sequence. The summer

accessions were sequenced with the help of an undergraduate honours student, Lucy Davie.

Overall eight single nucleotide polymorphisms (SNPs) were found. Four were non-synonymous substitutions, three synonymous substitutions and one a deletion of a single nucleotide. These SNPs were present in five different *FRI* alleles within the populations (Fig 4.1).

A summary of the *FRI* genotypes and flowering phenotypes for the families is shown in Table 4.1. Of the five alleles, the two most common; *FRI* I and *FRI* II, were found in both summer and winter annual populations. There was no obvious correlation between flowering time or vernalization response and these *FRI* alleles; for example the winter annual families 2E1 and 7B5 both carry *FRI* II and flower with 27 or 15 leaves respectively under unvernalized conditions and have a vernalization response involving reductions of 62% or 38% in leaf number (Chapter 3).

These sequences have been observed previously in worldwide accessions (Table 1.4). The *FRI* I sequence was identical to five previously characterised accessions, three of which (Ri-0, M7884S and HR15) carried loss of function *fri* alleles due to mutations in other regions of the gene, whereas the other two (Gr-3 and Bd-0) had been shown to possess functional copies of *FRI* (Lempe et al., 2005). Given the variation in flowering time observed within the local *FRI* I group it is possible that further relevant variation exists in regions that were not sequenced in this study and that more than one allele is represented by the *FRI* I sequence. Alternatively, the variation in vernalization response observed between local accessions could be due to genes other than *FRI*.

The *FRI* II allele had also been seen previously in three other accessions (Table 4.1) that carried both functional (Rsch-0) or non functional (An-1 and Pog-0) alleles of *FRI* (Lempe et al., 2005; Werner et al., 2005). Because local accessions carrying the *FRI* II allele differ in flowering behaviour, similar sequence variation elsewhere in *FRI* II alleles might contribute to differences in flowering behaviour.

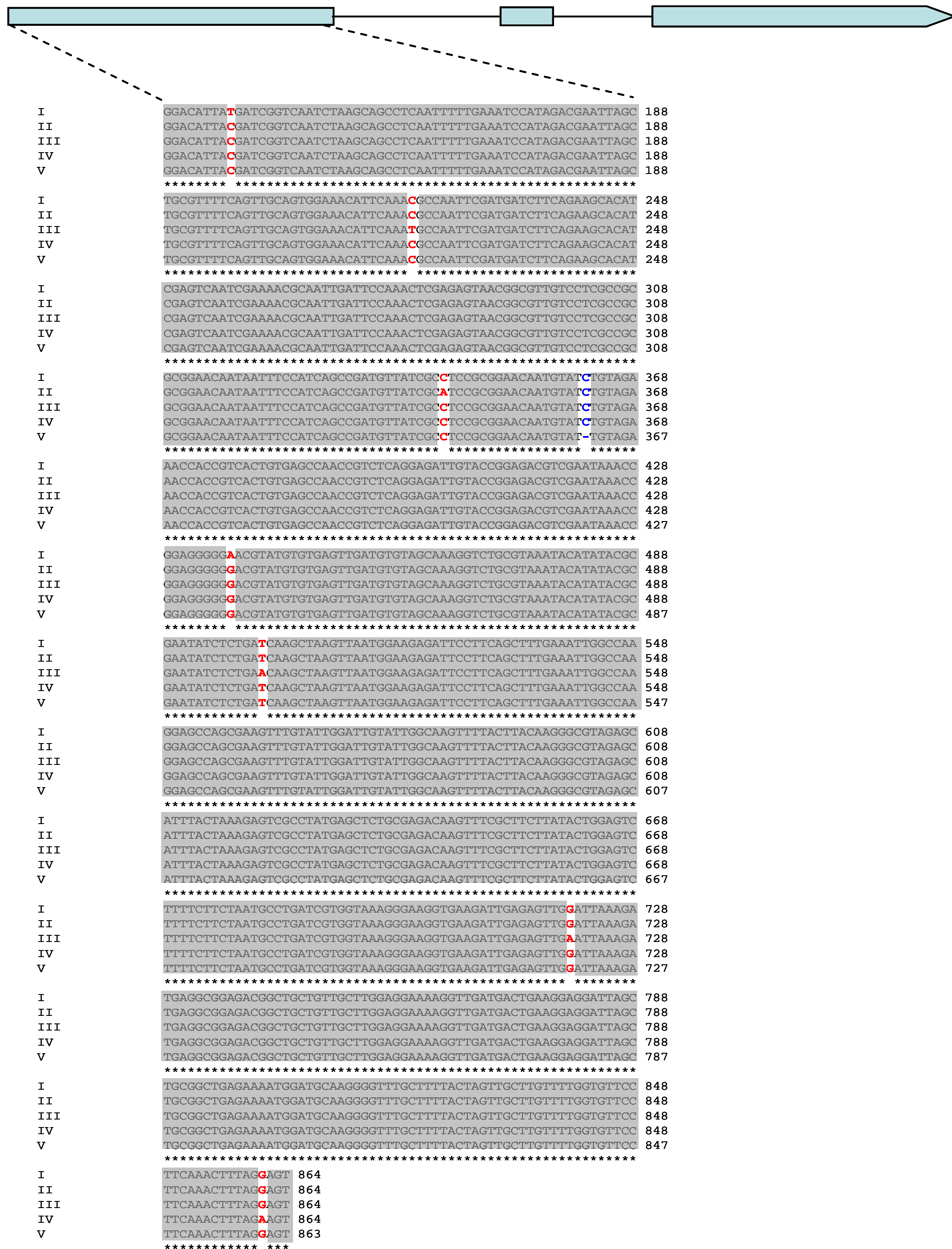


Figure 4.1. Sequence alignment of *FRI* alleles identified in local populations. Idenitcal sequences are highlighted in grey, SNPs are shown in red and deletions indicated in blue. Numbers indicate nucleotide position where the A of the ATG start codon is counted as 1. The structure of the *FRI* gene is shown above.

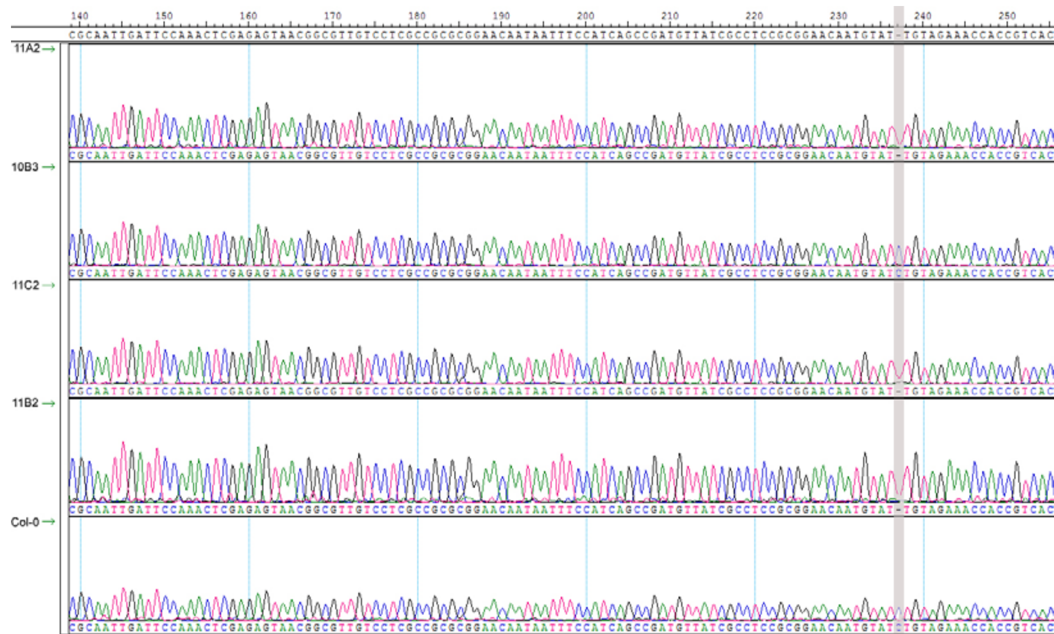


Figure 4.2 Sequence alignment and trace for *FRI* in local accessions and Col-0. Sequence data for three families from location 11 compared to a family from location 10 and the commonly used lab strain Col-0. In each case DNA was taken single plant from each family. This result was confirmed using biological and technical replicates (see main text for details). The SNP highlighted in grey was only founding families from location 11 and was found to result in a premature stop codon.

```
I               DIMIGQSKQPQFLKSIDELAAFSVAVETFKRQFDDLQKHIESIENAIDSKLESNGVVLAA 104
II              DITIGQSKQPQFLKSIDELAAFSVAVETFKRQFDDLQKHIESIENAIDSKLESNGVVLAA 104
III             DITIGQSKQPQFLKSIDELAAFSVAVETFKCQFDDLQKHIESIENAIDSKLESNGVVLAA 104
IV              DITIGQSKQPQFLKSIDELAAFSVAVETFKRQFDDLQKHIESIENAIDSKLESNGVVLAA 104
V               DITIGQSKQPQFLKSIDELAAFSVAVETFKRQFDDLQKHIESIENAIDSKLESNGVVLAA 104
                ** *************************** *****************************

I               RNNNFHQPMLSPPRNNVSVETTVTVSQPSQEIVPETSNKPEGERMCELMCSKGLRKYIYA 164
II              RNNNFHQPMLSHPRNNVSVETTVTVSQPSQEIVPETSNKPEGGRMCELMCSKGLRKYIYA 164
III             RNNNFHQPMLSPPRNNVSVETTVTVSQPSQEIVPETSNKPEGGRMCELMCSKGLRKYIYA 164
IV              RNNNFHQPMLSPPRNNVSVETTVTVSQPSQEIVPETSNKPEGGRMCELMCSKGLRKYIYA 164
V               RNNNFHQPMLSPPRNNVL------------------------------------------ 122
                *********** *****

I               NISDQAKLMEEIPSALKLAKEPAKFVLDCIGKFYLQGRRAFTKESPMSSARQVSLLILES 224
II              NISDQAKLMEEIPSALKLAKEPAKFVLDCIGKFYLQGRRAFTKESPMSSARQVSLLILES 224
III             NISEQAKLMEEIPSALKLAKEPAKFVLDCIGKFYLQGRRAFTKESPMSSARQVSLLILES 224
IV              NISDQAKLMEEIPSALKLAKEPAKFVLDCIGKFYLQGRRAFTKESPMSSARQVSLLILES 224
V               ------------------------------------------------------------

I               FLLMPDRGKGKVKIESWIKDEAETAAVAWRKRLMTEGGLAAAEKMDARGLLLLVACFGVP 284
II              FLLMPDRGKGKVKIESWIKDEAETAAVAWRKRLMTEGGLAAAEKMDARGLLLLVACFGVP 284
III             FLLMPDRGKGKVKIES-------------------------------------------- 240
IV              FLLMPDRGKGKVKIESWIKDEAETAAVAWRKRLMTEGGLAAAEKMDARGLLLLVACFGVP 284
V               ------------------------------------------------------------

I               SNFRS 289
II              SNFRS 289
III             -----
IV              SNFRS 289
V               -----
```

Figure 4.3 Alignment of FRI amino acid sequences identified in local populations. The sequences encoded by different *FRI* alleles from local populations are shown. Identical amino acids are highlighted in grey, substitutions are shown with the more common amino acid in yellow and the less common in pink.

Table 4.1 the distribution of alleles identified in this study locally and in worldwide accessions (refs). The country of origin for each worldwide accession is shown in brackets next to the name of the accession. Whether the local accessions are from the summer annual (SA) or winter annual (WA) collected populations is indicated next to the name of each local family. Where the variation was considered of particular interest then more than one family was sequenced from a given a location.

Allele	Local Accessions	Worldwide Accessions
I	114-02 (SA) 107-08 (SA) 1E5 (WA) 4A1 (WA)	Ri-0 (Canada) M7884S (US) HR15 (UK) Gr-3 (Austria) Bd-0 (Germany)
II	105-31 (SA) 7A4 (WA) 2E1	An-1 (Belgium) Rsch-0 (Russia) Pog-0 (Canada)
III	108-10 (SA)	None
IV	10B3 (WA)	None*
V	11A2 (WA) 11B2 (WA) 11C1(WA) 11C2 (WA)	None

*There were no identical sequences to this allele found in worldwide accessions however the SNP by which this allele differed from the others was found to result in a synonymous substation.

FRI IV was present only in the winter annuals from location 10. Although no previously identified *FRI* alleles were found to identical to *FRI* IV, sequences that differed by only one nucleotide (position 861), representing a synonomous substitution, were found in three other accessions.

FRI III was found only in the summer annual family 108-10. A G to A nucleotide substitution at position 720 results in a premature stop codon, shortening the FRI protein to only 196 amino acids and causing a potential loss of function. This disruption in *FRI* is a likely explanation for the early flowering and vernalization insensitivity of family 108-10.

FRI V was present in all families from location 11 and not found in any other local accessions or previously sequenced worldwide accessions. It carries a deletion of nucleotide 361 resulting in a frameshift and subsequent premature stop codon resulting in a truncated protein sequence of only 80 amino acids (Fig. 4.2). This allele is likely to be non-functional, which is surprising given that families from location 11 were collected in winter and were the most vernalization sensitive of all those used in this study. A possible explanation for this could be that these families also carry a mutation in an autonomous pathway gene. It has been shown previously that mutations in autonomous pathway components result in increased FLC expression even in the absence of functional FRI expression (REF). Additionally a natural variant of FY was identified in the Spanish accession Bla-6

4.3 Sequence variation in *FLOWERING LOCUS C*

Variation in *FLC* has also been shown to contribute to natural variation in flowering time. Unlike *FRI*, much of the relevant natural variation in *FLC* is present in the first intron, which is involved in epigenetic regulation in response to vernalization (Sheldon et al., 2002).

The common lab strain L*er* is known to contain a weak *FLC* allele caused by a 1.2 kb insertion in intron 1 (Michaels et al., 2003c). I initially screened for this insertion

using primers that would amplify across the insertion site (Appendix A). The insertion was not found in any of the populations used in this study.

I then amplified and sequenced a region of intron 1 to examine the diversity present in local populations. Sequencing of summer annuals was done with the help of undergraduate honours student, Jinty Urquhart. Significant sequence variation was found within local populations (Fig. 4.3). Seven polymorphisms were detected, consisting of five single nucleotide substitutions, two single nucleotide insertions or deletions (indels) and four longer indels. These were present in four different alleles, of which three were unique to single locations.

The most common allele identified; *FLC* I, was found in both summer and winter annuals (Table 4.2). This sequence was also found in ten worldwide accessions. Given the wide range of flowering phenotypes in accessions carrying this sequence, in this and previous studies, this polymorphism appeared unlikely to have a major effect in flowering time variation.

FLC II was only found at location 2 in the winter annual collection. It had not been found in previous studies, however it differed by a single nucleotide substitution from *FLC* I and therefore from a number of worldwide accessions.

FLC III was only found in the summer annual family 111-31. It carried a 5 bp deletion and a single nucleotide substitution relative to the other *FLC* alleles. It was identical to the sequence from the French accession Cen-0 (Lempe et al., 2005) had shown that crossing Cen-0 to an FLC^{+} genotype resulted in later flowering in the F1 generation, suggesting that Cen-0 might carry a weak *FLC* allele. Because only a small region of *FLC* in 111-31 was sequenced and found identical to Cen-0, more information would be needed to reveal whether the alleles were identical.

The fourth *FLC* allele to be identified was only found in the winter annual families from location 10. *FLC* IV contains two point mutations, two insertions (one of 2 bp the other of 3 bp) and an 11 bp deletion relative to other local accessions. Plants from

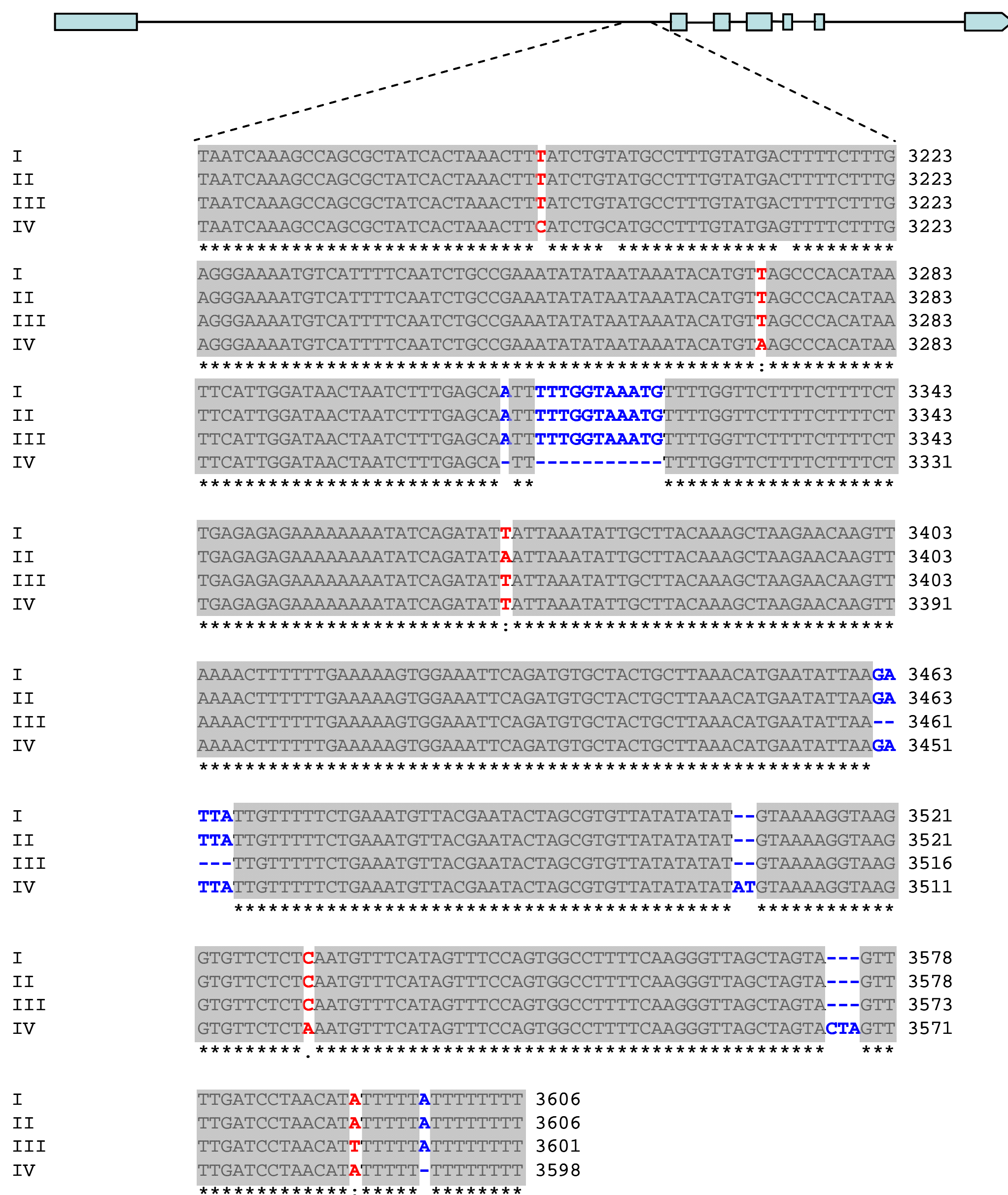


Figure 4.4 Sequence alignment of FLC alleles identified in local populations. Conserved sequences are highlighted in grey, SNPs are shown in red and insertions or deletions indicated in blue. Above is the structure of the FLC gene with exons represented as blocks and introns shown as lines.

Table 4.1 the distribution of *FLC* alleles identified in this study locally and in worldwide accessions (refs). The country of origin for each worldwide accession is shown in brackets next to the name of the accession. Whether the local accessions are from the summer annual (SA) or winter annual (WA) collected populations is indicated next to the name of each local family. Where the variation was considered of particular interest then more than one family was sequenced from a given a location.

Allele	Local Accessions	Worldwide Accessions
I	114-02 107-08 105-31 108-10 1E5 4A1 5_1 7A4	Col-0 (USA) Cvi-0 (Cape Verdi) PHW-35 (France) Co-1 (Portugal) Chi-1 (Russia) Bla-1 (Spain) Da(1)-12 (Czech Republic) Ll-2 (Spain) Kin-0 (USA) Gr-3 (Austria)
II	2E1	None
III	111-31	Cen-0 (France)
IV	10B1 10B2 10B3 10A1 10A2	None

location 10 were extremely late flowering under all conditions and the least vernalization responsive local accession. No other worldwide accession has been reported to carry any of the polymorphisms found in this sequence.

4.4 Flowering time in response to vernalization

From the sequence variation found in *FRI* and *FLC*, the *FLC* IV allele appeared to be an interesting candidate for further study. This allele was only found in the late flowering families from location 10 and carried three polymorphisms in a region of intron 1 that had previously been shown to be crucial for the expression of *FLC* and its stable repression during vernalization (Sheldon et al., 2002). It was therefore hypothesised that the polymorphisms might be responsible for the late flowering observed in accessions from location 10. These had been found to flower considerably later than all other accessions after an eight week vernalization treatment, with only a 20% reduction in leaf number - less than half the average response for the winter annual population (41%). One explanation is that eight weeks of vernalization is not sufficient for maximal stable repression of the *FLC* IV. To test this, an experiment was designed to analyse the vernalization requirement of an accession from location 10, compared to other families with different *FLC* alleles.

Families were chosen for comparison based on previous flowering behaviour (Chapter 3) and genotype. 10B2 was selected as an example from the late flowering population 10, which carried a deletion in the first intron of *FLC*, 1E2 was very early flowering and showed little response to vernalization, despite no obvious disruptions in the regions of *FLC* or *FRI* that had been sequenced, 11A2 had shown a strong vernalization response, but carried a premature stop codon in the protein coding region of *FRI* that was predicted to cause a loss of function *fri* mutation, and 5_3 had intermediate flowering time and vernalization responses and carried *FRI* and *FLC* alleles that were common locally. As controls the lab strains Col-0 (*fri*) and Col-0 introgressed with an active *FRI* allele (FRI^{+}) were used.

Seedlings were germinated on MS plates, transferred to soil and kept at 5°C under LD conditions. After different lengths of time in the cold, plants were transferred to

22°C LD conditions until flowering. Tissue for RNA extraction was collected immediately after removal from the cold and leaf counts were made weekly until each plant had flowered.

Figure 4.4 shows the flowering times observed. The control genotypes behaved as expected. Col-0 (*fri*) flowered early, producing an average of 12 leaves, and vernalization had no significant effect on its flowering time. In contrast, the presence of an active *FRI* allele was sufficient to restore later flowering in the absence of vernalization (83 leaves) and a response to cold. The vernalization response of Col-0 *FRI*$^{+}$ appeared to be saturated by 6 weeks because two additional weeks of cold had no further effect on flowering time.

Family 1E2 showed a very similar behaviour to that of Col-0 (*fri*); it flowered early (16 leaves) regardless of vernalization treatment. This lack of sensitivity to vernalization was consistent with a reduction in *FRI* activity. However no candidate loss of function mutation had been found in the coding region of *FRI* in 1E2, suggesting that this family carried a disruption elsewhere at the *FRI* locus or that other genes, including *FLC*, might be involved.

Family 5_3 took approximately eight weeks for the vernalization response to be saturated with a gradual but consistent reduction in flowering time after each increment of cold treatment. Reduction of flowering time in response to vernalization occurred faster than in the control line Col-0 *FRI*$^{+}$, although both carried similar *FRI* sequences. This suggested that the fast response of 5_3 was due to variation in genes other than *FRI*.

As expected, 10B2 flowered late in the absence of vernalization (with 130 leaves), and responded to cold. It took about 10 weeks for the vernalization response to be saturated; at least two weeks longer than any of the other families. This is consistent with the relatively late flowering previously observed for this family after only eight weeks of vernalization. It is also consistent with the hypothesis that the disruption to

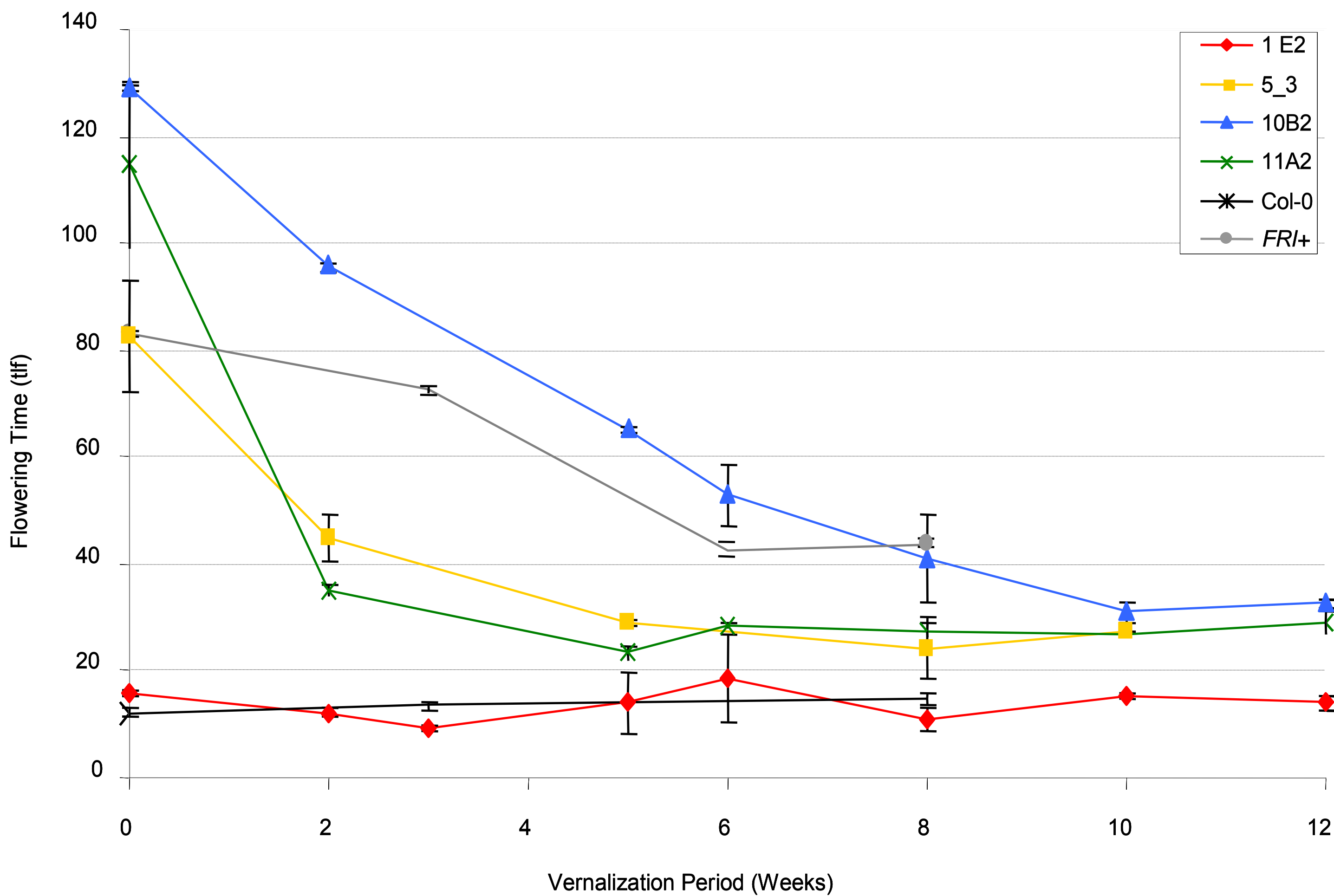


Figure 4.5 Relationship between flowering time and length of vernalization period. Flowering times were measured in total leaves to flower (tlf) for local accessions and two controls (Col-0 and Col-0 *FRI*$^+$) after increasing periods of vernalization. Each value is the means of five plants ± standard error.

intron 1 affects the stable silencing of the *FLC* allele in 10B2, giving this family the greatest requirement for vernalization.

The behaviour of family 11A2 was consistent with previous findings in that this family flowered very late without vernalization (115 leaves) and showed a very strong response to vernalization. This response occurred very quickly, saturating between two and five weeks of cold with a 70% reduction in leaves to flower. This response was surprising given that 11A2 carries a potential *fri* loss of function allele which was expected to cause early, vernalization-insensitive flowering.

4.5 Variation in expression of *FRIGIDA* and *FLOWERING LOCUS C*

To further investigate the potential effects of variation in *FRI* and *FLC* on the flowering responses observed in response to different periods of vernalization, I analysed the expression of these genes using reverse transcription and quantitative PCR (qRT-PCR; Fig. 4.5 and 4.6).

FRI expression was estimated in plants that had not been vernalized (0 weeks in the time-course experiment). Family 11A2 showed no significant expression of *FRI* (Fig. 4.5), suggesting that the premature stop codon within *FRI* V caused nonsense-mediated decay of the transcript (Isken and Maquat, 2008).

Despite the lack of detectable *FRI* expression, 11A2 was found to express *FLC* to a detectable level (Fig. 4.6). The *FLC* level in 11A2 without vernalization was about one quarter of that in the control line Col-0, which has relatively low *FLC* expression due to its null *fri* allele (Fig. 5.3; (Johanson et al., 2000). The level of *FLC* in 5_3 also declined in response to vernalization. However the progressive drop in *FLC* expression in 5_3 did not correlate with the effects of vernalization on flowering time, in which most of the decrease occurred in the first two weeks of cold. One explanation for this is that two weeks of vernalization are sufficient to stably repress *FLC* in 11A2, but not the other vernalization responsive accessions, and that further reduction in *FLC* transcript levels has no additional effect on flowering time.

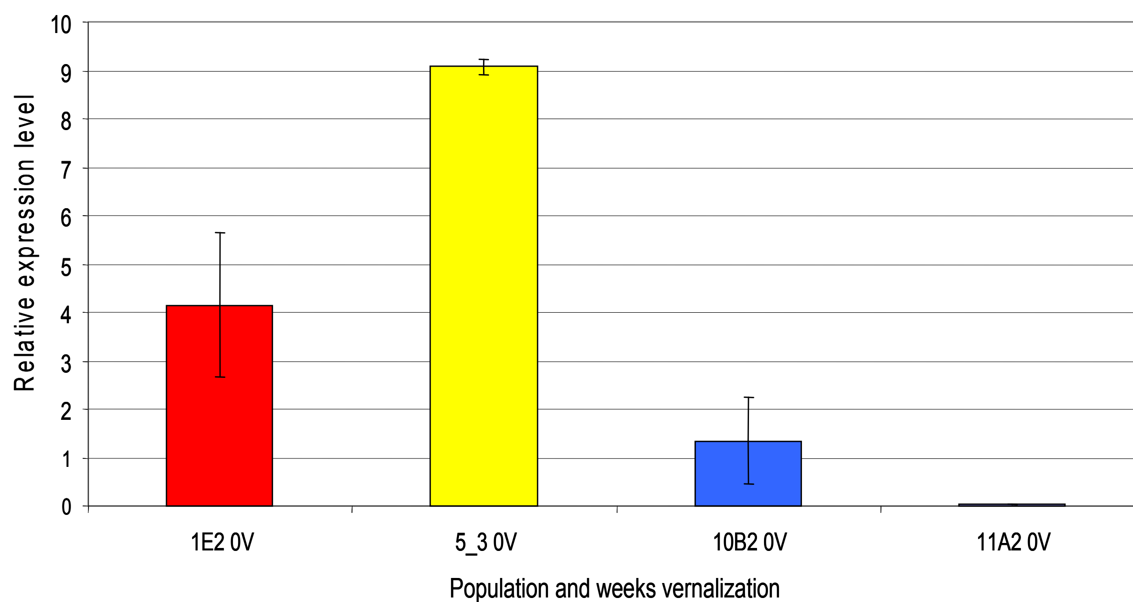


Figure 4.6 Relative expression of *FRI*. *FRI* transcript levels for four local accessions without cold treatment were assessed by qRT-PCR. Accessions are colour coded as in Figure 5.1. Mean values of three biological replicates are plotted, ± their standard-errors.

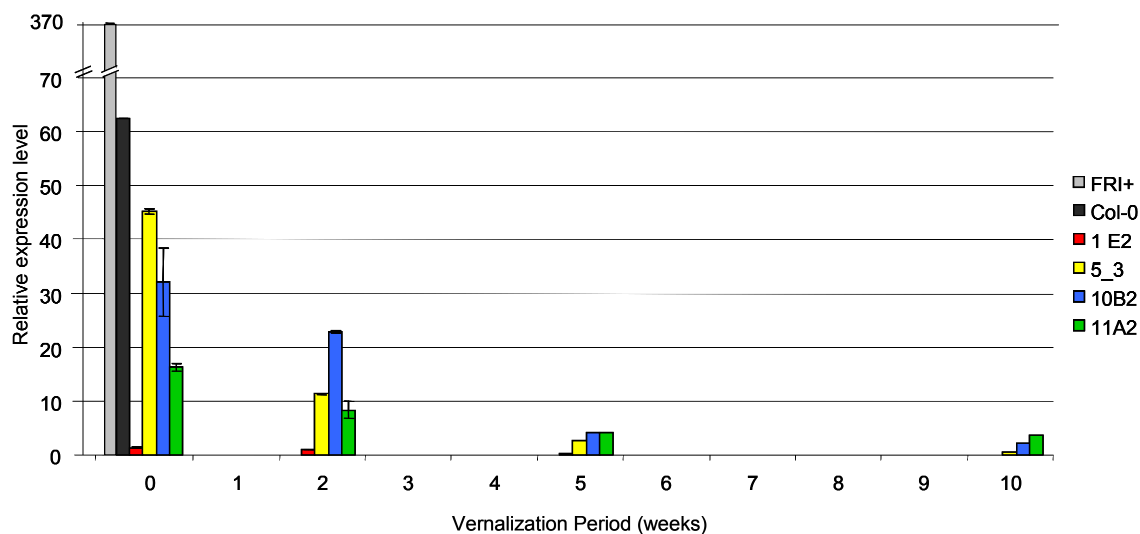


Figure 4.7 Relative expression of *FLC*. *FLC* transcript levels were estimated in four local accessions and Col-0 controls by qRT-PCR. The mean values for three biological replicates ± their standard-errors are shown. Accessions are coloured as in Figures 4.4 and 4.5

However there is still the question of how 11A2, a line with no detectable *FRI* expression, can exhibit such a strong vernalization response.

It was hypothesised that the *FLC* allele in accessions from location 10 contributed to the late flowering of this family. However expression of *FLC* in 10B2 without vernalization was lower than in the early flowering Col-0 (*fri*) control, and therefore *FLC* levels alone cannot explain the late flowering in 10B2. The relatively low levels of *FLC* expression in 10B2 were also consistent with the low levels of *FRI*

expression (Fig. 4.5). It had also been suggested that the *FLC* allele from location 10 might require the longest period of cold for stable repression. Cold reduced *FLC* expression in 10B2 more slowly than that in other lines, but subsequent differences in expression between families did not correlate with flowering times; for example by five weeks 10B2 showed the same *FLC* expression level as 11A2 but still flowered later. One explanation for this is that the reduction in *FLC* expression might differ in stability between accessions (Shindo et al., 2006), so *FLC* transcript levels might increase once 10B2 plants are returned to the warm after five weeks vernalization but remain low in 11A2. This possibility was not tested experimentally.

The earliest flowering population, 1E2, was found to have the lowest expression of *FLC* at all time points which is consistent with its early flowering and lack of vernalization response. However, it had intermediate levels of *FRI* expression compared to the other families and so its low *FLC* expression could not be attributed to relatively low FRI activity.

Family 5_3 showed the highest *FRI* expression and the highest expression of *FLC* of the local accession, but not compared to the controls. Its *FLC* expression decreased most rapidly on vernalization, dropping by about three quarters in the first two weeks. However, this contrasted to the effects of vernalization on flowering time, which decreased by only about one half after two weeks of cold.

4.6 Discussion

This chapter has examined whether variation in two well characterised candidate genes, *FRI* and *FLC*, can explain the variation in flowering time and vernalization responses between representative local accessions. This was done by comparing genotypes, flowering behaviour and gene expression levels. No overall correlation was found, suggesting that *FRI* and *FLC* are not responsible for most of the variation in flowering behaviour within local populations.

Consistent with other work that has examined worldwide variation in *FRI* and *FLC*, considerable sequence variation was found in local populations, with five different *FRI* alleles and four *FLC* alleles represented. Some of the sequences found in this study had been identified in other Arabidopsis accessions although three novel *FRI* sequence variants and two novel *FLC* alleles were detected.

For several families, the *FRI* genotypes and/or the *FRI* and *FLC* expression levels were consistent with flowering behaviour. For example, the early flowering summer annual accession 108-10 carries a likely *fri* loss of function mutation, while the early flowering, vernalization insensitive winter annual accession 1E2 shows low levels of *FLC* expression.

However, several aspects of the findings are inconsistent with the traditional view that high levels of *FLC* expression require an active *FRI* gene and that both are needed for late flowering and a vernalization response.

(1) Two likely *fri* loss of function alleles were detected (*FRI* III and *FRI* V). Although one was present in an early flowering summer annual accession (108-10), the other was present in all tested accessions from location 11, which are late flowering in the absence of vernalization and very vernalization sensitive. This suggests that genes are present in the local population which repress flowering and allow a vernalization response in the absence of FRI activity.

(2) *FRI* expression levels did not correlate with flowering behaviour. For example 11A2 had no detectable levels of *FRI* transcripts and 10B2 had relatively low levels, but both were late flowering and vernalization sensitive. Again this suggests that local accessions are not dependent on FRI activity for vernalization and late flowering. In the case of 10B2, which had been found to be insensitive to photoperiod (Chapter 3), a mutation in the photoperiod pathway might contribute to late flowering, though this cannot explain its sensitivity to vernalization.

(3) *FLC* expression within local accessions did not correlate with *FRI* expression or genotype; for example, 11A2 maintained higher levels of *FLC* than 1E2 despite having no detectable FRI activity. This suggests *FRI*-independent regulation of *FLC* occurs in local populations.

(4) *FLC* expression did not correlate with flowering behaviour. *FLC* transcript levels were found to be lower in all local accessions than in Col-0 *fri* under comparable conditions, but local accession ranged from early flowering and vernalization insensitive (1E2), to very late flowering with a strong vernalization requirement (10B2). These findings suggest that factors other than *FLC* transcript levels contribute to variation in flowering time variation locally.

These finding differ from comparisons of worldwide accessions, in which *FRI* and *FLC* can explain much of the variation in flowering time and vernalization response. For example, several studies have identified independent loss of function mutations at the *FRI* locus that results in early flowering and loss of the vernalization response in natural accessions (Gazzani et al., 2003; Johanson et al., 2000). A correlation between the *FRI* allele present in a given accession and its downstream *FLC* expression has also been reported in worldwide accessions (Shindo 2005).

The vernalization responses of local accessions also showed no overall correlation with the effects of cold on *FLC* expression. However this study did not examine if there was variation in the stability of *FLC* repression in response to varying cold periods. Shindo *et al.* (2006) have shown that the flowering time of different

worldwide accessions does not correlate with *FLC* expression before or during vernalization, but with the stability of *FLC* repression when plants had been returned to warm temperatures.

Because much of the variation in the flowering behaviour of local populations could be attributed to genes other than *FRI* and *FLC*, QTL mapping was carried out in hybrids between local accessions to identify other candidate loci. These experiments are described in Chapter 5.

5 Production and use of a hybrid population (1E1 X10B2) to identify candidate loci involved in flowering time variation locally

5.1 Introduction:

Studies by (Sheldon et al., 2002) showed that variation in particular regions of the first intron in *FLC* were important for the expression and repression of *FLC* after vernalization. As discussed in Chapter 4 this study has shown that all plants from one site (location 10) contain a disruption in one of these regions. It was also shown that families from this location require an unusually long vernalization treatment to maximally reduce flowering time. This is consistent with the hypothesis that disruption in *FLC* intron 1 is important for the flowering behaviour observed. This chapter aims to explore further the possibility that *FLC* is a candidate locus for the late flowering behaviour of families from location 10. This involved creation of a mapping population by crossing an individual from location 10 to an extremely early flowering family that was shown to be vernalization insensitive and following segregation of flowering time in the F2 hybrids. Flowering time variation was then mapped relative to SSLP markers located near candidate genes and to *FLC*.

The second parent of the cross came from location 1. Families from location 1 have been shown to flower early and to be insensitive to vernalization, though a candidate locus for this behaviour has not been confirmed. Therefore QTL analysis using this cross should also provide evidence about the loci are responsible for the extreme early flowering behaviour of families from location 1. In the previous chapter I showed that family 1E2 had significant *FRI* expression but little detectable *FLC* expression raising the possibility that it is a mutation in *FLC* itself that results in the early, vernalization insensitive flowering of 1E2.

It is unlikely that all of the flowering time variation observed between the parents is due to one locus, especially as they vary in several aspects of their flowering behaviour: flowering time, vernalization sensitivity and photoperiod sensitivity. Therefore this cross should also identify other loci that may be important for the variation in flowering behaviour seen between these the parental populations.

5.2 Production of mapping population

To further explore the role played by *FRI*, *FLC* and other loci in flowering variation in the local populations, crosses were made between plants that varied considerably in flowering behaviour. In this chapter a member of family 1E2 was crossed with pollen from a member of family 10B2. Families from location 1 are very early flowering and vernalisation insensitive whereas families from location 10 are extremely late flowering with a long vernalisation requirement of approximately 10 weeks. From sequence analysis it had been found that families from location 10 contained a disruption in a regulatory region of intron 1 in *FLC*, while families from location 1 carried common alleles of both *FRI* and *FLC*. However using qRT-PCR families from location 1 had been found to express virtually no *FLC* and family 10B2 to express *FLC* at moderate levels, suggesting that the late flowering of plants from location 10 was not due to strong *FLC* activity.

5.3 Late flowering in F1 and F2 hybrids

The F1 plants were grown under LD conditions without vernalization. Plants were very late flowering (Fig.5.1) and later than either parent. Where bolts were produced they formed aerial rosettes eventually leading to production of a pom-pom like shape. Counting leaves to assess flowering time was therefore not possible for these plants. To assess the response of these plants to different environments they were grown in four conditions, LD unvernalized, SD unvernalized, LD vernalized, SD vernalized. A four week vernalization period was chosen because it was the minimum necessary to significantly reduce flowering time in families from location 10. This was sufficient to induce flowering in the F1 hybrid. No difference was seen between the flowering behaviour of F1 plants grown in LD or SD (Table 5.1; Fig 5.1). Since the 10B2 parent was also insensitive to photoperiod, this suggests that insensitivity is determined by a dominant mutation in 10B2.

Seed was collected from F1 plants and an F2 population germinated on MS plates. One hundred seedlings were transplanted to soil, of which 97 survived, and flowering time was scored under LD unvernalized conditions. Of the 97 individuals 54 developed the pom-pom shape, as such leaf counts were discontinued and

Figure 5.1 Parents and F1 plants. (A) 10B2, flowering after being grown under LD conditions without a vernalization treatment. (B) 1E1, flowering after being grown under LD conditions without a vernalization treatment. (C) F1 plant, grown under LD conditions without vernalization treatment (the plant is ~12 months old). (D) F1 plant grown under LD conditions, after ~4months growth this plant was given a 4 week vernalization treatment (the plant is ~6 months old)

Table 5.1 Summary of F1 flowering times. Flowering times are given as the time between transfer of plants to the final growing conditions and opening of the first flower. Plants that were vernalized were 6 months old when treated; four plants in total were vernalized: two that had been grown in LD conditions before treatment and two that were grown in SD conditions before and after treatment. Three cuttings were taken from the same F1 plant and used for the glasshouse experiment. Two F1 plants (one that had been grown in LD growth room conditions and one that been grown in SD growth room conditions) were transplanted to the field site after 12 months.

	Flowering time (after transfer to given condition)	
	Unvernalized	**Vernalized**
Growth room (LD)	DNF	~4 weeks
Growth room (SD)	DNF	~4 weeks
Glasshouse (F1 cuttings)	~8 weeks	N/A
Field	~12 weeks	N/A

flowering dates were recorded. At approximately six months after transplanting to soil any plant that was yet to initiate flowering was recorded as “did not flower”

(DNF); these included 20 of the pom-pom plants. The remaining 34 pom-pom plants flowered over an interval of 55 days.

Because of the branched structure of the pom-pom plants it was possible to take cuttings from them and grow the same genotypes under different conditions to observe flowering behaviour. In glasshouse conditions, of 16°C - 22°C and 16 hours light per day, the cuttings flowered within approximately four weeks, all meristems initiated flowering and the vegetative rosettes senesced soon after bolting. Under field conditions it took approximately six weeks for flowering to be initiated, however this was not the case for all meristems and vegetative axillary rosettes were still present when seeds were collected. This gives rise to the possibility that these plants are not annual and are instead exhibiting a perennial behaviour. It would be interesting to test whether these plants can behave as perennials when grown from seeds in the field, rather than as cutting from mature plants.

As the pom-pom plants aged and grew larger their stems became woody, sections were taken from the thickest stems in the middle of the F1 plants and from younger stems of F1 and F2 plants and stained with toluidine blue. In the younger shoots from the F2 population, secondary thickening involves a very clear ring of lignified cells around the pith in the middle. However the appearance of the stem from the F1 plant is unusual, the entire stem has developed a lobed appearance with a bundle of lignified cells in each lobe as well as a ring of lignified cells around the central pith, which also appeared to contain lignified cells. A similar pattern of lignified cells was seen in younger F1 stems with evidence for the beginning of lobing (Fig. 5.2). This suggests that the lobed structure observed in the F1 plant develops with the age of the stem. More sections at different time points through the plants life would be needed to confirm this, and different stains could confirm whether vascular cells are present in the pith of the F1 plant.

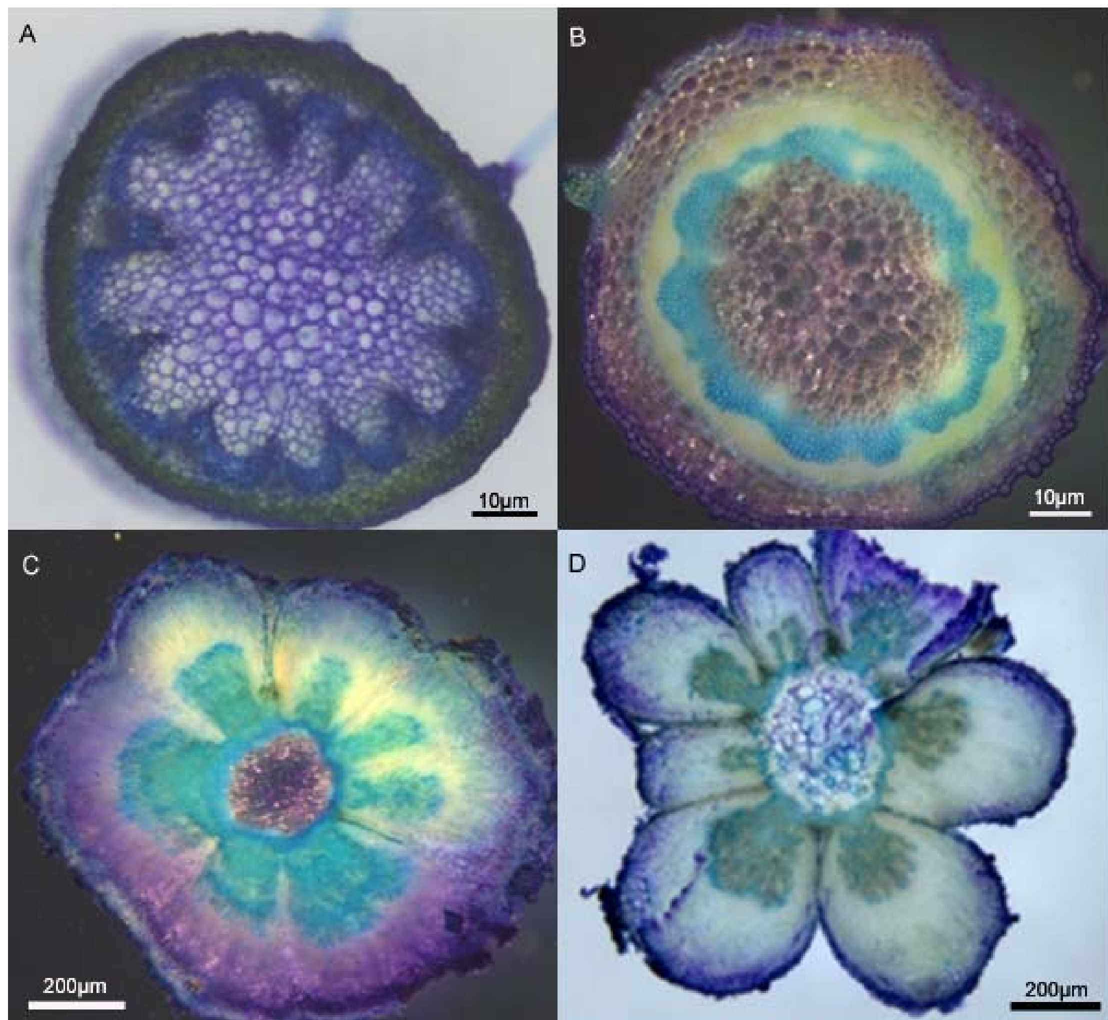


Figure 5.2 Transverse sections of stems stained with toluidine blue to identify lignified cells. (A) Col-0, (B) ~6 month old F2 plant, (C) section from the stem of a cutting taken from F1 plants at an earlier stage in development than the section taken from the thickest/woodiest stem of a 12 month old F1 plant (D).

5.4 Mapping candidate loci

In order to identify chromosomal regions that may contain candidate loci for the flowering behaviour observed in 1E1 and 10B2 primers were used to amplify SSLP markers and the *FLC* region known to differ between the parents (Fig 5.3).

One hypothesis suggested by the *FLC* genotype and vernalization response for families from population 10 was that the disruption in the first intron of *FLC* caused its late flowering. Primers that amplify this region of *FLC* were used to identify the *FLC* genotype of F2 plants. There appeared to be no correlation between *FLC* genotype and flowering time (Fig. 5.4).

In contrast, a correlation was found between flowering time and genotype at marker JV30/31, selected for its proximity to *FRI* (approximately 5 cM north). Of the 25 earliest flowering plants 20 were homozygous for the allele from parent 1E1 at this marker, three were heterozygous and two were homozygous for the allele from parent 10B2. This distribution of alleles is unlikely to have occurred by chance (p <0.0001 in a Chi-squared test). This further supports evidence that *FRI* plays a role in the earliness of families from location 1 and that there may be a loss-of-function mutation in a region of the *FRI* allele from 1E1 that had not been sequenced. The two plants that were shown to have the 10B2 allele at this marker might be explained in one of two ways. Firstly, recombination may have occurred between the marker and *FRI*; more markers in this region would help to test this. Alternatively epistasis could be involved, i.e., another segregating gene is causing the early flowering phenotype in these plants regardless of them having the 'late flowering' allele at *FRI*.

There was also a correlation between late flowering and the genotype at marker NGA139, selected for its proximity to *HUA2* in chromosome 5. Of the 20 plants that did not flower, 11 were homozygous for the allele from parent 10 and 9 were heterozygous, consistent with a dominant allele from 10 B2 causing late flowering. This distribution of genotypes is unlikely to have occurred by chance, since a Chi squared test gave a p value of less than 0.0001.

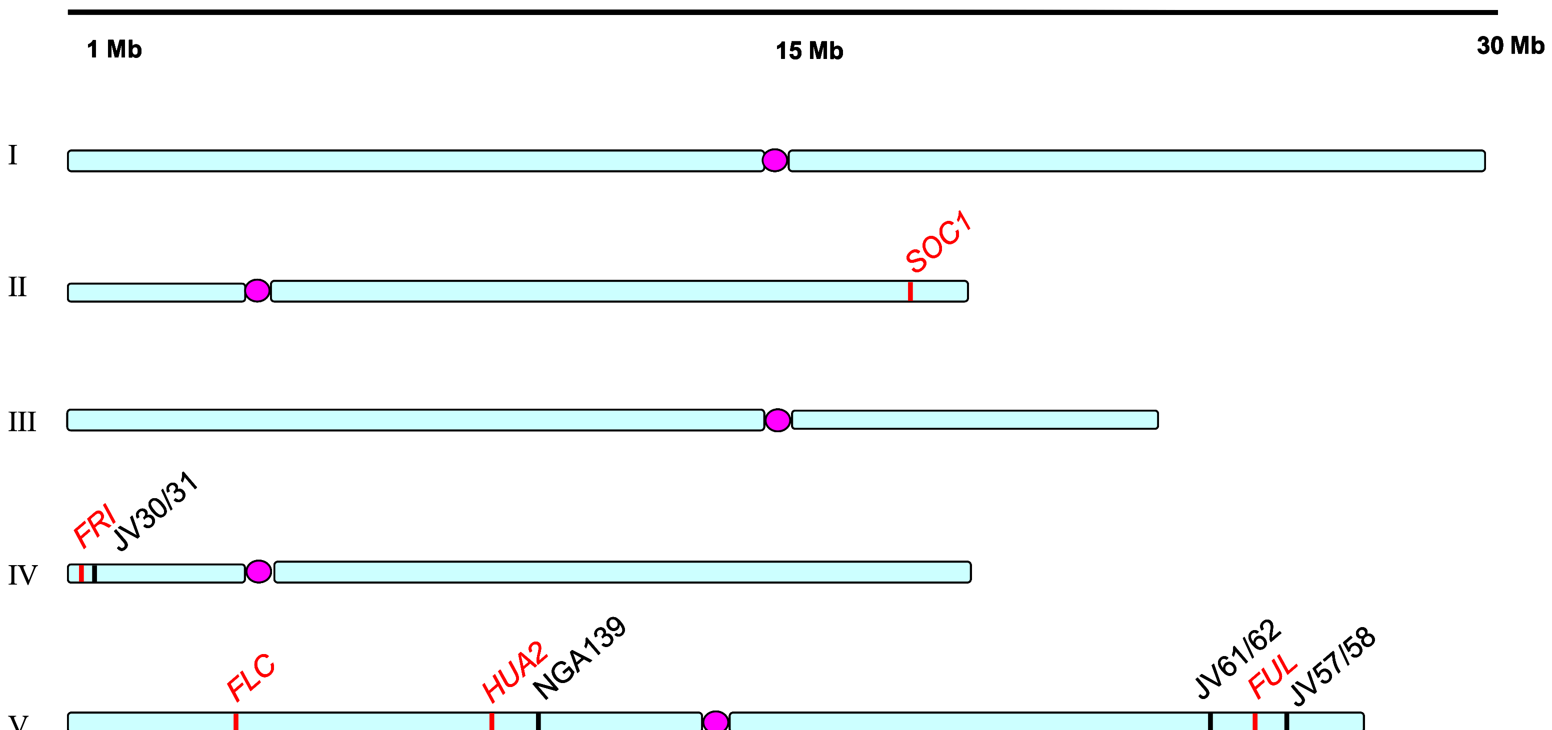


Figure 5.3 Approximate positions of flowering time genes and SSLP markers. Genes relevant to this chapter are shown in red and SSLPs are displayed in black. Roan numerals indicate the chromosome number, centromeres are represented as pink circles.

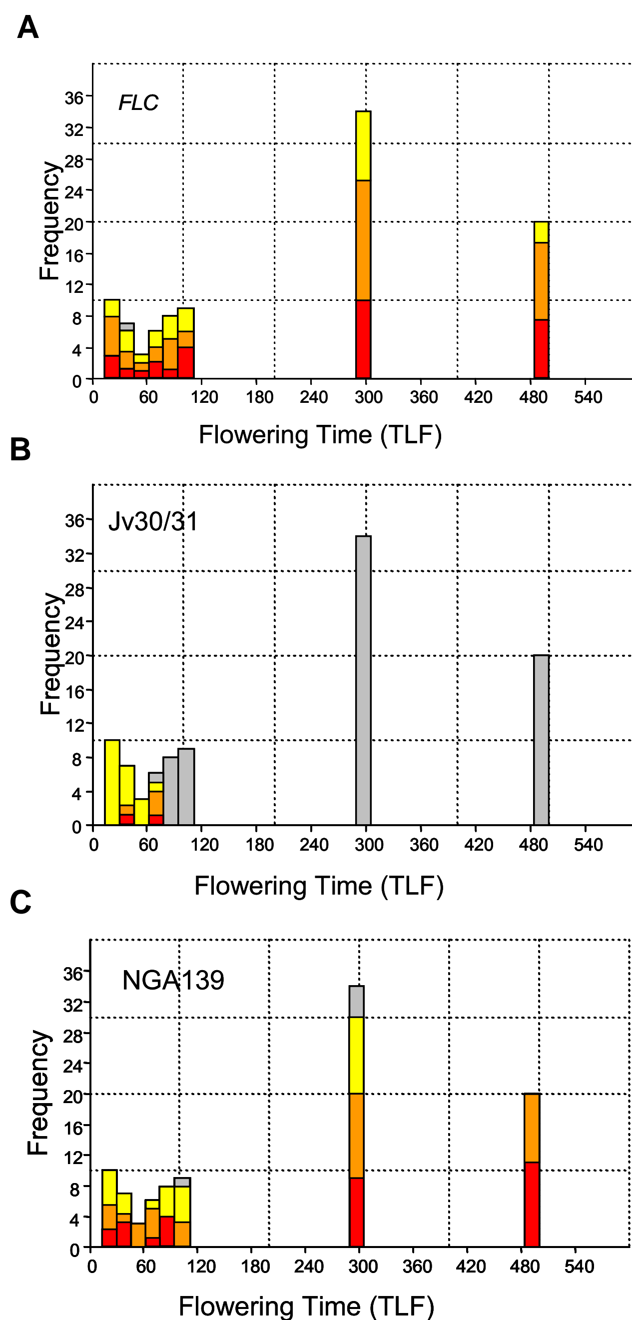


Figure 5.4 Relationship between genotype and flowering time at three markers. In all three cases flowering time is shown for unvernalized plants grown under LD conditions. (A) *FLC deletion F/R*, primers that amplify the region found to contain an 11bp deletion in families from location 10. (B) JV30/31, SSLP marker selected for its proximity to FRI (C) NGA139 selected for its proximity to HUA2. In all cases red shows plants that were homozygous for the 10B2 allele at that marker, yellow indicates homozygotes for the 1E1 allele and orange shows heterozygous plants, grey indicates plants that were not genotyped.

Further QTL analysis would be required to narrow down candidate regions containing the loci responsible for differences in flowering behaviours of 10B2 and 1E1 and candidate genes could then be confirmed by allelism tests. Given that 10B2 had been shown to be insensitive to photoperiod further investigation of loci from the photoperiodic pathway would be interesting.

As the F1 population flowered significantly later than either of the parents, both parents must have contributed to the extreme lateness of the progeny. However one of the parents, 1E1, is among the earliest flowering accessions under all conditions. If a *fri* mutation contributes to the earliness of 1E1, then this parent might carry "late" alleles at other loci that are not apparent because the *fri* allele has an epistatic effect on flowering time. Further analysis of *FRI* genotypes in the latest flowering plants could test this possibility.

5.6 Discussion

Results in this chapter suggest that *FLC* is unlikely to be responsible for the late flowering behaviour observed in population 10 or the early flowering of population 1. The rough QTL analysis does suggest that *FRI* might contribute to the early flowering in population 1, however no significant sequence variation was found in *FRI* from this population (Chapter 4). Clear *FRI* expression was also observed (Chapter 4), however it was significantly lower than that of vernalization sensitive family 5_3, raising the possibility that it was sufficiently low *FRI* to reduce flowering time. This idea is supported by the fact that population 1 has virtually no *FLC* expression (Chapter4) and yet flowering time variation does not map to *FLC*, suggesting that other genes that regulate *FLC* expression are involved.

NGA139 genotypes were shown to correlate with flowering time and therefore this marker could be close to a candidate locus. NGA139 is close to the region in which *HUA2* is found, and a dominant allele of *HUA2* is known to cause extreme late flowering by upregulating expression of *FLC* (Wang et al., 2007). The natural accession Sy-0 from the Isle of Skye was found to have an extremely late flowering phenotype and after floral initiation inflorescence meristems can revert to vegetative

growth producing aerial rosettes. This phenotype was caused by dominant alleles of two unlinked genes *AERIAL ROSETTE 1* (*ART1*) and *ENHANCER OF AERIAL ROSETTE* (*EAR*), which was later found to be allelic to *FRI*. Mapping of the *ART1* gene revealed that it was an allele of the floral regulator *HUA2* (*HUA2*-Sy-0) that enhanced the expression of *FLC*, especially within the shoot meristem, resulting in late flowering (Grbic and Bleecker, 1996; Wang et al., 2007).

All F2 plants that did not flower were found to either be homozygous for the 10B2 copy of NGA139 or to be heterozygous suggesting that a dominant allele from 10B2 was responsible for late flowering. *HUA2* has only been shown to cause this late flowering phenotype in one accession, Sy-0 from the Isle of Skye, it differs from the effect seen in F1 and some F2 plants in that Sy-0 flower very late whereas the F1 progeny of this hybrid population and F2 plants carrying the 10B2 allele of NGA139 in the F2 generation do not flower at all unless vernalized.

The extreme late flowering observed in the F1 generation cannot be explained solely by the dominant allele linked to NGA139 in the 10B2 parent, because the F1 flowered later than both its parents. This suggests that it involves additive effects of dominant alleles at other loci contributed by both parents.

Secondary growth and perennial type flowering behaviour has been seen in Arabidopsis however this was observed in a mutagenesis study and not a study using natural accessions. Melzer et al (2008) showed that mutations in *SOC1* and *FUL* resulted in the plants that were extremely late flowering and, after the induction of flowering, some vegetative meristems remained. It is likely that the longevity of the pom-pom plants has resulted in the unusual secondary growth patterns observed in these plants however further experiments would be needed to confirm this, for example by inducing earlier flowering in F1 plants by vernalization at the seedling stage. Given the role of *SOC1* in the age dependent control of flowering it would be interesting to see whether natural variation at this locus might be involved.

Chapter 6 Production and use of a hybrid population (11C1 X4D1) to identify candidate loci involved in flowering time variation locally

6.1 Introduction:

A naturally occurring *fri* loss of function allele had been found in families from location 11, which exhibits extremely late flowering without vernalization and a strong and rapid vernalization response that saturates within two weeks (Chapter 4). The most vernalization sensitive family from this location (11C1) was selected for the production of a hybrid population to perform a QTL screen. It was crossed to 4D1, a family that responded significantly less to vernalization (Chapter 3).

Previous studies that have used QTL analysis to identify loci underlying variation in flowering time and vernalization responses between *Arabidopsis* accessions have found several major effect loci that are often responsible for the majority of variation observed. These include the floral regulators *FRI* and *FLC* as well as loci involved in photoperiodism such as *PHYC*, the floral integrator *FT* and a commonly identified QTL at the bottom of chromosome 5 in a region known to contain several likely candidates for flowering time variation including the vernalization pathway gene *VIN3* and *VIP4* and four members of the MADS box class of genes *MADS AFFECTING FLOWERING* (*MAF2-MAF5*).

This chapter aims to analyse the variation in flowering time and vernalization response in an F4 population between parents that differ in their flowering time and vernalization responses. Additionally it will examine loci that are potential candidates for this observed variation.

6.2 Production of a mapping population

A hybrid population was produced by Andrew Hudson by fertilising a plant from family 11C1 family with pollen from 4D1 (fig 6.1). The parents for this cross had been selected because of their differences in vernalization response along with other traits of interest that were not related to flowering.

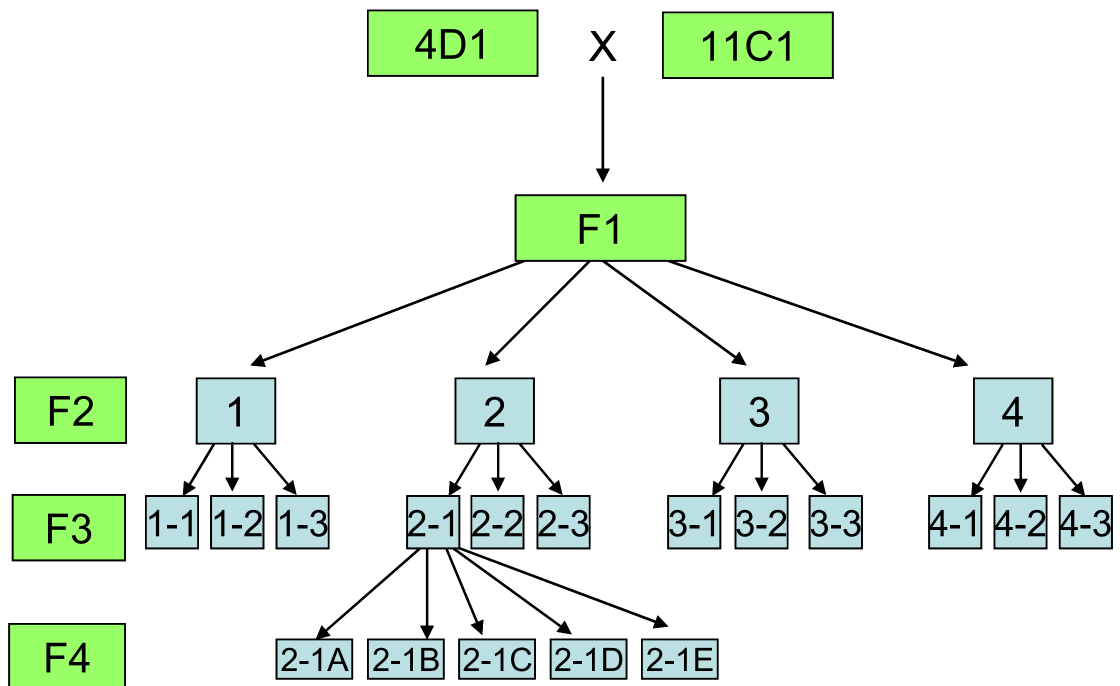


Figure 6.1 Creation of F4 mapping population. From the F2 generation onwards each plant was assigned a number and seed was collected from individual plants, this naming system continued so that the history of each line could be traced. For clarity in this figure only one F4 family is shown, although progeny from all F3 plants were grown and named in this manner.

Families from location 11 exhibit an extreme vernalization response (Chapter 3) and this response is saturated after just two weeks of cold treatment (Chapter 4). Families from location 11 were of additional interest because they were shown to carry a non-functional allele of *FRI* unusually for such a strong vernalization requirement, suggesting that suppressors of the *fri* mutant phenotype were also present.

In contrast, populations from location 4 had a below average vernalization responses relative to other families in the winter annual population with only a 35% decrease in leaf number in response to 6 weeks vernalization (the average decrease was 49%). Family 4D1 also differed from 11C1 in that it carried common alleles of both *FRI* and *FLC*.

6.3 Flowering time variation in F4 generation

A population of 89 F4 families (five individuals of each family) was grown under long day conditions either following a four week vernalization period or without vernalization. Total leaves to flower (TLF) was used as a measure of flowering time. Flowering times for all F4 families for which there was both vernalized and unvernalized data are shown in Figure 6.2.

Under both vernalized and unvernalized conditions some families flowered later than either parent. This suggests that both parents carry genes that delay flowering, so that some hybrids can inherit more genes for later flowering than were present in either parent. Under vernalized conditions no lines flowered earlier than 11C1, suggesting that genes conferring a strong vernalization response come only from 11C1.

Some of the families showed a large variance for flowering time, consistent with segregation of flowering time genes within these families. Around one quarter of the F4 families are expected to segregate for any locus that differed between the parents.

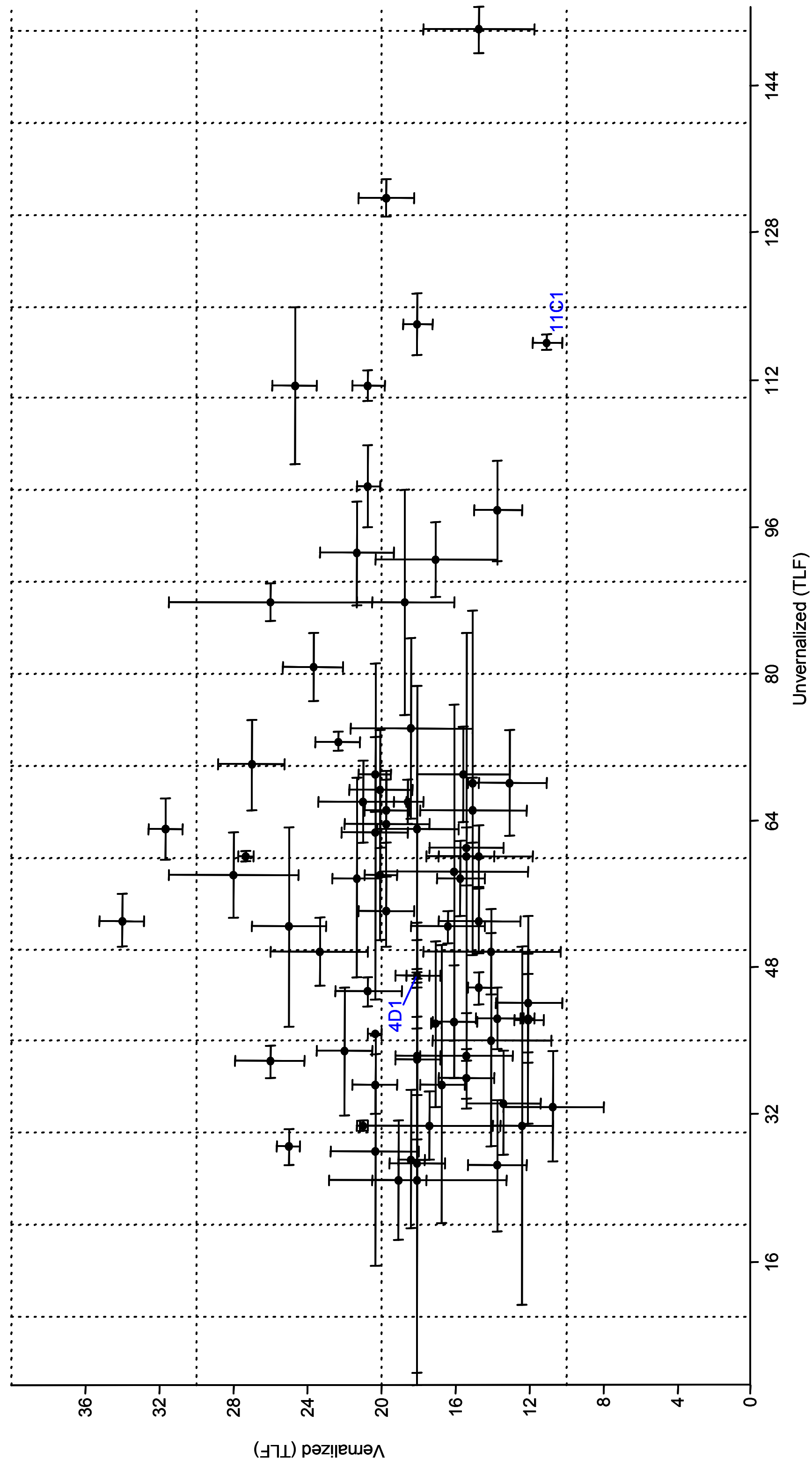


Figure 6.2 Flowering times of parents and F4 generation of a hybrid mapping population. Flowering times are given in mean total leaves until flowering (TLF) from ~5 plants per line ± their standard errors. The parents (4D1 and 11C1) are labelled.

6.4 Mapping candidate loci

From the F4 population the 12 earliest flowering and the 12 latest flowering families were selected and screened using SSLP (simple sequence length polymorphism or microsatellite) markers linked to known flowering time genes that were polymorphic between the parents (Fig. 6.4). This work was initially done with assistance from an Honours student, Emma Lindsay.

The first thing of note was that no correlation was found between the genotype of the marker selected for its proximity to *FRI* and flowering time or vernalization response (Fig. 6.5). This was surprising given that 11C1 was shown to carry a likely non-functional *fri* allele and showed no detectable *FRI* expression (Chapter 4). The apparent lack of correlation between flowering time and *FRI* genotype is consistent with an epistatic interaction in which other suppressor genes are inducing a vernalization requirement in 11C1 resulting in the *fri* mutation being irrelevant to flowering time. These suppressors might also be present in 4D1 and so would not segregate to reveal an effect of *fri* in the hybrid population. Alternatively, they might segregate but result in too few *fri* mutant plants lacking supression to detect the effect of *fri*.

There was also no correlation between flowering time under either condition and the genotype at the marker selected for its proximity to *FLC* (Fig. 6.6).

A correlation was however found between flowering behaviour and a marker situated at the bottom of chromosome 5 (JV75/76; Fig. 6.3). Under vernalized conditions the latest flowering plants were homozygous for the 4D1 allele at this marker (Fig. 6.6). This correlation was not detected under unvernalized conditions, suggesting that the candidate locus in this region is likely to be involved in the vernalization response, rather than flowering under both vernalized and unvernalized conditions. One known component of the vernalization pathway, *VIN3*, is present in this region.

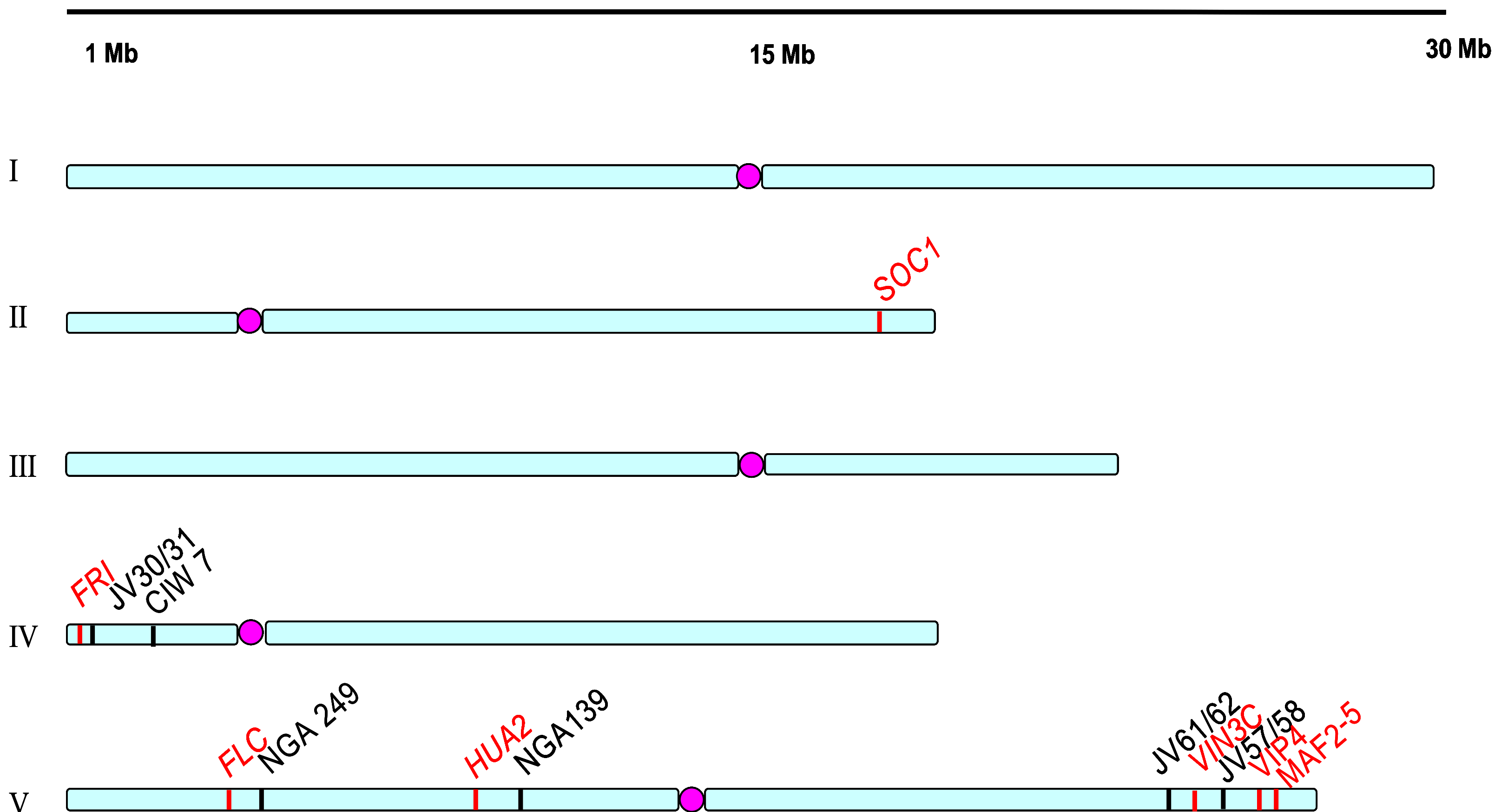


Figure 6.3 Approximate positions of flowering time genes and SSLP markers. Genes relevant to this chapter are shown in red and SSLPs are displayed in black. Roman numerals indicate the chromosome number, centromeres are represented as pink circles.

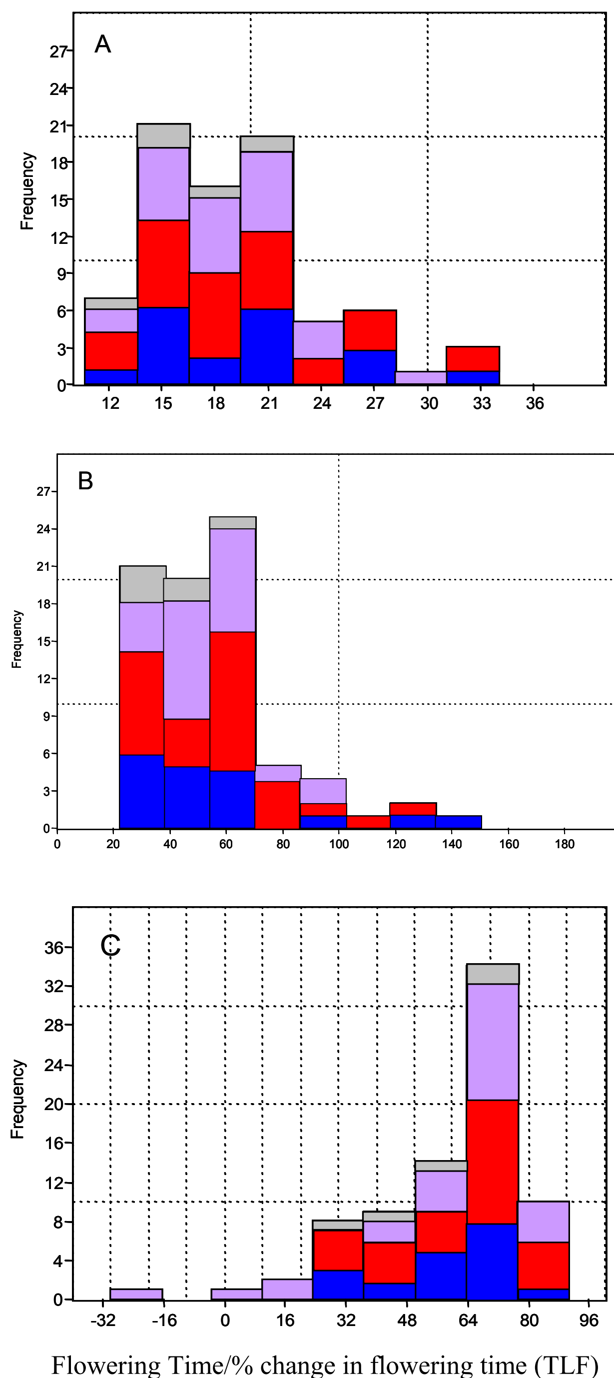


Figure 6.4 relationship between genotype and flowering time under vernalized and unvernalized conditions and with respect to overall vernalization response. Genotypes of F4 lines at marker NGA249, selected for proximity to *FLC* in relation to distribution of flowering times in total leaves to flower (TLF) after a 6 week vernalization treatment (A), without vernalization(B) or as % change in TLF in response to a 6 week vernalization treatment. Purple indicates heterozygous lines, blue shows lines that are homozygous for the 4D1 allele and red indicates lines that are homozygous for the 11C1 allele.

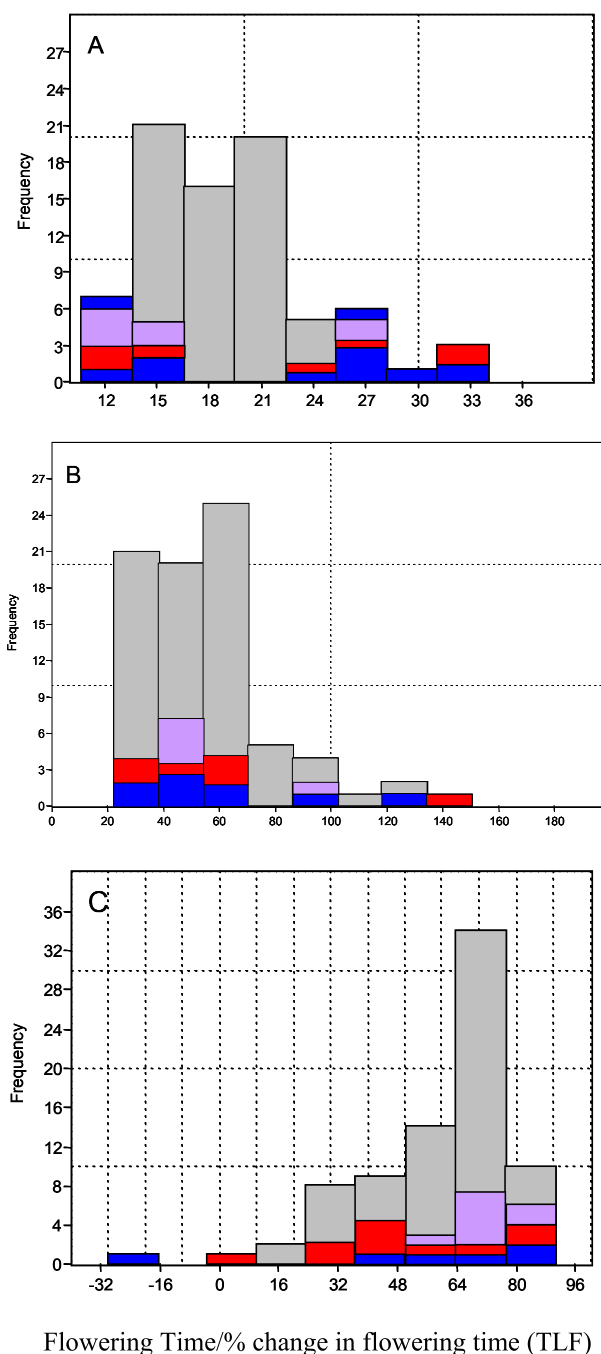


Figure 6.5 relationship between genotype and flowering time under vernalized and unvernalized conditions and with respect to overall vernalization response. Genotypes of F4 lines at marker JV30/31, selected for proximity to *FRI* in relation to distribution of flowering times in total leaves to flower (TLF) after a 6 week vernalization treatment (A), without vernalization(B) or as % change in TLF in response to a 6 week vernalization treatment. Purple indicates heterozygous lines, blue shows lines that are homozygous for the 4D1 allele and red indicates lines that are homozygous for the 11C1 allele.

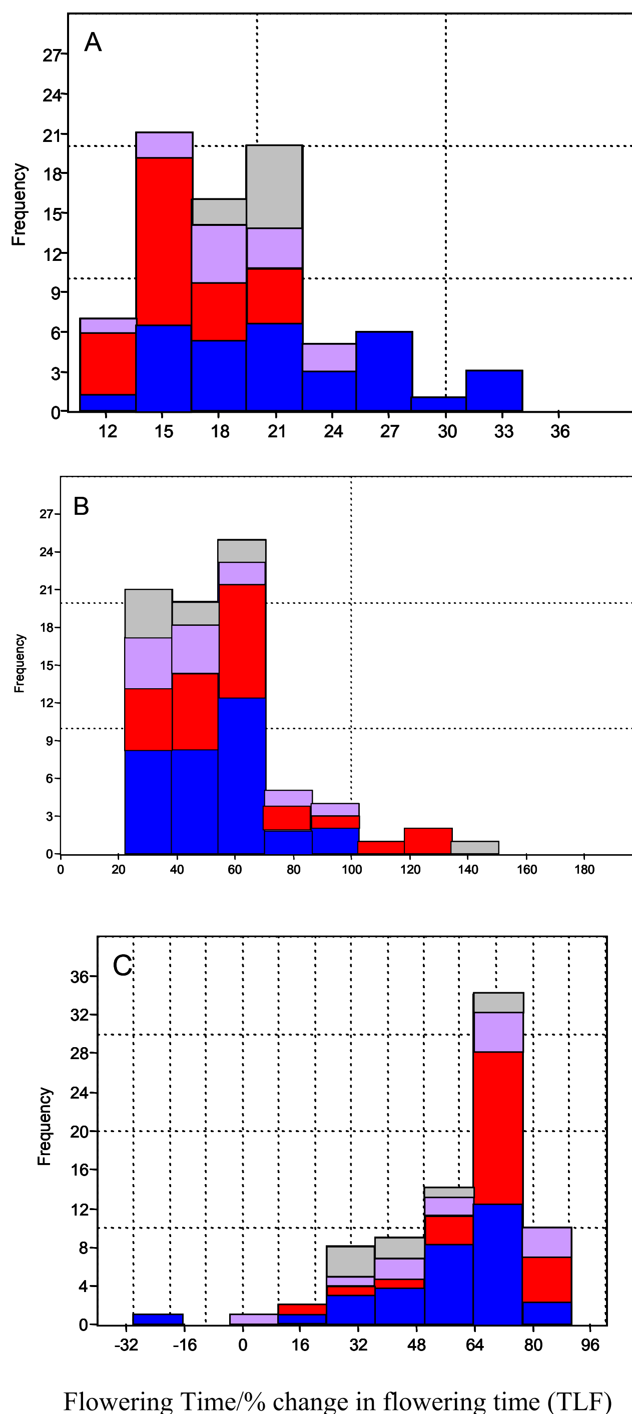


Figure 6.6 relationship between genotype and flowering time under vernalized and unvernalized conditions and with respect to overall vernalization response. Genotypes of F4 lines at marker JV75/76 in relation to distribution of flowering times in total leaves to flower (TLF) after a 6 week vernalization treatment (A), without vernalization(B) or as % change in TLF in response to a 6 week vernalization treatment. Purple indicates heterozygous lines, blue shows lines that are homozygous for the 4D1 allele and red indicates lines that are homozygous for the 11C1 allele.

Because *VIN3* is induced by low temperature and represses *FLC* as part of the vernalization response (Wood et al., 2006), variation in *VIN3* function might account for the differences in vernalization sensitivity between 4D1 and 11C1. Therefore the potential role of *VIN3* was investigated further.

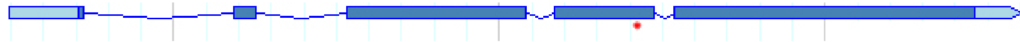
Additional markers were developed to distinguish the parental chromosomes in the region around *VIN3*, including an insertion-deletion (indel) polymorphism in *VIN3* itself (see below). These were used to identify F4 individuals which carried recombinations north of *VIN3* or between *VIN3* and other candidate genes (*VIP4* and *MAF2–MAF5*) further south. By measuring flowering and vernalization responses in the offspring of plants carrying these recombinant chromosomes it should be possible to determine whether the flowering behaviour co-segregates with *VIN3* or with other genes in this region of the chromosome.

6.5 *VIN3* sequence variation

A series of primers were designed to amplify and sequence the *VIN3* gene from 4D1 and 11C1 in overlapping sections. The region sequenced extended 1.6 kb upstream of the start of the open reading frame and therefore included all the sequences that had been shown to be needed to complement a *vin3* mutation (Wood et al., 2006). A total of 15 polymorphisms were identified; 11 single nucleotide polymorphisms and 4 indels (insertions or deletions) the majority of which were located in non-coding regions (Fig. 6.4). Only one SNP resulted in a non-synonymous substitution, with a difference between an alanine (A) and a threonine (T) residue at position 238 (Fig. 6.5). This substitution could potentially affect VIN3 protein function, although was not shown to be located in a conserved region of VIN3 and as such this variation was not studied further in this study.

The vast majority of sequence variation identified was found either in the promoter region or in introns. This could potentially cause variation in expression levels or splicing.

A



B

```
4D1       AGCATTTGCGTATATAATATCTTCCAAAGCACGTGAAGATATTTCCAACGTTTTTTTTAA 60
11C1      AGCATTTGCGTATATTATATCTTCCAAAGCACGTGAAGATATTTCCAACGTTTTTTTTAA 60
          *************** ********************************************

4D1       TATTAAAAAATACACAAAACCAAAATTCTAAAGAACTTAAAAAAAAATAGCTCTCATCAG 120
11C1      TATTAAAAAATACACAAAACCAAAATTCTAAAGAACTTAAAAAAAAAAAGCTCTCATCAG 120
          *********************************************** ************

4D1       AGCAGAAGTTTCATCCATCAGAGGGTTTCCTCCTTAGAAACATCTAGAAAAAACAAAAGG 180
11C1      AGCAGAAGTTTCATCCATCAGAGGGTTTCCTCCTTAGAAACATTTAGAAAAAACAAAAGG 180
          ******************************************* ****************

4D1       AGAGAGAGAGAGA--AGAACCATCCATCACAGAAACATCTTTTCCTTTTTCACTAAATCC 238
11C1      AGAGAGAGAGAGAGAAGAACCATCCATCACAGAAACATCTTTTCCTTTTTCACTAAATCC 240
          *************  *********************************************

4D1       --------GATCTCTAAAATAAACCTAGCAAAAAAAAAAATCAAGAGAAAAAAAAACGAAG 290
11C1      CCATAAAAGATCTCTAAAATAAACCTAGCAAAAAAACAAATCAAGAGAAAAAAAAACGAAG 300
                  **************************** ***********************

4D1       AACACGAAGAACGACAAACAAAAAAAAAATGCAAGCTGCTTCGGTAAACAGCGTTTTTTT 350
11C1      AACACGAAGAACGACAAACAAAAAAAAAATGCAAGCTGCTTCGGTAAACAGCTTTTTTTT 360
          **************************************************** *******

4D1       TTTATATAAATATTCTTTGTATGCTGTTTTTTTTTGAAGGCTTCTCATTCGGTGAACAAC 410
11C1      TTTATATAAATATTCTTTGTATGCTGTTTTTTTTTGAAGGCTTCTCATTCGGTGAACAAC 420
          ************************************************************

4D1       AACCAATCCTAATTTTGCATATATATTTACTCTCTGAGTATTTGCTTATTTTAATTAAAG 470
11C1      AACCAATTCTAATTTTGCATATATATTTACTCTCTGAGTATTTGTTTATTTTAATTAAAG 480
          ******* ************************************ ***************

4D1       AAAATTCTCAATGGTTTGTTCTTAAGACTGAGTCATCCTCGAATCTTTTGTCTAATAGTT 530
11C1      AAAATTCTCAATGGTTTGTTCTTAAGACTGAGTCATCCTCGAATCTTTTGTCTAATAGTT 540
          ************************************************************

4D1       ACAAATATAGGTTCTTACAAAGTGTATTTCATGGGTTTCTCTTTGGTTGTTTAAAGATTA 590
11C1      ACAAATATAGGTTCTTACAAAGTGTATTTCATGGGTTTCTCTTTGGTTGTTTAAAGATTA 600
          ************************************************************

4D1       AAACTTCAAAATGAAAATCGATAAAATCTGAATTTTTTTTTCGATCTTGTGTTCTTCATC 650
11C1      TAACTTCAAAATGAAAATCGATAAAATCTGAATTTTTTTTTCGATCTTGTGTTCTTCATC 660
           ***********************************************************

4D1       ATCGTAAGTGTTACATCTTTGAGAATTTATTTTTTGTCTGTACTTAAGTGTTATTGTCAG 710
11C1      ATCGTAAGTGTTA----------------------------------------TTGTCAG 680
          *************                                        *******

4D1       ATTGTTGTGATGGTTTATTTTTTTTGGTTCTTGCTTTAACTTTGACATAATTGAACCTTT 770
11C1      ATTGTTGTGATGGTTTTATTTTTTTGGTTCTTGCTTTAACTTTGACATAATTGAACCTTT 740
          ****************  ******************************************

4D1       TGTGTGAAATCGGATCTCG----GCAGCTCTCAAAGATCTGGCGTTTCGATGGTAATGTG 826
11C1      TGTGTGAAATCGGATCTCTCTCTGCAGCTCTCAAAGATCTGGCGTTTCGATGGTAATGTG 800
          ******************     *************************************

4D1       GGTCCAGAGAACATGGATTCTTCTTCGTTTGAAGGTACGATATTTTCTCGACGTGAACGT 886
11C1      GGTCCAGAGAACATGGATTCTTCTTCGTTTGAAGGTACGATATTTTCTCGACGTGAACGT 860
          ************************************************************

4D1       CTCTCTCACTGTTCCTAGTTTCATGCGTTTTAGGTCGTTAACACGACGACGTTTTTACAC 946
11C1      CTCTCTCACTGTTCCTAGTTTCATGCGTTTTAGGTCGTTAACACGACGACGTTTTTACAC 920
          ************************************************************

4D1       AATATCAACAAACGTTTTATTTCTCGCATTGGATTCGTCGTCATTGTTGATATTTCGTTT 984
11C1      AATATCAACAAACGTTTTATTTCTCGCATTGGATTCGTCGTCATTGTTGATATTTCGTTT 958
          ************************************************************

4D1       CTGGGTTGTTGAGATTCATTCAAAGTTTTATACTTTTGATCATTTTTACCCCAAAAAGAG 1044
11C1      CTGGGTTGTTGAGATTCATTCAAAGTTTTATACTTTTGATCATTTTTACCCCAAAAAGAG 1018
          ************************************************************

4D1       TCAGAATCTGTGATCCTCTCCTTTTTATCTTGAGGAGAATCTAGGGCTAAAGGATTGGTC 1104
11C1      TCAGAATCTGTGATCCTCTCCTTTTTATCTTGAGGAGAATCTAGGGCTAAAGGATTGGTC 1078
          ************************************************************
```

```
4D1     TTTTATGGTGCAGATAATGAATGCATTGAGACTTGTAAACCAAACGTTTTGAATGTAAGT 1164
11C1    TTTTATGGTGCAGATAATGAATGCATTGAGACTTGTAAACCAAACGTTTTGAATGTAAGT 1138
        ************************************************************

4D1     GAAAGGAGAGAATTGATCCACGCATTGTCTAACCAGCCTGAAGAAGCTTCGGAGCTTTTG 1224
11C1    GAAAGGAGAGAATTGATCCACGCATTGTCTAACCAGCCTGAAGAAGCTTCGGAGCTTTTG 1198
        ************************************************************

4D1     AATTCATGGAGCAGAAATGAGATCATGAAGATCATATGTGCTGAGATGGGTAAAGAGAGG 1284
11C1    AATTCATGGAGCAGAAATGAGATCATGAAGATCATATGTGCTGAGATGGGTAAAGAGAGG 1258
        ******************************************************** ***

4D1     AAGTACACTGGTCTTAACAAACCAAAGCTCATAGAGAATCTTCTGAATCTTGTGTCTCGT 1344
11C1    AAGTACACTGGTCTTAACAAACCAAAGCTCATAGAGAATCTTCTGAATCTTGTGTCTCGT 1318
        ************************************************************

4D1     CCTCTTGGAGAGACCTCTTGTTCTGACCGTAGAAACTCGAGGAAGAAGGAGAAGAAGATG 1404
11C1    CCTCTTGGAGAGACCTCTTGTTCTGACCGTAGAAACTCGAGGAAGAAGGAGAAGAAGATG 1378
        ************************************************************

4D1     ATCGGTTACATCATTTGCTGTGAGAATTTAGCTTGTAGAGCTGCGCTTGGATGCGATGAT 1464
11C1    ATCGGTTACATCATTTGCTGTGAGAATTTAGCTTGTAGAGCTGCGCTTGGATGCGATGAT 1438
        ************************************************************

4D1     ACGTTTTGCAGAAGGTGTTCTTGCTGCATCTGTCAAAAGTTTGATGATAATAAGGATCCT 1524
11C1    ACGTTTTGCAGAAGGTGTTCTTGCTGCATCTGTCAAAAGTTTGATGATAATAAGGATCCT 1498
        ************************************************************

4D1     AGTTTATGGCTTACTTGTGATGCTTGTGGATCGTCTTGTCATTTGGAATGTGGTTTGAAG 1584
11C1    AGTTTATGGCTTACTTGTGATGCTTGTGGATCGTCTTGTCATTTGGAATGTGGTTTGAAG 1558
        ************************************************************

4D1     CAAGATAGGTATGGGATTGGGAGTGATGATCTTGATGGTAGGTTTTATTGCGCGTATTGC 1644
11C1    CAAGATAGGTATGGGATTGGGAGTGATGATCTTGATGGTAGGTTTTATTGCGCGTATTGC 1618
        ************************************************************

4D1     GGTAAAGATAATGACTTGCTCGGGTAAGGATCTTTAGATATTGAAATCTATGGATATACT 1704
11C1    GGTAAAGATAATGACTTGCTCGGGTAAGGATCTTTAGATATTGAAATCTATGGATATACT 1678
        ************************************************************

4D1     ACAATGAAAAGAGAATGTAATGTAATGTGTGTGTGTATGCGAGTGCTTACAGATGCTGGA 1764
11C1    ACAATGAAAAGAGAATGTAATGTAATGTGTGTGTGTATGCGAGTGCTTACAGATGCTGGA 1738
        ************************************************************

4D1     GAAAACAAGTGAAGGTGGCGAAAGAGACGCGGCGTGTGGATGTACTTTGTTATCGTCTTT 1824
11C1    GAAAACAAGTGAAGGTGGCGAAAGAGACGCGGCGTGTGGATGTACTTTGTTATCGTCTTT 1798
        ************************************************************

4D1     CTTTAGGACAGAAGCTGTTGAGAGGTACCACGAAGTATCGGAATCTGTTGGAACTTATGG 1884
11C1    CTTTAGGACAGAAGCTGTTGAGAGGTACCACGAAGTATCGGAATCTGTTGGAACTTATGG 1858
        ************************************************************

4D1     ATGAGGCGGTGAAGAAGCTCGAAGGTGATGTGGGTCCGTTGTCGGGTTGGGCCATGAAGA 1944
11C1    ATGAGGCGGTGAAGAAGCTCGAAGGTGATGTGGGTCCGTTGTCGGGTTGGGCCATGAAGA 1918
        ************************************************************

4D1     TGGCTCGAGGCATCGTCAATAGACTTTCTTCGGGTGTGCATGTCCAGAAGCTGTGTTCTC 2004
11C1    TGGCTCGAGGCATCGTCAATAGACTTTCTTCGGGTGTGCATGTCCAGAAGCTGTGTTCTC 1978
        ************************************************************

4D1     AGGCAATGGAAGCTCTGGACAAAGTGGTCTCACCATCAGAATCTGTTTCAGGACAAGGTT 2064
11C1    AGGCAATGGAAACTCTGGACAAAGTGGTCTCACCATCAGAATCTGTTTCAGGACAAGGTT 2038
        *********** ************************************************

4D1     AGTGTGTTCATACATTACTTTGGTTTTAAAACCGAAGATGTTGACTTTAGTTCCTGTTAG 2124
11C1    AGTGTGTTCATACATTACTTTGGTTTTAAAACCGAAGATGTTGACTTTAGTTCCTGTTAG 2098
        ************************************************************

4D1     GTGACAAGATGACCGTGAGAGTAGAAGAGATTCAAGCAAGATCAGTCACTGTGAGAGTAG 2184
11C1    GTGACAAGATGACCGTGAGAGTAGAAGAGATTCAAGCAAGATCAGTCACTGTGAGAGTAG 2158
        ************************************************************

4D1     ACTCCGAGGAGCCGTCTTCTTCTACACAAAACAAGATCACAGGTTTCAGGTTGTTTTGTC 2244
11C1    ACTCCGAGGAGCCGTCTTCTTCTACACAAAACAAGATCACAGGTTTCAGGTTGTTTTGTC 2218
        ************************************************************

4D1     GAAAGTCGAAGGACGAAGAATGCTCGTCTCAGGGGAATTGTGTTGTTTATCTACCTGAGA 2304
11C1    GAAAGTCGAAGGACGAAGAATGCTCGTCTCAGGGGAATTGTGTTGTTTATCTACCTGAGA 2278
        ************************************************************

4D1     CGACGTCTGCCATCCAAGGACTTGAACCCGACACCGAGTTCTGTCTCAGAGTGGTTTCCT 2364
11C1    CGACGTCTGCCATCCAAGGACTTGAACCCGACACCGAGTTCTGTCTCAGAGTGGTTTCCT 2338
        ************************************************************
```

```
4D1                 TTAACGAGGAAGGTGACTTAGATGAGTCTGAGCTTCGGTTCACAACGTTGAAGGATGATG 2424
11C1                TTAACGAGGAAGGTGACTTAGATGAGTCTGAGCTTCGGTTCACAACGTTGAAGGATGATG 2398
                    ************************************************************

4D1                 GAGATGAAGCTGGGGACCAGCAAAGCCCTTTGACAAACTCAAGCAGTGGTCTTTGTAGTA 2484
11C1                GAGATGAAGCTGGGGACCAGCAAAGCCCTTTGACAAACTCAAGCAGTGGTCTTTGTAGTA 2458
                    ************************************************************

4D1                 ATCCATCTTTACCAGAAGATGAATCTAATAATGTCAATAAAAGCTGCAGCAAAGGAAATG 2544
11C1                ATCCATCTTTACCAGAAGATGAATCTAATAATGTCAATAAAAGCTGCAGCAAAGGAAATG 2518
                    ************************************************************

4D1                 GTGACAAGGACAACACTGAACACTGTAGTGCAGGAGAAGTAGAATCTGAGCTTGAAGAAG 2604
11C1                GTGACAAGGACAACACTGAACACTGTAGTGCAGGAGAAGTAGAATCTGAGCTTGAAGAAG 2578
                    ************************************************************

4D1                 AGAGGCTTGTAAAGAGGAAAGCAAACAAGATAGATGGAAGAGACTTGCTTGTAACACCCT 2664
11C1                AGAGGCTTGTAAAGAGGAAAGCAAACAAGATAGATGGAAGAGACTTGCTTGTAACACCCT 2638
                    ************************************************************

4D1                 GCAAGAGAGATATTTATAAAGGAAAGCAAGGAGGGAATAAAAGATTCAAATCAAGAACAG 2724
11C1                GCAAGAGAGATATTTATAAAGGAAAGCAAGGAGGGAATAAAAGATTCAAATCAAGAACAG 2698
                    ************************************************************

4D1                 TATCCTTGAACGAGAAACCTGAGATCAATAATGCCGCAAATGGAGTAGGAGATAAAGACT 2784
11C1                TATCCTTGAACGAGAAACCTGAGATCAATAATGCCGCAAATGGAGTAGGAGATAAAGACT 2758
                    ************************************************************

4D1                 TGGGTCATATTGTTAAGACGATTAGATGTTTAGAGGAAGAAGGACATATAGACAAGAGTT 2843
11C1                TGGGTCATATTGTTAAGACGATTAGATGTTTAGAGGAAGAAGGACATATAGACAAGAGTT 2818
                    ************************************************************

4D1                 TTAGGGAAAGGTTCTTGACATGGTATAGCTTAAGAGCTACTCACCGAGAAGTAAGAGTTG 2904
11C1                TTAGGGAAAGGTTCTTGACATGGTATAGCTTAAGAGCTACTCACCGAGAAGTAAGAGTTG 2878
                    ************************************************************

4D1                 TGAAGATCTTTGTTGAGACGTTTATGGAGGATCTGTCTTCTTTGGGACAACAGCTTGTGG 2964
11C1                TGAAGATCTTTGTTGAGACGTTTATGGAGGATCTGTCTTCTTTGGGACAACAGCTTGTGG 2938
                    ************************************************************

4D1                 ATACATTCTCAGAAAGTATACTGAGTAAGAGATCATCGACAAATGGTGTAGTACCTGCTG 3024
11C1                ATACATTCTCAGAAAGTATACTGAGTAAGAGATCATCGACAAATGGTGTAGTACCTGCTG 2998
                    ************************************************************

4D1                 GGATCTGCCTCAAGCTTTGGCATTAAAAGCTTTTGAAATCGGTCTATTTATTAACTCGCT 3084
11C1                GGATCTGCCTCAAGCTTTGGCATTAAAAGCTTTTGAAATCGGTCTATTTATTAACTCGCT 3058
                    ************************************************************

4D1                 TGATATTGTAAACCATAACGATAATTTTGGAATGAAATTATGTAAGTACGCATTTCCAAT 3144
11C1                TGATATTGTAAACCATAACGATAATATTGGAATGAAATTATGTAAGTACGCATTTCCAAT 3118
                    ************************* **********************************

4D1                 TTCTTTCTGATGCTTAGTGAAAAACTCAAAAATGAATAAGCCTGATTTCAAATGAACCAT 3204
11C1                TTCTTTCTGATGCTTAGTGAAAAACTCAAAAATGAATAAGCCTGATTTCAAATGAACCAT 3178
                    ************************************************************

4D1                 AGTCTGACAGTTAGAGCCGGGAGATTACACACAAACATTAACATTTTTAAATCTCGAAAA 3264
11C1                AGTCTGACAGTTAGAGCCGGGAGATTACACACAAACATTAACATTTTTAAATCTCGAAAA 3238
                    ************************************************************

4D1                 AGAAATAACTCACGAGTTGTTTCCAACCAAATTCAGAGAGTACAAGTTTGTTCTCTATAA 3324
11C1                AGAAATAACTCACGAGTTGTTTCCAACCAAATTCAGAGAGTACAAGTTTGTTCTCTATAA 3298
                    ************************************************************

4D1                 GACATCAGAACAGACTCTGGGGAAGAATTGGAGTAACTAGTAACTACTATCTGGGAAATG 3384
11C1                GACATCAGAACAGAGTCTGGGGAAGAATTGGAGTAACTAGTAACTACTATCTGGGAAATG 3358
                    ************** *********************************************
```

Figure 6.7 Structure of VIN3 gene (A) and sequence alignment of *VIN3* alleles from families 4D1 and 11C1(B). (A)The structure of the VIN3 gene is shown with exons being illustrated with blocks and darker blue sections indicating protein coding regions. The single SNP found in the protein coding region is shown with a red asterisk. (B)Identical sequences are highlighted in grey, SNPs are shown in red and deletions indicated in blue, start and stop codons are shown in yellow. Numbers indicate nucleotide position from the start of the sequence shown.

```
4D1             MQAASLSKIWRFDGNVGPENMDSSSFEDNECIETCKPNVLNVSERRELIHALSNQPEEAS 60
11C1            MQAASLSKIWRFDGNVGPENMDSSSFEDNECIETCKPNVLNVSERRELIHALSNQPEEAS 60
                ************************************************************

4D1             ELLNSWSRNEIMKIICAEMGKERKYTGLNKPKLIENLLNLVSRPLGETSCSDRRNSRKKE 120
11C1            ELLNSWSRNEIMKIICAEMGKERKYTGLNKPKLIENLLNLVSRPLGETSCSDRRNSRKKE 120
                ************************************************************

4D1             KKMIGYIICCENLACRAALGCDDTFCRRCSCCICQKFDDNKDPSLWLTCDACGSSCHLEC 180
11C1            KKMIGYIICCENLACRAALGCDDTFCRRCSCCICQKFDDNKDPSLWLTCDACGSSCHLEC 180
                ************************************************************

4D1             GLKQDRYGIGSDDLDGRFYCAYCGKDNDLLGCWRKQVKVAKETRRVDVLCYRLSLGQKLL 240
11C1            GLKQDRYGIGSDDLDGRFYCAYCGKDNDLLGCWRKQVKVAKETRRVDVLCYRLSLGQKLL 240
                ************************************************************

4D1             RGTTKYRNLLELMDEAVKKLEGDVGPLSGWAMKMARGIVNRLSSGVHVQKLCSQAMEALD 300
11C1            RGTTKYRNLLELMDEAVKKLEGDVGPLSGWAMKMARGIVNRLSSGVHVQKLCSQAMETLD 300
                *********************************************************:**

4D1             KVVSPSESVSGQGDKMTVRVEEIQARSVTVRVDSEEPSSSTQNKITGFRLFCRKSKDEEC 360
11C1            KVVSPSESVSGQGDKMTVRVEEIQARSVTVRVDSEEPSSSTQNKITGFRLFCRKSKDEEC 360
                ************************************************************

4D1             SSQGNCVVYLPETTSAIQGLEPDTEFCLRVVSFNEEGDLDESELRFTTLKDDGDEAGDQQ 420
11C1            SSQGNCVVYLPETTSAIQGLEPDTEFCLRVVSFNEEGDLDESELRFTTLKDDGDEAGDQQ 420
                ************************************************************

4D1             SPLTNSSSGLCSNPSLPEDESNNVNKSCSKGNGDKDNTEHCSAGEVESELEEERLVKRKA 480
11C1            SPLTNSSSGLCSNPSLPEDESNNVNKSCSKGNGDKDNTEHCSAGEVESELEEERLVKRKA 480
                ************************************************************

4D1             NKIDGRDLLVTPCKRDIYKGKQGGNKRFKSRTVSLNEKPEINNAANGVGDKDLGHIVKTI 540
11C1            NKIDGRDLLVTPCKRDIYKGKQGGNKRFKSRTVSLNEKPEINNAANGVGDKDLGHIVKTI 540
                ************************************************************

4D1             RCLEEEGHIDKSFRERFLTWYSLRATHREVRVVKIFVETFMEDLSSLGQQLVDTFSESIL 600
11C1            RCLEEEGHIDKSFRERFLTWYSLRATHREVRVVKIFVETFMEDLSSLGQQLVDTFSESIL 600
                ************************************************************

4D1             SKRSSTNGVVPAGICLKLWH 620
11C1            SKRSSTNGVVPAGICLKLWH 620
                ********************
```

Figure 6.8 Alignment of VIN3 amino acid sequences of 4D1 and 11C1. The protein sequences of the sequenced *VIN3* alleles, identical sequences are highlighted in grey, the presence of a substitution relative to the other sequence is identified in pink.

6.6 Variation in expression level of *VIN3* and its target *FLC*

Using qRT-PCR it was possible to assess variation in expression levels of *VIN3* and its target locus, *FLC*, between 4D1 and 11C1. Expression levels of the parental *VIN3* alleles were also examined in F1 plants.

4D1 had a substantially higher level of *FLC* expression than 11C1 under vernalized and unvernalized conditions (Fig. 6.8). Low *FLC* expression even under unvernalized conditions is consistent with the presence of non-functional *fri* allele in 11C1. It is also consistent with the low level of *FRI* expression seen in another family from location 11 (11A2) which also carried this *fri* allele (Chapter 4).

One explanation for the late flowering of 11C1 under unvernalized conditions is that vernalization is independent of *FLC* activity, because *FLC* expression is low in this family even before vernalization. However *FLC* expression in 11C1 is very sensitive to vernalization; one week of vernalization causes a vast reduction in expression of what *FLC* there is and this is consistent with saturation of the vernalization response of 11C1 within two weeks. This raises the alternative possibility that the low level of *FLC* in 11C1 is sufficient to delay flowering and the reduction of *FLC* is responsible for the vernalization response.

A reduction in *FLC* expression was also seen in 4D1. It was only apparent after two weeks of cold but the reduced level was still 4.7 times higher than the highest expression observed in 11C1.

The F1 population had little or no detectable *FLC* expression, even before vernalization, and this may explain why the F1 population flowered earlier than either parent. One explanation for the lack of *FLC* expression in the F1 is that the non-functional *fri* allele in 11C1 is dominant over the *FRI* allele from 4D1. However, this would require that 11C1 carries a recessive allele of a gene that is responsible for upregulating *FLC* to the degree seen in Figure 6.8.

The results obtained for *FLC* expression were consistent with those seen for *VIN3* (Fig. 6.9), increase in *VIN3* expression was observed in 4D1 until two weeks of cold treatment, corresponding to the reduction seen in *FLC* expression. In contrast, an increase in *VIN3* expression was seen in 11C1 after only one week of cold, corresponding with the decreased *FLC* expression by this time. This is therefore consistent with the idea that a difference in *VIN3* expression is responsible for the differences in the response of 11C1 and 4D1 to vernalization. The F1 population was observed to have increased *VIN3* expression after one week of vernalization, but less than that observed for 11C1. After two weeks of cold the expression level of *VIN3* was higher in the F1 population than in either of the parental lines.

One thing to note was that accurate estimates of *VIN3* expression could not be obtained for 4D1 under unvernalized conditions, because they were very low. Therefore samples of cDNA from 4D1, 11C1 and the F1 were amplified using *VIN3* specific primers and the products separated in an agarose gel. No expression of *VIN3* could be detected in 4D1 prior to vernalization in several repetitions of this experiment, although expression was consistently detected in 11C1 and the F1 hybrid.

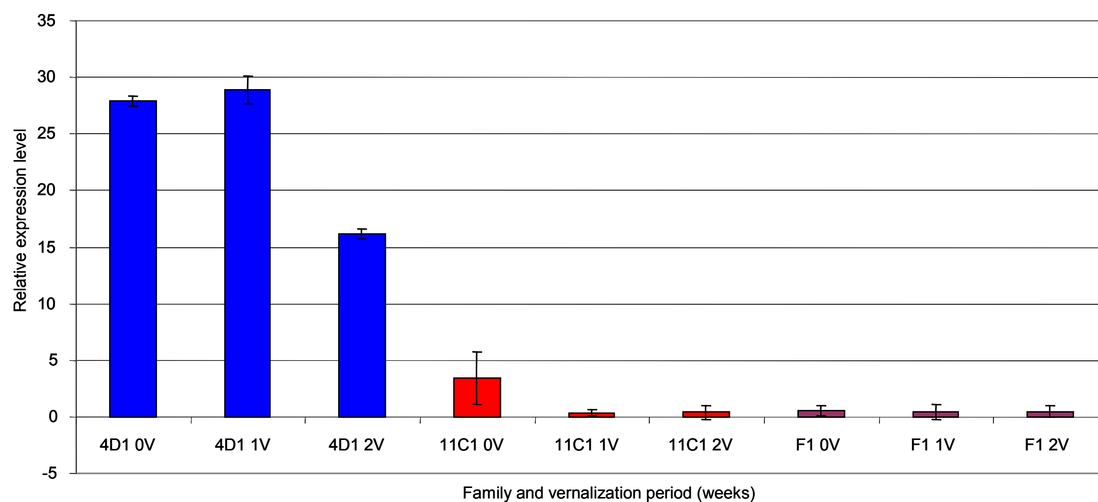


Figure 6.9 The relative expression levels of *FLC* in the parental lines 4D1 and 11C1 and the F1 generation of a hybrid mapping population. 4D1 is shown in blue, 11C1 is shown in red and F1 lines are indicated in purple. Relative expression levels were estimated using qRT-PCR and are shown ± their standard errors.

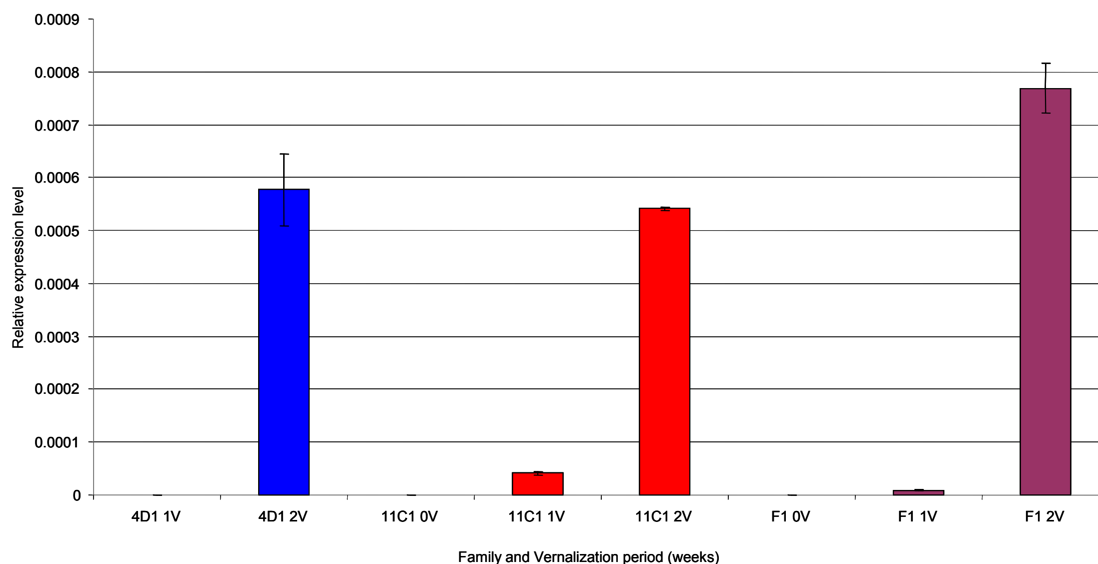


Figure 6.10 The relative expression levels of *VIN3* in the parental lines 4D1 and 11C1 and the F1 generation of a hybrid mapping population. 4D1 is shown in blue, 11C1 in red and F1 lines are indicated in purple. Relative expression levels were estimated using qRT-PCR and are shown ± their standard errors

6.7 Discussion

No apparent QTL at *FRI* or *FLC* was identified in the work described in this chapter. This is interesting given the likely *fri* loss of function allele identified in parent 11C1. Previous studies have found that *fri* loss of function alleles have a major effect on flowering time variation in natural populations of *Arabidopsis thaliana* (Gazzani et al., 2003; Johanson et al., 2000).

Variation in *FLC* expression was also found; 11C1 family members had a much lower level of *FLC* expression relative to 4D1. It is worth noting however that although *FLC* expression was very low, it still responded to vernalization in a dose dependent manner. No QTL was found near *FLC* therefore the observed difference in *FLC* expression is unlikely to be due to variation in *FLC* itself. It also seems unlikely that it reflects differences in *FRI* function, because no QTL was detected at *FRI* and the low levels of *FLC* expression did not behave as a recessive phenotype in the F1, as would be expected if a *fri* loss of function mutation was responsible.

Although work in this chapter goes some way to propose reasons for the vernalization response observed in population 11 it does not provide an explanation for why this family should be so late flowering without vernalization. Some F4 families were found to be very vernalization sensitive but homozygous for markers from 4D1 at the bottom of chromosome five. This suggests that genes elsewhere in the genome contribute to variation in the vernalization response. Potentially a gene acting downstream of *FRI* or epistatically to *FRI* might repress the *fri* mutant phenotype and confer a vernalization response. Also as discussed in chapter 4, disruption in an autonomous pathway component could confer a vernalization requirement in a plant that carries a *fri* loss of function allele. This could be something interesting to explore in future experiments.

Studies have shown that natural variation in flowering time can often be attributed to a few major effect QTL (Salome et al., 2011; Strange et al., 2011). These include, amongst others, *FRI*, *FLC* and genes at the bottom of chromosome five. This study

found a correlation between flowering time (under vernalized conditions) and a marker located in this region.

VIN3 is a well characterised component of the vernalization pathway is located at the bottom of chromosome five. *VIN3* was identified in a mutagenesis screen performed on Col-0 introgressed with a functional *FRI* allele as necessary for a vernalization response and silencing of *FLC* (Sung and Amasino, 2004; Wood et al., 2006).

Given that a correlation between genotype in this region and flowering time is seen only under vernalized conditions, it is likely the gene will be involved in vernalization and as such *VIN3* appeared to be a likely candidate.

Previous work examining the basis for differences in the vernalization responses of Swedish accessions, Col-0 and the Scottish accession Edi-0 concluded that natural variation in *FLC* silencing was not the result of variation in *VIN3* or the other well characterised components of the vernalization pathway, *VRN1* and *VRN2* (Shindo et al., 2006). This was suggested because no significant QTL was identified in the proximity of *VIN3* in hybrids between these lines and no significant sequence variation was detected in *VIN3*.

However significant QTL at the bottom of chromosome five have been found in several other studies (Caicedo et al., 2009; El-Lithy et al., 2004; El-Lithy et al., 2006; O'Neill et al., 2007; Salome et al., 2011; Simon et al., 2008; Strange et al., 2011; Ungerer et al., 2002) and so it is possible that natural variation in this region is important for differences in flowering behaviour elsewhere.

Variation in the timing of *VIN3* induction in response to cold was detected between 11C1 and 4D1. The earlier initiation of *VIN3* in 11C1 corresponded with an early decrease in *FLC* expression consistent with the more rapid saturation of the vernalization response observed in 11C1, although it is currently not known whether this repression of *FLC* is stable.

There are other candidate flowering time genes in this region, including the tandem array of *MAF2-MAF5*, which span approximately 24 kb to the south of *VIN3*. These genes appear to be susceptible to variation through deletion, duplication and recombinations that can give rise to chimeric genes (Rosloski et al., 2010). They are therefore likely candidates for natural variation in flowering time and a correlation between natural variants of *MAF2* and variation in flowering time has been reported (Caicedo et al., 2009). Loss of function mutations in *MAF2* suggest that this gene has a role in preventing floral initiation, independently of *FLC*, in response to short periods of cold (Caicedo et al., 2009; Rosloski et al., 2010). This is consistent with families from location 11 having reduced MAF2 activity, because they have a vernalization requirement that is saturated after two weeks of cold. However *maf2* mutants are early flowering in the absence of vernalization, which is not the case for families from location 11 which are among the latest flowering local accessions without vernalization.

Plants carrying chromosomes in which *VIN3* alleles from 11C1 are separated from other candidate genes and transgenic lines which carry either the 11C1 or 4D1 alleles of *VIN3* or reciprocal combination of their promoter and coding regions should clarify the role of *VIN3* in local variation.

7 Concluding Remarks

This thesis has demonstrated that there is extensive genetically determined variation in flowering time between local Scottish accessions of *Arabidopsis thaliana*, on a similar scale to that observed worldwide (Gazzani et al., 2003; Johanson et al., 2000; Shindo et al., 2005). In addition to this, variation has also been found locally in the timing of floral induction in response to photoperiod and to vernalization. Vernalization responses showed more continuous variation than photoperiod responses, as they do worldwide. Therefore this study has focussed predominantly on variation in response to vernalization.

Differences were found in the sequences of the *FRI* and *FLC* genes, which underlie much of the variation in flowering behaviour between worldwide accessions. However variation in these genes can explain only part of the variation in flowering behaviour observed locally.

Two mapping populations were produced from families shown to differ in their flowering behaviour. The first was created from parents that have different responses to vernalization and to photoperiod. QTL analysis on this population suggested that the early flowering, vernalization insensitive behaviour of the 1E1 parent could be attributed to variation in the region of *FRI*. *FRI* shows extensive sequence and functional variation worldwide. However I did not detect potential loss of function mutations in the *FRI* coding region in the early flowering 1E1 family, though *FRI* was expressed at a low level in 1E1 (compared to the *fri* mutant, Col-0) and no significant *FLC* expression was detected in 1E1, consistent with a *fri* loss of function mutation in families from location 1. Failure to identify a *fri* polymorphism at location 1 corresponding to a loss of function mutation does not rule out a mutation in another region of the gene that was not sequenced or epigenetic variation that would not be detected by DNA sequencing.

This mapping population was found to segregate a phenotype in which flowering would not occur without vernalization. The F1 generation had also been found to

flower significantly later than either of its parents. This suggests that the genes responsible for late flowering must have been contributed by both parents and that late flowering was determined by dominant genes of these loci. However the 1E1 parent flowered early, with or without vernalization, suggesting that the effects of its late flowering loci were suppressed. If 1E1 carries a *fri* mutation, these loci might act downstream of *FRI*, so that their effects are suppressed by the *fri* mutation. A correlation was also found between the extreme late flowering phenotype and a marker on chromosome five (NGA139). This region contains the *HUA2* gene that is involved in RNA processing and delays flowering by enhancing the repression of genes that delay flowering (Chen and Meyerowitz, 1999; Cheng et al., 2003). A naturally occurring allele of *HUA2*, found in the Scottish accession Sy-0 from the Isle of Skye, was shown to result in extremely late flowering (Grbic and Bleecker, 1996; Wang et al., 2007), making *HUA2* a likely candidate for the variation observed between the parents of this cross. Members of this mapping population also appeared to be insensitive to photoperiod, showed secondary growth and, under field conditions, behaved as perennials. Candidate genes for this behaviour include *SOC1*, because *soc1* mutants can cause perennial flowering behaviour and secondary growth (Melzer et al., 2008).

The second mapping population used for QTL analysis involved one parent that was very vernalization sensitive and surprisingly carried a likely *fri* loss of function allele and a second parent that was significantly less sensitive to vernalization. A QTL screen found no significant correlation between flowering time and *FRI* or *FLC* genotype. The 11C1 parent of this population had a much lower level of *FLC* expression compared to 4D1, presumably as a result of the *fri* loss of *FRI* function. However this family still had an extreme vernalization response which saturated rapidly and this was reflected in *FLC* expression which was reduced in a dose dependent manner in response to cold. Therefore it is possible that *FLC* is being upregulated, all be it to a low level, by a locus other than *FRI* and that this level of *FLC* is sufficient to repress flowering in this genetic background. Alternatively, the vernalization response seen in this population may be independent of *FLC*.

A correlation was found between flowering time (after vernalization) and the genotype at the bottom of chromosome five. I identified *VIN3* as a candidate locus and sequence variation in the promoter and subsequent expression variation after short periods of cold were consistent with this hypothesis. Additionally the *FLC* expression levels coordinated with *VIN3* expression levels, i.e. an increase of *VIN3* expression after one week of cold corresponded with a reduction of *FLC* expression after the same cold treatment. However *VIN3* is not the only potentially relevant gene at the bottom of chromosome five, and the tandem gene array of *MAF2 – MAF5* is also found in this region. The *MAF* genes are hypervariable and prone to rearrangement. Natural variation was found in *MAF2*, a gene believed to prevent flowering in response to short periods of cold, consistent with involvement of *MAF* gene variation locally. However plants carrying non functional alleles of *MAF2* are extremely early flowering without vernalization (Caicedo et al., 2009), in contrast to families from location 11.

Transgenic lines have been produced that carry the 11C1, 4D1 or reciprocal combinations of promoter and coding regions from each parent. These constructs were transformed in plants with a *vin3-1* mutant background (Sung and Amasino, 2004). This work is ongoing and will hopefully reveal whether *VIN3* is relevant to the differences in flowering behaviour seen between 11C1 and 4D1.

Consistent with studies of worldwide *Arabidopsis thaliana* accessions this study has found significant correlations between flowering time and genotype in only a small number of chromosomal regions (Salome et al., 2011; Strange et al., 2011). As has been seen in worldwide studies this work also found sequence and expression variation in *FRI* and *FLC* however these loci can account for only a small proportion of the flowering time variation found locally. Additionally, as with worldwide studies, QTL analysis also suggests that *HUA2* is a candidate for the behaviour observed in at least one family. However this study has found naturally occurring variation in *VIN3* which is potentially important. A role for *VIN3* in natural variation has not previously been reported.

7.1 Future Work

In order to further test whether *VIN3* underlies differences in the phenotypes of 11C1 and 4D1, a series of constructs of complementation of a *vin3* mutation have been made. The promoter and coding regions of *VIN3* were amplified separately with an overlap in part of the 5' UTR that did not contain any polymorphisms between the parents. Overlap-PCR was then used to fuse promoter and coding regions; the 11C1 promoter with its own coding sequence or the coding sequence of 4D1 and the 4D1 promoter with either the 4D1 or 11C1 coding sequence (Chapter 2). The four fusions were sequenced to ensure that no mutations had been introduced by PCR and inserted into a T-DNA binary vector (Chapter 2). *Agrobacterium* cultures carrying these constructs have been used to dip *vin3-1* mutants, which have a Col-0, *FRI*$^{+}$ background, have a deletion in *vin3* but no T-DNA that might trigger transgene silencing. I intend to select transgenic T0 plants and to determine their flowering times, with or without vernalization. This will test whether the *VIN3* allele from 4D1 and 11C1 differ in activity in the same genetic background and whether polymorphisms in the promoter or coding region are responsible.

I will also quantify the relative levels of transcripts from each parental allele in F1 plants, by direct sequencing of RT-PCR products carrying several SNPs within exons. If differences in the expression levels of the two alleles are found in the genetic background of the F1, it will suggest that the *cis*-acting regulatory mutations are responsible for the different levels of *VIN3* expression observed.

Another interesting issue that would be nice to resolve is that of the non functional *fri* allele in the very vernalization sensitive families from location 11. Firstly it would be useful to track this allele in further generations of the mapping population. This could be done by designing dCAPs markers specifically at the site where the point mutation has occurred.

Also it would be beneficial to confirm that the SNP found in families from location 11 does result in NMD. This could be done by produicing a hybrid population from a

representative from a family 11 with a plant mutant for NMD, for example an *upf* mutant plant (Arciga-Reyes et al., 2006).

Also the hybrid mapping populations could have been produced by crossing a members of the locally collected families with the lab strain Col-0. The benefit of this would be that the genome sequence of one of the parents would be known and much is already known about the alleles of flowering time genes in Col-0. This would have meant that variation in flowering time in the mapping population from what would be expected from Col-0 could reasonably be assumed to be as a result alleles present in the local accession. However by making the crosses with both parents from local accessions, this meant that a greater amount of genetic variation between the local accessions could be assessed.

Finally a sampling strategy that allowed for more accurate comparison between the flowering behaviours of plants shown to behave as summer or winter annual in the wild would allow me to address some of the other interesting questions raised in this study. Firstly having a winter and summer annual collection from the same sites with more detail about the microclimate and environment of each collection site would give a greater insight into locally derived flowering time variation. Additionally reciprocal transplant experiments between sites would go some way to further explore the question of local adaptation within these collections.

8 References Cited

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

Andersson,C.R., Helliwell,C.A., Bagnall,D.J., Hughes,T.P., Finnegan,E.J., Peacock,W.J., and Dennis,E.S. (2008). The FLX gene of Arabidopsis is required for FRI-dependent activation of FLC expression. Plant Cell Physiol *49*, 191-200.

Arciga-Reyes,L., Wootton,L., Kieffer,M., and Davies,B. (2006). UPF1 is required for nonsense-mediated mRNA decay (NMD) and RNAi in Arabidopsis. Plant J *47*, 480-489.

Ausin,I., Alonso-Blanco,C., Jarillo,J.A., Ruiz-Garcia,L., and Martinez-Zapater,J.M. (2004). Regulation of flowering time by *FVE*, a retinoblastoma-associated protein. Nat. Genet. *36*, 162-166.

Balasubramanian,S., Sureshkumar,S., Agrawal,M., Michael,T.P., Wessinger,C., Maloof,J.N., Clark,R., Warthmann,N., Chory,J., and Weigel,D. (2006). The PHYTOCHROME C photoreceptor gene mediates natural variation in flowering and growth responses of Arabidopsis thaliana. Nat. Genet. *38*, 711-715.

Baudry,A., Ito,S., Song,Y.H., Strait,A.A., Kiba,T., Lu,S., Henriques,R., Pruneda-Paz,J.L., Chua,N.H., Tobin,E.M., Kay,S.A., and Imaizumi,T. (2010). F-box proteins FKF1 and LKP2 act in concert with ZEITLUPE to control Arabidopsis clock progression. Plant Cell *22*, 606-622.

Bentsink,L., Jowett,J., Hanhart,C.J., and Koornneef,M. (2006). Cloning *DOG1* a quantitative trait locus controlling seed dormancy. Proc Natl. Acad. Sci U. S. A *103*, 17042-17047.

Bertani, G. Studies in lysogenesis .I. The mode of phage liberation by lysogenic *Escherichia coli*. Journal of Bacteriology [62], 293-300. 1951.
Ref Type: Generic

Brachi,B., Faure,N., Horton,M., Flahauw,E., Vazquez,A., Nordborg,M., Bergelson,J., Cuguen,J., and Roux,F. (2010). Linkage and association mapping of Arabidopsis thaliana flowering time in nature. PLoS. Genet. *6*, e1000940.

Brock,M.T., Dechaine,J.M., Iniguez-Luy,F.L., Maloof,J.N., Stinchcombe,J.R., and Weinig,C. (2010). Floral genetic architecture: an examination of QTL architecture underlying floral (co)variation across environments. Genetics *186*, 1451-1465.

Caicedo,A.L., Richards,C., Ehrenreich,I.M., and Purugganan,M.D. (2009). Complex rearrangements Lead to Novel Chimeric Gene Fusion Polymorphisms at the

Arabidopsis thaliana MAF2-5 Flowering Time Gene Cluster. Mol. Biol. Evol. *26*, 699-711.

Caicedo,A.L., Stinchcombe,J.R., Olsen,K.M., Schmitt,J., and Purugganan,M.D. (2004). Epistatic interaction between Arabidopsis FRI and FLC flowering time genes generates a latitudinal cline in a life history trait. Proc Natl. Acad. Sci U. S. A *101*, 15670-15675.

Chandler,J., Wilson,A., and Dean,C. (1996). Arabidopsis mutants showing an altered response to vernalization. Plant J. *10*, 637-644.

Chen,X. and Meyerowitz,E.M. (1999). HUA1 and HUA2 are two members of the floral homeotic AGAMOUS pathway. Mol. Cell. *3*, 349-360.

Cheng,Y., Kato,N., Wang,W., Li,J., and Chen,X. (2003). Two RNA binding proteins, HEN4 and HUA2 act in the processing of AGAMOUS pre-mRNA in Arabidopsis thaliana. Dev. Cell. *4*, 53-66.

Choi,K., Park,C., Lee,J., Oh,M., Noh,B., and Lee,I. (2007). Arabidopsis homologs of components of the SWR1 complex regulate flowering and plant development. Development *134*, 1931-1941.

Doyle, JJ and Doyle JL. Isolation of plant DNA from fresh tissue. Focus 12, 13-15. 1990.
Ref Type: Generic

Dubnau,J. and Struhl,G. (1996). RNA recognition and translational regulation by a homeodomain protein. Nature *379*, 694-699.

El-Din El-Assal,S., Alonso-Blanco,C., Peeters,A.J., Raz,V., and Koornneef,M. (2001). A QTL for flowering time in Arabidopsis reveals a novel allele of CRY2. Nat. Genet. *29*, 435-440.

El-Lithy,M.E., Clerkx,E.J.M., and Ruys,G.J. (2006). New Arabidopsis recombinant inbred line populations genotyped usig SNPWave and their use for mapping flowering-time quantitative trait loci. Genetics *172*, 1867-1876.

El-Lithy,M.E., Clerkx,E.J.M., Ruys,G.J., Koornneef,M., and Vreugdenhil,D. (2004). Quantitative Trait Locus Analysis of Growth-Related Traits in a New Arabidopsis Reombinant Inbred Population. Plant Physiol *135*, 444-458.

Eriksson,S., Bohlenius,H., Moritz,T., and Nilsson,O. (2006). GA4 is the active gibberellin in the regulation of LEAFY transcription and Arabidopsis floral initiation. Plant Cell *18*, 2172-2181.

Falconer,D.S. and Mackay,T.F. (1996). Introduction to Quantitative Genetics. (Essex: Addison wesley Longman).

Filiault,D.L., Wessinger,C.A., Dinneny,J.R., Lutes,J., Borevitz,J.O., Weigel,D., Chory,J., and Maloof,J.N. (2008). Amino acid polymorphisms in Arabidopsis

phytochrome B cause differential responses to light. Proc Natl. Acad. Sci U. S. A *105*, 3157-3162.

Fu,D., Dunbar,M., and Dubcovsky,J. (2006). Wheat *VIN3*-like PHD finger genes are up-regulated by vernalization. Molecular Genetics and Genomics *277*, 301-313.

Fujiwara,S., Wang,L., Han,L., Suh,S.S., Salome,P.A., McClung,C., and Somers,D.E. (2008). Post-translational regulation of the *Arabidopsis* circadian clock through selective proteolysis and phosphorylation of the pseudo-response regulator proteins. J. Biol. Chem *283*, 23073-23083.

Gassner,G. (1918). Beiträge zur physiologischen Charakteristik sommer- und winterannueller Gewächse, insbesondere der Getreidepflanzen. Z Bot *10*, 417-480.

Gazzani,S., Gendall,A.R., Lister,C., and Dean,C. (2003). Analysis of the molecular basis of flowering time variation in Arabidopsis accessions. Plant Physiol *132*, 1107-1114.

Gendall,A.R., Levy,Y.Y., Wilson,A., and Dean,C. (2001). The VERNALIZATION 2 gene mediates the epigenetic regulation of vernalization in Arabidopsis. Cell *107*, 525-535.

Grbic,B. and Bleecker,A.B. (1996). An altered body plan is conferred on Arabidopsis plants carrying dominant alleles of two genes. Development *122*, 2395-2403.

Greenup,A., Peacock,W.J., Dennis,E.S., and Trevaskis,B. (2009). The molecular biology of seasonal flowering-responses in Arabidopsis and the cereals. Ann Bot *103*, 1165-1172.

Griffiths,J., Murase,K., Rieu,I., Zentella,R., Zhang,Z.L., Powers,S.J., Gong,F., Phillips,A.L., Hedden,P., Sun,T.P., and Thomas,S.G. (2006). Genetic characterization and functional analysis of the GID1 gibberellin receptors in Arabidopsis. Plant Cell *18*, 3399-3414.

Hanada,K., Zhang,X., Borevitz,J.O., Li,W.H., and Shiu,S.H. (2007). A large number of novel coding small open reading frames in the intergenic regions of the Arabidopsis thaliana genome are transcribed and/or under purifying selection. Genome Res. *17*, 632-640.

Hanahan, D. Studies on transformation of *Escherichia coli* with plasmids. Journal of Molecular Biology 166[4], 557-580. 1983.
Ref Type: Generic

Harmer,S.L. (2009). The Circadian System in Higher Plants. Annu. Rev. Plant Biol *60*, 357-377.

He,Y., Doyle,M.R., and Amasino,R.M. (2004). PAF1-complex-mediated histone methylation of FLOWERING LOCUS C chromatin is required for the vernalization-responsive, winter-annual habit in Arabidopsis. Genes Dev. *18*, 2774-2784.

He,Y., Michaels,S.D., and Amasino,R.M. (2003). Regulation of flowering time by histone acetylation in *Arabidopsis*. Science *302*, 1751-1754.

Helliwell,C.A., Robertson,M., Finnegan,E.J., Buzas,D.M., and Dennis,E.S. (2011). Vernalization-repression of Arabidopsis FLC requires promoter sequences but not antisense transcripts. PLoS one *6*, e21513.

Heo,J.B. and Sung,S. (2011). Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. Science *331*, 76-79.

Hisamatsu,T. and King,R. (2008). The nature of floral signals in *Arabidopsis*. II. Roles for FLOWERING LOCUS T (FT) and gibberellin. J. Exp. Bot. *59*, 3821-3829.

Hornyik,C., Duc,C., Rataj,K., Terzi,L.C., and Simpson,G.G. (2010). Alternative polyadenylation of antisense RNAs and flowering time control. Biochem. Soc. Trans. *38*, 1077-1081.

Ishikawa,M., Kiba,T., and Chua,N.H. (2006). The Arabidopsis SPA1 gene is required for circadian clock function and photoperiodic flowering. Plant J. *46*, 736-746.

Isken,O. and Maquat,L.E. (2008). The multiple lives of NMD factors: balancing roles in gene and genome regulation. Nat. Rev. Genet. *9*, 699-712.

Johanson,U., West,J., Lister,C., Michaels,S., Amasino,R., and Dean,C. (2000). Molecular analysis of FRIGIDA, a major determinant of natural variation in Arabidopsis flowering time. Science *290*, 344-347.

Kardailsky,I., Shukla,V.K., Ahn,J.H., Dagenais,N., Christensen,S.K., Nguyen,J.T., Chory,J., Harrison,M.J., and Weigel,D. (1999). Activation tagging of the floral inducer FT. Science *286*, 1962-1965.

Kim,S., Choi,K., Park,C., Hwang,H.J., and Lee,I. (2006). SUPPRESSOR OF FRIGIDA4, encoding a C2H2-Type zinc finger protein, represses flowering by transcriptional activation of Arabidopsis FLOWERING LOCUS C. Plant Cell *18*, 2985-2998.

King,R.W. and Zeevaart,J.A.D. (1973). Floral stimulus movement in *Perilla* and flower inhibition caused by noninduced leaves. Plant Physiol *51*, 727-738.

Knott,J. (1934). Effect of a localized photoperiod on spinach. Proc Soc Hortic Sci 152-154.

Kobayashi,Y. and Weigel,D. (2007). Move on up, it's time for change--mobile signals controlling photoperiod-dependent flowering. Genes Dev. *21*, 2371-2384.

Korves,T.M., Schmid,K.J., Caicedo,A.L., Mays,C., Stinchcombe,J.R., Purugganan,M.D., and Schmitt,J. (2007). Fitness effects associated with the major flowering time gene FRIGIDA in Arabidopsis thaliana in the field. Am. Nat. *169*, E141-E157.

Laibach,F. (1943). Arabidopsis thaliana (L.) Heynh. als Objekt für genetische und entwicklungsphysiologische Untersuchungen. Bot Arch *44*, 439-455.

Lang,A. (1957). The effect of gibberellin upon flower formation. Proc Natl. Acad. Sci U. S. A *43*, 709-717.

Lang,A., Chailakhyan,K.H., and Frolova,I.A. (1977). Promotion and inhibition of flower formation in a day neutral plant in grafts with a short-day plant and a long-day plant. Proc Natl. Acad. Sci U. S. A *74*, 2412-2416.

Laubinger,S., Marchal,V., Le,G.J., Wenkel,S., Adrian,J., Jang,S., Kulajta,C., Braun,H., Coupland,G., and Hoecker,U. (2006). Arabidopsis SPA proteins regulate photoperiodic flowering and interact with the floral inducer CONSTANS to regulate its stability. Development *133*, 3213-3222.

Le,C., V, Roux,F., and Reboud,X. (2002). DNA polymorphism at the FRIGIDA gene in Arabidopsis thaliana: extensive nonsynonymous variation is consistent with local selection for flowering time. Mol. Biol. Evol. *19*, 1261-1271.

Lee,I., Aukerman,M.J., Gore,S.L., Lohman,K.N., Michaels,S.D., Weaver,L.M., John,M.C., Feldmann,K.A., and Amasino,R.M. (1994). Isolation of *LUMINIDEPENDENS*: a gene involved in the control of flowering time in *Arabidopsis*. Plant Cell *6*, 75-83.

Lee,J. and Lee,I. (2010). Regulation and function of SOC1, a flowering pathway integrator. J. Exp. Bot. *61*, 2247-2254.

Lee,J.H., Park,S.H., Lee,J.S., and Ahn,J.H. (2007). A conserved role of SHORT VEGETATIVE PHASE (SVP) in controlling flowering time of Brassica plants. Biochim. Biophys. Acta *1769*, 455-461.

Lempe,J., Balasubramanian,S., Sureshkumar,S., Singh,A., Schmid,M., and Weigel,D. (2005). Diversity of flowering responses in wild Arabidopsis thaliana strains. PLoS. Genet. *1*, 109-118.

Levy,Y.Y., Mesnage,S., Mylne,J., Gendall,A.R., and Dean,C. (2002). Multiple roles of *Arabidopsis* VRN1 in vernalization and flowering time. Science *297*, 243-246.

Li,D., Shen,L., Wu,Y., Chen,H., Robertson,M., Helliwell,C.A., Ito,T., Meyerowitz,E., and Yu,H. (2008). A repressor complex governs the integration of floweing signals in Arabidopsis. Dev. Cell. *15*, 110-120.

Lim,M.H., Kim,J., Kim,Y.S., Chung,K.S., Seo,Y.H., Lee,I., Kim,J., Hong,C.B., Kim,H.J., and Park,C.M. (2004). A new *Arabidopsis* gene, *FLK,* encodes an RNA binding protein with K homology motifs and regulates flowering time via FLOWERING LOCUS C. Plant Cell *16*, 731-740.

Liu,F., Quesada,V., Crevillen,P., Baurle,I., Swiezewski,S., and Dean,C. (2007). The Arabidopsis RNA-binding protein FCA requires a lysine-specific demethylase 1 homolog to downregulate FLC. Mol. Cell *28*, 398-407.

Liu,J., He,Y., Amasino,R., and Chen,X. (2004). siRNAs targeting an intronic transposon in the regulation of natural flowering behaviour in Arabidopsis. Genes Dev. *18*, 2873-2878.

Liu,L.J., Zhang,Y.C., Li,Q.H., Sang,Y., Mao,J., Lian,H.L., Wang,L., and Yang,H.Q. (2008). COP1-mediated ubiquitination of CONSTANS is implicated in crytptochrome regulation of flowering in *Arabidopsis*. Plant Cell *20*, 292-306.

March-Diaz,R., Garcia-Dominguez,M., Florencio,F.J., and Reyes,J.C. (2007). SEF, a new protein required for flowering repression in Arabidopsis, interacts with PIE1 and ARP6. Plant Physiol *143*, 893-901.

Martin-Trillo,M., Lazaro,A., Poethig,R.S., Gomez-Mena,C., Pineiro,M.A., Martinez-Zapater,J.M., and Jarillo,J.A. (2006). EARLY IN SHORT DAYS 1 (ESD1) encodes ACTIN-RELATED PROTEIN 6 (AtARP6), a putative component of chromatin remodelling complexes that positively regulates FLC accumulation in Arabidopsis. Development *133*, 1241-1252.

Mathieu,J., Warthmann,N., Küttner,F., and Schmid,M. (2007). Export of FT Protein from Phloem Companion Cells Is Sufficient for Floral Induction in *Arabidopsis*. Curr. Biol. *17*, 1055-1060.

Melzer,S., Lens,F., Gennen,J., Vanneste,S., Rohde,A., and Beeckman,T. (2008). Flowering-time genes modulate meristem determinacy and growth form in Arabidopsis thaliana. Nat. Genet. *40*, 1489-1492.

Michaels,S.D., Ditta,G., Gustafson-Brown,C., Pelaz,S., Yanofsky,M., and Amasino,R.M. (2003a). AGL24 acts as a promoter of flowering in Arabidopsis and is positively regulated by vernalization. Plant J. *33*, 867-874.

Michaels,S.D., He,Y., Scortecci,K.C., and Amasino,R.M. (2003b). Attenuation of FLOWERING LOCUS C activity as a mechanism for the evolution of summer-annual flowering behavior in Arabidopsis. Proc Natl. Acad. Sci U. S. A *100*, 10102-10107.

Michaels,S.D., He,Y., Scortecci,K.C., and Amasino,R.M. (2003c). Attenuation of FLOWERING LOCUS C activity as a mechanism for the evolution of summer-annual flowering behavior in Arabidopsis. Proc. Natl. Acad. Sci. U. S. A *100*, 10102-10107.

Monte,E., Alonso,J.M., Ecker,J.R., Zhang,Y., Li,X., Young,J., Austin-Phillips,S., and Quail,P.H. (2003). Isolation and characterization of phyC mutants in Arabidopsis reveals complex crosstalk between phytochrome signaling pathways. Plant Cell *15*, 1962-1980.

Moon,J., Suh,S.S., Lee,H., Choi,K.R., Hong,C.B., Paek,N.C., Kim,S.G., and Lee,I. (2003). The SOC1 MADS-box gene integrates vernalization and gibberellin signals for flowering in Arabidopsis. Plant J. *35*, 613-623.

Mouradov,A., Cremer,F., and Coupland,G. (2002). Control of Flowering Time: Interacting Pathways as a Basis for Diversity. The Plant Cell 111-113.

Murashige, T and Skoog F. A revised medium for rapid growth and bioassays with tobacco tissue cultures. Plant Physiology 15[3], 473-497. 1962.
Ref Type: Generic

Nakagawa,T., Takayuki,K., Hino,T., Tanaka,K., Kawamukai,M., Niwa,Y., Toyooka,K., Matsuoka,K., Jinbo,T., and Kimura,T. (2007). Development of Series of Gateway Binary Vectors, pGWBs, for Realizaing Efficient Construction of Fusion Genes for Plant Transformation. Journal of Bioscience and Bioengineering *104*, 34-41.

Noh,B., Lee,S.H., Kim,H.J., Yi,G., Shin,E.A., Lee,M., Jung,K.J., Doyle,M.R., Amasino,R.M., and Noh,Y.S. (2004). Divergent roles of a pair of homologous jumonji/zinc-finger-class transcription factor proteins in the regulation Arabidopsis flowering time. The Plant Cell *16*, 2601-2613.

Noh,Y.S. and Amasino,R.M. (2003). PIE1, an ISWI family gene, is required for FLC activation and floral repression in Arabidopsis. Plant Cell *15*, 1671-1682.

Nordborg,M. and Weigel,D. (2008). Next-generation genetics in plants. Nature *456*, 720-723.

O'Neill,C.M., Morgan,C., Kirby,J., Tschoep,H., Deng,P.X., Brennan,H., Rosas,U., Fraser,F., Hall,C., Gill,S., and Bancroft,I. (2007). Six new recombinant inbred populations for the study of quantitative traits in Arabidopsis thaliana. Theor. Appl. Genet. *116*, 623-634.

Oh,S., Zhang,H., Ludwig,P., and van,N.S. (2004). A mechanism related to the yeast transcriptional regulator Paf1c is required for expression of the Arabidopsis FLC/MAF MADS box gene family. Plant Cell *16*, 2940-2953.

Pittendrigh,C.S. (1960). Circadian rhythms and the circadian organization of living systems. Cold Spring Harb. Symp. Quant. Biol. *25*, 159-184.

Pittendrigh,C.S. (1972). Circadian surfaces and the diversity of possible roles of circadian organization in photoperiodic induction. Proc. Natl. Acad. Sci. U. S. A *69*, 2734-2737.

Platt,A., Horton,M., Huang,Y.S., Li,Y., Anastasio,A.E., Mulyati,N.W., Agren,J., Bossdorf,O., Byers,D., Donohue,K., Dunning,M., Holub,E.B., Hudson,A., Le,C., V, Loudet,O., Roux,F., Warthmann,N., Weigel,D., Rivero,L., Scholl,R., Nordborg,M., Bergelson,J., and Borevitz,J.O. (2010). The scale of population structure in Arabidopsis thaliana. PLoS. Genet. *6*, e1000843.

Purvis,O.N. and Gregory,F.G. (2011). The reversibility by high temperature of the vernalised condition in Petkus winter rye. Ann Bot *16*, 1-21.

Putterill,J., Robson,F., Lee,K., Simon,R., and Coupland,G. (1995). The CONSTANS gene of *Arabidopsis* promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. Cell *80*, 847-857.

Quesada,V., Macknight,R., Dean,C., and Simpson,G.G. (2003). Autoregulation of FCA pre-mRNA processing controls *Arabidopsis* flowering time. EMBO J *22*, 3142-3152.

Ratcliffe,O.J., Kumimoto,R.W., Wong,B.J., and Riechmann,J.L. (2003). Analysis of the *MADS AFFECTING FLOWERING* gene family:MAF2 prevents vernalization by short periods of cold. The Plant Cell *15*, 1159-1169.

Ratcliffe,O.J., Nadzan,G.C., Reuber,T.L., and Reichmann,J.L. (2001). Regulation of flowering in Arabidopsis by an *FLC* homologue. Plant Physiol *126*, 122-132.

Rinn,J.L., Kertesz,M., Wang,J.K., Squazzo,S.L., Xu,X., Brugmann,S.A., Goodnough,L.H., Helms,J.A., Farnham,P.J., Segal,E., and Chang,H.Y. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell *129*, 1311-1323.

Rivera-Pomar,R., Niessing,D., Schmitt-Ott,U., Gehring,W.J., and Jackle,H. (1996). RNA binding and translational suppression by bicoid. Nature *379*, 746-749.

Rosloski,S., Jali S.S, Balasubramanian,S., Weigel,D., and Grbic,V. (2010). Natural Diversity in Flowering Responses of Arabidopsis thaliana Caused by Variation in a Tandem Gene Array. Genetics *186*, 263-276.

Salisbury,F.B. (1985). Photoperiodism. Hortic. Rev (Am Soc Hortic Sci) *4*, 66-105.

Salome,P.A., Bomblies,K., Laitinen,R.A., Yant,L., Mott,R., and Weigel,D. (2011). Genetic architecture of flowering-time variation in Arabidopsis thaliana. Genetics *188*, 421-433.

Samis,K.E., Heath,K.D., and Stinchcombe,J.R. (2008). Discordant longitudinal clines in flowering time and phytochrome C in Arabidopsis thaliana. Evolution *62*, 2971-2983.

Sanda,S.L. and Amasino,R.M. (1996). Interaction of FLC and late-flowering mutations in Arabidopsis thaliana. Mol. Gen. Genet. *251*, 69-74.

Scarcelli,N., Cheverud,J.M., Schaal,B.A., and Kover,P.X. (2007). Antagonistic pleiotropic effects reduce the potential adaptive value of the FRIGIDA locus. Proc Natl. Acad. Sci U. S. A *104*, 16986-16991.

Schomburg,F.M., Patton D.A., Meinke,D.W., and Amasino,R.M. (2001). *FPA*, a gene involved in floral induction in *Arabidopsis*, encodes a protein containing RNA-recognition motifs. Plant Cell *13*, 777-787.

Schönrock,N., Bouveret,R., Leroy,O., Borghi,L., Köhler,C., Gruissem,W., and Hennig,L. (2006). Polycomb group proteins repress the floral activator *AGL19* in the *FLC*-independant vernalization pathway. Genes Dev. *20*, 1667-1678.

Searle,I., He,Y., Turck,F., Vincent,C., Fornara,F., Kröber,S., Amasino,R.A., and Coupland,G. (2006). The transcription factor FLC confers a flowering response to vernalization by repressing meristem competence and systemic signalling. Genes Dev. *20*, 898-912.

Sheldon,C.C., Conn,A.B., Dennis,E.S., and Peacock,W.J. (2002). Different regulatory regions are required for the vernalization-induced repression of FLOWERING LOCUS C and for the epigenetic maintenance of repression. Plant Cell *14*, 2527-2537.

Sheldon,C.C., Hills,M.J., Lister,C., Dean,C., Dennis,E.S., and Peacock,W.J. (2008). Resetting of FLOWERING LOCUS C expression after epigenetic repression by vernalization. Proc. Natl. Acad. Sci. U. S. A *105*, 2214-2219.

Shindo,C., Aranzana,M.J., Lister,C., Baxter,C., Nicholls,C., Nordborg,M., and Dean,C. (2005). Role of FRIGIDA and FLOWERING LOCUS C in determining variation in flowering time of Arabidopsis. Plant Physiol *138*, 1163-1173.

Shindo,C., Lister,C., Crevillen,P., Nordborg,M., and Dean,C. (2006). Variation in the epigenetic silencing of FLC contributes to natural variation in Arabidopsis vernalization response. Genes Dev. *20*, 3079-3083.

Simon,M., Loudet,O., Durand,S., Berard,A., Brunel,D., Sennesal,F.X., Durand-Tardif,M., Pelletier,G., and Camilleri,C. (2008). Quantitative trait loci mapping in five new large recombinant inbred line populations of Arabidopsis thaliana genotyped with consensus single-nucleotide polymorphism markers. Genetics *178*, 2253-2264.

Simpson,G.G. (2004). The autonomous pathway: epigenetic and post-transcriptional gene regulation in the control of *Arabidopsis* flowering time. Current Opinion in Plant Biology *7*, 570-574.

Simpson,G.G., Dijkwel,P.P., Quesada,V., Henderson,I., and Dean,C. (2003). FY is an RNA 3' end-processing factor that interacts with FCA to control the *Arabidopsis* floral transition. Cell *113*, 777-787.

Sironi,M., Menozzi,G., Comi,G.P., Cagliani,R., Bresolin,N., and Pozzoli,U. (2005). Analysis of intronic conserved elements indicates that functional complexity might represent a major source of negative selection on non-coding sequences. Hum. Mol. Genet. *14*, 2533-2546.

Srikanth,A. and Schmid,D. (2011). Regulation of flowering time: all roads lead to Rome. Cellular and Molecular Life Sciences 2013-2037.

Stinchcombe,J.R., Weinig,C., Ungerer,M., Olsen,K.M., Mays,C., Halldorsdottir,S.S., Purugganan,M.D., and Schmitt,J. (2004). A latitudinal cline in flowering time in

Arabidopsis thaliana modulated by the flowering time gene FRIGIDA. Proc Natl. Acad. Sci U. S. A *101*, 4712-4717.

Strange,A., Li,P., Lister,C., Anderson,J., Warthmann,N., Shindo,C., Irwin,J., Nordborg,M., and Dean,C. (2011). Major-Effect Alleles at Relatively Few Loci Underlie Distinct Vernalization and Flowering Variation in Arabidopsis Accessions. PLoS one *6*, 1-11.

Suarez-Lopez,P., Wheatley,K., Robson,F., Onouchi,H., Valverde,F., and Coupland,G. (2001). CONSTANS mediates between the circadian clock and the control of flowering in Arabidopsis. Nature *410*, 1116-1120.

Sung,S. and Amasino,R. (2004). Vernalization in Arabidopsis thaliana is mediated by the PHD finger protein VIN3. Nature *427*, 159-164.

The Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature *408*, 796-815.

Thomas,B. (2006). Light signals and flowering. J. Exp. Bot. *57*, 3387-3393.

Toomajian,C., Hu,T.T., Aranzana,M.J., Lister,C., Tang,C., Zheng,H., Zhao,K., Calabrese,P., Dean,C., and Nordborg,M. (2006). A nonparametric test reveals selection for rapid flowering in the Arabidopsis genome. PLoS. Biol. *4*, e137.

Trevaskis,B., Tadege,M., Hemming,M.N., Peacock,W.J., Dennis,E.S., and Sheldon,C. (2007). Short vegetative phase-like MADS-box genes inhibit floral meristem identity in barley. Plant Physiol *143*, 225-235.

Tsai,M.C., Manor,O., Wan,Y., Mosammaparast,N., Wang,J.K., Lan,F., Shi,Y., Segal,E., and Chang,H.Y. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. Science *329*, 689-693.

Ungerer,M.C., Halldorsdottir,S.S., Modliszewski,J.L., Mackay,T.F., and Purugganan,M.D. (2002). Quantitative trait loci for inflorescence development in Arabidopsis thaliana. Genetics *160*, 1133-1151.

Valverde,F., Mouradov,A., Soppe,W., Ravenscroft,D., Samach,A., and Coupland,G. (2004). Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. Science *303*, 1003-1006.

VanGuilder, HD. Twenty-five years of quantitative PCR for gene expression analysis. Vrana, KE and Freeman, WM. BioTechniques 44[5], 619-626. 2008.
Ref Type: Generic

Wang,J.W., Czech,B., and Weigel,D. (2009a). miR156-regulated SPL transcription factors define an endogenous flowering pathway in Arabidopsis thaliana. Cell *138*, 738-749.

Wang,Q., Sajja,U., Rosloski,S., Humphrey,T., Kim,M.C., Bomblies,K., Weigel,D., and Grbic,V. (2007). HUA2 caused natural variation in shoot morphology of A. thaliana. Curr. Biol. *17*, 1513-1519.

Wang,R., Farrona,S., Vincent,C., Joecker,A., Schoof,H., Turck,F., Alonso-Blanco,C., Coupland,G., and Albani,M.C. (2009b). PEP1 regulates perennial flowering in Arabis alpina. Nature *459*, 423-427.

Weigel,D., Ahn,J.H., Blazquez,M.A., Borevitz,J.O., Christensen,S.K., Fankhauser,C., Ferrandiz,C., Kardailsky,I., Malancharuvil,E.J., Neff,M.M., Nguyen,J.T., Sato,S., Wang,Z.Y., Xia,Y., Dixon,R.A., Harrison,M.J., Lamb,C.J., Yanofsky,M.F., and Chory,J. (2000). Activation tagging in Arabidopsis. Plant Physiol *122*, 1003-1013.

Weigel,D., Alvarez,J., Smyth,D.R., Yanofsky,M.F., and Meyerowitz,E.M. (1992). LEAFY controls floral meristem identity in Arabidopsis. Cell *69*, 843-859.

Weigel,D. and Glazebrook,J. (2010). Transformation of Agrobacterium Using the Freeze-Thaw Method. Cold Spring Harbor Protocols *2006*, db.

Weinig,C., Ungerer,M.C., Dorn,L.A., Kane,N.C., Toyonaga,Y., Halldorsdottir,S.S., Mackay,T.F., Purugganan,M.D., and Schmitt,J. (2002). Novel loci control variation in reproductive timing in Arabidopsis thaliana in natural environments. Genetics *162*, 1875-1884.

Werner,J.D., Borevitz,J.O., Uhlenhaut,N.H., Ecker,J.R., Chory,J., and Weigel,D. (2005). FRIGIDA-independent variation in flowering time of natural Arabidopsis thaliana accessions. Genetics *170*, 1197-1207.

Wigge,P.A. (2011). FT, a mobile developmental signal in plants. Curr. Biol. *21*, 374-378.

Wigge,P.A., Kim,M.C., Jaeger,K.E., Busch,W., Schmid,M., Lohmann,J.U., and Weigel,D. (2005). Integration of spatial and temporal information during floral induction in Arabidopsis. Science *309*, 1056-1059.

Wilczek,A.M., Roe,J.L., Knapp,M.C., Cooper,M.D., Lopez-Gallego,C., Martin,L.J., Muir,C.D., Sim,S., Walker,A., Anderson,J., Egan,J.F., Moyers,B.T., Petipas,R., Giakountis,A., Charbit,E., Coupland,G., Welch,S.M., and Schmitt,J. (2009). Effects of genetic perturbation on seasonal life history plasticity. Science *323*, 930-934.

Willige,B.C., Ghosh,S., Nill,C., Zourelidou,M., Dohmann,E.M., Maier,A., and Schwechheimer,C. (2007). The DELLA domain of GA INSENSITIVE mediates the interaction with the GA INSENSITIVE DWARF1A gibberellin receptor of Arabidopsis. Plant Cell *19*, 1209-1220.

Wilson,R.N., Heckman,J.W., and Somerville,C.R. (1992). Gibberellin Is Required for Flowering in Arabidopsis thaliana under Short Days. Plant Physiol *100*, 403-408.

Wood,C.C., Robertson,M., Tanner,G., Peacock,W.J., Dennis,E.S., and Helliwell,C.A. (2006). The Arabidopsis thaliana vernalization response requires a polycomb-like protein complex that also includes VERNALIZATION INSENSITIVE 3. PNAS *103*, 14631-14636.

Wu,G., Park,M.Y., Conway,S.R., Wang,J.W., Weigel,D., and Poethig,R.S. (2009). The sequential action of miR156 and miR172 regulates developmental timing in Arabidopsis. Cell *138*, 750-759.

Xing,D., Zhao,H., Xu,R., and Li,Q.Q. (2008). Arabidopsis PCFS4, a homologue of yeast polyadenylation factor Pcf11p, regulates FCA alternative processing and promotes flowering time. Plant J *54*, 899-910.

Yamaguchi,A., Kobayashi,Y., Goto,K., Abe,M., and Araki,T. (2005). TWIN SISTER OF FT (TSF) acts as a floral pathway integrator redundantly with FT. Plant Cell Physiol *46*, 1175-1189.

Yu,H., Tan,E., and Kumar,P. (2002). AGAMOUS-LIKE 24, a dosage-dependant mediator of the flowering signals. Proc Natl. Acad. Sci U. S. A *99*, 16336-16341.

Zhai,J., Liu,J., Liu,B., Li,P., Meyers,B.C., Chen,X., and Cao,X. (2008). Small RNA-directed epigenetic natural variation in Arabidopsis thaliana. PLoS. Genet. *4*, e1000056.

Zhang,H. and van,N.S. (2002). The VERNALIZATION INDEPENDENCE 4 gene encodes a novel regulator of FLOWERING LOCUS C. Plant J. *31*, 663-673.

Name	Sequence	Tm (°C)	Gene	Reference
Sequencing and screening for known polymorphisms				
FRI coding F	TCTGAACAGCGACGAAGAGA	57	*FRI*	
FRI coding R	GGATTTGGCTCAATCAGTCAT	56	*FRI*	
FRI UJ26	AGATTTGCTGGATTTGATAAGG	55	*FRI*	Johanson 2000
FRI UJ34	ATATTTGATGTGCTCTCC	49	*FRI*	
FRI Promoter F	AGTACTCACAAGTCACAA	52	*FRI*	
FRI Promoter R	GAAGATCATCGAATTGGC	51	*FRI*	
FLC insertion F	AAACAATCTGGACAGTAGAGGCTT	58	*FRI*	
FLC deletion R	CAGGCTGGAGAGATGACAAAA	58	*FRI*	
FLC deletion F	TGTATGACTTTTCTTTGAGGG	54	*FLC*	
FLC deletion R	GTTTTAACTTGTTCTTAGCTTTG	54	*FLC*	
Vin3 Prom Seq F	CATGGGATGCTAATTCAAAAA	55	*VIN3*	
Vin3 Prom Seq R	TTCACGTGCTTTGGAAGATAT	55	*VIN3*	
Vin3 Seq 1 F	GGTATATACAATAAGTCCAGC	55	*VIN3*	
Vin3 Seq 1 R	TGAATGAATCTCAACAACCC	55	*VIN3*	
Vin3 Seq 2 F	CGCATTGGATTCGTCGTCAT	55	*VIN3*	
Vin3 Seq 2 R	CATGCACACCCGAAGAAAGT	55	*VIN3*	
Vin3 Seq 3 F	GAAGAAGCTCGAAGGTGATGT	55	*VIN3*	
Vin3 Seq 3 R	ACCATGTCAAGAACCTTTCCC	55	*VIN3*	
Vin3 Seq 4 R	CTGCTCACCAACCTTTGAATA	55	*VIN3*	
VIN3 ins F	CGGATCTTGFGTTCTTCATCAT	55	*VIN3*	
VIN3 ins R	ATCCATGTTCTCTGGACCCAC	55	*VIN3*	
SSLPs				
CIW7 F	AATTTGGAGATTAGCTGGAAT	55		
CIW7 R	CCATGTTGATGATAAGCACAA	55		
JV30	CATTAAAATCACCGCCAAAAA	55		
JV31	TTTTGTTACATCGAACCACACA	55		
JV61	CGCTTTCCTTGTGTCATTCC	55		
JV62	AAATGCAAATATTGATGTGTGAAA	55		
JV75	CACAATCAGAGGGGGTTGAT	55		
JV76	AAATTTTGGGGGAAATGAAA	55		
JV57	TCCGATTGGTCTAAAGTACGAC	55		
JV58	TTTGATGGACTCTTACATTGGAAA	55		
NGA249 F	GGATCCCTAACTGTAAAATCCC	55		
NGA 249 R	TACCGTCAATTTCATCGCC	55		
NGA139 F	GGTTTCGTTTCACTATCCAGG	55		
NGA139 R	AGAGCTACCAGATCCGATGG	55		
RT-PCR				
FDH F	TTGGACTTGCTGTTGCCGA	57	*FDH*	J.G
FDH R	AAGTTCCCCATCCCTTGTGAC	57	*FDH*	J.G
FLCRtm-PCR F	CGGTCTCATCGAGAAAGCTC	57	*FLC*	J.G
FLCRtm-PCR R	CCACAAGCTTGCTATCCACA	57	*FLC*	J.G
FRI RT-PCR F	CGTGGTAAAGGGAAGGTGAA	57	*FRI*	
FRI RT-PCR R	TTGCCCTCTCAAATGACTCC	57	*FRI*	