Investigating the vernalisation requirement of *Brassica napus*

Eleri Hâf Tudor

Thesis submitted for the degree of Doctor of Philosophy

University of East Anglia

John Innes Centre

This page has been left intentionally blank.

Abstract

Many plant species have evolved to overwinter before flowering. In *Arabidopsis thaliana* a requirement for cold to flower is determined primarily by two genes *FRIGIDA (FRI)* and *FLOWERING LOCUS C (FLC)*. *FRI* upregulates the expression of the floral repressor *FLC*, which in turn inhibits the expression of floral promoting genes like *FLOWERING LOCUS T (FT)*. Prolonged cold temperatures overcome this inhibition through a process called vernalisation. Accessions of *A. thaliana* that do not need cold to flower often contain mutations at either *FRI* or *FLC*.

Winter, spring and biennial cultivars of *Brassica napus* are grown for vegetable and seed production and variation for vernalisation requirement is investigated. Vernalisation treatment had a significant effect on flowering time and inflorescence architecture. Four orthologues of *FRI* and nine orthologues of *FLC* have previously been characterised in *B. napus*, but how these genes contribute to the vernalisation requirement is not fully understood. Molecular characterisation of *FRI* from a diverse panel of *B. napus* accessions revealed the presence of non-synonymous allelic variation at all four orthologues that was significantly associated with crop type. However, natural and induced mutations at *FRI* had a minimal effect on flowering time and *FLC* expression. Two genetic mapping approaches, Associative Transcriptomics and QTL-seq, were subsequently used to identify the genetic variation responsible for variation in vernalisation requirement in *B. napus*. Associations were detected in genomic regions that encompass orthologues of *FLC* and *FT*, but not *FRI*.

Characterisation of the *FLC* and *FT* orthologues revealed the presence of DNA sequence and gene expression variation and I hypothesise this contributes to the flowering time differences measured. This work has enhanced our understanding of the vernalisation requirement of *B. napus* and could contribute to the improvement of cultivars with adapted flowering times and improved yields.

This page has been left intentionally blank.

Crynodeb

Mae rhywogaeth nifer o blanhigion wedi esblygu i oroesi oerfel y gaeaf cyn blodeuo. Yn y rhywogaeth *Arabidopsis thaliana* dau genyn, *FRIGIDA (FRI)* a *FLOWERING LOCUS C (FLC)* sy'n rheoli'r angen hwn am dymheredd oer cyn blodeuo. Mae *FRI* yn gweithio i gynyddu mynegiant *FLC* sy'n atal y broses flodeuo trwy rwystro mynegiant y genyn *FLOWERING LOCUS T (FT)*. Mae cyfnod estynedig o oerfel yn goresgyn ffwythiant *FRI* a *FLC* trwy broses a elwir "vernalisation". Heb *FRI* a *FLC* mae planhigion *A. thaliana* yn blodeuo'n gynnar heb dymereddau oer.

Tyf *Brassica napus* i gynhyrchu llysiau a hadau a dangoswyd yn y gwaith hwn bod amrywiaeth am angenrheidrwydd vernalisation. Dangoswyd bod vernalisation yn effeithio ar amser blodeuo ac ar nifer y canghennau. Cyn y gwaith hon darganfuwyd bod pedwar ortholog o *FRI* a naw ortholog o *FLC* yn *B. napus* ond ar hyn o bryd mae ansicrwydd ynghylch sut maent yn rheoli vernalisation. Ymchwilwyd *FRI* mewn poblogaeth amrywiol o *B. napus* a darganfuwyd gwahaniaethau rhwng yr amrywiaethau. Fodd bynnag chafodd y gwahaniaethau hyn ddim effaith ar amser blodeuo na mynegiant *FLC*. Er mwyn deall beth oedd yn rheoli anghenrwydd vernalisation yn *B. napus* defnyddiwyd dwy dechneg mapio genetig, "Associative Transcriptomics" a "QTL-seq". Darganfuwyd rhanbarthau genetig sy'n cynnwys *FLC*, *FT* ond nid *FRI*.

Dargafuwyd gwahaniaethau yn *FLC* a *FT* a rhagdybiwn taw'r rhain sy'n cyfrannu at amrywiaeth amser blodeo *B. napus*. Mae'r gwaith hwn wedi ehangu ar ein dealltwriaeth o broses vernalisation yn *B. napus* ac fe all y darganfyddiadau hyn gyfrannu at addasu amrywiaethau *B. napus* er mwyn cynyddu cyfradd cynhyrchu.

Acknowledgements

Thank you to all who have supported me over the last four years. First and foremost, I would like to thank my supervisors, Judith Irwin and Caroline Dean, for your valuable guidance and support. Thank you for your dedication to develop this project, your encouragement and enthusiasm have been unwavering throughout. I greatly appreciate the role you have played in my development as a scientist.

I acknowledge the BBSRC and iCASE partner Bayer CropScience for funding my PhD. Thank you to my iCASE supervisors Pieter Ouwerkerk and Mark Davey for your guidance. Thanks, are given to you and the Genetics team, especially Maria del Mar Perez Nicolas, Anne Olivier, and the glasshouse team in Deinze, for all your help during my visit to Gent, Belgium.

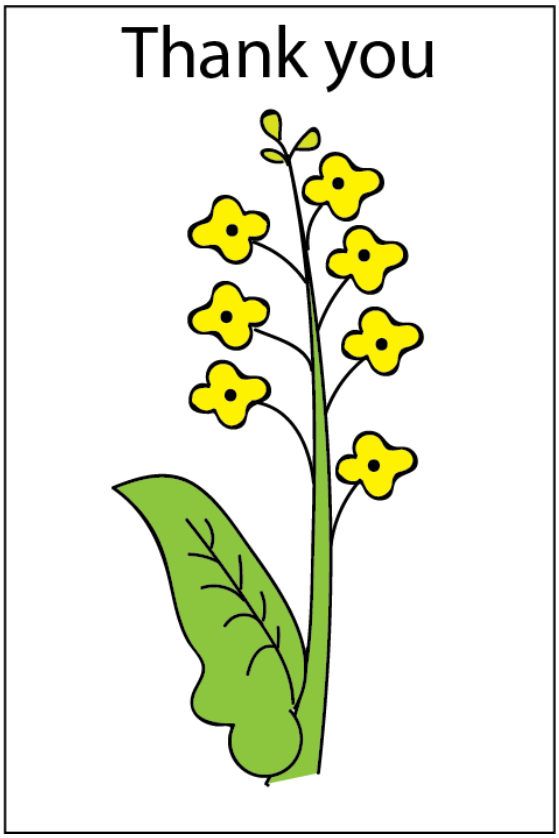
Thank you to members of the Dean Lab past and present, you have provided countless discussions which have helped shape this project. Special thanks to Rebecca Bloomer and Jo Hepworth for your helpful conversation about flowering time and beyond. Thank you for all your ideas, your scientific wisdom and for being an inspiring group of scientists.

Thank you to the *Brassica* Group past and present (Irwin, Morris, Penfield Labs, Martin Trick and Rachel Wells) for your knowledge, and for all the Brassica themed events both scientific and social. Special thanks to Rachel Wells for your endless encouragement and *Brassica* wisdom. A huge thanks to Marc Jones, a former fellow PhD student and coding wizard, for all your help with the bioinformatics analysis. Lastly, thank you to Emily Hawkes. I have missed not having you in the Lab this last year. Thank you for sharing plant materials and methods, and, most of all, for your friendship.

I am grateful to Lionel Perkins and the Horticultural Services team who have worked so diligently to look after my plants. Thank you also to the Laboratory Support team for your continued support in the Lab.

Thank you to those who have commented on and proofread drafts of this Thesis. I would most like to thank Judith Irwin for your feedback and comments, but also Hefin Miles and Carys Tudor for your proofreading and suggestions.

Finally, thank you to my family and friends for always being there with love, laughs and support. Thank you to my fantastic family, my Mam, Dad, Carys and Dafydd. for always being there for me. Thank you to my friends, especially to Cathy, for all the tea breaks, giggles and for helping to make the last four years memorable. My greatest thanks however must go to Hef, thank you for always being kind, making me laugh and for being your wonderful self.



Contents

List of Figures

List of Tables

Abbreviations

A/T/C/G	Adenine/Thymine/Cytosine/Guanine
ANOVA	Analysis of Variance
BLAST	Basic Local Alignment Search Tool
BnaFLC	*Brassica napus FLC*
BnaFRI	*Brassica napus FRI*
BnaFT	*Brassica napus FT*
BolFLC	*Brassica oleracea FLC*
BolFRI	*Brassica oleracea FRI*
BolFT	*Brassica oleracea FT*
BraFLC	*Brassica rapa FLC*
BraFRI	*Brassica rapa FRI*
BraFT	*Brassica rapa FT*
BRC1	*BRANCHED 1*
bp	Base pairs
BSA	Bulk Segregant Analysis
cDNA	Complementary DNA
ChIP	Chromatin immunoprecipitation
COOLAIR	Cold induced long antisense intragenic RNA
DFFS	Diversity Fixed Foundation Set
DH	Doubled Haploid
DNA	Deoxyribonucleic acid
DNF	Did not flower
FLC	*FLOWERING LOCUS C*
FRI	*FRIGIDA*
FT	*FLOWERING LOCUS T*
GEM	Gene Expression Marker
GWAS	Genome Wide Association Study
H3K27me3	Histone 3 lysine 27 trimethylation
H3K36me3	Histone 3 lysine 36 trimethylation
H3K4me3	Histone 3 lysine 4 trimethylation
InDel	Insertion/Deletion
JIC	John Innes Centre
MLM	Mixed Linear Model
Mya	Million years ago
mRNA	Messenger RNA
NBI	Norwich Bioscience Institute
NIL	Near Introgression Line
NV / NVERN	No vernalisation
OREGIN	Oilseed Rape Genetic Improvement Network
ORF	Open Reading Frame
OSR	Oilseed rape
PCR	Polymerase Chain Reaction
PHD	Plant homeodomain protein
PRC2	Polycomb Repressive Complex 2
qPCR	Quantitative PCR
QQ	Quantile Quantile

QTL	Quantitative Trait Locus
RNA	Ribonucleic acid
RPKM	Reads per kilobase per million aligned reads
RT-PCR	Reverse Transcription PCR
SAM	Shoot apical meristem
SEQ	Sequencing
SNP	Single nucleotide polymorphism
TGAC	The Genome Analysis Centre, now the Earlham Institute
UBC	Ubiquitin-C gene
UTR	Un-translated region
VERN	Vernalisation or vernalised
WT	Wild-type

Chapter 1: Introduction

1.1. Thesis overview

Timing of the floral transition in crop plants is a major target for plant breeding (Jung and Müller 2009). For seed crops flowering is essential for a harvestable product, while early flowering in vegetative crops negatively affects yield (Jung and Müller 2009). Four regulatory pathways integrate environmental and endogenous signals that influence flowering time. One environmental signal, known as vernalisation, leads to an acceleration of flowering in response to cold temperatures. Manipulation of the requirement for, and responsiveness to, vernalisation has facilitated the generation of novel crop cultivars that are adapted to local environments (Jung and Müller 2009).

In this Thesis I will describe my investigation of the genetic control of vernalisation requirement in the crop species *Brassica napus*. I analysed the flowering times, and inflorescence architecture traits of cultivars from a diverse collection of *B. napus*, including two subspecies *B. napus* ssp. *napus* and *B. napus* ssp. *rapifera*, under a range of vernalisation conditions I have identified and experimentally validated genetic variation likely important for variation in flowering time. My experimental results are described in Chapters 3 – 6 of this Thesis.

Allelic variation at *FRIGIDA (FRI)* is a major determinant of the vernalisation requirement in natural accessions of the model species *Arabidopsis thaliana* (Johanson *et al*., 2000, Shindo *et al*., 2005). I investigate the function of *FRI* orthologues in *B. napus* using a combination of genetic, transgenic and mutational approaches (Chapter 3). This analysis demonstrated that allelic variation at *FRI* orthologues are not the major determinant for variation in vernalisation requirement in *B. napus*.

To identify candidate genes that are important for variation in vernalisation requirement and response, I analysed flowering time measurements collected from a diversity set population of *B. napus* cultivars using an association genetics method known as Associative Transcriptomics (Harper *et al*., 2012; Chapter 4). In addition to this, I analysed flowering time measurements collected from a mapping population derived from two *B. napus* cultivars using a different association genetics method known as QTL-seq (Takagi *et al*., 2013; Chapter 5). The likely candidate genes, identified from the genomic regions that associated with variation for flowering time with and without vernalisation, were investigated and I discuss their possible role in determining flowering time. Interestingly, orthologues of *FLC* and *FT*, but not *FRI*, were identified as candidate genes important for variation in vernalisation requirement in *B. napus* (Chapter 4 and Chapter 5). In the final experimental Chapter of this Thesis I describe a preliminary study demonstrating that flowering time and branch number are significantly correlated in both *A. thaliana* and *B. napus*. I suggest flowering time, and therefore vernalisation requirement, has a pleiotropic effect on inflorescence architecture in both *A. thaliana* and *B. napus* (Chapter 6). As branch number is positively correlated with seed yield (Li *et al*., 2016) I discuss the consequence of modulating flowering time to influence seed yield.

Modelling predictions of climate change suggest, if carbon emissions remain unchanged, global temperature averages could increase by more than 4°C (Brown and Caldeira 2017). These increasing temperatures are predicted to couple with increasing inter-annual temperature variation, especially in temperate regions (Luterbacher *et al*., 2004). For *B. napus*, increasing average growth temperatures by 6°C significantly reduces plant biomass and therefore yield (Qaderi *et al*., 2006). Flowering time is extremely sensitive to variability in temperature therefore planning for future climate change is now a major challenge for plant breeding and crop improvement (Craufurd and Wheeler 2009). Mining germplasm for variation in sensitivity to environmental conditions will help breed new cultivars with robust flowering times in a changing climate (Irwin *et al*., 2016).

1.2 Oilseed rape: an economically important crop worldwide

Oilseed rape (*Brassica napus*, 2n=38, AACC) is grown for its seed oil and seed meal for both human and animal consumption. Second only to soybean, oilseed rape production represents nearly 15% of the global oilseed market (Carre and Pouzet 2014). According to the Food and Agriculture Organization of the United Nations (www.fao.org/faostat) over 41 million hectares of oilseed rape was harvested in 2016 with Canada, China, the EU and India among the major producers. Within the EU, production of this crop has exhibited sustained growth over the last 20 years (Carre and Pouzet 2014). A requirement for the mandatory use of biofuels by 2020 has led to an almost 17% increase in production of oilseed rape between 2006-2016 (https://ec.europa.eu/eurostat/, www.fao.org/faostat). To sustain this global demand for oilseed rape, a detailed understanding of the genetics underlying yield-related traits is required.

1.2.1. *Brassica* origins

Oilseed rape represents a crop type within the *B. napus* species and shares a close evolutionary history with other members of the *Brassica* genus. There are six cultivated *Brassica* species, three are diploid (*Brassica juncea, Brassica oleracea, Brassica rapa*) and three are allotetraploid (*Brassica carinata, B. napus, Brassica nigra*). Nagahara U first demonstrated in 1935 that the diploid species could inter-hybridise to form the allotetraploid species (Figure 1. 1 A). He hypothesised from this that *B. juncea, B. oleracea* and *B. rapa* were the ancestral progenitors of *B. carinata, B. napus*, and *B. nigra* and his theory has since been confirmed (Parkin *et al*., 1995).

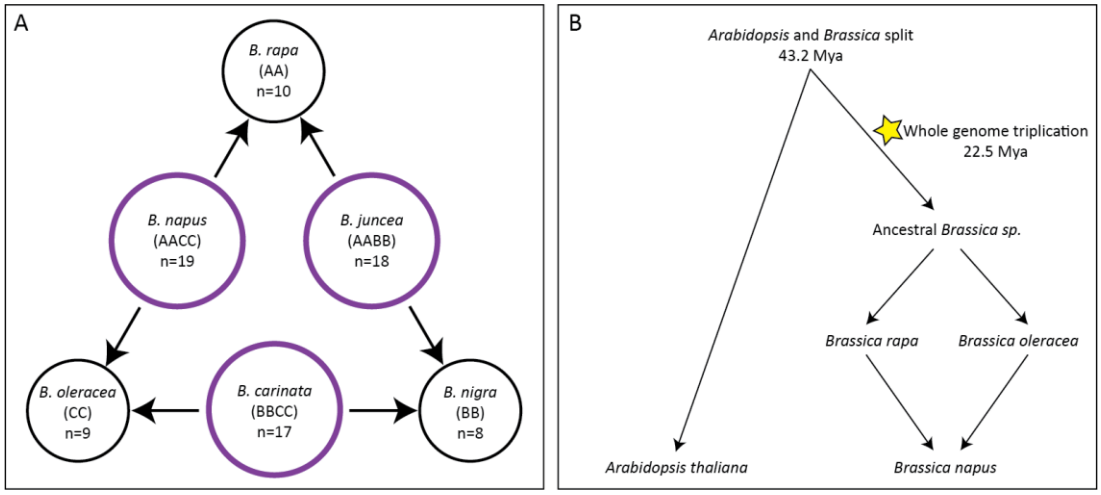


Figure 1. 1: The evolutionary history of *Brassica* spp. (A) The triangle of U (U 1935) illustrating the hybridisation between the diploid species (black) the allotetraploid species (purple). The genomes are annotated A, B or C and the number of haploid chromosomes listed. (B) The phylogenetic relationship between *A. thaliana*, *B. rapa*, *B. oleracea* and *B. napus* according to Beilstein *et al*., (2010), the predicted dates of the Arabidopsis-*Brassica* divergence and *Brassica* whole genome triplication events are highlighted.

Phylogenetic analysis of *B. juncea, B. oleracea, B. rapa* revealed two major lineages (Arias *et al*., 2014); the Nigra (containing *B. nigra*) lineage is found in North Africa and the Middle East while members of the Oleracea lineage (containing *B. oleracea* and *B. rapa*) are distributed across Europe and Asia. *B. juncea* is found in the Middle East at the contact zone between *B. nigra* and *B. rapa*, *B. napus* is found in Northern Europe at the contact zone between the *B. oleracea* and *B. rapa* native regions, while Kenya and Ethiopia are the native regions of *B. carinata*. According to Arias *et al*., (2014) *B. oleracea* and *B. rapa* diverged from *B. nigra* approximately 7.9 Mya. *B. rapa* then migrated throughout the Irana-Turanian region reaching central Asia approximately 2 Mya, while *B. oleracea* migrated across Europe. A complex evolutionary history involving hybridisation and migration events has allowed *Brassica* crops to colonise much of Europe, the Middle East and Asia.

Genomic analysis of *B. rapa* and *B. oleracea* has revealed the presence of an ancient hexaploidization event and therefore should be considered paleohexaploid and not diploid *per se* (Figure 1. 1 B; Lysak *et al*., 2005, Parkin *et al*., 2005). This involved a whole genome duplication event and multiple hybridisation events which gave rise

to an ancestral hexaploid *Brassica* species, dated 15.6-28.3 Mya (Arias *et al*., 2014). For every single gene present in the original *Brassica* genome, three are predicted for *B. rapa* and *B. oleracea*. However, the modern *B. rapa* and *B. oleracea* exhibit evidence of diploidisation, gene loss and rearrangement (Wang *et al*., 2011a, Cheng *et al*., 2012 Tang *et al*., 2012, Liu *et al*., 2014). *B. napus* is a neo-allotetraploid formed from multiple interspecific hybridisation events between *B. oleracea* and *B. rapa* (Kagale *et al*., 2014). The date of origin for *B. napus* is under debate, but multiple analyses have suggested that the hybridisation event that gave rise to *B. napus* occurred between 7,500 – 12,500 years ago (Schmidt and Bancroft 2011, Kagale *et al*., 2014, Chalhoub *et al*., 2014).

The *B. napus* genome contains 19 haploid chromosomes, 10 from the *B. rapa* progenitor genome and 9 from the *B. oleracea* progenitor genomes denoted the A and C genomes respectively (Figure 1. 1 A; Parkin *et al*., 1995). During meiosis, homologous chromosomes pair and recombine; A with homologous A or C with homologous C; ensuring the disomic inheritance of the A and C genomes, and ultimately the stability of the *B. napus* genome (Parkin et a 195, Udall *et al*., 2005). However, due to sequence and structural similarity between the A and C genomes, homeologous exchanges (A with C) have been documented to occur (Parkin *et al*.,1995, Sharpe *et al*.,1995, Osborn *et al*., 2003, Chalhoub *et al*., 2014, He *et al*., 2017). This phenomenon can impact allelic and phenotypic diversity within the *B. napus* species.

1.2.2. *Arabidopsis thaliana* as a model species

The publication of several *Brassica* genomes (*B. rapa* genome; Wang *et al*., 2011a, *B. oleracea* genome; Liu *et al*., 2014, Parkin *et al*,. 2014, *B. napus* genome; Chalhoub *et al*., 2014) has provided an invaluable resource for genetic research of the diploid and allotetraploid *Brassica* species. Hereafter when referring to *Brassica* this will denote three species; *B. rapa*, *B. oleracea* and *B. napus*.

Brassica and the model species *Arabidopsis thaliana* are members of the Brassicaceae family. The species diverged approximately 43.2 Mya (Figure 1. 1 B; Beilstein *et al*., 2010) and share up to 85% coding sequence identity (Trick *et al*., 2009). Comparative analyses between *A. thaliana*, *Brassica*, and other members of the Brassicaceae, have revealed extensive regions of noncollinearity between the species (Parkin *et al*., 2005, Schranz *et al*., 2006, Trick *et al*., 2009, Bancroft *et al*., 2011). Comparisons made between *A. thaliana, Arabidopsis lyrata, Capsella rubella* and *B. rapa* identified structural differences that account for the reduced genome size and low chromosome number of *A. thaliana* (Schranz *et al*., 2006). *A. lyrata* and *C. rubella* have 8 haploid chromosomes and highly similar genome structures while *A. thaliana* has 5 haploid chromosomes and extensive genome rearrangement in comparison. Schranz *et al*., (2006) proposed an ancestral karyotype for the Brassicaceae from which *A. thaliana, A. lyrata, C. rubella* and the *Brassica* crops have evolved. The karyotype is hypothesised to have 8 haploid chromosomes divided into 24 conserved blocks. These blocks when duplicated, re-arranged and deleted can be used to generate the genomes of all modern Brassicaceae (Schranz *et al*., 2006). The *A. thaliana* genome may therefore be an excellent model for identifying orthologous genes but not always an accurate reference for determining their genomic position in *Brassica*.

1.3. Flowering time as a yield trait

Many crop species have seen yield increase rates decline during the last thirty years (Huang *et al*., 2002). It is estimated that production of food and feed must continue to rise by 1.2% annually to satisfy demand from an ever-increasing global human population (Rosegrant *et al*., 1999, Huang *et al*., 2002). Many agronomic traits are considered important for yield and the timing of flowering is one trait that can affect the seed yield of oilseed rape, and the vegetable production of vegetative *B. napus* crops. The timing of flowering is considered so important that oilseed rape varieties are assessed for their earliness of flowering by the UK Agriculture and Horticulture Development Board (AHDB, https://cereals.ahdb.org.uk/) before they are approved for production and added to farmer Recommended Lists. A functional understanding

of the key genes involved in the floral transition would give Breeders a molecular toolkit for altering the timing of flowering of *B. napus*.

1.3.1. Floral integrator genes receive and convey flowering signals in *A. thaliana*

The transition from vegetative to reproductive growth is a major developmental decision in plants and its correct timing is crucial for reproductive success. Extensive studies on the control of flowering in *A. thaliana* have revealed many genes that function to promote or inhibit the floral transition (Koornneef *et al*., 1991, Clarke and Dean, 1994, Lee *et al*., 1994a, Lee *et al*., 1994b, Koornneef *et al*., 1994, Koornneef *et al*., 1998). Functional characterisation of these genes under various environmental conditions have revealed a complex network of proteins that can be assigned to four distinct regulatory pathways, integrating signals from endogenous and exogenous cues.

Signals from the autonomous, vernalisation, photoperiod, and gibberellin regulatory pathways converge on the floral integrator genes *FLOWERING LOCUS T (FT*) and *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1 (SOC1)* which activate the floral meristem identity genes *LEAFY (LFY), APETALA1 (AP1), CAULIFLOWER (CAL)* and *FRUITFULL (FUL)* (Figure 1. 2; Jung and Müller 2009, Blümel *et al*., 2015). Activation of these floral meristem identity genes promote the irreversible conversion from a vegetative to floral shoot apical meristem (Weigel *et al*., 1992, Bowman *et al*., 1993, Ferrándiz *et al*., 2000).

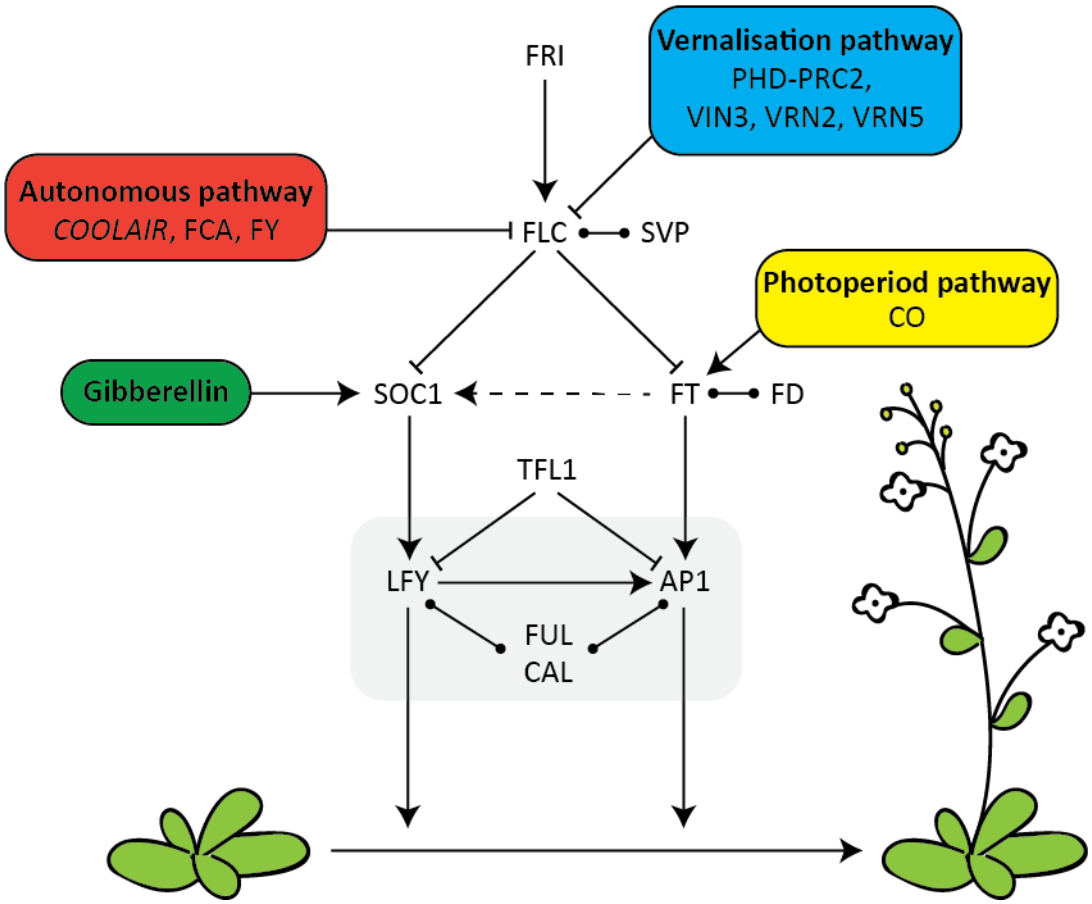


Figure 1. 2: A simplified diagram of the flowering network in *A. thaliana*. Signals from four regulatory pathways; autonomous, vernalisation, photoperiod and gibberellin are integrated to control flowering time. The floral meristem identity genes are highlighted in a grey box. Adapted from Jung and Müller 2009.

1.3.2. *FLOWERING LOCUS C* inhibits flowering

Two of the floral regulatory pathways, the autonomous and vernalisation pathways, promote flowering time by down-regulating the expression of a major floral repressor known as *FLOWERING LOCUS C (FLC)* (Figure 1. 2). *FLC* encodes a MADS-box transcription factor and its mRNA level is quantitatively correlated with flowering time (Chouard 1960, Michaels and Amasino, 1999, Sheldon *et al*., 2000). MADS-box transcription factors bind to a relatively conserved consensus sequence known as a CArG box with a core motif of $CC(A/T)_6GG$ (Riechmann *et al*., 1997). Chromatin immunopreciptation (ChIP) assays have revealed that FLC protein, in combination with SHORT VEGETATIVE PHASE (SVP), binds to the CArG box in the *SOC1* promoter and to the first intron of *FT*, inhibiting the expression of both genes (Hepworth *et al*., 2002, Helliwell *et al*., 2006, Li *et al*., 2008).

1.3.3. *FLC* expression is up-regulated by FRI

Vernalisation requirement in *A. thaliana* is a monogenic trait (Napp-Zinn 1957) chiefly determined by *FRIGIDA (FRI)* (Clarke and Dean 1994, Johanson *et al*., 2000). The FRI protein forms a transcriptional activator complex known as the FRIGIDA-complex (FRI-c) to promote high expression level of *FLC* (Figure 1. 2; Choi *et al*., 2011). Each protein within the FRI-c are reported to have a specific role in promoting *FLC* expression. FRI and FRIGIDA LIKE1 (FRL1) are essential for increasing the Histone H-3 Lysine-4 trimethylation (H3K4me3), histone marks that are associated with active transcription of the *FLC* locus, (Risk *et al*., 2010, Choi *et al*., 2011). FLC EXPRESSOR (FLX) and FRIGIDA-ESSENTIAL1 (FES1) are thought to function as transcriptional activators, while SUPPRESSOR OF FRIGIDA4 (SUF4) binds directly to the promoter region of *FLC* (Choi *et al*., 2011). Once formed, the FRI-c recruits histone modification factors of the SWR1-complex, the histone methyltransferase EFS which catalyses H3K36me3, the transcription factor *TAF14*, and the RNA-polymerase-associated factor 1 complex, to upregulate *FLC* expression (Choi *et al*., 2011 He *et al*., 2004). FRI is also implicated in the co-transcriptional regulation of *FLC* mRNA (Geraldo *et al*., 2009). Interaction between FRI and the cap binding complex (CBP20) increases the proportion of 5'-capped spliced *FLC* mRNA transcripts (Geraldo *et al*., 2009). Taken together FRI acts as a scaffold protein, recruiting many factors to the *FLC* locus, modifying the chromatin to increase the expression and stability of *FLC* transcription (***Figure 1. 2***He *et al*., 2004, Geraldo *et al*., 2009, Choi *et al*., 2011).

FRI belongs to a family of seven *FRI*-related genes, but despite sequence similarity, none have been reported to exhibit redundant biological functions (Michaels *et al*., 2004). The FRI protein is 609 amino acids long containing two regions that are predicted to form coiled-coil domains, and a stable, central domain that is conserved across all FRI-related proteins (Johanson *et al*., 2000, Michaels *et al*., 2004, Risk *et al*., 2010). Deletion of 151 residues from the C-terminus confers early flowering in *A. thaliana*, while wild-type FRI function is partially retained when 118 residues are deleted from the N-terminus (Risk *et al*., 2010). The C-terminus of the FRI protein is

therefore required to promote high *FLC* expression and, according to Choi *et al*., (2011), where direct interaction with FES1, FLX and SUF4 within the FRI-c takes place. Mutations that disrupt the C-terminal domain of FRI would certainly abolish protein function and result in early flowering. Interestingly loss of vernalisation requirement and early flowering in *A. thaliana* is often mapped to *FRI*, caused by nonsense mutations that disrupt the FRI protein (Burn *et al*., 1993, Clarke and Dean 1994, Lee *et al*., 1994b, Johanson *et al*.,2000, Ler Corre *et al*., 2002, Shindo *et al*.,2005). The consequence of allelic variation at *FRI* in *A. thaliana* and other Brassicaceae species is discussed in detail in Chapter 3.

1.3.4. The autonomous pathway and antisense transcription regulate *FLC*

Mutant genes of the autonomous pathway are late flowering but remain responsive to photoperiod and vernalisation signals (Koornneef *et al*., 1991). The late flowering phenotype of autonomous mutants is caused by high expression of *FLC*, independent of *FRI* (Figure 1. 2; Michaels and Amasino 2001). The genes involved in the autonomous pathway can broadly be divided into three functional groups that inhibit the expression of *FLC*; those that are proposed to regulate transcription (*LUMINIDEPENDENS* (*LD)*), those that function to modify the histones at the *FLC* locus (*FVE*, *FLOWERING LOCUS D* (*FLD*)), and those that have RNA binding and/or processing activity (FCA, FPA, FY) (Whittaker and Dean 2016). *LD* encodes a transcription factor whose protein is targeted to the nucleus to promote flowering (Lee *et al*., 1994a, Lee *et al*., 1994b, Aukerman *et al*., 1999). FVE and FLD are involved in histone de-acetylation and de-methylation levels, antagonising the FRI-c (He *et al*., 2004, Ausín *et al*., 2004, Liu *et al*., 2010). FCA and FPA proteins contain RNA binding domains (Macknight *et al*., 1997, Mockler *et al*., 2004), while FY is involved in the polyadenylation and 3' end processing of RNA (Simpson *et al*., 2003).

FCA, FPA and FY are required for the processing and alternative splicing of *COOLAIR*, a collection of non-coding, antisense transcripts transcribed from the *FLC* locus (Swiezewski *et al*., 2009). *COOLAIR* transcripts are divided into two

major classes based on their splicing patterns; proximal *COOLAIR* is spliced and polyadenylated using sites within intron 6 of *FLC*, while distal *COOLAIR* is spliced and polyadenylated using sites within the *FLC* promoter (Marquardt *et al*., 2014). Proximal *COOLAIR* splicing results in a reduction in H3K4me, facilitated by FLD, reducing transcriptional initiation, slowing transcriptional elongation, and ultimately resulting in low *FLC* expression (Liu *et al*., 2010, Wu *et al*., 2016). Together these factors contribute to the initial level of *FLC* expression, influencing the length of cold required to saturate the vernalisation requirement and accelerate flowering.

1.3.5. Vernalisation also down-regulates *FLC*

An acceleration of flowering in response to cold temperatures (<14°C) (Duncan *et al*., 2014) involves the epigenetic silencing of *FLC* through a process known as vernalisation (Figure 1. 3). Upon initial exposure to cold temperatures *COOLAIR* is rapidly up-regulated (Figure 1. 3) and VAL1 protein binds specifically to two RY cis-elements within the first intron of *FLC*, causing a shutdown of *FLC* transcription (Csorba *et al*., 2014, Rosa *et al*., 2016, Questa *et al*., 2016). Exposure to a prolonged period of cold temperature results in histone modifications at a site over the first exon and part of the first intron at the *FLC* locus known as the nucleation region (Finnegan and Dennis 2007, De Lucia *et al*., 2008). Prior to cold, the histones at the *FLC* locus carry H3K4me3 and H3K36me3 modifications (Li *et al*., 2007). During cold exposure and mediated by the Polycomb repressive complex 2 (PRC2), the H3K4me3 and H3K36me3 histone marks are replaced with H3K27me3; histone marks that are associated with inactive gene expression (Margueron and Reinberg, 2011, Young *et al*., 2011). VERNALISATION2 (VRN2), along with FIE, SWINGER and MSI1 proteins, function as components of the PRC2 which possess methyl-transferase activity and catalyse the methylation of H3K27me3 (Wood *et al*., 2006, Greb *et al*., 2007, De Lucia *et al*., 2008, Margueron and Reinberg, 2011). The PRC2 interacts with the plant homeodomain (PHD) finger proteins VERNALISATION INSENSITIVE3 (VIN3), VERNALISATION5 (VRN5) and VEL1 forming the PHD-PRC2 complex (Wood *et al*., 2006, Greb *et al*., 2007, De Lucia *et al*., 2008, Margueron and Reinberg, 2011) which is recruited to the *FLC* nucleation region by VAL1 (Questa *et al*., 2016). Association of the PHD-PRC2

complex at the *FLC* nucleation region results in de-acetylation and an accumulation of H3K27me3 (De Lucia *et al*., 2008).

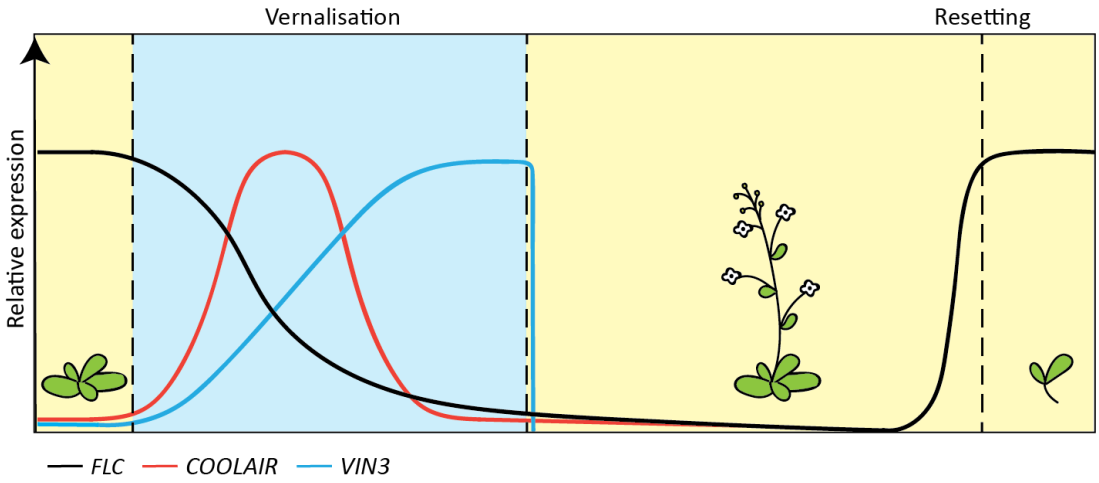


Figure 1. 3: Gene expression profiles of *FLC*, *COOLAIR* and *VIN3* with vernalisation. *FLC* (black), *COOLAIR* (red) and *VIN3* (blue) before, during, and after vernalisation. Warm temperature conditions are illustrated in yellow and vernalisation conditions illustrated in blue. Silencing of *FLC* by vernalisation promotes flowering in *A. thaliana*, however *FLC* expression levels must be re-set in the next generation. Adapted from Hepworth and Dean 2015.

When plants return to warm temperatures *VIN3* expression is rapidly down-regulated (Figure 1. 3) while the PHD-PRC2 complex, coupled with an increase of H3K27me3, is observed to spread over the entire *FLC* locus (De Lucia *et al*., 2008). It is thought that DNA replication is necessary for the spreading of H3K27me3 repressive marks (Finnegan and Dennis 2007, De Lucia *et al*., 2008), and LIKE HETEROCHROMATIN PROTEIN1 (LHP1) is required to maintain *FLC* repression through cell division (Mylne *et al*., 2006, Sung *et al*., 2006). Repression of *FLC* is maintained during multiple rounds of mitosis but is then re-set at meiosis (Sheldon *et al*., 2008). This ensures the maintenance of vernalisation requirement in the next generation.

The effect of vernalisation is quantitative; plants that have received long periods of cold will flower more rapidly than those that have received shorter periods. At the single cell level, *FLC* transcription is bistable and is either active “on” or inactive

"off" (Angel *et al*., 2011, Angel *et al*., 2015, Berry *et al*., 2015). Each cell is thought to respond to vernalisation independently, with an increasing proportion of cells switching to the "off" state with longer periods of cold (Angel *et al*., 2011, Berry *et al*., 2015). The quantitative nature of vernalisation is due to the average *FLC* state across a population of cells, and not due to quantitative changes at individual *FLC* loci (Angel *et al*., 2011, Berry *et al*., 2015).

1.3.6. Variation in vernalisation requirement and response underpins life history strategy

Allelic variation at *FRI* and *FLC* are often found to underlie the genetic basis for variation in reproductive strategy in *A. thaliana*, discussed in detail in Chapters 3 – 6 (Stinchcombe *et al*., 2004, Shindo *et al*., 2005). Mutations at *FRI* and *FLC* that disrupt protein function or expression level are often responsible for a rapid cycling flowering habit (Johanson *et al*., 2000, Michaels *et al*., 2003, Shindo *et al*., 2005) whereas accessions that are late flowering and require vernalisation possess functional FRI and highly expressed *FLC* alleles. For these late flowering accessions, non-coding variation at *FLC* is identified to be responsible for fine-tuning the vernalisation response and adaptation to local climates (Shindo *et al*., 2006, Strange *et al*., 2011, Ågren *et al*., 2013, Grillo *et al*., 2013, Li *et al*., 2014, Duncan *et al*., 2015). For example, a natural accession of *A. thaliana*, called Lov-1, requires three-months vernalisation to fully accelerate flowering (Shindo *et al*., 2006). Four single nucleotide polymorphisms (SNPs) within the non-coding region of *FLC* were found to be responsible for the late flowering phenotype of this accession, directly influencing H3K27me3 accumulation and epigenetic silencing at *FLC* (Coustham *et al*., 2012). Another example was found in the late flowering accession Var2-6 (Li *et al*., 2015). Naturally occurring splice site mutations in *COOLAIR* have been identified in Var2-6 (Li *et al*., 2015). The altered splicing of *COOLAIR* is correlated with elevated levels of *FLC* transcription prior to vernalisation (the pre-vernalisation levels) and thus impacts the length of vernalisation required to fully accelerate flowering.

Orthologues of *FRI* and *FLC* have been identified in closely related species of the Brassicaceae such as *A. lyrata* (Kuittinen *et al*., 2007, Kemi *et al*., 2013), *Arabidopsis halleri* (Aikawa *et al*., 2010), *Arabidopsis arenosa* (Baduel *et al*., 2016), *Arabis alpina* (Wang *et al*., 2009a, Lazaro *et al*., 2018), and *Capsella rubella* (Guo *et al*., 2012) and, like in *A. thaliana*, both genes are thought to be involved in the floral transition. Variation in *FLC* expression, in particular, has been demonstrated to contribute to differences in life history strategy both in annual and perennial Brassicaceae species (Kemi *et al*., 2013, Aikawa *et al*., 2010, Baduel *et al*., 2016, Wang *et al*., 2009a, Guo *et al*., 2012). For example, down-regulation of the *FLC* orthologue, *PEP1*, in the perennial species *A. alpina* is responsible for promoting the floral transition (Wang *et al*., 2009a). However unlike in *A. thaliana*, the silencing of *PEP1* expression is not stable in all meristems and re-activation of *PEP1* expression ensures some meristems flower in response to vernalisation while others remain vegetative (Wang *et al*., 2009a, Lazaro *et al*., 2018).

1.3.7. Long days promote flowering in *A. thaliana*

Day-length is perceived in the leaf by photoreceptors and the circadian clock (Turck *et al*., 2008). In *A. thaliana*, a facultative long-day plant, a key gene involved in the regulation of flowering by photoperiod is the plant-specific CCT domain transcription factor *CONSTANS (CO)* (Figure 1. 2, Putterill *et al*., 1995). *CO* is controlled by the circadian clock and has a 24-hour cyclical pattern in both gene and protein expression level (Suarez-Lopez *et al*.,2001). The external coincidence model proposes that the circadian clock generates an internal photoperiodic rhythm which is sensitive to light at certain phases (Bastow and Dean 2002). CO protein expression accumulation peaks during the evening (Suarez-Lopez *et al*., 2001). If CO protein expression peaks during daylight, which is more likely under long-day conditions, CO protein activity is stabilised, which promotes the activation of *FT* (Suarez-Lopez *et al*., 2001, Yanovsky and Kay 2002). CO protein binds cis-regulatory elements within the *FT* promoter via interaction with the Nuclear Factor-Y (NF-Y) proteins and the TGACG Motif Binding Factor 4 (TGA4) (Ben-Naim *et al*., 2006, Wenkel *et al*., 2006, Song *et al*., 2008, Tiwari *et al*., 2010, Adrian *et al*., 2010). Under short-day conditions *FT* expression is not induced due to the ubiquitination and degradation of CO protein which is influenced by the photoreceptor phyB (Valverde *et al*., 2004).

FT mRNA is only expressed in the vascular tissue (Takada and Goto 2003), however its protein functions in the shoot apical meristem (Wigge *et al*., 2005). Grafting experiments in *A. thaliana* revealed movement of the FT protein and not the mRNA from the leaves to meristematic tissue is required to induce flowering (Corbesier *et al*., 2007, Jaeger and Wigge 2007). FT protein function is dependent on a bZIP transcription factor *FLOWERING LOCUS D* (*FD*) (Abe *et al*., 2005, Wigge *et al*., 2005). The FT protein interacts with FD protein in the shoot apex (Abe *et al*., 2005, Wigge *et al*., 2005), likely mediated by the 14-3-3 protein (Taoka *et al*., 2011) to regulate the expression of *AP1* and *FUL* (Abe *et al*., 2005, Turck *et al*., 2008).

Six *FT*-related genes are found in *A. thaliana* and all encode proteins that contain a phosphatidylethanolamine binding domain (PEPB) (Turck *et al*., 2008). In addition to FT, the *A. thaliana* family of PEPB proteins include TWIN SISTER OF FT (TSF), BROTHER OF FT (BFT), ARABIDOPSIS THALIANA RELATIVE OF CENTRORADIALIS (ATC), MOTHER OF FT AND TFL1 (MFT), and TERMINAL FLOWER 1 (TFL1). Among them the function of TSF and TFL1 are best understood. TSF protein shows the highest similarity to the FT protein, with 81.3% amino acid sequence identity, both acting redundantly to promote flowering (Yamaguchi *et al*., 2005, Turck *et al*.,2008). TFL1 protein acts antagonistically to FT protein. Its expression is detected in the shoot apical meristem and functions to inhibit flowering (Shannon *et al*., 1991, Ahn *et al*., 2006). Interestingly a single amino acid substitution can convert TFL1, a repressor of flowering, to the floral promoting function of the FT protein suggesting specific residues have been selected for alternative protein function in *A. thaliana* (Hanzawa *et al*., 2005).

1.3.8. Gibberellin increases the expression of *SOC1*

The plant hormone gibberellin (GA) has long been known to promote flowering (Figure 1. 2; Langridge 1957). The GA-deficient mutant (*ga1-3*) flowers later than wild-type *A. thaliana* under both short and long days (Wilson *et al*., 1992) due to reduced expression of the floral integrator *SOC1* (Moon *et al*., 2003). *SOC1* encodes

a MADS box transcription factor, and like *FT*, integrates signals from multiple floral regulatory pathways (Figure 1. 2). Its expression is promoted by GA and by FT protein (Moon et a. 2003, Yoo *et al*., 2005, Wigge *et al*., 2005). SOC1 protein, in combination with AGAMOUS-LIKE 24 (AGL24), promote flowering by binding directly to the promoter of *LFY*, increasing its expression (Lee *et al*., 2008, Liu *et al*., 2008).

1.4. Homologous genes also control flowering time in *Brassica*

Variation for flowering time is present within *B. napus* and crop types are often grouped based on their vernalisation requirement and flowering time. The most commonly grown sub-species is *B. napus* ssp. *napus* (oilseed rape). Three main types are grown worldwide and each exhibit different vernalisation requirements; spring oilseed rape varieties exhibit no vernalisation requirement, Chinese semi-winter oilseed rape varieties exhibit a mild requirement, while winter oilseed rape varieties show a strong requirement for vernalisation. *B. napus* ssp. *napus* var. *pabularia* (kale) and *B. napus* ssp. *rapifera* (swede) are grown as vegetable crops and exhibit strong to obligate vernalisation requirements, flowering late or not at all without cold (Harper *et al*., 2012). For oilseed rape, the floral transition is necessary for seed production while for the vegetable crops flowering can limit yield (Jung and Müller 2009). The diversity of flowering habits within the *B. napus* species indicates the potential of identifying allelic variation that either promote, inhibit, or modulate flowering. Identification of the functional consequence of variation at those genes would provide an excellent source of genetic information that, in the long-term, could be exploited by marker assisted selection for crop improvement (Jung and Müller 2009).

1.4.1. *FRI* orthologous are identified in *Brassica*

FRI orthologues have been identified in the diploid species *B. rapa* (Fadina *et al*., 2013, Fadina & Khavkin 2014) and *B. oleracea* (Irwin *et al*., 2012) and the allotetraploid species *B. napus* (Wang *et al*., 2011b, Fadina & Khavkin 2014, Yi *et*

al., 2018) sharing up to 75% sequence identity with *FRI* from *A. thaliana* (Fadina *et al*., 2013). Two copies of *FRI* are present in both the *B. rapa* and *B. oleracea* genomes, originally named *BraA.FRI.a, BraA.FRI.b, Bol.C.FRI.a, Bol.C.FRI.b*. They were later mapped to syntenic regions on chromosomes A03 and A10 in *B. rapa* and to chromosomes C03 and C09 in *B. oleracea* (Schiessl *et al*., 2017) and hereafter referred as *BraFRI.A03, BraFRI.A10, BolFRI.C03, BolFRI.C09* respectively.

Interestingly *Brassica FRI* genes are found in regions that are syntenic to the *A. thaliana* chromosome 5 and not to chromosome 4 where *FRI* (AT4G00650) is located (Irwin *et al*., 2012). Interestingly the orthologue of *FRI* in *A. lyrata* also maps to the same region that is syntenic to chromosome 5 in *A. thaliana* (Kuittenen *et al*., 2007). A homology search of this region of chromosome 5 in *A. thaliana* identified a gene (AT5G51090) that is highly similar in sequence to *FRI* containing parts of intron 1 and exon 3, characterised a pseudogene due to its low expression level in *A. thaliana* (Irwin *et al*., 2012). It was proposed from this that during the evolution of the *A. thaliana* lineage, a genomic region containing *FRI* was translocated from chromosome 5 to the upper arm of chromosome 4 leaving behind a non-functional portion of the original gene (Irwin *et al*., 2012).

FRI genes exhibit high sequence homology within and between *Brassica* species (Fadina *et al*., 2013, Fadina & Khavkin 2014). Paralogous loci within the diploid species; *BraFRI.A03* versus *BraFRI.A10* in *B. rapa*, and *BolFRI.C03* versus *BolFRI.C09* in *B. oleracea*; exhibit the most sequence divergence (79-81% identity) while homologous loci between diploid species; *BolFRI.C03* versus *BraFRI.A03*, and *BraFRI.A10* versus *BolFRI.C09*; are more conserved (87-94% identity) (Fadina *et al*.,2013, Fadina & Khavkin 2014). Four orthologues of *FRI* have been identified in the allotetraploid *B. napus* (Wang *et al*., 2011b, Fadina *et al*., 2013). These were initially named *BnaX.FRI.a, BnaX.FRI.b, BnaX.FRI.c, BnaX.FRI.d* and later mapped to chromosomes A03, A10, C09 and C03 respectively (Schiessl *et al*., 2014) hereafter referred as *BnaFRI.A03, BnaFRI.A10, BnaFRI.C09* and *BnaFRI.C03* respectively. All four genes exhibit high sequence similarity (95-99% identity) with

their respective homologous genes in the diploid progenitor species (Fadina *et al*., 2013). Retention of multiple copies of *FRI* and the high conservation of gene sequence within and between species suggests conservation of function in controlling the floral transition in *Brassica*.

All published *Brassica FRI* genes are predicted to encode functional proteins (Wang *et al*., 2011b, Irwin *et al*., 2012, Fadina *et al*., 2013, Fadina & Khavkin 2014, Yi *et al*., 2018). Their predicted protein sequences are shorter than the 609 amino acids of the functional FRI protein in *A. thaliana* ranging from 569 to 596 amino acids (Wang *et al*., 2011b, Irwin *et al*., 2012, Fadina *et al*., 2013, Fadina & Khavkin 2014). BraFRI.A03, BnaFRI.A03, BolFRI.C03 and BnaFRI.C03 are predicted to contain two coiled-coil domains, while BraFRI.A10, BnaFRI.A10, BolFRI.C09 and BnaFRI.C09 contain one coiled-coil domain in the C-terminal region of the protein (Wang *et al*., 2011b, Irwin *et al*., 2012). *Brassica* and *A. thaliana* FRI exhibit extensive conservation of amino acid sequence within the central and C-terminal regions of their predicted protein (Wang *et al*., 2011b, Irwin *et al*., 2012, Fadina & Khavkin 2014). In *A. thaliana* the C-terminal domain has proved to be more important than the N-terminal domain for function (Risk *et al*., 2010), perhaps explaining the high degree of conservation at the C-terminus between species.

The exon-intron structure of *FRI* is conserved between *A. thaliana* and *Brassica* species (Wang *et al*., 2011b, Fadina *et al*., 2013, Fadina & Khavkin 2014). All four *FRI* homologues found in *B. napus*; *BnaFRI.A03, BnaFRI.C03, BnaFRI.A10* and *BnaFRI.C09* are expressed in spring, Chinese semi-winter and winter cultivars of oilseed rape (Wang *et al*., 2011b). Expression of *BnaFRI.A10* and *BnaFRI.C09* exhibit relatively constant levels in all plant tissues, while *BnaFRI.A03* and *BnaFRI.C03* exhibit a peak in expression in floral tissues (Wang *et al*., 2011b). The expression profile of *BnaFRI.A03* was later confirmed quantitatively and the highest expression was detected in seeds (Yi *et al*., 2018). Transient expression in tobacco suggests BnaFRI.A03 protein functions in the nucleus, possibly to regulate the expression of its target gene (Yi *et al*., 2018).

BolFRI.C03 and *BnaFRI.A03* both confer later flowering when transformed into and *A. thaliana fri* mutant (Irwin *et al*., 2012, Yi *et al*., 2018). When expressed under the *A. thaliana FRI* promoter two alleles of *BolFRI.C03* that associated with early and late flowering cultivars respectively conferred equally late flowering, suggesting the coding variation between the two alleles had no functional consequence on flowering time (Irwin *et al*., 2012). It was predicted from this that if allelic variation at *BolFRI.C03* was responsible for the flowering time variation between cultivars of *B. oleracea,* gene expression differences between the alleles of *BolFRI.C03* must be the causative difference (Irwin *et al*., 2012). Allelic variation at *BnaFRI.A03* is associated with flowering time variation in *B. napu*s (Wang *et al*., 2011b) and when transformed into *A. thaliana,* two alleles of *BnaFRI.A03* conferred significantly different flowering times (Yi *et al*., 2018). Interestingly allelic variation within promoter regions of *BnaFRI.A03* had no significant effect on flowering time or *FLC* expression in *A. thaliana,* while transformants carrying allelic variation within coding regions showed differential flowering times (Yi *et al*., 2018). Contrary to the results of *BolFRI.C03,* coding variation, and not expression variation at *BnaFRI.A03* was responsible for the flowering time differences associated with the allelic variation (Irwin *et al*., 2012, Yi *et al*., 2018). Interestingly, significantly different flowering times were detected between the alleles of *BnaFRI.*A03 when transformed into *A. thaliana*, conferring an average of 13.1 days difference (Yi *et al*., 2018). Due to the high sequence similarity between *BnaFRI.A03* and *BolFRI.C03* it is surprising that transgenic analyses of different alleles of both genes would give contrasting results, supporting a need for further investigation of the function of *FRI* orthologues in *Brassica* species.

1.4.2. *FLC* orthologues are identified in *Brassica*

Brassica FLC genes exhibit up to 75% sequence identity with *FLC* from *A. thaliana*, but a higher conservation of sequence is observed for coding regions (85% sequence identity) (Schranz *et al*., 2006). Comparisons made between *A. thaliana FLC* and *Brassica FLC* homologues revealed high conservation in both exon-intron structure and exon size (Schranz *et al*., 2006). Intron size and sequence however were diverged between the species (Schranz *et al*., 2006). Four *BraFLC* (Schranz *et al*., 2006) and four *BolFLC* (Schranz *et al*., 2006, Okazaki *et al*., 2007) genes have been

characterised and mapped to chromosomes A02, A03 and A10 in *B. rapa*, and to chromosomes C02, C03 and C09 in *B. oleracea* (Table 1. 1)

Table 1. 1: Four *FLC* clades are identified in diploid *Brassica*. Included in the table are the names and alternative names of the *FLC* homologues characterised in *B. rapa* and *B. oleracea.*

Clade	*B. rapa*		*B. oleracea*	
	Name	Alternative Name	Name	Alternative Name
FLC1	*BraFLC.A10*	*BrFLC1*	*BolFLC.C09*	*BoFLC1*
FLC2	*BraFLC.A02*	*BrFLC2*	*BolFLC.C02*	*BoFLC2*
FLC3	*BraFLC.A03a*	*BrFLC3*	*BolFLC.C03a*	*BoFLC3*
FLC5	*BraFLC.A03b*	*BrFLC5*	*BolFLC.C03b*	*BoFLC5*

Nine orthologues of *FLC* have been characterised in *B. napus* and mapped to six chromosomes; A02, A03, A10, C02, C03, and C09 (Table 1. 2; Zou *et al*., 2012). Most *BnaFLC* genes are found in genomic regions that are syntenic to the *A. thaliana* chromosome 5 where *FLC* is located. This includes *BnaFLC.C09* which appears as a tandem duplication on chromosome C09 (*BnaFLC.C09a* and *BnaFLC.C09b*). The exceptions to this are *BnaFLC.A03b* and *BnaFLC.C03b* which are mapped to a region that is syntenic to *A. thaliana* chromosome 2 (Schranz *et al*., 2006, Zou *et al*., 2012) suggesting subsequent duplication and translocation of *FLC* occurred within the evolution of *B. napus*. The nine *BnaFLC* genes are structured into seven exons and six introns and like other *Brassica FLC* genes can be divided into four clades. As is the case for *BnaFRI, BnaFLC* gene sequences show extremely high conservation (95-100%) to their homologous genes in *B. rapa* and *B. oleracea* (Zou *et al*., 2012). Due to this high sequence similarity between *FLC* homologues it is predicted that their functions are conserved between the three species.

Table 1. 2: Four *FLC* clades are identified in *B. napus*. Included in the table are the names and alternative names of the *FLC* homologues characterised in *B. napus.*

	B. napus			
	A genome		C genome	
Clade	Name	Alternative Name	Name	Alternative Name
FLC1	*BnaFLC.A10*	*BnFLC1*	*BnaFLC.C09a* *BnaFLC.C09b*	*BnFLC1*
FLC2	*BnaFLC.A02*	*BnFLC2*	*BnaFLC.C02*	*BnFLC4*
FLC3	*BnaFLC.A03a*	*BnFLC3*	*BnaFLC.C03a*	*BnFLC3*
FLC5	*BnaFLC.A03b*	*BnFLC5*	*BolFLC.C03b*	*BnFLC5*

Expression analysis of the nine *BnaFLC* genes showed that *BnaFLC.C03b* is likely a pseudogene as it is not expressed in oilseed rape, however expression of the remaining 8 *BnaFLC* genes was detected (Zou *et al*., 2012). Differential *BnaFLC* expression levels were detected between oilseed rape cultivars that exhibited variation for vernalisation requirement. Lateness of flowering correlated with *BnaFLC* transcript level and the latest cultivars exhibited the highest combined expression of *BnaFLC*. In late flowering winter oilseed rape, five *BnaFLC* genes namely *BnaFLC.A10*, *BnaFLC.A02*, *BnaFLC.A03a*, *BnaFLC.A03b* and *BnaFLC.C03a*, exhibited quantitative differences in expression after vernalisation (Zou *et al*., 2012). This difference in expression before and after vernalisation treatment suggests at least some *BnaFLC* genes could contribute to the vernalisation requirement and response of *B. napus*.

When expressed in *A. thaliana*, *Brassica FLC* homologues function as floral repressors (Tadege *et al*., 2001, Kim *et al*., 2007, Irwin *et al*., 2016). As an example, cDNA from five *BnaFLC* genes, namely *BnFLC1, BnFLC2, BnFLC3, BnFLC4 and BnFLC5*, were transformed into the *A. thaliana* accession Landsberg *erecta* (L*er*) which flowers early without vernalisation (Tadege *et al*., 2001). Intronic regions from the *BnaFLC* genes were not included in the constructs and expression was driven by the constitutive overexpressing promoter 35SCaMV therefore their

response to vernalisation could not be tested (Tadege *et al*., 2001). However, all five *BnaFLC* constructs delayed flowering by at least 3 weeks indicating the BnaFLC protein can repress flowering in *A. thaliana*. Constructs containing *BnFLC4* and *BnFLC5* exhibited the earliest flowering and therefore could be considered weak inhibitors of flowering time. Constructs containing *BnFLC1*, in contrast, flowered the latest, with some lines never flowering. The *BnaFLC* gene locaed on chromosomes A10 or C09 might therefore be a stronger inhibitor of flowering compared with the other *BnaFLC* genes tested. Retention of so many floral repressing *BnaFLC* genes within the *B. napus* genome poses questions about how each one contributes to the floral transition.

FLC homologues can perform similar functions when transferred between species of the Brassicaceae (Tadege *et al*., 2001, Kim *et al*., 2007, Irwin *et al*., 2016) however subtle differences in function and gene expression patterns have been identified. The *FLC* gene from *A. thaliana* conferred late flowering when transformed into a spring cultivar of *B. napus* confirming conservation of the floral initiation mechanism between both species (Tadege *et al*., 2001). Interestingly expression of *A. thaliana FLC* in *B. napus* impacted inflorescence development and fertility highlighting differences between *A. thaliana FLC* and flower development in *B. napus*.

Functional differences between the *A. thaliana FLC* and *BolFLC.C02* were highlighted by Lin *et al*., (2005). A BolFLC.C02::GUS fusion protein construct was not down-regulated by vernalisation when expressed in *A. thaliana*. Interestingly, *BolFLC.C02* is down-regulated by vernalisation in *B. oleracea* leading the authors to conclude the *FLC* regulatory machinery of *A. thaliana* is different from *B. oleracea* (Lin *et al*., 2005). Promoter-GUS fusion constructs containing the promoter regions from *BraFLC.A10, BraFLC.A02* or *BraFLC.A03a* exhibited differential tissue specific expression in *A. thaliana* seedlings (Hong *et al*., 2011). pBraFLC.A10::GUS expression was only detected in the apical meristem region, while pBraFLC.A02::GUS and pBraFLC.A03a::GUS expression were strongly detected in roots and in leaves of *A. thaliana* (Hong *et al*., 2011) suggesting the possibility that

FLC genes have undergone some sub-functionalisation resulting in differential expression between plant tissues.

Unpublished data from the Dean and Irwin Laboratories have revealed *FLC* expression in *A. thaliana* is regulated differently to *BolFLC.C02* (Woodhouse 2017). *BolFLC.C02* complements an *A. thaliana flc* mutation, conferring late flowering in the absence of vernalisation (Irwin *et al*., 2016). In the presence of *FRI*, *A. thaliana* pFLC::FLC::GUS fusion expression was detected in all plant tissues, but in contrast pBolFLC.C02::BolFLC.C02::GUS expression was detectable only in the apical meristem. Interestingly in separate transgenic lines that express *FLC* without GUS *BolFLC.C02* was expressed at significantly lower levels compared with the native *FLC* of *A. thaliana*. It is curious that such low quantitative levels of *BolFLC.C02* can confer late flowering phenotypes similar to an *A. thaliana* accession with a more highly expressed native *FLC*. It is clear from these data however that the interaction between the native *FRI* of *A. thaliana* and *BolFLC.C02* leads to a different overall spatial pattern of expression.

1.4.3. *COOLAIR* transcripts are identified in *Brassica*

COOLAIR transcripts influence the initial level of *FLC* expression in *A. thaliana* (Li *et al*., 2015). Transcripts that are homologous to the *A. thaliana COOLAIR* have been detected in *B. rapa* (Li *et al*., 2016, Hawkes *et al*., 2016), *B. oleracea* and *B. napus* (Hawkes 2017). Like in *A. thaliana*, *COOLAIR* transcripts identified in *Brassica* can be grouped into proximal poly-adenylated and distal poly-adenylated forms which are up-regulated in response to cold temperatures (Li *et al*., 2016, Hawkes 2017). *COOLAIR* has a complex secondary structure and mutations that disrupt this structure correlate with differences in *FLC* sense expression and flowering time (Li *et al*., 2015, Hawkes *et al*., 2016, Hawkes 2017). Despite low sequence conservation within the non-coding regions of *FLC* homologues, key secondary structural features of *COOLAIR* are conserved between species suggesting functional conservation of *COOLAIR* within the Brassicaceae (Hawkes *et al*., 2016, Hawkes 2017). Interestingly, overexpression of a *COOLAIR* transcript expressed from the *BraFLC.A02* locus in *B. rapa* resulted in early flowering in the absence of

vernalisation and a reduction in the expression of all *BraFLC* genes (Li *et al*., 2016). This *COOLAIR* transcript, when overexpressed, was able to influence the expression, and therefore likely the chromatin state, at multiple *BraFLC* genes implicating *COOLAIR* in the regulation of *FLC* in *B. rapa*.

1.4.4. *FT* orthologues are identified in *Brassica*

Six orthologues of *FT* have been characterised in *B. napus* (Wang *et al*., 2009b). Present on chromosomes A02, A07, C02 and C06 they will be referred as *BnaFT.A02*, *BnaFT.A07a*, *BnaFT.A07b*, *BnaFT.C02*, *BnaFT.C06a* and *BnaFT.C06b*. All six exhibit 85-87% coding sequence identity with *FT* from *A. thaliana* and are present within genomic regions that are syntenic to *A. thaliana* chromosome 1 where *FT* is located (Wang *et al*., 2009b). *BnaFT* genes are lowly expressed prior to the floral transition and in the winter oilseed rape cultivar Express-617 the expression of *BnaFT.A07a*, *BnaFT.A07b*, *BnaFT.C06a* and *BnaFT.C06b* increased rapidly after vernalisation, exhibiting a peak in expression at the floral stage (Guo *et al*., 2014). The expression of *BnaFT.A02* expression was detected at low levels, and *BnaFT.C02* was not detectable, even after vernalisation treatment (Guo *et al*., 2014), suggesting at least five *BnaFT* genes could contribute to the regulation of flowering time in *B. napus*. A CArG box domain located within the first intron of *FT* in *A. thaliana* is the likely binding site for *FLC* (Helliwell *et al*., 2006). CArG are predicted within introns 1 of *BnaFT.A07a*, *BnaFT.A07b*, *BnaFT.C06a* and *BnaFT.C06b* but not *BnaFT.A02* and *BnaFT.C02* (Wang *et al*., 2009b). This highlights a potential difference in the regulation mechanism between the multiple *BnaFT* genes by *BnaFLC*.

1.5. Aims

Despite high sequence conservation between *A. thaliana* and *B. napus* the presence of multiple orthologous genes complicates translation of the floral regulatory network from model to crop species (Jones *et al*., 2018). Multiple orthologues of flowering time genes have been preferentially retained and are expressed in *B. napus* (Schiessl *et al*., 2014, Jones *et al*., 2018). Due to their evolutionary history, each *A. thaliana* gene is predicted to occur as six orthologous genes in *B. napus*. In many cases the number of flowering time gene orthologues exceeds this (Schiessl *et al*., 2014, Jones *et al*., 2018) suggesting subsequent, more recent, duplication events have occurred.

Multiple studies have identified orthologues of *FLC* that perform a conserved role in the floral transition and vernalisation requirement in *B. rapa*, *B. oleracea*, and *B. napus* (Tadege *et al*., 2001, Kim *et al*., 2007, Irwin *et al*., 2016). Curiously, the function of *FRI* orthologues in *Brassica* is less clear (Wang *et al*., 2011b, Irwin *et al*., 2012, Fadina *et al*., 2013, Fadina & Khavkin 2014, Yi *et al*., 2018). Despite an association between allelic variation at *BolFRI.C03* and *BnaFRI.A03* with flowering time and crop type (Wang *et al*., 2011b, Irwin *et al*., 2012, Yi *et al*., 2018) these genes have a relatively minor effect on flowering time when expressed in *A. thaliana* (Irwin *et al*., 2012, Yi *et al*., 2018).

Using a candidate gene approach, and two association mapping methods, I identify and investigate genomic regions that are important for variation in the vernalisation requirement of *B. napus*. The experimental research I describe in this Thesis addresses three main aims:

1. Determine the function of *FRI* orthologues in controlling the vernalisation requirement of *B. napus*

I screened a diversity set population of *B. napus* for allelic variation at *BnaFRI* genes. Due to a near absence of variation that would disrupt the BnaFRI protein,

their functions were assessed by introducing induced mutations at *BnaFRI* and by a transgenic analysis of *BnaFRI.A03*.

2. Identify genetic regions that are important for controlling variation in vernalisation requirement and response in *B. napus* and investigate candidate genes

I used two association genetics methods; Associative Transcriptomics (Harper *et al*., 2012; described in Chapter 4) and QTL-seq (Takagi *et al*., 2013; described in Chapter 5) to identify candidate genes important for flowering time with and without vernalisation. DNA sequence and gene expression variation were detected at three candidate genes, *BnaFLC.A10*, *BnaFLC.A02* and *BnaFT.A02*, which associated with flowering time.

3. Determine whether flowering time and vernalisation requirement has a pleiotropic effect on inflorescence architecture in *A. thaliana* and *B. napus*

Flowering time correlates with branch number, likely impacting yield. I assessed the correlation between flowering time and inflorescence architecture with and without vernalisation in *A. thaliana* and *B. napus* and discuss a role for *FLC* in both traits.

Chapter 2: Methods

2.1. Genome Resources

The source of genome reference sequences I used to complete this work are listed in Table 2. 1.

Table 2. 1: Source of genome reference sequences used in this Thesis

Species	**Accession/Cultivar**	**Source**
Arabidopsis thaliana	Col-0	The Arabidopsis Genome Initiative (2000) (www.arabidopsis.org)
Brassica napus	Darmor-*bzh*	Chalhoub *et al*., 2014 (plants.ensembl.org)
	Cabriolet	In-house reference R. Wells unpublished
Brassica rapa	Chiifu	Wang *et al*., 2011a (plants.ensembl.org)
	R-0-18	In-house reference
Brassica oleracea	T01000	Parkin *et al*., 2014 (plants.ensembl.org)

2.2. DNA Sequence datasets

Details of *BnaFRI, BnaFLC* and *BnaFT* sequences I used as reference sequences are listed in Table 2. 2. Published gene sequences were originally downloaded from GenBank (www.ncbi.nlm.nih.gov/genbank). All other *BnaFRI, BnaFLC* and *BnaFT* sequences described in this text were generated by myself (the author; ET) or by the Judith Irwin Laboratory using targeted sequence capture (myBaits®; https://arborbiosci.com).

Table 2. 2: Source of *BnaFRI, BnaFLC* and *BnaFT* sequences used as reference sequences in this Thesis

Species	**Sequence (GenBank Accession code)**	**Cultivar**	**Source**
Brassica napus	*BnaFRI.A03* (JN936850) *BnaFRI.A10* (JN936851) *BnaFRI.C03* (JN936853) *BnaFRI.C09* (JN936852)	Express	Wang *et al*., 2011b
	BnaFLC.A02 (JQ255384) *BnaFLC.A03a* (JQ255386) *BnaFLC.A03b* (JQ255388) *BnaFLC.A10* (JQ255381) *BnaFLC.C02* (JQ255385) *BnaFLC.C03b* (JQ255389) *BnaFLC.C09a* (JQ255382) *BnaFLC.C09b* (JQ255383)	Tapidor	Zou *et al*., 2012
	BnaFLC.C03a (JQ255387)	Ningyou-7	Zou *et al*., 2012
	BnaFLC.A10 (JX901142)	Tapidor	Hou *et al*., 2011
	BnaFLC.A10 (JX901141)	Ningyou-7	Hou *et al*., 2011
	BnaFT.A02 (FJ848913) *BnaFT.A07a* (FJ848918) *BnaFT.A07b* (FJ848916) *BnaFT.C02* (FJ848914) *BnaFT.C06a* (FJ848915) *BnaFT.C06b* (FJ848917)	Tapidor	Wang *et al*., 2009b
	BnaFLC.A02	95 *B. napus* cultivars (Table 2.5)	Judith Irwin Laboratory (unpublished); (myBaits®; https://arborbiosci.com)
	BnaFT.A02		

2.3. Plant materials and growth conditions

2.3.1. Plant materials – *Arabidopsis thaliana*

The *A. thaliana* accessions, and their seed source, analysed are listed in Table 2. 3.

Table 2. 3: *A. thaliana* accessions, and their seed source, analysed in this Thesis

Species	Accession name	Accession details	Source
Arabidopsis thaliana	Col-0	Natural accession Columbia	In-house; Huamei Wang, Caroline Dean Laboratory
	Col-*FRI*^{SF2}	Natural accession Columbia carrying *FRI* locus introgressed from natural accession San Feliu 2 (Lee *et al*., 1994b)	
	Col-*FRI*^{SF2}(*flc-3*)	Generated by fast neutron radiation of Col-*FRI*^{SF2}; carries 104bp deletion over the start codon of *FLC* (Michaels and Amasino 1999)	
	Edi-0	Natural accession from Edinburgh, Scotland	
	Edi-NIL	Near Introgression Line of Col*FRI*^{SF2} carrying the *FLC* locus from accession Edi-0	
	Lov-1	Natural accession from Lövick, North Sweden	
	Lov-NIL	Near Introgression Line of Col*FRI*^{SF2} carrying the *FLC* locus from accession Lov-1	

2.3.2. Plant materials – *Brassica napus*

Unless otherwise described, *B. napus* cultivars were sourced from the JIC core diversity fixed foundation set (DFFS) maintained by the oilseed rape genetic improvement network (OREGIN; http://www.herts.ac.uk/oregin). Table 2. 5 lists the names and codes of each *B. napus* cultivar I used for the work described in this Thesis.

Table 2. 4*: B. napus* cultivar names and crop type described in this Thesis. Cultivars that have a "cultivar code" were included in the Associative Transcriptomics experiment described in Chapter 4, "N/I" indicates not included. DNA was extracted from all cultivars listed with a DNA sample number and used for genotyping by PCR, and for targeted sequence capture. OSR=oilseed rape.

Species	**Cultivar name**	**Cultivar code**	**Crop type**	**Cultivar (DNA) number**
Brassica napus	Aberdeenshire Prize	N/I	SWEDE	70
	Abukuma Natane	IB02_AbuN	WINTER OSR	11
	Altasweet	D3_ALT	SWEDE	36
	Amber X Commanche DH Line	D4_AMBxCOM	WINTER OSR	62
	Apex	D6_APE	MODERN WINTER OSR	14
	Apex-93_5 X Ginyou_3 DH Line	D7_APExGIN	WINTER OSR	n/a
	Assyst 078 (Express-617)	N/I	WINTER OSR	96
	Assyst 201 (Silona)	N/I	WINTER FODDER	30
	Assyst 291 (Campino)	N/I	SPRING OSR	29
	Assyst 452 (Magres Pajberg)	N/I	SWEDE	17
	Baltia	D10_BAL	WINTER OSR	53
	Bienvenu DH4	A11_Bie	WINTER OSR	48
	Bolko	D14_Bol	WINTER OSR	59
	Brauner Schnittkohl	IB15_BraS	SIBERIAN KALE	9
	Bronowski	A18_Bro	SPRING OSR	20
	Cabernet	IB21_Cab	WINTER OSR	81
	Cabriolet	IB22_Cab	WINTER OSR	82

Species	Cultivar name	Cultivar code	Crop type	Cultivar (DNA) number
Brassica napus	Canard	IB23_Can	WINTER FODDER	45
	Canberra X Courage DH Line	IB24_CanxCou	WINTER OSR	83
	Capitol	A25_Cap	MODERN WINTER OSR	13
	Capitol X Mohican DH Line	D26_CAPxMOH	WINTER OSR	60
	Castille	IB28_Cas	WINTER OSR	84
	Catana	IB29_Cat	WINTER OSR	85
	Ceska Krajova	D31_CES	SPRING OSR	24
	Chuanyou 2	A34_Chu	SEMIWINTER OSR	6
	Columbus X Nickel DH Line	IB37_ColxNic	WINTER OSR	86
	Coriander	D39_COR	WINTER OSR	61
	Couve Nabica	R_IB40_CouN	EXOTIC LEAFY	42
	Darmor	A42_Dar	WINTER OSR	51
	Dimension	IB45_Dim	WINTER OSR	87
	Dippes	D46_DIP	WINTER OSR	63
	Drakkar	R_A48_Dra	SPRING OSR	33
	Duplo	IB49_Dup	SPRING OSR	31
	Dwarf Essex	IB50_DwaE	WINTER FODDER	46
	Erglu	IB53_Erg	SPRING OSR	32
	Eurol	D55_EUR	WINTER OSR	64
	Excalibur	IB57_Exc	WINTER OSR	88
	Expert	IB58_Exp	MODERN WINTER OSR	16

Species	Cultivar name	Cultivar code	Crop type	Cultivar (DNA) number
Brassica napus	Flash	IB60_Fla	WINTER OSR	89
	Hanna	D69_HAN	SPRING OSR	25
	Hansen X Gaspard DH Line	IB70_HanxGas	WINTER OSR	90
	Huguenot	N/I	SWEDE	40
	Huron X Navajo DH Line	D74_HURxNAV	WINTER OSR	66
	Inca X Contact DH Line	D75_INCxCON	WINTER OSR	67
	Jaune A Collet Vert	R_IB78_JauCV	WINTER OSR	41
	Jetneuf	D79_JETN	SWEDE	69
	Judzae	N/I	WINTER OSR	47
	Karoo-057DH	D83_KAR	SPRING OSR	26
	Lembkes Malchower (Lenora)	IB86_LemM	WINTER OSR	91
	Licrown X Express DH Line	IB91_LicxExp	WINTER OSR	92
	Madrigal X Recital DH Line	IB99_MadxRec	WINTER OSR	93
	Major	IB100_Maj	WINTER OSR	71
	Matador	D102_MAT	WINTER OSR	54
	Moana, Moana Rape	IB106_MoaR	WINTER FODDER	43
	Monty-028DH	D108_MON	SPRING OSR	21
	N01D-1330	IB110_N01D_1330	SPRING OSR	27
	N02D-1952	IB111_N02D_1952	SPRING OSR	28

Species	Cultivar name	Cultivar code	Crop type	Cultivar (DNA) number
Brassica napus	Ningyou 7	NIN	SEMIWINTER OSR	7
	Nk Bravour	IB16_Bra	MODERN WINTER OSR	15
	Norin	D117_NOR	WINTER OSR	55
	Palmedor	IB120_Palm	WINTER OSR	72
	Palu	IB121_Palu	WINTER FODDER	44
	Poh 285, Bolko	IB125_POH285B	WINTER OSR	73
	Primor	D127_PRI	WINTER OSR	56
	Q100	D128_Q10	SYNTHETIC	10
	Quinta	IB130_Qui	WINTER OSR	74
	Rafal DH1	D131_Raf	WINTER OSR	57
	Ragged Jack	D132_RAGJ	RAPE KALE	1
	Ramses	IB133_Ram	WINTER OSR	75
	Rapid Cycling Rape (Crgc5)	D134_RAPCR	RAPE KALE	2
	Regent	A135_Reg	SPRING OSR	18
	Rocket	A136_Roc	WINTER OSR	49
	ROCKET (Pst) X Lizard DH Line	D137_ROCxLIZ	WINTER OSR	58
	Samourai	IB139_Sam	WINTER OSR	76
	Sensation New Zealand	IB141_SenNZ	SWEDE	37
	Shannon X Winner DH Line	IB143_ShaxWin	WINTER OSR	77
	Shengliyoucai	A144_She	SEMIWINTER OSR	3
	Siberische Boerenkool	D146_SIBB	SIBERIAN KALE	8
	Slapska, Slapy	IB150_SlaS	WINTER OSR	52

Species	Cultivar name	Cultivar code	Crop type	Cultivar (DNA) number
Brassica napus	Slovenska Krajova	IB151_SloK	UNSPECIFIED	78
	Stellar DH	A155_Ste	SPRING OSR	19
	Taisetsu	IB164_Tai	WINTER VEGETABLE	68
	Tapidor DH	TAP	WINTER OSR	95
	Temple	A168_TEM	WINTER OSR	50
	Tequilla X Aragon DH Line	IB169_TeqxAra	WINTER OSR	79
	Tina	D170_TIN	SWEDE	34
	Topas	D173_TOP	SPRING OSR	22
	Verona	A176_Ver	MODERN WINTER OSR	12
	Victor	R_IB177_Vic	WINTER OSR	94
	Vige DH1	IB178_Vig	SWEDE	38
	Vision	IB180_Vis	WINTER OSR	80
	Westar DH10	D181_WESD	SPRING OSR	23
	Wilhelmsburger-Reform	D182_WILR	SWEDE	35
	Xiangyou 15	A187_Xia	SEMIWINTER OSR	4
	York	N/I	SWEDE	39
	Zhongshuang II	A188_Zho	SEMIWINTER OSR	5

2.3.3. Plant growth conditions – glasshouse

Plants grown under glasshouse conditions were sown, watered, and chemically sprayed (when required) by the Norwich BioScience Institute (NBI) Horticultural Services Team. *A. thaliana* and *B. napus* seed were sown directly onto soil (Levington F2 compost, 600L peat, 100L 4mm grit, 196g Exemptor (Chloronicotinyl

Insecticide)) and glasshouse conditions were as follows; long-day (16-hours light/8-hours dark, 600W HPS lamps provided supplementary lighting when required), 18°C day temperature, 15°C night temperature and 70% humidity. Plant were transplanted to larger pots, staked and bagged (330 x 180mm glassine bags for *A. thaliana*, 380 x 900mm microperforated cellophane bags for *B. napus*) when required, and harvested at senescence.

2.3.4. Plant growth conditions – vernalisation

After sowing, *A. thaliana* seeds were placed in a vernalisation chamber (5°C, 8-hours light/16-hours dark, 70% humidity) for three days for stratification. They were then transferred to long-day glasshouse conditions for 7-days growth before being returned to vernalisation for 2-12 weeks. After vernalisation plants were returned to glasshouse conditions (16-hours light/8-hours dark, 18°C day temperature, 15°C night temperature and 70% humidity) or controlled environment conditions (20°C, 16 hours light, 8 hours dark, 70% humidity), transplanted to larger pots, staked and bagged (330 x 180mm glassine bags) when required. Plants were harvested, and seeds threshed at senescence.

No stratification treatment is needed to promote germination in *B. napus* therefore seeds were sown directly onto soil and grown under glasshouse conditions (16-hours light/8-hours dark, 18°C day temperature, 15°C night temperature and 70% humidity). 21-28 days after sowing plants were transferred to a vernalisation chamber for 4-12 weeks (5°C, 8-hours light/16-hours dark, 70% humidity). After vernalisation, plants were returned to glasshouse conditions, transplanted to larger pots, staked and bagged (380 x 900mm microperforated cellophane bags for *B. napus*) when required. Plants were harvested, and seeds threshed at senescence.

2.3.5. Plant growth conditions – tissue culture

A. thaliana seeds were vapour-phase sterilized (Clough and Bent 1998) and sown on GLM media without glucose (1x Murashige and Skoog salts, 2mg/L Glycine, 100mg/L myo-Inositol, 0.5mg/L Nicotinic acid, 0.5mg/L Pyridoxine HCl, 0.1mg/L Thiamine, 0.5g/L MES, 1% agar). For selection of transgenic plants phosphinothricin (PPT) was added to the media at a concentration of 10µg/ml. Seeds

on plates were stratified in the dark at 10°C for three days before being transferred to a controlled environment room (20°C, 16 hours light, 8 hours dark, 70% humidity). Seedlings were transplanted to soil 12 days after stratification and remained in a controlled environment room for the duration of growth. Plants were staked and bagged (330 x 180mm glassine bags) when required and harvested at senescence.

2.4. Phenotypic analysis

I assessed flowering time and branch number in both *A. thaliana* and *B. napus*. In all cases, plants were randomized into blocks and at least three replicate plants for each accession/cultivar per treatment were scored for their flowering and branching phenotypes.

For *A. thaliana*, I scored flowering time as the number of days until buds were visible in the apex of the rosette. Scoring commenced after the vernalisation treatment had ended, or after transplanting to soil. Primary rosette and cauline branch number were scored at senescence.

For *B. napus*, I scored flowering time as the number of days until buds were visible (BBCH51 according to Meier 2001) and until the first flower opened (BBCH60 according to Meier 2001). Scoring commenced after the vernalisation treatment had ended. Primary branch number and leaf node number were scored at senescence.

Mean flowering time and branch number were calculated from at least three replicate plants for each cultivar and treatment. All statistical analyses were performed using GenStat V8 (VSN International).

2.5. Genotyping by PCR and sequencing

2.5.1. DNA extraction

When genotyping a small number of samples, fresh leaf material was first homogenized in a 1.7ml Eppendorf tube using blue polypropylene pellet pestles

(Sigma Aldrich). I subsequently extracted DNA from ground leaf tissue following the Edwards DNA Extraction Method (Edwards *et al*., (1991); Supplementary Methods). For large numbers of DNA extractions, leaf material was submitted in 96-well racked collection tubes (Qiagen) to the JIC Genotyping and DNA Extraction Service for processing.

2.5.2. Gradient and Standard PCR

I designed primers for PCR and sequencing to amplify products of ~1000bp with at least one of the primer pair specific to the gene copy of interest. When the fragment to be sequenced was larger than 1000bp, I designed multiple primer pairs to amplify overlapping regions so that sequences could be aligned and assembled into larger fragments. I first tested all primer sets (Table 2. 5 and Table 2. 9) using a gradient PCR to determine the optimum annealing temperature. Gradient PCR was carried out in 0.2 ml PCR tubes or 96-well plates using the following protocol:

For one 20µl reaction
11.5µl of dH_20
2µl of 10X PCR Buffer (Invitrogen)
1.3µl of dNTPs (2mM)
2µl of Forward primer (2µM)
2µl of Reverse primer (2µM)
0.2 µl of Amplitaq Gold (5u/µl) (Invitrogen)
1µl of DNA at a concentration of 50ng/µl

All primers were designed to have a predicted annealing temperature of between 55°C and 65°C, therefore each gradient PCR reaction was carried out on an Eppendorf thermocycler using the following programme:

Heated lid 110°C
Hot start at 94°C for 10 minutes
followed by 40 cycles of:
    Denaturation at 94°C for 30 seconds
    Gradient at 55-65°C for 30 seconds

Elongation at 72°C for 30 seconds/500bp

Final extension at 72°C for 7 min

Store at 10°C

Following PCR, I checked the products for specificity using a 1% agarose gel electrophoresis and by capillary sequencing. For standard PCR, the gradient temperature step was replaced by the optimum annealing temperature (listed in Table 2. 5) according to the Gradient PCR results. The optimised PCR protocol was then used to test the accessions and cultivars of interest.

Table 2. 5: Sequencing primers for *BnaFRI.A03, BnaFRI.A10, BnaFRI.C03, BnaFRI.C09* and *BnaFLC.A10*. Primers used in this study previously published by others are referenced in the table. Tm is the annealing temperature.

Target	Forward primer name	Forward primer sequence 5' to 3'	Reverse primer name	Reverse primer sequence 5' to 3'	Tm (°C)	Source
BnaFRI.A03 exon 1	TILL_BnFRIA3_F1	CCCAATGGCCGTCCGTAAT	TILL_BnFRIA3_R1	GCTCAAGGCTTATGTAGTATGCA	60	
BnaFRI.A10 exon 1	TILL_BnFRIA10_F4	GCAATTCCCATGGCCTTTCGTAATGG	TILL_BnFRIA10_R4	AGCGAAGAAGCCGAAAACCTATCCTCT	60	
BnaFRI.C03 exon 1	YWFRI_F	CGCACATCGTCCATCAACAAG	FRIJ1_R2	ATCCTTCACCCACCAGCCT	60	Irwin *et al* 2012
BnaFRI.C09 exon 1	TILL_BnFRIC9_F2	TCCCATGGCCTTCCGTAATGA	TILL_BnFRIC9_R2	CCCTTTCATCGAACATTACTCAA	60	
BnaFRI.A03	N036	CGACTTCGGTTACTTTCGAG	A385	TGTGCAGCTTTACAACTTGTC	57	Wang *et al*., 2011b
BnaFRI.A03	N053	GAAGTTCGTGTTGGACTGCATC	N020	TGCTTGTTAAGCCCTAAGC	57	Wang *et al*., 2011b
BnaFRI.A10	Bna_FRIbF1	CGCGTATCCGCATGTCGAAATTTG	A386	AATGGTACGAATCAGAAGAGG	57	Wang *et al*., 2011b
BnaFRI.A10	N030	CGAAGAAGCCCTTTGATAGAG	N032	ATTCGGAGGATGAAGAGCC	57	Wang *et al*., 2011b
BnaFRI.C03	A156	CAAGAGCAAATGGCGAAGC	N010	GTACCTTCTTGAAACCGGAG	57	Wang *et al*., 2011b
BnaFRI.C09	Bna_FRIcF1	GGATCTTCATAAGATGGCTCAG	Bna_FRIcR3	GCAATCTCAGCAGTACCACTC	57	
BnaFRI.C09	A370	AGACAGTACCGGAGATTTCG	N004	GCATTATTCCTCGTCGTTCC	57	Wang *et al*., 2011b
BnaFLC.A10	FLCA10_m79_F	CGCTCAGTATCTCCGGCTAG	FLCA10_1011_R	TAACCGCAATCGATACTAGATCG	60	
BnaFLC.A10	FLCA10_m79_F	CGCTCAGTATCTCCGGCTAG	FLCA10_1011_R	TAACCGCAATCGATACTAGATCG	60	
BnaFLC.A10	FLCA10_829_F	CCGTGAGGTTATTACTTGATGCT	FLCA10_2092_R	CACACGTTTGTCCATCGACA	60	
BnaFLC.A10	FLCA10_1965_F	TGACTGTGTTCAGCTGTCGC	FLCA10_3093_R	GTTGTGCATGAGGATCCATCT	60	
BnaFLC.A10	FLCA10_2672_F	TCGATATGGAAAGCAACATGA	FLCA10_3738_R	CTCTTCCAGCAACTTCTCCTGT	60	
BnaFLC.A10	FLCA10_3581_F	GAGGTCGCAAGCCTATCTCC	FLCA10_4665_R	ATGCTTCTTCACCGTGAAGC	60	

2.5.3. Preparation of samples for Sanger sequencing

Preparation of PCR products for sequencing involved two protocols; first I used the ExoSAP protocol to remove residual primers and dNTPs from the reaction; secondly I prepared the cleaned PCR products for sequencing using the Big Dye V3.1 terminator kit.

For one ExoSAP reaction:
10µl PCR product
1µl Rapid alkaline phosphatase (Sigma Chemical)
0.5µl Exonuclease I (NEB)

The mixture was placed in an Eppendorf thermocycler and incubated at 37°C for 30 minutes, followed by 80°C for 10 minutes.

For each clean PCR product, the Big DyeV3.1 protocol was carried out as follows:
1.5µl of 5x Big Dye Buffer (Invitrogen)
1.0µl Big Dye V3.1 (Invitrogen)
1µl of primer*
5.5µl dH_20
1µl SAP cleaned PCR product

*Big Dye V3.1 reactions were prepared for both forward and reverse primers so that each PCR product was sequenced in both direction.

Reactions were placed in an Eppendorf thermocycler with the following programme:
Heated lid 110°C
Hot start at 96°C for 1 min
followed by 25 cycles of:
    Denaturation at 96°C for 10 seconds
    Annealing at 50°C for 5 seconds
    Elongation at 60°C for 4 minutes
terminated by a final temperature step at 4°C for 10 minutes
Store at 4°C

After preparation, the samples were sent to GATC, Germany (www.gatc-biotech.com) for capillary sequencing.

2.5.4. Analysis of sequencing results

I analysed the sequence trace files generated by GATC for quality using ABI Sequence Scanner Software 2 (Applied Biosystems). Following assessment of quality, sequences were assembled and aligned to identify polymorphic sites using the Invitrogen Software Contig Express and AlignX respectively. To interpret the effect of coding mutations, I converted the predicted mRNA sequence to amino acid sequence using the online ExPasy translate tool (https://web.expasy.org/translate/).

2.5.5. Genotyping by PCR

The following Section describes the methods I used to genotype for polymorphisms or Insertion/Deletion (InDel) variation at *BnaFLC.A10* (discussed in Chapter 4), *BnaFRI.C03* and *BnaFLC.A02* (both discussed in Chapter 5). The primers used are listed in Table 2.6.

Table 2. 6*:* Genotyping primers used to screen for polymorphisms at *BnaFLC.A10*, *BnaFRI.C03* and *BnaFLC.A02*. Primers used in this study that have previously been published are referenced in the table.

Target	Forward primer name	Forward primer sequence 5' to 3'	Reverse primer name	Reverse primer sequence 5' to 3'	Reference
BnaFRI.C03 InDel	FRIC3_GENOTYPE_F1	GTCATTTGAGAGGGCGAAACGTA	FRIC3_GENOTYPE_R2	CAAGCTTCGCCATTTGCTCTTG	
BnaFLC.A02 InDel	FLCA2_GENOTYPE_F1	TCCCTAAGTGTGCTTTATGAGC	FLCA2_GENOTYPE_R1	CTAACCAATCGGATCCCAA	
BnaFLC.A02 InDel	FLCA2_GENOTYPE_F2	TACCTTGTGTCGAGAGCCTCA	FLCA2_GENOTYPE_R2	AMCTAACCAATCGGATCCCAA	
BnaFLC.A10 exon 1- intron 1	FLCA10_44_F	GTAGCCGACAAGTTACCTTCTCT	FLCA10_1011_R	TAACCGCAATCGATACTAGATCG	
BnaFLC.A10 tourist-like mite transposon	InDelF (P4f)	GGTTCCTTTTCTTTTCGTTTGGG	InDelR (P4r)	GAAGTAAAGTCGGACAAGAAGG	Hou *et al*., 2011

2.5.5.1. CAPS marker design and analysis of *BnaFLC.A10*

I designed a CAPS marker to target a SNP at position 484 from the translational start of *BnaFLC.A10* (Chapter 4). PCR was performed on DNA from 95 *B. napus* cultivars (Table 2. 4) using primers FLCA10_44F and FLCA10_1011R (Table 2. 6). Approximately 100ng of PCR product was incubated with 1 unit of NspI in 1X CutSmart Buffer (NEB) at 37°C for 1 hour. NspI (NEB), which targets the specific sequence R-CATG-Y, introduces a cut site at SNP position 484C but not in DNA with a Thymine (T) base at that position. The products of each digest were visualised for the presence of restriction sites by agarose gel electrophoresis.

2.5.5.2. A PCR screen for a tourist-like mite transposon at *BnaFLC.A10*

Using primers InDelF (P4f) and InDelR (P4r) (Table 2. 6, Hou *et al*., 2012) I screened the DNA from 95 *B. napus* cultivars (Table 2. 4) for the presence of a tourist-like mite transposon upstream of *BnaFLC.A10* (Chapter 4). Due to the A/T rich nature of the transposon, the PCR was performed using GoTaq polymerase using the following protocol:

For one 20µl reaction
8.7µl of dH_20
4µl of 5X PCR Buffer (Promega)
1.2µl $MgCl_2$
1µl of dNTPs (2mM)
2µl of F primer (2µM)
2µl of R primer (2µM)
0.1 µl of GoTaq (5u/µl) (Promega)
1µl of DNA at a concentration of 50ng/µl

PCR reaction was carried out on an Eppendorf thermocycler using the following programme:
Heated lid 110°C
Hot start at 94°C for 5 minutes

followed by 40 cycles of:

Denaturation at 94°C for 30 seconds

Anneal at 52°C for 30 seconds

Elongation at 65°C for 3 minutes

Final extension at 72°C for 7 min

Store at 10°C

PCR products were visualised by agarose gel electrophoresis and cultivars were scored for the presence or absence of the transposon.

2.5.5.3. A PCR screen for a deletion at *BnaFRI.C03*

I designed primers to detect the presence of insertion/deletion (InDel) variation at *BnaFRI.C03*; FRIC3_GENOTYPE_F1 and FRIC3_GENOTYPE_R1 (Table 2. 6 Chapter 5). The PCR products were visualised by high percentage (3%) agarose gel electrophoresis and zygosity determined by presence of InDel variation. Heterozygous plants would carry both alleles of *BnaFRI.C03* and therefore produce two bands of different sizes on an agarose gel.

2.5.5.4. A PCR screen for a deletion at *BnaFLC.A02*

I designed primers to detect the presence of InDel variation in intron 5 of *BnaFLC.A02*; FLCA2_GENOTYPE_F1 and FLCA2_GENOTYPE_R1, FLCA2_GENOTYPE_F2 and FLCA2_GENOTYPE_R2 (Table 2. 6). These primers were designed to amplify a region from intron 5 to intron 6 of *BnaFLC.A02*. PCR products were visualised by gel electrophoresis and cultivars were scored for the presence or absence of the deletion.

2.6. Gene expression analysis by RT-PCR

2.6.1. RNA extraction

For RNA extraction I collected leaf material into 1.7ml RNase/DNase free Eppendorf tubes containing one 3mm Tungsten carbide bead (Qiagen) and flash frozen on liquid nitrogen.

For *A. thaliana*, leaf material was collected to assess gene expression before and after vernalisation conditions. Whole seedlings were pooled 12 days after sowing for no vernalisation conditions, and at T0 (the last day of vernalisation treatment, sampled in the cold), T10 (10 days after vernalisation treatment, sampled in a controlled environment room) and T30 (30 days after vernalisation treatment, sampled in a controlled environment room). Leaf material was homogenized using the Geno/Grinder® by shaking at 13,000rpm for 30 seconds. For large amounts of leaf material, homogenization was carried out using a pestle and mortar cooled with liquid nitrogen.

For *B. napus*, leaf material was taken from the newest expanded true leaf before and after vernalistion treatment; at 21-28 days after sowing for no vernalisation conditions; at T0 (the last day of vernalisation treatment, sampled in the cold); at T16 (16 days after vernalisation treatment, sampled in the glasshouse); and at T32 (32 days after vernalisation treatment, sampled in the glasshouse). The leaf material was later homogenized using the Geno/Grinder® by shaking at 13,000rpm for 30 seconds.

I extracted total RNA from individual and pooled leaf samples using the E.Z.N.A. Plant RNA Kit (Omega BioTek) and contaminating DNA was removed using the on-column RNase-free DNase Set I (Omega BioTek), or using the TURBO DNase kit, according to the manufacturer's instructions. Quality and quantity of the RNA was determined by measuring the absorbance at 260nm using the Nanodrop® 1000 spectrophotometer (ThermoScientific).

2.6.2. cDNA Synthesis

1-2µg RNA was converted to cDNA using Superscript III (Invitrogen) and oligo dT (Promega), or gene-specific reverse primers, according to the manufacturer's instructions.

2.6.3. Primers

For RT-PCR, I designed primers that would amplify a product <120bp in size, that were selective for the gene copy of interest, and that had an optimum annealing temperature (Tm) of 60°C. Although not always possible, I ensured at least one primer was located over an exon-exon junction to avoid amplification of DNA. Primers used for RT-PCR are listed in Table 2. 7.

.

2.6.4. qPCR Primer efficiency testing

I tested all primers used for quantitative expression analysis for specificity and efficiency first by amplifying from cDNA in a qualitative RT-PCR. 1µl of cDNA was amplified using Amplitaq Gold PCR protocol as described in Section 2.5.2. Each reaction was carried out in an Eppendorf thermocycler using the following programme:

Heated lid 110°C
Hot start at 94°C for 10 minutes
followed by 40 cycles of:
    Denaturation at 94°C for 30 seconds
    Annealing at 60°C for 30 seconds
    Elongation at 72°C for 30 seconds
Final extension at 72°C for 7 min
Store at 10°C

After PCR amplification, products were visualized by 2% agarose gel electrophoresis and submitted for sequencing by GATC, Germany (www.gatc.com) to check for sequence specificity.

The primer efficiencies and melt curves were determined using qPCR on a four-fold dilution series, with a total of eight dilutions, of cDNA. Each 10µl reaction was prepared as follows:
5µl of SYBR Green (Roche Diagnostics Ltd)
0.125µl of Forward primer (10µM)
0.125µl of Reverse primer (10µM)
0.75µl of dH_20
4µl of cDNA

Quantitative expression was analysed on a Roche LightCycler® 480 using the following conditions:

Pre-incubation	95°C for 5 minutes
50 cycles of Amplification	95°C for 15 seconds, 60°C for 20 seconds, 72°C for 30 seconds
Followed by a Melting curve step at 97°C	

Efficiencies were calculated using a standard curve whereby the CT values were plotted against the log10 of the cDNA dilution (Ginzinger 2002). I used only primers with efficiencies of 95-100% for quantitative RT-PCR in this Thesis.

2.6.5. Expression Analysis by Quantitative RT-PCR

For quantitative gene expression analysis, I diluted cDNA ten-fold in RNase free water and pipetted 4µl by hand into 384-well white LightCycler plates (Roche Diagnostics Ltd). Then 6µl of SYBR Green Master Mix (SYBR Green (Roche Diagnostics Ltd), Forward primer, Reverse primer, RNase free water as in Section 2.6.4.) was added to each cDNA sample by hand using an electronic pipette. The reaction was then placed in a Roche LightCycler® 480 with the conditions as described in Section 2.6.4. For each biological replicate, three technical replicates were performed. Positive samples were included on each plate to compare between multiple LightCycler® runs. Negative samples, water and no reverse transcriptase (-RT) controls, were also included to test for DNA contamination.

Second derivative maximum values were calculated using the LightCycler® Software to give absolute expression values. Expression values of the target genes were then normalized using the 2-ΔΔCT method (Livak and Schmittgen 2001). For *A. thaliana*, expression values were normalized to the internal reference gene *UBC*, for *B. napus* normalized to *UBC21*. In all cases mean expression values were calculated for at least two biological replicates.

Table 2. 7: Primers used for quantitative RT-PCR. Previously published primers are referenced in the table.

Target Gene	Forward primer name	Forward primer sequence 5' to 3'	Reverse primer name	Reverse primer sequence 5' to 3'	Source
AtUBC	UBC_qPCR_F	CTGCGACTCAGGGAATCTTCTAA	UBC_qPCR_R	TTGTGCCATTGAATTGAACCC	Dean Lab
AtFLC	FLC spliced F	AGCCAAGAAGACCGAACTCA	FLC spliced R	TTTGTCCAGCAGGTGACATC	
AtFT	FT_cDNA_372F	CTGGAACAACCTTTGGCAAT	FT_cDNA_590R	AGCCACTCTCCCTCTGACAA	
AtBRC1	At_RT-PCR-BRC1-AF	TTCCCAGTGATTAACCACCAT	At_RT-PCR-BRC1-BR	TCCGTAAACTGATGCTGCTC	Aguilar-Martínez *et al*., 2007
BnaUBC	BnUBC21_qF2	GTCCTCTCAACTGCGACTCA	BnUBC21_qR2	GTGTGTACATGCGTGCCATT	Orsel *et al*., 2014
BnaFLC.A10	q15	GTCTCCGCCTCCGGGAA	q16	AGCTTTTGACTGACGATCCAA	Hawkes 2017
BnaFLC.C09A	q29	GTCTCCGCCTCCAGAAAACT	q30	TTGACTGACGATCCAAGGCT	
BnaFLC.A02	q41	CCTCCGGCAAGCTTTACAAC	q42	GAGCTTTTGACTGAAGATCCAGA	
BnaFLC.C02	q47	GGCTAGCCAGATGGAGAAGAATA	q48	GTGGGAGAGTTACCGGACAA	
BnaFLC.A03A	q62	GCTTGTGGAATCAAATGTCGGT	q63	GCTTCAACATTAGTTCTGTCTTCCT	
BnaFLC.C03A	q70	GACGCAGCGGTCTCGT	q71	CAAGGATCCTGACCAGGTTATC	
BnaFLC.A03B	q80	TCTCTGCGAGGCATCTGTTG	q81	GATCTTCTCCAGTCTATCCCCG	
BnaFLC.C03B	q88	AGGAAAACAGTAGCAGACAAGTTA	q89	CCAGTCTATCACCGGAGGAGAA	
BnaFT.A02	BnC2FT_BnA2FT_Guo_Fw	GTTGTAGGAGACGTTCTTGAATGT	BnA2FT_Guo_Rev	TCTGGATCCACCATAACCAAAGTA	Guo *et al*., 2014
BnaFRI.A03	BNFRIA3_SET5_F	GCGTGGAATGCATATTGAAGCT	BNFRIA3_SET6_R	AGCTTGGCCATTTGCTCTTT	
BnaFRI.C03	BNFRIC3_SET5_F	GCGTGGAATGCATATTGAAGCA	BNFRIC3_SET6_R	AGCTTCGCCATTTGCTCTTG	
BnaFRI.A10	BNFRIA10C9_SET7_F	GGATTTGTTTGATCTGGTAC	BNFRIA10_SET7_R	CAAGGAAAGGTGACCGTTTA	
BnaFRI.C09	BNFRIA10C9_SET7_F	GGATTTGTTTGATCTGGTAC	BNFRIA10_SET7_R	CAAGGAAAGGTGACCGTTTT	

2.7. A Mutation Analysis of *BnaFRI*

The following section describes methods that are specific to Chapter 3. Winter oilseed rape lines carrying EMS mutations at *BnaFRI* had previously been generated by Bayer CropScience NV, Belgium (Table 2. 8). I assessed all lines listed in Table 2. 8 for flowering time with and without vernalisation. Plants were sown on soil and grown under glasshouse conditions (18°C daytime temperature/15°C night-time temperature, 16-hour light/8-hour dark) for 24 days before being transferred to vernalisation conditions (5°C, 8-hour light, 16-hour dark) for 6, 4, or 2 weeks. Sowing was staggered so that all plants were transferred back to glasshouse conditions on the same day. Plants that received no vernalisation remained in the glasshouse for the duration of the experiment. Leaf material was taken for RNA extraction (Section 2.6.1) and I analysed the quantitative expression of *BnaFLC* by quantitative RT-PCR (see Section 2.6. for full method details).

Table 2. 8: *B. napus* lines generated by Bayer CropScience NV, Belgium assessed for flowering time and *BnaFLC* expression. Included in the table are line names and whether *BnaFRI.A10, BnaFRI.A03, BnaFRI.C09* or *BnaFRI.C03* were mutated. Westar-10, a spring cultivar, was also included for comparison.

Species	**Line name**	***BnaFRI* Mutant allele**	**Source**
Brassica napus	134 (15MBBN000134)	*BnaFRI.A10* (FRI-A1-EMS02) *BnaFRI.A03* (FRI-A2-EMS02) *BnaFRI.C09* (FRI-C1-EMS01) *BnaFRI.C03* (FRI-C2-EMS09)	Bayer CropScience NV, Belgium
	135 (15MBBN000135)	*BnaFRI.A10* (FRI-A1-EMS02) *BnaFRI.A03* (FRI-A2-EMS02) *BnaFRI.C09* (FRI-C1-EMS01) *BnaFRI.C03* (WT)	
	136 (15MBBN000136)	*BnaFRI.A10* (FRI-A1-EMS02) *BnaFRI.A03* (FRI-A2-EMS02) *BnaFRI.C09* (WT) *BnaFRI.C03* (FRI-C2-EMS09)	
	137 (15MBBN000137)	*BnaFRI.A10* (FRI-A1-EMS02) *BnaFRI.A03* (WT) *BnaFRI.C09* (FRI-C1-EMS01) *BnaFRI.C03* (FRI-C2-EMS09)	Bayer CropScience NV, Belgium
	138 (15MBBN000138)	*BnaFRI.A10* (WT) *BnaFRI.A03* (FRI-A2-EMS02) *BnaFRI.C09* (FRI-C1-EMS01) *BnaFRI.C03* (FRI-C2-EMS09)	

Species	**Line name**	***BnaFRI* Mutant allele**	**Source**
Brassica napus	139 (15MBBN000139)	*BnaFRI.A10* (FRI-A1-EMS02) *BnaFRI.A03* (FRI-A2-EMS02) *BnaFRI.C09* (WT) *BnaFRI.C03* (WT)	
	140 (15MBBN000140)	*BnaFRI.A10* (FRI-A1-EMS02) *BnaFRI.A03* (WT) *BnaFRI.C09* (FRI-C1-EMS01) *BnaFRI.C03* (WT)	
	141 (15MBBN000141)	*BnaFRI.A10* (FRI-A1-EMS02) *BnaFRI.A03* (WT) *BnaFRI.C09* (WT) *BnaFRI.C03* (FRI-C2-EMS09)	
	142 (15MBBN000142)	*BnaFRI.A10* (WT) *BnaFRI.A03* (FRI-A2-EMS02) *BnaFRI.C09* (FRI-C1-EMS01) *BnaFRI.C03* (WT)	
	143 (15MBBN000143)	*BnaFRI.A10* (WT) *BnaFRI.A03* (FRI-A2-EMS02) *BnaFRI.C09* (WT) *BnaFRI.C03* (FRI-C2-EMS09)	
	144 (15MBBN000144)	*BnaFRI.A10* (WT) *BnaFRI.A03* (WT) *BnaFRI.C09* (FRI-C1-EMS01) *BnaFRI.C03* (FRI-C2-EMS09)	
	145 (15MBBN000145)	*BnaFRI.A10* (FRI-A1-EMS02) *BnaFRI.A03* (WT) *BnaFRI.C09* (WT) *BnaFRI.C03* (WT)	
	146 (15MBBN000146)	*BnaFRI.A10* (WT) *BnaFRI.A03* (FRI-A2-EMS02) *BnaFRI.C09* (WT) *BnaFRI.C03* (WT)	
	147 (15MBBN000147)	*BnaFRI.A10* (WT) *BnaFRI.A03* (WT) *BnaFRI.C09* (FRI-C1-EMS01) *BnaFRI.C03* (WT)	
	148 (15MBBN000148)	*BnaFRI.A10* (WT) *BnaFRI.A03* (WT) *BnaFRI.C09* (WT) *BnaFRI.C03* (FRI-C2-EMS09)	
	Wild-type line 149 (15MBBN000149)	*BnaFRI.A10* (WT) *BnaFRI.A03* (WT) *BnaFRI.C09* (WT) *BnaFRI.C03* (WT)	
	Winter oilseed rape parental line 150 (15MBBN000150)	*BnaFRI.A10* (WT) *BnaFRI.A03* (WT) *BnaFRI.C09* (WT) *BnaFRI.C03* (WT)	
	Westar-10	-	OREGIN

2.8. A Transgenic Analysis of *BnaFRI.A03*

To investigate the functional consequence of allelic variation at *BnaFRI.A03*, a gene proposed to underlie a QTL for flowering time in *B. napus* (Wang *et al*., 2011b), I performed a transgenic analysis of *BnaFRI.A03* in *A. thaliana*. The following section describes methods used in this analysis and are specific to Chapter 3.

2.8.1. Sequencing, synthesis and transformation of *BnaFRI.A03*

To generate an allelic series of constructs for *BnaFRI.A03* I amplified and sequenced the open reading frame +/- 2kb of *BnaFRI.A03* from eight *B. napus* cultivars (Stellar, Westar-10, Ningyou-7, Shengliyoucai-11, Express-617, Major, Sensation New Zealand, and Wilhelmsburger-Reform) using primers listed in Table 2. 9. DNA was extracted from each cultivar using the Edwards DNA extraction method (Edwards *et al*. 1991; Supplementary Methods) and after PCR optimisation, PCR was performed on each cultivar (as described in Section 2.5.2). PCR and capillary sequencing were performed according to methods described in Sections 2.5.3 and 2.5.4. To reduce sequencing error, multiple PCR products were submitted for sequencing per cultivar.

To determine the functional consequence of variation at coding and non-coding regions of *BnaFRI.A03* I designed each construct as four level 0 modules (promoter, ORF1, ORF2, terminator) that could be assembled into a complete *BnaFRI.A03* construct by Golden Gate assembly (Engler *et al*. 2009). To ensure specific and seamless assembly of all four modules, *BsaI* digestion and fusion sites were incorporated according to Weber *et al*. (2011) and as illustrated in Figure 2. 1. I designed the modules so that all constructs generated would include the open read frame plus 1.2kb of the 5'UTR (Promoter) and 1.5kb of the 3'UTR (Terminator). The open reading frame was split in two; ORF1 included the translational start codon to 1099bp downstream, ORF2 included from 1100bp downstream from the translational start codon to the translational stop codon. All modules for *BnaFRI.A03* were synthesised in the pL0M vector by the ENSA DNA synthesis service (www.ensa.ac.uk).

Table 2. 9: Primers used to amplify and sequence *BnaFRI.A03*. Primers previously published by others are referenced in the table. Tm denotes the annealing temperature.

Target	Forward primer name	Forward primer sequence 5' to 3'	Reverse primer name	Reverse primer sequence 5' to 3'	Tm (°C)	Source
BnaFRI.A03 5'UTR	A3F_promoter_1	GTGTGGATCTTACGGCACGA	A3R_promoter_1	CTGGTTCGGCCTGCTTGATA	60	
BnaFRI.A03 5' UTR	A3F_promoter_2	CTATCAAGCAGGCCGAACCA	A3R_promoter_2	GGCTTCGCTGAGTGGATCTT	60	
BnaFRI.A03 5' UTR	A3F_promoter_3	TATCAAGCAGGCCGAACCAG	A3R_promoter_3	CGAGAACCACAAAAGCCGTG	60	
BnaFRI.A03 5'UTR	A3F_promoter_4	AAGATCCACTCAGCGAAGCC	A3R_promoter_4	TCGTTCCTCCTCTCGTGACT	60	
BnaFRI.A03 5'UTR	A3F_promoter_5	CACGGCTTTTGTGGTTCTCG	A3R_promoter_5	AACCATTACGGACGGCCATT	60	
BnaFRI.A03 5'UTR	A3F_promoter_6	ACAGGAAACCCTTCGGTTCG	A3R_promoter_6	ACGGTTTCGACAGTAGCCTG	60	
BnaFRI.A03 5' UTR sequencing primer	A3F_promoter_7_seq	AGTTGCAATTTCTCAGCCCTT			n/a	
BnaFRI.A03 5' UTR sequencing primer			A3R_promoter_7_seq	AAGGGCTGAGAAATTGCAACT	n/a	
BnaFRI.A03 5' UTR sequencing *primer*	A3F_promoter_8_seq	TCTGCGATTTTGCGGTTAATC			n/a	
BnaFRI.A03 5' UTR			A3R_promoter_8_seq	GATTAACCGCAAAATCGCAGA	n/a	
BnaFRI.A03 5'UTR sequencing primer for Spring cultivars	SPRING_SEQ_PROM_F1	CCAAGCAAAAATGTGACCGGT			n/a	
BnaFRI.A03 5'UTR Sequencing primer for Spring cultivars	SPRING_SEQ_PROM_F2	TCCCRATCCCCTTTTCCGAT			n/a	
BnaFRI.A03 open reading frame	N036	CGACTTCGGTTACTTTCGAG	A385	TGTGCAGCTTTACAACTTGTC	57	Wang *et al*., 2011bb
BnaFRI.A03 open reading frame	N053	GAAGTTCGTGTTGGACTGCATC	N020	TGCTTGTTAAGCCCTAAGC	57	Wang *et al*., 2011bb
BnaFRI.A03 3'UTR	A3F_3UTR_1	CTCGGCCCATCATCAGGTTT	A3R_3UTR_1	CAGCCTCCAACTTCGGTTCT	60	
BnaFRI.A03 3' UTR	A3F_3UTR_2	TTCGACTCTTCGGTTCTGCG	A3R_3UTR_2	ACCCACCCAGCTTTTCGATT	60	
BnaFRI.A03 3'UTR	A3F_3UTR_1	CTCGGCCCATCATCAGGTTT	A3R_3UTR_3	CACTGATGCAGACAAGGCGAA	60	

For Golden Gate assembly I used a protocol adapted from Engler *et al*. 2009 (see Supplementary Methods). The level 0 modules for the Promoter, ORF1, ORF2 and Terminator regions of *BnaFRI.A03* were combined according to Figure 2. 1 C. After assembly by Golden Gate (Engler *et al*., 2009) I transformed the *BnaFRI.A03* constructs into sub-cloning efficiency DH5α competent cells (Invitrogen) according to the manufacturer's instructions and incubated at 37°C under selection (10µg/ml kanamycin, 0.08µl/ml IPTG, and 1.28µl/ml X-gal) overnight on LB plates for blue-white screening of colonies. After incubation, I selected single white colonies to be multiplied overnight under selection in liquid LB before purification using the QIAprep Spin Miniprep Kit (Qiagen) according to the manufacturer's instructions. *BnaFRI.A03* constructs were then ligated with pSLJ-DEST using the LR clonase enzyme (Invitrogen) see Supplementary Methods for details. The pSLJ-DEST binary vector is a version of binary vector SLJ-75515 (a gift from Professor Jonathan Jones to the Caroline Dean Laboratory (Jones *et al*. 1992)) modified to carry Gateway® recombination sites (originally generated by C. Lister, Caroline Dean Laboratory).

I transformed the pSLJ-DEST plasmids containing the *BnaFRI.A03* constructs into *Agrobacterium tumefaciens* strain C58 using the triparental mating method with helper plasmid HB101. I then transformed the nine constructs into *A. thaliana* accession Col-0 by an *Agrobacterium*-mediated floral spray method adapted from Clough and Bent (1998). Flowers of forty *A. thaliana Col-0* plants were sprayed twice at weekly intervals per construct. Plants were bagged and T_1 seeds harvested. Details of the triparental mating and *Agrobacterium*-mediated floral spray methods are described in Supplementary Methods.

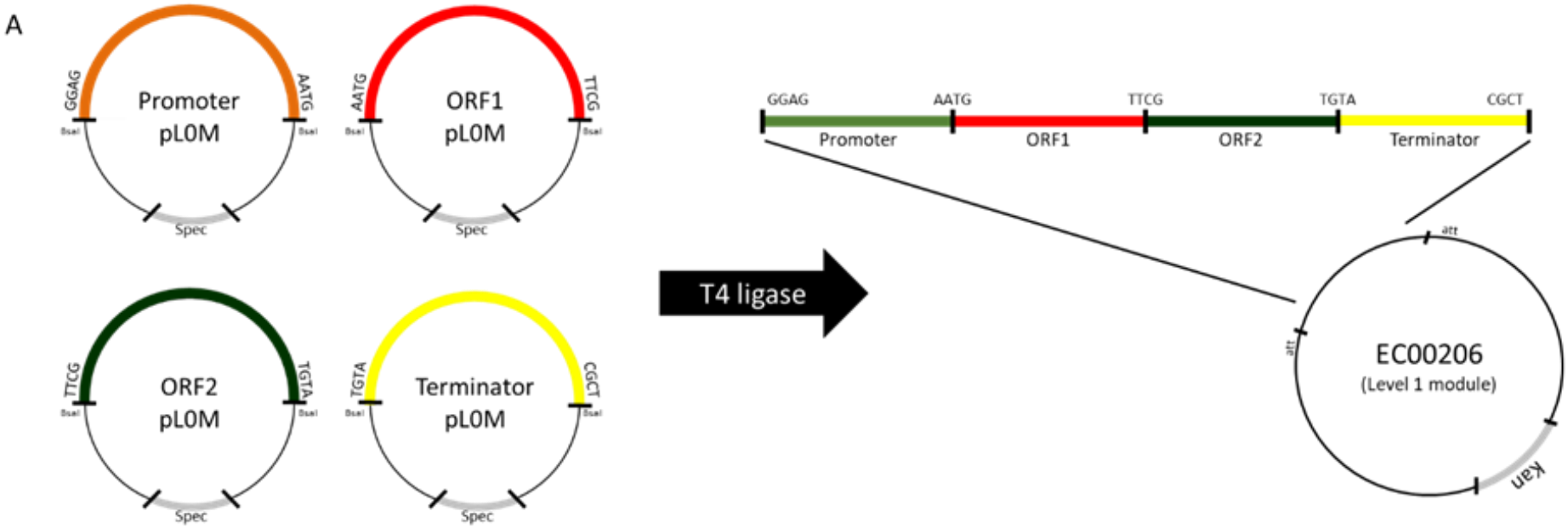
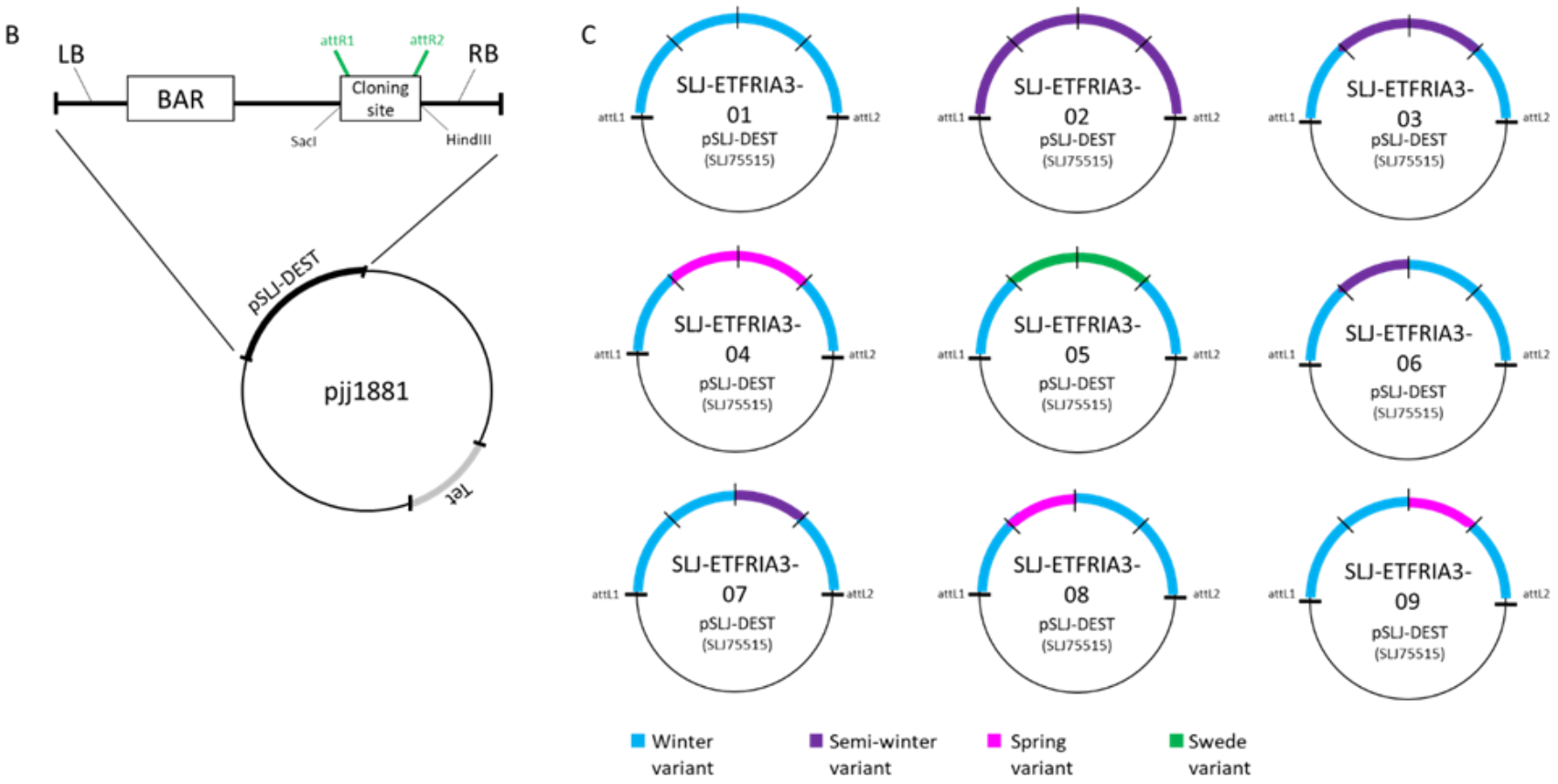


Figure 2. 1: *BnaFRI.A03* constructs designed, assembled and transformed into *A. thaliana* accession Col-0. (A) Four Level 0 modules were designed for DNA synthesis. The Promoter, ORF1, ORF2 and Terminator sequences were synthesized with specific *BsaI* digestion and fusion sites, the four base-pair overhangs are highlighted. The modules were synthesized in pL0M vectors by ENSA, and the assembled Level 1 module is illustrated. (B) A schematic diagram of PSLJ-DEST with the Gateway ligation sites highlighted in green. (C) Nine constructs were assembled using Promoter, ORF1, ORF2 and Terminator sequences from winter oilseed rape (blue), Chinese semi-winter oilseed rape (purple), spring oilseed rape (pink), and swede (green) cultivars.

Figure 2. 1 continued



2.8.2. Selection of transgenic lines and phenotypic analysis

Hemizygous T_1 seeds were sown on soil by the NBI Horticultural Services Team and sprayed with BASTA® (BASF, contains phosphinothricin) for selection of positive transformants. After selection, I collected leaf material from all positive transformants and submitted these to iDNA Genetics (http://www.idnagenetics.com/) for DNA extraction and copy number analysis.

T_1 plants containing single copy *BnaFRI.A03* constructs were bagged and T_2 seeds harvested. To assess the flowering time of each *BnaFRI.A03* construct, T_2 seed from eight single copy lines per construct, and seed of wild-type *A. thaliana* accessions Col-0 (*fri*) and Col-*FRI*^{SF2} (Col-0 with functional *FRI* introgressed from accession San Feliu 2; Lee *et al*., 1994b) were vapour phase sterilised and sown on GM plates without glucose and with selection, if required (see Section 2.3.5 for method details). After stratification, as described in Section 2.3.5, seedlings were transferred to controlled environment conditions (20°C, 16-hour light, 8-hour dark, 70% humidity) and transplanted onto soil on the 12th day after stratification. I recorded the flowering time of each plant when buds were visible in the apex of the rosette from the date of transplantation to soil. To assess the quantitative expression of *BnaFRI.A03* and the native *FLC*, on the 12th day after stratification whole *A. thaliana* seedlings were pooled to generate three biological replicates. For the T_2 transgenic lines, five seedlings from each single copy line were combined to make a pool of 40 seedlings per biological replicate. Method details for RNA extraction and quantitative RT-PCR are described in Section 2.6.

2.8.3. PCR amplification and sequencing of a mutator-like transposon at *BnaFRI.A03*

I designed primers A3F_promoter_4 and A3R_promoter_6 (Table 2. 9) to amplify a transposon from spring oilseed rape lines Stellar and Westar-10. The transposon was detected in the 5'UTR region of *BnaFRI.A03* and the PCR was performed using GoTaq polymerase using the following protocol:

For one 20µl reaction

8.7µl of dH_20

4µl of 5X PCR Buffer (Promega)

1.2µl $MgCl_2$

1.0µl of dNTPs (10mM)

2µl of Forward primer (2µM)

2µl of Reverse primer (2µM)

0.1 µl of GoTaq (5u/µl) (Promega)

1µl of DNA at a concentration of 50ng/µl

PCR reaction was carried out on an Eppendorf thermocycler using the following programme:

Heated lid 110°C

Hot start at 94°C for 5 minutes

followed by 40 cycles of:

- Denaturation at 94°C for 30 seconds
- Anneal at 60°C for 30 seconds
- Elongation at 65°C for 4 minutes
- Final extension at 72°C for 7 min

Store at 10°C

I visualised the PCR products by gel electrophoresis and purification was carried out using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research). After purification PCR products were cleaned using the ExoSap protocol and sent to GATC, Germany for capillary sequencing (as described in Sections 2.5.3 and 2.5.4).

2.9. An Associative Transcriptomics Analysis of Flowering Time in *B. napus*

The following section describes methods that are specific to Chapter 4. 88 cultivars from the OREGIN DFFS population of *B. napus* were assessed for flowering time with and without a vernalisation treatment in a trial in 2014. This trail was carried out prior to the start of my PhD project and the flowering time dataset was made

available to me by Dr J. Irwin. To identify genetic variation important for determining vernalisation requirement in the OREGIN DFFS population of *B. napus* I analysed the flowering time dataset by Associative Transcriptomics using methods previously published by Harper *et al*., 2012.

2.9.1. Plant Material and Growth Conditions

A diversity set comprising 88 *B. napus* cultivars from the OREGIN DFFS population (Table 2. 4), were grown for flowering time analysis during the spring and summer of 2014. Within the panel of lines were 5 Chinese semi-winter oilseed rape, 14 spring oilseed rape, 52 winter oilseed rape, 6 swede, 4 winter fodder, 4 kale, 1 winter vegetable, 1 synthetic and 1 exotic cultivars.

Seed of all cultivars were sown on soil and grown under glasshouse conditions (16-hour light/8-hour dark, 18°C daytime temperature/15°C night-time temperature, 70% humidity). Plants received one of two treatments; vernalisation (VERN) or no vernalisation (NVERN). For the VERN treatment, plants were transferred to a vernalisation chamber (5°C, 8-hour light/16-hour dark, 70% humidity) at 21 days after sowing for 6 weeks. After the vernalisation treatment, plants were transferred to a glasshouse for one week for acclimation, before being transplanted to a soil poly-tunnel (Keder; protected environment, no supplementary heating or lighting) at JIC. Plants receiving the NVERN treatment were grown in the glasshouse for 28 days (21 days pre-growth treatment plus 7 days) before being transplanted to the poly-tunnel. Sowing dates were staggered to ensure all plants from both treatments were transplanted on the same day. Plants were arranged in a complete, randomized block design with one plant from each line included in each of three blocks. Tiny-tag® data loggers recorded the temperature at one-hour intervals throughout the duration of the experiment. Flowering time was measured twice per plant according to Meier *et al*., 2001; on the day when buds were visible from above (BBCH51) and on the day when the first flower had opened (BBCH60). Scoring of flowering commenced after the end of the vernalisation treatment; 63 and 21 days after sowing for the VERN and NVERN treatments respectively.

2.9.2. mRNA sequencing dataset

The mRNA-sequencing (mRNA-seq) dataset used to carry out the Associative Transcriptomics analysis was previously published by Lu *et al*., (2014b) Briefly, the first true leaf from four plants per cultivar were pooled at 21 days after sowing. RNA was extracted from these pooled leaf samples and sequenced by Illumina®. This mRNA-seq dataset is publicly available under accession numbers: ERA122949, ERA036824, ERA063602.

In silico expression levels for *BnaFLC* genes were generated by Drs J. Irwin, R. Wells, N. Pullen, and M. Trick. To correct for mis-assignment of sequencing reads the mRNA-seq dataset (Lu *et al*., 2014b) were aligned to homeologue specific sequences located in exon 7 of *BnaFLC* (Figure 2. 2). The aligned sequence reads for each *BnaFLC* gene were extracted and counted to produce quantitative *in silico* expression levels for each *BnaFLC* from all 88 *B. napus* cultivars within the diversity population (Figure S. 2).

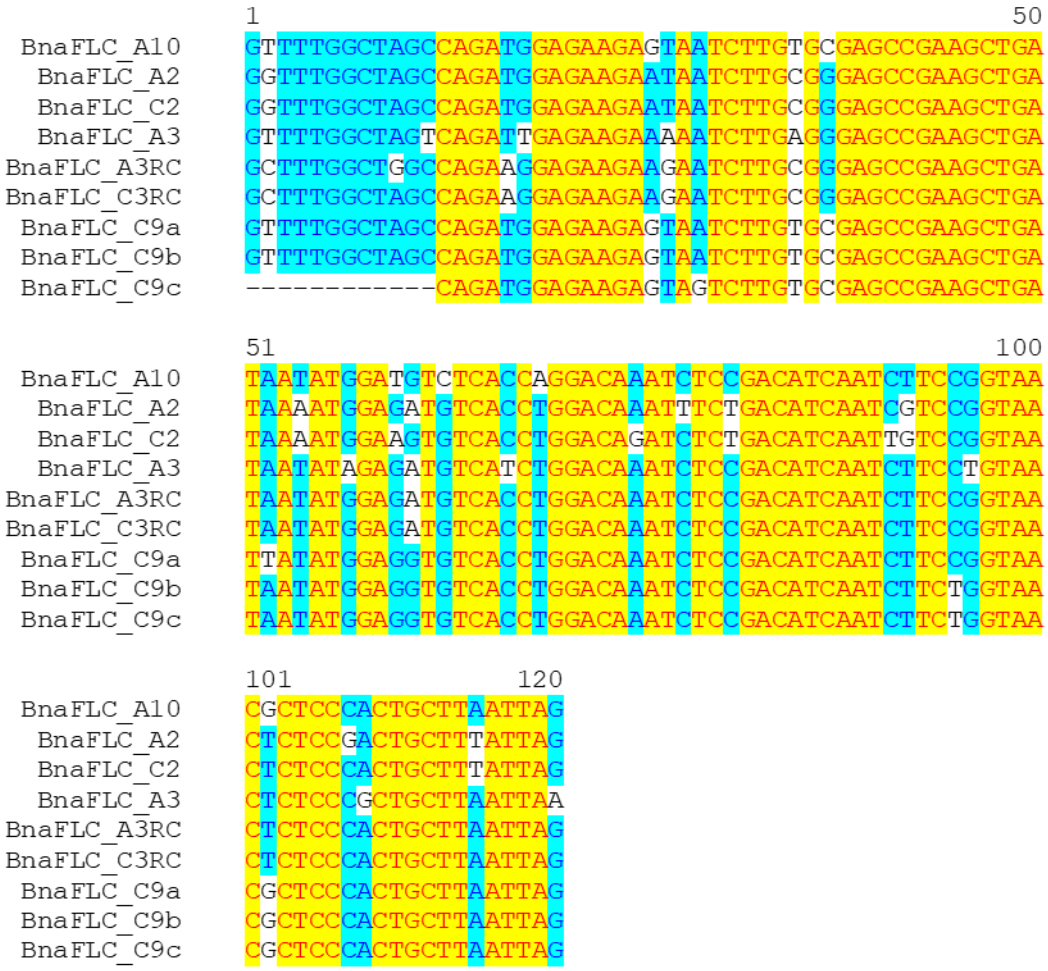


Figure 2. 2: Homeologue specific sequences identified in exon 7 of *BnaFLC* that distinguish all nine *BnaFLC* genes. The mRNA-seq dataset was aligned and the number of aligned reads for each *BnaFLC* gene was counted to give *in silico* expression levels.

2.9.3. Linear Regression for Gene expression marker analysis

To identify variation in gene expression that associated with flowering time, I used a published Associative Transcriptomics method called Gene Expression Marker (GEM) analysis (Harper *et al*., 2012) to identify correlations between sequence read depth and the flowering time measurements collected for the OREGIN DFFS population (experimental details in Section 2.9.1). For this method, the sequence read depth of each unigene (determined from expressed sequence tag (EST) assemblies by Trick *et al*., 2009) in the mRNA-seq dataset published by Lu *et al*., (2014b) is used as a proxy for gene expression measurements. After alignment of mRNA-seq reads to the *B. napus* pseudomolecule reference sequence (Harper *et al*., 2012, Lu *et al*., 2014b) the transcript level of each unigene was quantified as reads

per kb per million aligned reads (RPKM) and this dataset was made publicly available by Lu *et al*., (2014b).

Following filtering of transcripts with a RPKM value of less than 0.4, I performed a linear regression using published R scripts developed by Harper *et al*., 2012 (R version 3.3.3. "Another Canoe"). To identify correlations between mean flowering time and gene expression, R^2 and significance values were calculated between the flowering time and RPKM values for each unigene to generate a P-value for each flowering time-unigene association. Each P-value was converted into $-\log_{10}P$ and plotted against the physical position in the *B. napus* pseudomolecules to generate a Manhattan plot. Bonferroni significance threshold was calculated according to Lu *et al*., 2014b by P = 1/100,534 GEMs; $-\log_{10}P = 5.0$. I used this method to identify associations between flowering time measurements of the OREGIN DFFS population (data collected in Section 2.9.1) and gene expression, under both the NVERN and VERN treatments

2.9.4. Assessment of population structure

Assessment of population structure of the 88 *B. napus* cultivars was performed by Dr R. Wells. Sequencing reads from the mRNA-seq dataset (Lu *et al*., 2014b) were aligned to the unigene pseudomolecule reference sequence and, using methods described by Trick *et al*., 2009, 225,001 SNP markers were identified. For assessment of population structure, SNP markers were selected at 500kb intervals along all chromosomes. SNP markers present within 1000kb of centromeric regions, as defined by Mason *et al*, 2015 and SNPs showing allele frequencies of less than 5% were excluded. After SNP selection, 680 unlinked SNPs were used in an analysis by STRUCTURE (Evanno *et al*., 2005) as quoted by Harper *et al*., (2012) and Lu *et al*., (2014b). The optimal number of K populations was determined and used to assign a Q-matrix score to each *B. napus* cultivar; a proportion of the SNPs in each cultivar that can be assigned to each K population. For Associative Transcriptomics, K-1 Q-matrix scores for each cultivar were used as fixed effects in the linear model performed.

2.9.5. Associative Transcriptomics analysis for cis- and trans-factors that correlate with gene expression

In order to identify genetic variation important for determining the expression level of *BnaFLC.A10* I used the Associative Transcriptomics (Harper *et al*., 2012) method to analyse the *in silico* expression of *BnaFLC.A10* (Section 2.9.2). SNP data for each *B. napus* cultivar (published by Lu *et al*., 2014b) and the *in silico* expression of *BnaFLC.A10* were analysed using the programme TASSEL (V4.0) (Bradbury *et al*., 2007) according to methods described by Harper *et al*., (2012). To identify the model of best fit, I performed both General Linear Model (GLM) and Mixed Linear Model (MLM). Quantile-quantile (QQ) plots were generated for each linear model, and after assessment of the observed versus expected marker-trait associations I selected the model of best fit for further analysis.

For GLM, the SNP data, K-1 Q-matrix and *BnaFLC.A10 in silico* expression data were loaded into TASSEL. I removed minor allele states below 5% from the SNP dataset as they are assumed to arise from sequencing error. I subsequently performed a GLM which generated an output of P-values that describe the correlation between SNP and *BnaFLC.A10* expression variation. These P-values were plotted against the expected values of no association in a QQ plot for model assessment.

MLM has previously been shown to minimise false-positive associations by inclusion of a kinship matrix as a random effect in the analysis. Again, the SNP data, K-1 Q-matrix and *BnaFLC.A10 in silico* expression data were loaded into TASSEL. I again removed minor allele states below 5% from the SNP dataset and used this filtered SNP dataset to estimate the pairwise relatedness between cultivars using the TASSEL software. I subsequently performed a MLM using the filtered SNP dataset, the K-1 Q-matrix, the kinship matrix and *BnaFLC.A10 in silico* expression data. The P-values generated by the model were plotted against the expected values in a QQ plot for model assessment.

After selection of the model of best fit, the $-\log_{10}$ P-values of each SNP-trait association were ordered along the *B. napus* pseudomolecule and displayed as a Manhattan plot using R scripts developed by Harper *et al*., 2012 (R version 3.3.3. "Another Canoe").

2.10. QTL-seq to identify candidate genes for vernalisation requirement and response in winter oilseed rape

The following section describes methods that are specific to Chapter 5. To identify genetic regions important for vernalisation requirement in winter oilseed rape I selected two cultivars, Cabriolet and Darmor, with contrasting flowering time phenotypes for analysis by QTL-seq, a bulk segregant analysis (BSA) approach originally published by Takagi *et al*., 2013.

2.10.1. Generation of an F_2 population

I performed reciprocal crosses between three individual plants of Cabriolet and Darmor to generate heterozygous F_1 seed (Table 2. 10). Oilseed rape is hermaphroditic therefore each female flower was first emasculated using forceps that had been sterilised with 70% ethanol to prevent self-fertilisation. After emasculation, pollen from the male donor was transferred to the stigma of the female flower and bagged using microperforated cellophane bags. The fruit from each cross were left to mature before harvesting. Leaf tissue was sampled into foil from each individual parent, flash frozen in liquid nitrogen and stored at -80°C until required.

Table 2. 10: Reciprocal crosses were performed between three Cabriolet and three Darmor individuals named Cab53, Cab144, Cab203, Dar98, Dar124, Dar208.

Cabriolet♀ x Darmor♂	**Darmor♀ x Cabriolet♂**
Cab53 x Dar98	Dar98 x Cab53
Cab144 x Dar124	Dar124 x Cab144
Cab203 x Dar208	Dar208 x Cab203

Six seed from each reciprocal cross were sown on soil and after 21 days growth under glasshouse conditions (18°C day/15°C night, 16-hour light/8-hour dark, 70% humidity) leaf tissue was sampled for DNA extraction by Edwards Method (see Supplementary Methods for details). 50ng of DNA from each line was genotyped using primers FRIC3_GENOTYPE_F1 and FRIC3_GENOTYPE_R1 (Table 2. 6) that I had designed to target a 12bp InDel variation between Cabriolet and Darmor at *BnaFRI.C03* (see Section 2.5.5.3. for details).

Based on the genotyping data, ten heterozygous F_1 plants (Table 2. 11); representing five F_1 sibling plants from cross Cab203 x Dar208 and five F_1 sibling plants from cross Dar208 x Cab203 were carried forward to the F_2 generation. At 28 days growth all ten F_1 plants were transferred from the glasshouse to vernalisation (5°C, 8-hour light/16-hour dark, 70% humidity) for 12 weeks. After vernalisation, plants were returned to glasshouse conditions and transplanted to 2-litre pots. Individual plants were staked and bagged before flowering to prevent outcrossing and seeds were harvested after senescence.

Table 2. 11: Ten F_1 plants from reciprocal cross Cab203 x Dar208 carried forward to F_2 generation

Species	Line name	Source
Brassica napus	F1_CD_203_208_1	F1 seed, Cab203 x Dar208
	F1_CD_203_208_2	F1 seed, Cab203 x Dar208
	F1_CD_203_208_4	F1 seed, Cab203 x Dar208
	F1_CD_203_208_5	F1 seed, Cab203 x Dar208
	F1_CD_203_208_6	F1 seed, Cab203 x Dar208
	F1_DC_208_203_1	F1 seed, Dar208 x Cab203
	F1_DC_208_203_2	F1 seed, Dar208 x Cab203
	F1_DC_208_203_4	F1 seed, Dar208 x Cab203
	F1_DC_208_203_5	F1 seed, Dar208 x Cab203
	F1_DC_208_203_6	F1 seed, Dar208 x Cab203

2.10.2. Phenotypic analysis of an F_2 population

I assessed F_2 plants for flowering time variation under two vernalisation treatments; 6-weeks vernalisation (VERN) and no vernalisation (NVERN). For both vernalisation treatments, 72 F_2 seed were randomly selected from each of ten F_1

plants (Table 2. 12) giving a total F_2 population of 720 individuals per treatment. Parental lines were included as controls for both vernalisation treatments; 12 plants each of Cabriolet, and Darmor giving a total of 24 parent plants (Table 2. 12).

Table 2. 12: *Brassica napus* lines assessed for flowering time under VERN and NVERN treatments

Species	Line name	Number plants sown	Source
Brassica napus	Cabriolet	12	OREGIN, selfed, mother plant vernalised for 12 weeks
	Darmor	12	OREGIN, selfed seed, mother plant vernalised for 12 weeks
	F2_CD_203_208_1	72	F2 seed, selfed from F1_CD_203_208_1
	F2_CD_203_208_2	72	F2 seed, selfed from F1_CD_203_208_2
	F2_CD_203_208_4	72	F2 seed, selfed from F1_CD_203_208_4
	F2_CD_203_208_5	72	F2 seed, selfed from F1_CD_203_208_5
	F2_CD_203_208_6	72	F2 seed, selfed from F1_CD_203_208_6
	F2_DC_208_203_1	72	F2 seed, selfed from F1_DC_208_203_1
	F2_DC_208_203_2	72	F2 seed, selfed from F1_DC_208_203_2
	F2_DC_208_203_4	72	F2 seed, selfed from F1_DC_208_203_4
	F2_DC_208_203_5	72	F2 seed, selfed from F1_DC_208_203_5
	F2_DC_208_203_6	72	F2 seed, selfed from F1_DC_208_203_6

For both vernalisation treatments, seed were sown on soil in the glasshouse (18°C day/15°C night, 16-hour light/8-hour dark, 70% humidity); for the VERN treatment plants were grown for 28 days under glasshouse conditions then transferred to a vernalisation growth chamber (5°C, 8-hour light/16-hour dark, 70% humidity) for six weeks; for the NVERN treatment plants were grown for 28 days under glasshouse conditions. After each treatment all plants were transferred to a concrete floored poly-tunnel (Keder house; protected environment, no supplementary heating or lighting) at the John Innes Centre, Norwich in April 2017. Sowing of each

treatment was staggered so that plants from both treatments were transferred to the poly-tunnel on the same day. After one week of acclimation, plants were transplanted to 1-litre pots and positioned according to their vernalisation treatment in a randomised complete block design containing four blocks at a density of 36 plants/m^2. One hundred and eighty F_2 plants and three replicates of each parent line were included in each block.

Plants were watered twice daily by automatic irrigation and chemically sprayed by the Horticultural Services Team when required. Temperature and humidity within the poly-tunnel was recorded by Tinytag at 30-minute intervals for the duration of the experiment (Figure S. 3). I measured flowering time as the number of days to first flower open (BBCH60) according to Meier *et al*., 2001. Scoring of flowering commenced on the day plants were transferred to the poly-tunnel; 28 days after sowing for NVERN and 70 days after sowing for VERN.

2.10.3. DNA extraction and Next Generation sequencing

For DNA extraction, I sampled leaf tissue from all plants in the F_2 population prior to transplantation to the poly-tunnel. Whole leaves and 1cm leaf disks were sampled separately and stored in Qiagen Collection Microtubes and Greiner (Sigma-Aldrich) 96-well round bottom plates respectively. Leaf material was stored at -20°C until required.

For whole genome sequencing I extracted DNA using a CTAB (cetyl trimethylammonium bromide) method optimised to generate high quality DNA from *Brassica* for Illumina sequencing (see Supplementary Methods for specific details). For the parent lines, a single leaf sample originally taken from plants Cab203 (Cabriolet) and Dar208 (Darmor) were ground to a fine powder in liquid N_2 using a pestle and mortar prior to DNA extraction. For the F_2, lines were selected, and leaf material pooled based on flowering time phenotype. Leaf material was pooled prior to DNA extraction; one 1cm leaf disk from each individual plant to be included in each pool were combined and ground to a fine powder in liquid N_2 using a pestle and

mortar. Each DNA pool contained approximately 36 individuals, representing 5% of the population. Two DNA pools were generated for the NVERN treatment; the early pool contained approximately the 36 earliest flowering individuals of the population; the late pool contained 36 of the latest flowering individuals of the population. To identify candidate genes that conferred the extreme early and late flowering phenotypes within the F_2 population, four pools were generated for the VERN treatment. Two early pools were generated; one which contained approximately the 36 earliest lines; and a second pool which contained lines with a less extreme early phenotype flowering 37^{th} -72^{nd} in the population. I adopted a similar strategy for the late pools, generating two late pools; one which contained the 36 latest lines; and a second pool containing lines with less extreme phenotypes flowering 37^{th} -72^{nd} from the latest line.

Following extraction, DNA was quantified using a Nanodrop® spectrophotometer (ThermoScientific) and diluted to ~1200ng in a total volume of 25µl of H_20. DNA was then sent on dry ice to Novogene Co., Ltd., Hong Kong (HK) for paired-end whole genome sequencing by Illumina® at an average of 30x coverage. Sample quality control, library construction and sequencing were performed by Novogene Co., Ltd., HK.

2.10.4. Bioinformatic analysis

In collaboration with Dr D. M. Jones, the whole genome sequence data from both parents; Cabriolet (Cab203) and Darmor(Cab208), and from six samples of pooled leaf material representing the six DNA pools; NVERN_Early, NVERN_Late, VERN_Early_1, VERN_Early_2, VERN_Late_3, VERN_Late_4; were analysed according to methods similar to those previously described in rice by Takagi *et al*., (2013). All scripts associated with this analysis are available at https://github.com/marc-jones/brassica-napus-bulk-segregant.

Alignment to the Darmor-*bzh* reference genome

The Darmor-*bzh B. napus* genome sequence (Chalhoub *et al*., 2014, plants.ensembl.org, Genome Assembly: AST_PRJEB5043_v1, originally downloaded by Dr D. M. Jones in 2014) was used as a reference for sequence order. The sequencing reads from the eight datasets (Cabriolet, Darmor, NVERN_Early, NVERN_Late, VERN_Early_1, VERN_Early_2, VERN_Late_3, VERN_Late_4) were aligned to the Darmor-*bzh* genome reference sequence using Bowtie-2 v2.2.3 (Langmead *et al*., 2009, software available from http://bowtie-bio.sourceforge.net/bowtie2/index.shtml) to create eight separate alignment files in Sequence Alignment/Map (SAM) format. Using SAMtools v1.5 (Li *et al*., 2009, https://github.com/samtools/samtools) all SAM (.sam) files were converted to BAM (.bam) file format for use in all further analyses.

For stringency, sequence data from non-uniquely mapped reads were excluded. All non-uniquely mapped reads were filtered from the alignment using the "-q 42" parameter in SAMtools, ensuring that all aligned reads had a mapping quality score of at least 42. After alignment to the Darmor-*bzh* reference, and after filtering for non-uniquely mapped reads, a sequence PileUp was generated using the "mpileup" command in SAMtools (Li *et al*., 2009).

Filtering parameters and QTL-seq analysis

The following section describes the filtering parameters used to analyse the sequencing data to identify candidate genes that associated with flowering time variation.

First, to ensure accurate base calling, genomic positions where the read depth was <20 were excluded from the analysis. Excluding all genomic positions with a read depth of <20 removed 80-85% of the sequence data but allowed us to score SNPs with a high degree of accuracy. This filtering parameter was applied to all eight sequence datasets. Next, after filtering for adequate read depth, genetic differences

were identified in Cabriolet compared with Darmor and their genomic positions were scored according to the Darmor-*bzh* reference genome. To ensure simple SNPs and not hemi-SNPs (Trick *et al*., 2009) were scored, SNP and small InDels (<9bp) in Cabriolet were scored if the base call was found in >95% of sequence reads. All other base calls, including those from regions of heterozygosity within the parent lines, were excluded from further analysis. Genomic positions that were genetically identical i.e. no alternative base call was detected in the Cabriolet sequence compared with Darmor, were also excluded.

The sequence data from all six DNA pools were then analysed and SNP index values calculated. At each genomic position where a SNP or small InDel (<9bp) had been identified in Cabriolet compared with Darmor the total number of sequencing reads at that genomic position was counted for each DNA pool. The proportion of reads at each genomic position detected with an alternative base call, matching the Cabriolet sequence, was calculated to give a SNP index value. SNP index values equal to +1.0 represented a genomic position that was genetically identical to Cabriolet. In contrast, SNP index values that were equal to 0.0 represented a genomic position that was genetically identical to Darmor. Sequence calls that matched neither Cabriolet or Darmor were excluded from the total count.

A delta (Δ) SNP index value was then calculated to identify differences in allele frequencies between the DNA pools using the following equations:

SNP index NVERN_Late – SNP index NVERN_Early = ΔSNP index NVERN

SNP index VERN_Late_4 – SNP index VERN_Early_1 = ΔSNP index VERN

SNP index VERN_Late_3 – SNP index VERN_Early_2 = ΔSNP index VERN_ INTERMEDIATE

ΔSNP index values are expected to range between -1.0 and +1.0. When ΔSNP index= +1.0, the late flowering pool is genetically identical to Cabriolet and the early flowering pool identical to Darmor. When ΔSNP index= -1.0, the late flowering pool is genetically identical to Darmor and the early flowering pool like Cabriolet. To

reduce noise in the data, SNP index values that were <0.3 in both bulks were excluded.

SNP index and ΔSNP index values were visualised in R (R Studio, R version 3.5.0. "Joy in Playing") using the ggplot2 (Wickham 2016) and circlize (version 0.4.4., Gu *et al*., 2014, https://github.com/jokergoo/circlize) packages.

Bioinformatic analysis of *BnaFLC.A02* and *BnaFT.A02*

Read depth and sequence information for *BnaFLC.A02* and *BnaFT.A02* were extracted from the Cabriolet and Darmor sequence PileUp by Dr D. M. Jones. The read depth count was plotted against each base position of the Darmor-*bzh* reference for both *BnaFLC.A02* and *BnaFT.A02*.

To assess for DNA polymorphisms, sequences for *BnaFLC.A02* and *BnaFT.A02* were extracted from the Cabriolet and Darmor sequence PileUp by Dr D. M. Jones. I aligned these against *BnaFLC.A02* and *BnaFT.A02* sequences from Darmor-*bzh* (downloaded from plants.ensembl.org) and Cabriolet (in-house reference, JIC/TGAC Assembly, Dr R. Wells unpublished) using the AlignX (Invitrogen) software. By comparing these sequences, I could detect DNA polymorphisms between Cabriolet, Darmor and Darmor-*bzh*.

2.10.5. KASP marker: Design and Analysis

I designed twelve KASP markers to target a QTL region identified on chromosome A02 in *B. napus*. The SNP positions according to the Darmor-*bzh* reference genome for each marker designed are listed in Table 2. 13. At each SNP position a 1kb surrounding sequence was extracted from the Cabriolet and Darmor sequence PileUp by Dr D. M. Jones. After aligning these sequences using the Align X (Invitrogen) software I designed primers to target each SNP according to recommendations by LGC Group (https://www.lgcgroup.com/products/kasp-genotyping-chemistry). Each target amplicon was no larger than 120bp and the predicted annealing temperature

according to NetPrimer Software (Permier Biosoft) was approximately 58°C±4°C. For each target SNP I designed three primers; a conserved primer that was found in both Cabriolet and Darmor; and two allele specific primers that were fluorescently labelled; one FAM-labelled that would amplify in Darmor; and another HEX-labelled that would amplify in Cabriolet. To check for specificity, all designed primers were BLASTed against the JIC/TGAC Cabriolet genome assembly (R. Wells unpublished) using a short query parameter. If multiple hits were found, then I re-designed the primers. I designed four KASP markers at *BnaFLC.A02* (SNP positions 136045, 136554, 13832, 137493) and three designed at *BnaFT.A02* (SNP positions 6375505, 6376139, 6378254), the remaining five markers were positioned between the two genes on chromosome A02.

Table 2. 13: KASP markers designed for fine-mapping the QTL region on chromosome A02. The FAM and HEX sequence added to each allele specific primer is highlighted bold.

Target	Name	SNP position	Direction	Fluorescence tail	KASP marker sequence
CONSERVED	C_A02_136045	136045	R		GAAGCTTGTTGATAGCCTCAAAG
DARMOR	X_A02_136045	136045	F	FAM	**GAAGGTGACCAAGTTCATGCT**TGAGGCTCTCGACACAAGGTAG
CABRIOLET	Y_A02_136045	136045	F	HEX	**GAAGGTCGGAGTCAACGGATT**TGAGGCTCTCGACACAAGGTAA
CONSERVED	C_A02_136554	136554	R		TGCCTAGAGGGCTCCATACA
DARMOR	X_A02_136554	136554	F	FAM	**GAAGGTGACCAAGTTCATGCT**CGAACTAATGGCTCATGGGTGC
CABRIOLET	Y_A02_136554	136554	F	HEX	**GAAGGTCGGAGTCAACGGATT**CGAACTAATGGCTCATGAGTGG
CONSERVED	C_A02_136832	136832	R		TGCCATGGATTTATTGCTTGA
DARMOR	X_A02_136832	136832	F	FAM	**GAAGGTGACCAAGTTCATGCT**GTTCAACACCATTGTAATGGCA
CABRIOLET	Y_A02_136832	136832	F	HEX	**GAAGGTCGGAGTCAACGGATT**GTTCAACACCATTGTAATGGCG
CONSERVED	C_A02_137493	137493	F		CCGAAGTGCATTGCATTACAAA
DARMOR	X_A02_137493	137493	R	FAM	**GAAGGTGACCAAGTTCATGCT**GCTGTGATGCATACTTGATGTTGTT
CABRIOLET	Y_A02_137493	137493	R	HEX	**GAAGGTCGGAGTCAACGGATT**GCTGTGATGCATACTTGATGTTGTG
CONSERVED	C_A02_1290192	1290192	R		TGTCTGAGTGCCTCCTTGCA
DARMOR	X_A02_1290192	1290192	F	FAM	**GAAGGTGACCAAGTTCATGCT**TCTGCTCTCTGAATTTTGAGATTA
CABRIOLET	Y_A02_1290192	1290192	F	HEX	**GAAGGTCGGAGTCAACGGATT**TCTGCTCTCTGAATTTTGAGATTG
CONSERVED	C_A02_1797526	1797526	F		AGCTCTTGTTATCACCCTGCA
DARMOR	X_A02_1797526	1797526	R	FAM	**GAAGGTGACCAAGTTCATGCT**TCGTGACCTTACAGCTTCTTCT
CABRIOLET	Y_A02_1797526	1797526	R	HEX	**GAAGGTCGGAGTCAACGGATT**TCGTGACCTTACAGCTTCTTCA

Table 2. 13. continued

Target	Name	SNP position	Direction	Fluorescence tail	KASP marker sequence
CONSERVED	C_A02_1997465	1997465	R		AGCCAAGGATCCACATAATCA
DARMOR	X_A02_1997465	1997465	F	FAM	**GAAGGTGACCAAGTTCATGCT**AGCAAGGAGAATGCATATTCTCA
CABRIOLET	Y_A02_1997465	1997465	F	HEX	**GAAGGTCGGAGTCAACGGATT**AGCAAGGAGAATGCATATTCTCT
CONSERVED	C_A02_5703292	5703292	R		CCTACTCTTCTCGGCGTCTCTA
DARMOR	X_A02_5703292	5703292	F	FAM	**GAAGGTGACCAAGTTCATGCT**TCGGATCCTCTGTTTGATCA
CABRIOLET	Y_A02_5703292	5703292	F	HEX	**GAAGGTCGGAGTCAACGGATT**TCGGATCCTCTGTTTGATCC
CONSERVED	C_A02_6375505	6375505	R		CGTCTCCGACTTGTAACCCA
DARMOR	X_A02_6375505	6375505	F	FAM	**GAAGGTGACCAAGTTCATGCT**GTTATGATCACCGATCCGAACT
CABRIOLET	Y_A02_6375505	6375505	F	HEX	**GAAGGTCGGAGTCAACGGATT**TATGATCACCGATCCGAACC
CONSERVED	C_A02_6376139	6376139	R		TGCATGGCCAAGTTATAGTAGAA
DARMOR	X_A02_6376139	6376139	F	FAM	**GAAGGTGACCAAGTTCATGCT**GGATCTAAGGCCTTCTCAAGTTA
CABRIOLET	Y_A02_6376139	6376139	F	HEX	**GAAGGTCGGAGTCAACGGATT**GGATCTAAGGCCTTCTCAAGTTC
CONSERVED	C_A02_6378254	6378254	R		CCAACCTGTAATAAGTGTACCCAT
DARMOR	X_A02_6378254	6378254	F	FAM	**GAAGGTGACCAAGTTCATGCT**AATAATGTTGCTCTTACACCGGA
CABRIOLET	Y_A02_6378254	6378254	F	HEX	**GAAGGTCGGAGTCAACGGATT**AATAATGTTGCTCTTACACCGGC

I first tested all twelve KASP markers (Table 2. 13) according to LGC Group recommendations (https://www.lgcgroup.com/products/kasp-genotyping-chemistry). Three DNA samples; Cabriolet, Darmor, a mixed sample of equal concentrations of Cabriolet and Darmor; and a no template control (NTC) were used as test samples.

The reaction mix was prepared as follows:

2.5µl DNA (20ng/µl) (or H_2O for NTC)

2.5µl 2x KASP Master mix

0.07µl Primer mix*

*The primer mix was prepared as follows:

12µl Allele specific primer 1(100µM)

12µl Allele specific primer 2 (100µM)

30µl Common primer (100µM)

46µl H_2O

Each reaction mix was pipetted into a 384 well plate and placed in an Eppendorf Thermocycler using the following cycling conditions:

94°C for 15 minutes

10 cycles of: 94°C for 20 seconds, 61-55°C for 60 seconds (dropping 0.6°C per cycle)

26 cycles of: 94°C for 20 seconds, 55°C for 60 seconds

Hold at 15°C

After testing, I selected five markers that worked successfully. These were used to screen 384 DNA samples per vernalisation treatment, which represented 94 early

flowering lines, 94 late flowering lines, plus leaf material from Cabriolet and Darmor parental lines. The leaf material and KASP markers were sent to the JIC Genotyping and DNA Extraction Service for DNA extraction and KASP assaying.

Chapter 3: Analysis of natural and induced variation at *BnaFRI*

3.1. Introduction

FRI is often identified as a major determinant of vernalisation requirement and flowering time in *A. thaliana* (Johanson *et al*., 2000, Shindo *et al*., 2005, Strange *et al*., 2011). In a panel of 192 *A. thaliana* accessions considerable variation was detected at *FRI* (Shindo *et al*., 2005). At least twenty *FRI* haplotypes were predicted to encode non-functional FRI proteins and as a consequence many of these accessions were early flowering in the absence of vernalisation due to low *FLC* expression levels (Shindo *et al*., 2005). Allelic variation that is predicted to disrupt FRI protein function by introducing a premature stop codon will hereafter be referred to as "non-functional variation".

Two major non-functional *FRI* haplotypes are exemplified by the early flowering accessions Columbia (Col-0) and Landsberg *erecta* (L*er*) (Johanson *et al*., 2000, Shindo *et al*., 2005). Col-0 carries a 16bp frameshift mutation in exon 1 that leads to a premature stop codon at *FRI* (Johanson *et al*., 2000). Using *FRI* deletion constructs Risk *et al*., (2010) demonstrated the C-terminus of the FRI protein is required for protein function, while the N-terminus has minimal effect on *FLC* expression. Deletion of the C-terminus domain of the FRI protein would disrupt FRI-c formation, impacting FRI protein interaction with FLX, FES1 and SUF4, with the latter shown to bind directly to the *FLC* promoter (Choi *et al*., 2011). Therefore, the mutation found in Col-0 would disrupt the C-terminal end of the FRI protein and abolish FRI function, causing the early flowering phenotype. Landsberg *erecta* (L*er*) carries a 376bp deletion combined with a 31bp insertion in the promoter that removes the translational start codon of *FRI* (Johanson *et al*., 2000, Shindo *et al*., 2005, Schmalenbach *et al*., 2014). Despite the 5' deletion most of the coding sequence remains and the L*er* allele of *FRI* has been shown to have some retained function (Schmalenbach *et al*., 2014). Due to the presence of an in-frame alternative start codon, a FRI protein that has lost only 42 amino acids from its N-terminus has

been detected (Schmalenbach *et al*., 2014). When expressed under a wild-type *FRI* promoter the L*er* allele of *FRI* conferred later flowering in Col-0. As *FLC* in L*er* contains a transposon insertion, the early flowering phenotype of L*er* is therefore thought to be due to low expression of *FRI* caused by the deletion, combined with the presence of a weak allele at *FLC* (Schmalenbach *et al*., 2014). This study (Schmalenbach *et al*., 2014) was the first to demonstrate the occurrence and function of weak alleles of *FRI* in *A. thaliana*.

Due to the strong association between mutations at *FRI* and flowering time in *A. thaliana*, orthologues of *FRI* have been investigated in other members of the Brassicaceae (Kuittinen *et al*., 2008, Wang *et al*., 2011b, Guo *et al*., 2012, Irwin *et al*., 2012, Fadina *et al*., 2013, Fadina & Khavkin 2014, Yi *et al*., 2018). To date there is little evidence to suggest that variation at *FRI* in other Brassicaceae is responsible for variation in vernalisation requirement. For example, in an attempt to understand the genetic cause of variation in flowering time in *C.*, 20 accessions were analysed for sequence variation at *FRI* (Guo *et al*., 2012). Non-synonymous mutations were detected, but none were predicted to encode premature stop codons like those found in *A. thaliana* accessions (Guo *et al*., 2012). Sequence variation at *FRI* orthologues have also been investigated in natural populations of *Arabidopsis lyrata* (Kuittinen *et al*., 2008). 295 *FRI* sequences were analysed but only three carried mutations that would give rise to non-functional FRI proteins. Two common alleles were detected which exhibited insertion/deletion (InDel) variation that resulted in a 14-amino acid difference at the protein level. Both were transformed into *A. thaliana* Col-0 (*fri*) under the constitutive overexpressing promoter (35S::CaMV) and conferred late flowering. However, the allelic variation at the *A. lyrata FRI* orthologue contributed to relatively small differences in flowering times (an average difference of 9.2 days) suggesting the allelic variation at *FRI* orthologues do not contribute strongly to the large differences in flowering times (an average of >29 days) detected between *A. lyrata* populations (Kuittinen *et al*., 2008). Interestingly studies in *C. rubella* (Guo *et al*., 2012), *A. lyrata* (Kemi *et al*., 2013), *Arabis alpina* (Wang *et al*., 2009a, Albani *et al*., 2012), and *Arabidopsis arenosa* (Baduel *et al*., 2016) suggest variation at *FLC*, and not *FRI*, is associated with loss of vernalisation requirement and variation in flowering time.

Allelic variation has been reported at *FRI* orthologues of *B. rapa* (Fadina *et al*., 2013, Fadina & Khavkin 2014) *B. oleracea* (Irwin *et al*., 2012, Fadina *et al*., 2013, Fadina & Khavkin 2014) and *B. napus* (Wang *et al*., 2011b, Yi *et al*., 2018). When transformed into *A. thaliana* Col (*fri*) under the *A. thaliana FRI* promoter two major alleles of *BolFRI.C03*, distinguishing early from late *B. oleracea* cultivars, exhibited no functional variation and both conferred equally late flowering (Irwin *et al*., 2012). When expressed in *A. thaliana*, the non-synonymous variation at *BolFRI.C03* did not contribute to variation in vernalisation requirement.

A QTL co-localised with *BnaFRI.A03* in an oilseed rape (*B. napus*) mapping population derived from late flowering Tapidor, a winter oilseed rape, and Ningyou-7, a Chinese semi-winter oilseed rape that has a weak vernalisation requirement (Wang *et al*., 2011b). Non-synonymous mutations at the *BnaFRI.A03* gene were identified which significantly associated with flowering time (Wang *et al*., 2011b). Prior to 2018 no functional analysis of *BnaFRI.A03* had been carried out, and the function of other *FRI* orthologues in *A. lyratat* and *B. oleracea* (Kuittinen *et al*., 2008, Irwin *et al*., 2012) had not been assessed under their native promoters. Very recently however, a transgenic functional complementation experiment of *BnaFRI.A03* was published by Yi *et al*. (2018). This study demonstrated expression *BnaFRI.A03* in *A. thaliana*, under its native promoter, could confer late flowering. The allele of *BnaFRI.A03* found in the winter oilseed rape cultivar Tapidor conferred later flowering and had higher *FLC* expression compared with the allele found in Ningyou-7. Analysis of promoter swap constructs also demonstrated that coding sequence variation, and not gene expression variation, was more important for controlling flowering time however, the causal SNPs were not identified. In addition, phenotypic comparisons were made with Col-0 (*fri*) and not with an equivalent *A. thaliana* accession with a functional *FRI*, such as Col*FRI*^{SF2} (Lee *et al*., 1994b; Table 2. 3). From this publication we cannot therefore determine whether *BnaFRI.A03* controls variation for vernalisation requirement between *B. napus* cultivars.

Sequence variation has also been detected at *BnaFRI.C03*, and *BnaFRI.A10*in 5 spring oilseed rape and 5 winter oilseed rape cultivars and, unlike in *A. thaliana*, no nonsense mutations were identified (Yi *et al*., 2018). 174 cultivars representing spring oilseed rape, semi-winter oilseed rape, and winter oilseed rape, were subsequently analysed for InDel variation that associated with *BnaFRI* haplotypes and a significant association between *BnaFRI* haplotype and crop type was detected (Yi *et al*., 2018). To date, allelic variation at *BnaFRI* in other *B. napus* crop types, such as swede, fodder and kale, have not been investigated.

My aim for the present study was to investigate allelic variation at *BnaFRI* in a diverse population of *B. napus* cultivars including the both subspecies *ssp. napus* and *ssp. pabularia*. I was curious that non-synonymous mutations at *FRI* orthologues are nearly absent in many Brassicaceae. To investigate this, I analysed winter oilseed rape lines, kindly provided by Bayer CropScience NV, Belgium, that carried induced nonsense mutations at all four *BnaFRI* genes. The experiments discussed in this Chapter reveal *BnaFRI* as a weak controller of flowering time and promoter of *FLC* expression. Due to the association between *BnaFRI.A03* and flowering time I carried out a detailed sequence analysis of *BnaFRI.A03* which revealed coding and non-coding variation, in addition to the presence of a mutator-like transposon in the promoter region of spring oilseed rape. I undertook a transgenic complementation assay of *BnaFRI.A03* to determine the functional consequence of this allelic variation. *BnaFRI.A03* constructs were generated using promoter, open read frame, and terminator sequences from Chinese semi-winter oilseed rape, winter oilseed rape, spring oilseed rape, and swede alleles to determine which region of the *BnaFRI.A03* gene was important for flowering time variation between crop types. It is clear from my results that *BnaFRI* is not the primary driver of a vernalisation requirement in *B. napus*. The possibility that this is the case for most Brassicaceae is discussed.

3.2. Results

3.2.1. Coding and non-coding variation, but no expression variation, is identified at *BnaFRI*

Allelic variation at *BnaFRI* has previously been reported in cultivars of Chinese semi-winter, spring and winter oilseed rape (Wang *et al*., 2011b, Yi *et al*., 2018). Here I investigated the allelic variation present at *BnaFRI* genes within the OREGIN DFFS population. 95 cultivars (listed in Table 2. 4) including 5 Chinese semi-winter oilseed rape, 15 spring oilseed rape, 52 winter oilseed rape, 11 swede, 5 winter fodder, 4 kale, 1 exotic (leafy type cultivar), 1 synthetic, and 1 winter vegetable crop types were assessed. Using primers targeting exon 1 (Table 2. 5) all 95 cultivars were screened for allelic variation and the number of alleles at each *BnaFRI* locus were counted. Based on DNA sequence nine, five, four, and two alleles of *BnaFRI.A03*, *BnaFRI.A10*, *BnaFRI.C03* and *BnaFRI.C09* respectively were identified in the *B. napus* population (Figure 3. 1). At least one allele at each *BnaFRI* gene, denoted *BnaFRI.A03*-a, *BnaFRI.A10*-a, *BnaFRI.C03*-a, *BnaFRI.C09*a, was identical in sequence to the winter oilseed rape Darmor-*bzh* reference sequence (Chalhoub *et al*. 2014, plants.ensembl.org) and to the published *BnaFRI* sequence from the winter oilseed rape cultivar Express-617 (Wang *et al*., 2011b). The remaining alleles identified for *BnaFRI.A03*, *BnaFRI.A10*, *BnaFRI.C03* and *BnaFRI.C09* exhibited DNA polymorphisms compared with their respective *BnaFRI* reference sequences.

A

	Amino acid position																					
BnaFRI.A03	7	35	36	37	38	39	40	41	46	47	63	69	85	108	109	115	188	216	219	285	286	300
BnaA.FRI.a_protein_JN936850	S	E	T	V	P	T	N	I	E	Q	L	V	E	P	N	V	K	K	K	A	E	I
BnaFRI.A03-a	S	E	T	V	P	T	N	I	E	Q	L	V	E	P	N	V	K	K	K	A	E	I
BnaFRI.A03-b	S	-	-	-	-	-	-	-	G	P	F	V	E	Q	N	I	K	K	K	A	E	I
BnaFRI.A03-c	S	E	T	V	P	T	N	I	E	Q	L	A	G	Q	N	I	N	GE	K	A	E	I
BnaFRI.A03-d	S	E	T	V	P	T	N	I	E	Q	L	A	G	Q	N	I	N	K	K	A	E	I
BnaFRI.A03-e	S	E	T	V	P	T	N	I	E	Q	L	A	G	Q	N	I	K	K	K	D	E	M
BnaFRI.A03-f	S	E	T	V	P	T	N	I	E	Q	L	A	G	Q	N	I	N	GE	K	D	E	I
BnaFRI.A03-g	T	E	T	V	P	T	N	I	E	Q	L	A	G	P	D	I	N	K	Q	A	E	I
BnaFRI.A03-h	S	E	T	V	P	T	N	I	E	Q	L	A	G	P	D	I	K	K	K	A	D	I
BnaFRI.A03-i	S	E	T	V	P	T	N	I	E	Q	L	A	G	P	N	V	K	K	K	D	E	I

B

	Amino acid position																								
BnaFRI.A10	39	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	201	209	212	215	233	234	243	261
BnaX.FRI.b_protein_JN936851	Q	V	A	A	T	Q	S	P	P	K	E	T	C	E	T	V	A	E	S	K	F	V	M	S	A	S
BnaFRI.A10-a	Q	V	A	A	T	Q	S	P	P	K	E	T	C	E	T	V	A	E	S	K	F	V	M	S	A	S
BnaFRI.A10-b	Q	V	A	A	T	Q	S	P	P	-	-	-	-	-	-	V	A	E	S	K	F	V	M	S	A	T
BnaFRI.A10-c	H	A	T	A	V	A	A	T	Q	SPQK	E	T	C	E	T	V	A	D	C	-	L	G	I	N	P	S
BnaFRI.A10-d	Q	V	A	E	I	S	-	-	-	-	-	-	-	-	-	-	-	D	S	K	F	V	M	S	A	S
BnaFRI.A10-e	Q	V	A	A	T	Q	S	P	P	-	-	-	-	-	-	V	A	E	S	-	L	G	M	S	A	S

C

	Amino acid position							
BnaFRI.C03	36	90	95	99	100	101	102	115
BnaX.FRI.d_protein_JN936853	T	I	A	S	P	N	K	E
BnaFRI.C03-a	T	I	A	S	P	N	K	E
BnaFRI.C03-b	VPTNIETT	V	V	PKDA	S	G	E	Q
BnaFRI.C03-c	T	I	A	S	P	N	K	E
BnaFRI.C03-d	VPTNIETT	V	V	PPKDA	S	G	E	E

D

	Amino acid position				
BnaFRI.C09	288	289	290	291	292
BnaX.FRI.c_protein_JN936852	R	K	S	G	T
BnaFRI.C09-a	R	K	S	G	T
BnaFRI.C09-b	Q	W	Y	C	STOP

Figure 3. 1: Allelic variation identified at *BnaFRI*. Predicted amino acid substitutions for the alleles detected in 95 *B. napus* cultivars for (A) BnaFRI.A03, (B) BnaFRI.A10, (C) BnaFRI.C03, and (D) BnaFRI.C09. Sequence comparisons were made to the *BnaFRI* genes previously published by Wang *et al*., (2011b) for the cultivar Express-617 available from GenBank, accession numbers: JN936850 (BnaFRI.A03), JN936851 (BnaFRI.A10), JN936853 (BnaFRI.C03), and JN936852 (BnaFRI.C09). Amino acid substitutions were predicted based on published BnaFRI protein sequences (Wang *et al*. 2011b)

Non-functional variation at *BnaFRI* has not previously been reported (Wang *et al*., 2011b, Yi *et al*., 2018). However, in the present study I detected a frameshift mutation in one cultivar from the OREGIN DFFS population. A swede cultivar called Jaune a Collet Vert carries a 5bp deletion that is predicted to cause a premature stop codon at *BnaFRI.C09*. This premature stop codon reduces the predicted BnaFRI.C09 protein sequence from 584 amino acids in length to 291 amino acids (Figure 3. 1 D) and deletes the C-terminus domain predicted to be essential for FRI protein function in *A. thaliana*. No other cultivar carried this allele of *BnaFRI.C09* and the remaining 94 cultivars carried *BnaFRI.C09-a*. Sequence conservation of *BnaFRI.C09* was also observed by Yi *et al*., (2018) possibly

suggesting this gene may have an important function but that this gene is unlikely to contribute to variation in flowering time between cultivars.

One allele of *BnaFRI.*C03, denoted *BnaFRI.C03*-c, contained polymorphisms in exon 1 but all were synonymous leading to no changes in amino acid sequence (Figure 3. 1C). Winter oilseed rape cultivar Cabriolet carried this allele of *BnaFRI.C03*. An in-house genome sequence assembly is available for Cabriolet (R. Wells, unpublished), therefore I extracted the *BnaFRI.C03* gene sequence from the Cabriolet genome assembly and aligned to the reference sequence for *BnaFRI.C03* generated from the winter oilseed rape cultivar Express-617 (Wang *et al*., 2011b). Non-synonymous mutations located in exon 3 distinguish Cabriolet from Express-617 at *BnaFRI.C03* (Figure 3. 2). These mutations are predicted to cause six amino acid substitutions and produce an amino acid sequence with 97% identity to the E1 allele of *BolFRI.C03*, an allele associated with early flowering *B. oleracea* cultivars (Irwin *et al*., 2012).

```
                                1                                                50
BnaX.FRI.d_FRIGIDA_JN936853    (1) MAVRNGYAHRPSTREEEQPSSAMIRRREAQATVETTIEQSNDPQFLKSIV
 CABRIOLET_BnaFRIC3_PROTEIN    (1) MAVRNGYAHRPSTREEEQPSSAMIRRREAQATVETTIEQSNDPQFLKSIV
                               51                                               100
BnaX.FRI.d_FRIGIDA_JN936853   (51) DLTALAAAVNAFKRRYDELQSHMDYIENAIDSNLKTNGIIEIAAASPPSP
 CABRIOLET_BnaFRIC3_PROTEIN   (51) DLTALAAAVNAFKRRYDELQSHMDYIENAIDSNLKTNGIIEIAAASPPSP
                              101                                               150
BnaX.FRI.d_FRIGIDA_JN936853  (101) NKTATAIACQSPPKEKSEAERLCESMCSKELRRYMFVNISERAKLIEELP
 CABRIOLET_BnaFRIC3_PROTEIN  (101) NKTATAIACQSPPKEKSEAERLCESMCSKELRRYMFVNISERAKLIEELP
                              151                                               200
BnaX.FRI.d_FRIGIDA_JN936853  (151) GALKLAKDPAKFVLDCIGKFYLQGRKAFANDSPAITARKVSLLVLECYLL
 CABRIOLET_BnaFRIC3_PROTEIN  (151) GALKLAKDPAKFVLDCIGKFYLQGRKAFANDSPAITARKVSLLVLECYLL
                              201                                               250
BnaX.FRI.d_FRIGIDA_JN936853  (201) TFDPEGEKKQVGSSVKDEAEAAAVAWKKRLVGEGWLGAAEAVDARGLLLL
 CABRIOLET_BnaFRIC3_PROTEIN  (201) TFDPEGEKKQVGSSVKDEAEAAAVAWKKRLVGEGWLGAAEAVDARGLLLL
                              251                                               300
BnaX.FRI.d_FRIGIDA_JN936853  (251) VACFGIPESFKSMDLLDLIRQSGTAEIVGALKRSPFLVPMMSGIVDSSIK
 CABRIOLET_BnaFRIC3_PROTEIN  (251) VACFGIPESFKSMDLLNLIRQSGTAEIVGALKRSPFLVPMMSGIVDSSIK
                              301                                               350
BnaX.FRI.d_FRIGIDA_JN936853  (301) RGMHIEALEMVYTFGMEDRFSPSSILTSFLRMSKESFERAKRKAQAPMAS
 CABRIOLET_BnaFRIC3_PROTEIN  (301) RGMHIEALEMVYTFGMEDRFSPSSILTSFLRMSKESFERAKRKAQAPMAS
                              351                                               400
BnaX.FRI.d_FRIGIDA_JN936853  (351) KTANEKQLDALSSVMKCLEAHKLDPAKEVPGWQIQEQMAKLEKEIVQLDK
 CABRIOLET_BnaFRIC3_PROTEIN  (351) KTANEKQLDALSSVMKCLEAHKLDPVKEVPGWQIQEQMAKLEKEIVQLDK
                              401                                               450
BnaX.FRI.d_FRIGIDA_JN936853  (401) QMEEARSISRMEEARSISRMEEAAISQRLYNQQMKRPRLSEREMPPTASL
 CABRIOLET_BnaFRIC3_PROTEIN  (401) QMEEARSISRMEEARSISLREEAAISERLYNQQMKRPRLSEREMPPTASL
                              451                                               500
BnaX.FRI.d_FRIGIDA_JN936853  (451) SYSPMYRDQSFPSHREGDADEISALVSSYLGPSAGFPHRSSLRRSPEYMV
 CABRIOLET_BnaFRIC3_PROTEIN  (451) SYSPMYRDQSFPSHREGDADEISALVSSYLGPSAGFPHRSSLRRSPEYMV
                              501                                               550
BnaX.FRI.d_FRIGIDA_JN936853  (501) PPGGLGRSVSAYDHQPPTSYSPVSRRYSPVHGQRLPQEYSLPVHGQHQMP
 CABRIOLET_BnaFRIC3_PROTEIN  (501) PPGGLGRSVSAYDHQPPNSYSPVSRRYSPVHGQRLPQEYSLPVHGQHQMP
                              551                                583
BnaX.FRI.d_FRIGIDA_JN936853  (551) YGLYRHSPSVERYLALSNHRTPRNLSQDRIGGM
 CABRIOLET_BnaFRIC3_PROTEIN  (551) YGLYRHSPSVERYLALSNHRTPRNLSQDRIGGM
```

Figure 3. 2: The predicted protein sequence of BnaFRI.C03. Illustrated are the sequences for winter oilseed rape cultivar Cabriolet which carries allele *BnaFRI.C03-c* compared with the *BnaFRI.C03-a* allele sequence published from the winter oilseed rape cultivar Express-617 (GenBank accession JN936853; Wang *et al*., 2011b)

Using InDel variation, the alleles of *BnaFRI* could be grouped into haplotypes (as per Yi *et al*., 2018). From this I detected three haplotype groups for *BnaFRI.A03* (A03HAP1: *BnaFRI.A03-a*, *-d*, *-e*, *-g*, *-h*, *-i*, A03HAP2: *BnaFRI.A03-b*, A03HAP3: *BnaFRI.A03-c* and *-f*), four haplotype groups for *BnaFRI.A10* (A10HAP1: *BnaFRI.A10-a*, A10HAP2: *BnaFRI.A10-b* and *-e*, A10HAP3: *BnaFRI.A10-c*, A10HAP4: *BnaFRI.A10-d*), two haplotype groups for *BnaFRI.C03* (C03HAP1: *BnaFRI.C03-a* and *-c*, C03HAP2: *BnaFRI.C03-b* and *-d*), and two haplotype groups for *BnaFRI.C09* (C09HAP1: *BnaFRI.C09-a*, C09HAP2: *BnaFRI.C09-b*). Cultivars were grouped according to their crop type (information provided by OREGIN; Table 2. 4) and their flowering habit was predicted from this crop type classification. I classified the crop types Chinese semi-winter oilseed rape, spring oilseed rape and synthetic as early flowering without vernalisation; and I classified crop types exotic, kale, swede, winter fodder, winter oilseed rape, winter vegetable as late flowering without vernalisation. I then assessed the frequency of each haplotype for both flowering groups using a Chi-square analysis (as per Yi *et al*., 2018) to determine whether *BnaFRI* haplotype was significantly associated with early or late flowering (Table 3. 1). *BnaFRI.C09* was excluded from the analysis as *BnaFRI.C09-b* was detected in only one cultivar

Table 3. 1: Chi-square analysis of *BnaFRI* haplotype distribution within two fowering habit groups of *B. napus*.

Chi-square test for *BnaFRI.A03* haplotypes		A03HAP1[a]		A03HAP2		A03HAP3			X^2
Flowering habit	Total number of lines	Observed	Expected	Observed	Expected	Observed	Expected		(2 df)
EARLY	19	9	6.33	7	6.33	3	6.33		2.95
LATE	72	62	24	3	24	7	24		90.58[c]***

Chi-square test for *BnaFRI.A10* haplotypes		A10HAP1		A10HAP2		A10HAP3		A10HAP4		X^2
Flowering habit	Total number of lines	Observed	Expected	Observed	Expected	Observed	Expected	Observed	Expected	(3 df)
EARLY	20	6	5	6	5	8	5	0	5	7.2
LATE	72	56	18	9	18	4	18	3	18	108.11***

Chi-square test for *BnaFRI.C03*haplotypes		C03HAP1		C03HAP2			X^2
Flowering habit	Total number of lines	Observed	Expected	Observed	Expected		(1 df)
EARLY	21	3	10.5	18	10.5		10.71*
LATE	72	47	36	25	36		6.72*

H_0: *BnaFRI* haplotypes follow the same distribution across the two flowering habit groups of *B. napus*.

[a] haplotype based on InDel markers, [b] degrees of freedom, [c] *** $p<0.001$, * $p<0.05$

A significant association between late flowering *B. napus* cultivars and *BnaFRI* haplotype was detected ($p<0.05$; χ^2 values of 90.58, 108.11, and 6.72 for *BnaFRI.A03*, *BnaFRI.A10, BnaFRI.C03* respectively). A high percentage of late flowering cultivars; 86%, 78%, 65%; carried haplotypes A03HAP1, A10HAP1, and C03HAP1 respectively perhaps suggesting relative conservation at all *BnaFRI* loci might be important for vernalisation requirement and late flowering.

A significant association between early flowering *B. napus* cultivars and *BnaFRI* haplotype was detected for *BnaFRI.C03* ($p<0.05$; χ^2 value of 10.71) with a majority (86%) of early flowering cultivars carrying C03HAP2. No significant association was detected between early flowering cultivars and haplotype variation at *BnaFRI.A03* and *BnaFRI.A10* (χ^2 values of 2.95 and 7.2 respectively). This is unsurprising as we know from *A. thaliana* that multiple alleles of *FRI* can give rise to similar early flowering phenotypes. When analysing many *A. thaliana* accessions with different mutations at *FRI* the association between *FRI* and flowering is not significant (Atwell *et al*. 2010). Multiple mutations at *FRI* can confound the association as many have a similar effect on flowering. Perhaps it is similar here, several different haplotypes of *BnaFRI.A03* and *BnaFRI.A10* could give rise to an early flowering phenotype, possibly explaining a lack of association between their respective haplotypes and flowering time.

To confirm the absence of non-functional variation at *BnaFRI* I used a combination of previously published primers (Wang *et al*., 2011b) and some designed by the myself (Table 2. 5), to amplify by PCR and sequence the *BnaFRI.A03* full gene sequence, the *BnaFRI.A10* full gene sequence, a partial fragment covering exon 3 from *BnaFRI.C03* and the *BnaFRI.C09* full gene sequence. In the small selection of cultivars analysed coding and non-coding variation was detected at *BnaFRI.A03* (Figure 3. 3 A), *BnaFRI.A10* (Figure 3. 3 B) and *BnaFRI.C03* (Figure 3. 3 C) whereas no additional sequence variation was detected at *BnaFRI.C09*, in agreement with previous findings (Yi *et al*., 2018). Within the sample of cultivars tested for each *BnaFRI* gene, no non-functional variation was detected.

A

BnaFRI.A03	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128
Bronowski	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Hanna	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Stellar	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Wetar-10	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Bienvenue	G	T	G	C	C	T	A	C	A	A	A	C	A	T	C	G	A	A	A	C	C
Express 617	G	T	G	C	C	T	A	C	A	A	A	C	A	T	C	G	A	A	A	C	C
Major	G	T	G	C	C	T	A	C	A	A	A	C	A	T	C	G	A	A	A	C	C
Tapidor	G	T	G	C	C	T	A	C	A	A	A	C	A	T	C	G	A	A	A	C	C
Shengliyoucai-11	G	T	G	C	C	T	A	C	A	A	A	C	A	T	C	G	A	A	A	C	C
Sensation NZ	G	T	G	C	C	T	A	C	A	A	A	C	A	T	C	G	A	A	A	C	C
Wilhelmsburger R	G	T	G	C	C	T	A	C	A	A	A	C	A	T	C	G	A	A	A	C	C
	EXON 1																				

BnaFRI.A03 (Position from ATG)	136	137	139	188	197	205	253	281	322	323	342	347	416	488	563	645	761	800	853	899	957	971	972	973	974	1041	1223	1701	1789	1948	1979	2083	2157
Bronowski	A	A	A	A	A	C	G	T	A	A	A	T	C	G	C	A	A	G	C	C	G	-	-	-	-	A	A	C	G	G	C	G	A
Hanna	G	G	C	T	G	T	A	C	A	G	A	G	C	G	A	A	A	G	C	C	G	-	-	-	-	A	A	C	A	T	C	G	A
Stellar	G	G	C	T	G	T	A	C	A	G	A	G	C	G	A	A	A	G	C	C	G	-	-	-	-	A	A	C	A	T	C	G	A
Wetar-10	G	G	C	T	G	T	A	C	A	G	A	G	C	G	A	A	A	G	C	C	G	-	-	-	-	A	A	C	A	T	C	G	A
Bienvenue	A	A	A	A	G	T	A	C	C	G	G	G	T	G	A	A	C	C	C	C	A	A	G	A	A	T	T	C	G	G	C	G	A
Express 617	A	A	A	A	G	T	A	C	C	G	G	G	T	G	A	A	C	C	C	C	A	A	G	A	A	T	T	C	G	G	C	G	A
Major	A	A	A	A	G	T	A	C	C	G	G	G	T	G	A	A	C	C	C	C	A	A	G	A	A	T	T	C	G	G	C	G	A
Tapidor	A	A	A	A	G	T	A	C	C	G	G	G	T	G	A	A	C	C	C	C	A	A	G	A	A	T	T	C	G	G	C	G	A
Shengliyoucai-11	A	A	A	A	A	C	G	T	A	A	A	T	C	G	A	A	A	G	A	G	G	-	-	-	-	A	T	T	G	G	C	G	A
Sensation NZ	A	A	A	A	A	C	G	T	A	A	A	T	T	T	C	GGAG	C	G	C	C	G	-	-	-	-	A	A	C	G	G	A	C	G
Wilhelmsburger R	A	A	A	A	A	C	G	T	A	A	A	T	T	T	C	GGAG	C	G	C	C	G	-	-	-	-	A	A	C	G	G	A	C	G
	EXON 1																					INTRON 1					EXON 2	EXON 3					

B

BnaFRI.A10	44	83	116	310	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352
Bronowski	CACGAGGGAAAATCAATCATC	T	A	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Hanna	CACGAGGGAAAATCAATCATC	T	A	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Stellar	CACGAGGGAAAATCAATCATC	G	T	AGCCACCGC	A	G	A	A	G	G	A	G	A	C	T	T	G	T	G	A	G	A	C
Westar-10	CACGAGGGAAAATCAATCATC	G	T	AGCCACCGC	A	G	A	A	G	G	A	G	A	C	T	T	G	T	G	A	G	A	C
Bienvenue	-	G	A	-	C	A	A	A	G	G	A	G	A	C	T	T	G	T	G	A	G	A	C
Express 617	-	G	A	-	C	A	A	A	G	G	A	G	A	C	T	T	G	T	G	A	G	A	C
Major	-	G	A	-	C	A	A	A	G	G	A	G	A	C	T	T	G	T	G	A	G	A	C
Tapidor	-	G	A	-	C	A	A	A	G	G	A	G	A	C	T	T	G	T	G	A	G	A	C
Ningyou-7	-	G	A	-	C	A	A	A	G	G	A	G	A	C	T	T	G	T	G	A	G	A	C
	EXON 1																						

BnaFRI.A10 (Posistion from ATG)	362	572	596	601	623	624	625	626	636	643	668	698	700	719	726	779	781	893	966	1094	1146	1152	1233	1277	1311	1352	1375	1456	1582
Bronowski	G	G	G	C	G	A	A	G	T	T	G	G	G	G	G	G	C	T	C	T	T	T	A	C	C	G	G	T	C
Hanna	G	G	G	C	G	A	A	G	T	T	G	G	G	G	G	G	C	T	C	T	T	T	A	C	C	G	G	T	C
Stellar	T	C	C	G	-	-	-	C	C	G	C	C	A	C	C	C	G	C	G	C	A	G	C	T	T	T	T	A	C
Westar-10	T	C	C	G	-	-	-	C	C	G	C	C	A	C	C	C	G	C	G	C	A	G	C	T	T	T	T	A	C
Bienvenue	G	G	G	C	G	A	A	G	T	T	G	G	G	G	G	C	G	T	C	T	T	T	A	C	C	G	G	T	G
Express 617	G	G	G	C	G	A	A	G	T	T	G	G	G	G	G	C	G	T	C	T	T	T	A	C	C	G	G	T	G
Major	G	G	G	C	G	A	A	G	T	T	G	G	G	G	G	C	G	T	C	T	T	T	A	C	C	G	G	T	G
Tapidor	G	G	G	C	G	A	A	G	T	T	G	G	G	G	G	C	G	T	C	T	T	T	A	C	C	G	G	T	G
Ningyou-7	G	G	G	C	G	A	A	G	T	T	G	G	G	G	G	C	G	T	C	T	T	T	A	C	C	G	G	T	G
	EXON 1																	INTRON 1			EXON 2			INTRON 2		EXON 3			

C

BnaFRI.C03 (Position relative to ATG)	1614	1647	1648	1661	1664	1684	1932	1953	1958
Hanna	A	G	C	G	G	G	C	T	A
Express	G	C	A	T	T	C	A	A	C
Major	G	C	A	T	T	C	A	A	C
Tapidor	A	G	C	G	G	G	C	T	A
Ningyou-7	A	G	C	G	G	G	C	T	A
Shengliyoucai-11	A	G	C	G	G	G	C	T	A
	EXON 3								

Figure 3. 3: DNA sequence polymorphisms detected at *BnaFRI* in *B. napus*. (A) *BnaFRI.A03*, (B) *BnaFRI.A10*, and (C) *BnaFRI.C03*. SNPs and InDels are listed according to their position from the translational start codon of the published *BnaFRI* sequence (Wang *et al*., 2011b), GenBank accession numbers: JN936850 (*BnaFRI.A03*), JN936851 (*BnaFRI.A10*), JN936853 (*BnaFRI.C03*)

Due to the absence of non-functional variation within the amino acid sequence of BnaFRI, I determined the quantitative gene expression in a winter oilseed rape cultivar, Tapidor, and a spring oilseed rape cultivar, Westar-10 (see Section 2. 6 for experimental methods). In agreement with previous reports (Wang *et al.*, 2011b) *BnaFRI.A03*, *BnaFRI.A10*, *BnaFRI.C03*, and *BnaFRI.C09* were expressed in both cultivars, but although expressed at higher levels in the winter oilseed rape cultivar Tapidor, no significant quantitative difference in expression were detected between the cultivars (Figure 3. 4). At the developmental timepoint analysed (7 days after germination) the quantitative expression of *BnaFRI* is unlikely to have a significant effect on flowering time.

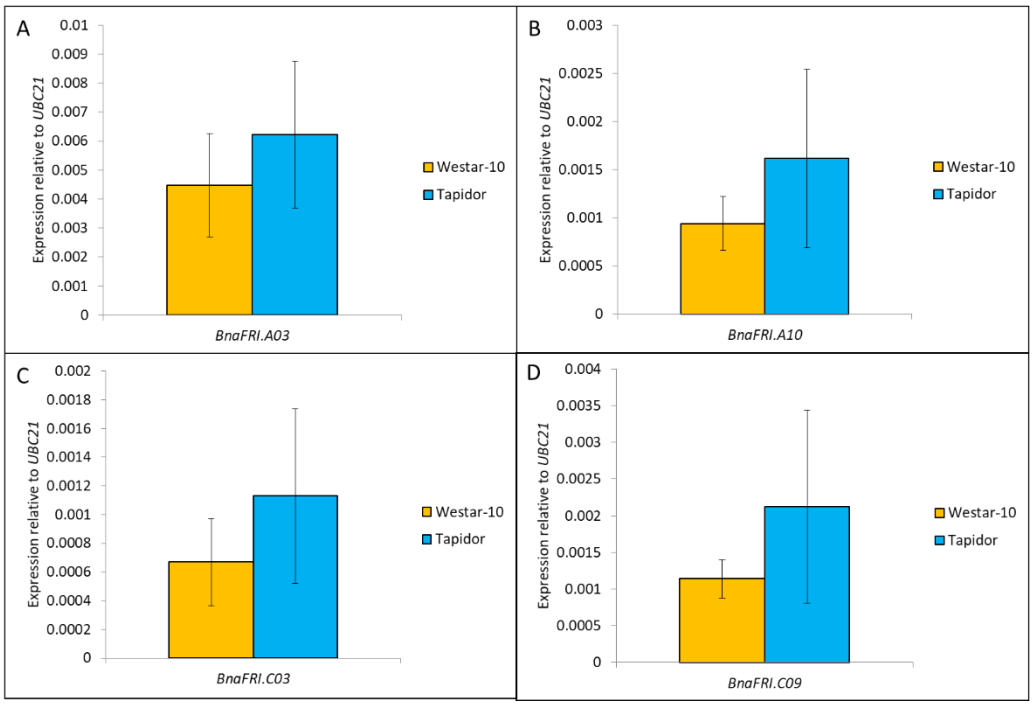


Figure 3. 4: Quantitative expression of *BnaFRI*. (A) *BnaFRI.A03*, (B) *BnaFRI.A10*, (C) *BnaFRI.C03*, and (D) *BnaFRI.C09*; in spring oilseed rape cultivar Westar-10 (orange) and winter oilseed rape cultivar Tapidor (blue). Samples were taken seven days after germination. Mean expression was calculated from three biological replicates normalised to the internal reference gene *UBC21*; error bars denote 95% confidence interval around the mean.

3.2.2. Induced mutations at *BnaFRI* have no significant effect on flowering time and *BnaFLC* expression

Due to the near absence of nonsense mutations at *BnaFRI* I investigated the function of the four *BnaFRI* genes using a population of winter oilseed rape breeding lines that carried chemically induced mutations at *BnaFRI*. Prior to this work a chemically mutagenized (Ethyl methanesulfonate; EMS) spring oilseed rape line generated by Bayer CropScience NV, Belgium, was genotyped for mutations at *BnaFRI*. Four mutations were identified (Figure 3. 5); SNP448:C>Tin *BnaFRI.A03*, SNP373:C>T in *BnaFRI.A10*, SNP343:G>T in *BnaFRI.C03*, and SNP355C>T in *BnaFRI.C09*. All are predicted to encode premature stop codons, and translated into proteins of 150, 115, 125, and 119 amino acids in length for BnaFRI.A03, BnaFRI.C03, BnaFRI.A10 and BnaFRI.C09 respectively. These mutations were backcrossed at Bayer into an early flowering winter oilseed rape breeding line to generate single, double, triple and quadruple mutants. It must be noted that the early flowering winter oilseed rape breeding line used as the parental line was generated by introgression of a region of chromosome A10, which contains *BnaFLC.A10*, from a spring oilseed rape breeding line.

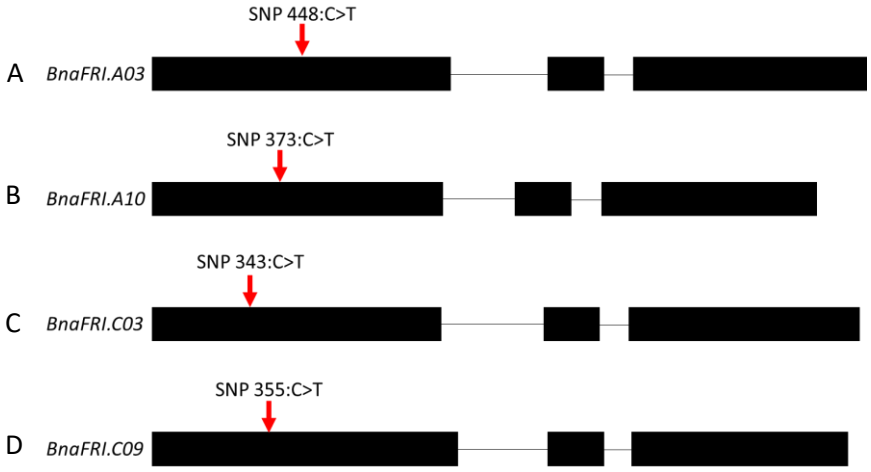


Figure 3. 5: A schematic illustration of the location of mutations generated by EMS that encode for premature stop codons at *BnaFRI*. (A) *BnaFRI.A03*, (B) *BnaFRI.A10*, (C) *BnaFRI.C03*, and (D) *BnaFRI.C09*. Black boxes denote exons, black lines denote introns, red arrows highlight relative position of the nonsense mutation detected.

A total of twelve *Bnafri* mutants, one wild-type line (line 149), and the original winter oilseed rape parental line (line 150) were available from Bayer CropScience NV, Belgium (see Table 2. 8 for specific details) and I assessed these lines with and without vernalisation treatment. I assessed the flowering times (scored as the number of days to BBCH60 according to Meier *et al*., 2001) of the quadruple (line 134) and triple (lines 135, 136, 137 and138) *Bnafri* mutant lines. These were compared with the wild-type line (line 149 hereafter referred as *BnaFRI* wild-type line) and the spring oilseed rape cultivar Westar-10 under glasshouse conditions (16 light/8 hours dark, 18°C day temperature, 15°C night temperature, and 70% humidity) with (6, 4, and 2-weeks at 5°C, 8 hours light/16 hours dark, 70% humidity) and without vernalisation. The results of this experiment are illustrated in Figure 3. 6. All lines flowered early (<60 days) without vernalisation and showed an acceleration in flowering time with increasing lengths of vernalisation. Under all treatments the *BnaFRI* wild-type line, and the mutants (lines 134, 135, 136, 137, and 138) exhibited comparable and non-significant flowering times ($0.167<p<0.998$, ANOVA). In addition to the quadruple and triple *Bnafri* mutants, all single and double *Bnafri* mutant lines (lines 139-148; Table 2. 8) were assessed for flowering time (BBCH60) under glasshouse conditions without vernalisation (Figure S. 1). No *Bnafri* mutation had a significant effect on flowering time ($p= 0.238$, ANOVA) compared with the wild-type *BnaFRI line* therefore, under the conditions tested, nonsense mutations at *BnaFRI* genes had no significant effect on flowering time.

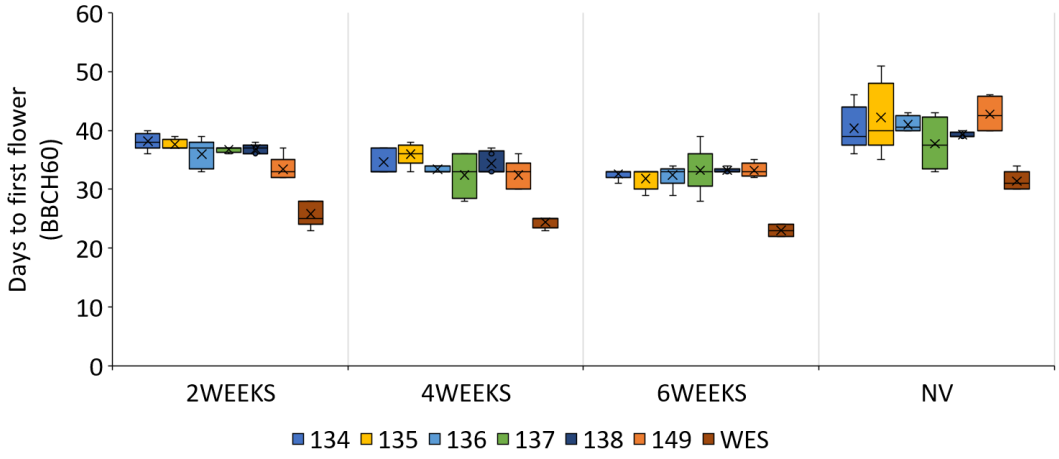


Figure 3. 6: Assessment of flowering time in *Bnafri* mutants. *Bnafri* mutant lines 134, 135, 136, 137 and 138; *BnaFRI* wild-type line 149; and spring oilseed rape cultivar Westar-10. Vernalisation treatments were: 2-weeks, 4-weeks, 6-weeks and no vernalisation (NV). Boxplots illustrate the median, mean (x), and range values calculated from five replicate plants per line and per treatment.

For comparison Westar-10, an early flowering spring oilseed rape cultivar, was included in my experiment. The *Bnafri* mutant lines and the *BnaFRI* wild-type line flowered significantly later than Westar-10 after vernalisation ($p<0.001$ all vernalisation treatments, ANOVA). Although not significant ($p=0.133$, ANOVA) In the absence of vernalisation (NV) Westar-10 flowered 10.4 days earlier on average than the *BnaFRI* wild-type line and 9.6 days earlier than the earliest *Bnafri* mutant line (line 136). Increased variability between replicate plants under no vernalisation conditions was observed; a possible explanation for not detecting significant difference in flowering times between lines. This experiment has revealed that non-functional alleles of *BnaFRI* do not confer early, spring oilseed rape-like flowering. Therefore, it is assumed from this that mutations at loci independent from *BnaFRI* contribute to the early flowering phenotype of Westar-10.

Quantitative expression of *BnaFLC* was measured before and after four weeks of vernalisation in the quadruple (line 134), and triple mutants (lines 135, 136, 137, and 138), and compared with the *BnaFRI* wild-type line and Westar-10 (Wes); Figure 3. 7 (see Section 2. 6 for method details). Expression of seven copies of *BnaFLC* (*BnaFLC.A02*, *BnaFLC.A03a*, *BnaFLC.A03b*, *BnaFLC.A10*, *BnaFLC.C02*,

BnaFLC.C03b and *BnaFLC.C09*) was detected in all lines. Significant differences in expression were detected using the Bonferroni Multiple Comparisons test with an α value = 0.05.

At all timepoints analysed after 4-weeks vernalisation treatment; timepoints T0 at the end of vernalisation, T16 at 16 days after return to glasshouse conditions, T32 at 32 days after return to glasshouse conditions *BnaFLC.A02* (Figure 3. 7 A), *BnaFLC.A03a* (Figure 3. 7B), *BnaFLC.A03b* (Figure 3. 7 C) and *BnaFLC.A10* (Figure 3. 7 D) were expressed at significantly higher levels in the *BnaFRI* wild-type line (149) compared with Westar-10. *BnaFLC.C02*, *BnaFLC.C03b* and *BnaFLC.C09* genes in contrast exhibited similar expression levels in both lines before and after vernalisation, suggesting lower expression of A genome *BnaFLC* copies may contribute more strongly to the earliness of flowering in Westar-10 compared with the *BnaFRI* wild-type line.

In all cases *BnaFLC.A02*, *BnaFLC.A03a*, *BnaFLC.A03b* and *BnaFLC.A10* were stably down-regulated by four weeks of vernalisation treatment, while expression of *BnaFLC.C03b* exhibited re-activation phenotypes in response to four weeks vernalisation treatment. Although not significant, *BnaFLC.C02*, exhibited higher levels of expression in Westar-10 before vernalisation compared with the *BnaFRI* wild-type line, expression of this gene in Westar-10 was responsive to vernalisation whereas expression levels of *BnaFLC.C02* in the *BnaFRI* wild-type line remained relatively unchanged by vernalisation. In all lines, expression of *BnaFLC.C09*, measured as the combined expression of both *BnaFLC.C09a* and *BnaFLC.C09b* genes, increased with cold (Figure 3. 7 G).

No significant difference in *BnaFLC* gene expression were detected between the *BnaFRI* wild-type line and any *Bnafri* mutant line (lines 134, 135, 136, 137, and 138) before and after four weeks vernalisation (4-weeks vernalisation; T0, T16, T32). It can be concluded from this that nonsense mutations at *BnaFRI* have minimal downstream effect on *BnaFLC* expression. This poses questions about whether

Bnafri does indeed regulate *BnaFLC* in *B. napus*, or due to the polyploid nature of *B. napus* whether another gene is able to compensate for the loss of *BnaFRI* in the *Bnafri* mutants.

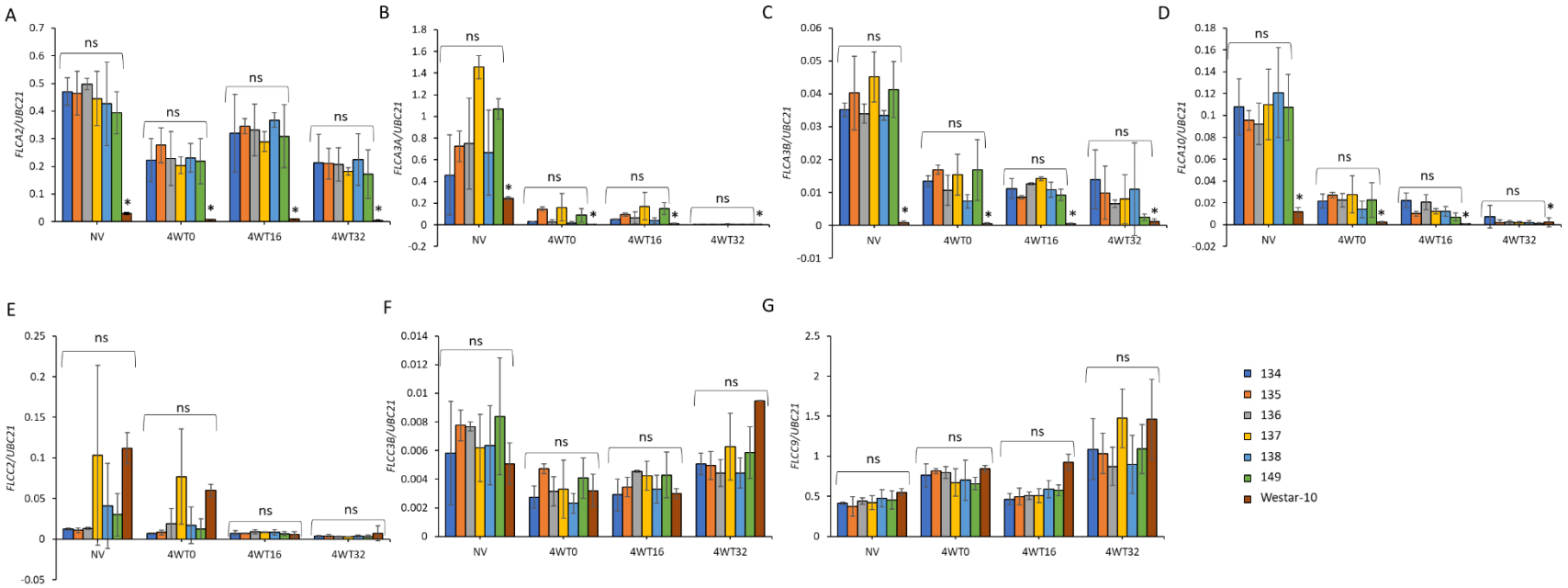


Figure 3. 7: Quantitative expression level of *BnaFLC* in *Bnafri* mutant lines. (A) *BnaFLC.A02*, (B) *BnaFLC.A03a*, (C) *BnaFLC.A03b*, (D) *BnaFLC.A10*, (E) *BnaFLC.C02*, (F) *BnaFLC.C03b*, (G) *BnaFLC.C09*; without vernalisation (NV); after 4-weeks vernalisation sampled in the cold (T0); 16 days after vernalisation, sampled in the glasshouse (T16); and 32 days after vernalisation, sampled in the glasshouse (T32). *BnaFLC* expression was normalized to the internal reference gene *UBC21*, mean values were calculated from three biological replicates, error bars denote 95% confidence interval. Significant differences were detected using the Bonferroni Multiple Comparisons test with an α value = 0.05; significant differences in expression are indicated with "*"; "n.s" = not significant.

3.2.3. Allelic variation at *BnaFRI.A03* has a weak and inconsistent effect on flowering time and *FLC* expression

My data in the previous sections of this Chapter suggest allelic variation at *BnaFRI* associates with crop type, however chemically induced non-functional variation at all four *BnaFRI* genes has no significant effect on flowering time or *BnaFLC* expression. Nevertheless, allelic variation at *BnaFRI.A03* has been suggested as the most likely candidate responsible for variation in flowering time in oilseed rape (Wang *et al*., 2011b). A QTL for flowering time variation was previously detected over *BnaFRI.A03* in a mapping population derived from the winter oilseed rape cultivar Tapidor and the Chinese semi-winter oilseed rape cultivar Ningyou-7 (Wang *et al*., 2011b). To further assess the role of this gene across *B. napus* crop types, I chose eight cultivars representing four crop types and four alleles of *BnaFRI.A03* (from the analysis in 3.2.1.; *BnaFRI.A03-a*, *BnaFRI.A03-b*, *BnaFRI.A03-c and BnaFRI.A03-e*). Two Chinese semi-winter oilseed rape (Ningyou-7 and Shengliyoucai), two spring oilseed rape (Stellar and Westar-10), two winter oilseed rape (Express-617 and Major) and two swede (Sensation New Zealand and Wilhelmsburger-Reform) cultivars were analysed in detail for DNA polymorphisms at the *BnaFRI.A03* gene, including 1.5kb upstream (5'UTR) and downstream (3'UTR) of the translation start and stop codons (the primers used in this analysis are listed in Table 2. 9). The results of this analysis are given in Figure 3. 3 A and Figure 3. 8.

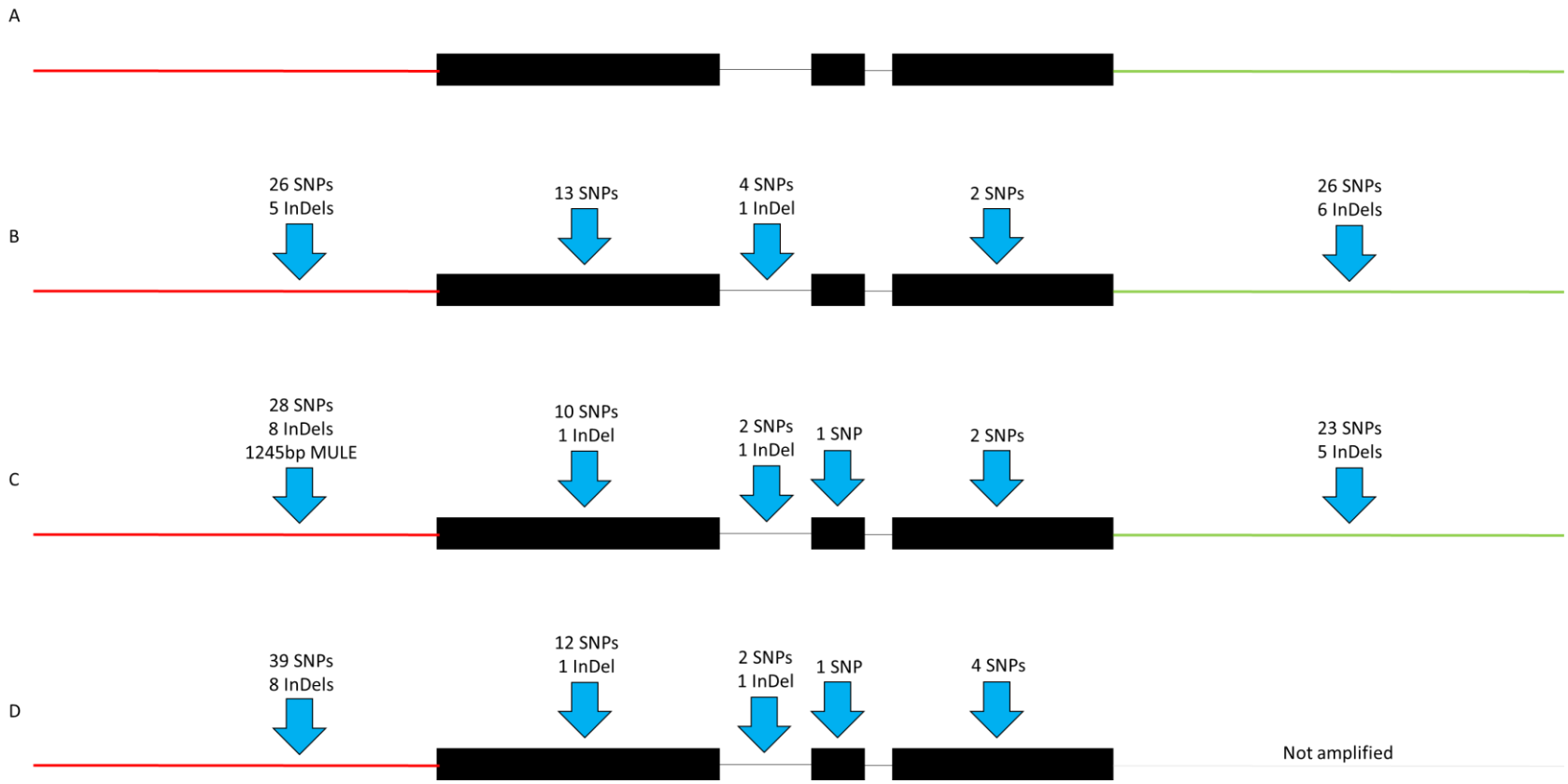


Figure 3. 8: Schematic illustration of the location of SNPs and InDels identified at *BnaFRI.A03*. Sequences were compared with the winter oilseed rape reference allele sequence *BnaFRI.A03-a* (GenBank accession JN936850) of (A) cultivars Express-617 and Major, (B) cultivars Ningyou-7 and Shengliyoucai representing allele *BnaFRI.A03-e*, (C) Stellar and Westar-10 representing allele *BnaFRI.A03-b*, (D) Sensation New Zealand and Wilhelmsburger; Reform representing allele *BnaFRI.A03-c*. Black boxes denote exons, black lines denote introns, red lines denote 5'UTR and green lines denote 3'UTR. The 3'UTR could not be amplified from cultivars Sensation New Zealand and Wilhelmsburger-Reform.

Sequence comparisons were made with Express-617 and Major, both of which carry the same allele of *BnaFRI.A03* denoted *BnaFRI.A03-a* in Section 3.2.1. Swede cultivars Sensation New Zealand and Wilhelmsburger; Reform both carried allele *BnaFRI.A03*-c and exhibited the greatest coding sequence diversity compared with Express-617 and Major (Figure 3. 8 D). 17 SNPs and a 3bp insertion result in 9 amino acid substitutions and 1 amino acid insertion distinguishing *BnaFRI.A03*-c from *BnaFRI.A03-a*. Sequence polymorphisms were also detected in intron 1, and in the promoter and terminator regions. Intron 2 was found to be conserved between *BnaFRI.A03*-c and *BnaFRI.A03-a*. Due to polymorphisms within primer regions, I was not able to amplify and sequence the 3'UTR region of *BnaFRI.A03-c* from Sensation New Zealand and Wilhelmsburger-Reform. Currently the 3'UTR sequence of *BnaFRI.A03-c* from these cultivars is unknown.

Ningyou-7 and Shengliyoucai carry allele *BnaFRI.A03-e* (Figure 3. 3 A) confirming the sequence polymorphisms previously published for Ningyou-7 by Wang *et al*., 2011b. Six amino acid substitutions are predicted for *BnaFRI.A03-e* (Figure 3. 8 B) that distinguish it from *BnaFRI.A03-a*. Sequence polymorphisms were also detected in intron 1 but again intron 2 sequence was conserved between the alleles. SNPs and InDel variation were detected in the promoter and terminator regions (Figure 3. 8 B).

Spring oilseed rape cultivars Stellar and Westar-10 exhibit the fewest number of coding SNPs (Figure 3. 8 C) compared with the winter oilseed rape cultivars Express-617 and Major and carry allele *BnaFRI.A03-b* (Figure 3. 3 A). Eight amino acid substitutions and a seven-amino acid deletion distinguish *BnaFRI.A03-b* from *BnaFRI.A03-a*. Polymorphisms were also detected in intron 1, and in the 3'UTR. Like *BnaFRI.A03-c and BnaFRI.A03-e* the intron 2 sequence is conserved between *BnaFRI.A03-b* and *BnaFRI.A03-a*.

I initially found amplification and sequencing of the 5'UTR from spring oilseed rape cultivars Stellar and Westar-10 challenging. After PCR optimisation, a 1,255bp A/T

rich insertion was detected approximately 863bp upstream from the translational start codon of *BnaFRI.A03-b* (Figure 3. 8 C). The insertion contained a 10bp target site duplication, a 232bp terminal inverted repeat and a 772bp internal sequence, not dissimilar in structure to a mutator-like transposon (Figure 3. 10, Robertson 1978, Ferguson and Jiang 2011). I detected a similar A/T rich insertion at *BraFRI.A03* in the early flowering *B. rapa* cultivar R-o-18 (of which we have an In-house reference sequence). An association between early flowering *B. napus* cultivars and the presence of a transposable element upstream of *BnaFRI.A03* remains to be explored.

B

CTTCATAATCGGGAAAATCGCATTTTAAACCATGAAGGTGTAAGAAGGTAGCACTTTAAACCTCAATGCTCAAAAACTCACACTTCAAACCCTC
AAACCAATTATATTAACACATTAAACCTTGAAACTAACTGACTTCTTTACCAAGCAAAAATGTGACCGGTACTGTTCATCCGTGAGTAAAAAGG
TATGACGTGTCAGTTGTTTTTTTTTTAATTAATAAATAAAAAAATATTTCAAATCACAAAAAAAATAAAAAATAAATGAAAAAGGTATAAAACTTAT
AAAAATAAATAAATAAAATTTGAAATTTATAAATAAATTAAAATAAATTTTAAAATAAAATTAAAATGAATAAAAATCTAAATTTATCAATATAATTA
AATAATTTAATACTTAAGAATAAAATTTAAAATAAATTAAAAAAATTAAAATAAAATTTAAAAACTAAAAAATAATAAAATAAACTTAAATATAATT
CAAAATTTTTTTAAAAAAATCAAAATAAGTAAATAAATCCCAAAATTTAATATACGCAACTAAAAATTTTCAAAATATTTTTTACATGTGAAAATT
TTCCGATCCCTTTTCCGATATATTAATAGTAACTATTTCTGAAATATGACCCTTGATGTTGATGGATTTTGGCATAAACCATTAAACATGATATATC
GGAAATCACCACCAATGTGTTATCTTCTTACATTGTCATGGTTTTCACTAGCATTTGTTGAATTTATTTACTTATTTTGATTTTTTATAATTATTTTA
AATTTATTTTTAAATTTTAATTTAATTTATTTTTATTTTTGGATTTATTTAAGAGTTTTTTTTAAAATTTGTTAGTTAATTAAAAATTTTATTAATTTAT
TTTTGATTTTTTAAATAAATCCAAAAAATAAATTTACTTTTAATTTATTAATAAATTTAAATTTTATTCATTTATTTTTAAAGTTTTATTTTTTTTGAGT
TTTTTTTGCATTTATCTCTGATTTTTAAATTAATTTTTTTTTTTTATTTCAAATATTTTTTTATTTATTAATTAAAAAAAAACAACTGACACGTCATAC
CTTTTTACCCACGGATGAACAGTACCGGTCACATTTTTGCTTGGTAAAGAAGTCAGTTAGTTTCAAGGTTTAATGTGTTAATATAATTGGTTTG
AGGGTTTGAAGTGTGAGTTTTTGAGCACTGAGGTTTAAAGTGCTACCTTCTTACACCTTCATGGTTTAAAATGCGATTTTCCCCTTCATAATC

Figure 3. 9: A mutator-like transposon was detected in spring oilseed rape cultivars Stellar and Westar-10. (A) A schematic illustration of the structure of the transposon identified, TSD=target site duplication, TIR=terminal inverted repeat. (B) The sequence of the transposon detected in Stellar and Westar-10, the target site duplication is highlighted with red text, the terminal inverted repeat is underlined and the internal sequence is represented in black text.

Due to the polymorphisms detected at *BnaFRI.A03*, I performed a transgenic complementation assay to determine the functional consequence of allelic variation at *BnaFRI.A03*. I designed nine constructs (see Figure 2. 1 and Section 2. 8 for specific method details) and transformed into the *A. thaliana* accession Col-0, which carries a non-function *fri*, but a functional *FLC* allele (Table 2. 3). After selection and copy number analysis eight independent, single insertion, T_2 transgenic lines were grown under controlled environment conditions (20°C, 16 hours light, 8 hours dark, 70% humidity) without vernalisation treatment. I scored their flowering time as

the number of days from sowing to when floral buds were visible. The results of this analysis are illustrated in Figure 3. 10.

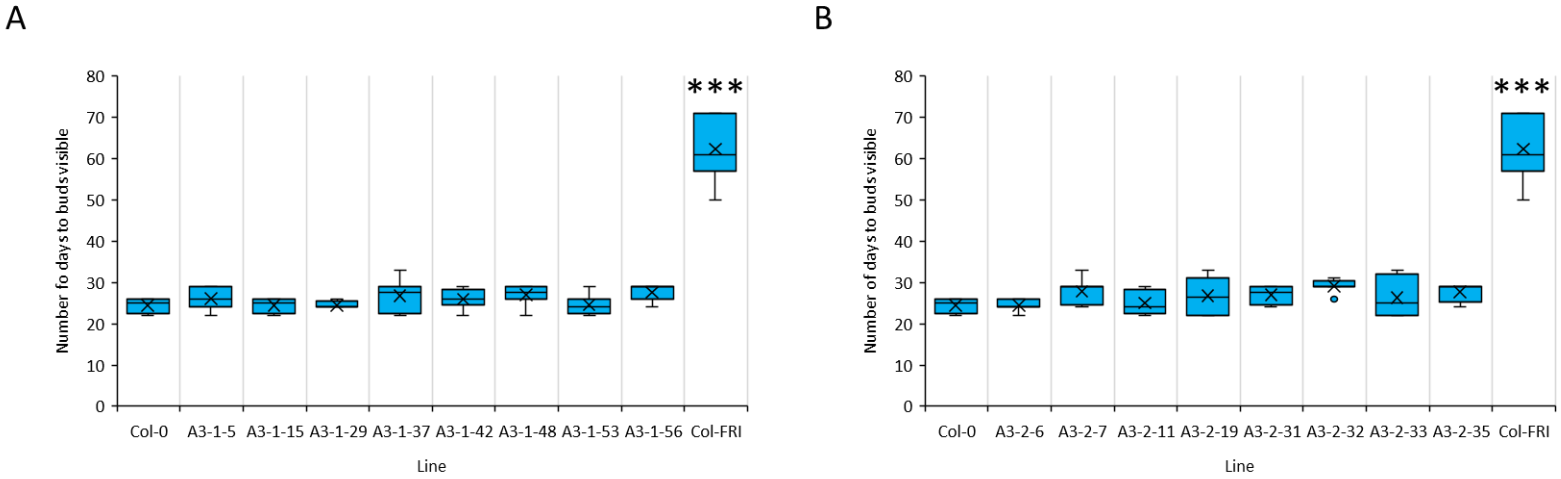


Figure 3. 10: Analysis of single insertion T_2 *BnaFRI.A03* transgenic lines. (A-I) The flowering times of eight independent T_2 lines for construct (A) A3-1, (B) A3-2, (C) A3-3, (D) A3-4, (E) A3-5, (F) A3-6, (G) A3-7, (H) A3-8, (I) A3-9, compared with *A. thaliana* ecotypes Col-0 and Col-*FRI*. Box-plots illustrate the median and range values calculated for eight plants per T_2 line. Significant differences in flowering times were assessed by ANOVA, "***" denote lines that exhibit significant differences compared with Col-0; α-value = 0.005. (J) Quantitative expression of endogenous *FLC* in Col-0, Col-*FRI*^{SF2} and pools of T_2 transgenic lines. Pools were made up of 40 seedlings from eight independent T_2 lines per construct and whole seedlings were sampled 12 days after stratification. *FLC* expression was normalised to the internal reference gene *UBC*, mean expression values were calculated from three biological replicates, error bars denote 95% confidence interval.

Figure 3.10 continued

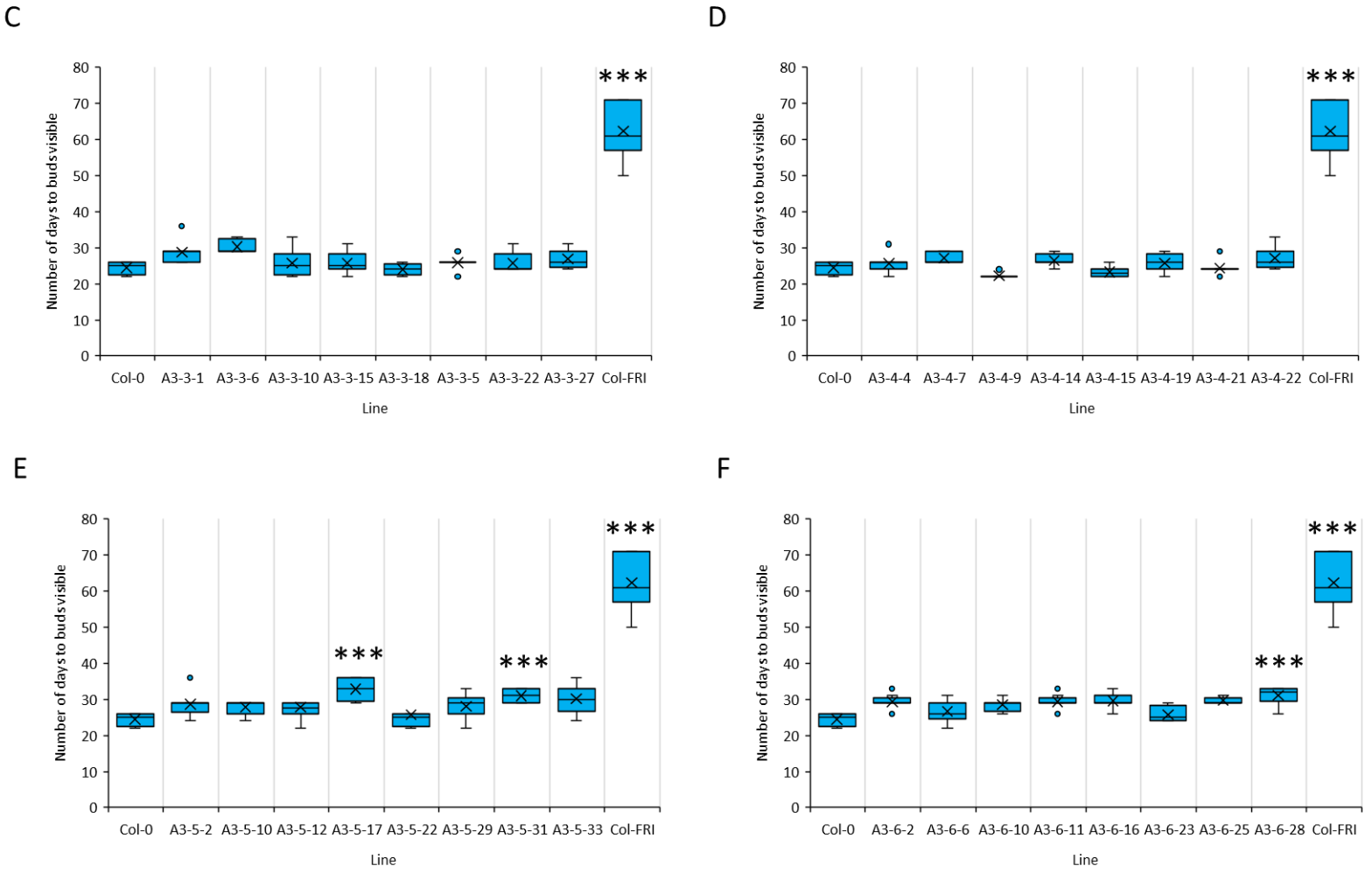
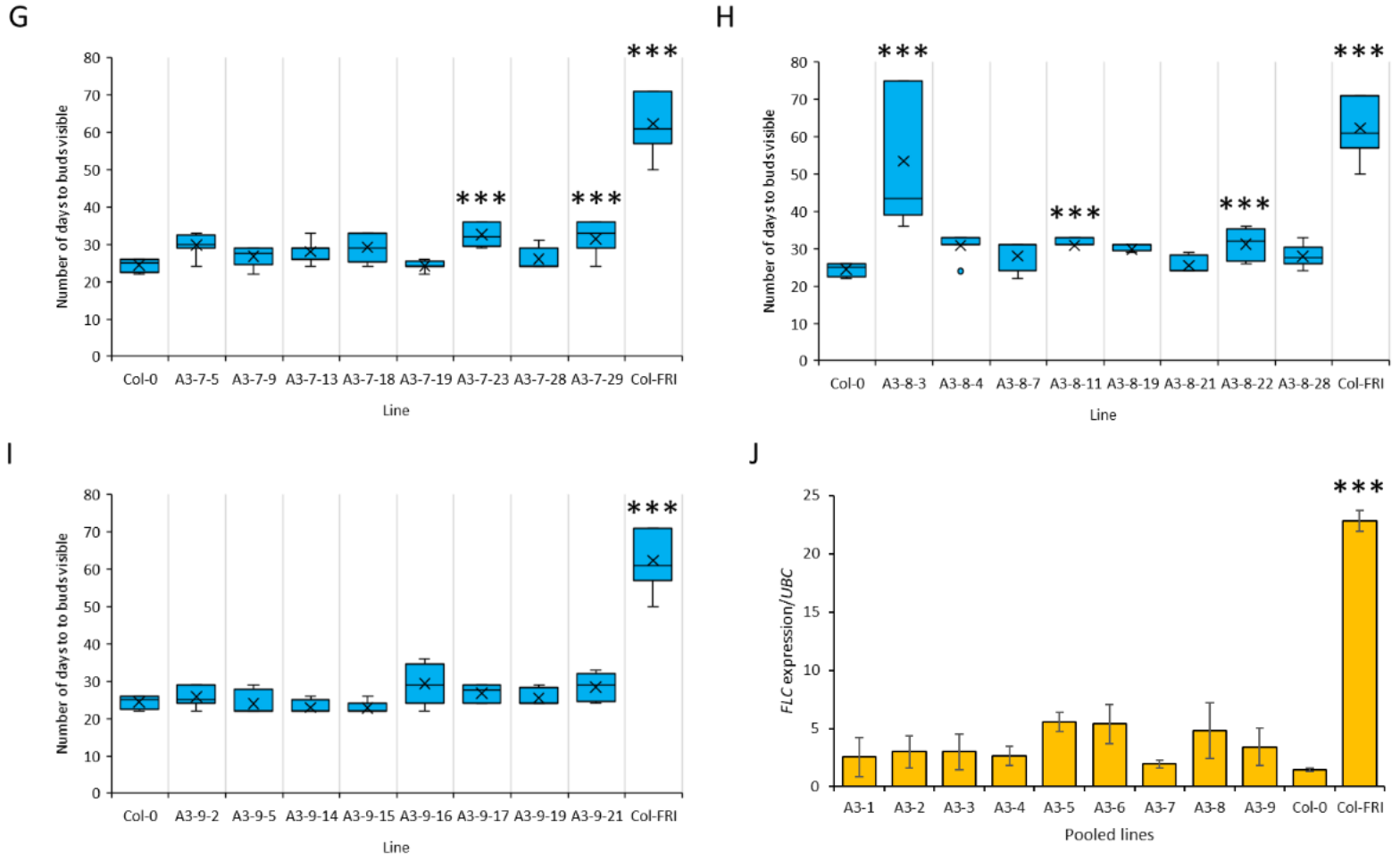


Figure 3.10 continued



The flowering times of all T_2 lines from each of the nine constructs were compared with wild-type Col-0 (*fri*) and Col-*FRI*^{SF2}. Eight independent T_2 lines (A3-5-17, A3-5-31, A3-6-28, A3-7-23, A3-7-29, A3-8-3, A3-8-11, and A3-8-22) flowered significantly ($p<0.05$,Bonferroni multiple comparisons) later than Col-0. However, no single *BnaFRI.A03* construct exhibited consistently significant different flowering time phenotypes compared with Col-0 across all independent T_2 lines analysed. With the exception of line A3-8-3, all T_2 lines flowered significantly earlier ($p<0.001$, ANOVA) than Col-*FRI*^{SF2} indicating that the *BnaFRI.A03* construct transformed into Col-0 did not complement the loss of function *fri* mutation.

I genotyped all T_2 transgenic lines for the presence of the *BnaFRI.A03* construct and the flowering times shown in Figure 3. 10 are from those carrying the *BnaFRI.A03* DNA. The constructs were designed to include 1275-1309bp of the 5'-UTR and partially overlapped a 190bp region of the gene upstream from *BnaFRI.A03*, a homologue of AT5G51070. The constructs also contained 1443-1455bp of the 3'-UTR. However, I was unable to confirm the quantitative expression of *BnaFRI.A03* in any of the T_2 lines. I had originally designed qPCR primers (Table 2. 7) to detected *BnaFRI.A03* expression in *B. napus*, but these amplified non-specifically in *A. thaliana*. Quantification of *BnaFRI.A03* will need to be determined to confirm that the construct is expressed in *A. thaliana*, however due to time constraints this was not done. Alternative primers, such as those published by Yi *et al*. (2018), could be used to complete this experiment.

Endogenous *FLC* expression in *A.* thaliana was elevated in pooled T_2 lines for all constructs (Figure 3. 10 J) compared with Col-0 wild-type however no significant difference in *FLC* expression was detected (Bonferroni Multiple Comparisons α-value = 0.05). The *BnaFRI.A03* constructs designed and transformed into Col-0 appear to retain some function in promoting endogenous *FLC* expression however, *FLC* levels were significantly lower to those measured in Col-*FRI*^{SF2} (Bonferroni Multiple Comparisons α-value = 0.05).

Constructs A3-6, A3-7, A3-8 and A3-9 were designed to test the functional significance of variation at the N-terminal and C-terminal domains of the predicted BnaFRI.A03 protein (Figure 2. 1). The open read frame (ORF) for these constructs were designed in two parts, one including exon 1 and the other including exons 2 and 3 (Figure 2. 1). Although no consistent pattern was observed between the constructs three independent T_2 lines containing construct A3-8 where the N-terminal domain of the BnaFRI.A03 protein was encoded by the spring oilseed rape sequence *(BnaFRI.A03-b*) and the C-terminal domain encoded by the winter sequence (*BnaFRI.A03-a*) exhibited significantly later flowering when compared with Col-0 (Bonferroni Multiple Comparisons α-value = 0.05). For example, I observed late flowering (55.5+/-19.5 days) for T_2 line A3-8-3, and three out of eight plants had not flowered within 75 days after sowing making their flowering times comparable with Col-FRI^{SF2}. In contrast the N-terminal domain encoded by the winter sequence (*BnaFRI.A03-a*) in combination with the C-terminal domain encoded by the spring sequence (*BnaFRI.A03-b*) in construct A3-9 exhibited no significant delay in flowering time compared with Col-0. As previously reported in *A. thaliana* (Risk *et al*., 2010) the C-terminal domain of BnaFRI.A03 protein appears to be more important for delaying flowering. However, as the effect was not consistent between all T_2 lines perhaps a positional effect of the construct insertion site rather than *BnaFRI.A03* gene function itself was responsible for the late flowering phenotype in lines A3-8.

3.3. Discussion

FRI is a major determinant of vernalisation requirement in *A. thaliana* (Shindo *et al*., 2005). Prior to the start of this project, four orthologues had been identified in *B. napus*, and expression of all four had been detected qualitatively in cultivars of oilseed rape (Wang *et al*., 2011b). Using genetic markers for *BnaFRI.A03* Wang *et al*., (2011) identified an association between flowering time and allelic variation at *BnaFRI.A03*. Due to the close evolutionary relationship between *A. thaliana* and *Brassica* crop species, I investigated orthologues of *FRI* for their role in the vernalisation requirement of *B. napus*. My work described in this Chapter in part supports a recent publication by Yi *et al*., (2018). Allelic variation was detected at *BnaFRI.A03*, *BnaFRI.A10*, and *BnaFRI.C03* that associated significantly with crop type. Expression of all four *BnaFRI* genes were detected in both winter and springoilseed rape. Induced mutations that are predicted to abolish BnaFRI protein function had no significant effect on flowering time or *BnaFLC* expression perhaps suggesting that *BnaFRI* may have an alternative function in *B. napus*. Further confirmation of this hypothesis was revealed by my transgenic analysis of *BnaFRI.A03*, a gene reported to underlie a QTL for flowering time (Wang *et al*., 2011b). The allelic variation present had no consistent significant effect on flowering time or *FLC* expression when expressed in *A. thaliana*which is in contrast to claims made by Yi *et al*., (2018). This leads me to the speculate that *BnaFRI* has evolved a subtly different function in *B. napus* which requires further exploration.

I detected multiple alleles at *BnaFRI.A03*, *BnaFRI.A10*, and *BnaFRI.C03* and *BnaFRI.C09*, some of which have previously been reported in *B. napus* and for homologous genes in *B. rapa* and *B. oleracea* (Wang *et al*., 2011b, Irwin *et al*., 2012, Fadina *et al*., 2013, Fadina and Khavkin *et al*., 2014). Conservation of *BnaFRI* sequence between species, and the maintenance of multiple alleles within the *B. napus* population, suggests *FRI* orthologues perform a similar function within all three species. As has been reported (Yi *et al*., 2018) allelic variation at *BnaFRI* associated significantly with crop type however my mutation and transgenic approaches suggests this allelic variation does not strongly control flowering time.

In *A. thaliana*, non-functional variation at *FRI* is associated with early flowering (Shindo *et al*., 2005). Non-synonymous polymorphisms were detected at all four *BnaFRI* gene however, of the 380 *BnaFRI* genes sequenced from 95 *B. napus* cultivars only one was predicted to encode a non-functional BnaFRI protein; *BnaFRI.C09b*. This allele of *BnaFRI.C09* is present in only one *B. napus* cultivar called Jaune a Collet Vert, a late flowering swede cultivar, and therefore an unlikely candidate gene for variation in flowering time. The remaining 94 cultivars carried the same allele, *BnaFRI.C09-a*, predicted to encode a functional BnaFRI.C09 protein. Lack of allelic variation at *BnaFRI.C09* agrees with recently published data (Yi *et al*., 2018) indicating this gene might perform an essential function in *B. napus* but is not responsible for variation in flowering time. According to the *B. napus* genome (Darmor-*bzh*, Chalhoub *et al*., 2014, plants.ensembl.org) *BnaFRI.C09* is found within an island of non-coding sequence on chromosome C09. The nearest annotated genes are found approximately 18-20kb up- and downstream. This contrasts with *BnaFRI.A03*, *BnaFRI.A10* and *BnaFRI.C03* where the nearest annotated genes are found within approximately 0-8kb, perhaps suggesting genomic location of *BnaFRI* may influence allelic diversity.

Due to the apparent lack of non-functional variation at *BnaFRI* I investigated the phenotypic and molecular consequence of inducing premature stop codons at all four copies of *BnaFRI* using winter oilseed rape breeding lines kindly provided by Bayer CropScience, NV, Belgium. Loss of BnaFRI protein had no significant effect on flowering time under the growth conditions tested. It seems in nature functional BnaFRI protein has been maintained throughout evolution, however loss of BnaFRI has no obvious phenotypic effect in *B. napus*. The near absence of non-functional variation at *BnaFRI* is similar to reports in other Brassicaceae such as *A. lyrata* (Kuittinen *et al*., 2008), *C. rubella* (Guo *et al*.,2012) and *B. oleracea* (Irwin *et al*., 2012) suggesting some *FRI* function is indispensable in many species of the Brassicaceae, but optional in *A. thaliana* (Kuittinen *et al*., 2008). However, from the data described in this Chapter *FRI* in *B. napus* is unlikely to be responsible for loss of vernalisation requirement. It is possible that *FRI* acquired a novel function during the evolution of *A. thaliana*. If so, what is its function in *B. napus*? Is it involved in the floral regulatory pathway or does it control another trait?

Loss of BnaFRI protein had no significant effect on *BnaFLC* expression. The expression levels of *BnaFLC* genes were not reduced in the absence of BnaFRI suggesting loci independent to *BnaFRI* contribute to the low *BnaFLC* expression levels detected in spring oilseed rape cultivars. We cannot exclude the possibility that another gene, maybe multiple *BnaFRI-like* genes or another previously unidentified homologue of *BnaFRI*, could compensate for the loss of function at BnaFRI in these lines. As *BnaFLC* is not the target of BnaFRI protein function, a comparative transcriptome analysis of *B. napus* with and without BnaFRI protein may reveal which genes, if any, are transcriptionally regulated by BnaFRI.

BnaFRI.A03 exhibited the most allelic diversity, and I identified nine alleleswithin the OREGIN DFFS population. No mutation detected at *BnaFRI.A03* exhibited parallels with those reported in *A. thaliana* (Shindo *et al*., 2005) therefore the functional consequence of these mutations could not be predicted. A mutator-like transposon was detected upstream of *BnaFRI.A03* in spring oilseed rape cultivars. It is interesting to note that the early flowering *B. rapa* cultivar R-o-18 also carries a mutator-like transposon within a region upstream of *BraFRI.A03* but carries an alternative allele that is similar in sequence to *BnaFRI.A03-c*. Although the presence of the transposon had no significant effect on the quantitative expression of *BnaFRI.A03*, I measured expression at only one timepoint. In *A. thaliana FRI* expression is essential for promoting *FLC* expression in developing embryos (Caroline Dean Laboratory, unpublished). Could the presence of the transposon affect the spatial and temporal expression of *BnaFRI.A03*? A possible correlation between presence of a transposon at *BnaFRI.A03* and crop type was not investigated but is worth exploring.

My transgenic analysis of *BnaFRI.A03* provides further evidence that *BnaFRI* function might not be conserved between *B. napus* and *A. thaliana*. Previous reports have shown that expression of *FRI* orthologues from *A. lyrata* and *B. oleracea* can significantly delay flowering (Kuittinen *et al*., 2008, Irwin *et al*., 2012). However, in both experiments *FRI* expression was controlled by non-native promoter sequences.

In the present study constructs of *BnaFRI.A03* were generated which included the native promoter and terminator sequence from Chinese semi-winter and winter oilseed rape cultivars. Single insertion, T_2 transgenic lines carrying *BnaFRI.A03* allelic variants did not flower consistently differently from Col-0 (*fri*) and in most cases significantly earlier than Col-*FRI*^{SF2}. Expression of *FLC* was not significantly affected by expression of the *BnaFRI.A0*3 constructs indicating *BnaFRI.A03* could not fully complement the loss of function *fri* mutation in Col-0. These data must be interpreted with caution however as the expression of *BnaFRI.A03* was not quantified. Yi *et al*., (2018) report a significant delay in flowering time (between 3.9 and 23.0 days) in their *BnaFRI.A03* transgenic lines compared with Col-0. Data from only two transgenic lines per construct were presented and the number of transgene inserts in each line was not determined. In my study, eight independent transgenic lines carrying a single insertion of *BnaFRI.A03* were analysed for flowering time variation and *FLC* expression and nine different constructs were tested. Variation in flowering time between transgenic lines was detected, and in some cases conferred significantly later flowering than wild-type Col-0 (*fri*). This effect was not consistent between independent T_2 lines carrying the same *BnaFRI.A03* construct possibly suggesting that the position of transgene insertion site can influence flowering time and interpretation of results.

Perhaps most interesting is to compare the outcome of the transgenic analysis of *BnaFRI.A03* presented in this Chapter (and published by Yi *et al*., 2018) with the transgenic analysis of *BolFRI.C03* published by Irwin *et al*., (2012). *BolFRI.C03,* which exhibits 97-99% sequence identity with *BnaFRI.A03* (Fadina & Khavkin 2014) conferred late flowering similar to Col-*FRI*^{SF2} in *A. thaliana*. However, the *BolFRI.C03* constructs were expressed under the *FRI* promoter sequence from *A. thaliana.* Due to their high sequence similarity one would hypothesise that both *BnaFRI.A03* and *BolFRI.C03* genes would perform highly similar functions however, when expressed under its native promoter *BnaFRI.A03* conferred early flowering. Functional differences may have arisen through sequence variation at the promoter regions of *FRI*, a possible explanation why *BnaFRI.A03* did not confer late flowering in *A. thaliana*. Are there temporal or spatial expression differences

between *FRI* in *A. thaliana* and *FRI* orthologues in *Brassica*? This remains to be tested.

The results described in this Chapter suggests allelic variation at BnaFRI has not significant consequence on flowering time or *FLC* expression. The reported association between *BnaFRI.A03* allelic variation and flowering time in *B. napus* is therefore curious (Wang *et al*. 2011). Perhaps another gene that is genetically closely linked to *BnaFRI.A03* confers variation in flowering time detected (Wang *et al*. 2011). Many questions remain with regards to the function of *FRI* orthologues in *B. napus*. Striking differences are identified between *FRI* function in *A. thaliana* and other members of the Brassicaceae (Kuittinen *et al*., 2008, Irwin *et al*., 2012, Yi *et al*., 2018). Does *BnaFRI* regulate *BnaFLC*? This work suggests not, however research in the Judith Irwin Laboratory is currently underway to further address this question.

Chapter 4: Variation at *BnaFLC.A10* associates with flowering time in *Brassica napus*

4.1. Introduction

I demonstrated in Chapter 3 of this Thesis that orthologues of *FRI* do not strongly influence the vernalisation requirement and flowering time in *B. napus.* The aim of the work in this Chapter was to identify the genes that are important for vernalisation requirement. Using a genetic mapping approach, known as Associative Transcriptomics (Harper *et al*., 2012) which uses mRNA-seq data to identify SNPs and gene expression differences that associate with a trait of interest, I analysed the flowering times of 88 diverse *B. napus* cultivars, including spring oilseed rape, Chinese semi-winter oilseed rape, winter oilseed rape, swede, kale, winter fodder, winter vegetable, exotic, and synthetic crop types under two vernalisation treatments for marker-trait associations.

Genome-wide association studies (GWAS) have been used extensively in human genetics studies, and in plants such as maize, rice, and *A. thaliana* (Atwell *et al*., 2010, Tian *et al*., 2011, Zhao *et al*., 2011) to identify causative mutations that are in linkage disequilibrium with a trait. Large and diverse populations are analysed in GWAS to identify historical recombination between loci (Harper *et al*., 2012). This is in contrast to QTL mapping whereby mapping populations are most commonly generated from biparental crosses to identify the causative variation present between the parental lines. Use of diverse germplasm for GWAS can give a more accurate representation of the variation within species and have revealed candidate genes that are under strong selection (Harper *et al*., 2012, Lu *et al*., 2014b, Havlickova *et al*., 2017).

In 2010 an extensive GWAS of *A. thaliana* was published by Atwell *et al*., (2010). Over one hundred phenotypes were measured in two diverse populations and correlated against genotypic data using a mixed linear model approach to correct for

population structure. Analysis of multiple flowering phenotypes identified *FRI*, *FLC*, *SVP* and *DOG1* as strong candidates for flowering time across the diverse collection of accessions. However flowering time was severely confounded by geographic area (Atwell *et al*., 2010). It was clear from this, that population structure must be considered when conducting GWAS, especially when analysing flowering time. False discoveries due to population structure is often a concern when identifying genetic polymorphisms that are important for a trait of interest (Zhao *et al*., 2007, Zhang *et al*., 2010). General Linear Models (GLM) were initially used to correct for population structure, but now Mixed Linear Models (MLM), that incorporate kinship among individuals as a co-variate, is the preferred statistical approach (Yu *et al*., 2006, Zhang *et al*., 2010).

Since the publication of the *B. napus* genome (Chalhoub *et al*., 2014) several GWA studies have investigated the flowering time variation present within the species (Schiessl *et al*., 2015, Xu *et al*., 2015, Raman *et al*., 2016, Wang *et al*., 2016a, Zhou *et al*., 2017). Population sizes for these analyses ranged from 158 (Schiessl *et al*., 2015) to 523 (Xu *et al*., 2015) individuals and plants were grown under an array of field and experimental conditions. GWAS in *B. napus* has identified many promising candidate genes for variation in flowering time (Schiessl *et al*., 2015, Xu *et al*., 2015, Raman *et al*., 2016, Wang *et al*., 2016a, Zhou *et al*., 2017), however the underlying associating variation is highly dependent on the population analysed and the environmental conditions tested. For example, Xu *et al*., (2015) analysed a population of 523 *B. napus* cultivars. This population was predominated by semi-winter oilseed rape cultivars and included some spring and winter oilseed rape, making up only 8.4% and 7.6% of the population respectively. Over 40 SNP markers were detected in eight field environments. Candidate genes for flowering time identified by Xu *et al*., (2015) included orthologues of *FLC, CO*, in addition to *FY*, *VIN3*, *FRL1*, and *SVP*. In contrast, and unsurprisingly when Schiessl *et al*., (2015) analysed 158 European winter *B. napus* inbred lines different candidate genes were detected. Associations between orthologues of *CO* and *FLC* were not detected while interesting candidates included orthologues of *FRI*, *FUL*, and *TFL1*.

For GWAS a large number of ordered molecular markers are required and, until fairly recently, development of diverse genetic resources has been mainly restricted to model plant species with published genome sequences such as *A. thaliana*, maize and rice (Atwell *et al*., 2010, Tian *et al*., 2011, Zhao *et al*., 2011). However, computational technologies are available for analysis of large SNP datasets in species without reference genome sequences. Prior to the publication of the *B. napus* genome (Chalhoub *et al*., 2014) many datasets had been generated using transcriptome sequencing (mRNA-seq) for SNP discovery (Trick *et al*., 2009), linkage mapping and genome characterisation (Bancroft *et al*., 2011) and transcript quantification (Higgins *et al*., 2012). One such mRNA-seq dataset formed the basis for an association genetics technique known as Associative Transcriptomics (Harper *et al*., 2012). This technique (Harper *et al*., 2012) has been used to identify both SNP markers and gene expression markers (GEM) that associate with variation for a trait of interest. Utilisation of mRNA-seq from a diverse panel of *B. napus* cultivars not only provides a large dataset of coding SNP markers, but sequence read depth values (Reads Per Kilobase per Million aligned read or RPKM values) of genes that can be used to infer a measure for gene expression.

Before the Associative Transcriptomics pipeline could be used, a hypothetical *B. napus* genome was assembled into ordered “pseudomolecules" based on the genomes of the progenitor species *B. rapa* and *B. oleracea* (Harper *et al*., 2012). *Brassica* expressed sequence tags, or unigenes (Trick *et al*., 2009), were positioned onto these pseudomolecules to identify the genomic positions of expressed genes and to generate a gene expression-based map of the *B. napus* genome. Orthologous of *A. thaliana* gene models for each unigene were identified by sequence similarity and each unigene was assigned an orthologous Arabidopsis Genome Initiative (AGI) code. Illumina mRNA-seq was performed on 84 cultivars and, using the pseudomolecule as a reference, the data generated were assayed for SNP markers and transcript abundance. It was first published as a proof of concept method to identify marker-trait associations for seed glucosinolate content in 84 *B. napus* cultivars (Harper *et al*., 2012) but has since been used to study seed characteristics (Lu *et al*., 2014b, Havlickova *et al*., 2018) and anion status in *B. napus* (Koprivova

et al 2014); stem strength characteristics in bread wheat (Miller *et al*., 2016); and disease resistance markers for ash dieback in European ash (Harper *et al* 2016).

GWA studies in *A. thaliana*, maize, and rice have involved the analysis of large diversity sets (Atwell *et al*., 2010, Tian *et al*., 2011, Zhao *et al*., 2011). The small size of the diversity panel used for Associative Transcriptomics (84 cultivars; Harper *et al*., 2012) may mean that complex traits such as flowering time, that is known to be controlled by many genes with small effects, cannot be analysed accurately or successfully (Spencer *et al*., 2009). Since its publication (Harper *et al*., 2012) the Associative Transcriptomics dataset for *B. napus* has been expanded to 101 cultivars by Lu *et al*., (2014b) (the set used in the present study) and more recently to 383 cultivars by Havlickova *et al*., (2018). Within the 101-cultivar set, 144,131 SNP markers were mapped onto the *B. napus* pseudomolecule reference (Lu *et al*., 2014b). Due to high sequence similarity between the A and C genomes, only 10.6% of all SNP markers could be mapped to a specific genome location, the genome location of the majority of SNP calls could not be determined and therefore were denoted hemi-SNPs (Trick *et al*., 2009, Harper *et al* 2012, Lu *et al*., 2014b). The GEMs however could be more easily mapped to specific genomes and a total of 49,599 and 50,935 GEMs could be assigned to the A and C genomes respectively. Based on SNP marker data, all *B. napus* cultivars could be divided into two clusters (Harper *et al*., 2012, Lu *et al*., 2014b) and overall mean Linkage Disequilibrium (LD) was relatively low at r^2=0.0209 across the whole genome (Lu *et al*., 2014b).

It is clear from the results in Chapter 3 that orthologues of *FRI* do not contribute strongly to variation in vernalisation requirement and flowering time in *B. napus* therefore I performed an Associative Transcriptomics analysis on 88 *B. napus* cultivars that exhibited diverse flowering times. The aim of this work was to identify genetic variation that is important for controlling vernalisation requirement and response in *B. napus*. Variation at an orthologue of *FLC*, *BnaFLCA10*, was found to be important for flowering time. I detected a significant association between *BnaFLC.A10* expression level and flowering time, and cis-variation at *BnaFLC.A10*, not variation at *BnaFRI*, was found to be important for regulating its expression. I

detected polymorphisms within cis-regulatory regions, known from *A. thaliana* to be important for expression and the epigenetic silencing of *FLC*, were detected at *BnaFLC.A10* (Coustham *et al*., 2012, Li *et al*., 2014, Li *et al*., 2015, Hawkes *et al*., 2016, Questa *et al*., 2016). This work provides evidence that changes at cis-regulatory regions of *BnaFLC.A10* have been selected by Breeders to alter flowering time, much like they have been selected by nature in *A. thaliana*.

4.2. Results

4.2.1. *Brassica napus* exhibits a diverse range of flowering times

Two flowering time traits (the number of days to when buds were visible from above (BBCH51); and the number of days to when the first flower opened (BBCH60)) of 88 *B. napus* cultivars were measured under two vernalisation treatments; after a 6-week vernalisation treatment (VERN); and without vernalisation (NVERN) as described in Section 2.9.1. Scoring of flowering time commenced at the end of the vernalisation treatment; 21 and 63 days after sowing for NVERN and VERN respectively.

Temperature was recorded at hourly intervals during the experiment and from this, the average daily temperatures were calculated (Figure 4. 1). During the experiment, temperatures fluctuated by an average of 24.47°C daily. The lowest recorded temperature was 1.7°C and the highest 47.0°C. During the experiment, thirteen days had average daily temperatures below 15°C possibly suggesting plants received an additional thirteen days of vernalising temperatures (Duncan *et al*., 2015).

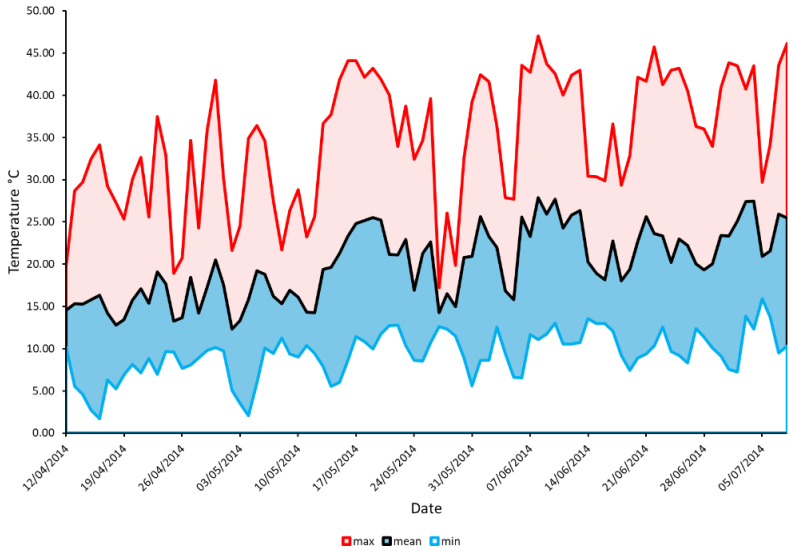


Figure 4. 1: Temperature profile from within the poly-tunnel. Temperature was recorded at hourly intervals by Tinytag and the daily average (black), maximum (red) and minimum (blue) recordings are plotted.

A broad range of flowering times were measured under both vernalisation treatments. The mean flowering times calculated from three plants per cultivar arelisted in Table 4. 2. Flowering time was not normally distributed ($p<0.001$, Shapiro-Wilke test) and transformations did not improve the distribution. Significant differences in flowering times were detected between the vernalisation treatments ($p<0.001$, Kruskal Wallis ANOVA), and flowering times were significantly correlated between treatments (Table 4. 1). BBCH51 and BBCH60 measurements were also significantly correlated (Table 4. 1) therefore, all further analyses were performed on the BBCH51 data only. Hereafter when referring to flowering time this will denote BBCH51 measurements only.

Table 4. 1: Spearman's rank correlation coefficient and significance values were calculated between the flowering time measurements under NVERN and VERN conditions, and for BBCH51 and BBCH60.

Spearman's rank correlation	**NVERN BBCH51**	**NVERN BBCH60**	**VERN BBCH51**	**VERN BBCH60**
NVERN BBCH51	1.000 p<0.001			
NVERN BBCH60	0.987 p<0.001	1.000 p<0.001		
VERN BBCH51	0.844 p<0.001	0.833 p<0.001	1.000 p<0.001	
VERN BBCH60	0.845 p<0.001	0.836 p<0.001	0.986 p<0.001	1.000 p<0.001

Table 4. 2: Flowering time measurements recorded for 88 *B. napus* cultivars. For each vernalisation treatment (NVERN and VERN) the mean number of days (+/- 95% confidence interval) to buds visible (BBCH51) and first flower open (BBCH60) were calculated from three replicate plants per cultivar and measurement of flowering time commenced after the vernalisation treatment had ended. For each cultivar, their cultivar code for the associative transcriptomics analysis and crop type are included. DNF= did not flower

				Treatment NVERN		**Treatment VERN**	
Cultivar Name	**Cultivar (DNA) Number**	**Cultivar code**	**Crop type**	**Mean number of days to BBCH51**	**Mean number of days to BBCH60**	**Mean number of days to BBCH51**	**Mean number of days to BBCH60**
Abukuma Natane	11	IB02_AbuN	Winter OSR	70.00 +/- 2.99	87.00 +/- 4.08	32.67 +/- 1.73	51.00 +/- 0.00
Altasweet	36	D3_ALT	Swede	DNF +/- 0.00	DNF +/- 0.00	DNF +/- 0.00	DNF +/- 0.00
Amber X Commmanche DH Line	62	D4_AMBxCOM	Winter OSR	58.67 +/- 5.10	74.33 +/- 4.28	37.67 +/- 7.53	55.00 +/- 8.98
Apex	14	D6_APE	Winter OSR	70.00 +/- 4.93	89.67 +/- 8.49	35.00 +/- 13.91	53.67 +/- 14.06
Apex-93_5 X Ginyou_3 DH Line	n/a	D7_APExGIN	Winter OSR	25.67 +/- 0.65	43.00 +/- 1.13	28.33 +/- 12.46	45.00 +/- 12.75
Baltia	53	D10_BAL	Winter OSR	102.67 +/- 3.97	129.67 +/- 15.69	35.33 +/- 1.73	53.00 +/- 2.26
Bienvenu DH4	48	A11_Bie	Winter OSR	76.00 +/- 11.48	105.67 +/- 39.60	33.00 +/- 5.19	51.67 +/- 4.57
Bolko	59	D14_Bol	Winter OSR	138.00 +/- 3.92	DNF +/- 0.00	65.67 +/- 4.71	86.33 +/- 6.23
Brauner_Schnittkohl	9	IB15_BraS	Siberian kale	132.00 +/- 15.68	DNF +/- 0.00	46.33 +/- 2.36	60.67 +/- 0.65
Bravour (NK)	15	IB16_Bra	Winter OSR	70.67 +/- 0.65	87.67 +/- 2.61	27.67 +/- 0.65	46.33 +/- 0.65
Bronowski	20	A18_Bro	Spring OSR	7.00 +/- 0.00	25.00 +/- 1.13	26.33 +/- 2.36	44.67 +/- 2.36
Cabernet	81	IB21_Cab	Winter OSR	84.00 +/- 1.96	116.33 +/- 28.12	27.67 +/- 1.73	46.67 +/- 0.65

				Treatment NVERN		Treatment VERN	
Cultivar Name	**Cultivar (DNA) Number**	**Cultivar code**	**Crop type**	**Mean number of days to BBCH51**	**Mean number of days to BBCH60**	**Mean number of days to BBCH51**	**Mean number of days to BBCH60**
Cabriolet	82	IB22_Cab	Winter OSR	37.33 +/- 9.15	65.00 +/- 2.26	25.00 +/- 1.96	44.33 +/- 1.73
Canard	45	IB23_Can	Winter fodder	103.00 +/- 27.81	120.67 +/- 23.56	48.33 +/- 1.31	63.00 +/- 2.99
Canberra X Courage DH Line	83	IB24_CanxCou	Winter OSR	64.33 +/- 1.31	78.33 +/- 1.31	34.67 +/- 7.95	49.33 +/- 6.23
Capitol	13	A25_Cap	Winter OSR	70.33 +/- 3.27	87.67 +/- 4.71	33.00 +/- 2.99	49.00 +/- 2.99
Capitol X Mohican DH Line	60	D26_CAPxMOH	Winter OSR	65.33 +/- 2.85	81.33 +/- 1.73	30.67 +/- 2.85	49.33 +/- 1.73
Castille	84	IB28_Cas	Winter OSR	29.00 +/- 4.08	51.00 +/- 5.66	27.00 +/- 2.26	44. 33 +/- 1.73
Catana	85	IB29_Cat	Winter OSR	52.00 +/- 17.64	69.00 +/- 11.81	29.00 +/- 1.96	46.33 +/- 0.65
Ceska Krajova	24	D31_CES	Spring OSR	22.67 +/- 0.65	37.00 +/- 1.13	24.00 +/- 1.13	40.00 +/- 1.13
Chuanyou 2	6	A34_Chu	Semi-winter OSR	27.67 +/- 3.64	44.67 +/- 4.71	20.67 +/- 2.85	37.33 +/- 2.85
Columbus X Nickel DH Line	86	IB37_ColxNic	Winter OSR	102.00 +/- 5.19	120.33 +/- 5.35	38.33 +/- 8.49	56.33 +/- 6.53
Coriander	61	D39_COR	Winter OSR	64.67 +/- 5.35	82.00 +/- 4.53	31.33 +/- 2.61	49.67 +/- 4.28
Couve_nabica	42	R_IB40_CouN	Exotic	85.00 +/- 2.26	98.67 +/- 0.65	32.00 +/- 2.99	48.00 +/- 1.96
Darmor	51	A42_Dar	Winter OSR	138.33 +/- 3.27	DNF +/- 0.00	50.00 +/- 5.66	66.00 +/- 7.92
Dimension	87	IB45_Dim	Winter OSR	81.33 +/- 3.97	114.67 +/- 30.37	32.00 +/- 4.93	48.67 +/- 3.46
Dippes	63	D46_DIP	Winter OSR	130.33 +/- 18.95	137.33 +/- 15.03	33.00 +/- 1.96	49.33 +/- 0.65
Drakkar	33	R_A48_Dra	Spring OSR	4.33 +/- 2.61	17.00 +/- 5.66	19.00 +/- 5.99	36.33 +/- 2.85

				Treatment NVERN		Treatment VERN	
Cultivar Name	**Cultivar (DNA) Number**	**Cultivar code**	**Crop type**	**Mean number of days to BBCH51**	**Mean number of days to BBCH60**	**Mean number of days to BBCH51**	**Mean number of days to BBCH60**
Duplo	31	IB49_Dup	Spring OSR	6.33 +/- 1.31	22.00 +/- 2.26	25.33 +/- 0.65	43.00 +/- 0.00
Dwarf Essex	46	IB50_DwaE	Winter fodder	78.33 +/- 2.85	90.67 +/- 1.73	36.67 +/- 10.51	54.67 +/- 14.15
Erglu	32	IB53_Erg	Spring OSR	7.67 +/- 0.65	29.67 +/- 1.73	26.67 +/- 1.31	43.00 +/- 1.13
Eurol	64	D55_EUR	Winter OSR	130.00 +/- 12.60	DNF +/- 9.80	41.67 +/- 5.58	57.00 +/- 5.88
Excalibur	88	IB57_Exc	Winter OSR	74.00 +/- 9.26	90.00 +/- 13.05	32.67 +/- 1.73	49.33 +/- 1.73
Expert	16	IB58_Exp	Winter OSR	134.00 +/- 11.76	DNF +/- 0.00	34.00 +/- 12.85	51.00 +/- 11.81
Flash	89	IB60_Fla	Winter OSR	78.00 +/- 9.80	82.00 +/- 31.01	29.67 +/- 4.28	48.00 +/- 3.92
Hanna	25	D69_HAN	Spring OSR	3.00 +/- 0.00	14.33 +/- 2.36	22.00 +/- 1.96	37.00 +/- 0.00
Hansen X Gaspard DH Line	90	IB70_HanxGas	Winter OSR	91.67 +/- 5.70	108.67 +/- 6.82	35.00 +/- 8.16	52.00 +/- 7.92
Huron X Navajo DH Line	66	D74_HURxNAV	Winter OSR	95.67 +/- 4.71	109.33 +/- 5.35	29.33 +/- 1.73	46.33 +/- 0.65
Inca X Contact DH Line	67	D75_INCxCON	Winter OSR	86.67 +/- 3.97	101.67 +/- 3.64	30.00 +/- 1.96	47.33 +/- 0.65
Jaune a Collet Vert	41	R_IB78_JauCV	Swede	DNF +/- 0.00	DNF +/- 0.00	61.33 +/- 3.64	80.67 +/- 4.28
Jet-neuf	69	D79_JETN	Winter OSR	116.67 +/- 24.59	129.33 +/- 17.54	43.67 +/- 6.43	61.00 +/- 8.98
Karoo-057	26	D83_KAR	Spring OSR	-4.00 +/- 0.00	3.00 +/- 0.00	2.00 +/- 0.00	10.33 +/- 0.65
Lembkes Malchower	91	IB86_LemM	Winter OSR	117.67 +/- 24.38	135.33 +/- 18.95	40.00 +/- 4.93	56.67 +/- 1.73
Licrown X Express DH Line	92	IB91_LicxExp	Winter OSR	70.67 +/- 7.70	87.67 +/- 10.51	31.67 +/- 7.53	48.33 +/- 5.58

				Treatment NVERN		Treatment VERN	
Cultivar Name	**Cultivar (DNA) Number**	**Cultivar code**	**Crop type**	**Mean number of days to BBCH51**	**Mean number of days to BBCH60**	**Mean number of days to BBCH51**	**Mean number of days to BBCH60**
Madrigal X Recital DH Line	93	IB99_MadxRec	Winter OSR	69.00 +/- 0.00	87.00 +/- 2.26	27.33 +/- 0.65	46.00 +/- 1.13
Major	71	IB100_Maj	Winter OSR	97.67 +/- 4.71	119.33 +/- 25.46	30.67 +/- 1.31	50.00 +/- 1.13
Matador	54	D102_MAT	Winter OSR	97.33 +/- 7.53	117.67 +/- 2.85	46.00 +/- 9.80	62.00 +/- 8.84
Moana, Moana Rape	43	IB106_MoaR	Winter fodder	DNF +/- 0.00	DNF +/- 0.00	65.00 +/- 3.92	82.33 +/- 4.57
Monty-028 DH	21	D108_MON	Spring OSR	-4.00 +/- 0.00	1.00 +/- 0.00	6.33 +/- 1.31	20.00 +/- 1.96
N01D-1330	27	IB110_N01D_1330	Spring OSR	-1.33 +/- 2.61	9.00 +/- 1.13	15.00 +/- 4.93	32.33 +/- 1.73
N02D-1952	28	IB111_N02D_1952	Spring OSR	-4.00 +/- 0.00	3.00 +/- 0.00	7.67 +/- 0.65	28.00 +/- 5.88
Ningyou 7	7	NIN	Semi-winter OSR	14.67 +/- 9.08	34.67 +/- 10.27	17.00 +/- 3.92	34.67 +/- 1.31
Norin	55	D117_NOR	Winter OSR	102.67 +/- 1.31	113.33 +/- 4.71	55.00 +/- 8.54	70.33 +/- 9.42
Palmedor	72	IB120_Palm	Winter OSR	53.00 +/- 25.71	71.33 +/- 20.13	31.67 +/- 1.31	48.00 +/- 1.13
Palu	44	IB121_Palu	Winter fodder	101.33 +/- 6.63	123.33 +/- 21.24	58.67 +/- 2.85	72.67 +/- 2.36
Poh 285, Bolko	73	IB125_POH285B	Winter OSR	115.00 +/- 25.28	128.67 +/- 16.49	47.00 +/- 7.84	63.33 +/- 9.22
Primor	56	D127_PRI	Winter OSR	71.67 +/- 2.85	87.67 +/- 2.36	32.33 +/- 4.28	50.33 +/- 3.46
Q100	10	D128_Q10	Synthetic	22.67 +/- 7.53	42.67 +/- 6.53	22.67 +/- 0.65	40.33 +/- 1.31
Quinta	74	IB130_Qui	Winter OSR	76.00 +/- 4.53	92.33 +/- 6.82	42.00 +/- 2.26	56.33 +/- 2.61
Rafal DH1	57	D131_Raf	Winter OSR	91.33 +/- 1.73	112.67 +/- 5.70	28.67 +/- 1.31	47.33 +/- 1.73

				Treatment NVERN		Treatment VERN	
Cultivar Name	**Cultivar (DNA) Number**	**Cultivar code**	**Crop type**	**Mean number of days to BBCH51**	**Mean number of days to BBCH60**	**Mean number of days to BBCH51**	**Mean number of days to BBCH60**
Ragged Jack	1	D132_RAGJ	Kale	DNF +/- 0.00	DNF +/- 0.00	84.00 +/- 12.75	100.67 +/- 10.51
Ramses	75	IB133_Ram	Winter OSR	107.00 +/- 7.84	125.33 +/- 8.19	49.00 +/- 1.13	66.00 +/- 1.96
Rapid Cycling Rape (Crgr5)	2	D134_RAPCR	Kale	7.00 +/- 1.13	23.67 +/- 2.36	24.00 +/- 1.96	41.33 +/- 1.73
Regent	18	A135_Reg	Spring OSR	4.67 +/- 1.31	16.33 +/- 7.53	21.33 +/- 3.27	37.00 +/- 2.26
Rocket	49	A136_Roc	Winter OSR	93.67 +/- 3.46	113.00 +/- 1.13	34.67 +/- 6.43	53.67 +/- 6.23
Rocket (Pst) X Lizard DH Line	58	D137_ROCxLIZ	Winter OSR	94.33 +/- 2.61	108.67 +/- 1.73	27.33 +/- 0.65	44.67 +/- 1.31
Samourai	76	IB139_Sam	Winter OSR	71.00 +/- 1.13	88.67 +/- 2.85	27.67 +/- 2.61	46.00 +/- 2.26
Sensation_NZ	37	IB141_SenNZ	Swede	137.33 +/- 5.23	DNF +/- 0.00	48.00 +/- 1.96	61.33 +/- 3.64
Shannon X Winner DH Line	77	IB143_ShaxWin	Winter OSR	70.67 +/- 6.23	86.33 +/- 5.58	27.67 +/- 0.65	46.67 +/- 0.65
Shengliyoucai	3	A144_She	Semi-winter OSR	4.67 +/- 1.31	18.33 +/- 4.28	16.67 +/- 2.85	34.33 +/- 1.73
Siberische_Boerenkool	8	D146_SIBB	Kale	DNF +/- 0.00	DNF +/- 0.00	60.33 +/- 1.31	78.00 +/- 1.96
Slapska, Slapy	52	IB150_SlaS	Winter OSR	126.00 +/- 15.31	DNF +/- 0.00	55.00 +/- 6.30	77.67 +/- 5.58
Slovenska_Krajova	78	IB151_SloK	Winter OSR	138.33 +/- 3.27	DNF +/- 0.00	41.67 +/- 2.85	59.00 +/- 2.99
Stellar DH	19	A155_Ste	Spring OSR	-2.00 +/- 0.00	8.00 +/- 2.99	18.33 +/- 1.31	36.67 +/- 1.31
Taisetsu	68	IB164_Tai	Winter vegetable	65.00 +/- 3.92	81.67 +/- 4.71	28.00 +/- 0.00	45.33 +/- 1.31
Tapidor DH	95	TAP	Winter OSR	71.33 +/- 4.57	89.33 +/- 3.46	28.67 +/- 1.73	45.67 +/- 0.65

				Treatment NVERN		Treatment VERN	
Cultivar Name	**Cultivar (DNA) Number**	**Cultivar code**	**Crop type**	**Mean number of days to BBCH51**	**Mean number of days to BBCH60**	**Mean number of days to BBCH51**	**Mean number of days to BBCH60**
Temple	50	A168_TEM	Winter OSR	52.00 +/- 22.72	73.33 +/- 16.57	29.00 +/- 1.96	45.67 +/- 0.65
Tequila X Aragon DH Line	79	IB169_TeqxAra	Winter OSR	DNF +/- 0.00	DNF +/- 0.00	48.33 +/- 3.64	68.00 +/- 4.53
Tina	34	D170_TIN	Swede	DNF +/- 0.00	DNF +/- 0.00	45.33 +/- 3.27	59.67 +/- 4.71
Topas	22	D173_TOP	Spring OSR	4.00 +/- 0.00	11.33 +/- 1.73	18.00 +/- 3.39	36.33 +/- 2.36
Verona	12	A176_Ver	Winter OSR	113.00 +/- 26.61	123.67 +/- 20.94	33.00 +/- 5.19	49.33 +/- 3.46
Victor	94	R_IB177_Vic	Winter OSR	90.00 +/- 15.72	120. 33 +/- 24.20	34.33 +/- 3.27	50.67 +/- 2.36
Vige DH1	38	IB178_Vig	Swede	DNF +/- 0.00	DNF +/- 0.00	78.00 +/- 13.05	114.67 +/- 30.91
Vision	80	IB180_Vis	Winter OSR	82.00 +/- 9.67	101.00 +/- 10.79	28.33 +/- 0.65	46.33 +/- 0.65
Westar DH10	23	D181_WESD	Spring OSR	-1.00 +/- 0.00	7.00 +/-2.26	12.33 +/- 3.27	30.33 +/- 1.73
Wilhelmsburger-Reform	35	D182_WILR	Swede	DNF +/- 0.00	DNF +/- 0.00	46.33 +/- 5.58	60.33 +/- 5.81
Xiangyou 15	4	A187_Xia	Semi-winter OSR	DNF +/- 0.00	DNF +/-0.00	57.67 +/- 11.56	72.67 +/- 11.45
Zhongshuang II	5	A188_Zho	Semi-winter OSR	22.67 +/- 0.65	38.33 +/- 1.31	19.33 +/- 1.73	35.00 +/- 1.13

Under VERN conditions all cultivars, except for the swede cultivar Altasweet, flowered (Table 4. 2) suggesting Altasweet requires longer than 6-weeks vernalisation treatment to promote flowering. Excluding Altasweet, mean flowering times, counted from the end of the vernalisation treatment, ranged between 2-84 dayswith the earliest flowering cultivar, a spring oilseed rape Karoo-057, flowering in an average of two days and the latest flowering cultivar, a kale type Ragged Jack, flowering in an average of 84 days.

Variation for vernalisation requirement was detected in the OREGIN DFFS diversity population of *B. napus*. Under NVERN treatment flowering times, counted from the end of the NVERN treatment at 21 days after sowing, ranged between -4-140 days on average. Therefore, spring oilseed rape cultivars Karoo-057, Monty-028, N01D-1330, N02D-1952, Stellar and Westar 10 flowered within an average of 17-20 days after sowing and flowered before being transplanted to the poly-tunnel (Table 4. 2). Ten cultivars did not flower within the time-frame of the experiment demonstrating an obligate requirement for vernalisation. These cultivars were given a did not flower (DNF) score of 140 days and included one winter oilseed rape, one Chinese semi-winter oilseed rape, one winter fodder, five swede and two kale crop types (Table 4. 2).

Significant differences in flowering times were detected between crop types under both VERN and NVERN treatments (Figure 4. 2, $p<0.001$ for both treatments, Kruskal-Wallis ANOVA) and certain crop types flowered consistently earlier than others. For example, spring oilseed rape cultivars flowered earlier on average compared with winter oilseed rape, indicating spring oilseed rape have been bred for early flowering regardless of vernalisation treatment. Winter oilseed rape cultivars however flowered earlier than swede cultivars, indicating that late flowering has been selected in swede cultivars. Swede cultivars are grown for their vegetative tissue therefore late flowering ensures extended vegetative development for vegetable crop production.

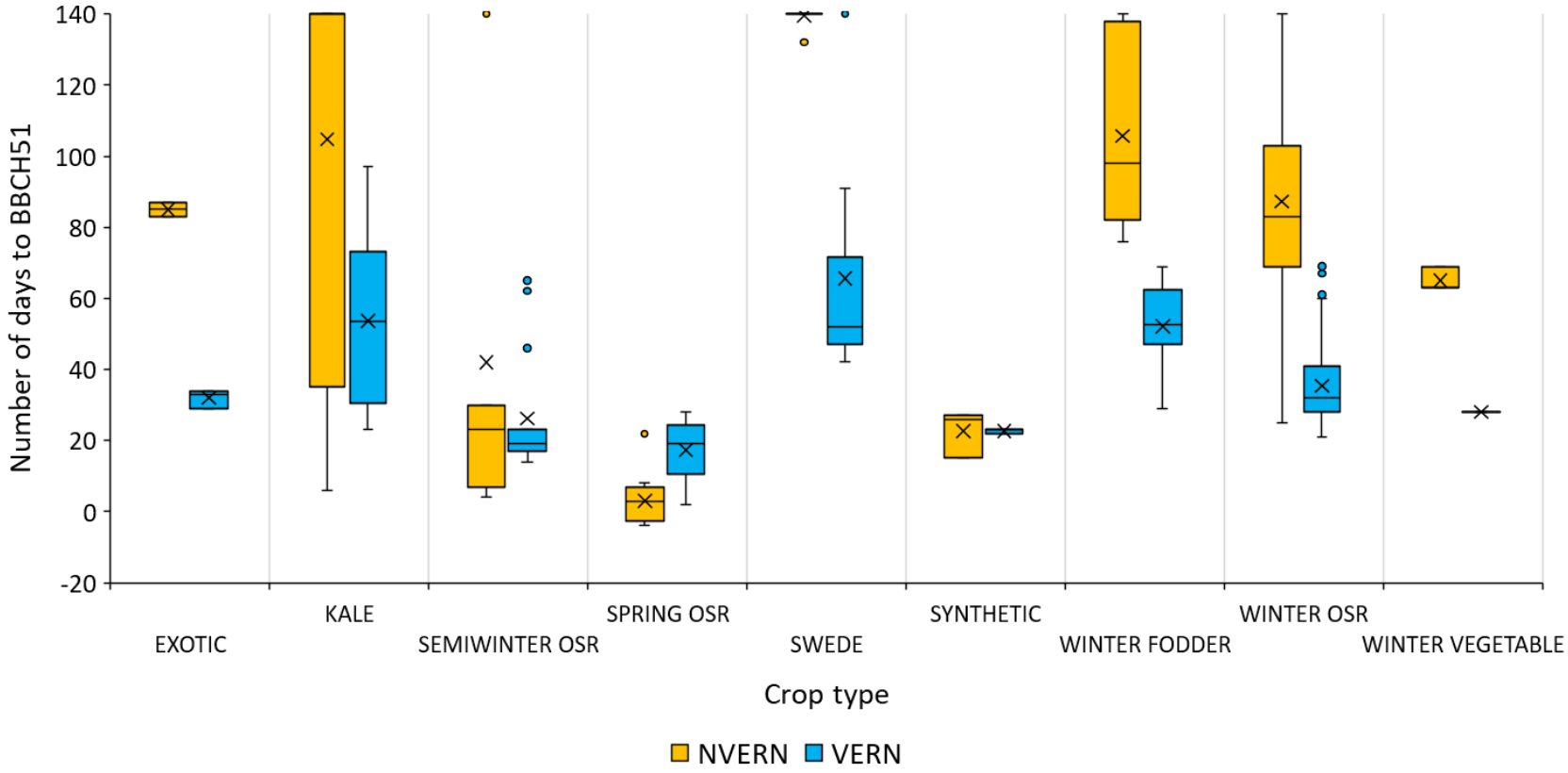


Figure 4. 2: The vernalisation requirement of *B. napus* crop types. The flowering time measurements collected under NVERN (orange) and VERN (blue) treatments are plotted as boxplots illustrating the median, mean (x) and range of measurements grouped by crop type, OSR=oilseed rape.

The flowering time response to vernalisation varied between crop type and plants could be divided into three groups accordingly (Figure 4. 2). Chinese semi-winter oilseed rape and synthetic crop types exhibited no significant difference in flowering time with (VERN) and without vernalisation (NVERN) ($p=0.559$, $p=0.007$ for Chinese semi-winter and synthetic crop types respectively; Mann-Whitney U test). This indicates these crop types have no vernalisation requirement and their flowering time is not affected by vernalisation. Despite detecting large differences in mean flowering times in exotic (mean difference of 53 days) and vegetable (mean differences of 33 days) crop types, no significant difference in flowering times ($p=0.1$, $p=0.1$ for exotic and vegetable crop types respectively; Mann-Whitney U test) were detected between vernalisation treatments. Kale, swede, winter fodder, and winter oilseed rape crop types exhibited a significant acceleration in flowering time in response to vernalisation treatment ($p=0.03$ for Kale, $p<0.001$ for swede, $p<0.001$ for winter fodder, $p<0.001$ for winter oilseed rape; Mann-Whitney U test). These cultivars therefore have a requirement for vernalisation and flowered significantly later without vernalisation. Spring oilseed rape cultivars flowered significantly later with vernalisation (VERN) compared to without vernalisation (NVERN) ($p<0.001$; Mann-Whitney U test), suggesting extended periods of cold temperatures negatively affect flowering time, possibly by slowing growth rate in spring oilseed rape.

Kale is grown for its vegetative tissue, and therefore is expected to be late flowering and exhibit a requirement for vernalisation. Classified within the kale crop types according to OREGIN is the cultivar Rapid Cycling Rape which flowers on average 17 days earlier NVERN compared with VERN. The remaining three kale cultivars were late flowering without vernalisation and exhibited an acceleration in flowering in response to vernalisation treatment. According to its response to vernalisation, Rapid Cycling Rape should be grouped with spring oilseed rape and not kale crop types.

Winter oilseed rape cultivars exhibited a vernalisation requirement and flowered significantly earlier with vernalisation ($p<0.001$; Mann-Whitney U test). A broad range of flowering times were measured under NVERN and VERN conditions

suggesting, although vernalisation treatment accelerated flowering time on average, variation for vernalisation requirement and response is present within winter oilseed rape (Figure 4. 3). For example, under NVERN conditions the winter oilseed rape cultivar Cabriolet flowered 101 days on average earlier than the winter oilseed rape cultivar Darmor indicating variation for vernalisation requirement. Under VERN conditions Cabriolet flowered 25 days on average earlier than Darmor indicating variation for flowering time after vernalisation treatment between both cultivars.

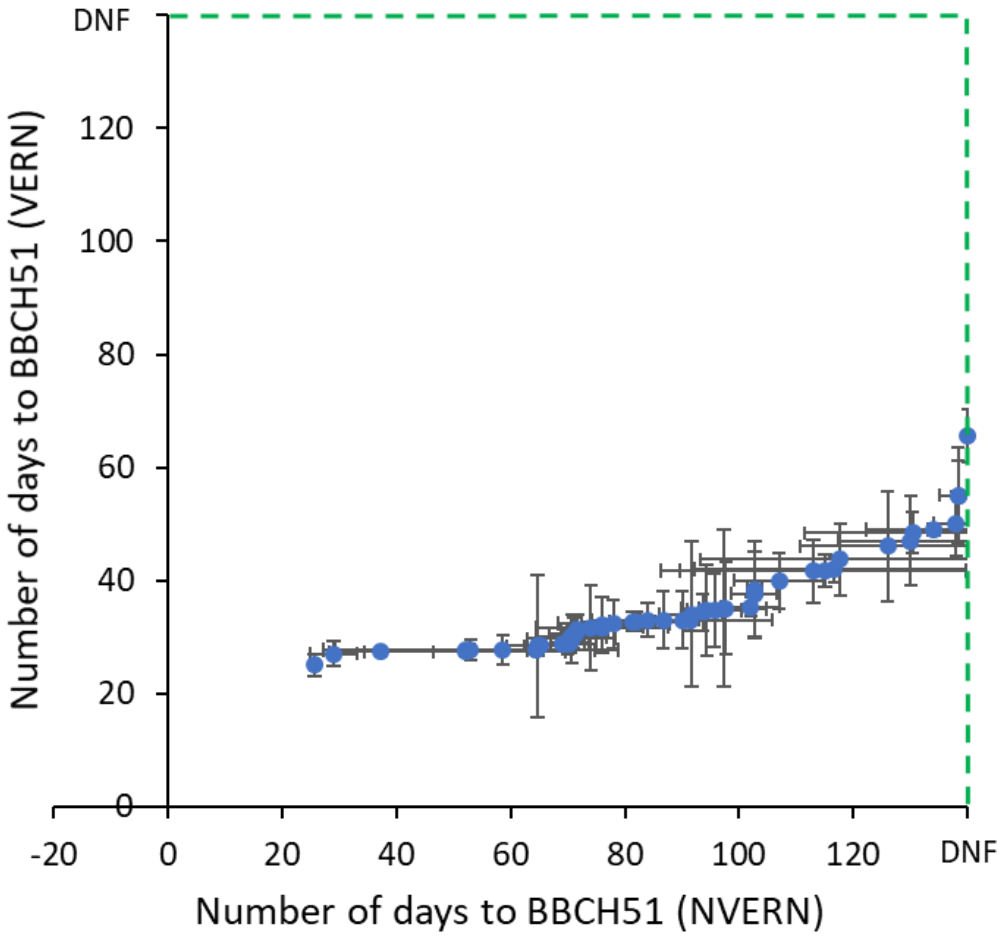


Figure 4. 3: The relationship between flowering time with and without vernalisation in winter oilseed rape. A positive correlation is detected between flowering time measurements of winter oilseed rape cultivars under no vernalisation (NVERN) and after vernalisation (VERN) conditions. The green dashed line represents a DNF value of 140 days. Mean values were calculated from three plants per cultivar, error bars denote the 95% confidence around the mean.

4.2.2. Expression variation at *BnaFLC.A10* associates significantly with flowering time in *Brassica napus*

To identify genes that are important for variation in flowering time with (VERN) and without (NVERN) vernalisation, I analysed the flowering time datasets described in Section 4.3.1. by Gene Expression Marker (GEM) analysis (methods published by Harper *et al*., 2012 and described in Section 2.9.3). GEM analysis identified significant correlations between variation in unigene expression value (plotted as RPKM; originally published by Harper *et al*., 2012 and Lu *et al*., 2014b) and flowering time. The results of these analyses are displayed as Manhattan plots (Figure 4. 4) and the most significant GEMs for each vernalisation treatment that scored aboe the Bonferroni threshold of association are listed in Table S. 1 and Table S. 2.

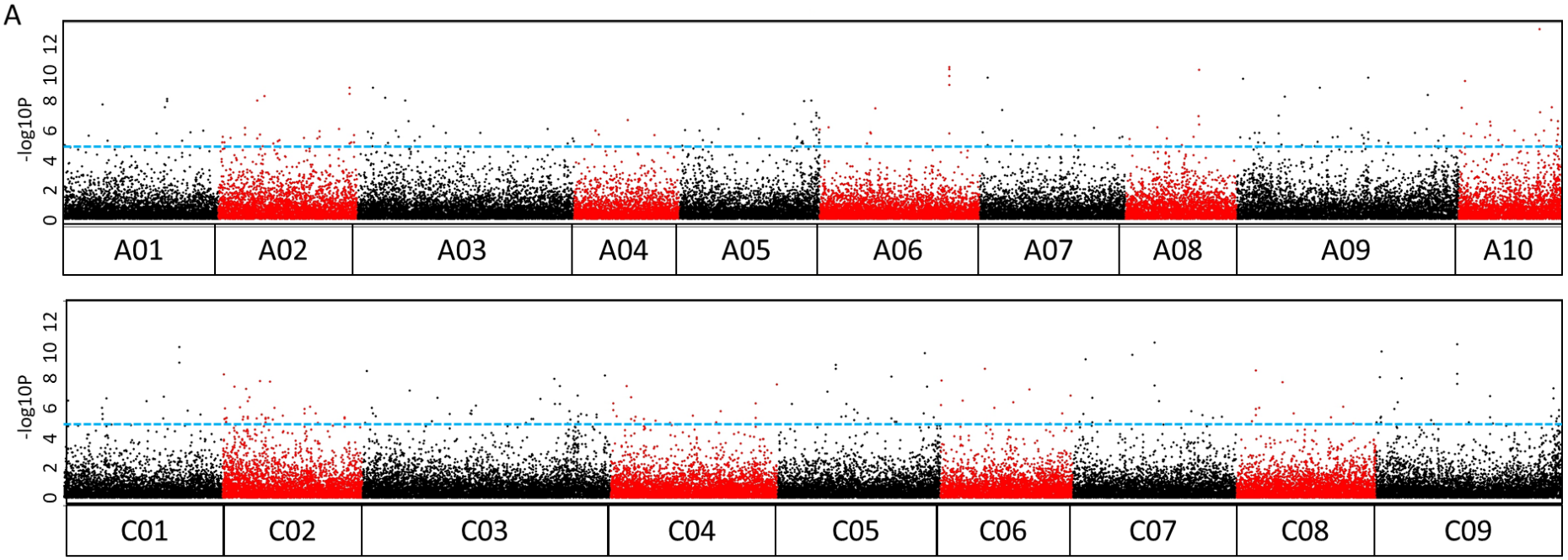
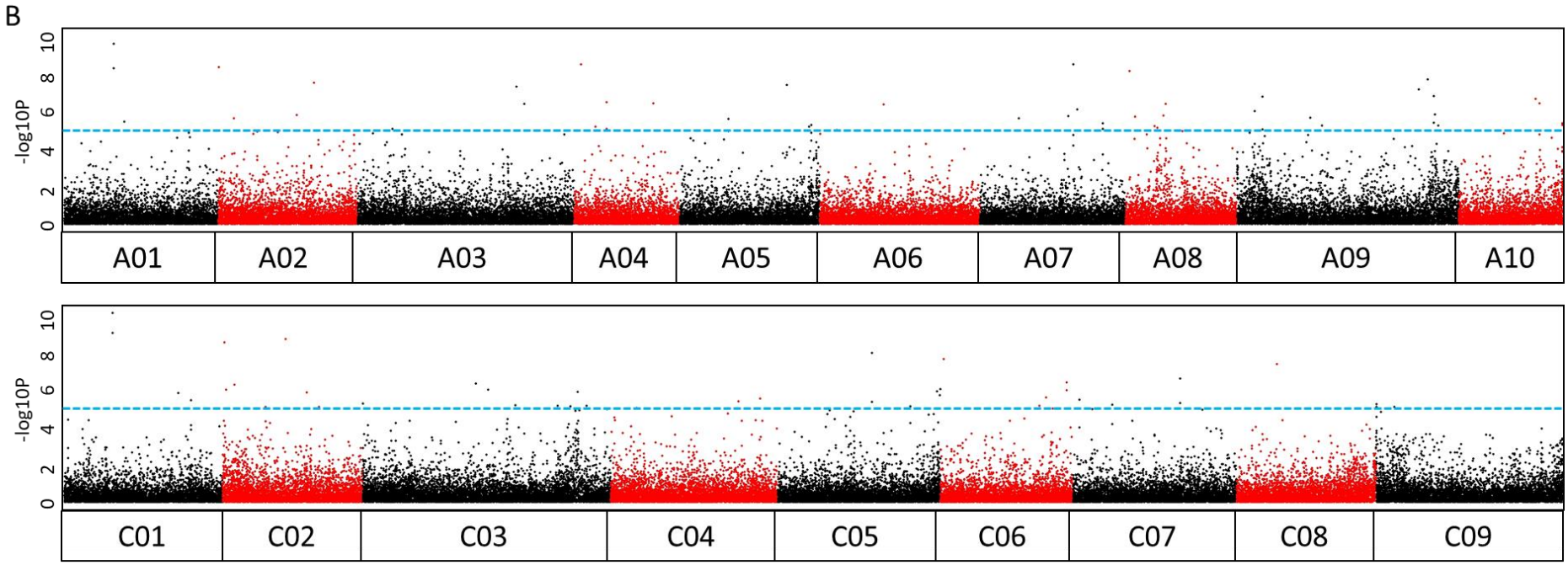


Figure 4. 4: GEM analysis results for flowering time (BBCH51) under (A) NVERN and (B) VERN treatments. Gene expression markers are plotted in pseudomolecule order, their position within a chromosome plotted along the x-axis and the strength of association with flowering time is plotted along the y-axis. Each genome is represented separately. The Bonferroni threshold of association is denoted by a blue dashed line.

Figure 4. 4 continued



The most significant GEM that associated with flowering time NVERN was JCVI_22584, the unigene that corresponds to *BnaFLC.A10* (Figure 4. 5 A, -log10P =12.84). The expression level of JCVI_22584 also formed a significant association under VERN conditions, but to a lower significance level (Figure 4. 5 B, -log10P =2.56E-07) and plotted separately in Figure 4. 6 A and B. The mRNA-seq dataset was generated from leaf material of 21-day old *B. napus* plants. This suggests that variation in the pre-vernalisation expression level of *BnaFLC.A10* is significantly correlated with variation in flowering time with and without vernalisation, possibly suggesting that, like in *A. thaliana* (Li *et al*., 2015), variation in *BnaFLC.A10* expression before (pre-)vernalisation influences expression level after (post-)vernalisation.

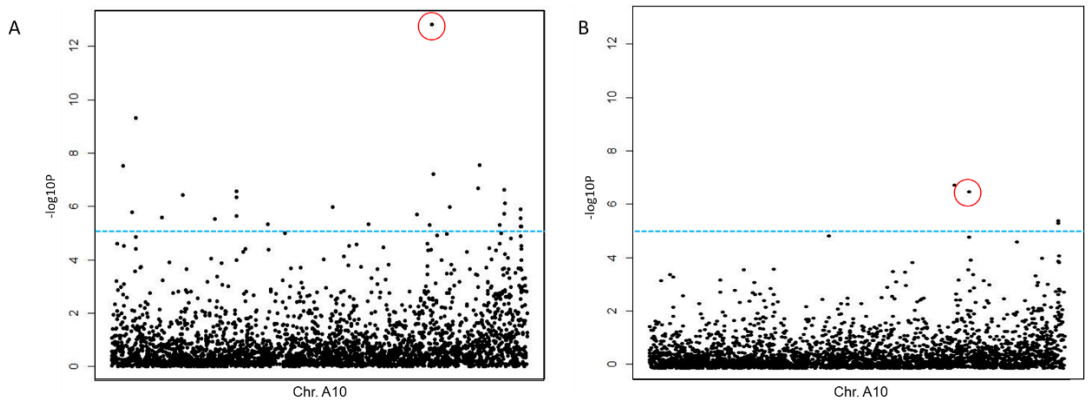


Figure 4. 5: Zoom of GEMs on chromosome A10 associated with flowering time variation under NVERN (A) and VERN (B) treatments. Each GEM is plotted at -log10P significance along the y-axis and in pseudomolecule order along the x-axis. The significantly associated GEM JCVI_22584 is highlighted in red. The Bonferroni threshold of association is denoted with a blue dashed line.

The unigene JCVI_22584 corresponds to *BnaFLC.A10* however, *BnaFLC* genes share up to 85% sequence identity within their coding regions (Schranz *et al*., 2006, Zou *et al*., 2012). To correct for miss-assignment of sequencing reads caused by high sequence similarity between *BnaFLC* genes, sequencing reads for each *BnaFLC* gene were counted and graphed using a homeologue specific sequence (See Section 2.9.2 for method details). The *in silico* expression levels of *BnaFLC* for each *B. napus* cultivar (Figure S. 2) were compared with the RPKM values for unigene JCVI_22584. Although correlated ($p<0.001$ for both NVERN and VERN treatments,

Spearman's rank correlation; Figure 4 .6 A and Figure 4. 6 B) a discordance between the RPKM levels of unigene JCVI_22584 and the *in silico* expression level of *BnaFLC.A10* was apparent suggesting the quantitative expression level of JCVI_22584 may not truly represent the actual expression level of *BnaFLC.A10*. Nevertheless, a significant correlation between flowering time and *BnaFLC.A10 in silico* expression is detected ($p<0.001$ for both NVERN and VERN treatments, Spearman's rank correlation; Figure 4. 6 C and Figure 4.6 D) indicating that variation in the pre-vernalisation expression level of *BnaFLC.A10* is likely a significant determinant of flowering time under NVERN and VERN conditions.

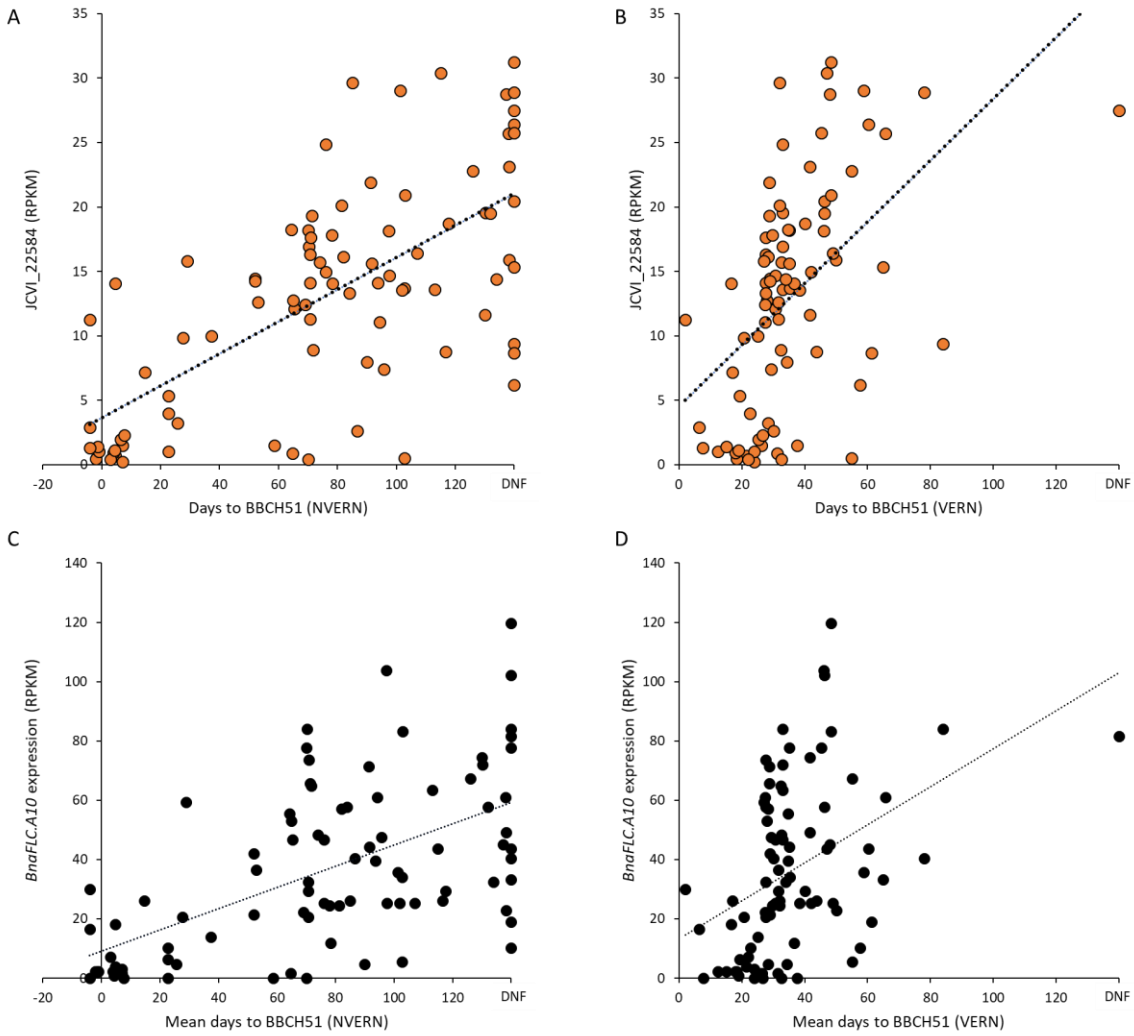


Figure 4. 6: Relationship between *BnaFLC.A10* expression and flowering time. The relationship between RPKM values of unigene JCVI_22584 and flowering time (A) NVERN and (B VERN. The relationship between the *in silico* expression values of *BnaFLC.A10* with flowering time (C) NVERN and (D VERN.

The mRNA-seq data indicates that *BnaFLC.A10* expression varied between *B. napus* cultivars (Figure S. 2) and this expression variation contributed significantly to variation in flowering time (Figure 4. 5 and Figure 4. 6). To confirm this, I measured the quantitative expression of *BnaFLC.A10* by qRT-PCR in seven *B. napus* cultivars before (NV) and after six-weeks (6WT0) of vernalisation treatment (using methods described in Section 2.6). I selected cultivars Darmor, Major, Altasweet and Sensation New Zealand were selected as they exhibited late flowering without vernalisation, and Karoo-057, Chuanyou-2 and Ningyou-7 as they exhibited early flowering without vernalisation (Table 4. 2).

The highest quantitative expression levels of *BnaFLC.A10* were detected in the winter oilseed rape cultivars Darmor and Major and in the swede cultivars Altasweet and Sensation New Zealand (Figure 4. 7). Spring and Chinese semi-winter oilseed rape cultivars Karoo-057, Chuanyou-2 and Nigyou-7 in contrast exhibited much lower expression levels (Figure 4. 7). Interestingly *BnaFLC.A10* expression was also higher in the winter oilseed rape and swede cultivars after a six-week vernalisation treatment compared with the spring and Chinese semi-winter cultivars. These late flowering cultivars therefore had high *BnaFLC.A10* expression before and after cold while the early flowering cultivars had low *BnaFLC.A10* expression under both treatments. According to the GEM analysis, the expression level of *BnaFLC.A10* in 21-day old plants correlated significantly with flowering time under both NVERN and VERN treatments (Figure 4. 5 and Figure 4. 6), perhaps suggesting the pre-vernalisation level of *BnaFLC.A10* can influence flowering time with and without cold. The higher quantitative expression of *BnaFLC.A10* in late flowering cultivars before and after six-weeks vernalisation support this.

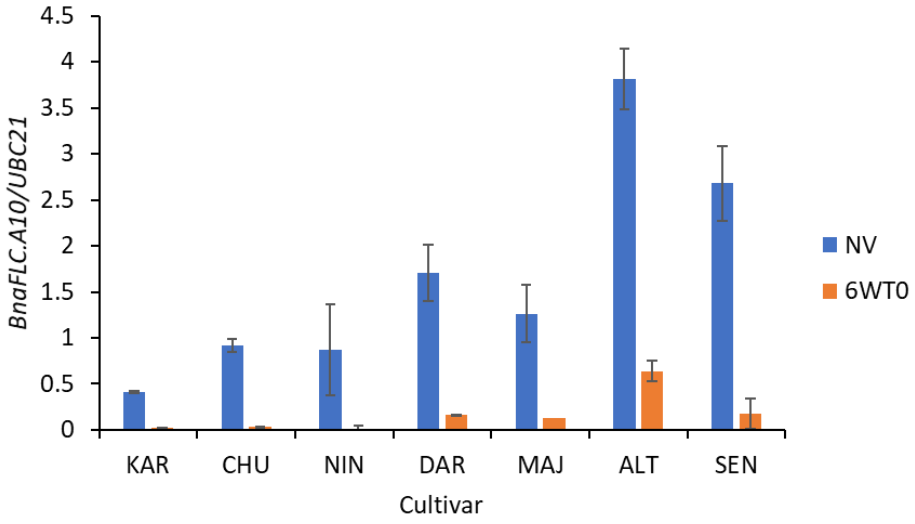


Figure 4. 7: Expression of *BnaFLC.A10* in *B. napus* cultivars. Analysis of gene expression by qRT-PCR confirms that *BnaFLC.A10* expression varies between *B. napus* cultivars before (NV) and after vernalisation (6WT0). Mean expression values were calculated from at least two biological replicates. Error bars denote 95% confidence intervals.

4.2.3. Expression of five *BnaFLC* genes significantly correlate with flowering time without vernalisation

At least nine *BnaFLC* genes have been characterised in *B. napus* (Zou *et al*., 2012) and at least five of these are thought to act as floral inhibitors (Tadege *et al*., 2010). Due to the significant correlation between *BnaFLC.A10* expression and flowering time, I analysed the *in silico* expression of 9 *BnaFLC* genes for correlations with flowering; *BnaFLC.A02*, *BnaFLC.A03a*, *BnaFLCA03b*, *BnaFLC.C02*, *BnaFLC.C03a*, *BnaFLC.C09a*, *BnaFLC.C09b*, *BnaFLC.C09c* (Figure S. 2).

In agreement with previous reports (Zou *et al*., 2012) expression of *BnaFLC.C03b* was not detected and assumed to be a pseudogene. Correlation coefficients were calculated and significant associations were found between flowering time (NVERN) and *BnaFLC.A03a*, *BnaFLCA03b*, *BnaFLC.A10*, *BnaFLC.C02* and *BnaFLC.C09b* expression ($p=0.019$, $p<0.001$, $p<0.001$, $p=0.001$, $p=0.018$ respectively, Table 4. 3 and Figure 4. 8). Four *BnaFLC* genes exhibited a positive

correlation with flowering time suggesting high expression of *BnaFLC.A03a*, *BnaFLCA03b*, *BnaFLC.A10* and *BnaFLC.C02* conferred late flowering. Interestingly *BnaFLC.C09b* exhibited a significant negative correlation with flowering time suggesting this gene is highly expressed in early flowering *B. napus* cultivars. The total sum of *BnaFLC* expression from all nine *BnaFLC* genes did not correlate significantly with flowering time NVERN (p=0.080; -0.187 Spearman's Rank Correlation) possibly suggesting the expression of *BnaFLC* does not contribute additively to flowering time.

Table 4. 3: Relationship between *BnaFLC* expression and flowering time (NVERN). Calculated by Spearman's rank correlation analysis.

	BnaFLC expression								
Correlation coefficient and significance value	***A02***	***A03a***	***A03b***	***A10***	***C02***	***C03a***	***C09a***	***C09b***	***C09c***
NVERN	0.167	0.250	0.457	0.589	0.351	0.064	-0.043	-0.252	-0.203
BBCH51	P=0.119	P=0.019 *	P<0.001 ***	P<0.001 ***	P=0.001 ***	P=0.552	P=0.692	P=0.018 *	P=0.058

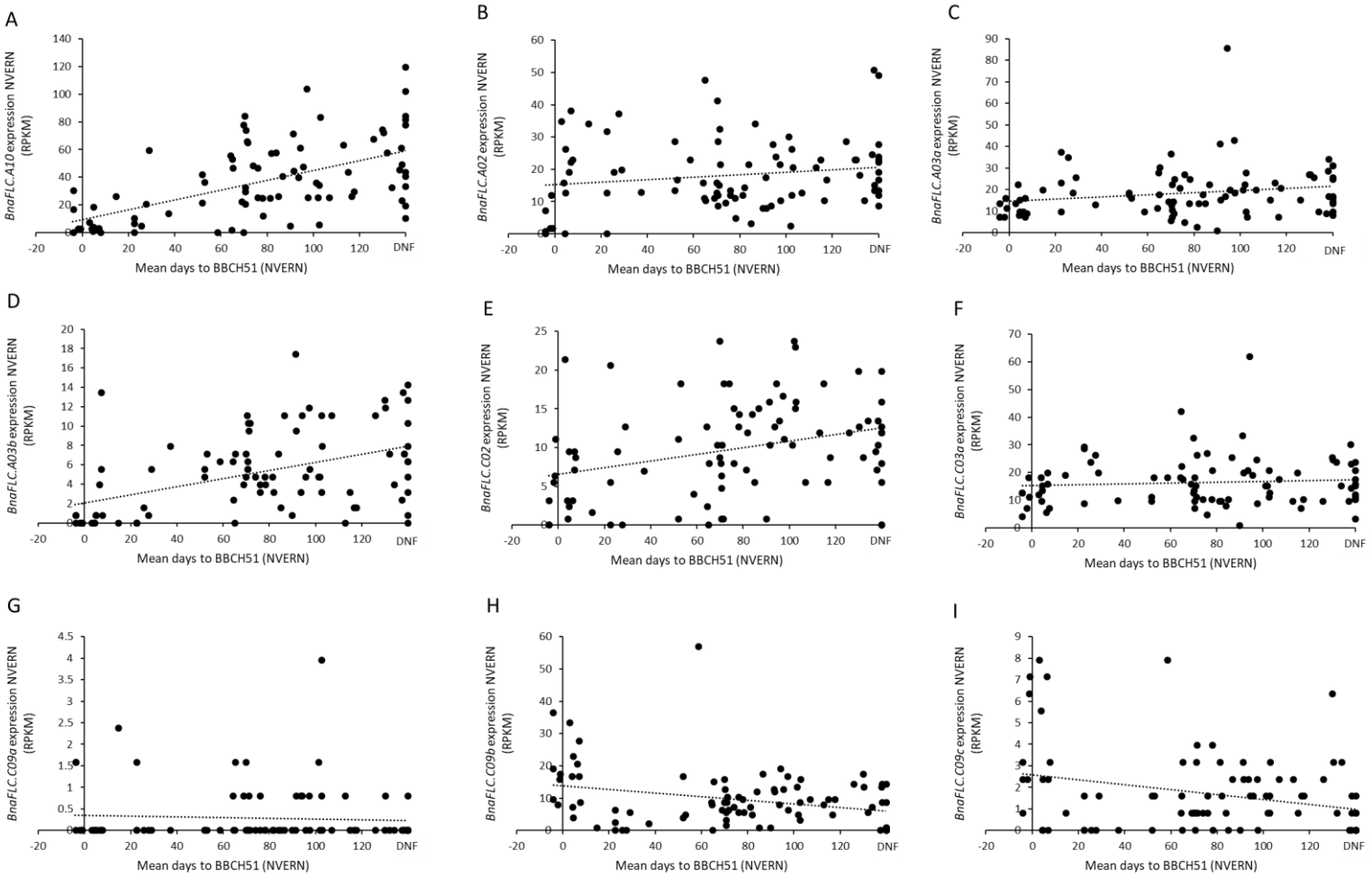


Figure 4. 8: Relationship between *in silico BnaFLC* expression and flowering time (BBCH51) NVERN. (A) *BnaFLC.A10*, (B) *BnaFLC.A02*, (C) *BnaFLC.A03a*, (D) *BnaFLC.A03b*, (E) *BnaFLC.C02*, (F) *BnaFLC.C03a*, (G) *BnaFLC.C09a*, (H) *BnaFLC.C09b*, (I) *BnaFLC.C09c.*

4.2.4. Cis-variation is associated with expression variation of *BnaFLC.A10*

I detected a significant association between the pre-vernalisation expression level of *BnaFLC.A10* and flowering time in the OREGIN DFFS diversity population of *B. napus* (Section 4.2.2.). In *A. thaliana*, variation in *FLC* expression is often associated with variation at *FRI* (Shindo *et al*., 2005, Atwell *et al*., 2010), however results in Chapter 3 suggest orthologues of *FRI* in *B. napus* might not be essential for determining *FLC* expression. To identify genes that are important for controlling the pre-vernalisation expression level of *BnaFLC.A10* I performed an Associative Transcriptomics analysis according to methods first described by Harper *et al*., 2012 (method details in Sections 2.9.4 and 2.9.5). *In silico* gene expression values for *BnaFLC.A10* (Figure S. 2) were analysed to identify SNP markers that associate significantly with variation in *BnaFLC.A10* expression level.

To correct for population structure within the OREGIN DFFS diversity panel a STRUCTURE (Evanno *et al*., 2005) analysis of 680 unlinked SNP markers was performed by Dr R. Wells. The results of this analysis are illustrated in Figure 4. 9. The highest deltaK value as calculated by STRUCTURE shows the *B. napus* population used in the present study has a structure of K=2 (Figure 4. 9 A) which agrees with previously published work (Harper *et al*., 2012, Lu *et al*., 2014b, Henklikova *et al*., 2018). However, a smaller peak in deltaK is also detected at K=4 which suggests the population can be further divided into four ancestral groups and sub-structure exists within the diversity population.

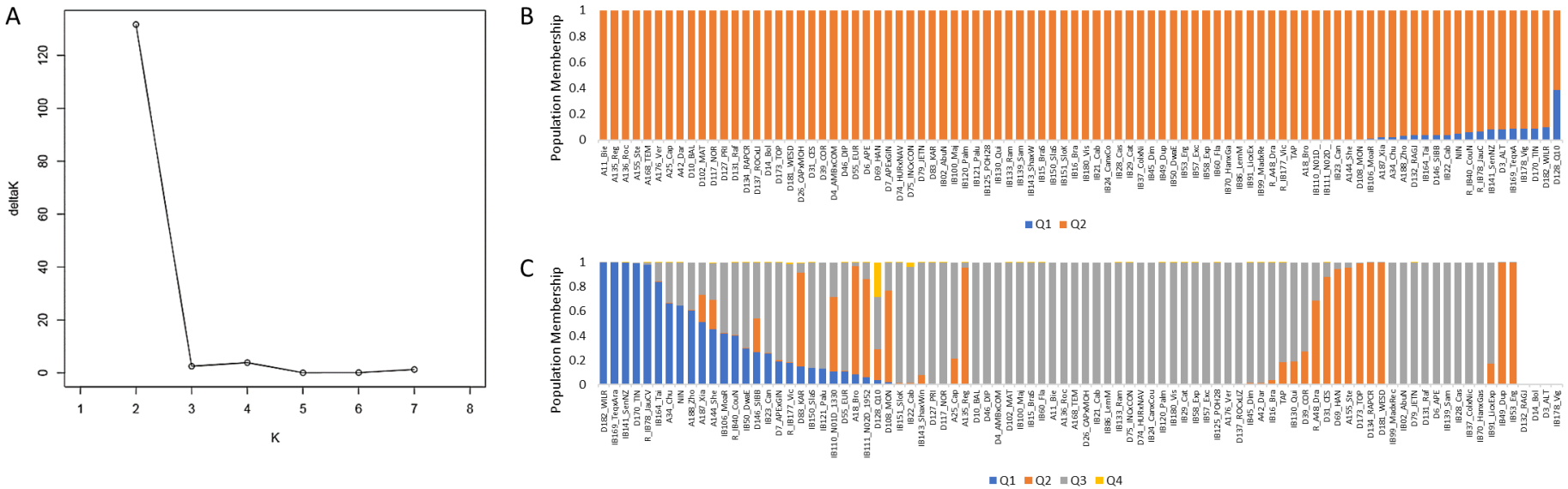


Figure 4. 9: Results of STRUCTURE analysis carried out in *B. napus*. The analysis was performed according to methods described by Evanno *et al*., (2005). (A) The most likely estimate for the value of K=2 (as reported by Harper *et al*., 2012) but the population can be further subdivided to K=4. (B and C) each *B. napus* cultivar is plotted with the proportion of each population (Q1, Q2, Q3 and Q4) from which they are derived as estimated by STRUCTURE (B) K=2, (C) K=4.

After assessing population structure within the OREGIN DFFS diversity population I performed an Associative Transcriptomics analysis on the *in silico* expression of *BnaFLC.A10* using both K=2 and K=4; linear regression was performed by GLM and MLM. To ensure all confounding variables were accounted for QQ plots were generated for each analysis; the expected distribution (x-axis) of association across the 144,131 SNPs were plotted against the observed (y-axis) distribution. A model that accounts for all confounding variables would generate SNP associations that followed an expected x=y until it curves sharply towards the end, only a small number of true significant associations should be identified (Atwell *et al*., 2010, Voorman *et al*., 2011). A deviation from the expected line of x=y demonstrates that a large proportion of the SNPs are significantly associated with the trait, revealing unaccounted bias within the dataset.

As displayed in Figure 4. 10 there is a deviation from the expected x=y line present in the GLM analyses performed with K=2 (Figure 4. 10 A) and K=4 (Figure 4. 10 B). There is a reduction in false positive associations with an MLM analysis (Figure 4. 10 C and Figure 4. 10 D), most likely due to the incorporation of a kinship matrix in addition to the Q-matrix in an MLM. The model with the least evidence of unexplained confounding variables for *BnaFLC.A10* expression is an MLM with a population structure of K=2 (Figure 4. 10 C). Incorporation of population structure as K=4 into the model increases the number of false positive associations and a greater deviation from the expected of x=y (Figure 4. 10 D).

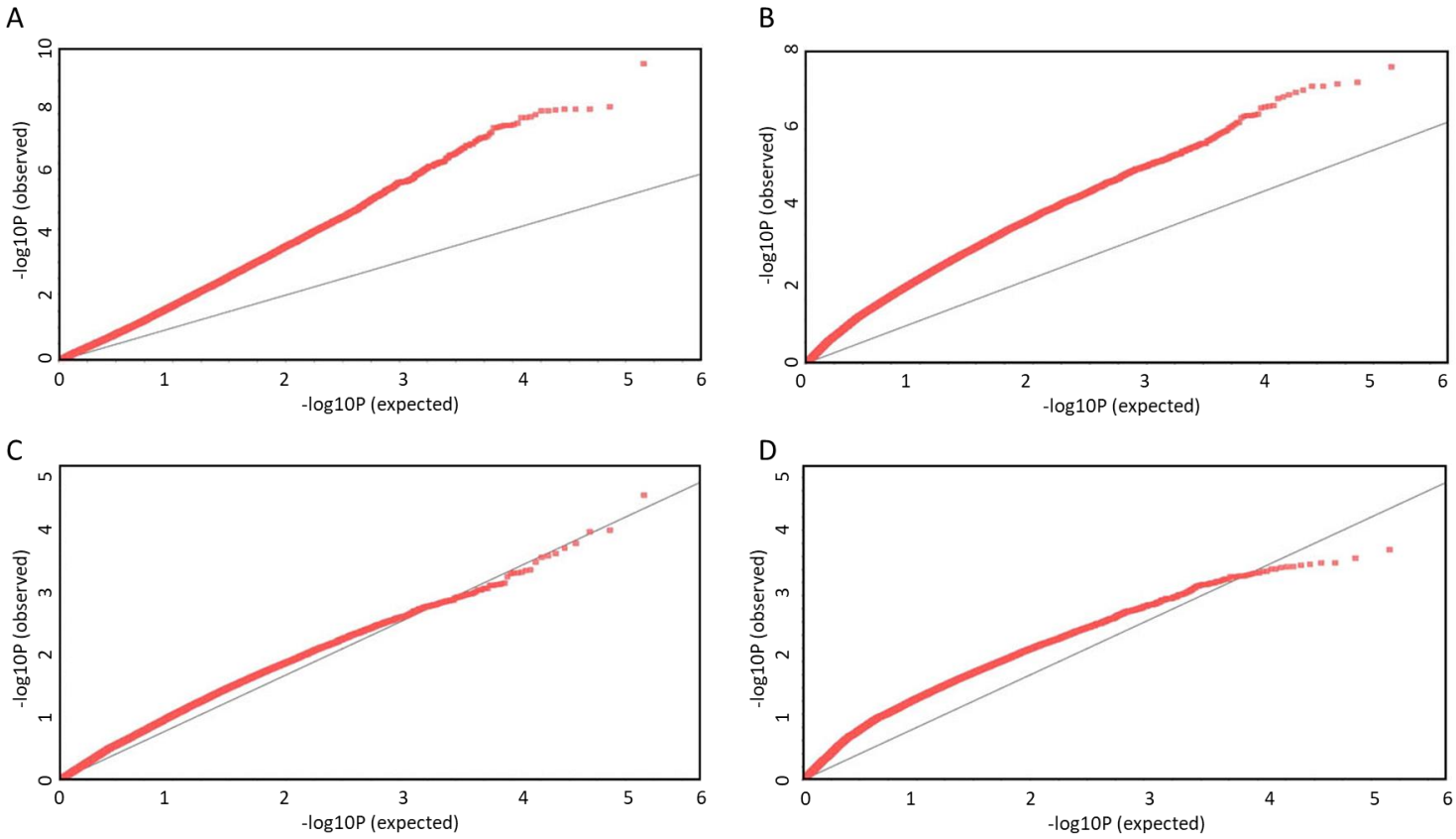


Figure 4. 10: Quantile-Quantile plots for SNP-association analysis. Model (A) GLM and K=2, (B) GLM and K=4, (C) MLM and K=2, (D) MLM and K=4. In each plot the grey line denotes the hypothetical line of no association when x=y, the red points denote the expected distribution of P-values (-log10P) plotted against the observed distribution of P-value (-log10P).

From these results, I performed an Associative Transcriptomics analysis using an MLM with K=2 to identify SNP markers that associated with the *in silico* expression of *BnaFLC.A10*. A Manhattan plot of the SNP associations is displayed in Figure 4 .11. Four peaks were detected (Figure 4. 11); a major peak detected on chromosome A10/C09, and three minor peaks on chromosomes A02/C02, A03/C03, and A09/C08 however none were found above the Bonferroni threshold of significant association. The unigene marker at the top each peak is listed in Table 4. 4.

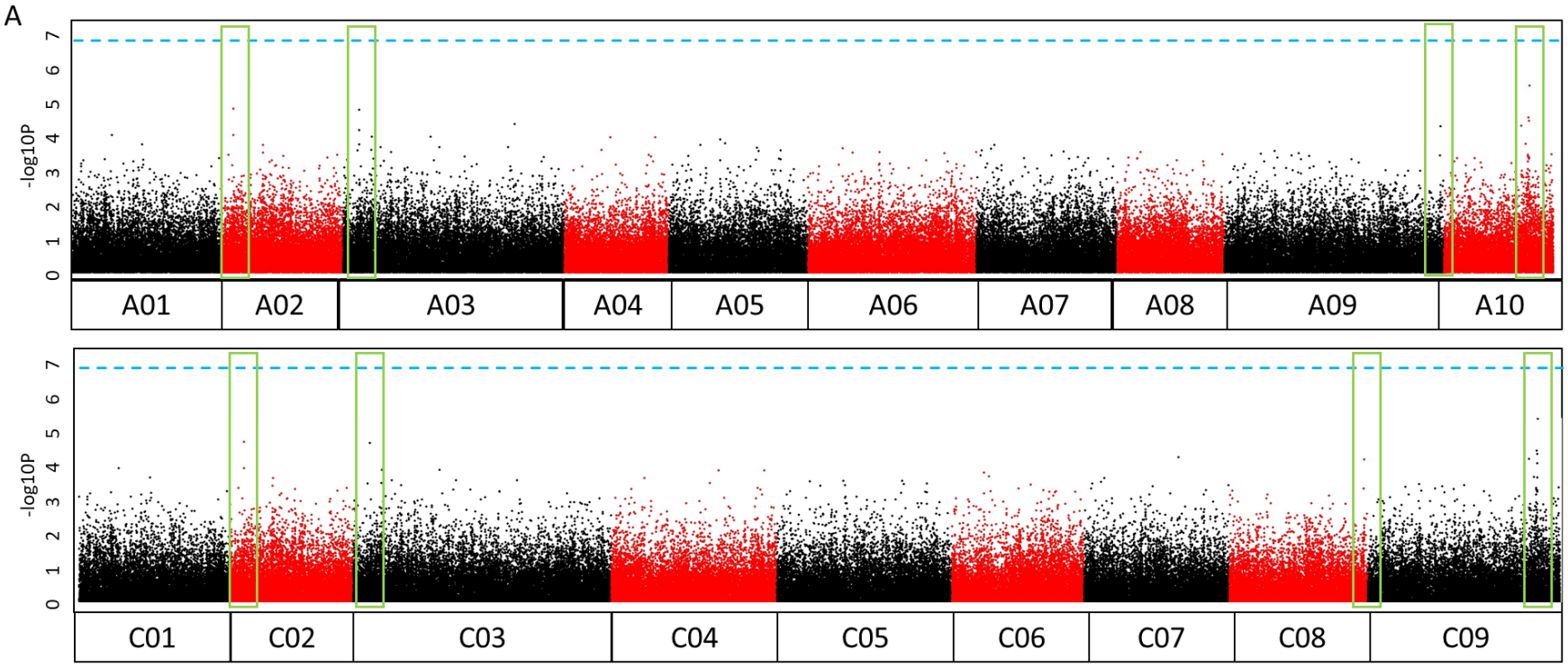


Figure 4. 11: Associative transcriptomics SNP association analysis results for *in silico* expression of *BnaFLC.A10*. (A) Manhattan plot for SNP-trait association, each SNP marker is plotted in pseudomolecule order along the x-axis with the strength of association (-log10P) along the y-axis. Each genome is plotted separately. (B) Zoom of SNP association peaks on chromosomes A02/C02, A03/C03, A09/C08 and A10/C09. The peak on A10/C09 shows the strongest association, SNP markers in JCVI_22584 (the unigene with homology to *BnaFLC.A10*) are highlighted in magenta. The Bonferroni threshold of association is denoted with a blue dashed line.

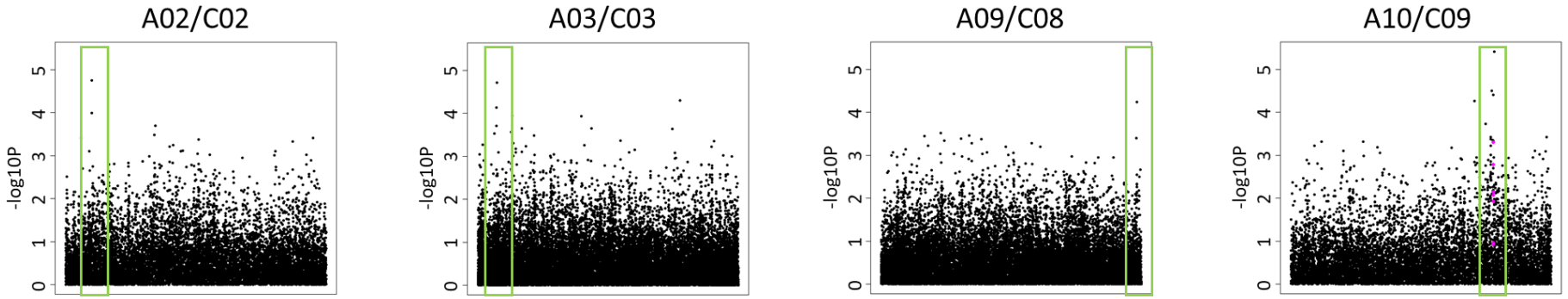


Figure 4. 11 continued

Table 4. 4: The most significantly associated unigene SNP marker in each of four peaks detected for variation in expression level of *BnaFLC.A10*, their pseudomolecule position, the alleles type, the number of cultivars carrying each allele, the effect of each allele, the significance of association between unigene and *BnaFLC.A10* expression, and the *A. thaliana* gene with highest BLAST homology.

Chr.	**Unigene Marker**	**Pseudomolecule position (kb) (A/C genomes)**	**Allele**	**#Cultivars**	**Effect**	**-log10P**	**Homology to *A. thaliana* gene**
A10/C09	JCVI_10962:251	12391526/ 60153661	Y	50	36.44	4.414	AT5G09850; Transcription elongation factor (TFIIS) family protein
			T	28	0.34		
			N	10			
A02/C02	JCVI_1049:424	1580570/ 2251144	C	64	36.35	4.743	AT5G10360; EMBRYO DEFECTIVE 3010
			Y	24	0.00		
			N	0			
A03/C03	EV216481:142	1786792/ 2112579	T	51	29.68	4.712	AT5G10360; EMBRYO DEFECTIVE 3010
			Y	37	0.00		
			N	0			
A09/C08	JCVI_21063:284	40962165/ 38740725	A	60	22.04	4.236	AT1G02660; PLASTID LIPASE2
			R	10	41.32		
			N	18			

4.2.5. Allelic variation is detected at *BnaFLC.A10*

The peak on chromosome A10/C09 corresponds to a SNP within the unigene JCVI_10962 which exhibits homology to the *A. thaliana* gene AT5G09850, part of the Transcription Elongation Factor (TFIIS) family of proteins (Table 4. 4). The association suggests SNP variation at this gene, or a gene that is closely linked to it, may influence *BnaFLC.A10* expression. Perhaps most interesting however, is that unigene JCVI_10962 is within the predicted LD range (<0.15Mb, Lu *et al*., 2014b) from unigene JCVI_22584, or *BnaFLC.A10*. Strikingly SNP markers within unigene JCVI_22584 form significant associations with *BnaFLC.A10* expression variation (-log10P<4; highlighted in magenta in Figure 4. 11 B) suggesting that allelic variation at *BnaFLC.A10* itself, or its homeologue *BnaFLC.C09*, is important for *BnaFLC.A10* expression. To confirm the SNPs at unigene JCVI_22584 were in *BnaFLC.A10* and not *BnaFLC.C09*, I explored sequence variation at *BnaFLC.A10* in seven *B. napus* cultivars, representing four crop types; 1 spring oilseed rape (Karoo-057), 2 Chinese semi-winter oilseed rape (Ningyou-7 and Shengliyoucai), 3 winter oilseed rape (Darmor, Major and Tapidor) and 1 swede (Sensation New Zealand).

I designed four primer pairs to amplify and sequence overlapping regions of the entire *BnaFLC.A10* locus from the translational start to stop codons (Table 2. 5). Sequence traces were analysed for quality and assembled into full gene sequences before being compared to the *BnaFLC.A10* sequence extracted from the winter oilseed rape Darmor-*bzh* reference (plants.ensembl.org) and to the *BnaFLC.A10* sequences of *B. napus* cultivars Tapidor (GenBank accession code JQ255381, Zou *et al*., 2012) and Ningyou-7 (GenBank accession code JX901101, Huo *et al*., 2012). To ensure that *BnaFLC.A10*, and not its homeologue *BnaFLC.C09*, had been sequenced from the cultivars I also compared their sequences to the *BnaFLC.C09a* sequence from *B. napus* cultivar Tapidor (GenBank accession code JQ255832, Zou *et al*., 2012). The sequencing results are illustrated in Figure 4. 12.

A

		BnaFLC.A10 gene structure																						
		exon1							intron1													exon2	intron2	intron5	exon7
		Position from translation start site																						
		9	29	30	58	66	154	156	306	444	461	472	484	595	1008	1360	1401	1834	2039	2180	2187	2692	2833	3600	4137
Reference	Darmor-bzh *BnaFLC.A10*	A	C	A	A	T	G	G	G	--	T	A	T	A	G	T	T	C	--	A	A	A	T	C	-
	JQ255381_*BnaFLC.A10*	G	G	A	A	T	G	G	G	--	-	-	T	A	G	T	T	C	--	A	A	A	T	C	-
	JX901141_*BnaFLC.A10*	G	G	A	A	T	G	G	T	CTTTTCTTCTT	-	-	C	T	A	C	C	G	ACGATT	G	G	A	T	C	A
	JQ255382_*BnaFLC.C9a*	A	C	G	A	C	A	A														G			
Line	Darmor	G	G	A	A	T	G	G	G	--	-	-	T	A	G	T	T	C	--	A	A	A	T	C	-
	Major	G	G	A	A	T	G	G	G	--	-	-	T	A	G	T	T	C	--	A	A	A	T	C	-
	Tapidor	G	G	A	A	T	G	G	G	--	-	-	T	A	G	T	T	C	--	A	A	A	T	C	-
	Sensation New Zealand	G	G	A	A	T	G	G	G	--	-	-	T	A	G	T	T	C	--	A	A	A	T	C	-
	Karoo-057	G	G	A	C	C	G	G	T	CTT	-	-	C	T	A	C	T	G	ACGATT	N	N	N	N	N	N
	Ningyou-7	G	G	A	A	C	G	G	T	CTTTTCTTCTT	-	-	C	T	A	C	T	G	ACGATT	A	G	A	G	T	-
	Shengliyoucai-11	G	G	A	A	C	G	G	T	CTTTTCTTCTT	-	-	C	T	A	C	T	G	ACGATT	A	G	A	G	T	-

B

```
                Arabidopsis thaliana (WT)   TGCATGG-ATTTCATTATTTCCTTGGAAAAAAATTGCATG
                Arabidopsis thaliana (vrn8) TGTATGG-ATTTCATTATTTCCTTGGAAAAAAATTGCATG
Darmor-bzh_chrA10:14999145..14999181        TGCATGGTATTT--TTTTTT--AACAAAAAAAATTGTATG
                JQ255381_BnaFLC.A10         TGCATGG-ATTT--TTTTTT---ACAAAAAAAATTGTATG
                JX901141_BnaFLC.A10         TGCATGG-ATTT--TTTTTT---ACAAAAAAAATTGTATG
                Karoo-057                   TGCATGG-ATTT--TTTTTT---ACAAAAAAAATTGCATG
                Ningyou-7                   TGCATGG-ATTT--TTTTTT---ACAAAAAAAATTGCATG
                Shengliyoucai-11            TGCATGG-ATTT--TTTTTT---ACAAAAAAAATTGCATG
                Darmor                      TGCATGG-ATTT--TTTTTT---ACAAAAAAAATTGTATG
                Major                       TGCATGG-ATTT--TTTTTT---ACAAAAAAAATTGTATG
                Tapidor                     TGCATGG-ATTT--TTTTTT---ACAAAAAAAATTGTATG
                Sensation New Zealand       TGCATGG-ATTT--TTTTTT---ACAAAAAAAATTGTATG
```

Figure 4. 12: Sequence variation at *BnaFLC.A10* identified by capillary sequencing. (A) Seven *B. napus* cultivars were selected for sequencing and their sequences compared to published reference sequences from Darmor-*bzh* (plants.ensembl.org), Tapidor *BnaFLC.A10* (GenBank accession JQ255381, Zou *et al*., 2012), Ningyou-7 *BnaFLC.A10* sequence (GenBank accession JX501141, Hou *et al*., 2012), Tapidor *BnaFLC.C09a* sequence (GenBank accession JQ255382, Zou *et al*., 2012). Sequence variation observed within the cultivars is listed in the table according to their position from the translational start site according to Zou *et al*., (2012), "-" denotes deletion, "N" denotes missing sequence. (B) Sequence comparisons of the RY elements in intron 1 of *BnaFLC.A10*. Both RY elements are highlighted in bold text, SNP484: T>C is highlighted in red/black, the intergenic region between the RY elements is also included. The *A. thaliana RY* elements found in wild-type (WT) and *vrn8* mutants (as reported in Questa *et al*., 2016) are included for comparison.

The winter cultivars Damor, Major and Tapidor and the swede cultivar Sensation New Zealand exhibit no sequence variation at *BnaFLC.A10* compared with the published *BnaFLC.A10* sequence from Tapidor (Zou *et al*., 2012). However, two SNPs (SNP9:A>G and SNP29:C>G) and two single base InDels (InDel461:T>- and InDel372:A>-) distinguish Darmor-*bzh* from these four cultivars. Interestingly Schiessl *et al*., (2017a) identified the SNP29:C>G, which causes an amino acid change R10P, using targeted bait sequencing. This SNP29:C>G was found to be correlated with the winter-spring crop type split in *B. napus* however this SNP is not present within the seven cultivars tested. Although I analysed a small sample of cultivars, an explanation for this could be that, due to high sequence similarity between *BnaFLC* paralogues, miss-mapping of sequencing reads could erroneously identify SNPs. The SNP9:A>G and SNP29:C>G are two of seven coding SNPs that distinguish *BnaFLC.A10* from *BnaFLC.C9a* (Zou *et al*., 2012, Figure 4. 15 A). This requires further exploration; sequence analysis of more cultivars would reveal whether the SNP29:C>G identified by Schiessl *et al*., (2017a) is present within the OREGIN DFFS diversity set population and whether it correlates with flowering time. Due to the polymorphisms identified at Darmor-*bzh*, and all further sequence comparisons between the *B. napus* cultivars will be made to the published Tapidor *BnaFLC.A10* sequence (GenBank accession code JQ255381, Zou *et al*., 2012).

I detected three alleles of *BnaFLC.A10* within the seven cultivars. Darmor, Major, Tapidor and Sensation New Zealand carry identical alleles of *BnaFLC.A10* (hereafter referred to as *BnaFLC.A10-a*). Ten SNPs, and two InDels were detected within the Chinese semi-winter cultivars Ningyou-7 and Shengliyoucai-11 compared with the Tapidor reference sequence (*BnaFLC.A10-b*; Figure 4. 15 A), and all but one SNP had been previously published (Hou *et al*., 2012). An additional SNP66:T>C was detected in Ningyou-7 and Shengliyoucai that had not been previously identified, while SNP2180:A>G from the published Ningyou-7 (GenBank accession JX901101; Hou *et al*., 2012) was not detected. The spring oilseed rape cultivar Karoo-057 carried a different allele of *BnaFLC.A10* (*BnaFLC.A10-c*; Figure 4. 15 A). *BnaFLC.A10-c* exhibits high sequence similarity to *BnaFLC.A10-b* but

carries an additional SNP58:A>C and variation for insertion size within intron 1. The 11bp insertion located in intron 1 of *BnaFLC.A10-b* is 8bp shorter than in *BnaFLC.A10-c* and absent in *BnaFLC.A10-a*. The sequencing results suggest that cultivars without a vernalisation requirement, such as Karoo-057, Ningyou-7 and Shengliyoucai, carry similar alleles at *BnaFLC.A10* that are different from cultivars with a vernalisation requirement like Darmor, Major, Tapidor and Sensation New Zealand.

A notable SNP identified at *BnaFLC.A10* was SNP484:T>Cfound in *BnaFLC.A10-b* and *BnaFLC.A10-c* (Figure 4. 12**)**. This SNP is located within the second RY element of intron 1 and is similar a polymorphism characterised in *A. thaliana* (Figure 4. 12 B**;** SNP585:C>T in the first RY element of *FLC* intron 1; Questa *et al*., 2016). To determine the frequency of this mutation within the *B. napus* diversity set I designed a CAPs marker analysis to target the SNP484:T>C; cultivars that carry a Thymine (T) at this position would not be digested by the NspI enzyme, compared with cultivars that carry a Cytosine (C) at the same position (see Section 2.5.5.1 for method details). 95 cultivars were screened for the SNP (cultivars listed in Table 2. 4; Figure 4. 13 A, Figure 4. 13 C and Figure 4. 13 E). A total of 66 *B. napus* cultivars within the OREGIN DFFS diversity set population were found to carry a Thymine at position 484 (Figure 4. 13 A, Figure 4. 13 C and Figure 4. 13 E). PCR products amplified from 29 cultivars were either fully or partially digested by the NspI enzyme indicating they carry a Cytosine in a homozygous or heterozygous state at position 484 from the translational start codon of *BnaFLC.A10* (Figure 4. 13 A, Figure 4. 13 C and Figure 4. 13 E).

Cultivars that carry the SNP484:T>C were digested by the NspI enzyme to produce two products approximately 500bp in size. One cultivar, a kale type called Siberische Boerenkool, (Cultivar 8; Figure 4. 13 A) was partially digested by the NspI enzyme, however the product was larger than the expected size; approximately 750bp and not 500bp. This perhaps indicates the presence of the SNP484:T>C in Sibberische Boerenkool but also the possibility that an insertion is present somewhere within the exon 1 and intron 1 region and requires further exploration.

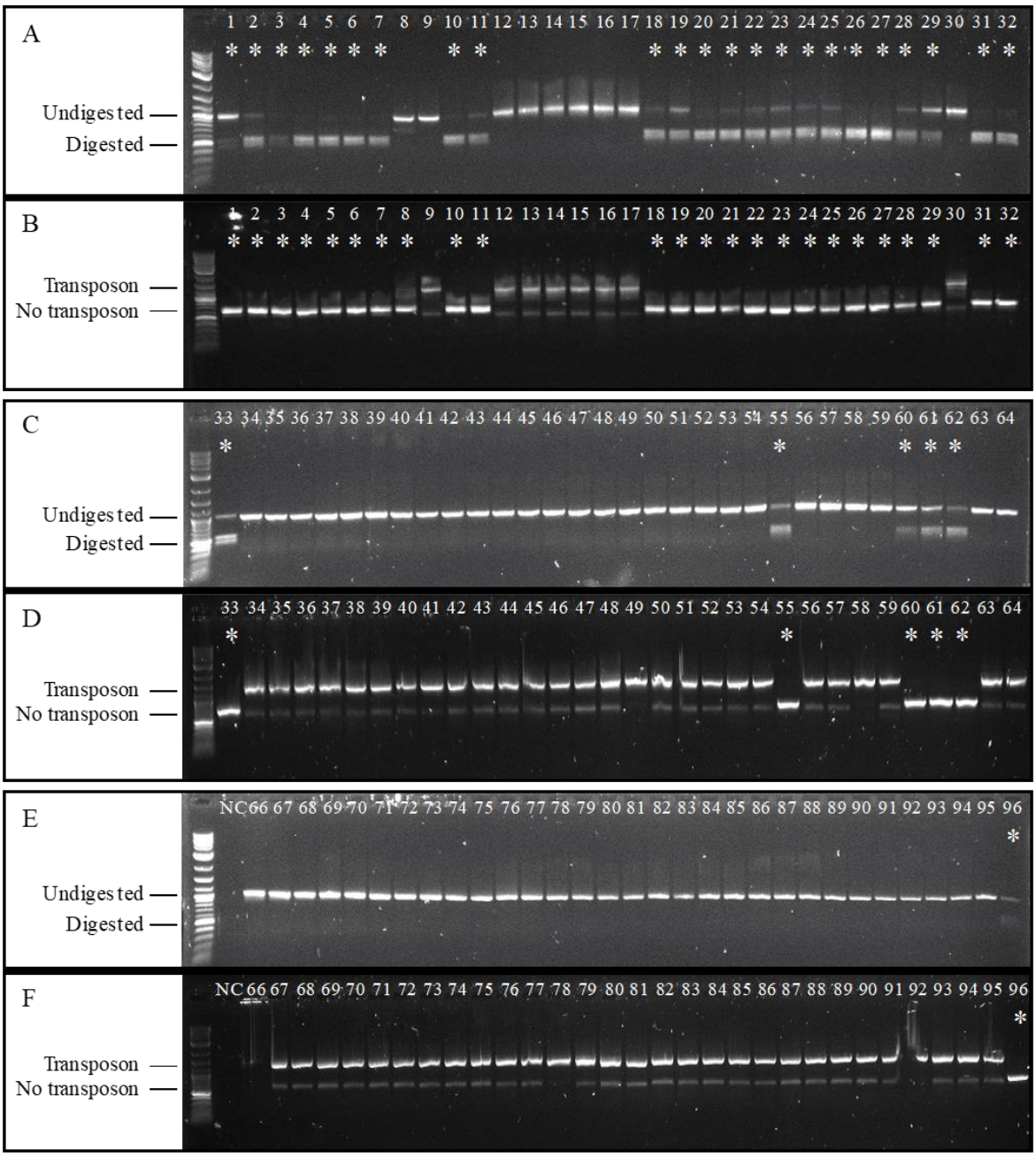


Figure 4. 13: Allelic variation is detected at *BnaFLC.A10* in the OREGIN DFFS population. 95 *B. napus* cultivars were analysed and are labelled 1-96 according to Table 2. 4, a negative control is included (NC). (A, C, E) PCR screen for the T484C SNP. PCR products from *B. napus* cultivars with the SNP484:T>C and digested by the NspI enzyme are labelled with *. (B, D, F) PCR screen for a tourist-like mite transposon. *B. napus* cultivars that do not carry a transposon upstream of *BnaFLC.A10* produce smaller PCR products compared with cultivars that carry the transposon; cultivars without the transposon are labelled with *.

4.2.6. A Tourist-like mite transposon is in linkage with SNP484:T>C

Hou *et al*., (2012) published the presence of a tourist-like mite transposon upstream of *BnaFLC.A10* and demonstrated a strong correlation between the presence of the transposon and vernalisation requirement in *B. napus*. I used the published PCR marker for this transposon to screen the same 95 *B. napus* cultivars as analysed in Section 4.2.5. for presence/absence of the transposon at *BnaFLC.A10* (see Section 2.5.5.2 for method details; Figure 4. 13 B, Figure 4. 13 D, Figure 4. 13 F). The transposon was detected in 64 cultivars, while not detected in 30 cultivars. Huron x Navajo, (Cultivar 66; Figure 4. 13 F) was excluded from further analysis as the PCR reaction failed.

Analysis of the distribution of the presence/absence of the transposon and the SNP484:T>C detected a close association between both features (χ^2=113.92, df=1, p<0.001). Excluding Siberische Boerenkool, all cultivars that do not have the transposon have the T484C SNP. The sequencing results in Figure 4. 12 indicate all cultivars that have the T484C SNP have nine additional SNPs compared with *BnaFLC.A10-a*. These nine SNPs are therefore also in linkage with the absence of the transposon. Cultivars that carry the transposon, and the linked SNPs, flower significantly later (p<0.001, Mann-Whitney U test, Figure 4. 14 B) and have significantly higher *in silico BnaFLC.A10* expression (p<0.001, Mann-Whitney U test, Figure 4. 14 A) compared with cultivars that do not have the transposon.

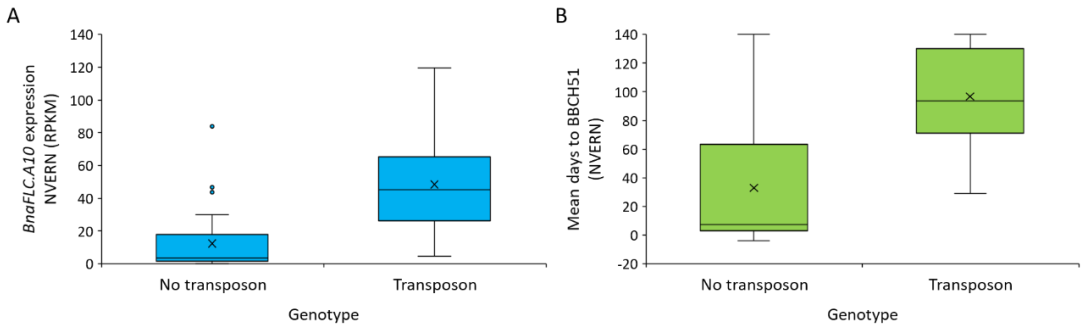


Figure 4. 14: The tourist-like mite transposon associates with flowering time and *BnaFLC.A10* expression. (A) *In silico* expression level variation observed between the cultivars with and without the transposon. (B) Flowering time (BBCH51 NVERN) variation observed between the cultivars with and without the transposon (DNF=140 days). Boxplots illustrate the median, mean (x) and range of values measured.

The SNP marker that most significantly associated with variation in *BnaFLC.A10* expression, as identified by the Associative Transcriptomics analysis (Table 4. 4), was JCVI_10962:251. I therefore analysed the sequence of unigene JCVI_10962 from all cultivars (Lu *et al*., 2014b) for SNP variation. Of the 77 *B. napus* cultivars that had adequate sequence coverage for unigene JCVI_10962, 28 cultivars carried a Thymine (T) and 49 cultivars a Cytosine (C) at position 251 in unigene JCVI_10962. Distribution of SNP JCVI_10962:251 among the cultivars was therefore compared with the pattern of segregation detected for the presence/absence of the transposon and SNP484:T>C. All three sequence features exhibited significantly similar segregation patterns, and therefore are assumed to be genetically linked (Table 4. 5, χ^2=141.08, $p<0.01$, H_0=random assortment of markers).

Table 4. 5: Chi-square analysis of linkage between unigene marker JCVI_10962:251, the tourist-like mite transposon and the SNP484:T>C.

Genotype	JCVI_10962:251	Transposon	SNP484:T>C	#Observed Cultivars	#Expected Cultivars	χ^2
1	T	NO	C	24	12.5	10.580
2	C	NO	C	2	12.5	8.820
3	T	YES	T	2	12.5	8.820
4	C	YES	T	46	12.5	89.780
5	T	NO	T	0	12.5	12.500
6	C	NO	T	1	12.5	10.580
					TOTAL	141.08

4.2.7. Coding variation at *BnaFLC.A10* associate with *BnaFLC.A10* expression variation

In addition to the SNP marker at unigene JCVI_10962:251, SNP markers present within JCVI_22584, the unigene that corresponds to *BnaFLC.A10*, were also detected by Associative Transcriptomics but to a lower significance level (Figure 4. 11 B). I compared the JCVI_22584 unigene sequence from all *B. napus* cultivars from the mRNA-seq database (Lu *et al*., 2014b) with seven *B. napus* cultivars sequenced by capillary sequencing in Figure 4. 12 A. Eight SNPs were detected at unigene JCVI_22584, two of which (SNP58:A>C and SNP66:T>C) were confirmed in the seven cultivars (Figure 4. 15). The remaining six SNPs were not detected however all are SNPs that distinguish *BnaFLC.A10* from *BnaFLC.C09a*. This further confirms the presence of mis-mapping of *BnaFLC* paralogues within the Associative Transcriptomics pipeline.

A

		Position							
	From *BnaFLC.A10* ATG	-19	9	29	30	58	66	154	156
	JCVI_22584	137	164	184	185	213	221	309	311
	JCVI_22584 SNP call	A/C	G/A	G/C	A/G	A/C	T/C	G/A	G/A
	JQ255381_*BnaFLC.A10*	A	G	G	A	A	T	G	G
	JQ255382_*BnaFLC.C9a*	C	A	C	G	A	C	A	A
Line	Karoo-057	A	G	G	A	C	C	G	G
Line	Ningyou-7	-	G	G	A	A	C	G	G
Line	Shengliyoucai-11	A	G	G	A	A	C	G	G
Line	Darmor	A	G	G	A	A	T	G	G
Line	Major	A	G	G	A	A	T	G	G
Line	Tapidor	A	G	G	A	A	T	G	G
Line	Sensation New Zealand	A	G	G	A	A	T	G	G
						*	*		

B

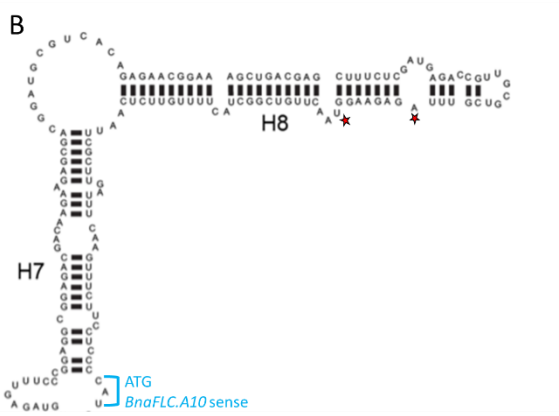


Figure 4. 15: SNPs at unigene JCVI_22584 (A) The unigene JCVI22584 sequence was compared with *BnaFLC.A10* and *BnaFLC.C09a* (Zou *et al*., 2012) and seven *B. napus* cultivars. Eight SNPs were detected for unigene JCVI_22584 across the *B. napus* diversity panel, their position relative to the *BnaFLC.A10* translational start site is indicated. SNPs confirmed by PCR and capillary sequencing are highlighted with *. (B) The distal terminal exon of *COOLAIR* overlaps with *BnaFLC.A10* sense exon 1. The position of SNPs 58:A.C and 66:T>C found within the H8 helix of distal *COOLAIR* are highlighted with *. The figure was taken and modified from Dr E. Hawkes Thesis (Hawkes (2017)).

Both SNPs, 58:A>C and 66:T>C, are located within exon 1 of *BnaFLC.A10*. SNP58:A>C causes a non-synonymous mutation, converting a Threonine to Proline at amino acid position 30. SNP66:T>C however is synonymous and causes no amino acid change. According to Hawkes (2017), exon 1 of *BnaFLC.A10* sense overlaps with the terminal exon of distal *COOLAIR*. Both SNPs, 58:A>C and 66:T>C, appear within highly reactive regions of the H8 helix of *BnaFLC.A10* distal *COOLAIR* (Figure 4.15 B). The secondary structure of *COOLAIR* is relatively conserved among the Brassicaceae (Hawkes *et al*., 2016) and allelic variation at *COOLAIR* is correlated with variation in pre-vernalisation expression level of *FLC* and flowering time (Li *et al*., 2014, Hawkes *et al*., 2016, Hawkes 2017). The possible association between allelic variation at regions that are important for *COOLAIR* secondary structure and the pre-vernalisation expression level of *BnaFLC.A10,* , provides a mechanism for how cis-variation at *BnaFLC.A10* may be important for modifying its expression.

Taken together five sequence features; JCVI_10962:251, a tourist-like mite transposon, SNPs 58:A>C, 66:T>C and 484:T>C; form significant associations with variation in *BnaFLC.A10* expression. After combining all information for the five sequence features the *B. napus* diversity set population can be divided into at least seven groups (Table 4. 6). Analysis of the *BnaFLC.A10* expression in these groups revealed two higher order groupings or haplotypes (Figure 4. 16); groups 1, 2 and 3 exhibit low *BnaFLCA10* expression and groups 4, 5, 6 and 7 exhibit high expression levels in comparison. The sequence feature that distinguishes these two haplotype groups is SNP66:T>C in *BnaFLC.A10* (Table 4. 6) providing a strong case that polymorphisms within *COOLAIR* correlate with expression variation at *BnaFLC.A10*.

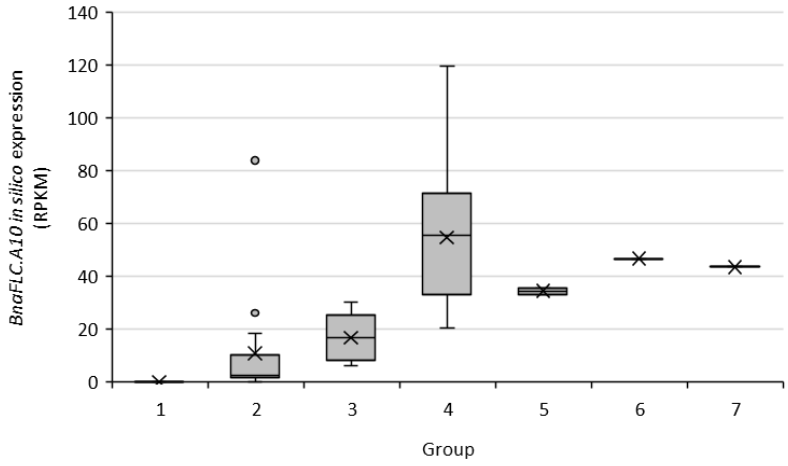


Figure 4. 16: *BnaFLC.A10 in silico* expression by sequence feature group. Groups were assigned according to the genotype of unigene marker JCVI_10962:251, the presence/absence of a tourist-like mite transposon, SNP58:A>C, SNP66:T>C, and SNP484:T>C.

Table 4. 6: The presence and nature of five sequence features on chromosome A10; Unigene SNP markers JCVI_10962:251, presence (YES)/absence (NO) of a tourist-like mite transposon; SNP58:A>C (A58C); SNP66:T>C (T66C); and SNP484:T>C (T484C) in 70 *B. napus* cultivars.

Name	JCVI_10962:251	TRANSPOSON	A58C	T66C	T484C	Group
IB111_N02D_1952	C	NO	A	C	C	1
D132_RAGJ	T	NO	A	C	C	2
D134_RAPCR	T	NO	A	C	C	2
A144_She	T	NO	A	C	C	2
NIN	T	NO	A	C	C	2
A135_Reg	T	NO	A	C	C	2
A18_Bro	T	NO	A	C	C	2
D173_TOP	T	NO	A	C	C	2
D181_WESD	T	NO	A	C	C	2
D31_CES	T	NO	A	C	C	2
IB110_N01D_1330	T	NO	A	C	C	2
IB49_Dup	T	NO	A	C	C	2
IB53_Erg	T	NO	A	C	C	2
D128_Q10	T	NO	A	C	C	2
D117_NOR	T	NO	A	C	C	2
D4_AMBxCOM	T	NO	A	C	C	2
A187_Xia	T	NO	C	C	C	3
A188_Zho	T	NO	C	C	C	3
A34_Chu	T	NO	C	C	C	3
D108_MON	T	NO	C	C	C	3
D83_KAR	T	NO	C	C	C	3
IB15_BraS	C	YES	A	T	T	4
D170_TIN	C	YES	A	T	T	4
D182_WILR	C	YES	A	T	T	4
D3_ALT	C	YES	A	T	T	4
IB141_SenNZ	C	YES	A	T	T	4
IB23_Can	C	YES	A	T	T	4
A11_Bie	C	YES	A	T	T	4
A136_Roc	C	YES	A	T	T	4
A168_TEM	C	YES	A	T	T	4
A176_Ver	C	YES	A	T	T	4
A25_Cap	C	YES	A	T	T	4
A42_Dar	C	YES	A	T	T	4
D10_BAL	C	YES	A	T	T	4
D102_MAT	C	YES	A	T	T	4
D127_PRI	C	YES	A	T	T	4
D131_Raf	C	YES	A	T	T	4
D137_ROCxLIZ	C	YES	A	T	T	4
D14_Bol	C	YES	A	T	T	4
D46_DIP	C	YES	A	T	T	4
D55_EUR	C	YES	A	T	T	4
D6_APE	C	YES	A	T	T	4
IB100_Maj	C	YES	A	T	T	4
IB120_Palm	C	YES	A	T	T	4
IB125_POH285B	C	YES	A	T	T	4
IB130_Qui	C	YES	A	T	T	4
IB133_Ram	C	YES	A	T	T	4
IB139_Sam	C	YES	A	T	T	4
IB143_ShaxWin	C	YES	A	T	T	4
IB150_SlaS	C	YES	A	T	T	4
IB151_SloK	C	YES	A	T	T	4
IB16_Bra	C	YES	A	T	T	4
IB169_TeqxAra	C	YES	A	T	T	4
IB180_Vis	C	YES	A	T	T	4
IB21_Cab	C	YES	A	T	T	4
IB24_CanxCou	C	YES	A	T	T	4
IB28_Cas	C	YES	A	T	T	4
IB29_Cat	C	YES	A	T	T	4
IB57_Exc	C	YES	A	T	T	4
IB58_Exp	C	YES	A	T	T	4
IB60_Fla	C	YES	A	T	T	4
IB70_HanxGas	C	YES	A	T	T	4
IB86_LemM	C	YES	A	T	T	4
IB91_LicxExp	C	YES	A	T	T	4
TAP	C	YES	A	T	T	4
IB164_Tai	C	YES	A	T	T	4
IB106_MoaR	T	YES	A	T	T	5
IB121_Palu	T	YES	A	T	T	5
D26_CAPxMOH	C	NO	A	T	C	6
D146_SIBB	C	NO	A	T	T	7

4.3. Discussion

In the present Chapter I have demonstrated that a large proportion of the NVERN flowering time variation that exists in *B. napus* can be explained by variation in the pre-vernalisation expression level of *BnaFLC.A10*. I used Associative Transcriptomics to identify gene expression variation, in the form of GEMs, and DNA sequence variation, in the form of SNP markers, that associated significantly with flowering time. This analysis revealed *BnaFLC.A10* as a candidate gene for flowering time variation in *B. napus*. *BnaFLC.A10* expression was shown to vary between cultivars of *B. napus* and this variation correlated significantly with flowering time. Several SNP markers formed a significant association with *BnaFLC.A10* expression, the strongest association was observed on chromosome A10, with linkage to *BnaFLC.A10*. This suggests that cis-variation at *BnaFLC.A10* may contribute to the pre-vernalisation expression level of *BnaFLC.A10*. The polymorphisms I identified at *BnaFLC.A10* strike parallels with those characterised in *A. thaliana* (Coustham *et al*., 2012, Li *et al*., 2014, Li *et al*., 2015, Hawkes *et al*., 2016, Questa *et al*., 2016) which provide a mechanistic understanding for how *BnaFLC.A10* expression has been fine-tuned to regulate flowering.

Many GWA studies have described the variation in flowering time present within *B. napus* (Xu *et al*., 2015, Schiessl *et al*., 2015, Raman *et al*., 2016, Wang *et al*., 2016a, Zhou *et al*., 2017). These studies have focused on the flowering time variation present within populations of predominantly Chinese semi-winter oilseed rape (Xu *et al*., 2015, Raman *et al*., 2016, Wang *et al*., 2016a, Zhou *et al*., 2017), Australian oilseed rape (Raman *et al*., 2016) or in the case of Schiessl *et al*., (2015) a population of European winter type cultivars only. Here I have measured the flowering times of spring, Chinese semi-winter and winter oilseed rape, in addition to other *B. napus* crop types like winter exotic, kale, swede, synthetic, winter fodder, and winter vegetable cultivars. As has been reported by others (Xu *et al*., 2015, Schiessl *et al*., 2015, Raman *et al*., 2016, Wang *et al*., 2016a, Zhou *et al*., 2017), variation for vernalisation requirement and flowering time is present within *B. napus*. All spring oilseed rape and most Chinese semi-winter oilseed rape were early flowering without vernalisation. Winter oilseed rape cultivars in contrast exhibited a vernalisation

requirement and an acceleration in flowering time in response to vernalisation was detected. A broad range of flowering times were detected in winter oilseed rape under both NVERN and VERN treatments indicating variation for vernalisation requirement and possibly vernalisation response respectively. We could speculate from results published by Schiessl *et al*., 2015, that variation at *BnaFT*, *BnaCCA1* and/or *BnaFUL* may influence flowering time responses of winter oilseed rape, this is investigated further using a winter oilseed rape mapping population in Chapter 5. Most cultivars of the kale, winter fodder and swede crop types flowered late with vernalisation and did not flower without vernalisation. These cultivars are harvested for their vegetative tissue and have therefore been bred to have an obligate requirement for vernalisation. Flowering in these crops is disadvantageous, affecting the quality and yield of the product (Jung and Müller 2009).

Polyploidisation has occurred multiple times throughout plant evolution (Adams and Wendel 2005). Duplicated genes with redundant functions may undergo neo- or sub-functionalisation or be lost through pseudogenisation, as was demonstrated by *BnaFLC.C03b* (Moore and Purugganan 2005, Zou *et al*., 2012). In *B. napus*, nine copies of *BnaFLC* have been characterised (Zou *et al*., 2012) providing an interesting system to study the role of multiple *FLC* genes in the control of flowering time. Significant correlations were detected between the expression of five *BnaFLC* genes and flowering time. *BnaFLC.A10*, *BnaFLC.A03a*, *BnaFLC.A03b*, and *BnaFLC.C02*, exhibited positive correlations between expression level and flowering time supporting previous studies (Tadege *et al*., 2001, Schranz *et al*., 2002, Long *et al*., 2007, Hou *et al*., 2012, Zou *et al*., 2012, Ridge et a. 2015, Irwin *et al*., 2016, Schiessl *et al*., 2017a). Generally, cultivars that flowered early without vernalisation exhibited low expression of these four *BnaFLC* genes while cultivars that were late flowering cultivars exhibited high expression. The expression of *BnaFLC.A10*, *BnaFLC.A03a*, *BnaFLC.A03b*, *BnaFLC.C02* may therefore contribute to flowering in a dosage dependent manner. *BnaFLC.C09b* however exhibited a significant negative correlation with flowering time, and for this gene the early flowering cultivars exhibited the highest expression. This difference in the effect of *BnaFLC* expression level for *BnaFLC.C09b* perhaps indicates this gene has undergone neo- or sub-functionalisation, no longer acting to inhibit flowering. Perhaps most interesting

was the revelation that the cumulative expression of all *BnaFLC* genes did not significantly correlate with flowering time, suggesting not all *BnaFLC* genes contribute additively to control flowering time.

Of all *BnaFLC* genes, variation in *BnaFLC.A10* expression associated most strongly with variation in flowering time suggesting the pre-vernalisation expression level of *BnaFLC.A10* can have a significant influence on flowering time, even after vernalisation. Cultivars with contrasting flowering time phenotypes exhibited differences in their pre- and post-vernalisation expression levels of *BnaFLC.A10.* As described in *A. thaliana* (Coustham *et al* 2012, Li *et al*., 2014a, Questa *et al*., 2016) and in *B. oleracea* (Irwin *et al*., 2016) variation in *BnaFLC.A10* expression dynamics may therefore have contributed to the variation in flowering times between cultivars after vernalisation.

In agreement with previous reports (Tadege *et al*., 2001), *BnaFLC.A10* may play a more dominant role in regulating flowering time compared with other *BnaFLC* genes in *B. napus*. According to Hawkes (2017), *COOLAIR*, the antisense transcript expressed at the *BnaFLC.A10* locus, has evolved differently to *COOLAIR* found at other *BnaFLC* genes. The *COOLAIR* expressed from the *BnaFLC.A10* locus contains a splice site mutation that gives rise to a *COOLAIR* isoform similar to those found in the late flowering *A. thaliana* accession Var2-6. This isoform of *COOLAIR* found in Var2-6 is hypothesised to give rise to higher *FLC* sense expression in *A. thaliana.* Interestingly the *BnaFLC.A10-a* allele, when transformed into *A. thaliana*, conferred late flowering, however mutation of the *COOLAIR* splicing pattern at *BnaFLC.A10* conferred early flowering (Hawkes 2017). As no other *BnaFLC* gene was observed to produce this isoform of *COOLAIR* Hawkes (2017) suggested that *BnaFLC.A10* could have a stronger influence on flowering time. The difference in *COOLAIR* secondary structure, and therefore its regulation of *BnaFLC.A10* sense expression, would impact flowering time. I did not detect variation for *COOLAIR* predicted splice sites (according to Hawkes 2017) between cultivars of *B. napus* that exhibited high and low expression of *BnaFLC.A10*, however two SNPs 58:A>C and 66:T>C formed an association with gene expression by Associative Transcriptomics.

Both mutations occur within exon 1 of *BnaFLC.A10* sense and therefore within the terminal exon of distal *COOLAIR.* SNP58:A>C causes an amino acid change from Threonine to Proline which may impact *BnaFLC.A10* protein structure and/or function but SNP66:T>C is synonymous and therefore unlikely to affect protein function. Both mutations occur within highly reactive regions of the H8 helix of distal *COOLAIR* secondary structure (Hawkes 2017) and may therefore alter *BnaFLC.A10* expression.

The tourist-like mite transposon at *BnaFLC.A10*, as first published by Hou *et al*., (2012), was detected in 64 *B. napus* cultivars of the OREGIN DFFS population and its presence was significantly associated with vernalisation requirement and high *BnaFLC.A10* expression. There were some exceptions to this such as the kale type Ragged Jack which exhibited high levels of *BnaFLC.A10* expression and flowered late without vernalisation, but did not carry the transposon. I hypothesis from this that high *FLC* expression and late flowering can be conferred independent of the transposon. According to Hou *et al*., (2012), the transposon arose during the evolution of *B. napus* and is absent in the homologue in *B. rapa*. Like *B. napus*, *B. rapa* exhibits a diverse range of flowering times and crop types therefore incorporation of the transposon upstream of *BnaFLC.A10* is perhaps not necessary for late flowering and high *BnaFLC.A10* expression, but likely genetically linked to a polymorphism that affects *BnaFLC.A10* expression.

Hou *et al*., (2012) did not detect significant differences in expression before vernalisation between Tapidor which carries the transposon and Ningyou-7 which does not but did report variation in expression after vernalisation. I found another mutation; SNP484:T>C within the second RY element of intron 1; that associated with early flowering. The mutation is similar to the *vrn8* mutation published by Questa *et al*., (2016) in *A. thaliana*. The wild-type sequence for the RY element is TGCATG and this sequence is recognised by the B3 binding domain protein VAL1. When mutated in *A. thaliana* (TGTATG) VAL1 binding was disrupted leading to lower accumulation of H3K27me3 chromatin marks and reduced down-regulation of *FLC* during cold (Questa *et al*., 2016). I speculate therefore that the SNP484:T>C

detected in *B. napus* would result in a similar consequence for *BnaFLC.A10* expression and H3K27me3 accumulation. Late flowering cultivars that carry an RY element like the *vrn8* mutant in *A. thaliana* (TGTATG) would exhibit a recution in the downregulation of *BnaFLC.A10* expression during cold. I observed four late flowering cultivars with a mutated RY element (TGTATG) exhibited higher expression levels of *BnaFLC.A10* after six-weeks vernalisation compared with those cultivars that carried wild-type (TGCATG) RY elements. This possibly indicates that *BnaFLC.A10* expression dynamics vary between the different alleles during vernalisation.

BnaFLC.A10 has been identified as a candidate for controlling flowering time in several *B. napus*, and its orthologue *BraFLC.A10* in *B. rapa*, biparental populations (Kole *et al*., 2001, Tadege *et al*., 2001, Long *et al*., 2007, Yuan *et al*., 2009, Hou *et al*., 2012, Zou *et al*., 2012). It is not surprising therefore that I identified *BnaFLC.A10* as a potential candidate for flowering time variation, even within the diverse panel of cultivars tested here. My analysis has revealed that cis-variation at *BnaFLC.A10* may contribute to expression variation which correlates significantly with variation in flowering time and vernalisation requirement in *B. napus*. I detected several polymorphisms at *BnaFLC.A10* and all were genetically closely linked. After grouping *B. napus* cultivars according to their polymorphisms one SNP, 66:T>C, differentiated high and low *BnaFLC.A10* expressing cultivars. However, due to the significant correlation between the sequence features identified, mutational and transgenic work is required to identify the causative polymorphism. Based on work in *A. thaliana*, the combination of SNPs at *BnaFLC.A10* are likely to influence the fine tuning of gene expression and ultimately flowering time. Each SNP at *BnaFLC.A10* will contribute to the regulation of gene expression before, during and after vernalisation and further work is required to elucidate the contribution of each individually. Identifying the cis-variation at *BnaFLC.A10* which likely contribute to variation in expression level would provide a mechanistic understanding why Breeders have selected for allelic variation at *BnaFLC.A10* to modify flowering time.

Chapter 5: Variation at *BnaFLC* and *BnaFT*, but not *BnaFRI*, associate with vernalisation requirement in winter oilseed rape

5.1 Introduction

In my previous Chapters I have shown *BnaFLC.A10* to be important for variation in flowering time with and without vernalisation (Chapter 4), while orthologues of *FRI* exhibited an inconsistent and minimal effect on flowering time in *B. napus* (Chapter 3). Analysis of a diverse population of *B. napus* for flowering time with and without vernalisation (Chapter 4) revealed winter oilseed rape cultivars exhibited variation for vernalisation requirement; some cultivars flowered early, and others flowered late with and without vernalisation. Interestingly winter oilseed rape exhibited allelic variation at *BnaFRI.C03* (Chapter 3), however based on previous results (Chapter 3 and Chapter 4) it is unlikely to contribute significantly to flowering time. The aim of my work in this Chapter was to identify and investigate the genetic variation important for variation in flowering time in the absence of vernalisation (the vernalisation requirement) and variation in flowering time after vernalisation (the vernalisation response) in winter oilseed rape.

Variation for vernalisation requirement and response have been investigated in three *Brassica* species; *B. rapa*, *B. oleracea* and *B. napus*. In *B. rapa* previous studies (Teutonico & Osborn 1995, Schranz *et al*., 2002, Ajisaka *et al*., 2001, Kole *et al*., 2001, Yuan *et al*., 2009, Zhao *et al*., 2010, Wu *et al*., 2012, Xiao *et al*., 2013) have identified orthologues of *FLC*, in particular *BraFLC.A10* and *BraFLC.A02*, as drivers for variation in vernalisation requirement and response. For example, a splice site mutation at *BraFLC.A10*, was found to be associated with flowering time variation in *B. rapa* as the mutation leads to the translation of a non-functional BraFLC.A10 protein (Yuan *et al*., 2009). At *BraFLC.A02* allelic variation (Wu *et al*., 2012) and expression variation (Zhao *et al*., 2010, Xiao *et al*., 2013) were associated with variation for vernalisation requirement and response.

Variation at *BolFLC.C02* is associated with variation in vernalisation requirement and response in *B. oleracea*, responsible for up to 65% of variation in flowering time (Camargo and Osborn 1996, Lin *et al*., 2005, Okazaki *et al*., 2007, Razi *et al*., 2008, Ridge *et al*., 2014, Irwin *et al*., 2016). A frame-shift mutation and copy number variation at *BolFLC.C02* is associated with variation in vernalisation requirement (Okazaki *et al*., 2007, Ridge *et al*., 2015). *BolFLC.C02* is also associated with variation in vernalisation response (Irwin *et al*., 2016). In late flowering cultivars of *B. oleracea* both alleles of *BolFLC.C02* produce functional proteins and promote late flowering in the absence of vernalisation, however variation in flowering time was detected after vernalisation. The variation in flowering time in response to vernalisation was attributed to cis-variation at *BolFLC.C02* influencing gene expression dynamics, similar to reports of cis-variation at *FLC* in natural accessions of *A. thaliana* (Coustham *et al*., 2012, Li *et al*., 2014, Questa *et al*., 2016).

Variation at *BnaFRI.A03* (Wang *et al*., 2011b, Yi *et al*., 2018), *BnaFLC.A10* (Long *et al*., 2007, Hou *et al*., 2012, Chapter 4), *BnaFLC.A02* (Raman *et al*., 2016). *BnaCO.C09* (Xu *et al*., 2015), *BnaFT.A02* and *BnaFT.C06* (Wang *et al*., 2009b) for example have been identified as probable candidates for variation in flowering time in *B. napus* however a mechanism for how these genes control flowering time is not well understood. Several studies have demonstrated variation for vernalisation requirement between different crop types of *B. napus* (Murphy *et al*., 1994, Ferreira *et al*., 1995, Tadege *et al*., 2001, Long *et al*., 2007, Hou *et al*., 2012, Raman *et al*., 2013, Schiessl *et al*., 2015, Raman *et al*., 2016), however to date none have investigated the vernalisation requirement within winter oilseed rape. In order for me to identify genes that are important for vernalisation requirement and response in winter oilseed rape an association mapping technique would be required. An alternative to traditional QTL analysis, which can be time-consuming and labour intensive, is QTL-seq (Takagi *et al*., 2013). Like other bulk segregant analysis (BSA) approaches (Giovannoni *et al*., 1991; Michelmore *et al*., 1991) a segregating population is analysed for deviations in allele frequencies in two pooled DNA samples made up of a sub-set of individuals that show contrasting phenotypes.

Unlike traditional BSA approaches, QTL-seq uses whole genome sequencing to generate SNP markers in order to identify genetic variation that is linked to a trait of interest. Since the proof of concept method was published in rice, QTL-seq has been used to identify QTLs for flowering time in cucumber (Lu *et al*., 2014a), seed weight in chickpea (Das *et al*., 2015), fruit weight and locule number in tomato (Illa-Berenguer *et al*., 2015) and branch angle in *B. napus* (Wang *et al*., 2016b).

The aim of my work in this Chapter was to identify and investigate the genetic variation important for variation in flowering time in the absence of vernalisation (vernalisation requirement) and flowering time after vernaisation (vernalisation response) in winter oilseed rape. I assessed the flowering times of five winter oilseed rape cultivars under a range of vernalisation treatments, confirming the presence of variation for flowering time in the absence of vernalisation and after vernalisation. To investigate the genetic variation important for the phenotypic variation, I generated an F_2 population from a cross between an early flowering cultivar called Cabriolet and a late flowering cultivar called Darmor. QTL-seq was used to identify associations between genetic variation and flowering time with and without vernalisation. A major QTL was identified on chromosome A02, a region that includes *BnaFLC.A02* and *BnaFT.A02*. DNA sequence and gene expression variation were detected at both genes, revealing parallels with natural accessions of *A. thaliana*. My work described in this Chapter begins to provide a mechanism for how *BnaFLC.A02* and *BnaFT.A02* may contribute to flowering time variation in winter oilseed rape.

5.2. Results

5.2.1. Winter oilseed rape exhibit variation for vernalisation requirement

I grew five winter oilseed rape cultivars named Cabriolet, Darmor, Major, Norin and Primor were grown under six vernalisation (5°C, 8-hour light/16-hour dark, 70% humidity) treatments; 12, 10, 8, 6, 4 and 0 (no vernalisation = NV) weeks. Plants were grown under glasshouse conditions (18°C day/15°C night, 16-hour light/8-hour dark, 70% humidity) for 28 days before being transferred to a vernalisation chamber. Sowing was staggered so that all plants were removed from vernalisation on the same day and subsequently grown under glasshouse conditions. The cultivars were measured for two flowering time traits according to Meier *et al*., 2001; number of days to buds visible (BBCH51) and the number of days to opening of the first flower (BBCH60). Scoring commenced after the vernalisation treatment had ended and continued for 125 days, and plants that did not flower during the time-frame of the experiment were given a did not flower (DNF) value of 125 days (Figure 5. 1).

Flowering times ranged from 14 days to a did not flower score of 125 days. The number of days between buds visible (BBCH51) and first flower opening (BBCH60) after vernalisation (4-12 weeks) was consistent between cultivars (an average of16± 9 days). Three cultivars (Damor, Major and Norin) did not flower without vernalisation suggesting these lines have a strong requirement for vernalisation (Figure 5. 1 F). Cabriolet and Primor reached BBCH51 without vernalisation, however Cabriolet was the only cultivar to reach BBCH60 under NV conditions. An average of 21.8 days was measured between BBCH51 and BBCH60 for Cabriolet under no vernalisation conditions, indicating possibly a reduction in rate of flower development under these conditions. Although Primor did flower without vernalisation, reaching BBCH51 within an average of 119± 6 days, Cabriolet exhibited the weakest vernalisation requirement and flowered significantly earlier than all other cultivars without vernalisation (56.3±14.5 days for BBCH51 and 87±13 days for BBCH60; $p<0.001$, ANOVA Bonferroni Multiple Comparisons).

These data confirm my previous results (Chapter 4) that variation for vernalisation requirement is present in cultivars of winter oilseed rape.

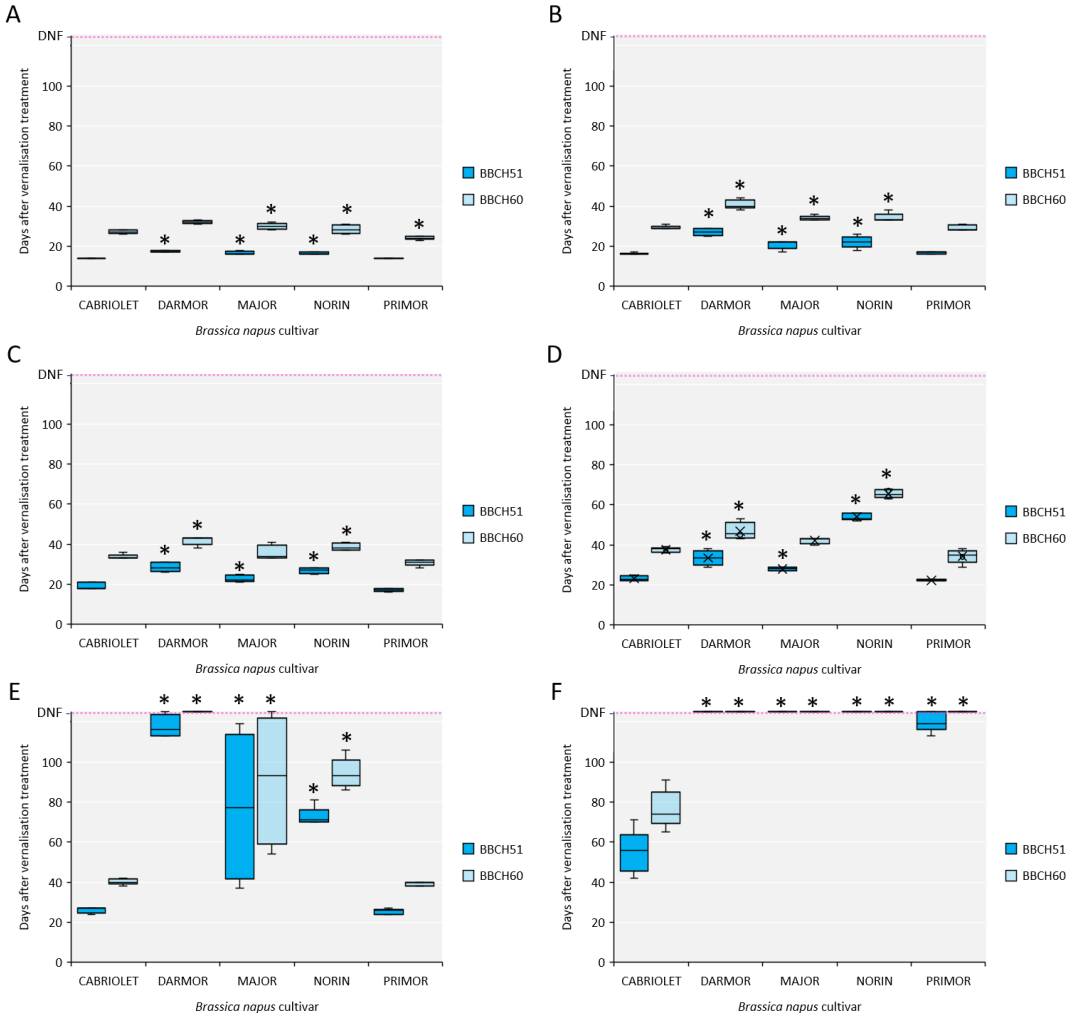


Figure 5. 1: Flowering times of five winter oilseed rape cultivars under six vernalisation treatments. (A) 12 weeks, (B) 10 weeks, (C) 8 weeks, (D) 6 weeks, (E) 4 weeks and (F) 0 weeks; vernalisation. Flowering times (BBCH51 and BBCH60) are displayed as boxplots indicating the median and range values for five replicate plants of each cultivar under each treatment. Both flowering time traits of each cultivar were compared with Cabriolet under each vernalisation treatment, cultivars that flowered significantly different ($p<0.05$; ANOVA Bonferroni Multiple comparisons) than Cabriolet are indicated with "*".

All winter oilseed rape cultivars exhibited an acceleration in flowering in response to vernalisation treatment, however the strength of the response varied between cultivars. Cultivars could be divided into two groups based on their BBCH51 flowering responses to vernalisation; with 4-12 weeks vernalisation Cabriolet and Primor flowered consistently earlier and more synchronous than Darmor, Major and

Norin ($p<0.001$; ANOVA and Bonferonni Multiple comparisons, Figure 5. 1). Of all five cultivars, Darmor flowered the latest after 4-weeks vernalisation treatment, reaching BBCH51 in an average of 119±6 days, while Major and Norin flowered earlier, reaching BBCH51 within an average of 89.5±35.5 and 91±5 days respectively. This suggests Darmor has the strongest requirement for vernalisation out of all three late flowering cultivars; Darmor, Major and Norin.

Cabriolet and Primor exhibited rapid flowering after all vernalisation treatments, possibly indicating weak vernalisation responses. Cabriolet flowered (BBCH51) 12 days later with 4-weeks compared with 12-weeks vernalisation, while Primor flowered (BBCH51) 11 days later with 4-weeks compared with 12. In contrast Darmor, Major and Norin exhibited a greated delay in flowering times between vernalisation treatments, possibly indicating a stronger vernalisation response. These cultivars flowered (BBCH51) on average 100, 61, and 51 days later respectively after 4-weeks compared with 12-weeks vernalisation treatment. Four weeks was therefore sufficient to promote rapid flowering in Cabriolet and Primor, but not for Darmor, Major and Norin demonstrating that variation for the length of cold required to promote early flowering is present in winter oilseed rape.

Variation in flowering times measured between replicate plants of the same cultivar increased with decreasing lengths of cold. This is especially apparent for Major after 4-weeks of vernalisation (Figure 5. 1 E) where flowering measurements between replicate plants ranged by more than 71 days for both BBCH51 and BBCH60. This may be due to my observation that plants of this cultivar exhibited loss of apical dominance after four weeks vernalisation. Plants developed multiple flowering stems and the main apical meristem could not be determined. Incomplete vernalisation, in this case, had an impact on flowering time and plant architecture.

5.2.2. Variation for flowering time maps to chromosome A02

Cultivars Cabriolet and Darmor exhibited contrasting flowering time phenotypes with and without vernalisation (Section 5.2.1.). Darmor flowered consistently and significantly later than Cabriolet under all vernalisation treatments tested. I wanted to identify and investigate the candidate genes underlying this variation for vernalisation requirement and response, therefore I generated an F_2 population and assessed flowering time under two vernalisation treatments. I used QTL-seq (Takagi *et al*., 2013) to identify genetic variation important for variation in vernalisation requirement and response between both cultivars. From this analysis I could identify candidate genes. All methods for this part of my work are described in Section 2.10.

5.2.2.1. Sequence variation is identified between Cabriolet, Darmor and Darmor-*bzh*

Following methods described in Section 2.10.4 DNA from Cabriolet and Darmor were sent for paired-end sequencing by Illumina® by Novogene Co., Ltd., HK at an average of 30x coverage. All sequencing quality control was performed by Novogene (see Table S. 3). 121,756,072 and 106,033,425 sequencing reads were generated for Cabriolet and Darmor respectively, both with a base error rate of 0.01% and a Q20 Phred score of approximately 96%. The GC content for both samples were approximately 37%, similar to previous reports (Wang *et al*., 2016b).

After aligning to the Darmor-*bzh* reference (Chalhoub *et al*., 2014, plants.ensembl.org), 12,522 high confidence (read depth>20, >95% base call) SNPs, small InDels (<9bp) and replacements bases ("N" call in Darmor-*bzh*) were identified in Darmor compared with Darmor-*bzh*. 9,561 of these could be anchored to a chromosomal position (outer circle of Figure 5. 2). Genetic variation between Darmor-*bzh* and Darmor was detected genome-wide but the chromosomes with the highest density of SNPs (and InDels) were chromosomes A04, A06, A08, A09, C01, C02, C03, and C09 (Figure 5. 3).

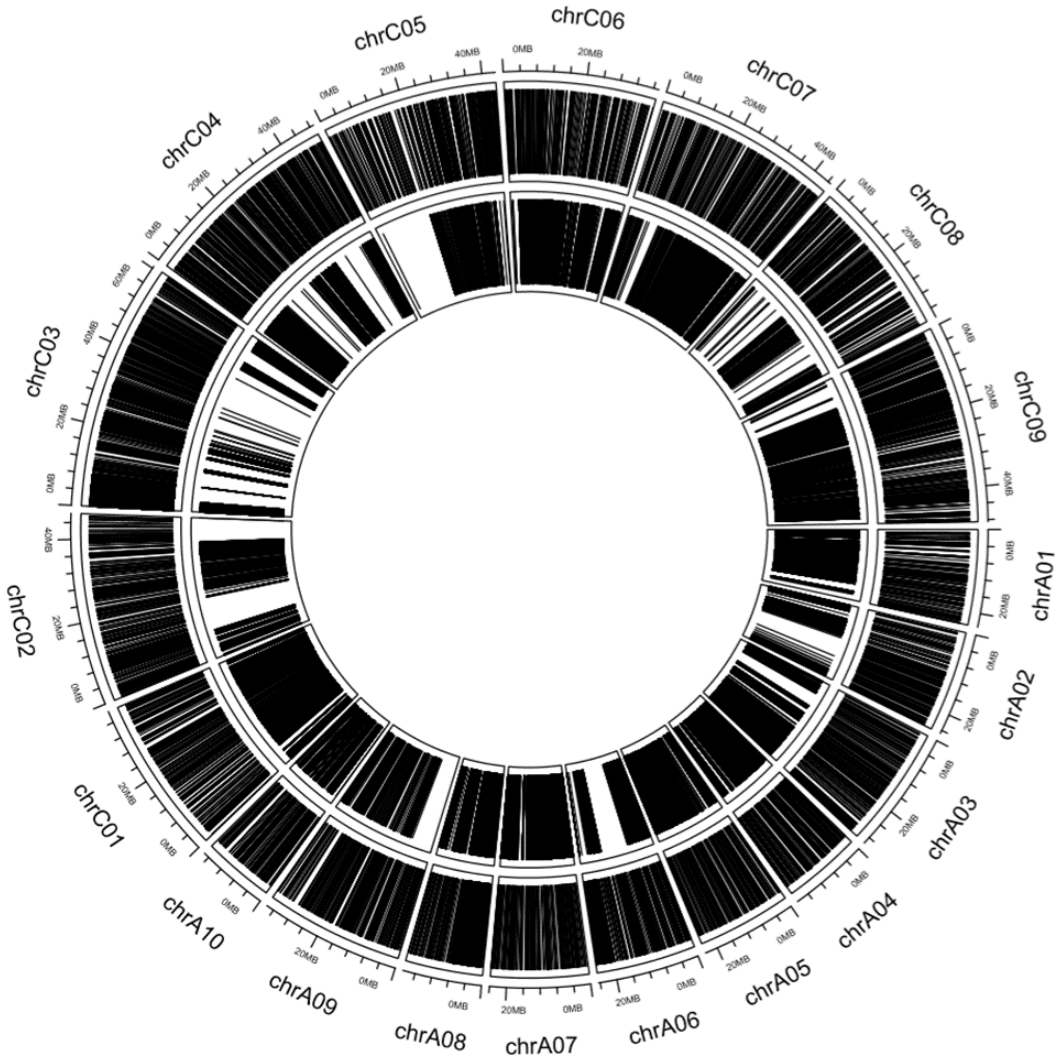


Figure 5. 2: Circlize plot illustrating the genomic distribution of anchored SNPs and small InDels in Darmor-*bzh*, Darmor and Cabriolet. (outer circle) Darmor compared with Darmor-*bzh*; (inner circle) Cabriolet compared with Darmor. Each SNP, small InDel is represented by a vertical line plotted against its chromosomal position according to the Darmor-*bzh* reference. Regions that are genetically identical are coloured white.

The variation detected between Darmor and Darmor-*bzh* was used to generate a pseudo-reference sequence identical to Darmor. Sequencing reads for Cabriolet were then aligned and scored for SNPs and small InDels (<9bp). A total of 39,717 high confidence SNPs and small InDels were identified in Cabriolet compared with Darmor and 32,773 could be anchored to a chromosomal position (inner circle of Figure 5. 2). The variation between Darmor and Cabriolet was not distributed genome-wide and appeared as blocks detected on all chromosomes. As is observed

in Figure 5. 2 (inner circle) multiple and sometimes large regions are present that are genetically identical between Darmor and Cabriolet. The filtering parameters used during this analysis were stringent; bases were called when read depth was above 20 and when the variant detected in Cabriolet was present in more than 95% of those reads. Regions of absolute genetic similarity may therefore be a consequence of insufficient read depth and/or variant data, not absence of variation *per se*.

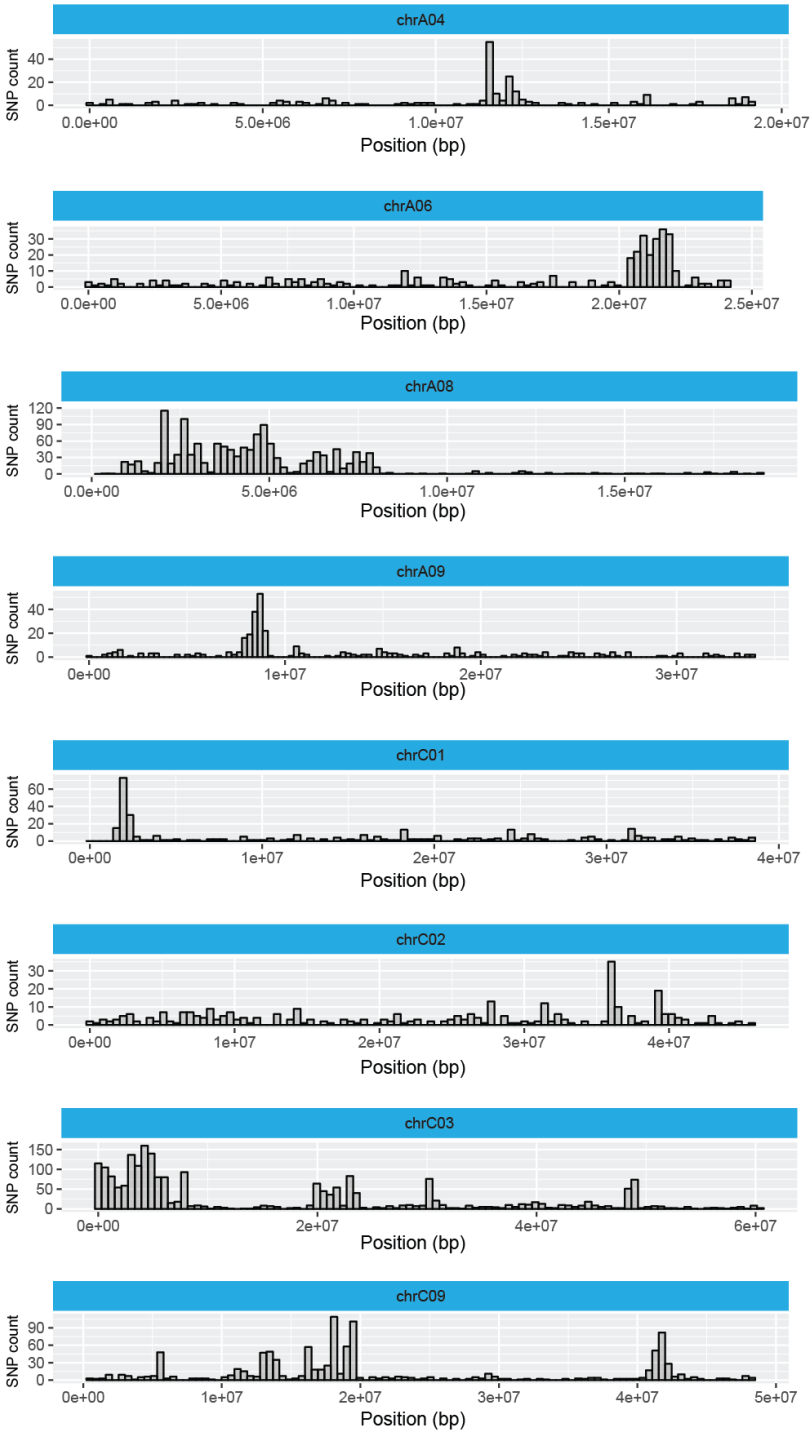


Figure 5. 3: Evidence of genetic variation between Darmor and Darmor-*bzh*. Each plot illustrates the number of SNPs detected per chromosome in Darmor compared with Darmor-*bzh*. Chromosomes A04, A06, A08, A09, C01, C02, C03 and C09 exhibited the highest density in SNP variation of all 19 chromosomes and each chromosome is plotted separately. The position (in bp) according to the Darmor-*bzh* is plotted along the x-axis, and the number of SNPs detected in Darmor compared with Darmor-*bzh* is plotted along the y-axis; note difference in y-axis scale between plots.

5.2.2.2. Variation for vernalisation requirement maps to chromosome A02

To determine the genetic variation important for variation in vernalisation requirement and response, I grew plants from an F_2 population, generated from a cross between the cultivars Cabriolet and Darmor under poly-tunnel conditions during the spring and summer of 2017 as described in Section 2.10.1 and Section 2.10.2. Replicate plants of cultivars Cabriolet and Darmor were also included in the experiment for comparison of flowering time. No vernalisation treatment was applied. 708 F_2 plants and 12 replicate plants each of Cabriolet and Darmor were measured for flowering time (number of days to first open flower; BBCH60 according to Meier *et al*., 2001 Figure 5.4 A**)**. Scoring of flowering time commenced 28 days after sowing, when plants were transferred to the poly-tunnel, and ended 170 days later. All plants that had not flowered within the time-frame of the experiment were given a did not flower (DNF) score of 170 days. On day 170 all plants that had not flowered were inspected for visible buds in the apex (BBCH51 according to Meier *et al*., 2001), giving an indication of whether the plant had undergone the floral transition or not.

As expected, from my previous data (Section 5.2.1) Darmor flowered significantly later than Cabriolet ($p<0.001$, Mann-Whitney U test) and did not flower within the time-frame of the experiment. Cabriolet plants in contrast flowered within 79.5±4.5 days while all replicate plants of Darmor were given a DNF of 170 days. No buds were visible within the apex of the Darmor plants.

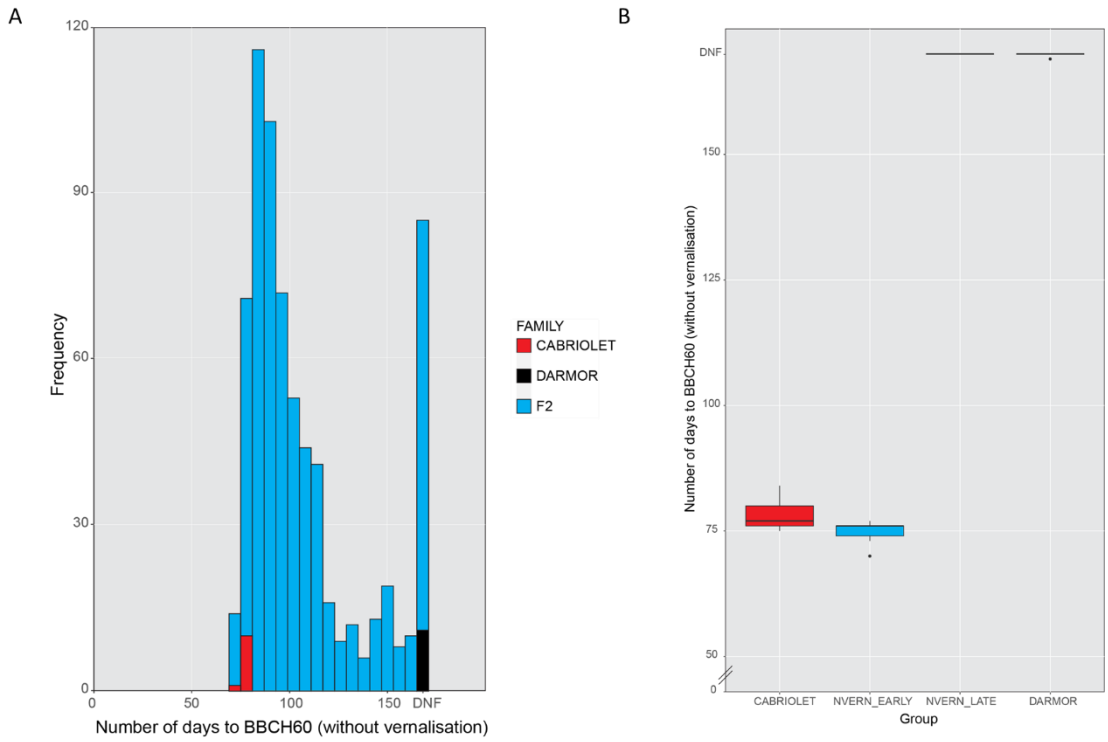


Figure 5. 4: Flowering time measurements under NVERN treatment scored as the number of days to BBCH60 (A) The flowering time distribution of 708 F_2 lines (blue), 12 replicate plants of Cabriolet (red), 12 replicate plants of Darmor (black). (B) Two pools of F_2 lines were submitted for whole genome sequencing; the flowering time measurements of these pools are illustrated compared with Cabriolet (red) and Darmor (black). Boxplots denote the median and range of flowering time measurements.

Flowering time measurements exhibited a bimodal distribution due to the arbitrary value given to those plants that did not flower (Figure 5. 4 A). Flowering time measurements within the F_2 population ranged between 70 and more than 170 days. Plants that did not flower within the timeframe of the experiment were given a did not flower (DNF) score of 170 days. 122 individual F_2 plants exhibited an early flowering phenotype like Cabriolet and flowered within 84 days while 86 individual F_2 plants exhibited a late flowering phenotype like Darmor and did not flower during the timeframe of the experiment. 500 individuals exhibited an intermediate phenotype suggesting that flowering time is controlled by multiple genes.

I used the flowering time data in Figure 5. 4 A to select F_2 individuals for sequencing as described in Section 2.10.3. Two DNA pools were generated and approximately

36 F_2 individuals, representing roughly 5% of the population, were included in each pool. The early flowering DNA pool (NVERN_EARLY) contained the earliest flowering lines while the late flowering pool (NVERN_LATE) contained only plants that had no buds visible at 170 days. Significant differences in flowering time measurements ($p<0.001$, Mann-Whitney U test) were detected between both pools. Their flowering time phenotypes compared with Cabriolet and Darmor are illustrated in Figure 5. 4 B.

The DNA from two pooled leaf samples (NVERN_EARLY and NVERN_LATE) were sent for paired-end sequencing by Illumina® by Novogene Co., Ltd., HK at an average 30x coverage. All sequence quality control results provided by Novogene Co., Ltd., HK, for each sequenced sample is listed in Table S. 3. 129,129,964 and 132,335,109 sequencing reads were generated for NVERN_EARLY and NVERN_LATE respectively and the GC content for both samples were approximately 37%.

After aligning the reads from each pool to the cured Darmor-*bzh* reference, Cabriolet variants were called, and a QTL-seq analysis was carried out (methods in Section 2.10.4) to compare NVERN_EARLY with NVERN_LATE (hereafter referred to as the NVERN comparison). 26,947 high confidence (read depth >20, base call >95% reads) SNPs and small InDels (<9bp) were detected in the NVERN comparison and 22,214 of these could be anchored to a chromosomal position. The results of this analysis are illustrated in Figure 5. 5 and Figure S. 4 - S. 43.

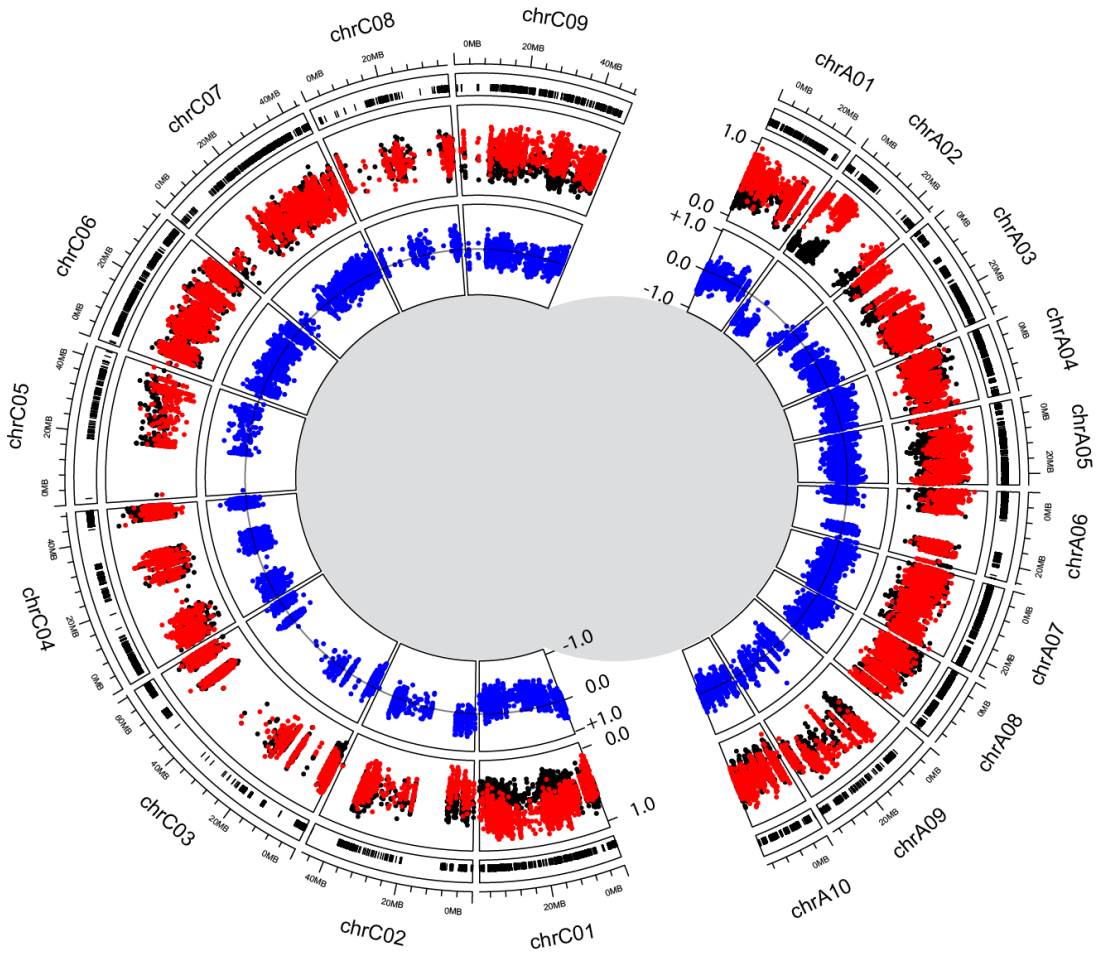


Figure 5. 5: QTL-seq results for the NVERN comparison. Circlize plot illustrating results of the QTL-seq analysis performed on pools NVERN_EARLY and NVERN_LATE. Outer circle: the distribution of Cabriolet SNPs and InDels detected in the pools plotted against the chromosomal position according to Darmor-*bzh.* Middle circle: the SNP index values calculated for each SNP and InDel in the Early pool (red) and the Late pool (black) plotted against the chromosomal position according to Darmor-*bzh*. Inner circle: the difference between SNP index values between the pools plotted as the ΔSNP index (blue) against the chromosomal position according to Darmor-*bzh*, a ΔSNP index equal to zero representing no deviation in allele segregation between the pools is plotted as a horizontal line. Each genome is plotted as separate half circles.

Deviation in SNP index values were detected between the DNA pools on chromosome A02 (Figure 5. 5). The direction of SNP index values suggests the late flowering pool (NVERN_LATE) contained sequence reads identical to Darmor (64 SNPs had a SNP index = 0.0) and the early flowering pool (NVERN_EARLY) contained sequence reads identical to the Cabriolet variant (37 SNPs had a SNP index = +1.0). The ΔSNP index values confirmed this by deviating away from the

expected value of no association (ΔSNP = 0.0) towards a value of -1.0. Allelic contribution from Darmor on chromosome A02, therefore conferred late flowering under NVERN conditions.

Analysis of chromosome A02 in more detail revealed a region approximately 10Mbp in size that associates with flowering time variation in the NVERN comparison (Figure 5. 6). This region encompasses, among other genes, *BnaFLC.A02* and *BnaFT.A02*. By calculating a sliding window average over 100 ΔSNP index values, the region surrounding *BnaFT.A02* appears to to associate more strongly with flowering time variation under NVERN conditions compared with *BnaFLC.A02* and is further investigation in Section 5.2.2.4. A list of all flowering genes located within the 10Mbp region on chromosome A02 are listed in Table 5. 1.

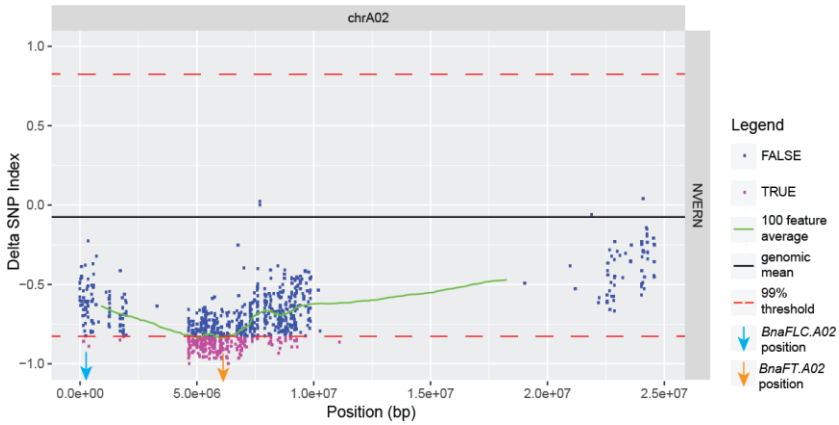


Figure 5. 6: A close-up image of the associating QTL on chromosome A02. Each point on the graph represents a Delta(Δ)SNP value calculated between DNA pools NVERN_EARLY and NVERN_LATE. Each ΔSNP value is plotted against their chromosomal position along the x-axis. ΔSNP index values that fall outside the 99% threshold of no association (red dashed line) are highlighted in pink, while the remaining ΔSNP index values are highlighted in blue. A sliding window average of 100 ΔSNP index values illustrated by the green line. The black line denotes the genomic mean ΔSNP index value. The relative position of *BnaFLC.A02* and *BnaFT.A02* are illustrated with blue and orange arrows respectively.

Table 5. 1: Candidate flowering time genes located in the genomic region associated with flowering time on chromosome A02. The chromosomal position, name according to Darmor-*bzh*, and the genetic status in Cabriolet compared with Darmor of each gene is listed. The homologous gene and function in *A. thaliana* is described for reference.

Position (bp)	**Gene name**	**Genetic status in Cabriolet**	**Homologue in *A. thaliana***	**Function in *A. thaliana***
134159-138121	BnaA02g00370d	Polymorphic	AT5G10140, *FLOWERING LOCUS C*, *FLC*	MADS-box transcription factor, repressor o flowering, responsive to vernalisation
408979-413943	BnaA02g01030D	Conserved	AT5G11530, *EMBRYONIC FLOWER 1*, *EMF1*	Involved in reproductive development
565722-567762	BnaA02g01270D	Conserved	AT5G12840, Nuclear transcription factor Y subunit A-1, *NFYA1*	Expressed in reproductive tissue
773658-778149	BnaA02g01670D	Conserved	AT5G13480, *FY*	Involved in regulation of flowering time, affects *FCA* mRNA processing
893325-894144	BnaA02g01960D	Conserved	AT5G14010, *KNUCKLES*, *KNU*	Transcription factor, mediates repression of *WUS* in floral meristem determinacy control
2115552-2118359	BnaA02g04770D	Conserved	AT5G20240, *PISTILLATA*, *PI*	Floral homeotic gene, MADS domain transcription factor, required for specification of petal and stamen identities
3111280-3113426	N/A	Conserved	AT5G60100, *Pseudo-response regulator 3*, *PRR3*	*PRR3* transcript levels vary in circadian pattern
3320312-3321741	BnaA02g07010D	Conserved	AT5G59560, *SENSITIVITY TO RED LIGHT REDUCED 1*, *SRR1*	Required for normal oscillator function during circadian rhythm
3685147-3687730	BnaA02g07770D	Conserved	AT5G58230, *MSI1*	Required for the transition to flowering

Position (bp)	**Gene name**	**Genetic status in Cabriolet**	**Homologue in *A. thaliana***	**Function in *A. thaliana***
3861836-3864833	BnaA02g08140D	Conserved	AT5G57380, *VERNALIZATION INSENSITIVE 1 VIN3*	Plant homeodomain protein, part of polycomb group complex of proteins, has a role in establishing *FLC* repression during vernalisation
5948452-5953119	BnaA02g11340D	Polymorphic	AT5G51230, *EMBRYONIC FLOWER 2, EMF2*	Polycomb group protein, a negative regulator of reproductive development
6375937-6378901	BnaA02g12130D	Polymorphic	AT1G65480, *FLOWERING LOCUS T, FT*	A promoter of flowering, expressed in leaves and is induced by long day treatment
8998899-9001658	BnaA02g15530D	Polymorphic	AT1G71800, *CSTF64*	RNA 3'-end-processing factor of antisense *FLC* transcript, mediates silencing of *FLC* gene
10804098-10805192	N/A	Conserved	AT4G20370, *TWIN SISTER OF FT, TSF*	A promoter of flowering and a homologue of *FT*, *FT* and *TSF* play overlapping roles in the transition to flowering

A minor QTL is also detected on chromosome C01 for the NVERN comparison Figure 5. 5 and Figure S. 14. The direction of SNP index values calculated suggests sequence reads in the NVERN_EARLY pool match Cabriolet loci and the sequence reads in the NVERN_LATE pool match Darmor loci. This suggests allelic contribution from Darmor on chromosome C01 conferred late flowering in the absence of vernalisation. The flowering time genes located on chromosome C01 are listed in Table 5. 2. Although allelic variation at *BnaFRI.C03* is detected between Cabriolet and Darmor (Chapter 3), this variation formed no significant association with flowering time under NVERN conditions.

Table 5. 2: Candidate flowering time genes located in the genomic region associated with flowering time on chromosome C01. The chromosomal position, name according to Darmor-*bzh*, and the genetic status in Cabriolet compared with Darmor of each gene is listed. The homologous gene and function in *A. thaliana* is described for reference.

Position (bp)	Gene name	Genetic status in Cabriolet	*A. thaliana* homolog	Function
872007-877899	BnaC01g01710D	Polymorphic	AT4G36920, *APETALA 2*, *AP2*	Floral homeotic gene, transcription factor involved in the specification of floral organ identity
1446633-1448582	BnaC01g02720D	Conserved	AT4G35900, *FD*	bZIP protein interacts with *FT*, mutants are late flowering
4360599-4364355	BnaC01g08200D	Conserved	AT4G30200, *VERNALISATION5/VIN3-LIKE 1*, *VEL1*	Part of the PRC2, involved in epigenetic silencing of *FLC*
7097688-7103238	BnaC01g11460D	Conserved	AT4G18960, *AGAMOUS*, *AG*	MADS domain transcription factor, specifies floral meristem identity
15020758-15025200	BnaC01g21540D	Polymorphic	AT4G16845, *VERNALISATION2*, *VRN2*	Involved in epigenetic silencing of *FLC*, mutants are late flowering
15309703-15318370	BnaC01g21860D	Conserved	AT4G16280, *FLOWERING CONTROL LOCUS A*, *FCA*	Involved in RNA-mediated silencing of *FLC*
32886488-32893069	BnaC01g33680D	Polymorphic	AT3G18990, *VRN1*	Required for vernalisation
38292886-38295651	BnaC01g39760D	Conserved	AT3G07650, *CONSTANS-LIKE 9*, *COL9*	Part of the CONSTANS gene family, downregulates CO, *FT* and SOC1

5.2.2.3. Variation for flowering time after vernalisation also maps to chromosome A02

In parallel to the experiment detailed in Section 5.2.2.2. the F_2 population generated from a cross between cultivars Cabriolet and Darmor were treated with 6-weeks vernalisation prior to growth under poly-tunnel conditions during the spring and summer of 2017 according to methods described in Section 2.10.1 and Section 2.10.2. Replicate plants of cultivars Cabriolet and Darmor were included in the experiment for comparison of flowering time measurements. I measured flowering time (days to first flower open, or BBCH60) for 704 F_2 individuals and 12 replicate plants each of Cabriolet and Darmor (Figure 5. 7 A). Scoring of flowering commenced at the end of the vernalisation treatment, 70 days after sowing when plants were transferred to poly-tunnel conditions and continued until all plants had flowered.

As expected from my previous data (Section 5.2.1) Cabriolet flowered significantly earlier than Darmor after 6-weeks vernalisation ($p<0.001$; Mann-Whitney U test). Cabriolet flowered within 43.5+/-2.5 days while Darmor exhibited a broader range of flowering times (69.5+/-13.5 days) indicative of incomplete vernalisation. The flowering time distribution of the F_2 population exhibited a right-hand skew, with flowering times ranging between 42 and 81 days (Figure 5. 7 A). 150 individual F_2 plants flowered early like Cabriolet, reaching BBCH60 within 46 days, while 104 individual F_2 plants flowered late like Darmor, reaching BBCH60 later than 55 days. 450 individuals, representing 64% of the population, exhibited an intermediate phenotype again indicating flowering time was controlled by more than one major gene.

Individual F_2 plants were selected for DNA extraction based on their flowering time phenotype to generate four pools (VERN_EARLY_1, VERN_EARLY_2, VERN_LATE_3 and VERN_LATE_4) that captured the earliest and latest flowering individuals of the population (see Section 2.10.3 for method details). Each pool represented approximately 5% of the population and their flowering time phenotypes are illustrated in Figure 5. 7 B. The pools exhibited significantly different flowering

times ($p<0.001$; Mann-Whitney U test) and the range of flowering time measurements in each pool did not overlap. Pools VERN_EARLY_2 and VERN_LATE_3 exhibited a less extreme flowering time phenotypes compared with VERN_EARLY_1 and VERN_LATE_4 respectively, this was to ascertain which, if any, regions of the genome were important for the extreme flowering times observed within the population.

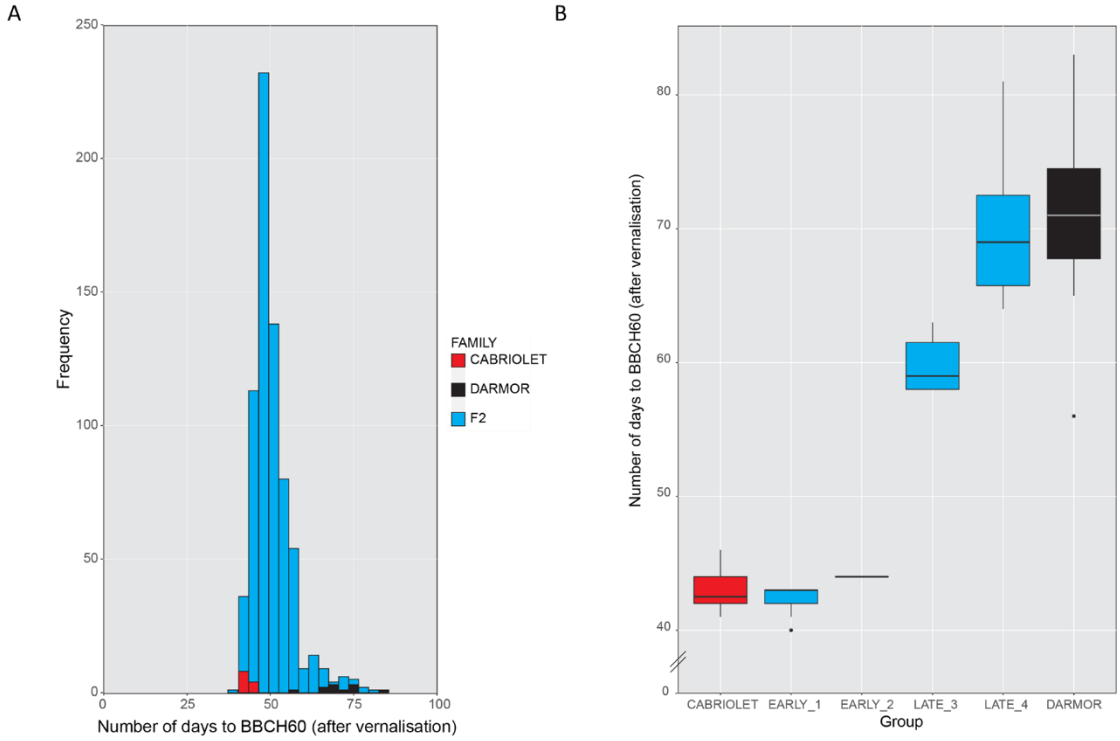


Figure 5. 7: Flowering time measurements under VERN treatment scored as the number of days to BBCH60. (A) The flowering time distribution of 704 F_2 individuals (blue), 12 replicate Cabriolet plants (red), and 12 replicate Darmor plants (black). (B) Four DNA pools were submitted for whole genome sequencing; the flowering time measurements of these four pools are illustrated compared with Cabriolet (red) and Darmor (black). Boxplots denote the median and range of flowering time measurements.

DNA from four pooled leaf samples (VERN_EARLY_1, VERN_EARLY_2, VERN_LATE_3, and VERN_LATE_4) were sent for paired-end sequencing by Illumina® by Novogene Co., Ltd., HK at an average 30x coverage. Sequencing

quality control was performed by Novogene Co., Ltd., HK (see Table S. 3 for details). 106,467,543, 113,252,999, 110,030,775, 110,166,633 sequencing reads were generated for VERN_EARLY_1, VERN_EARLY_2, VERN_LATE_3, and VERN_LATE_4 respectively and the GC content for all samples were approximately 38%.

After aligning the reads from each pool to the Darmor pseudo-reference, Cabriolet variants were called, and a QTL-seq analysis was carried out (according to methods described in Section 2.10.4) to compare VERN_EARLY_1 with VERN_LATE_4 (hereafter referred to as the VERN comparison), and to compare VERN_EARLY_2 with VERN_LATE_3 (hereafter referred to as the VERN INTERMEDIATE comparison). 17,754 and 20,018 high confidence (read depth >20, base call >95% reads) SNPs and small InDels (<9bp) were detected in the VERN, and the VERN INTERMEDIATE, comparisons respectively. For the VERN comparison 14,673 of these could be anchored to a chromosomal position, for the VERN INTERMEDIATE 16,542 could be anchored. The results of these analyses are illustrated in Figure 5. 8 and Figure 5. 9 and Figure S. 4 – S. 43.

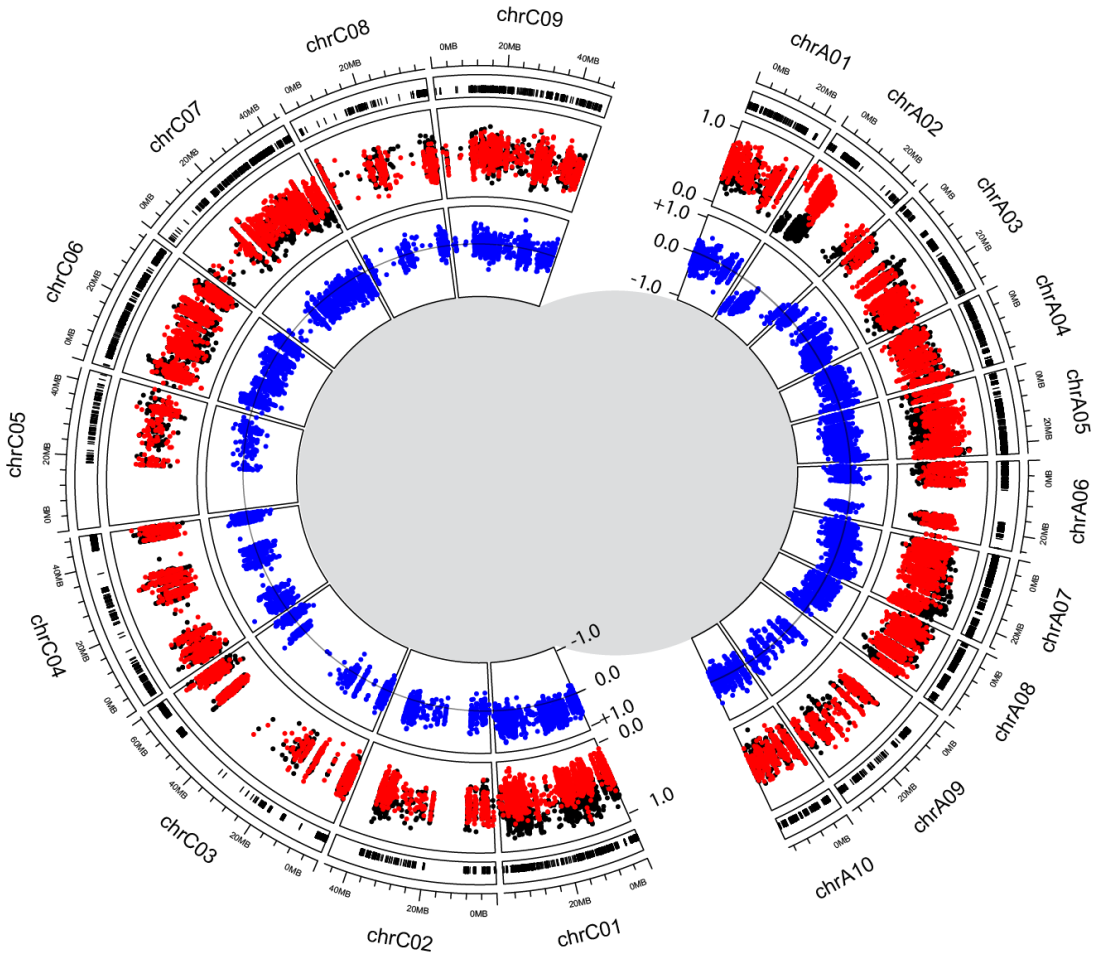


Figure 5. 8: QTL-seq results for the VERN comparison. Circlize plot illustrating results of the QTL-seq analysis performed on pools VERN_EARLY_1 and VERN_LATE_4 for the VERN comparison. Outer circle: the distribution of Cabriolet SNPs and InDels detected in the pools plotted against the chromosomal position according to Darmor-*bzh*. Middle circle: the SNP index values calculated for each SNP and InDel in the Early flowering pool (red) and the Late flowering pool (black) plotted against the chromosomal position according to Darmor-*bzh*. Inner circle: the difference between SNP index values between the pools plotted as the Delta(Δ)SNP index (blue) against the chromosomal position according to Darmor-*bzh*. A ΔSNP index equal to zero representing no deviation in allele segregation between the two pools is plotted as a horizontal line. Each genome is plotted as separate half circles.

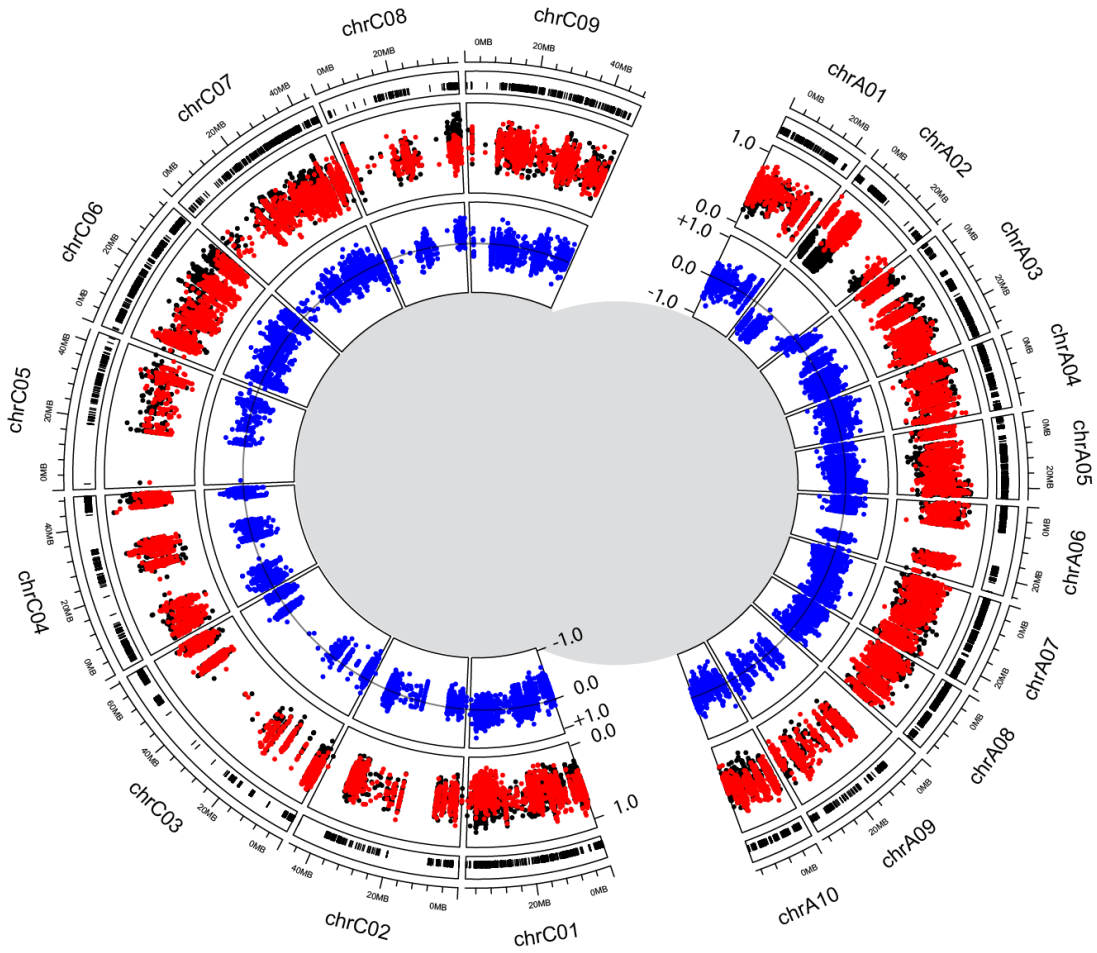


Figure 5. 9: QTL-seq results for the VERN_INTERMEDIATE comparison. Circlize plot illustrating results of the QTL-seq analysis performed on pools VERN_EARLY_2 and VERN_LATE_3. Outer circle: the distribution of Cabriolet SNPs and InDels detected in the pools plotted against the chromosomal position according to Darmor-*bzh*. Middle circle: the SNP index values calculated for each SNP and InDel in the Early flowering pool (red) and the Late flowering pool (black) plotted against the chromosomal position according to Darmor-*bzh*. Inner circle: the difference between SNP index values between the pools plotted as the Delta(Δ)SNP index value (blue) against the chromosomal position according to Darmor-*bzh*. A ΔSNP index equal to zero representing no deviation in allele segregation between the two pools is plotted as a horizontal line. Each genome is plotted as separate half circles.

Like in Section 5.2.2.2. deviations in SNP index values were detected between the pools on chromosome A02 for both the VERN and VERN INTERMEDIATE comparisons (Figure 5. 8 and Figure 5. 9). The early and late flowering lines segregated for loci on chromosome A02 and the direction of SNP index values suggest the late flowering pools (VERN_LATE_3 and VERN_LATE_4) contained sequence reads that matched Darmor (9 SNPs and 11 SNPs had SNP index values = 0.0 in the VERN and VERN INTERMEDIATE comparisons respectively) and the early flowering pools (VERN_EARLY_1 and VERN_EARLY_2) contained sequence reads that matched Cabriolet loci (8 SNPs and 1 SNP had SNP index values = +1.0 in the VERN and VERN INTERMEDIATE comparisons respectively). The ΔSNP index values confirmed this by deviating away from the expected value of no association (ΔSNP = 0.0) towards a value of -1.0. Allelic contribution from Darmor on chromosome A02, therefore confers late flowering after vernalisation. Analysis of chromosome A02 in more detail (Figure 5. 10) revealed the same 10Mbp region as identified in Section 5.2.2.2. associating with flowering time variation for both the VERN and VERN INTERMEDIATE comparisons. Again, this region encompasses, among other genes, *BnaFLC.A02* and *BnaFT.A02*. Interestingly for the VERN comparison, the sliding window average of 100 ΔSNP index values suggests the region at *BnaFLC.A02* associates more strongly with flowering time after vernalisation, while the region at *BnaFT.A02* associated less strongly. Variation at both *BnaFLC.A02* and *BnaFT.A02* contributed relatively equally to variation in flowering time in the VERN_INTERMEDIATE comparison. A list of flowering genes located within the region on chromosome A02 are listed in Table 5. 1.

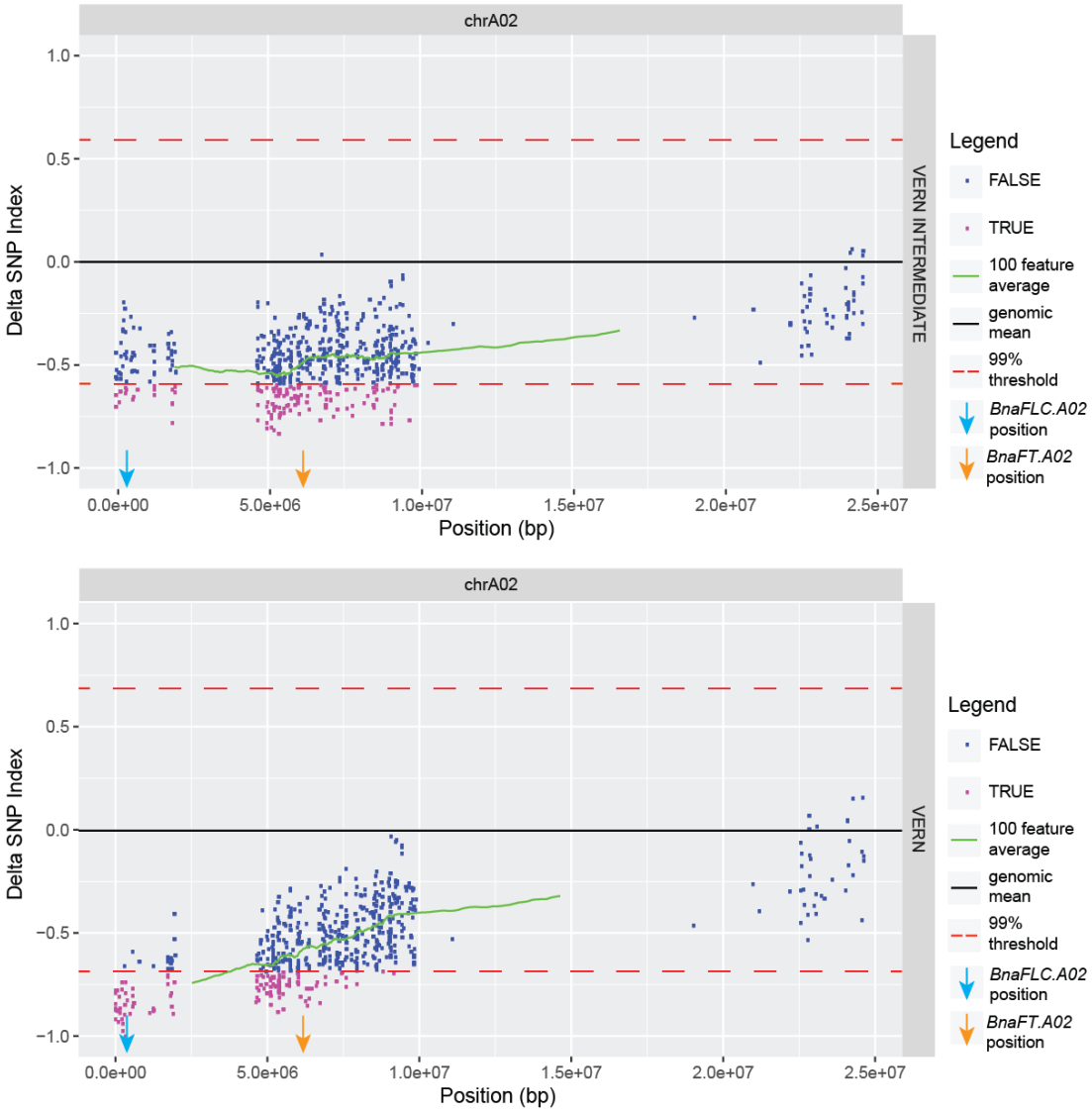


Figure 5. 10: A close-up image of the associating QTL on chromosome A02. Upper panel: VERN INTERMEDIATE comparison. Lower panel: VERN comparison. Each point on the graph represents a ΔSNP index value calculated between two DNA pools. Each ΔSNP index value is plotted against their chromosomal position along the x-axis. ΔSNP index values that fall outside the 99% threshold of no association (red dashed line) are highlighted in pink, while the remaining ΔSNP index values are highlighted in blue. A sliding window average of 100 ΔSNP index values illustrated by the green line. The black line denotes the genomic mean ΔSNP index value. The relative position of *BnaFLC.A02* and *BnaFT.A02* are illustrated with blue and orange arrows respectively.

Like in Section 5.2.2.2. a minor QTL is detected on chromosome C01 however deviations in SNP index values were detected in the VERN comparison (Figure 5.8) but not in the VERN INTERMEDIATE comparison (Figure 5. 9 and Figure S. 14). This suggests the effect of genetic variation on chromosome C01 is important for the

extreme flowering time differences represented in the VERN_EARLY_1 and VERN_LATE_4 pools. In the NVERN comparison, lateness was conferred by allelic contribution from Darmor on chromosom C01 however under VERN conditions, the direction of ΔSNP index values calculated by QTL-seq, suggest sequence reads in the VERN_EARLY_1 pool are like Darmor and the sequence reads in the VERN_LATE_4 pool are like Cabriolet. With vernalisation, allelic contribution from Cabriolet on chromosome C01 conferred late flowering. Flowering genes that are present on chromosome C01 are listed in Table 5. 2. Again, allelic variation at *BnaFRI.C03* formed no significant association with flowering time under VERN conditions.

5.2.2.4. Genotypic data suggests *BnaFT.A02* is important for vernalisation requirement while *BnaFLC.A02* is important for flowering time after vernalisation

To confirm the association between genes in a 10Mbp region on chromosome A02 and variation for flowering time I designed KASP markers to target five SNPs (at positions 136554, 1290192, 1997465, 4654252, 6375505) on chromosome A02 (see methods Section 2.10.5). Two SNPs, SNP 136554 and SNP 6375505, were located within *BnaFLC.A02* and *BnaFT.A02* respectively. For both vernalisation treatments 188 F_2 lines representing 94 of the earliest and 94 of the latest flowering lines were genotyped. KASP genotyping results are listed in Table S. 4 and Table S. 5.

Under NVERN conditions, of the lines that were succefully screened with all five KASP markers, a majority (84%) of early flowering F_2 lines were homozygous for the 6375505-Cab SNP allele while no early flowering line was homozygous for the 6375505-Dar SNP allele (Table 5. 3). A region including *BnaFT.A02* in Cabriolet therefore conferred early flowering. The converse was true for the late flowering F_2 lines; a majority (84%) were homozygous for the 6375505-Dar SNP allele while none carried a homozygous 6375505-Cab SNP allele suggesting the *BnaFT.A02* region in Darmor conferred late flowering. A reduction in homozygosity was detected towards *BnaFLC.A02*. For the 136554 SNP, located within *BnaFLC.A02*, both heterozygous (46%) and homozygous (46%) alleles of the 136554-Cab SNP

allele contributed to early flowering while heterozygous (34%) and homozygous (57%) alleles of the 136554-Dar SNP allele contributed to late flowering. This suggests perhaps that homozygosity at *BnaFLC.A02* is not essential for determining flowering time, while homozygosity at *BnaFT.A02* is more important. The genotyping results highlights *BnaFT.A02* as the more likely candidate gene for flowering time under NVERN conditions and the major determinant of vernalisation requirement in this F_2 population.

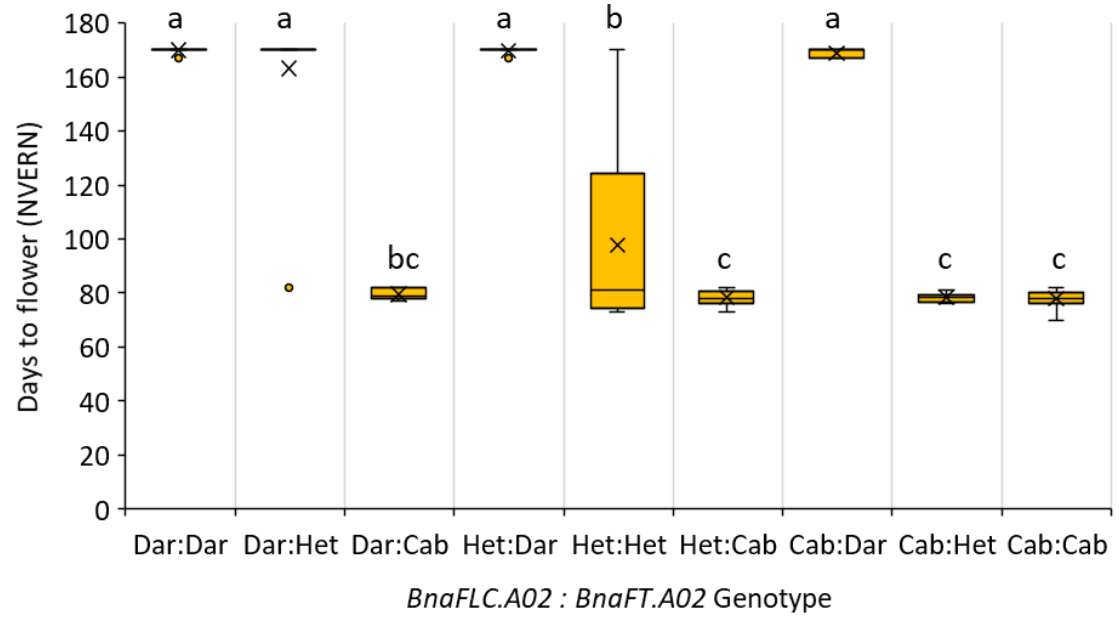
Table 5. 3: A summary of the genotyping results for F_2 lines from the NVERN treatment. The SNP location of five KASP markers 136554, 1290192, 1997465, 4654252, 6375505 are listed and those that are in *BnaFLC.A02* and *BnaFT.A02* are indicated. The proportion of each genotype (Cabriolet-like, Darmor-like, or Heterozygous) found at each SNP position within F_2 lines that were either early or late flowering was calculated and are listed below each SNP position.

			Gene				
			BnaFLC.A02				***BnaFT.A02***
			SNP position				
Phenotype	**Genotype**		**136554**	**1290192**	**1997465**	**4654252**	**6375505**
NVERN EARLY	Cabriolet	Number of lines:	40	43	51	65	73
		Population proportion:	(0.46)	(0.49)	(0.59)	(0.75)	(0.84)
	Heterozygous	Number of lines:	40	38	32	22	14
		Population proportion:	(0.46)	(0.44)	(0.37)	(0.25)	(0.16
	Darmor	Number of lines:	7	6	4	0	0
		Population proportion:	(0.08)	(0.07)	(0.05)	(0.00)	(0.00)
NVERN LATE	Cabriolet proportion	Number of lines:	8	5	4	1	0
		Population proportion:	(0.09)	(0.06)	(0.04)	(0.01)	(0.00)
	Heterozygous proportion	Number of lines:	30	29	23	19	14
		Population proportion:	(0.34)	(0.33)	(0.26)	(0.21)	(0.16)
	Darmor proportion	Number of lines:	51	55	62	69	75
		Population proportion:	(0.57)	(0.62)	(0.70)	(0.78)	(0.84)

Based on the genotypic data, the *BnaFLC.A02* (SNP marker 136554) and *BnaFT.A02* (SNP marker 6375505) genes were either homozygous for Cabriolet, homozygous for Darmor, or heterozygous. Nine genotypic combinations of

BnaFLC.A02 and *BnaFT.A02* alleles were therefore detected in the F_2 population (Figure 5. 11). Significant differences in flowering time measurements were detected between the genotypes ($p<0.001$, Mann-Whitney U test; Figure 5. 11 A). Lines homozygous for the *BnaFT.A02-Dar* allele flowered significantly later than lines homozygous for *BnaFT.A02-Cab* allele, regardless of *BnaFLC.A02* genotype (Figure 5. 11 A). Lines that were heterozygous for both *BnaFLC.A02* and *BnaFT.A02* exhibited intermediate flowering times while the flowering times of lines heterozygous for *BnaFT.A02* were dependent on *BnaFLC.A02* genotype. *BnaFLC.A02-Cab* conferred early flowering, but not when in combination with *BnaFT.A02-Dar*, while *BnaFLC.A02-Dar* conferred late flowering but not when in combination with *BnaFT.A02-Cab*. This suggests a possible epistatic interaction between the genes.

A



B

		Gene	
		BnaFLC.A02	*BnaFT.A02*
		SNP position	
		136554	6375505
Treatment	Phenotype	Genotype	
NVERN	EARLY	Cab	Cab
		Het	Cab
		Cab	Het
		Dar	Cab
	INTERMEDIATE	Het	Het
	LATE	Dar	Het
		Cab	Dar
		Het	Dar
		Dar	Dar

Figure 5. 11: The phenotypes of *BnaFLC.A02* and *BnaFT.A02* genotype in the F_2 population under NVERN treatment. Nine genotypic combinations of *BnaFLC.A02* and *BnaFT.A02* were detected within the F_2 population. The flowering times of each genotyped F_2 line are plotted in (A). Boxplots illustrate the median, mean (x) and range of flowering times measured for each *BnaFLC.A02/BnaFT.A02* genotypic combination. Letters indicate significant differences determined by pair-wise multiple comparison test using Mann-Whitney U test ($p<0.05$). (B) A summary of (A); the flowering time phenotypes detected for each *BnaFLC.A02* and *BnaFT.A02* genotypic combination. Cab (red)= Homozygous Cabriol*et al*lele, Dar (yellow)= Homozygous Darmor allele, Het (green)= Heterozygous allele.

Using the same five KASP markers, 188 F_2 individuals representing 94 of the earliest and 94 of the latest flowering individuals from the VERN treatment were genotyped. KASP genotyping results for all individual lines are listed in Table S. 5.

Again, I observed a clear segregation pattern within the lines that were successfully genotyped using the five KASP markers (Table 5. 4). Unlike for the NVERN analysis, which identified *BnaFT.A02* as important for determining flowering time, this time a majority (74%) of the early flowering lines were homozygous for the 136554-Cab SNP, located in *BnaFLC.A02.* No early flowering line was homozygous for the 136554-Dar SNP. Homozygosity deteriorated towards *BnaFT.A02* and approximately half of all early flowering lines were either homozygous or heterozygous for the 6375505-Cab SNP, and 6% of the early flowering lines were homozygous for 6375505-Dar SNP. This suggests homozygosity at *BnaFLC.A02* is important for determining early flowering after vernalisation. For the late flowering lines 68% were homozygous for the 136554-Dar SNP and 68% were also homozygous for 6375505-Dar SNP suggesting homozygosity for Darmor variation at both *BnaFLC.A02* and *BnaFT.A02* is important to confer late flowering after vernalisation. Taken together the genotyping results highlighte *BnaFLC.A02*, and not *BnaFT.A02*, as the major regulator of flowering time after vernalisation in this F_2 population.

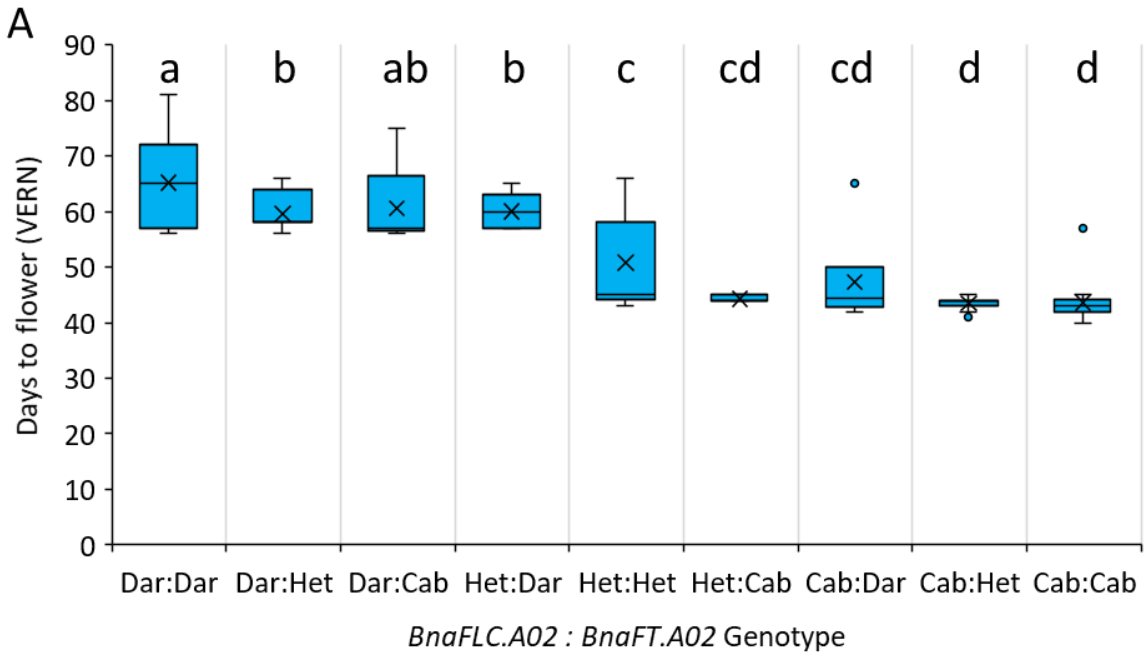
Table 5.4: A summary of the genotyping results for F_2 lines from the VERN treatment. The SNP location of five KASP markers 136554, 1290192, 1997465, 4654252, 6375505 are listed and those that are in *BnaFLC.A02* and *BnaFT.A02* are indicated. The proportion of each genotype (Cabriolet-like, Darmor-like, or Heterozygous) found at each SNP position within F_2 lines that were either early or late flowering was calculated and are listed below each SNP position.

			Gene				
			BnaFLC.A02				***BnaFT.A02***
			SNP position				
Phenotype	**Genotype**		**136554**	**1290192**	**1997465**	**4654252**	**6375505**
VERN EARLY	Cabriolet	Number of lines:	61	56	52	44	38
		Population proportion:	(0.74)	(0.68)	(0.63)	(0.54)	(0.46)
	Heterozygous	Number of lines:	21	23	28	35	39
		Population proportion:	(0.26)	(0.28)	(0.34)	(0.43)	(0.48)
	Darmor	Number of lines:	0	3	2	3	5
		Population proportion:	(0.00)	(0.04)	(0.02)	(0.04)	(0.06)
VERN LATE	Cabriolet	Number of lines:	2	3	4	4	6
		Population proportion:	(0.02)	(0.03)	(0.04)	(0.04)	(0.07)
	Heterozygous	Number of lines:	27	29	27	23	23
		Population proportion:	(0.30)	(0.32)	(0.30)	(0.25)	(0.25)
	Darmor	Number of lines:	62	59	60	64	62
		Population proportion:	(0.68)	(0.65)	(0.66)	(0.70)	(0.68)

Again, based on the genotypic data, nine combinations of *BnaFLC.A02* and *BnaFT.A02* were detected in the VERN treatment. I analysed the flowering time measurements for each genotyped F_2 line (Figure 5. 12) and significant differences in flowering time measurements were detected between the *BnaFLC.A02* and *BnaFT.A02* genotypes ($p<0.001$, Mann-Whitney U test). Lines that were homozygous for *BnaFLC.A02-Dar* flowered significantly later than lines that were homozygous for *BnaFLC.A02-Cab* (Figure 5. 12 A). Lines that were heterozygous for both *BnaFLC.A02* and *BnaFT.A02* exhibited intermediate flowering times while the flowering time phenotype of lines heterozygous for *BnaFLC.A02* was dependent on *BnaFT.A02* genotype. Again, a possible epistatic interaction between the genes was detected after vernalisation. *BnaFT.A02-Cab* conferred early flowering, but not

when in combination with *BnaFLC.A02-Dar*, while *BnaFT.A02-Dar* conferred late flowering but not when in combination with *BnaFLC.A02-Cab*.

Interestingly, when comparing the results of both vernalisation treatments, the epistatic interaction detected between *BnaFLC.A02* and *BnaFT.A02* resulted in contrasting phenotypes that were dependent on vernalisation treatment. For example, *BnaFLC.A02-Cab* in combination with *BnaFT.A02-Dar* conferred late flowering NVERN (Figure 5. 11) but early flowering with VERN (Figure 5. 12). The alternative scenario was also detected where *BnaFLC.A02-Dar* in combination with *BnaFT.A02-Cab* conferred early flowering NVERN (Figure 5. 11) but late flowering with VERN (Figure 5. 12).



B

		Gene	
		BnaFLC.A02	*BnaFT.A02*
		SNP position	
		136554	6375505
Treatment	**Phenotype**	**Genotype**	
VERN	EARLY	Cab	Cab
		Cab	Het
		Cab	Dar
		Het	Cab
	INTERMEDIATE	Het	Het
	LATE	Het	Dar
		Dar	Het
		Dar	Cab
		Dar	Dar

Figure 5. 12: The phenotypes of *BnaFLC.A02* and *BnaFT.A02* genotype in the F_2 population under VERN treatment. Nine genotypic combinations of *BnaFLC.A02* and *BnaFT.A02* were detected within the F_2 population. The flowering times of each genotyped F_2 line is plotted in (A). Boxplots illustrate the median, mean (x) and range of flowering times measured for each *BnaFLC.A02/BnaFT.A02* genotypic combination. Letters indicate significant differences determined by pair-wise multiple comparison test using Mann-Whitney U test ($p<0.05$). (B) A summary of (A); the flowering time phenotypes detected for each *BnaFLC.A02* and*BnaFT.A02* genotypic combination. Cab (red) = Homozygous Cabriol*et al*lele, Dar (yellow) = Homozygous Darmor allele, Het (green) = Heterozygous allele.

5.2.3. DNA sequence and expression variation are detected at *BnaFLC.A02* and *BnaFT.A02*

Due to the association between genetic variation at *BnaFLC.A02* and *BnaFT.A02* and flowering time with and without vernalisation, I investigated both genes for sequence polymorphisms and quantitative gene expression differences in Cabriolet and Darmor.

5.2.3.1. Non-coding polymorphisms are detected at *BnaFLC.A02*

Sequence read depth generated by the Illumina® sequencing (Section 5.3.2.1.) were plotted against the genomic position for *BnaFLC.A02* according to the Darmor-*bzh* reference for both Cabriolet and Darmor (Figure 5. 13 A). Sequence read depth can give an indication of InDel variation between Cabriolet and Darmor, but also between both cultivars and the Darmor-*bzh* reference. InDels can be identified from the presence of crevices or dips in the read depth values plotted. A large deletion was detected in both Cabriolet and Darmor, compared with Darmor-*bzh*, approximately 2,500bp from the translational start site, within intron 5 of *BnaFLC.A02*. This deletion in Cabriolet and Darmor compared with Darmor-*bzh* was confirmed using two sets of primers that were designed to amplify a region from intron 5 to intron 6, encompassing the deletion (methods described in Section 2.5.5.4). Amplification of this region in Cabriolet and Darmor gave products of equal size (Figure 5. 13 B), but approximately 250bp smaller than the predicted region from Darmor-*bzh*. Curiously, analysis of the insertion in Darmor-*bzh* (Darmor-*bzh BnaFLC.A02* sequence retrieved from plants.ensembl. org) revealed a 256bp insertion in intron 5 of *BnaFLC.A02* which was an exact duplication of part of intron 5, exon 6 and part of intron 6 indicating a possible sequence assembly error in Darmor-*bzh*, although, as DNA from Darmor-*bzh* was not available to me for this analysis, the results are unconfirmed. Three additional deletions were detected in Cabriolet compared with Darmor-*bzh*; two within intron 1 and a third within intron 4 (Figure 5. 13 A).

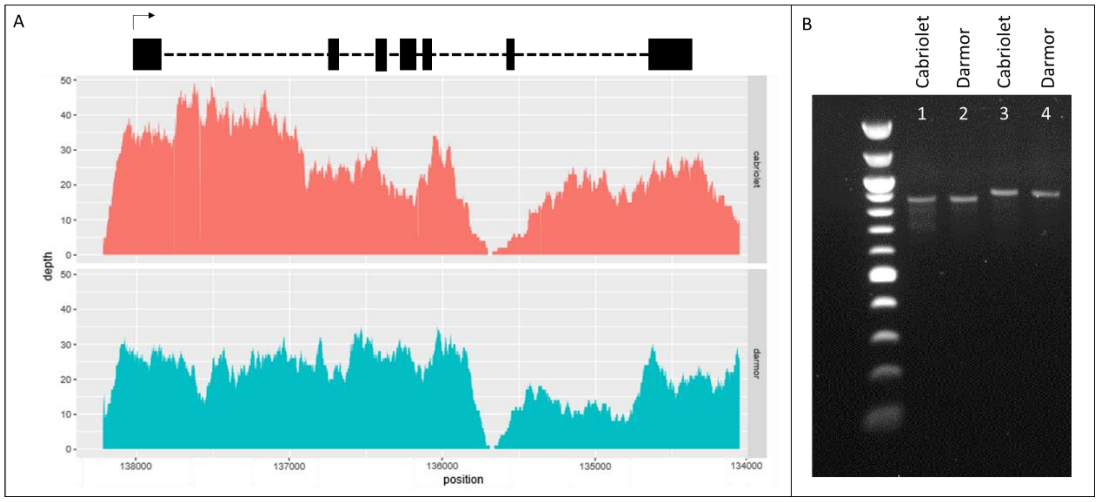


Figure 5. 13: InDel variation at *BnaFLC.A02* between Darmor-*bzh*, Carbiolet and Darmor. **(**A) Sequence read depth of Cabriolet (red) and Darmor (blue) at *BnaFLC.A02* plotted against genomic position according to Damor-*bzh*, the relative position of *BnaFLC.A02* is illustrated schematically, black boxes represent exons and dashed lines represent introns, the black arrow indicates the direction of transcription. (B) PCR marker for a 256bp deletion identified in both Cabriolet and Darmor compared with Darmor-*bzh* (Samples 1 and 2= primers *FLC*A2_GENOTYPE_F1+ *FLC*A2_GENOTYPE_R1, Samples 2 and 3= *FLC*A2_GENOTYPE_F2+ *FLC*A2_GENOTYPE_R2). Amplification of this region gave products of equal sizes in Cabriolet and Darmor, approximately 250bp smaller than the product predicted from Darmor-bzh (100bp DNA ladder reference).

Alignment of *BnaFLC.A02* from Cabriolet (three sequence sources: Illumina® sequences generated in Section 5.2.2.; targeted sequence capture of *BnaFLC.A02* by myBaits®; and the JIC/TGAC Cabriolet genome assembly (R. Wells unpublished)) with Darmor (two sequence sources: Illumina® sequences generated in Section 5.2.2.; targeted sequence capture of *BnaFLC.A02* by myBaits®) and Darmor-*bzh* I was able to identify sequence polymorphisms (Figure 5. 14).

A

Line	Position from ATG	297	337	408	409	492	494	496	504	558	560	835	950	1103	1106	1114	1159	1219	1362	1424	1482	1497	1501	1549	1568	1696	1741	1918
Darmor-bzh		C	C	A	G	A	C	2bp DEL	A	T	T	C	G	A	7bp DEL	T	T	T	T	T	C	G	C	A	T	G	T	T
Darmor		C	C	A	G	A	C	2bp DEL	A	T	T	C	G	A	7bp DEL	T	T	T	T	T	C	G	C	A	T	G	T	T
Cabriolet		1bp DEL	G	G	A	T	T	TT	G	G	A	A	A	G	7bp INS	C	C	C	A	A	T	C	T	G	C	T	G	C
		intron1																	intron2							intron3		intron4

Line	Position from ATG	1983	2006	2029	2037	2075	2076	2080	2121	2122-2978	2583	2588	2603	2635	2650	2690	2692	2756	2767	2790	2854	2889	2893	2898	2899	2913	3073	3179
Darmor-bzh		C	C	T	T	C	C	G	T	256bp INS	A	G	T	A	C	T	T	T	T	C	T	T	T	G	G	A	A	G
Darmor		C	C	T	T	C	C	G	T	256bp DEL	A	G	T	A	C	T	T	T	T	C	T	T	T	G	G	A	A	G
Cabriolet		T	T	A	C	G	G	A	G	256bp DEL	G	A	A	G	T	A	1bp DEL	G	A	A	G	C	A	A	A	T	C	T
		intron5									intron6																	

Figure 5. 14: DNA sequence variation at *BnaFLC.A02* in Cabriolet and Darmor compared with the reference Darmor-*bzh*. All positions are plotted according to the reference Darmor-*bzh*. (A) Listed in the table are all sequence polymorphisms found at *BnaFLC.A02* in Cabriolet and Darmor compared with Darmor-*bzh*. Genetic position of the sequence variation are relative to the translational start site predicted in Darmor-*bzh*, Exon/Intron location is indicated. (B) The distribution of SNPs and InDels across the *BnaFLC.A02* gene in Cabriolet compared with Darmor are illustrated schematically. Black boxes represent exons and dashed lines represent introns, the black arrow indicates the direction of transcription, red arrows indicate where sequence variation is located. (C) Alignment of the RY element region of intron 1 in *A. thaliana* accessions Col-0 and Lov-1, and in *B. napus* cultivars Darmor and Cabriolet. The 598 SNP in Lov-1 (Coustham *et al*., 2012) is highlighted with *, polymorphisms between Cabriolet and Darmor highlighted with *, the RY elements highlighted with blue boxes.

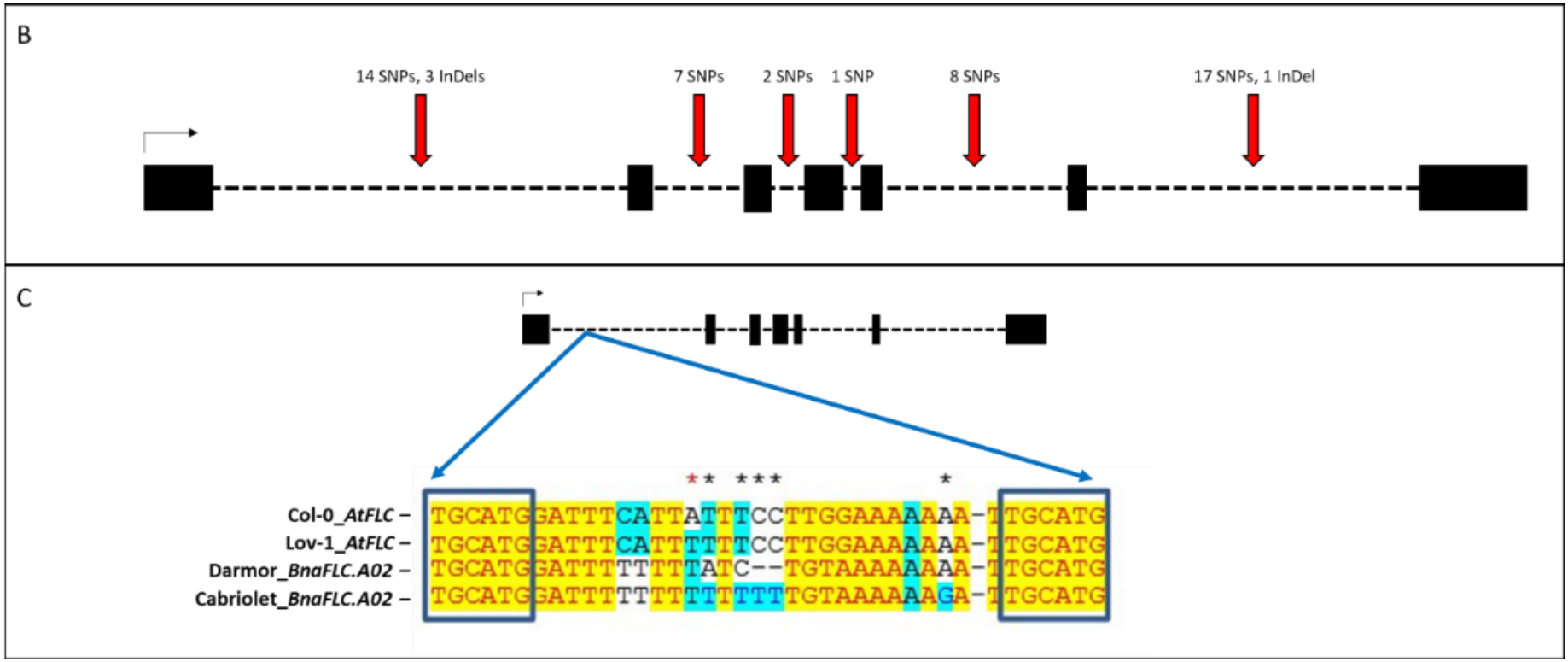


Figure 5. 14 continued

Apart from a 256bp deletion, confirmed in Figure 5. 13, Darmor carries an allele of *BnaFLC.A02* that is identical in sequence to Darmor-*bzh* (Figure 5. 14 A; hereafter referred to as *BnaFLC.A02-Dar*). Compared with Darmor, Cabriolet has 49 SNPs, 2 deletions and 2 insertions at *BnaFLC.A02* and all are located within introns (hereafter referred to as *BnaFLC.A02-Cab*; Figure 5. 14 A and Figure 5. 14 B). Three notable polymorphisms are detected within an A/T rich region found between the RY elements of intron 1 (predicted from Questa *et al*., 2016; Figure 5. 14 C). This region has been identified in *A. thaliana* as important for promoting nucleation and histone modification at the *FLC* locus (Coustham *et al*., 2012, Questa *et al* 2016). Polymorphisms have been detected in this region between the RY elements in a natural accession of *A. thaliana*, Lov-1, known to have a strong vernalisation requirement and response (Figure 5. 14 C). Three SNPs;492:A>T, 494:C>T and 504:A>G, and a 2bp deletion, were detected within this region of intron in *BnaFLC.A02-Cab* compared with *BnaFLC.A02-Dar*.

5.2.3.2. Quantitative expression differences are detected at *BnaFLC.A02*

Due to the presence of non-coding sequence variation at *BnaFLC.A02*, I analysed the expression of this gene in a time-series experiment under difference vernalisation treatments (methods in Section 2.6; Figure 5. 15). No difference in expression was detected without vernalisation therefore the pre-vernalisation level of *BnaFLC.A02* did not vary between Cabriolet and Darmor (Figure 5. 15 F). The absence of expression variation without vernalisation possibly confirms that *BnaFLC.A02* genotype is less important under no vernalisation conditions. *BnaFLC.A02* expression dynamics with vernalisation, in contrast, differ between Cabriolet and Darmor (Figure 5. 15 A-E) providing evidence that *BnaFLC.A02* expression could contribute to the variation in flowering time after vernalisation, and therefore the vernalisation response between both cultivars.

With 12-weeks vernalisation treatment *BnaFLC.A02-Cab* and *BnaFLC.A02-Dar* expression down-regulated to similar levels (Figure 5. 15 A). Although a peak in *BnaFLC.A02* expression was detected in Darmor after plants return to warm conditions (12WT16), the gene exhibited reduced expression again after 32 days

growth in warm glasshouse conditions, reaching levels that were similar to those detected in Cabriolet. After shorter periods of vernalisation however (Figure 5. 15 B-E) *BnaFLC.A02-Dar* expression was always detected at higher levels after vernalisation compared with *BnaFLC.A02-Cab*. With 6-weeks vernalisation treatment, for example (Figure 5. 15 D), the expression of *BnaFLC.A02-Dar* re-activated, reaching levels similar to those observed under no vernalisation (NV) conditions. Four-weeks vernalisation treatment appeared to have a minimal effect on *BnaFLC.A02-Dar* down-regulation (Figure 5. 15 E), possibly contributing to the late flowering time phenotype detected in my previous data (Figure 5. 1 E). Under all vernalisation treatments tested *BnaFLC.A02-Cab* expression appeared to be stably silenced, even after 4-weeks vernalisation treatment (Figure 5. 15 E).

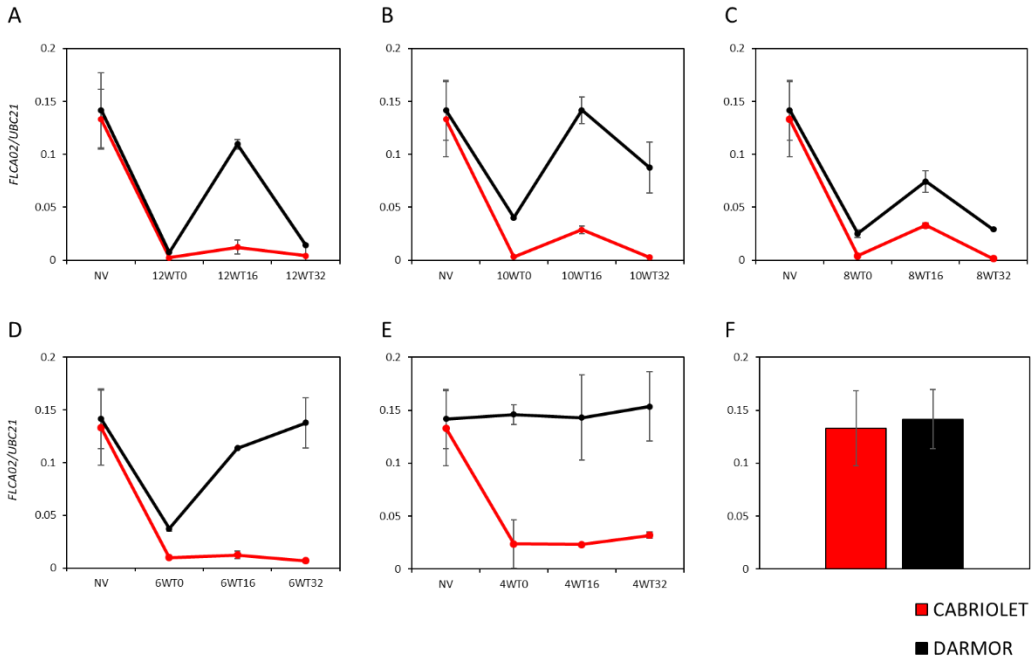


Figure 5. 15: Quantitative expression of *BnaFLC.A02* in Cabriolet and Darmor across time under different vernalisation treatments. Cabriolet is plotted in red, Darmor plotted in black; NV= no vernalisation and sampled 28 days after sowing, T0= at the end of vernalisation treatment and sampled in the cold, T16= 16 days after the end of vernalisation treatment and sampled in the glasshouse, T32= 32 days after the end of vernalisation treatment and sampled in the glasshouse. (A) 12-weeks vernalisation, (B) 10-weeks vernalisation, (C) 8-weeks vernalisation, (D) 6-weeks vernalisation, (E) 4-weeks vernalisation. (F) No vernalisation, a single timepoint was taken for this plot at 28-days after sowing. All expression values are plotted relative to *UBC21*, error bars denote standard error of 2-3 biological replicates.

5.2.3.3. Coding and non-coding variation is present at *BnaFT.A02*

Sequence read depth generated by the Illumina® sequencing (Section 5.2.2.1) for both Cabriolet and Darmor were plotted against the genomic position for *BnaFT.A02* according to the Darmor-*bzh* reference (Figure 5. 16). Like in Section 5.2.3.1. sequence read depth can give an indication of InDel variation between Cabriolet and Darmor, and between both cultivars and the Darmor-*bzh* reference. InDels were identified by the presence of crevices or dips in the read depth values. Sequence read depth coverage indicated the presence of an InDel in Cabriolet and Darmor compared with Darmor-*bzh* approximately 1,500bp from the translational start site, and within intron 2 of *BnaFT.A02*. Two additional deletions were detected in Cabriolet compared with Darmor-*bzh*, all within intronic regions.

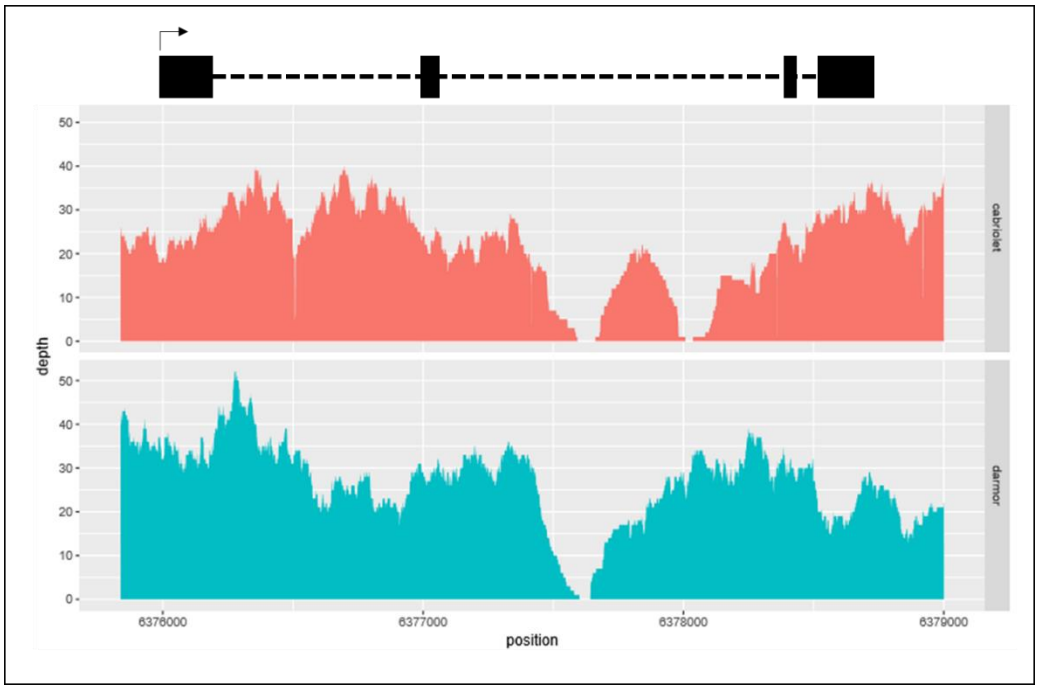


Figure 5. 16: InDel variation at *BnaFT.A02* in Cabriolet and Darmor compared with Darmor-*bzh.* (A) Sequence read depth of Cabriolet (red) and Darmor (blue) at *BnaFT.A02* plotted against the genomic position according to Darmor-*bzh*. Dips in sequence read depth indicate InDel variation. The relative position of *BnaFT.A02* is illustrated schematically, black boxes represent exons and dashed lines represent introns, the black arrow indicates the direction of transcription

Alignment of *BnaFT.A02* from Cabriolet (three sequence sources: Illumina® sequences generated in section 5.2.2.1.; targeted sequence capture of *BnaFT.A02* by myBaits®; and the JIC/TGAC Cabriolet genome assembly (R. Wells unpublished)) Darmor (two sequence sources: Illumina® sequences generated in section 5.3.2.1.; targeted sequence capture of *BnaFT.A02* by myBaits®) with Darmor-*bzh* (sequence retrieved for *BnaFT.A02* from plants.ensembl.org) I was able to identify sequence polymorphisms (Figure 5. 17).

Relative to Darmor-*bzh*, the *BnaFT.A02* in Darmor (hereafter referred to as *BnaFT.A02-Dar*) is almost identical in sequence to Darmor-*bzh*, however a 76bp insertion was detected in intron 2, located 1565bp from the translational start codon of *BnaFT.A02-Dar* (Figure 5. 17 A) confirming the InDel variation identified in Figure 5. 16.

Compared with *BnaFT.A02-Dar*, three SNPs were identified in the Cabriolet *BnaFT.A02* sequence (Figure 5.17 A; hereafter referred to as *BnaFT.A02-Cab*). One SNP145:A>C appears in exon 1 and is non-synonymous, causing an amino acid change from Isoleucine to Leucine (I48L) (Figure 5. 17 C). The remaining two SNPs; 286:C>T and 370:A>G; were detected within intron 1 (Figure 5. 17 A and Figure 5. 17 B). Three deletions were detected in the *BnaFT.A02-Cab* sequence compared with *BnaFT.A02-Dar* (Figure 5. 17 A and Figure 5. 17 B); one 15bp deletion in intron 1, a 6bp in intron 2, and a 287bp deletion in intron 2.

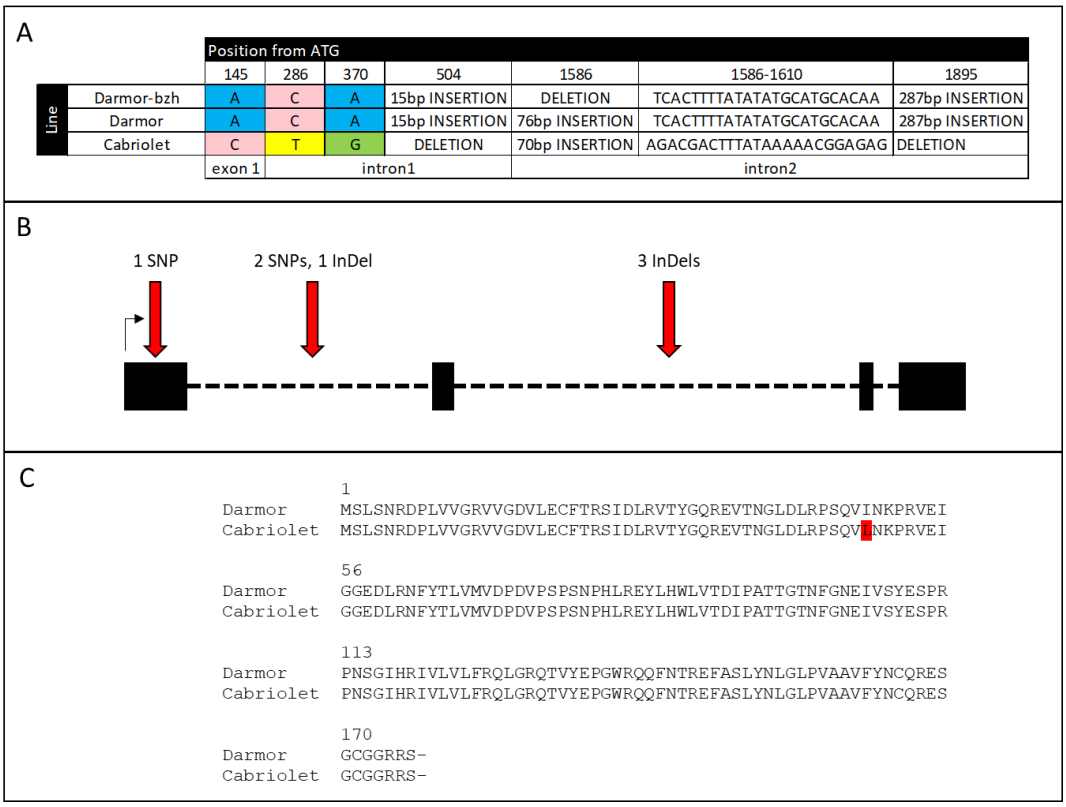


Figure 5. 17: DNA sequence variation at *BnaFT.A02* in Cabriolet and Darmor and Darmor-*bzh.* All genomic positions are plotted according to the Darmor-*bzh* reference genome. (A) The table includes a list of sequence polymorphisms found at *BnaFT.A02* in Cabriolet and Darmor compared with Darmor-*bzh*. The genetic positions of all polymorphisms listed are relative to the translational start site predicted in Darmor-*bzh*. Exon/Intron locations of all polymorphisms are indicated. (B) The location of SNPs and InDels across the *BnaFT.A02* gene in Cabriolet relative to Darmor. The *BnaFT.A02* the gene is illustrated schematically, black boxes represent exons and dashed lines represent introns, the black arrow indicates the direction of transcription (C) The amino acid sequence of BnaFT.A02 in Darmor and Cabriolet, the amino acid substitution, caused by SNP145:A>C, is highlighted in red.

5.2.3.4. Quantitative expression differences are detected at *BnaFT.A02*

Cis-regulatory regions are important for controlling *FT* expression in *A. thaliana* (Adrian *et al*., 2010) therefore I analysed the expression of *BnaFT.A02* in the same time-series experiment as detailed in Section 5.2.3.2.. Under all vernalisation treatments tested, expression of the allele associated with late flowering, *BnaFT.A02-Dar*, was not detected (Figure 5. 18). In contrast *BnaFT.A02-Cab* expression was detectable under no vernalisation (NV) conditions, at 28 days after sowing (Figure 5. 18 F), and exhibited increased expression after vernalisation (Figure 5. 18 A-E). The quantitative expression level of *BnaFT.A02-Cab* increased

with longer periods of vernalisation and the highest quantitative expression was detected 32 days after a 12-week vernalisation treatment (Figure 5. 18 A).

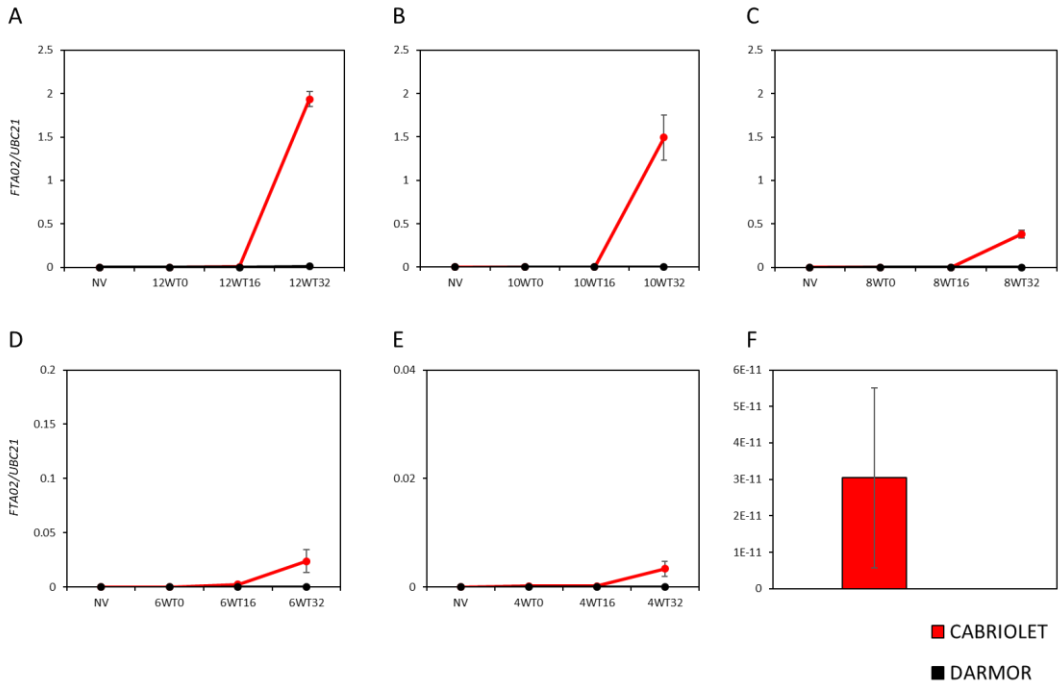


Figure 5. 18: Quantitative expression of *BnaFT.A02* in Cabriolet and Darmor across time under different vernalisation treatments. Cabriolet is plotted in red, Darmor plotted in black; NV= no vernalisation sampled 28 days after sowing, T0= at the end of vernalisation treatment and sampled in the cold, T16= 16 days after the end of vernalisation treatment and sampled in the glasshouse, T21= 32 days after the end of vernalisation treatment and sampled in the glasshouse. (A) 12 weeks vernalisation, (B) 10 weeks vernalisation, (C) 8 weeks vernalisation, (D) 6 weeks vernalisation, (E) 4 weeks vernalisation. (F) Expression of *BnaFT.A02* under NV conditions. Note different y-axis scale in D, E and F. All expression values are plotted relative to UBC21, error bars denote standard error of 2-3 biological replicates.

5.2.3.5 Variation at *BnaFLC.A02* and *BnaFT.A02* is found within diverse winter oilseed rape cultivars

Due to the association between *BnaFLC.A02* and *BnaFLC.A02* and flowering time I screened 47 winter oilseed rape cultivars from the OREGIN DFFS population for *BnaFLC.A02* and *BnaFT.A02* allelic variation by analysing sequences generated by targeted sequence capture of *BnaFLC.A02* and *BnaFT.A02* (myBaits®; https://arborbiosci.com; Table 2. 2). Both the *BnaFLC.A02-Cab* and *BnaFLC.A02-*

Dar alleles were detected in the population (Table 5. 9) and 37 winter oilseed rape cultivars carried *BnaFLC.A02-Cab* while 10 carried *BnaFLC.A02-Dar*. Three alleles of *BnaFT.A02* were found in the population; *BnaFT.A02-Cab*, *BnaFT.A02-Dar*, and a third hereafter referred to as *BnaFT.A02-Var*. Five winter oilseed rape cultivars carried *BnaFT.A02-Cab*, 38 carried *BnaFT.A02-Dar*, and 4 carried *BnaFT.A02-Var*. As the majority of cultivars carry *BnaFLC.A02-Cab* and *BnaFT.A02-Dar* genotypes, I hypothesise these alleles provide a yield advantage in the field.

Results from Section 5.2.2.4. indicated that the allelic combination of *BnaFLC.A02* and *BnaFT.A02* was an important determinant for vernalisation requirement and response. In the OREGIN DFFS population, five winter oilseed rape cultivars exhibited the *BnaFLC.A02-Cab/BnaFT.A02-Cab* genotype, ten cultivars exhibited the *BnaFLC.A02-Dar/BnaFT.A02-Dar*, four exhibited the *BnaFLC.A02-Cab/BnaFT.A02-Var* genotype, while the majority, 28 cultivars, exhibited the *BnaFLC.A02-Cab/BnaFT.A02-Dar*. No cultivar was detected with genotype *BnaFLC.A02-Dar/BnaFT.A02-Cab*. According to Figure 5. 11 and Figure 5. 12, and based solely on their *BnaFLC.A02* and *BnaFT.A02* genotype, the data in Table 5. 5 suggests the majority of winter oilseed rape cultivars in the OREGIN DFFS population have been bred to have a strong vernalisation requirement (conferred by *BnaFT.A02-Dar*), but flower early in response to vernalisation (conferred by *BnaFLC.A02-Cab*).

Table 5. 5: Distribution of *BnaFLC.A02* and *BnaFT.A02* alleles within winter oilseed rape (OSR) cultivars from the OREGIN DFFS population. Cultivars with *BnaFLC.A02-/ BnaFT.A02-Cab* alleles highlighted red, cultivars with *BnaFLC.A02-/ BnaFT.A02-Dar* alleles highlighted black, *BnaFT.A02-Var* highlighted green.

	Type	*BnaFLC.A02*	*BnaFT.A02*
CABRIOLET	Early parent		
DARMOR	Late parent		
ABUKUMA NATANE	Winter OSR		
AMBER X COMMANCHE DH LINE	Winter OSR		
ASSYST 078 (EXPRESS-617)	Winter OSR		
Baltia	Winter OSR		
BIENVENU DH4	Winter OSR		
Bolko	Winter OSR		
Cabernet	Winter OSR		
Cabriolet	Winter OSR		
CANBERRA x COURAGE DH LINE	Winter OSR		
Capitol	Winter OSR		
CAPITOL x MOHICAN DH LINE	Winter OSR		
Castille	Winter OSR		
COLUMBUS X NICKEL DH LINE	Winter OSR		
Coriander	Winter OSR		
Darmor	Winter OSR		
Dippes	Winter OSR		
EUROL	Winter OSR		
Excalibur	Winter OSR		
Expert	Winter OSR		
Flash	Winter OSR		
HANSEN X GASPARD DH LINE	Winter OSR		
HURON x NAVAJO DH LINE	Winter OSR		
INCA x CONTACT DH LINE	Winter OSR		
JetNeuf	Winter OSR		
LICROWN X EXPRESS DH LINE	Winter OSR		
MADRIGAL x RECITAL DH LINE	Winter OSR		
Major	Winter OSR		
Matador	Winter OSR		
NK Bravour	Winter OSR		
NORIN	Winter OSR		
Palmedor	Winter OSR		
POH 285, Bolko	Winter OSR		
PRIMOR	Winter OSR		
Quinta	Winter OSR		
RAFAL DH1	Winter OSR		
Rocket	Winter OSR		
ROCKET (pst) x LIZARD DH LINE	Winter OSR		
Samourai	Winter OSR		
SHANNON x WINNER DH LINE	Winter OSR		
SLAPSKA, SLAPY	Winter OSR		
Slovenska Krajova	Winter OSR		
TAPIDOR DH	Winter OSR		
Temple	Winter OSR		
TEQUILLA x ARAGON DH LINE	Winter OSR		
Verona	Winter OSR		
Victor	Winter OSR		
Vision	Winter OSR		

5.3. Discussion

Variation for vernalisation requirement and response in winter oilseed rape highlights striking parallels with natural accessions of *A. thaliana* (Shindo *et al*., 2005, Shindo *et al*., 2006, Li *et al*., 2014, Duncan *et al*., 2015). My assessment of flowering times in a vernalisation series experiment identified two cultivars, Cabriolet and Darmor that had contrasting phenotypes. To investigate the genetic loci important for determining variation in vernalisation requirement and response I performed a QTL-seq analysis on an F_2 population generated from a reciprocal cross of these cultivars. A major QTL region on chromosome A02, including *BnaFLC.A02* and *BnaFT.A02*, was detected. I identified DNA sequence and gene expression variation at both *BnaFLC.A02* and *BnaFT.A02*, which likely contributed to the variation in vernalisation requirement and response. Like in previous Chapters, allelic variation at orthologues of *FRI* did not significantly influence flowering time.

The approach used to identify SNPs and small InDels for the QTL-seq analysis was stringent; detecting only those SNPs and small InDels that had a read depth coverage of at least 20 reads and the SNP or InDel had to be present in more than 95% of those reads. This meant that at least 80% of genomic positions were filtered from the QTL-seq analysis. Inclusion of this filtered data would possibly allow the identification of more QTL that associates with flowering time. Nevertheless, this stringent approach in SNP and InDel detection allowed me to identify more than 17,754 SNPs in Cabriolet compared with Darmor, and 12,522 SNPs and small InDels in Darmor compared with Darmor-*bzh*. SNPs and small InDels in Darmor compared with Darmor-*bzh* were fund genome-wide, however the density and location of these SNPs and InDels indicated evidence of large genomic blocks that varied between Darmor and Darmor-*bzh*. According to Foisset *et al*., 1995 Darmor-*bzh* was generated by introgressing the dwarf BREIZH (*bzh*) locus from a line called B192 into Darmor. The *bzh* locus was first identified in a mutagenized population of the winter oilseed rape cultivar Primor, which was subsequently crossed with another winter oilseed rape cultivar Jet-Neuf to generate line B192. The *bzh* locus was then introgressed into other cultivars such as Darmor. The genetic variation detected between Darmor and Darmor-*bzh* support this difference in pedigree history.

Previous reports (Nelson *et al*., 2014, Schiessl *et al*., 2015, Raman *et al*., 2016, Xu *et al*., 2016) have identified associations between regions on chromosome A02 and flowering time of *B. napus* under field, glasshouse and controlled environment conditions. In the present study, variation for flowering time in the absence of vernalisation (vernalisation requirement) and flowering time after vernalisation (vernalisation response) mapped to the same 10Mbp region on chromosome A02. Although this genomic region was large, genotypic data suggested that allelic variation at *BnaFT.A02* contributed to the vernalisation requirement while allelic variation at *BnaFLC.A02* contributed to the vernalisation response. With and without vernalisation, *BnaFT.A02-Cab* and *BnaFLC.A02-Cab* alleles conferred an early flowering phenotype while *BnaFT.A02-Dar* and *BnaFLC.A02-Dar* alleles conferred a late flowering phenotype.

A previous report (Raman *et al*., 2016) has suggested that variation in the pre-vernalisation level of *FLC2* i.e. the combined expression of *BnaFLC.A02* and *BnaFLC.C02* before vernalisation, associates with flowering time and contributes up to 22% of the variation recorded in *B. napus*. Difference in pre-vernalisation expression level of *BnaFLC.A02* was not detected between early and late flowering cultivars Cabriolet and Darmor respectively. However striking differences in the expression profiles of *BnaFLC.A02-Cab* and *BnaFLC.A02-Dar* alleles were detected in response to vernalisation. The expression of the *BnaFLC.A02-Cab* allele was stably silenced even with short periods (4-weeks) of vernalisation while repression of the *BnaFLC.A02-Dar* allele to levels similar to those detected in Cabriolet was only achieved with 12-weeks vernalisation. Like in *A. thaliana* variation in the length of cold required to induce epigenetically stable silencing of *FLC* (Coustham *et al*., 2012) is likely a major determinant of flowering timein *B. napus*.

Variation in the epigenetic silencing of *BnaFLC.A02-Cab* and *BnaFLC.A02-Dar* highlights striking parallels with previous reports in *A. thaliana* and *B. oleracea* (Coustham *et al*., 2012, Li *et al*., 2014, Irwin *et al*., 2016). Cis-variation is an important determinant for the epigenetic silencing of *FLC* in both species (Coustham

et al., 2012, Li *et al*., 2014, Li *et al*., 2015, Hawkes *et al*., 2016, Irwin *et al*., 2016, Questa *et al*., 2016). Natural *A. thaliana* accessions and in *B. oleracea* that exhibit variation in *FLC* silencing also exhibit sequence variation within non-coding regions of the gene (Li *et al*., 2014, Irwin *et al*., 2016). Similarly, I found sequence variation at *BnaFLC.A02* at non-coding regions, while the BnaFLC.A02 protein in Cabriolet and Darmor exhibited 100% amino acid sequence identity. The polymorphisms responsible for the variation in *BnaFLC.A02* expression are yet to be determined however reports in *A. thaliana* suggest cis-regulatory regions in intron 1 of *FLC* are important for regulating the vernalisation response phenotype in *A. thaliana* (Coustham *et al*., 2012, Li *et al*., 2014). Important cis-regulatory elements found in intron 1 include the nucleation region and the VRE (vernalization response element) which are important for *FLC* silencing and maintenance of silencing respectively (Sheldon *et al*., 2002; Finnegan and Dennis 2007; Angel *et al*., 2011, Li *et al*., 2014). Due to conservation between *FLC* in *A. thaliana* and *Brassica* species (Tadege *et al*., 2001, Schranz *et al*., 2002, Yuan *et al*., 2009, Zhao *et al*., 2010, Wu *et al*., 2012, Xiao *et al*., 2013, Irwin *et al*., 2016) it is reasonable to hypothesise that sequence variation within intron 1 of *BnaFLC.A02* might contribute to the differential expression dynamics detected. For example, DNA polymorphisms that distinguishes *BnaFLC.A02-Cab* from *BnaFLC.A02-Dar* are found between the two RY elements located within intron 1 of *BnaFLC.A02*. These cis-elements are recognised by the B3 domain binding protein VAL1 (Questa *et al*., 2016) and mutations within and between the RY elements have been documented to change the *FLC* silencing and reactivation phenotype in *A. thaliana* (Coustham *et al*., 2012, Li *et al*., 2014, Questa *et al*., 2016). Transgenic complementation experiments are underway in the Judith Irwin Laboratory to determine the functional consequence of polymorphisms within intron 1 of *BnaFLC.A02*.

BnaFT.A02 and its homologue in *B. rapa*, *BraFT.A07*, have previously been identified as a candidate gene for a flowering time (Long *et al*., 2007, Wang *et al*., 2009b, Zhang *et al*., 2015, Raman *et al*., 2016). In the present study I identified allelic variation at *BnaFT.A02* which associated with variation for flowering time, with and without vernalisation. Allelic variation was detected in the coding region of *BnaFT.A02*, and one non-synonymous SNP differentiated *BnaFT.A02-Cab* from

BnaFT.A02-Dar. The SNP encoded an amino acid substitution from an Isoleucine to a Leucine at amino acid position 48 in the early flowering cultivar Cabriolet. The same amino acid substitution has been reported in *B. rapa* BraFT.A07 (Zhang *et al*., 2015, Schiessl *et al*., 2017b) but detected in a late flowering cultivar. As the I48L mutation is associated with contrasting flowering time phenotypes in *B. rapa and B. napus*, it is unlikely to confer the flowering time variation detected but this remains to be tested.

Like previous reports (Coustham *et al*., 2012, Irwin *et al*., 2016) vernalisation length had a quantitative effect on *FT* expression. *BnaFT.A02-Cab* expression increased 100-fold after 12-weeks vernalisation compared with 6-weeks, most likely as a consequence of the epigenetic silencing of *BnaFLC*. In contrast, *BnaFT.A02-Dar* expression was not detectable at any timepoint suggesting a difference in the regulation of *BnaFT.A02* expression between the two cultivars. InDel variation was detected within intronic regions of *BnaFT.A02-Cab* and *BnaFT.A02-Dar*. Perhaps most interesting was an insertion detected within intron 2 of *BnaFT.A02* in the late flowering cultivar Darmor compared with Cabriolet. A previous report in *B. rapa* identified an insertion within intron 2 of *BraFT.A07* which was associated with late flowering due to failed transcription of *BraFT.A07* (Zhang *et al*., 2015) similar to the absence of *BnaFT.A02* expression in Darmor. I hypothesise that variation at cis-regulatory regions of *BnaFT.A02*, possibly within intron 2, also resulted in the failed transcription of *BnaFT.A02-Dar* and ultimately late flowering.

Genotypic data confirmed that *BnaFLC.A02-Cab* and *BnaFT.A02-Cab* conferred early flowering, while *BnaFLC.A02-Dar* and *BnaFT.A02-Dar* conferred late flowering, with and without vernalisation. Analysis of F_2 lines segregating for *BnaFLC.A02* and *BnaFT.A02* indicated a genetic interaction between both genes that was dependent on genotype. *BnaFLC.A02-Cab*, when in combination with *BnaFT.A02-Dar*, conferred late flowering in the absence of vernalisation, but early flowering after vernalisation. In contrast, *BnaFLC.A02-Dar*, when in combination with *BnaFT.A02-Cab*, conferred early flowering in the absence of vernalisation, but late flowering after vernalisation. This suggests that *BnaFT.A02* is important for

determining flowering time without vernalisation, while *BnaFLC.A02* is important after vernalisation. Interestingly the majority of winter oilseed rape cultivars within the OREGIN DFFS diversity set contained *BnaFLC.A02-Cab* in combination with *BnaFT.A02-Dar*. I suggest Breeders have selected for winter oilseed rape cultivars with a strong vernalisation requirement, which would inhibit flowering until after winter, but a relatively weak vernalisation response, which would promote early flowering even after relatively short winter growth.

It's worth noting that a minor QTL was detected on chromosome C01, contributing to the most extreme flowering time phenotypes with and without vernalisation. Several potential candidate genes were detected, and the direction of SNP index values suggested Cabriolet loci contributed to earliness without vernalisation but contributed to lateness after vernalisation. It's possible that one gene has an alternative role with and without vernalisation, but also likely that two different genes, that are both located on chromosome C01, contributed to the flowering time phenotype separately with and without vernalisation. For example, based on their function in *A. thaliana*, orthologues of *FCA* and *FD* on chromosome C01 could contribute to flowering time without vernalisation, while orthologues of *VRN1*, *VRN2*, and *VEL1* could contribute to flowering time with vernalisation. The possibility that allelic variation at *FD*, which interacts with *FT* in the apical meristem, helps coordinate flowering time in *B. napus* would also be an interesting avenue to explore.

It is apparent that cis-variation at orthologues of *FLC* and *FT* in *B. napus* are likely to have a major influence on gene expression and ultimately the flowering time phenotype of a plant (Coustham *et al*., 2012, Li *et al*., 2014, Li *et al*., 2015, Irwin *et al*., 2016, Hawkes *et al*., 2016, Questa *et al*., 2016). Like in *A. thaliana*, targeting cis-variation, resulting in changes to transcription factor binding sites, has played an important role in flowering time evolution in *B. napus*.

Chapter 6: Flowering time and inflorescence architecture: a pleiotropic role for *FLC*?

6.1. Introduction

In previous Chapters I have demonstrated that allelic variation at orthologues of *FLC* is a major determinant of flowering time in *B. napus*. *FLC* has been shown to contribute to traits other than flowering in *A. thaliana* and *Brassica* (Grbic and Bleecker 1996, Kalinina *et al*., 2002, Poduska *et al*., 2003, Lou *et al*., 2007, Wang *et al*., 2007, Chiang *et al*., 2009, Mei *et al*., 2009, Willmann and Poethig, 2011, Huang *et al*., 2013, Fletcher *et al*., 2014). One trait known to be associated with allelic variation at *BnaFLC* which contributes to yield in *B. napus* is plant architecture (Li *et al*., 2016). In this Chapter I describe a preliminary analysis of the relationship between flowering time, plant architecture and *FLC* in both *A. thaliana* and *B. napus*.

The overall architecture and morphology of a flowering plant is determined by the activity of clusters of mitotically active cells, known as meristems, that give rise to above- and below-ground organs (Greb *et al*., 2003). During embryogenesis the establishment of a primary apical-basal axis gives rise to the shoot and root apical meristems which produce the primary shoot and root respectively (Grbić and Bleecker, 2000). Nearly all above-ground organs are derived from the shoot apical meristem (SAM) (Müller and Leyser, 2011). In *A. thaliana*, the SAM grows indeterminately, with the number, pattern, and timing of each new organ determined by the environment and genetic status of the plant (Müller and Leyser, 2011). During post-embryonic growth, secondary growth axes are established due to the initiation of new meristems (Grbić and Bleecker, 2000, Greb *et al*., 2003). These new meristems, after a vegetative phase, switch to reproductive development and, under ideal conditions, elongate into a branch (Grbić and Bleecker, 2000). Secondary meristems are not dissimilar to the main SAM (Greb *et al*., 2003). New growth axes are initiated from a cluster of mitotically active cells and form leaves, flowers, additional meristems and branches in an iterative pattern (Müller and Leyser, 2011).

It is the location and activation of these additional meristem initials that determines the branching architecture and thus the overall morphology of a plant (Grbić and Bleecker, 2000).

Meristem initials that give rise to shoot branches form within the axils of leaf primordia (Grbić and Bleecker, 2000) and will hereafter be referred to as axillary meristems. During prolonged vegetative growth, axillary meristems form in the axils of rosette leaves (i.e. basal leaves), during reproductive growth they form in the axils of cauline leaves (i.e. leaves associated with the main stem). It has been argued that cauline leaves are the leaf primordia present at the SAM at the floral transition and a floral stimulus leads to accelerated axillary meristem development and a reduction in leaf size (Hempel and Feldman, 1994, Hempel *et al*., 1998).

The pattern of meristem initiation differs between vegetative and reproductive growth stages (Grbić and Bleecker, 2000). During vegetative development, axillary meristem formation is initially inhibited but can develop in the axils of rosette leaves during prolonged vegetative growth (Grbić and Bleecker, 2000). Axillary meristems of vegetative leaves show an acropetal pattern in their initiation and form in the axils of the oldest rosette leaves first. Meristem initiation then proceeds in a wave of activation from the oldest to youngest leaf (Grbić and Bleecker, 2000). At the transition from vegetative to reproductive growth, axillary meristems develop at the base of the SAM (Grbić and Bleecker, 2000). These axillary meristems form in the axil of the youngest cauline-leaf primordia first, and the pattern of axillary meristem activation switches to a basipetal pattern of initiation, from the youngest to oldest leaf primordia (Grbić and Bleecker, 2000).

The decision to remain dormant or grow out into a branch is heavily dependent on auxin signalling and secondary signals given by strigolactones and by cytokinins (Müller and Leyser, 2011). In *A. thaliana* these hormone signals are integrated within the axillary bud by *BRANCHED1 (BRC1)* and *BRC2,* genes that are closely related to the TCP transcription factor identified in maize (*Zea mays*) *teosinte*

branched1 (Aguilar-Martínez *et al*., 2007). *BRC1* was shown to be highly expressed in developing axillary meristems and inhibits branch outgrowth. As the axillary meristem develops, *BRC1* expression is down-regulated and allowing branch outgrowth (Aguilar-Martínez *et al*., 2007).

Using yeast two-hybrid and *in vivo* assays Niwa *et al*., (2013) demonstrated that the BRC1 protein can interact with FT and TWIN SISTER OF FT (TSF) proteins in plant cells, and this protein-protein interaction is localised in the nucleus. FT protein was also found to move from the subtending leaf to the axillary meristem. The *brc1* mutant is not earlier flowering compared with wild-type, however expressing *BRC1* in the SAM inhibited *FT*, and significantly delayed flowering (Niwa *et al*., 2013). Flowering time control and branching are therefore developmentally tightly linked.

Several studies in *A. thaliana* (Grbic and Bleecker 1996, Kalinina *et al*., 2002, Poduska *et al*., 2003, Wang *et al*., 2007, Huang *et al*., 2013) have identified a correlation between branch phenotypes and flowering time. For example, using *A. thaliana* Multi-Parent Recombinant Inbred Line populations (developed by Huang *et al*., (2011)) eight major QTL were mapped for a "Reduced Stem Branching" phenotype (RSB) calculated as the proportion of cauline leaf axils that developed branches (Huang *et al*., 2013). Individuals that exhibited an absence of branches were always late flowering leading to the conclusion that early flowering was epistatic to RSB (Huang *et al*., 2013). Three of the QTL for RSB co-localised with the flowering time genes *FRI, FLC* and *FT* suggesting allelic variation at all three genes can contribute to branching phenotypes in *A. thaliana* (Huang *et al*., 2013)

An *A. thaliana* accession called Sy-0, originally collected from the Isle of Skye in Scotland, has an unusual branching morphology. This accession is late flowering without vernalisation and produces rosettes of leaves in the axils of cauline leaves on the primary stem. Histological analysis of aerial rosette development revealed the axillary meristems of Sy-0 initiate many more leaf primordia compared with accessions that are early flowering without vernalisation such as L*er*. Grbic and

Bleecker (1996) termed this unusual branching morphology, aerial rosettes. The altered morphology of the plant results in a prolonged vegetative phase that continues even after transition to flowering. Interestingly when Sy-0 plants receive a vernalisation treatment, aerial rosettes do not form and "wild-type" branches develop (Grbic and Bleecker, 1996).

A cross between Sy-0 and the early flowering accession L*er* revealed late flowering and aerial rosettes are genetically dominant. Two loci were found to be required for the Sy-0 morphology which mapped to *FRI* and *AERIAL ROSETTE 1*, which later became known as *HUA2* (Grbic and Bleecker, 1996). Both genes act independently to increase the expression of *FLC* (Grbic and Bleecker, 1996, Poduska *et al*., 2003, Wang *et al*., 2007). It was therefore concluded that allelic variation within functional alleles of *FRI, HUA2* and *FLC* can lead to the production of aerial rosettes, but only in the absence of vernalisation (Grbic and Bleecker, 1996, Poduska *et al*., 2003, Wang *et al*., 2007).

Another *A. thaliana* accession reported to have unusual branching morphology is Zu-0, originally collected from Germany (Kalinina *et al*., 2002). The accession is vernalisation responsive and after vernalisation Zu-0 exhibits accelerated flowering and develops branches normally. In the absence of vernalisation, however Zu-0 is late flowering and does not form branches in the axils of cauline leaves. Histological analysis of non-vernalized Zu-0 plants show arrested axillary meristem development. Zu-0 plants carry functional alleles at *FRI* and exhibit high *FLC* expression, again implicating a role for vernalisation in determining plant architecture in *A. thaliana* (Kalinina *et al*., 2002).

FLC is a major determinant of flowering time, but it has been shown to contribute to a number of other traits including branching, seed germination, and the juvenile to adult transition in *A. thaliana* (Chiang *et al*., 2009, Willmann and Poethig, 2011, Huang *et al*., 2013). Orthologues of *FLC* are also reported to underlie major QTLs for drought tolerance, turnip formation and plant height in *B. rapa* and *B. napus* (Lou

et al., 2007, Mei *et al*., 2009, Fletcher *et al*., 2014). Genome-wide FLC Chromatin Immuno Precipitation-sequencing (ChIP-seq) data published by Deng *et al*., (2009) showed that FLC protein binds to over 500 targets within the *A. thaliana* genome. Interestingly, one of the FLC protein targets identified in this analysis was a region upstream of the gene *BRC1*. The binding of FLC protein to *BRC1* is proposed to be weak (Deng *et al*., 2011, Huang *et al*., 2013), nonetheless this hypothesised interaction is an interesting avenue which I have preliminarily explored in this Chapter.

Like *A. thaliana, B. napus* plants generate primary and secondary inflorescences bearing flowers and fruit (Li *et al*., 2016). Plant height, branch number, branch distribution, and inflorescence length contribute to the overall architecture of a *B. napus* plant (Li *et al*., 2016). The number of primary branches has been shown to positively correlate with seed yield (Chen *et al*., 2014, Li *et al*., 2016) and an analysis of 472 *B. napus* cultivars demonstrated primary branch number in *B. napus* is influenced by genetic and environmental factors (Li *et al*., 2016). Candidate genes proposed by Li *et al*., (2016) that associated with primary branch number in a diverse collection of *B. napus* include *FRI, FLC, APETALLA 2, CONSTANS-like 4, AGAMOUS-LIKE 3, AGAMOUS-like 79* and *FLOWERING PROMOTING FACTOR 1* suggesting flowering time genes may also contribute to plant architecture in *B. napus*.

The aim of my work in this Chapter was to investigate the phenotypic effect of vernalisation on flowering time and primary branch number in *A. thaliana* and *B. napus*. Using a combination of natural accessions and introgression lines of *A. thaliana* I show vernalisation had a quantitative effect on flowering time, and on branch number and branch distribution. In agreement with previous findings (Grbic and Bleecker 1996, Kalinina *et al*., 2002, Poduska *et al*., 2003, Wang *et al*., 2007, Chiang *et al*., 2009, Willmann and Poethig, 2011, Huang *et al*., 2013). Vernalisation promoted an acceleration in flowering time, an increase in rosette branch number and a decrease in cauline branch number in *A. thaliana*. This was in part due to allelic contribution from *FLC*. In *B. napus* I found flowering time to be positively

correlated with leaf number but negatively correlated with branch number. Late flowering cultivars produced fewer branches compared with early flowering cultivars. Vernalisation treatment influenced branch number in a crop dependent manner suggesting that the vernalisation requirement of a plant is a determinant of flowering time and plant architecture, traits that could be manipulated for higher yields.

6.2. Results

6.2.1. *FRI* and *FLC* delay flowering and increase branch number in *A. thaliana*

I analysed three *A. thaliana* accessions, Col-0, Col-*FRI*^{SF2} and Col-*FRI*^{SF2}(*flc-3*) for their flowering time and branch architecture phenotypes (as illustrated in Figure 6. 1 A) without vernalisation and under controlled environment conditions (20°C, 16-hour light/8-hour dark, 70% humidity; Figure 6. 1). The *A. thaliana* accession Columbia, or Col-0, carries a non-functional *FRI* allele (described in Chapter 3), and thus is early flowering without vernalisation, expressing *FLC* at low levels. Col-*FRI*^{SF2} (hereafter referred to as Col-*FRI*) is an introgression line derived from Col-0 and carries a dominant allele of *FRI*. The *FRI* allele was introgressed from the *A. thaliana* accession San Feliu-2 (SF2, Lee *et al*., 1994b) and as a result it is late flowering without vernalisation and *FLC* is expressed at high levels (described in Chapter 3). Col-*FRI*^{SF2}(*flc-3*) (hereafter referred to as Col-*FRI*(*flc-3*)) was generated by fast neutron radiation of Col-*FRI* (Michaels and Amasino 1999) and carries a 104bp deletion over the start codon of *FLC* making it non-functional, and early flowering without vernalisation. See Table 2 .3 for *A. thaliana* accession details.

As expected, when grown under no vernalisation conditions, Col-*FRI* flowered significantly later compared with Col-0 (Figure 6. 1 B, $p<0.001$, Mann-Whitney U test). Although more variable, Col-*FRI* also produced significantly more branches (Figure 6. 1 C, $p<0.001$, Mann-Whitney U test), including both cauline (Figure 6. 1 D, $p<0.001$, Mann-Whitney U test) and rosette (Figure 6. 1 E, $p=0.001$, Mann-Whitney U test) branches, compared with Col-0. Col-*FRI* also flowered significantly later compared with Col-*FRI*(*flc-3*) (Figure 6. 1 B, $p<0.001$, Mann-Whitney U test) produced significantly more branches (Figure 6. 1 C, $p<0.001$, Mann-Whitney U test), including both cauline (Figure 6. 1 D, $p<0.001$, Mann-Whitney U test) and rosette (Figure 6. 1 E, $p=0.032$, Mann-Whitney U test) branches. Cauline branch number is known to positively correlate with flowering time (Hempel and Feldman, 1994, Hempel *et al*., 1998) and the difference in flowering times between Col-*FRI*, Col-0 and Col-*FRI*(*flc-3*) (74+/-10, 24+/-1, and 22+/-2 days on average respectively)

might in part explain the increase in cauline branch number in Col-*FRI*. Of the total branches produced by these accessions a significantly ($p<0.003$, Mann-Whitney U test) smaller proportion of branches were activated from rosette leaves in Col-FRI compared with Col-0 and Col-*FRI*(*flc-3*). Functional alleles of both *FRI* and *FLC* are therefore required to delay flowering, increase the number of cauline branches, and inhibit rosette branch activation.

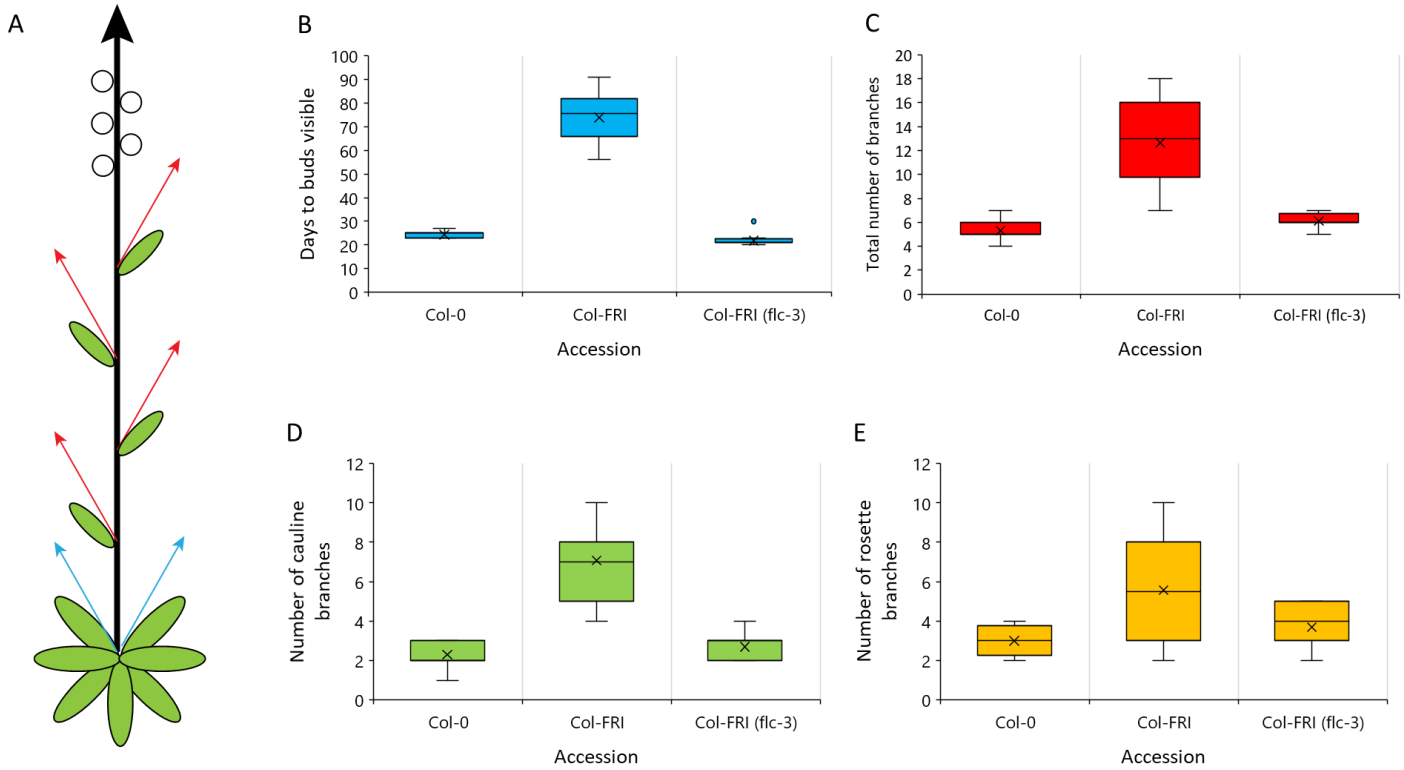


Figure 6. 1: Flowering time and branching phenotypes of *A. thaliana*. (A) Schematic representation of an *A. thaliana* plant. tThe black arrow represents the primary indeterminate stem, green ellipses represent leaves, red arrows represent cauline branches, blue arrows represent rosette branches, white circles represent flowers. (B-E) A phenotypic analysis of *A. thaliana* accessions Col-0, Col-*FRI* and Col-*FRI*(*flc-3*) under no vernalisation conditions, (B) number of days to buds visible, (C) total number of branches (cauline branches plus rosette branches), (D) number of cauline branches, (E) number of rosette branches. Plants were grown under controlled environment conditions, data are plotted as boxplots representing the median, mean (x) and range values from at least 14 plants.

6.2.2. Vernalisation length affects flowering time and branch phenotype in *A. thaliana*

To test the effect of vernalisation length on flowering time and branch number I grew *A. thaliana* Col-*FRI* plants under five vernalisation (5°C, under 8-hour light and 16-hour dark, 70% humidity) treatment lengths; 12, 8, 4, and 0 weeks at. Sowing was staggered so that all plants were transferred to glasshouse conditions (16-hour light and 8-hour dark, 18°C daytime temperature and 15°C night-time temperature, 70% humidity) on the same day. I scored plants for their flowering time, total branch number, cauline branch number and rosette branch number (as illustrated in Figure 6. 1 A). Increasing the length of vernalisation treatment resulted in an acceleration of flowering time as expected (Figure 6. 2 A). The total number of branches i.e. the sum of cauline and rosette branches, were similar across all vernalisation treatments tested (Figure 6. 2 B). Although measurements were variable between replicate plants, a reduction in cauline branch number (Figure 6. 2 C) and an increase in rosette branch number (Figure 6. 2 D) were recoded with increasing lengths of vernalisation treatment. Cauline branch number is known to positively correlate with flowering time (Hempel and Feldman, 1994, Hempel *et al*., 1998) explaining why cauline branch number decreased with increasing vernalisation length. Rosette branch activation was dependent on vernalisation length, implicating a possible role for the quantitative expression of *FLC* in rosette branch activation.

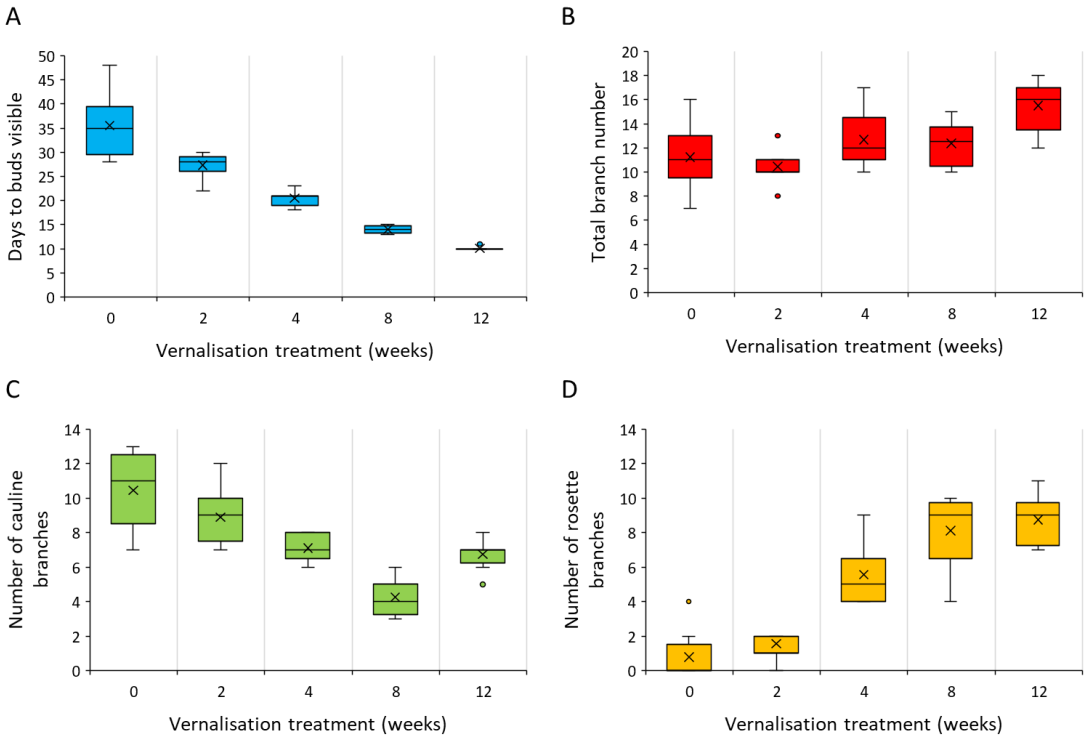


Figure 6. 2: Flowering time and branching phenotypes of *A. thaliana* accession Col-*FRI* with vernalisation. (A) number of days to buds visible, (B) total number of branches (cauline + rosette branches), (C) number of cauline branches, (D) number of rosette branches. Vernalisation treatments given were 12, 8, 4, 2, and 0 weeks, plants were grown under glasshouse conditionsafter vernalisation treatment, data are plotted as boxplots representing the median, mean (x) and range values from at least 8 plants.

Interestingly after a 12-week vernalisation treatment and subsequent growth under long day glasshouse conditions, I observed a reduction in cauline branches compared with 8-weeks vernalisation in Col-*FRI*. When repeated under controlled environment conditions (16-hour light and 8-hour dark, 20°C, 70% humidity; Figure 6. 3) this reduction in cauline branches in response to 12-weeks vernalisation was not detected. Plant architecture is an environmentally sensitive trait (Müller and Leyser, 2011). Perhaps the more variable temperature environment of the glasshouse contributed to the difference in cauline branch number patterning measured after 12 weeks.

The branching phenotype of Col-*FRI* with vernalisation was compared, in a separate experiment, to two other *A. thaliana* accessions; Edi-0 an accession from Edinburgh; and Lov-1 an accession from Sweden. Their phenotypes were compared with two near isogenic lines (NIL) Edi-NIL, an introgression line carrying the *FLC* allele from Edi-0 in a Col-*FRI* genomic background; and Lov-NIL, an introgression line carrying the *FLC* allele from Lov-1 in a Col-*FRI* genomic background (***Figure 6. 3***see Table 2. 3 for *A. thaliana* accession details). The accessions can be divided according to Li *et al*., (2014) into two groups; the rapid vernalising accessions Col-*FRI* and Edi-0, and the slow vernalising accession Lov-1.

Both Col-*FRI* and Edi-0 exhibited a facultative vernalisation requirement and flowered late without cold (Figure 6. 3 A). Similar to the previous experiment with Col-*FRI* (Figure 6. 2), increasing lengths of vernalisation treatment lead to an acceleration in flowering (Figure 6. 3 A) a decrease in the number of cauline branches (Figure 6. 3 B) and an increase in the number of rosette branches (Figure 6. 3 C). Edi-0 flowered later than Col-*FRI* under all vernalisation treatments and as a result produced more cauline branches and fewer rosette branches. Edi-NIL produced an intermediate phenotype for both flowering time and branch number indicating allelic variation at *FLC* might contribute to but is not the only gene regulating both traits.

The majority of Lov-1 plants that had received less than 6 weeks vernalisation did not flower under controlled growth conditions and therefore did not produce branches (Figure 6. 3 D, E, F). Under these non-saturating vernalisation conditions (≤ 6 weeks) I observed some plants of Lov-1 which bolted but exhibited floral reversion phenotypes and produced aerial rosettes as reported in *A. thaliana* accession Sy-0 (Grbic and Bleecker, 1996) in place of cauline leaves. This made assessing flowering time and branch number difficult and therefore I excluded the flowering time and branching data for Lov-1 after ≤ 6 weeks vernalisation treatment from further analysis. Lov-1 flowered later (Figure 6. 3 D) and produced more cauline (Figure 6. 3 E) and fewer rosette branches (Figure 6. 3 F) compared with Col-*FRI*. The Lov-NIL line carrying the *FLC* allele from Lov-1 exhibited an

intermediate phenotype, flowering with at least 4 weeks vernalisation. Lov-NIL flowered earlier than Lov-1, produced fewer cauline branches and more rosette branches indicating again that allelic variation at *FLC* can in part contribute to both flowering time and branch phenotypes in *A. thaliana*.

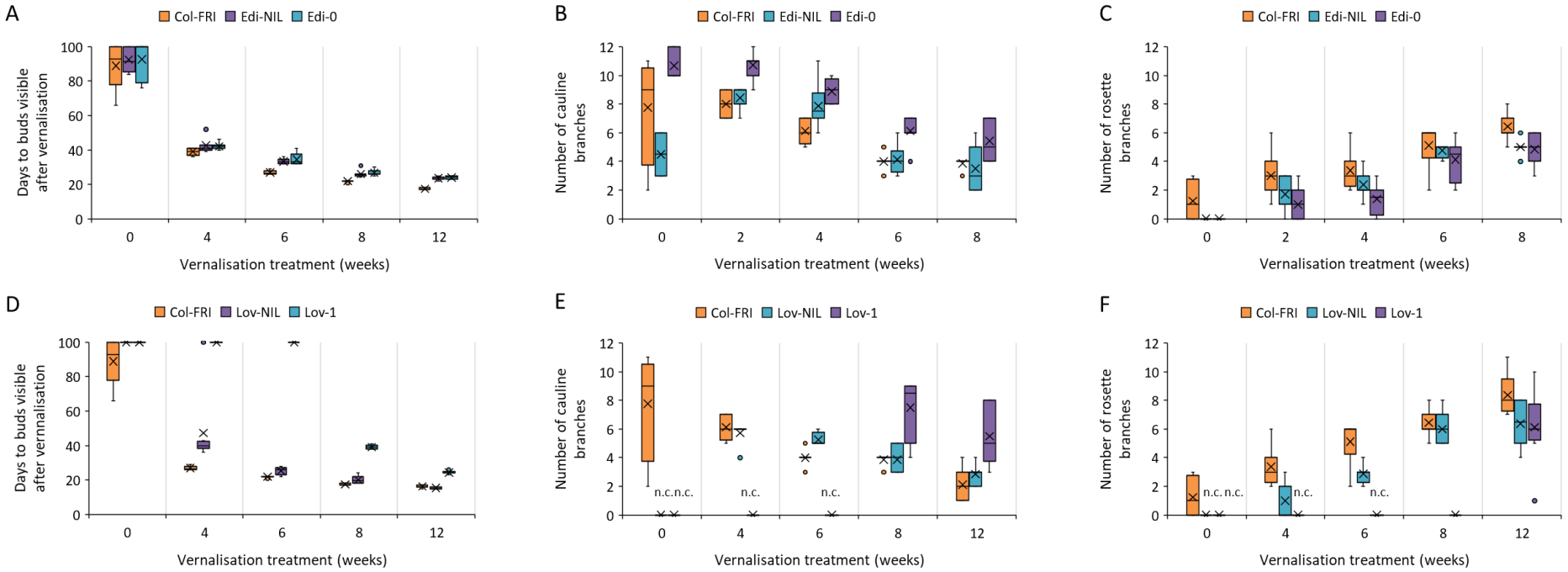


Figure 6. 3: Flowering time and branching phenotypes of *A. thaliana* accessions with vernalisation. (A-C) Col-*FRI*, Edi-0 and Edi-NIL, (D-F) Col-*FRI*, Lov-1 and Lov-NIL. (A + D) number of days to buds visible, (B + E) number of cauline branches, (C + F) number of rosette branches. Vernalisation treatments given were 12, 8, 6, 4, 2, and 0 weeks, plants were grown under controlled environment conditions, data is plotted as boxplots representing the median, mean (x) and range values from at least 8 plants. Plants that flowered after 100 days were given a did not flower measurement of 100 days, plants that did not flower were not scored for branches and are indicated by "n. c." (not counted).

6.2.3. Vernalisation length correlates with *BRANCHED1* activation

FLC has been reported to bind upstream of the branching inhibitor gene *BRANCHED 1* (*BRC1)* (Deng *et al*., 2009) who's protein has been shown to interact with FT protein in the axillary meristems of *A. thaliana* (Niwa *et al*., 2013). I therefore investigated the qualitative expression of *BRC1*, *FT*, *FLC*, and the control gene *UBC* (primer sequences in Table 2. 7 and method details described in Section 2. 6) in Col-*FRI* without vernalisation (NV) and after 4-weeks (4W) and 8-weeks (8W) vernalisation treatment (Figure 6. 4). RNA was extracted from leaf material at 12 days after stratification (NV), at the end of the vernalisation treatment (4WT0 and 8WT0), and at ten days after return to warm conditions (4WT10 and 8WT10). *FLC* expression was detectable in all timepoints tested. Previous reports (Angel *et al*., 2011, Berry *et al*., 2015) have shown that *FLC* expression in Col*FRI* is reduced after 8-weeks compared with 4-weeks vernalisation treatment, although I did not test this quantitatively, it can be assumed in my analysis that *FLC* expression level is also reduced after 8-weeks compared with 4-weeks vernalisation. *FT* expression was detectable only after 8 weeks vernalisation (8WT0 and 8WT10; Figure 6. 4 D and Figure 6. 4E). Eight weeks vernalisation was therefore sufficient to down-regulate *FLC* to levels that released the inhibition of *FT*, resulting in increased *FT* expression. Four weeks vernalisation in contrast was not sufficient length to promote *FT* (4WT0 and 4WT10; Figure 6. 4 B and Figure 4. 6 C). *FT* expression was detected NV (Figure 6.4 A), indicating 12 days in long-day conditions was sufficient to promote some *FT* expression even when *FLC* expression is at its highest.

BRC1 expression was detectable after 8-weeks vernalisation treatment (Figure 6. 4 D and Figure 6. 4 E), but not after 4 weeks. This suggests, like *FT*, more than 4-weeks vernalisation treatment is required to promote the formation of axillary meristems that express *BRC1*. Unlike *FT* however, *BRC1* expression was not detectable under NV (Figure 6. 4 A) conditions suggesting both genes are regulated independently. According to Aguilar-Martínez *et al*., (2007) *BRC1* expression accumulates in the axillary meristems after flowering is initiated. In Col-*FRI*, 8-weeks vernalisation was sufficient to downregulate *FLC* allowing *FT* to be expressed and the floral transition to occur during the vernalisation treatment. Due to the detection of *BRC1* expression at 8-weeks T0, we can deduce that axillary meristems have started to initiate during

vernalisation treatment. This initiation of axillary meristems in the cold may explain the increase in rosette branch number measured in Col-*FRI* with increasing lengths of vernalisation.

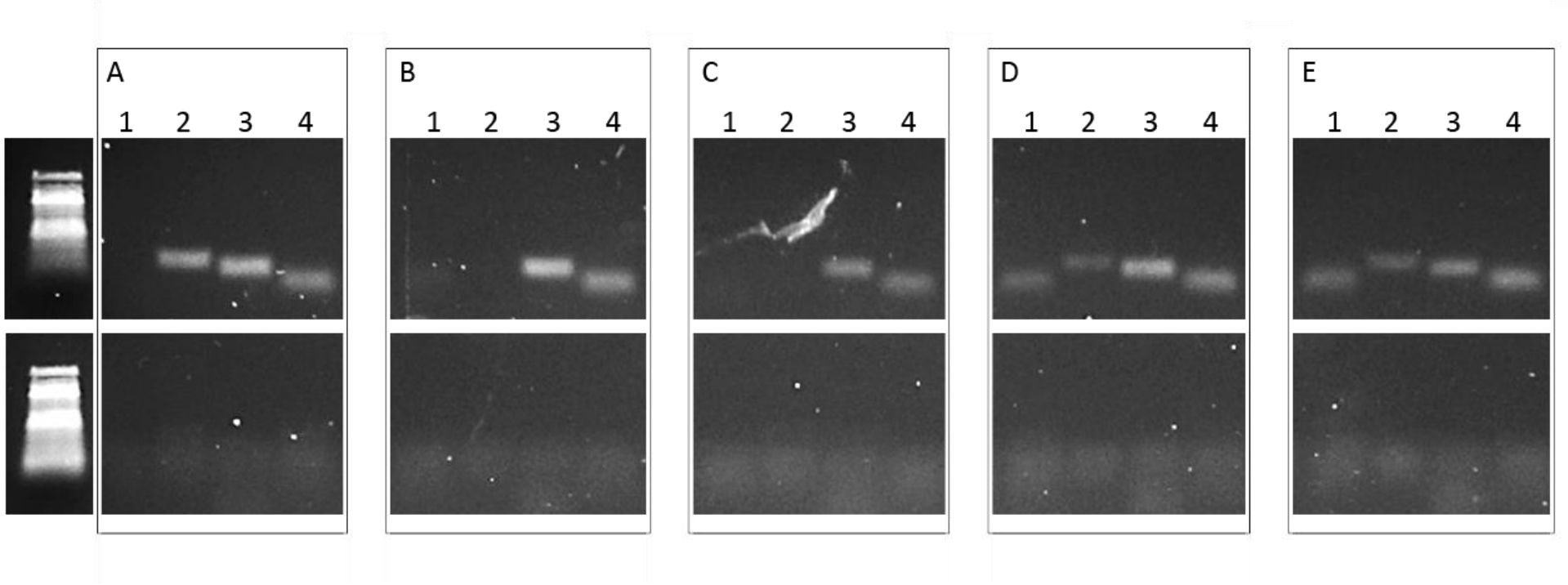


Figure 6. 4: A qualitative analysis of *BRC1*, *FT*, *FLC* and *UBC* in *A. thaliana* accession Col-*FRI*. RT-PCR was performed on pooled seedlings at (A) NV, (B) 4WT0, (C) 4WT10, (D) 8WT0, (E) 8WT10. Bands are visible on the gel when expression of the gene is detectable by RT-PCR. Lane 1: *BRC1*, Lane 2: *FT*, Lane 3: *FLC*, Lane 4: *UBC*. 100bp DNA ladder is included for reference.

6.2.4. Vernalisation increases branch number in winter oilseed rape

B. napus exhibits an indeterminate branching habit like *A. thaliana*. However, unlike *A. thaliana*, primary branches are produced from leaves on primary and secondary stems, no rosette leaves or rosette branches form (Figure 6. 5 A). I previously analysed five winter oilseed rape cultivars named Cabriolet, Darmor, Major, Norin and Primor were previously analysed for flowering time under six vernalisation treatments; 12, 10, 8, 6, 4, and 0 weeks (Chapter 5; Figure 5. 1). In addition to flowering time measurements, I also analysed plants from this experiment for leaf number (the number of leaves and leaf scars on the primary stem) and branch number, as illustrated in Figure 6. 5 A.

Winter oilseed rape plants produced the greatest number of leaves under no vernalisation conditions, indicative of the late flowering phenotypes measured under these conditions (Chapter 5; Figure 5. 1). Three cultivars (Darmor, Major, Norin) did not flower without vernalisation (0 weeks; Figure 5. 1 F). They produced no branches and many leaves, and therefore leaf number was not counted under this treatment (Figure 6. 5 C, D, E). Cultivars Cabriolet and Primor exhibited the smallest difference in leaf number between treatments. Vernalisation treatment had no significant effect on leaf number in Cabriolet, while a significant increase in leaf number was detected in Primor only without vernalisation (0 weeks) (Figure 6. 5 B and F, $p<0.05$, ANOVA Bonferroni multiple comparison). For Darmor, Major and Norin, cultivars with strong vernalisation requirements (described in Chapter 5), a significant increase in leaf number was detected with 4-weeks ($p<0.05$, ANOVA Bonferroni multiple comparisons) in Darmor and Major (Figure 6. 5 C, D), and with 6-weeks in Norin ($p<0.05$, Bonferroni multiple comparisons; Figure 6. 5 E). However, this increase in leaf number may be a consequence of the delay in flowering time measurements recorded for these cultivars under those vernalisation conditions (described in Chapter 5).

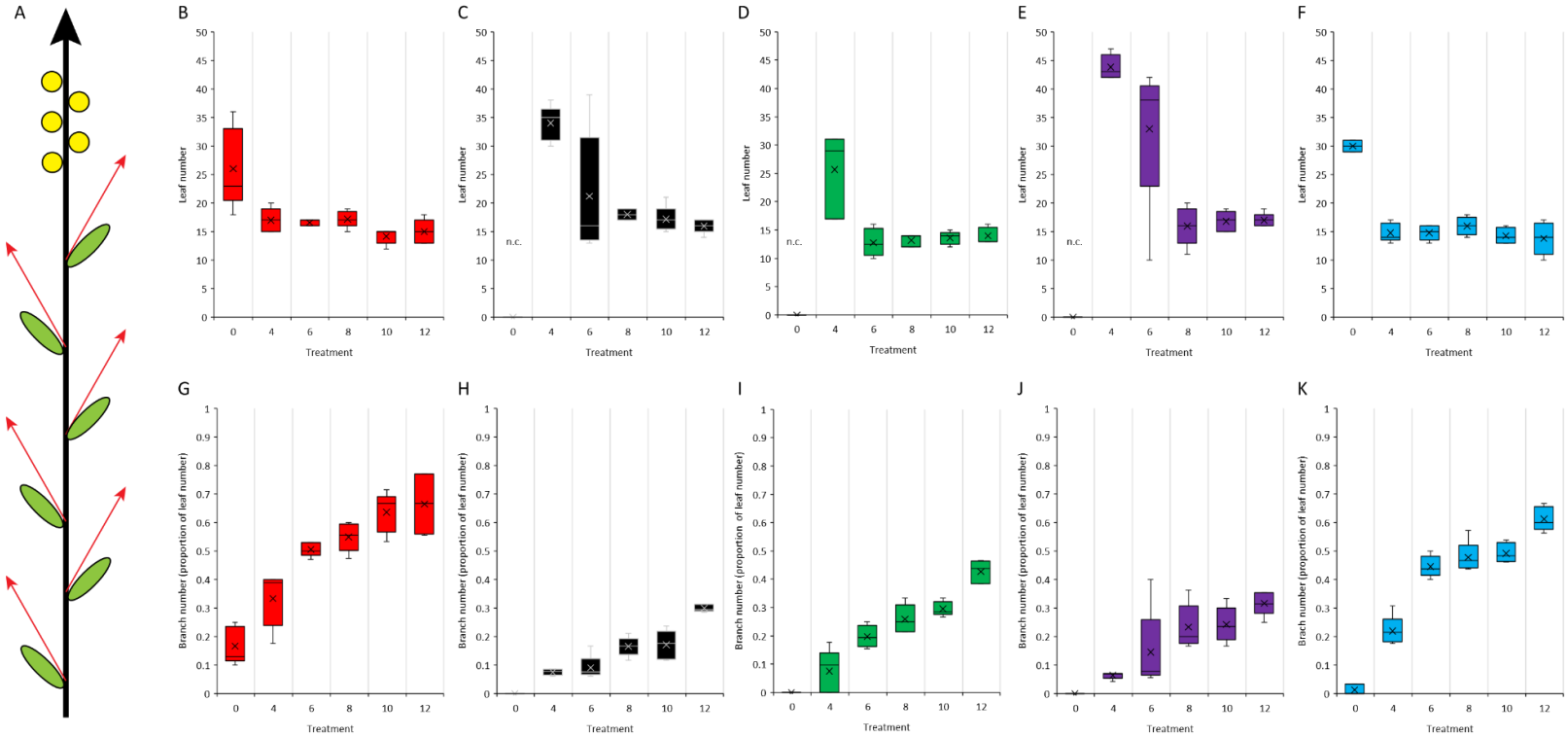


Figure 6. 5: Branching phenotypes of winter oilseed rape cultivars with vernalisation. (A) Schematic representation of a *B. napus* plant, the black arrow represents the primary indeterminate stem, green ellipses represent leaves, red arrows represent branches, yellow circles represent flowers. (B-F) Number of leaves (n.c. = not counted), (G-K) Number of branches (plotted as a proportion of the leaf number) (B+G) Cabriolet, (C+H) Darmor, (D+I) Major, (E+J) Norin, (F+K) Primor. Vernalisation treatments given were 12, 10, 8, 6, 4, 2, and 0 weeks, plants were grown under glasshouse conditions after vernalisation treatment. The data are plotted as boxplots representing the median, mean (x) and range values from 5 plants.

The number of primary stem branches is plotted in Figure 6. 5 G-K. In all cases branch number (calculated as a proportion of the total number of leaves along the primary stem) increased with increasing lengths of vernalisation. Branches were activated basipetally (from the apex downwards) in response to vernalisation (Figure 6. 6) and all plants produced the greatest number of branches after 12-weeks vernalisation. Comparing the number of branches produced after a 12-week vernalisation treatment with other treatments, Cabriolet, Norin and Primor produced significantly fewer branches after 6-weeks vernalisation treatment, while Darmor produced significantly fewer branches after 10-weeks and Major after 8-weeks vernalisation ($p<0.05$, Bonferroni multiple comparisons) suggesting variation in branch number in response to vernalisation is present between cultivars. Under all vernalisation treatments tested Darmor, Major and Norin produced significantly fewer branches compared with Cabriolet and Primor ($p<0.05$, Bonferroni multiple comparisons). Different responses between cultivars could be due to the difference in vernalisation requirement I detected previously (Figure 5. 1; Chapter 5) which possibly indicates a pleiotropic role for flowering time on plant architecture which requires further investigation.

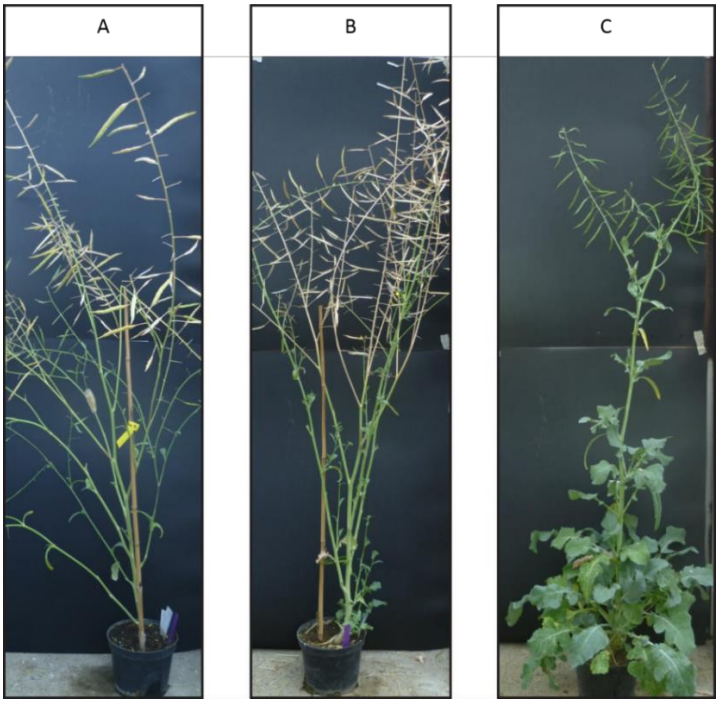


Figure 6. 6: Infloresence architecture of winter oilseed rape cultivar Cabriolet after (A) 12-weeks, (B) 6-weeks, (C) 0-weeks vernalisation treatment. Branches were activated in a basipetal pattern from the apex downwards.

6.2.5. The effect of vernalisation on branch number is dependent on crop type

I have described the flowering times of 88 *B. napus* cultivars from the OREGIN DFFS diversity population (Chapter 4). In the same experiment, these cultivars were also measured for leaf and branch number). Three replicate plants from each cultivar were grown with (6-weeks vernalisation; VERN) and without (0-weeks, NVERN) vernalisation and grown under poly-tunnel conditions during the spring and summer of 2014 as defined in Chapter 4 and in Section 2.9.1. Each plant was scored for the number of leaves (sum of leaves and leaf node scars) and separately for the number of branches as is illustrated in Figure 6. 5 A. The leaf number and branch number data were then correlated with the flowering times measured for each plant (data taken from Chapter 4) and are illustrated in Figure 6. 7. The number of branches are plotted as a proportion of the total leaf number (0= no leaves produced a branch, 1= all leaves produced a branch).

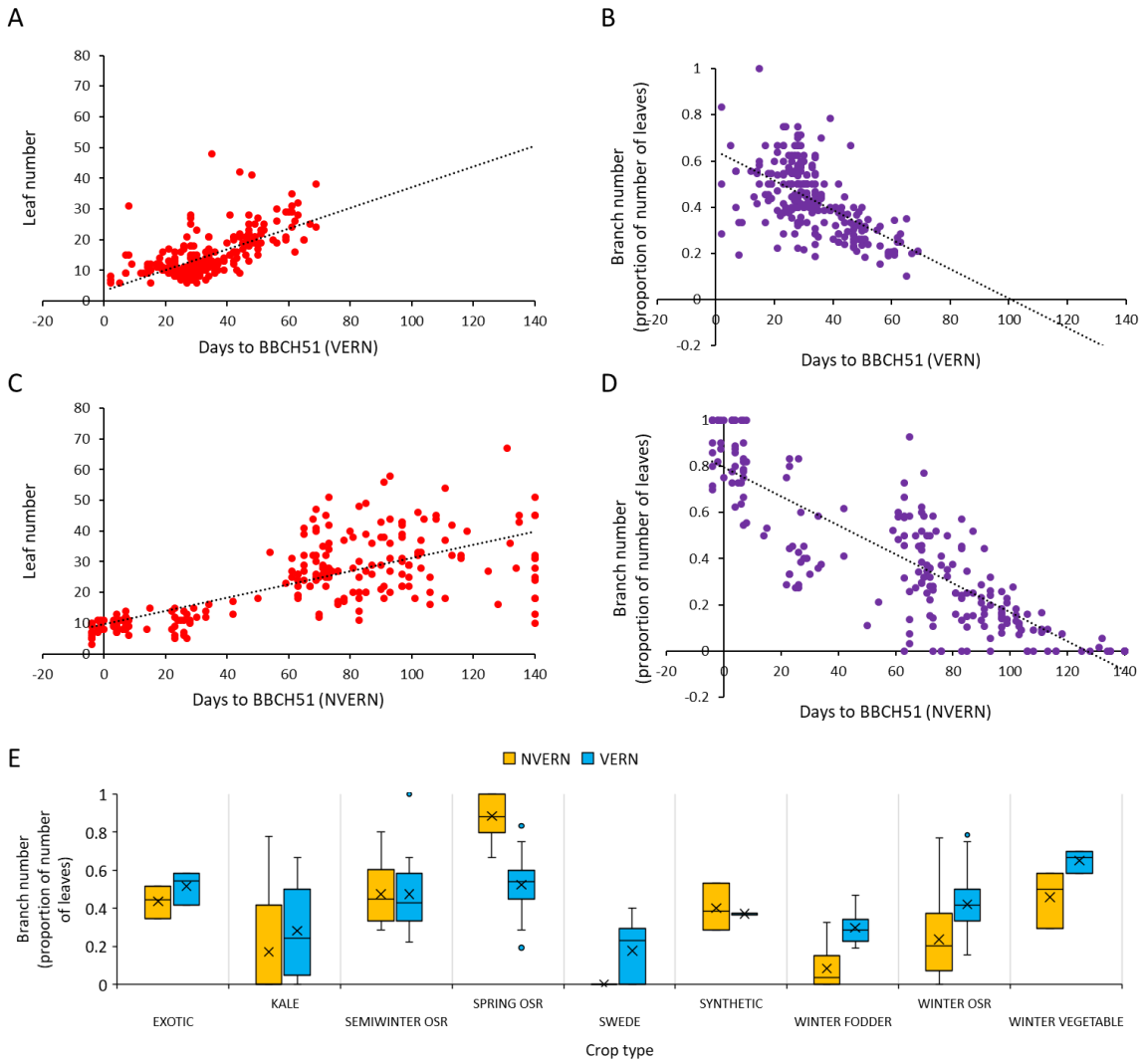


Figure 6. 7: Flowering time and branching phenotypes of 88 *B. napus* cultivars with (VERN) and without (NVERN) vernalisation. (A) Under VERN conditions flowering time exhibits a significant positive correlation ($p<0.001$; Spearman's rank correlation) with the number of leaves, (B) Under VERN conditions flowering time exhibits a significant negative correlation ($p<0.001$; Spearman's rank correlation) with branch number, (C) Under NVERN conditions flowering time exhibits a significant positive correlation ($p<0.001$; Spearman's rank correlation) with the number of leaves, (D) Under NVERN conditions flowering time exhibits a significant negative correlation ($p<0.001$; Spearman's rank correlation) with the number of branches. (E) Branch number associates with crop type under NVERN (yellow) and VERN (blue) conditions. Data is plotted as boxplots representing the median, mean (x) and range values from three replicates per cultivar. Branch number was calculated as the number of branches formed from the primary stem as a proportion of the total number of leaves.

There was a significant positive correlation between flowering time (number of days to BBCH51) and leaf number in *B. napus* (Table 6. 1; Figure 6. 7 A, C) indicating

the latest flowering plants produced the highest number of leaves. Without vernalisation (Figure 6. 7A) plants produced an average of 35±32 leaves, while after a 6-week vernalisation treatment (Figure 6. 7B) plants produced significantly fewer leaves (27±21 leaves, $p<0.001$, Mann-Whitney U test). There was a significant negative correlation between flowering time and branch number with and without vernalisation (Table 6. 1; Figure 6. 7 B, D). This suggests the latest flowering individuals produced the highest number of leaves but most of these leaves did not develop a branch, regardless of vernalisation treatment. Significant differences ($p=0.001$, Mann-Whitney U test) were observed in branch number between treatments, with significantly more plants producing fewer branches under no vernalisation conditions.

Table 6. 1: Spearman's rank correlation coefficient for leaf number and branch number correlated with flowering time. Correlation coefficient and t-statistic probability values of significance are listed. Branch number was calculated as the proportion of leaves that produced a branch.

Trait		Leaf number		Branch number	
	Treatment	VERN	NVERN	VERN	NVERN
Number of days to BBCH51	VERN	0.631 p<0.001		-0.865 p<0.001	
	NVERN		0.719 p<0.001		-0.591 p<0.001

The effect of vernalisation on branch number was specific to crop type (Figure 6. 7E). No significant effect of vernalisation on branch number was detected in exotic, semi-winter oilseed rape, synthetic, and winter vegetable crop types ($p= 0.400$, $p=0.857$, $p= 0.700$, $p= 0.200$ respectively, Mann-Whitney U test). Vernalisation also had no significant ($p=0.797$, Mann-Whitney U test) effect on branch number in kale crop types. Accoding to the OREGIN DFFS population, the early flowering cultivar Rapid Cycling Rape is classified as a kale crop type, along with the late flowering Siberian and Japanese kale cultivars. Removal of this cultivar from the analysis resulted in the detection of a significant ($p=0.071$; Mann Whitney U test) positive relationship between vernalisation treatment and branch number. In contrast, vernalisation had a significant ($p<0.001$, Mann Whitney U test) negative effect on branch number in spring oilseed rape, with more branches produced without vernalisation than with vernalisation. The vernalisation treatment may therefore have impacted the fitness of spring oilseed rape plants, delaying flowering time (discussed in Chapter 4) and negatively affecting branch number. Due to the positive correlation between branch number and seed yield, spring cultivars are therefore expected to exhibit reduced yields following a vernalisation treatment compared with if they had been grown without vernalisation, although I did not test this.

Swede, winter fodder and winter oilseed rape crop types exhibited a significant increase in branch number with vernalisation compared to without ($p<0.001$ in all cases, Mann-Whitney U test). These lines have a vernalisation requirement and exhibit an acceleration in flowering time after a 6-week vernalisation treatment (discussed in Chapter 4). This acceleration in flowering time is coupled with an increase in branch number and, although untested, my data suggests an increase in seed yield would be expected. Taken together the results of this analysis suggest the vernalisation status of a *B. napus* plant can influence both flowering time and overall plant architecture, and therefore most likely, yield.

6.3. Discussion

Vernalisation is known to have a quantitative effect on flowering time in *A. thaliana* (Shindo *et al*., 2006, Angel *et al*., 2011, Li *et al*., 2014, Berry *et al*., 2015, Duncan *et al*., 2015). In this Chapter I investigated a role for vernalisation in determining both flowering time and inflorescence architecture in both *A. thaliana* and *B. napus*. Phenotypic analyses of accessions and cultivars from both species revealed duration of vernalisation not only had a quantitative effect on flowering time, but directly impacted branch number and the overall architecture of a plant. In *A. thaliana*, insufficient vernalisation led to delayed flowering, an increase in cauline branch number and an inhibition of rosette branch outgrowth. The strength of the response to vernalisation treatment however was accession dependent. In *B. napus* insufficient vernalisation resulted in axillary meristems remaining vegetative and a reduction in branch number, however this effect was crop type specific. In *A. thaliana*, complete vernalisation was coupled with *FT* and *BRC1* expression activation during vernalisation. While time constraints did not allow this to be tested in *B. napus*, from the preliminary study presented here, I propose a role for *FLC* expression in determining the developmental fate of axillary meristems. High *FLC* expression inhibits the transition from vegetative to floral development, and axillary meristems thus remain vegetative. Stable repression of *FLC* during prolonged vernalisation would lead to activation of *FT* and *BRC1* expression and a switch in axillary meristem fate from vegetative to reproductive development.

The results described here show parallels to a recently published study (Lazaro *et al*., 2018) in *Arabis alpina*, a perennial member of the Brassicaceae family. Due to its perennial nature branches in *A. alpina* form from axillary meristems that are present prior to vernalisation. Meristems that form during or after vernalisation are not competent to receive vernalisation signals and remain dormant until the following season. Lazaro *et al*., (2018) demonstrated by increasing the duration of vernalisation, flowering time was accelerated, and branch number was increased. Lazaro *et al*., (2018) also detected an increase in inflorescence length and silique number, suggesting an increase in reproductive fitness in response to vernalisation.

In *A. alpina* axillary meristem outgrowth was associated with increased expression of *FRUITFUL*, *LEAFY* and *APETALA1* homologues, and a decrease in *PERPETUAL FLOWERING 1 (PEP1)* and the *TERMINAL FLOWER1* homologue (Lazaro *et al*., 2018). Interestingly *PEP1*, the homologue of *FLC*, was found to repress both flowering and inflorescence outgrowth and exhibited a basipetal expression pattern. Differential *PEP1* expression was detected between floral and vegetative axillary meristems, with the highest *PEP1* expression detected in vegetative meristems. Due to the functional homology between *FLC* and *PEP1*, it is therefore reasonable to hypothesise that *FLC* and its homologues in *B. napus* could contribute to axillary meristem activation and overall plant architecture in *A. thaliana* and *B. napus*. I have shown that non-saturating vernalisation conditions inhibited rosette branch outgrowth in *A. thaliana*, and like *A. alpina* the lowermost axillary meristems remained vegetative. The strength of inhibition on rosette branch outgrowth was in part dependent on *FLC* allelic variation, and therefore its presumed epigenetic state (Angel *et al*., 2010, Berry *et al*., 2015).

In *A. alpina* axillary meristems remain vegetative due to high *PEP1* and low *FT* expression (Lazaro *et al*., 2018). Insufficient vernalisation lead to a delay in flowering time due to reactivation of *PEP1* expression, expression profiles that are similar to those observed for *FLC* in the *A. thaliana* accession Lov-1 (Coustham *et al*., 2012, Li *et al*., 2014) and for *BolFLC.C02* in *B. oleracea* (Irwin *et al*., 2016). *A. alpina* plants that had not received sufficient vernalisation treatment length exhibited floral reversion phenotypes and low seed set (Lazaro *et al*., 2018). I observed floral reversion phenotypes in one accession of *A. thaliana* with short vernalisation treatments. Aerial rosettes formed in Lov-1 plants that had received non-saturating vernalisation treatments (≤6 weeks). This phenotype was previously reported in the Scottish accession Sy-0 (Grbic and Bleecker, 1996, Poduska *et al*., 2003, Wang *et al*., 2007) and in transgenic *A. thaliana* that constitutively overexpress *FLC* (Wang *et al*., 2007). Taken together these data highlight a correlation between *FLC* expression, plant architecture and reproductive strategy.

I also detected a correlation between duration of vernalisation and plant architecture in winter oilseed rape. A higher proportion of axillary meristems remained vegetative, inhibiting branch outgrowth, with decreasing lengths of vernalisation. As branch number is known to positively correlate with seed yield (Li *et al.*, 2016), I predict increasing the vernalisation length would increase yield. The response to vernalisation treatment was dependent on cultivar, and therefore the genetic status of the plant. Cabriolet had the weakest response to vernalisation treatment producing similar numbers of branches with 8-, 10- and 12-weeks vernalisation, however significantly fewer branches after 6-, 4- and 0-weeks vernalisation. Darmor exhibited the strongest response and produced significantly fewer branches after 10-weeks vernalisation compared with 12-weeks. Due to the significant correlation between flowering time and branch number in *B. napus* I hypothesise that *BnaFLC.A02* and *BnaFT.A02*, candidate genes I identified in Chapter 5 to associate with flowering time variation, might also contribute to the variation in branch number detected between the Cabriolet and Darmor but this requires further testing.

Branch number correlated significantly with flowering time in *B. napus*, however the effect of vernalisation on branch number was crop type dependent. For example, Chinese semi-winter oilseed rape cultivars exhibited no significant vernalisation response in branch number. Winter oilseed rape cultivars exhibited a significant increase in branch number in response to vernalisation, while spring oilseed rape cultivars in contrast produced significantly fewer branches after vernalisation. Vernalisation treatment could therefore have a positive or negative effect on yield depending on the vernalisation requirement of a plant.

It is apparent from my preliminary study that genetic variation for vernalisation requirement not only affects flowering time, but also has pleiotropic effects on plant architecture and ultimately yield. If these traits could be uncoupled the resulting genetic information could enhance a Breeder's toolkit when developing new varieties for increased yields and food security.

Chapter 7: Discussion and Conclusions

7.1. Variation for vernalisation requirement in *Brassica napus*: a tale of three genes

B. napus has an almost worldwide distribution and loss of vernalisation requirement has facilitated adaptation to different growing environments. In the present study I have characterised the vernalisation requirement of diverse *B. napus* cultivars were characterised in the present study. Associated with crop type, variation for vernalisation requirement affected flowering time (Chapter 4 and Chapter 5) and inflorescence architecture (Chapter 6) and is therefore predicted to influence yield in *B. napus*. Due to extensive prior research carried out in the Caroline Dean Laboratory (Clarke and Dean 1994, Johanson *et al*., 2000, Shindo *et al*., 2005) orthologues of *FRI* were hypothesised to be responsible for variation in vernalisation requirement in *B. napus* therefore this was investigated. However, characterisation of *FRI* orthologues (Chapter 3) and separate genetic mapping experiments (Chapter 4 and Chapter 5) identified alternative candidate genes, including orthologues of *FLC* and *FT*, as important for variation in vernalisation requirement and vernalisation response in *B. napus* (Chapter 4 and Chapter 5).

Natural and induced allelic variation at *BnaFRI* genes had no consistent significant effect on flowering time or *FLC* expression (Chapter 3). Unlike in *A. thaliana*, allelic variation at *FRI* orthologues in *B. napus* is unlikely to be the main factor responsible for variation in vernalisation requirement detected between crop types. Selection by Breeders at loci distinct from *FRI* such as orthologues of *FLC* and *FT* have resulted in loss of vernalisation requirement and facilitated adaptation to alternative growing environments. The characterisation of allelic variation at major flowering genes like orthologues of *FRI*, *FLC* and *FT* should provide Breeders with a catalogue of genetic markers that confer specific flowering time phenotypes. The targeted introduction of genetic markers that have known functional consequences will help generate new cultivars with predictable flowering times, reducing crop losses and improving yields.

7.2. Is there a role for *FRIGIDA* in determining variation in flowering time of *Brassica napus*?

In *A. thaliana* sequence variation at *FRI* can account for approximately 70% of variation in flowering time observed in worldwide collections of natural accessions (Shindo *et al*., 2005). Variation at *FRI* has provided a selective advantage for earliness of flowering, increasing reproductive fitness in certain environmental niches (Le Corre *et al*., 2002, Shindo *et al*., 2005, Werner *et al*., 2005, Korves *et al*., 2007). Many haplotypes have been identified, and loss of vernalisation requirement via non-functional FRI protein has evolved independently at least 20 times during the evolution of *A. thaliana* (Shindo *et al*., 2005). In other Brassicaceae such as *A. lyrata* (Kuittinen *et al*., 2008), *C. rubella* (Guo *et al*., 2012), *B. oleracea* (Irwin *et al*., 2012, Fadina *et al*., 2013), *B. rapa* (Fadina *et al*., 2013) and more recently *B. napus* (Yi *et al*., 2018, Chapter 3) there is a striking absence of loss of function mutations found at *FRI* orthologues, but many *FRI* haplotypes have been described. This poses two questions (1) What is the consequence of natural allelic variation at *BnaFRI* on the flowering time of *B. napus*? (2) What consequence would non-functional variation at *BnaFRI* genes have on flowering time and *BnaFLC* expression in *B. napus*?

Prior to the start of this project no detailed characterisation of the allelic variation at *BnaFRI.A03*, *BnaFRI.C03*, *BnaFRI.A10* and *BnaFRI.C09* had been carried out. A single publication (Wang *et al*., 2011b) detailed the number and sequence of *BnaFRI* genes present in *B. napus* and expression of all four genes were detected in three cultivars of *B. napus*. Variation at one gene, *BnaFRI.A03*, was found to have an association with flowering time (Wang *et al*., 2011b) however this had not been characterised until recently (Chapter 3, and by Yi *et al*., 2018). As in previous reports (Wang *et al*., 2011b, Yi *et al*., 2018), there was a near absence of non-functional mutations at *BnaFRI* in the OREGIN diversity set population of *B. napus* (Chapter 3). Of the 380 *BnaFRI* genes sequenced from 95 cultivars, multiple alleles were detected but only one was predicted to be non-functional - *BnaFRI.C09-b* found in a late flowering swede cultivar - an unlikely candidate for conferring early

flowering. The remaining *BnaFRI* sequences detected in the population were predicted to encode complete and functional BnaFRI proteins, perhaps suggesting a requirement for BnaFRI function in *B. napus*.

Introduction of a functional *FRI* allele into an *A. thaliana* accession with non-functional *fri* significantly delays flowering (Lee *et al*., 1994b) and increases the quantitative expression of *FLC*. The translocation event that moved *FRI* from its ancestral location, where it is found in other members of the Brassicaceae, to its current position in the *A. thaliana* genome (Kuittinen *et al*., 2008, Irwin *et al*., 2012) may have contributed to the overall high frequency of non-functional variation at *FRI*. In *A. thaliana*, *FRI* is located roughly 0.269Mbp from the tip of chromosome 4 sandwiched between the nucleolar organisation region and a heterochromatic region, called the "knob," (Fransz *et al*., 2000, Drouaud *et al*., 2006). This region was identified as a "hot spot" for crossover rates during meiosis (Drouaud *et al*., 2006) providing a possible explanation for the recent positive selection for earliness of flowering (Le Corre *et al*., 2002) conferred by variation at *FRI* in *A. thaliana*.

My transgenic and mutational approaches carried out in the present study support recent findings (Yi *et al*., 2018) that *BnaFRI* genes of *B. napus* are weak controllers of flowering time and *BnaFLC* expression. Constructs containing variants of *BnaFRI.A03* did not fully complement the *fri* loss of function mutation of Col-0 and induced mutations at *BnaFRI* genes in *B. napus* had no significant effect on flowering time or *FLC* expression (Chapter 3). Although able to confer late flowering, variation at *FRI* orthologues in *A. lyrata* and *B. oleracea* too had minimal (Kuittinen *et al*., 2008), if any (Irwin *et al*., 2012), effect on flowering time. *BolFRI.C03* and *BnaFRI.A03* exhibit extremely high sequence conservation within coding regions (Fadina *et al*., 2013, Fadina & Khavkin 2014) however transformation of both genes into *A. thaliana* Col-0 resulted in contrasting flowering time phenotypes (Irwin *et al*., 2012 versus Chapter 3, and Yi *et al*., 2018). In a construct containing the *A. thaliana FRI* promoter sequence, *BolFRI.C03* conferred late flowering (Irwin *et al*., 2012) while a construct containing the native *BnaFRI.A03* promoter sequence *BnaFRI.A03* conferred early flowering (Chapter 3,

and in Yi *et al*., 2018). We could speculate that the different effects on flowering time are most likely due to differences in expression conferred by promoter variation. Although work has not been carried out to determine how *BnaFRI* gene expression differs from *A. thaliana*, both in quantitative level and localisation throughout development, it is interesting to consider that *FRI* in *A. thaliana* may have acquired an alternative role in determining the vernalisation requirement through promoter variation. This is currently being investigated by the Judith Irwin Laboratory.

The striking parallels between *B. napus*, *B. oleracea*, *A. lyrata* and *C. rubella* highlight *A. thaliana* as an anomaly. *FRI* is indispensable in determining the vernalisation requirement of *A. thaliana* (Shindo *et al*., 2005). However, my work and others (Kuittinen *et al*., 2008, Irwin *et al*., 2012, Yi *et al*., 2018) suggest *FRI* orthologues are not essential for determining the vernalisation requirement of other members of the Brassicaceae family. One could hypothesise that multiple genes are responsible for determining the vernalisation requirement in these species, genes which may have been lost during the evolution of *A. thaliana*. Compensation from other genes such as *FRI-like* or autonomous pathway genes might reduce the selective advantage for the loss of *FRI*. It is also possible in the paleo- and neo-polyploid *Brassica* species that the presence of multiple copies of genes has facilitated the emergence of novel gene function (Conant *et al*., 2014, Jones *et al*., 2018) allowing *FRI* to evolve an alternative role in controlling flowering time. My work demonstrates *BnaFRI* is not a major regulator of flowering time in *B. napus* however, we are no closer in understanding what other possible function *FRI* orthologues serve in *Brassica* species. A transcriptomic or proteomic analysis of *B. napus* with loss of function *Bnafri* might reveal targets, alternative to *BnaFLC*, that are regulated or compensated for by *BnaFRI*.

7.3. Photoperiod and vernalisation requirement

Associations between mutations at *FT* and loss of vernalisation requirement are reported in *Medicago truncatula* (Jaudal *et al*., 2013), and *Lupinus angustifolius*

(Nelson *et al*., 2017). In both species, a de-repression of *FT* expression was correlated with InDel variation within non-coding regulatory regions of *FT* and proposed to be responsible for the early flowering phenotype (Jaudal *et al*., 2013, Nelson *et al*., 2017). In the present study I detected an association between *BnaFT.A02* and vernalisation requirement (Chapter 5). Polymorphisms at *BnaFT.A02* were found predominately in non-coding regions known from *A. thaliana* for their importance in controlling the temporal and spatial expression of *FT* (Helliwell *et al*., 2006, Adrian *et al*., 2010). Interestingly an amino acid substitution I48L detected in the early flowering cultivar Cabriolet is perhaps unlikely to be responsible for the early flowering phenotype as the same mutation was found to be associated with late flowering in an orthologous gene in *B. rapa* (Zhang *et al*., 2015, Schiessl *et al*., 2017b).

In *A. thaliana FT* integrates signals from the vernalisation and photoperiod pathways and an increase in its expression is associated with the floral transition (Hepworth *et al*., 2002, Wigge *et al*., 2005, Helliwell *et al*., 2006, Wigge, 2011). In *A. thaliana* the expression of *FT* is inhibited by FLC protein and promoted by CO protein (Hepworth *et al*., 2002, Helliwell *et al*., 2006, Wenkel *et al*., 2006, Tiwari *et al*., 2010). In *B. napus* however, the mechanism by which *BnaFT* is regulated remains relatively unknown. In my study, inhibition of *BnaFT.A02* expression correlated with late flowering (Chapter 5) which parallels a study in *B. rapa* (Zhang *et al*., 2015). In both cases (Chapter 5, and Zhang *et al*., 2015) the *BraFT.A07* and *BnaFT.A02* alleles found in the late flowering cultivar were not expressed, even after the floral transition, while expression of the allele found in the early flowering cultivar was detected before, and in response to vernalisation.

My characterisation of the phenotypic consequence of allelic variation at multiple flowering time genes has contributed to our understanding of the floral network in *B. napus*. I detected a complex genetic interaction between *BnaFLC.A02* and *BnaFT.A02* which resulted in contrasting flowering time phenotypes that were dependent on genotype and vernalisation treatment (Chapter 5). Interestingly, genotypic data for *BnaFLC.A02* and *BnaFT.A02* across the OREGIN diversity set

suggests Breeders have selected for cultivars with a strong vernalisation requirement, that flower rapidly after vernalisation. Most winter oilseed rape cultivars were found to carry an allele of *BnaFT.A02* which associated with a strong vernalisation requirement and an early allele of *BnaFLC.A02* ensuring the plants flower early with short periods of vernalisation. By selecting for specific alleles of *BnaFLC.A02* and *BnaFT.A02*, Breeders have ensured that many winter oilseed rape cultivars are prevented from flowering until after winter has passed, but flower rapidly even after short winters.

It is crucial for reproductive success, and therefore crop yields, that winter oilseed rape flowering is inhibited until after winter, but global average temperatures are predicted to increase over the next 100 years (Brown & Caldeira 2017). Milder and more variable winter temperatures are predicted in mid-high latitude countries (Luterbacher *et al*., 2004), posing important challenges for winter oilseed rape breeding. Developing new varieties that can complete their life cycle at higher temperatures, or with shorter lengths of vernalisation, will ensure that farmers can continue to grow this crop in its current geographical range.

7.4. Orthologues of *FLOWERING LOCUS C* are master regulators of flowering time

Unlike *FRI*, orthologues of *FLC* are often found to associate with flowering time in the Brassicaceae (Wang *et al*., 2009a, Yuan *et al*., 2009, Aikawa *et al*., 2010, Zhao *et al*., 2010, Hou *et al*., 2012, Guo *et al*., 2012, Wu *et al*., 2012, Kemi *et al*., 2013, Xiao *et al*., 2013, Yi *et al*., 2014, Baduel *et al*., 2016, Irwin *et al*., 2016, Schiessl *et al*., 2017a, Taylor *et al*., 2017, Lazaro *et al*., 2018), but also in other dicot and monocot species (Reeves *et al*., 2006, Ruelens *et al*., 2013, Kumar *et al*., 2016). Nine orthologues of *FLC* have previously been characterised in *B. napus* (Zou *et al*., 2012) and in the present study variation at *BnaFLC.A02* and *BnaFLC.A10* were significantly correlated with variation in flowering time (Chapter 4 and Chapter 5). In agreement with previous publications (Hou *et al*., 2012, Schiessl *et al*., 2017a) *BnaFLC.A10* could be considered a major determinant of flowering time and an

important gene for controlling the vernalisation requirement. Here I demonstrated that cultivars of *B. napus* that required vernalisation for flowering exhibited the highest expression levels of *BnaFLC.A10* before and after vernalisation (Chapter 4). Interestingly *BnaFLC.A10* was often the most highly expressed of the *BnaFLC* homologues. Variation in *BnaFLC.A10* expression between cultivars accounted for a higher proportion of variation in flowering time compared with other *BnaFLC* homologues perhaps suggesting that *BnaFLC.A10* plays a more dominant role in controlling the vernalisation requirement. Interestingly, after accounting for variation at *BnaFLC.A10*, contribution from another *BnaFLC* homologue, *BnaFLC.A02*, became important for flowering time (Chapter 5). Whether *BnaFLC* in *B. napus* act independently or additively to control flowering requires further exploration but it is clear that the importance of each *BnaFLC* that contribute to control flowering time is dependent on genotype.

Non-coding sequence variation influences gene expression before, during and after vernalisation, altering H3K27me3 deposition and silencing rates at *FLC* (Coustham 2014, Li *et al*., 2014, Li *et al*., 2015, Hawkes *et al*., 2016, Questa *et al*., 2016). I detected interesting parallels between *A. thaliana* and *B. napus* at *FLC* orthologues. I identified polymorphisms at non-coding regions known from *A. thaliana* to be important for influencing *FLC* expression (Coustham 2014, Li *et al*., 2014, Li *et al*., 2015, Hawkes *et al*., 2016, Questa *et al*., 2016), providing a mechanistic explanation for their associations with flowering time (Chapter 4 and Chapter 5). Variation in expression level of *BnaFLC.A10* was significantly correlated with variation in flowering time with and without vernalisation (Chapter 4). This variation in expression level was associated with cis-variation at *BnaFLC.A10* and two haplotypes were identified that had contrasting effects on flowering time. *BnaFLC.A10* contains a mutation that alters a splice site within distal *COOLAIR* (Hawkes 2017). This mutation is associated with late flowering and elevated levels of *FLC* (Li *et al*., 2015, Hawkes *et al*., 2016). Other *BnaFLC* homologues appear not to contain this splice site mutation (Hawkes 2017), which explain how *BnaFLC.A10* might play a more dominant role in controlling flowering time. Although this alternative distal *COOLAIR* splice site at *BnaFLC.A10* is conserved between spring and winter crop types of *B. napus* (Hawkes 2017) I identified additional mutations

within key structural regions of *COOLAIR* which could modulate the expression of *BnaFLC.A10* in warm conditions (Chapter 4). It is also possible that mutations at other regions of *BnaFLC.A10* affect its function to inhibit flowering time. For example, a splice-site mutation at *BraFLC.A10*, an orthologue of *BnaFLC.A10*, has been reported to associate with flowering time in *B. rapa* (Yuan *et al*., 2009). In this case, variation that disrupted the translation of a functional BraFLC.A10 protein resulted in earlier flowering. The presence of such a mutation at *BnaFLC.A10* in the OREGIN DFFS population remains to be explored.

BnaFLC.A10 and *BnaFLC.A02* exhibited polymorphisms at the VAL1 binding domain regions of intron 1. Similar mutations have previously been characterised in *A. thaliana*, affecting *FLC* expression dynamics during and after vernalisation (Coustham *et al*., 2012, Li *et al*., 2014, Questa *et al*., 2016). Allelic variation within important cis-regulatory regions that are hypothesised to affect *BnaFLC* expression dynamics provide a mechanism for variation in flowering time in *B. napus*.

A large proportion of the variation for vernalisation response and life history strategy in natural *A. thaliana* accessions can be explained by sequence polymorphisms at cis-regulatory regions of *FLC* (Li *et al*., 2014). Interestingly, Breeders have also selected for cis-regulatory variation at *BnaFLC* genes to modify the vernalisation requirement and response of *B. napus* cultivars. Work is currently underway in the Judith Irwin Laboratory to determine the consequence of these cis-regulatory polymorphisms at *BnaFLC.A10* and *BnaFLC.A02* on flowering time, *FLC* expression regulation and histone modification.

7.5 Multiple alleles of flowering time genes are maintained and conserved in *Brassica* spp.

B. napus evolved from inter-specific hybridisation between *B. rapa* and *B. oleracea* (Kagale *et al*., 2014) and high levels of sequence identity between homologous genes is reported (Schiessl *et al*., 2017b). The work described in this Thesis highlights not

only that homologous gene sequences are conserved between three *Brassica* species, but often the same multiple alleles and sequence features that have been reported in *B. rapa* and *B. oleracea* are also present within *B. napus*. For example, an insertion within the same region of intron 2 in *BraFT.A07* as detected in *BnaFT.A02* was associated with delayed flowering in *B. rapa* (Zhang *et al*., 2015). This allele, like *BnaFT.A02-Dar*, was not expressed in a late flowering *B. rapa* cultivar (Zhang *et al*., 2015). *Brassica FT* sequence and gene expression dynamics therefore appear to be conserved between *B. rapa* and *B. napus*. In addition to *BnaFT.A02*, all four alleles of *BnaFRI.C03* detected in *B. napus* (Chapter 3) were also reported for *BolFRI.C03* in *B. oleracea* (Irwin *et al*., 2012). I detected the key amino acid substitutions identified by Irwin *et al*., (2012) and found to be associated with early and late flowering crop types of *B. oleracea*, in *B. napus*. In addition, *BraFLC.A10* and *BraFLC.A02*, the homologues of *BnaFLC.A10* and *BnaFLC.A02* respectively, are associated with variation for vernalisation response in *B. rapa* similar to their *B. napus* homologues described in Chapter 4 and Chapter 5 (Yuan *et al*., 2009, Zhao *et al*., 2010, Wu *et al*., 2012, Xiao *et al*., 2013).

Maintenance of multiple alleles in *B. napus* that exhibit high conservation, and in many instances 100% sequence identity, with those alleles reported in *B. rapa* (Zhang *et al*., 2015) and *B. oleracea* (Irwin *et al*., 2012, Fadina *et al*., 2013, Fadina & Khavkin, 2014) provides further evidence that *B. napus* arose from multiple hybridisation events between *B. rapa* and *B. oleracea* (Kagale *et al*., 2014). Flowering time variation appears to be controlled by homologous genes, often by similar allelic variation. Characterising the allelic variation present at one gene in *B. rapa* or *B. oleracea* could be used to infer the function of its orthologue in *B. napus* and *vice versa*. Mining germplasm for functional allelic variation in *B. rapa*, *B. oleracea* and *B. napus* will undoubtedly contribute to the genetic improvement of all three species.

7.6. Targeting flowering time could have a pleiotropic effect on yield

I demonstrated in a preliminary study the pleiotropic effect of vernalisation requirement on flowering time and on inflorescence architecture in *A. thaliana* and *B. napus* (Chapter 6). Like previous reports (Grbic and Bleecker 1996, Poduska *et al*., 2003, Kalinina *et al*., 2002, Wang *et al*., 2007, Huang *et al*., 2013) variation for vernalisation requirement had a significant effect on branch number and inflorescence architecture in both species. Expression of functional *FRI* in *A. thaliana* resulted in an increase in overall branch number but, possibly as a consequence of high *FLC* expression, an inhibition of the lowermost branches. The strength of branch inhibition was dependent on flowering time and *FLC* allelic variation. Branch inhibition was reduced with increasing lengths of vernalisation.

B. napus exhibited a similar inflorescence phenotype in response to vernalisation. Cultivars that required vernalisation to flower exhibited an inhibition of the lowermost branches, but this inhibition decreased with increasing lengths of vernalisation. Spring and Chinese semi-winter oilseed rape cultivars are early flowering in the absence of vernalisation due to low *FLC* expression levels (Chapter 4). These cultivars exhibited very little branch inhibition under no vernalisation conditions and had inflorescence architecture phenotypes that were similar to completely vernalised winter oilseed rape. Similar to a recent finding in *A. alpina* (Lazaro *et al*., 2018) I propose that *FLC* expression inhibits the floral transition in all meristems, including those that develop into branches. This pattern of inhibition acts basipetally from the apical meristem downwards and reduction in *FLC* expression by vernalisation leads to a switch from vegetative to reproductive development and consequently branch elongation. A plant's response to vernalisation, controlled by *FLC*, therefore affects its overall phenotype, influencing both flowering time and inflorescence architecture.

Response to vernalisation, controlled by *FLC* expression, affects flowering time and inflorescence architecture, but also pod and seed size (Dr M Brueser and Dr S

Penfield, personal communications) in *B. napus*. It is therefore important to consider the pleiotropic effects of altering *FLC* on traits that might influence yield. Under field conditions, spring cultivars of oilseed rape that do not require vernalisation tend to produce lower seed yield compared with vernalisation requiring winter cultivars (3 t/ha for spring oilseed rape versus 5 t/ha for winter oilseed rape; https://cereals.ahdb.org.uk). Interestingly spring cultivars of oilseed rape germinate poorly when the mother plant is grown under vernalising conditions (Dr S Penfield/C O'Neil, personal communications). Targeting *FLC* to breed for early flowering winter oilseed rape with weak vernalisation requirements and response (Bayer CropScience, personal communications) may negatively impact fitness by reducing tolerance to cold temperatures, ultimately leading to low seed set. The alternative is also possible and targeting *FLC* for strong vernalisation requirements may also reduce seed yield, resulting in an extremely late cultivar that is ideal for vegetable production but not ideal for producing seed. Understanding how multiple copies of *BnaFLC* genes, and the alleles detected at them, contribute to flowering time will enable Breeders to fine-tune the phenotypes of *B. napus* cultivars. My work described in this Thesis has revealed variation at two *FLC* genes, *BnaFLC.A10* and *BnaFLC.A02*, can contribute to variation in vernalisation requirement and response to vernalisation. However due to the tight correlation between genetic variation at *FLC*, flowering time, inflorescence architecture and seed yield, uncoupling these traits for yield increases may be challenging. A balance between selecting for optimal flowering time and increasing yield must be established for the continued genetic improvement of *Brassica* crops.

7.7. Conclusion

My work demonstrates that *BnaFLC* has been a major target facilitating the evolution of diverse *B. napus* crop types with contrasting vernalisation requirements and responses. Unlike *A. thaliana*, genetic changes at *BnaFRI* have not contributed significantly to the diversification of flowering time phenotypes in *B. napus* suggesting adaptation to new environments have evolved through subtly different processes. Genetic changes at genes downstream of *BnaFLC* in the floral network,

like *BnaFT*, have contributed further to the flowering times of *B. napus* cultivars, with pleiotropic consequences.

To continue this research (1) Further elucidation of *BnaFRI* function in *B. napus* and *FRI* orthologues in related Brassicaceae species would help unravel this elusive protein. If allelic variation at *BnaFRI* genes has no effect on flowering time, is the variation a consequence of genetic drift or does it serve another function? (2) The sequence variation at *BnaFLC* genes requires further characterisation. Are the important cis-regulatory regions known in *A. thaliana* also important in *Brassica*? This knowledge could be used to predict the functionality of many more *BnaFLC* genes, creating a catalogue of alleles that could be used in breeding programmes for new varieties. (3) Allelic variation at *FLC* contributes to flowering time and inflorescence architecture in both *B. napus* and *A. thaliana*. A mechanistic understanding of this process may reveal additional loci that could be targeted to alter the flowering time and inflorescence architecture of *B. napus*.

Chapter 8: Appendix

8.1. Supplementary Figures

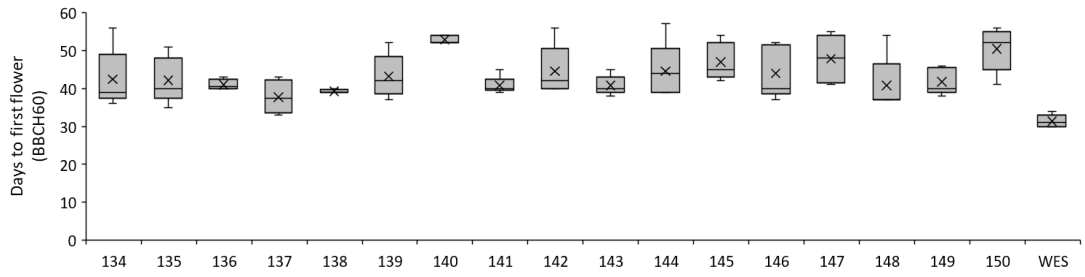


Figure S. 1: Flowering time under no vernalisation conditions of all *Bnafri* mutants (134-148), the *BnaFRI* wild-type line (149), the original wild-type winter oilseed rape parent line (150) and the spring oilseed rape cultivar Westar-10 (Wes). No significant difference was detected between the *BnaFRI* wild- line (149) and *Bnafri* mutant lines (134-148). Flowering times are displayed as boxplots, illustrating the median, mean (x) and range of values measured from five plants per line grown under glasshouse conditions.

Figure S. 2: *In silico* expression of nine *BnaFLC* for 101 *B. napus* cultivars. The quantitative expression level of *BnaFLC* are plotted for each cultivar using sequencing reads from the mRNA-seq dataset (Harper *et al*. 2012, Lu *et al*. 2014). Each *BnaFLC* gene (in this case A3= *BnaFLC.A03b*, A3RC= *BnaFLC.A03a*, C3RC= *BnaFLC.C03a*) is plotted along the x-axis and their expression value along the y-axis. This data was used as a trait in the Associative Transcriptomics analysis detailed in Chapter 4.

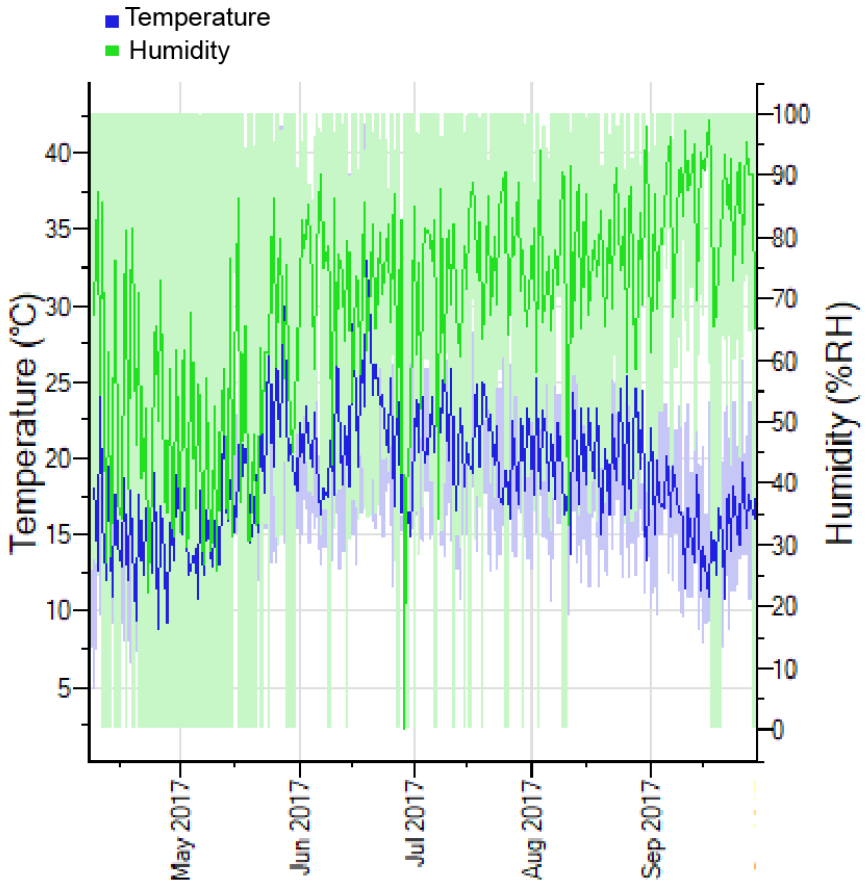


Figure S. 3: Temperature and humidity inside Keder poly-tunnel during April-September 2017 (Chapter 5). Temperature and humidity recorded at half-hourly intervals for the duration of the experiment by Tinytag® data logger.

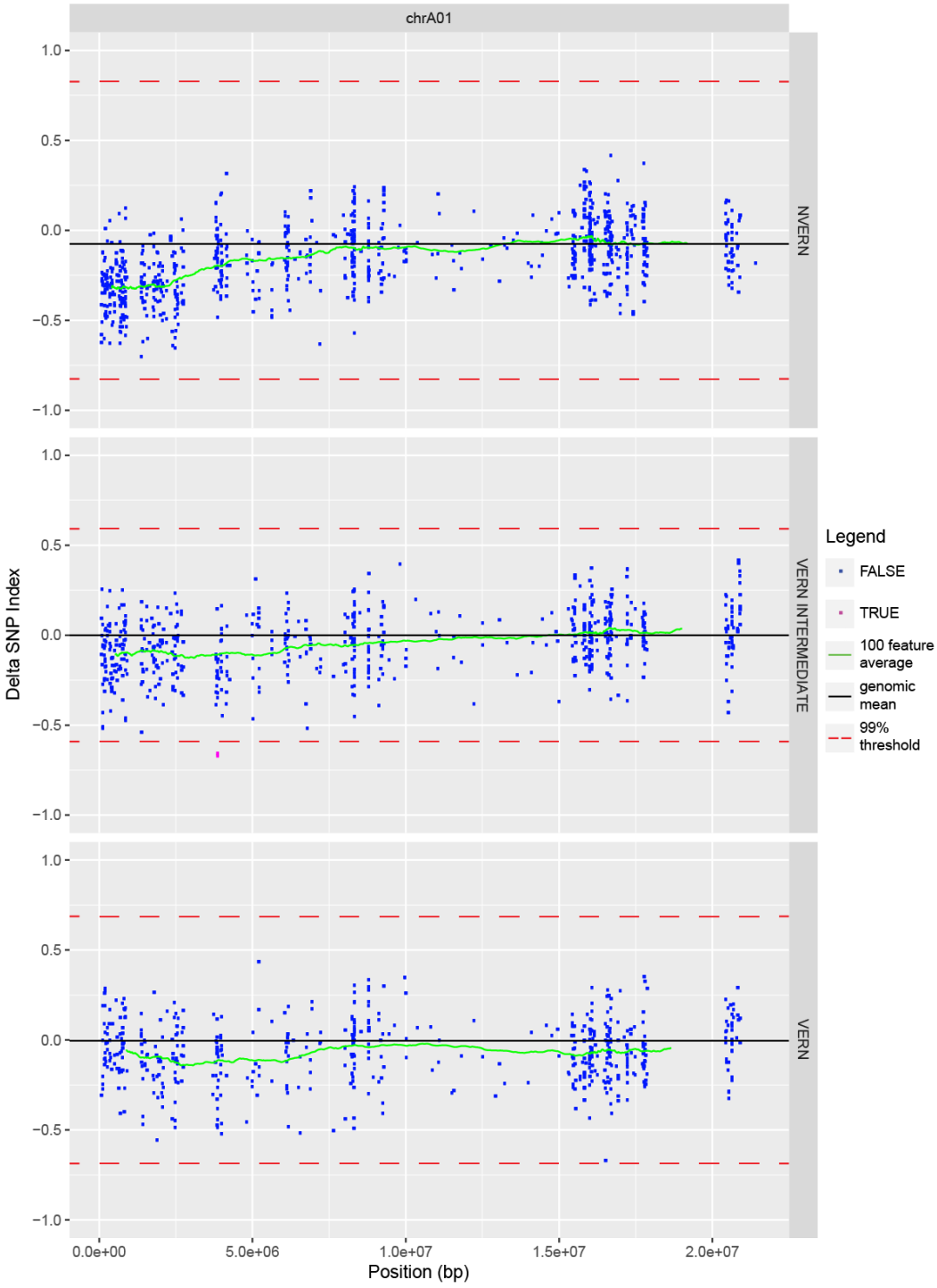


Figure S. 4: Chromosome A01. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in pink, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

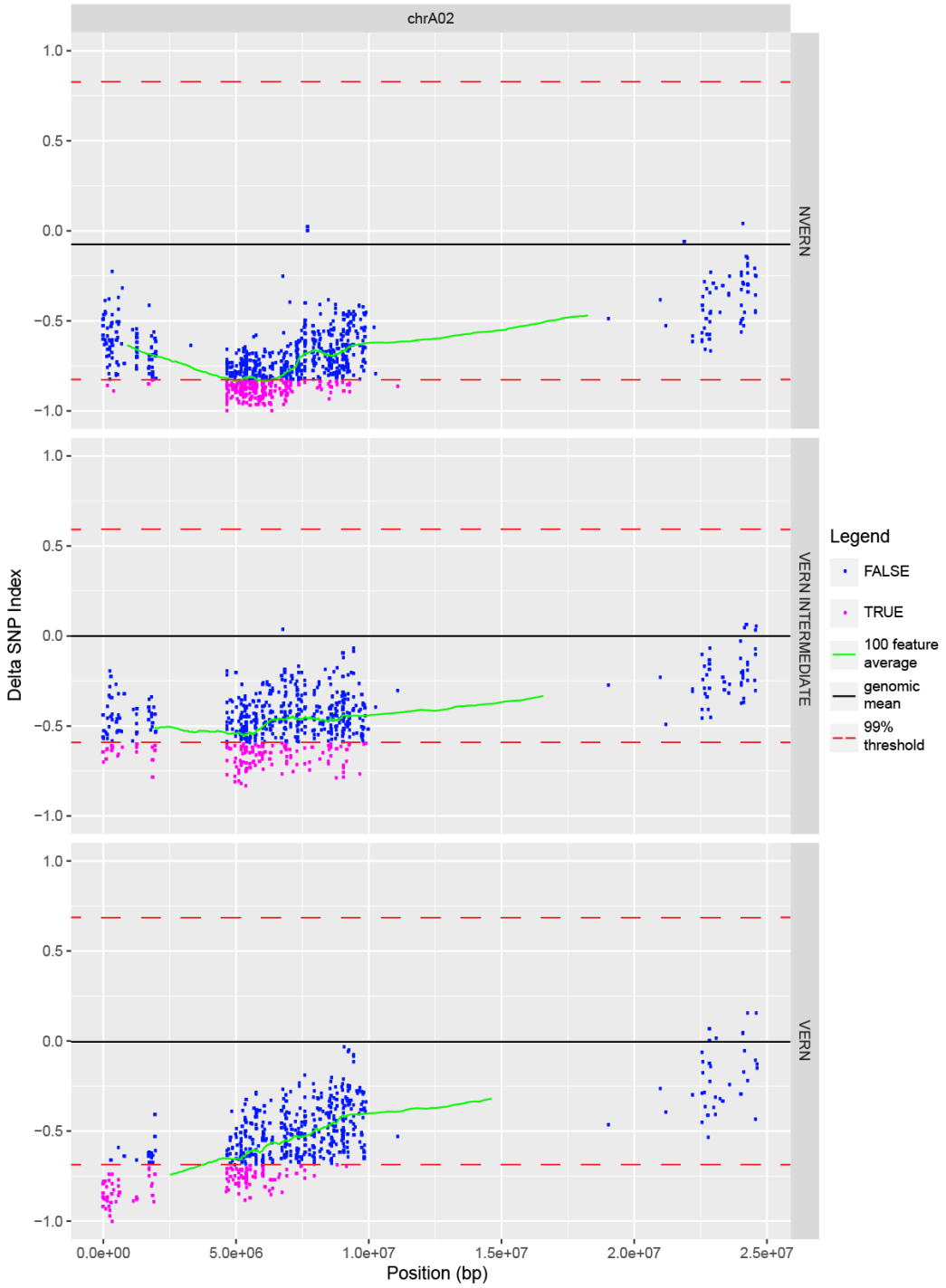


Figure S. 5: Chromosome A02. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in pink, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

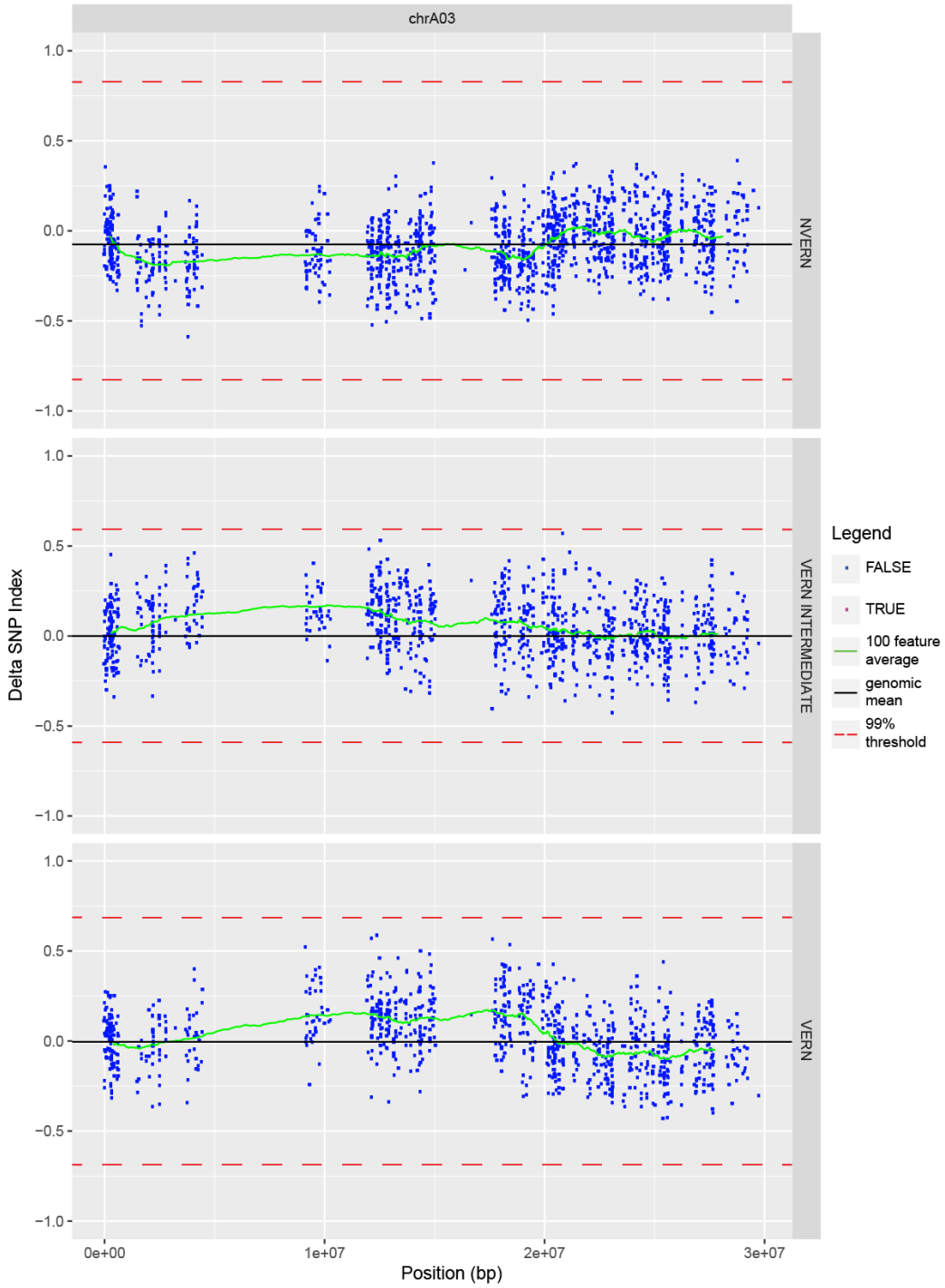


Figure S. 6: Chromosome A03. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in pink, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

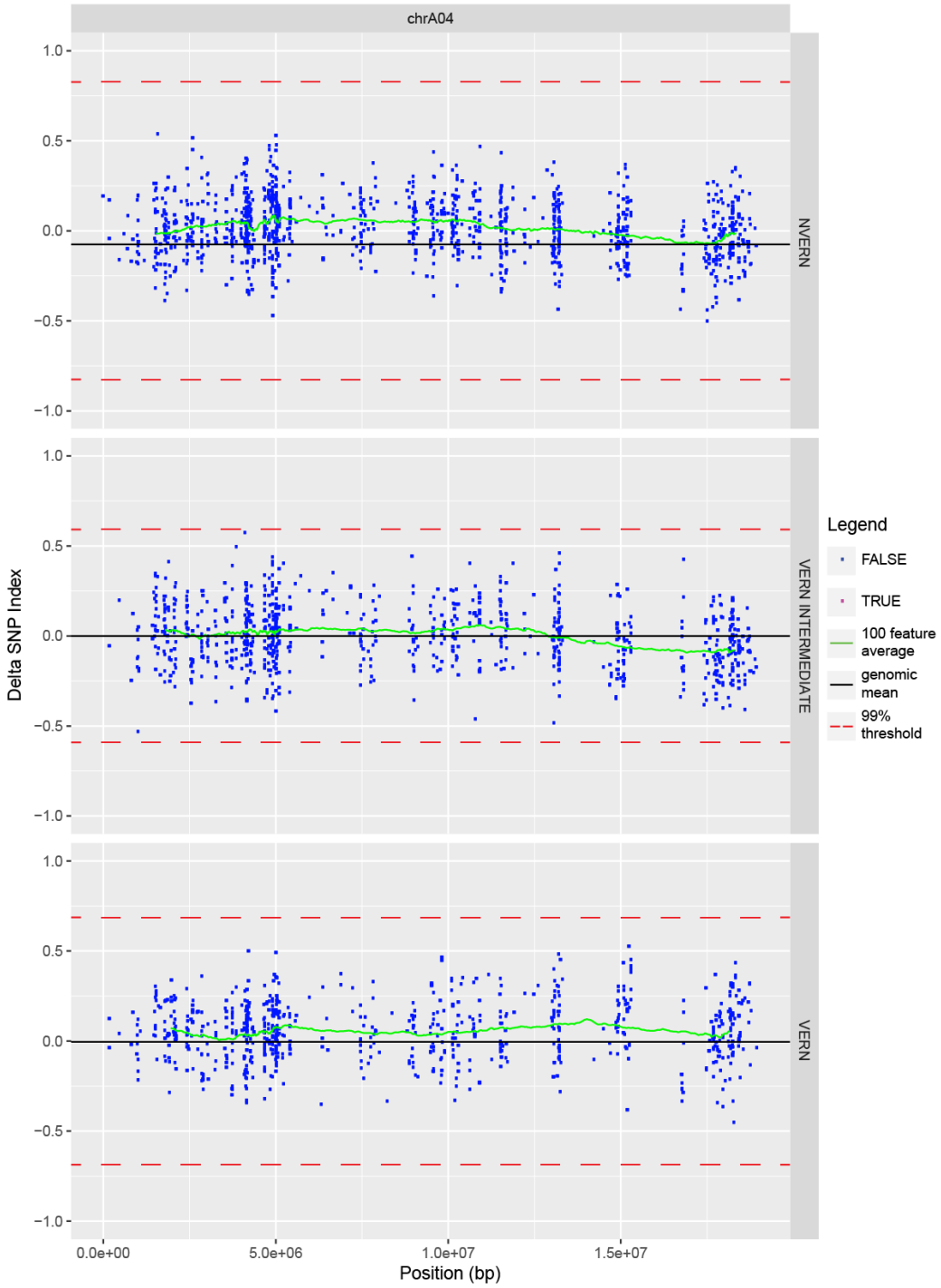


Figure S. 7: Chromosome A04. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in pink, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

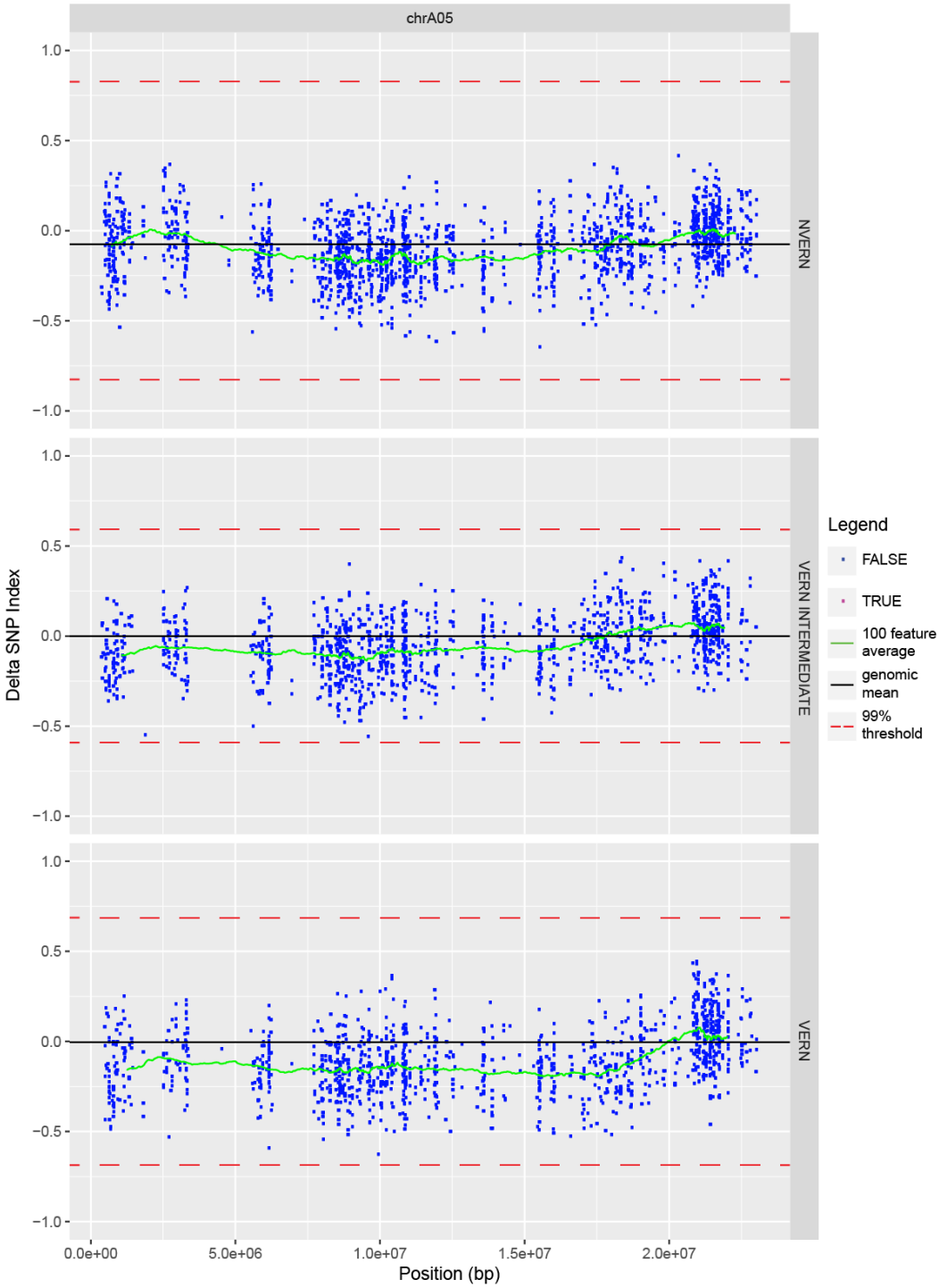


Figure S. 8: Chromosome A05. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in pink, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

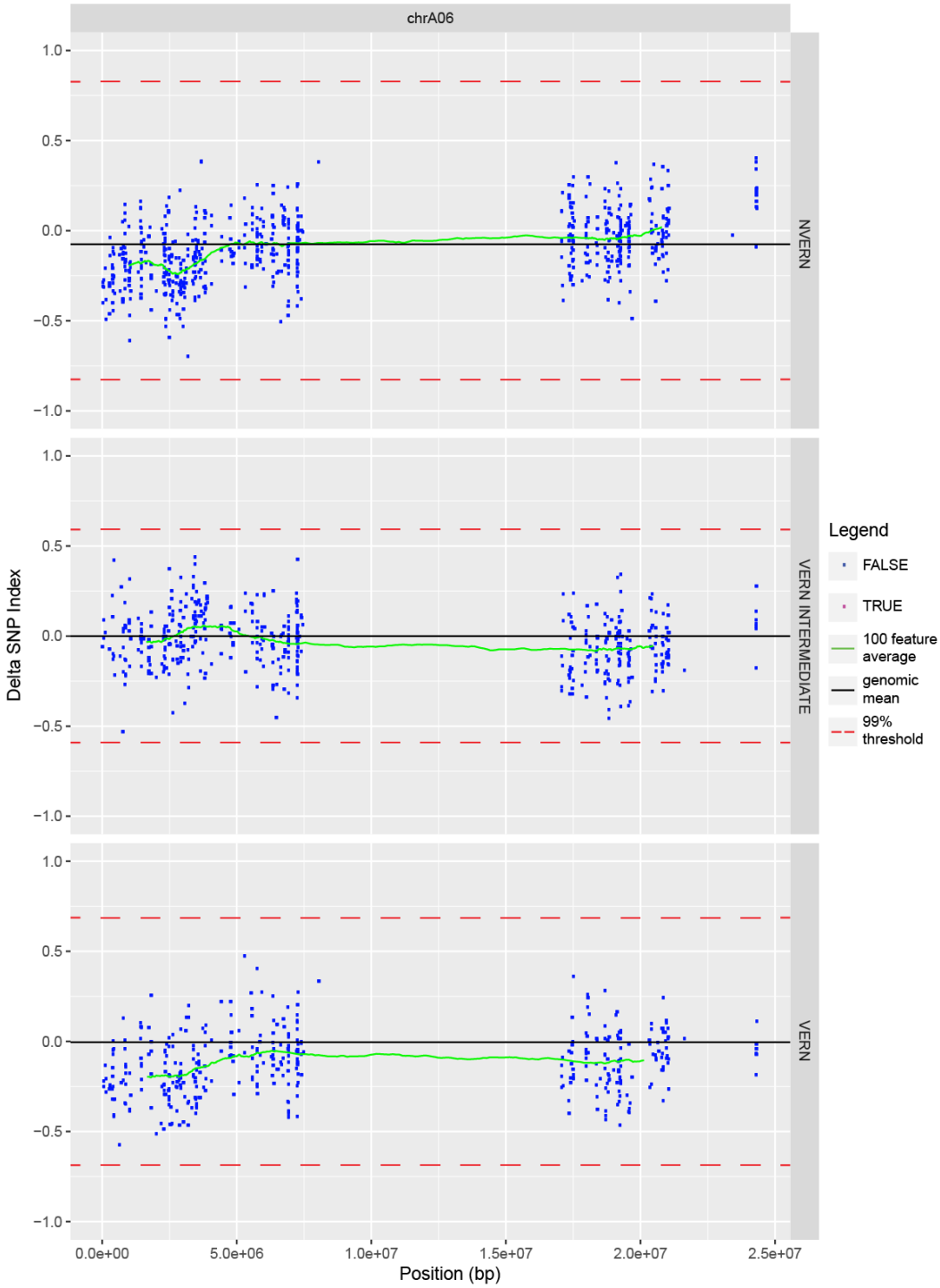


Figure S. 9: Chromosome A06. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in pink, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

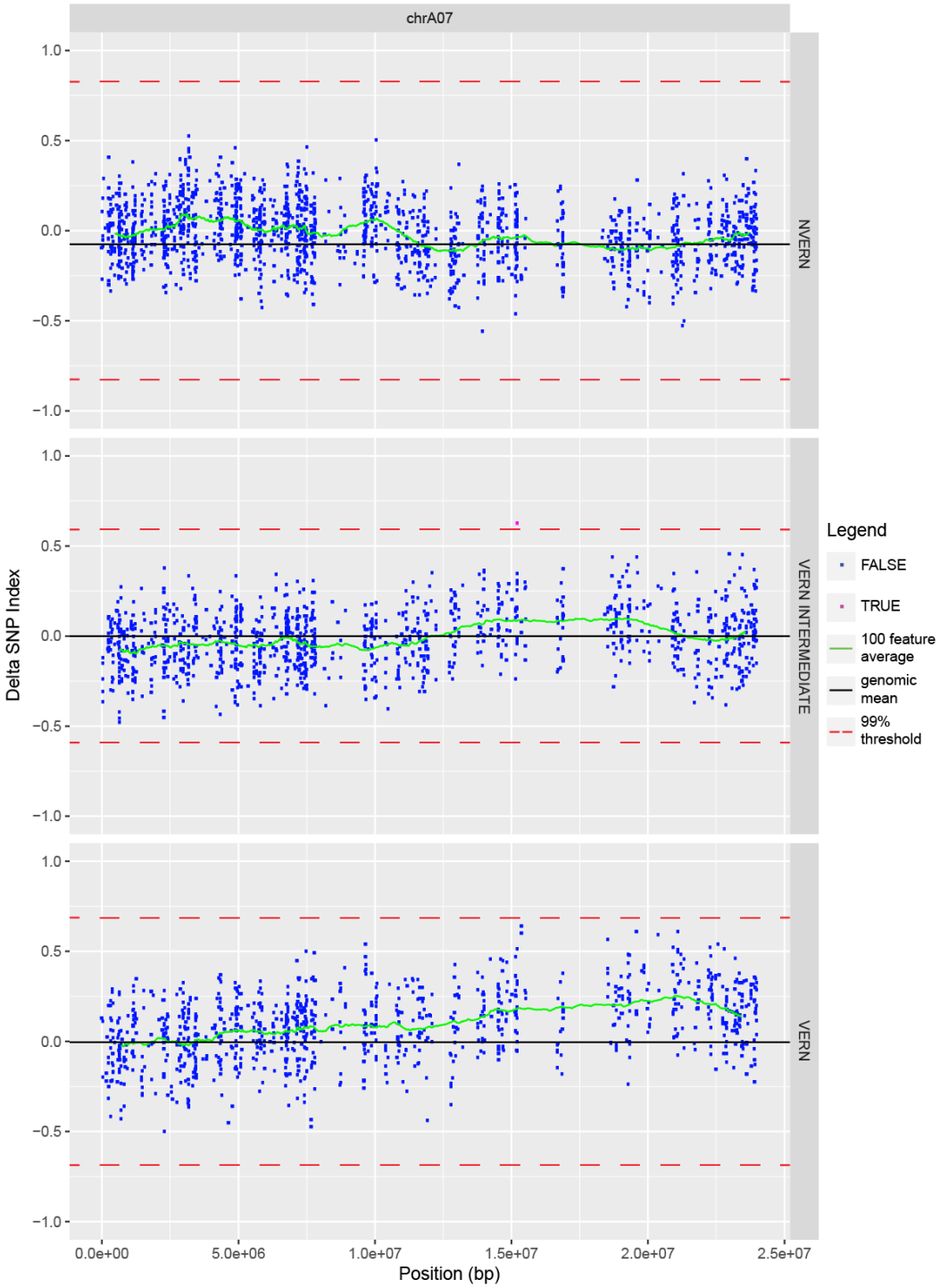


Figure S. 10*:* Chromosome A07. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in pink, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

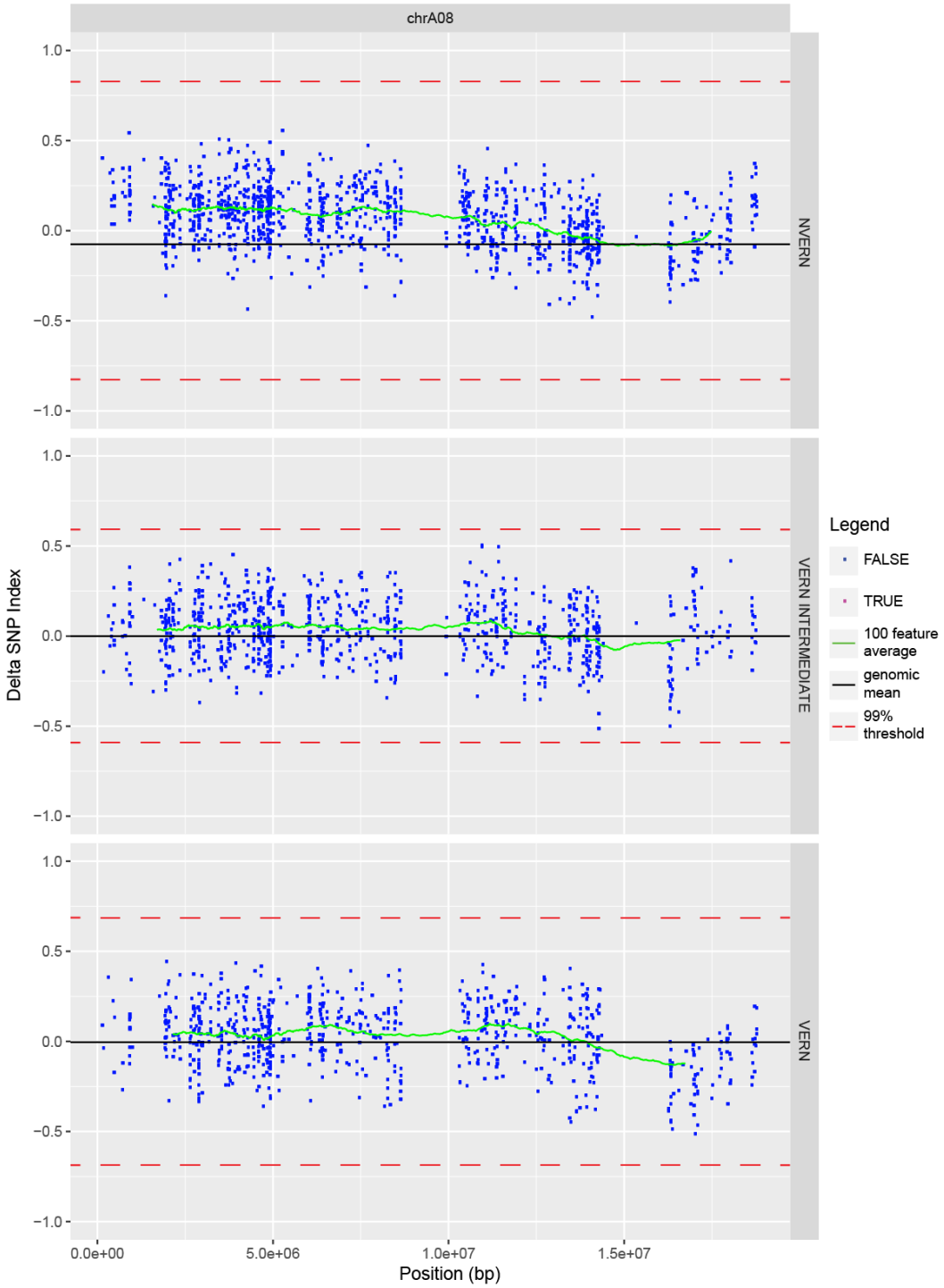


Figure S. 11: Chromosome A08. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in pink, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

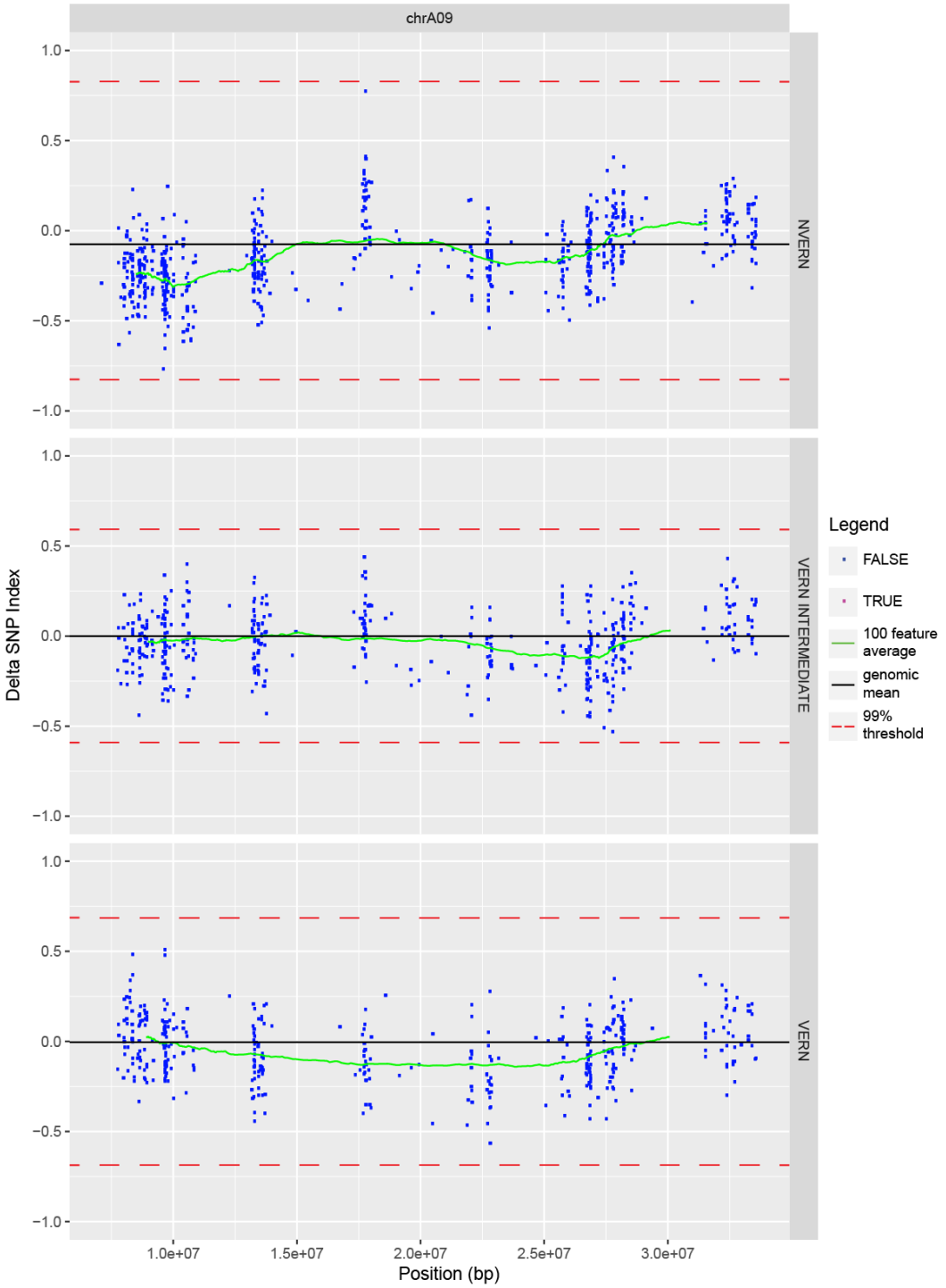


Figure S. 12: Chromosome A09. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in pink, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

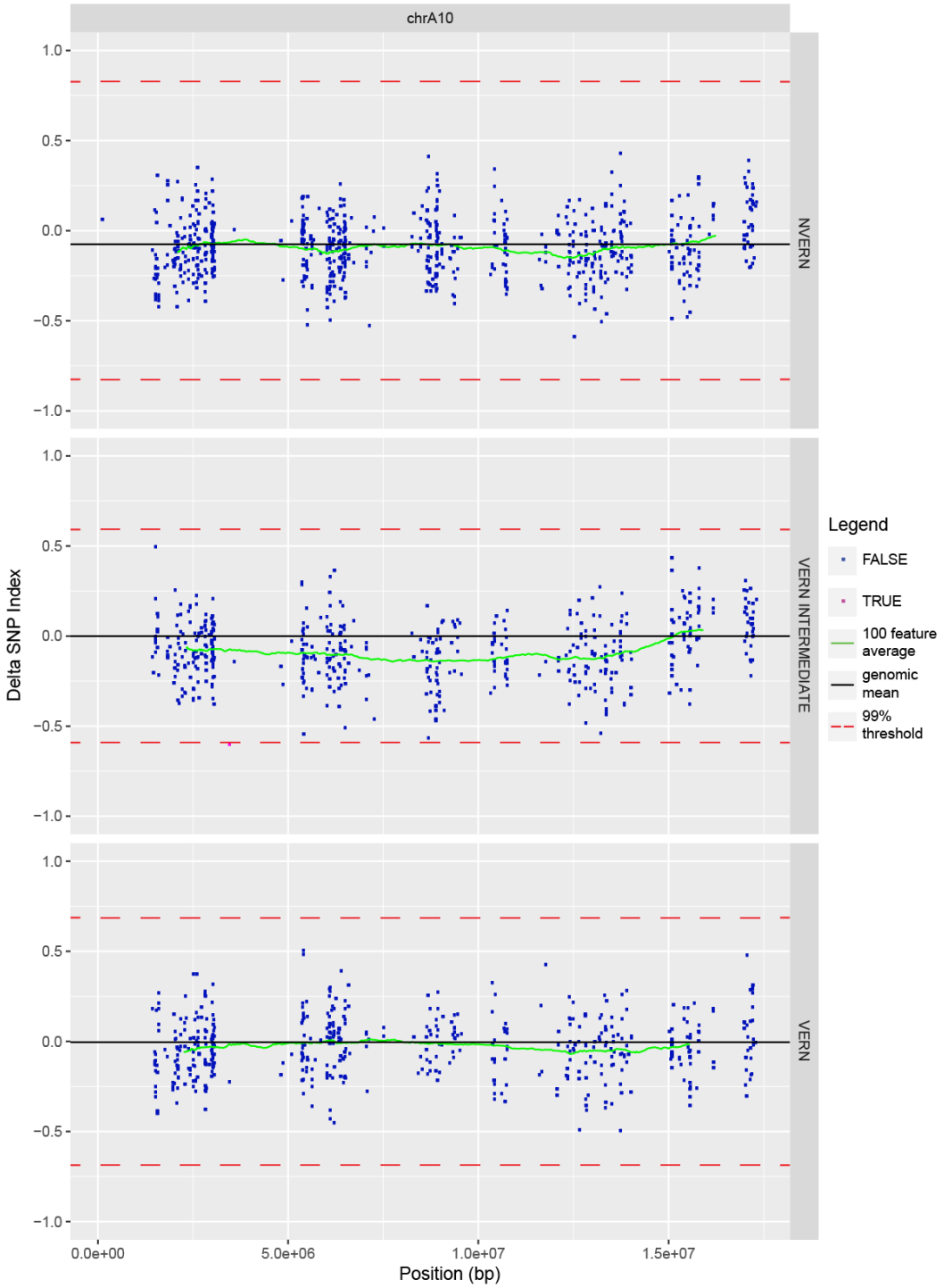


Figure S. 13: Chromosome A10. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in pink, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

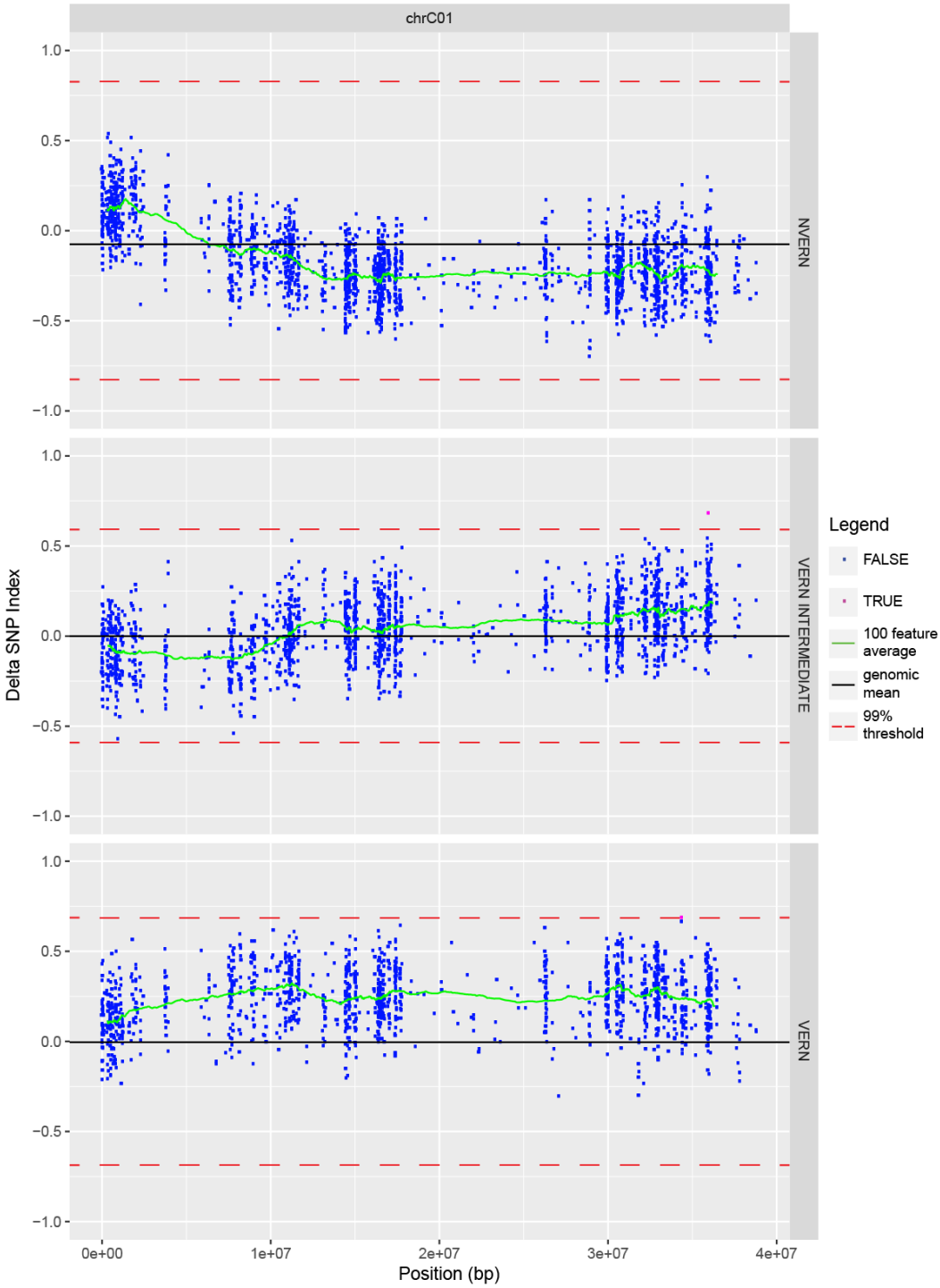


Figure S. 14: Chromosome C01. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in pink, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

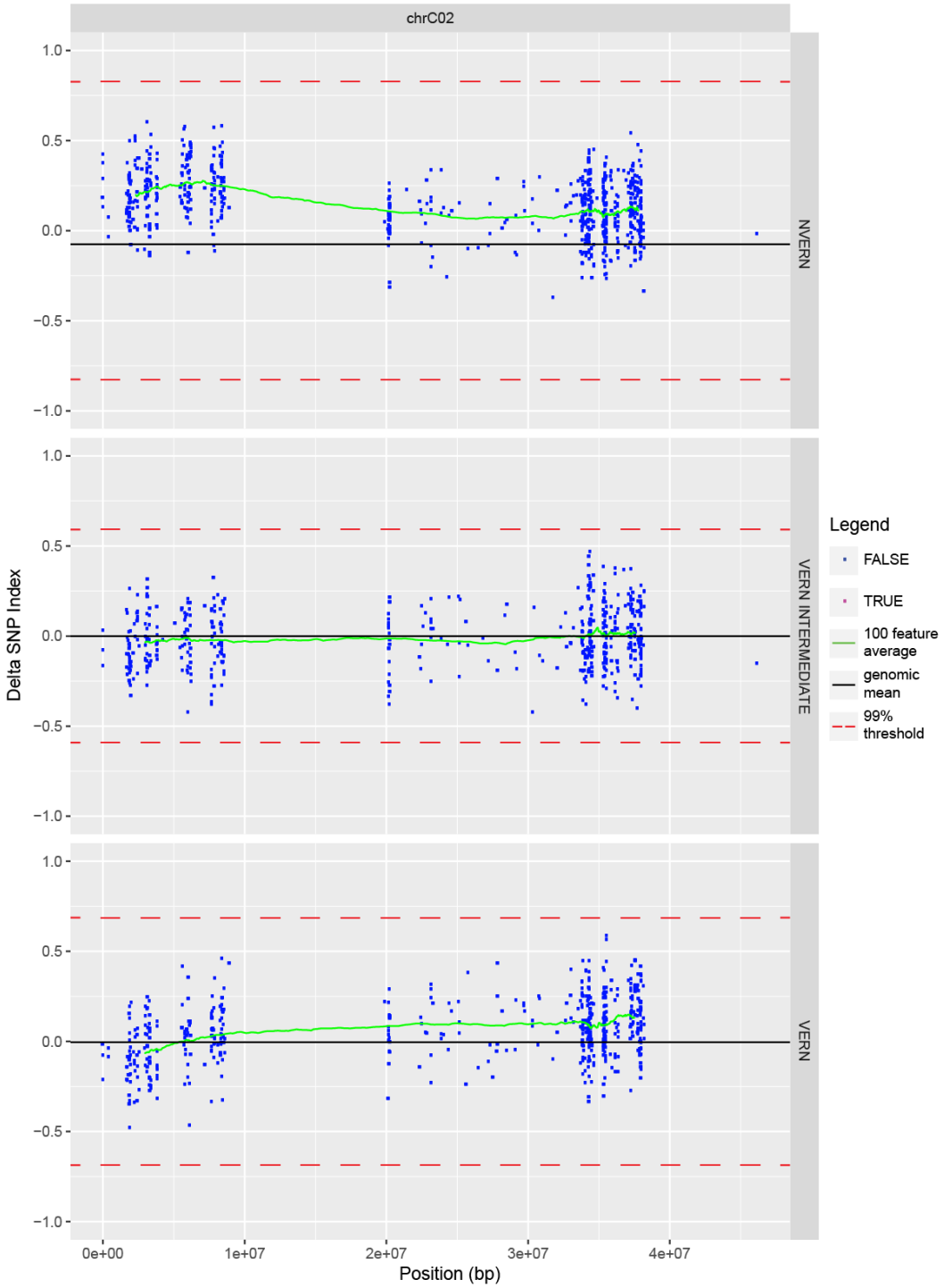


Figure S. 15: Chromosome C02. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in pink, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

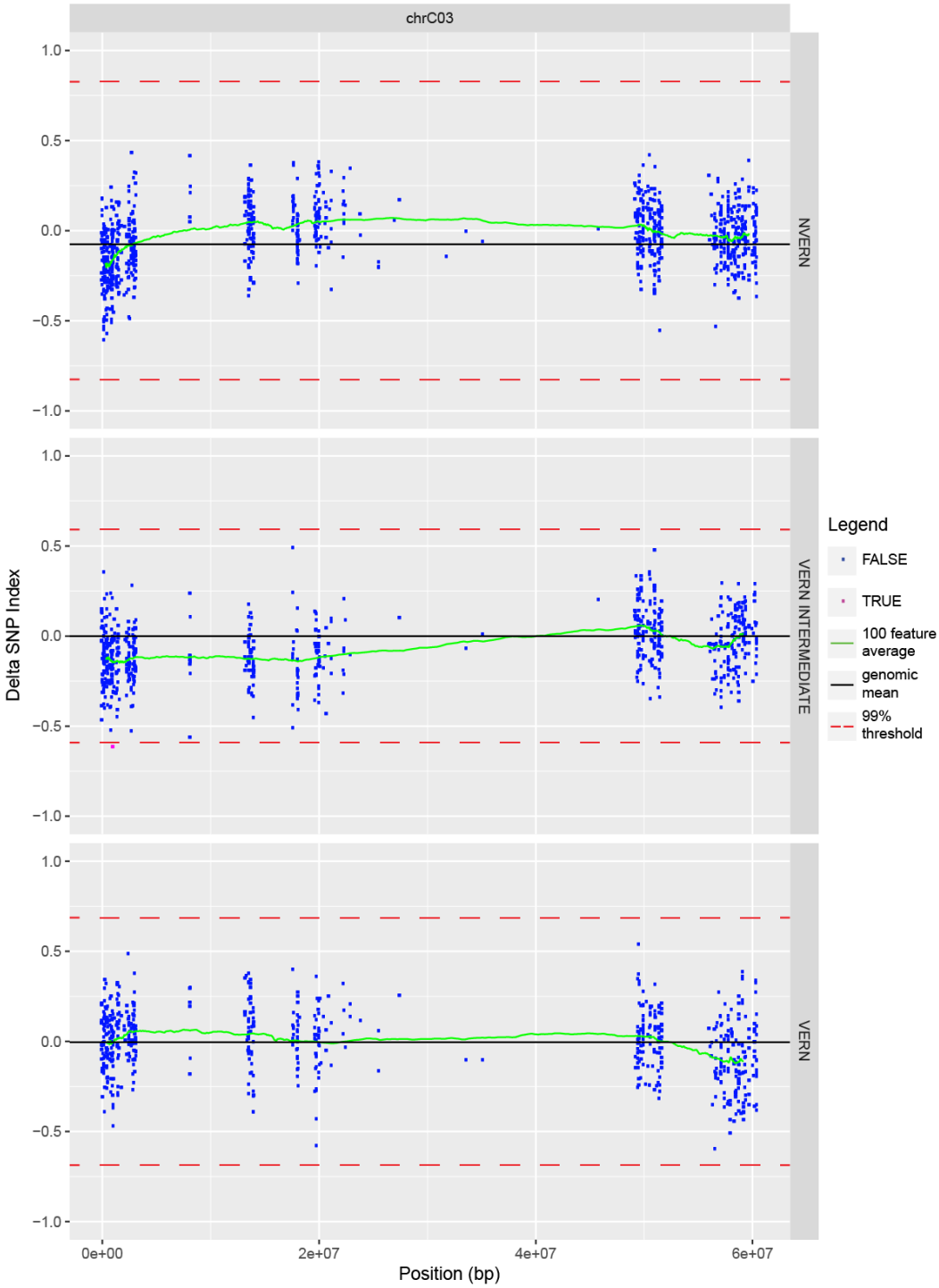


Figure S. 16: Chromosome C03. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in pink, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

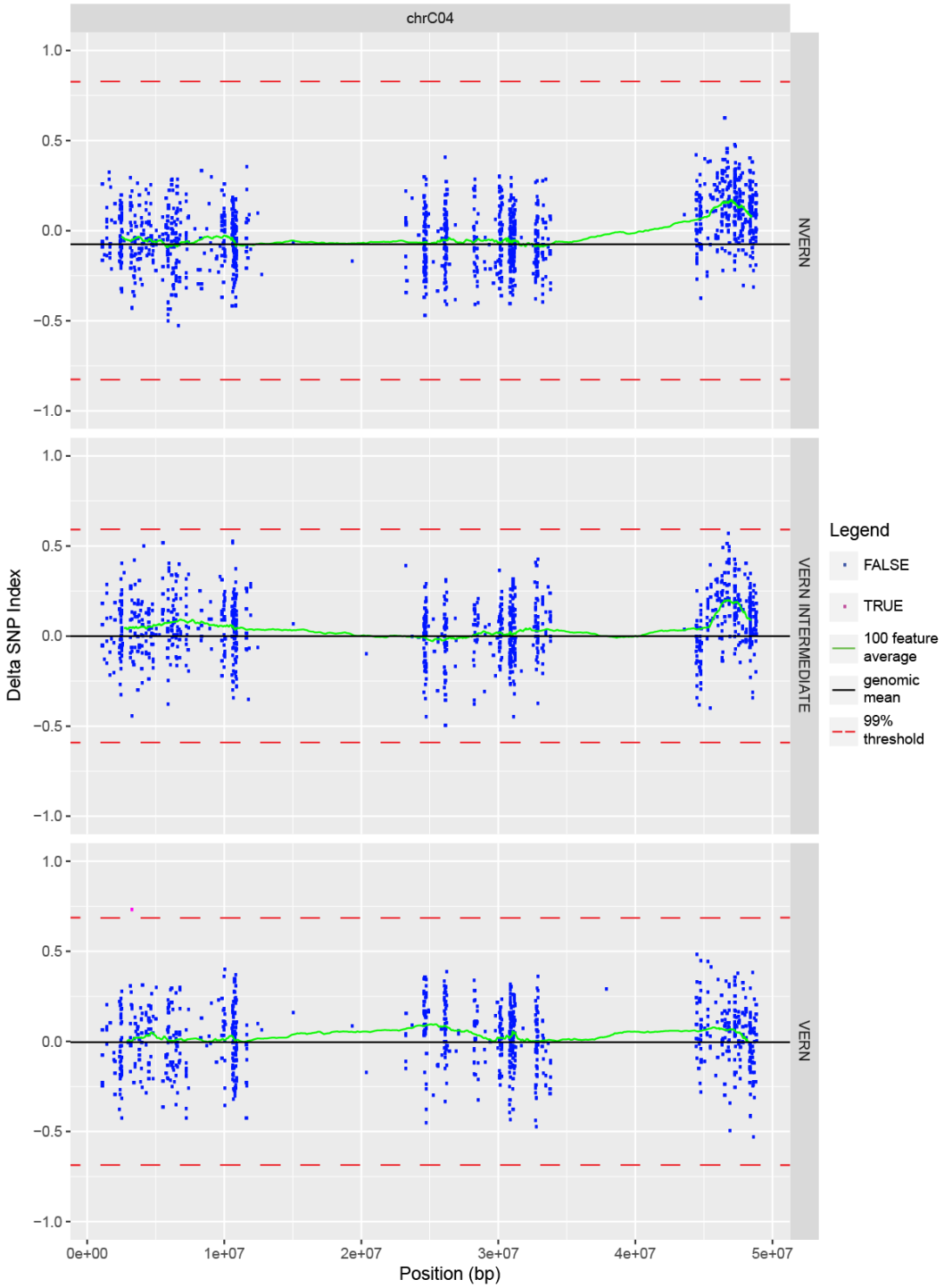


Figure S. 17: Chromosome C04. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in pink, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

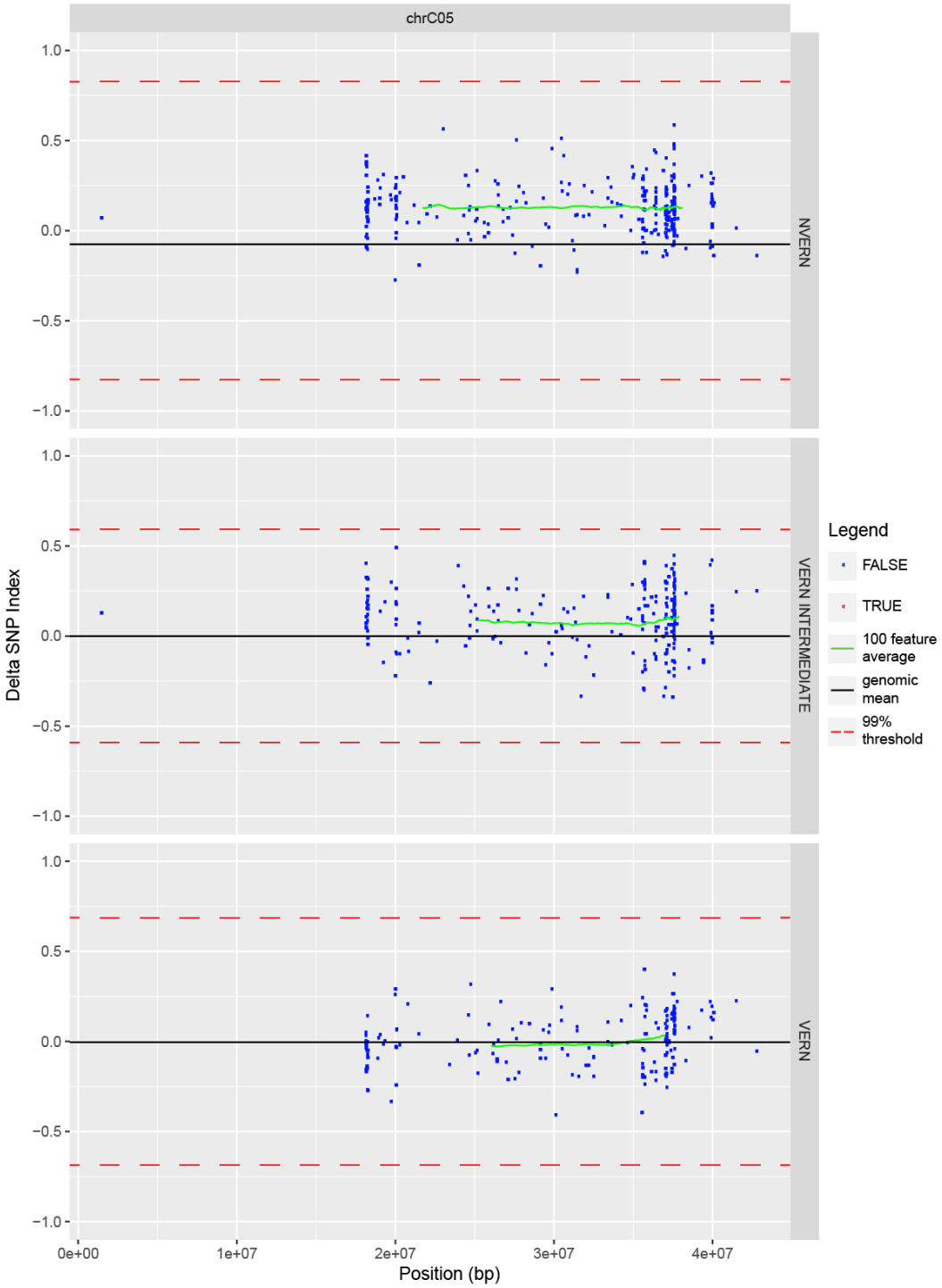


Figure S. 18: Chromosome C05. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in pink, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

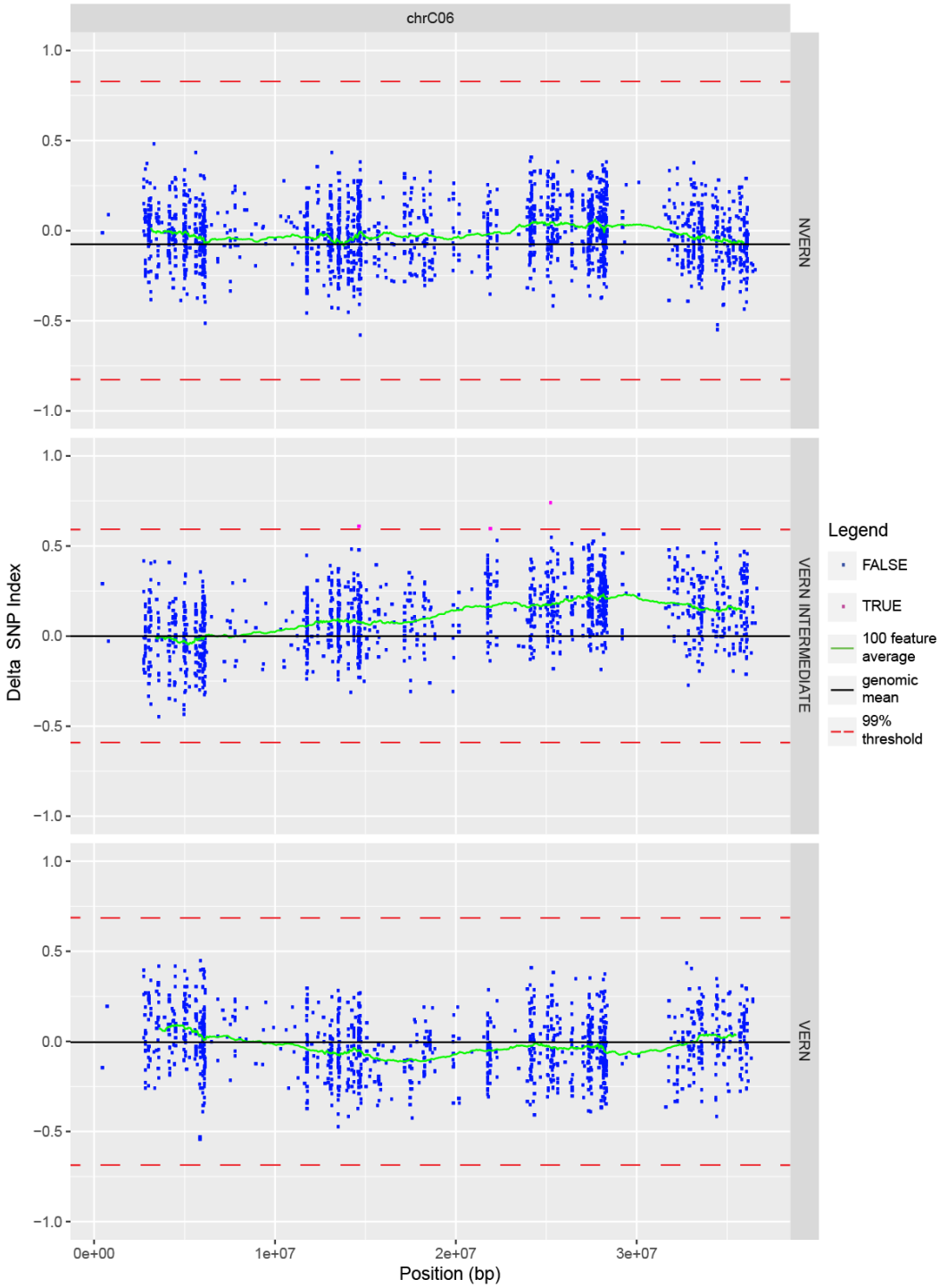


Figure S. 19: Chromosome C06. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in pink, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

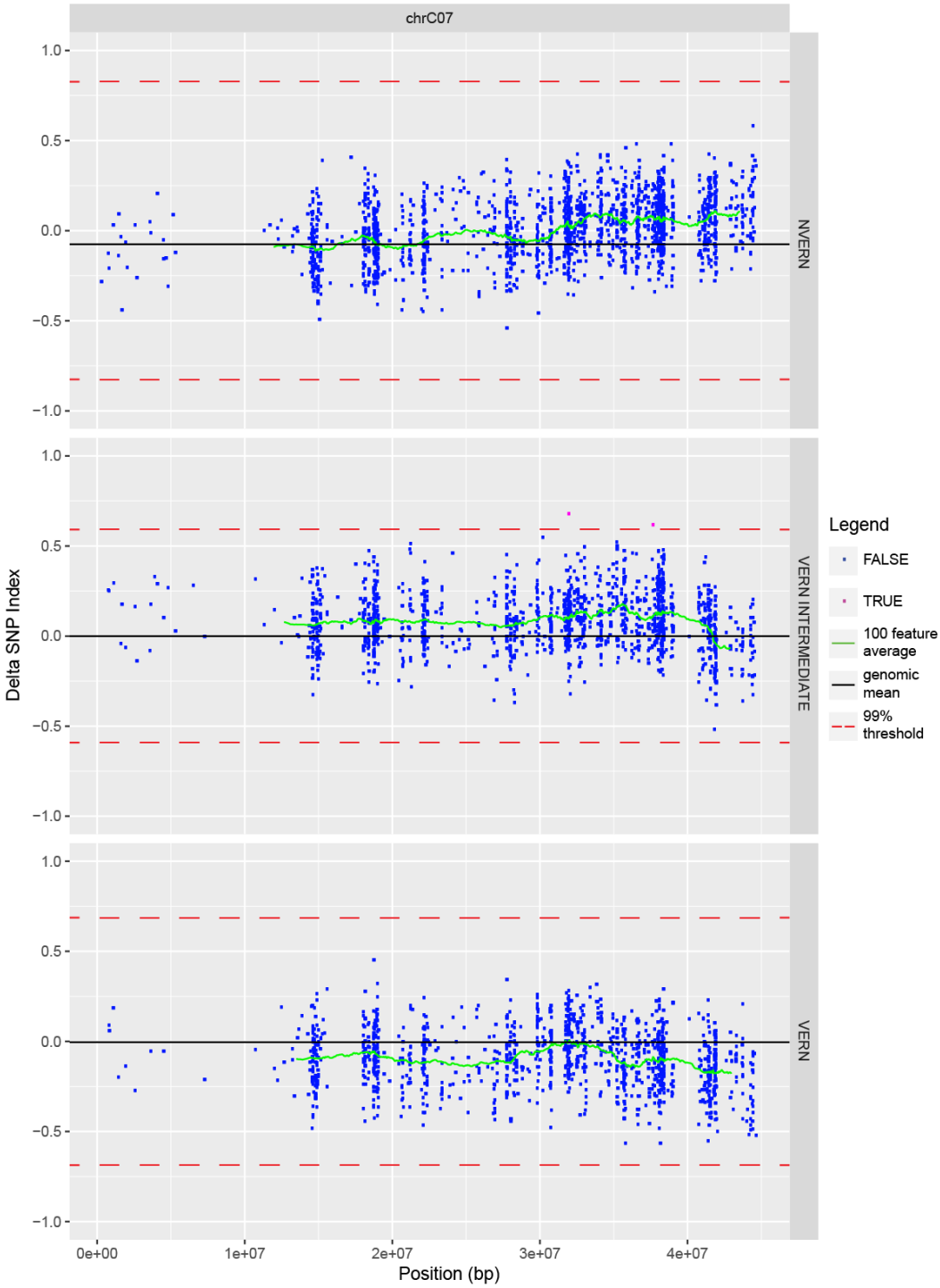


Figure S. 20: Chromosome C07. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in pink, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

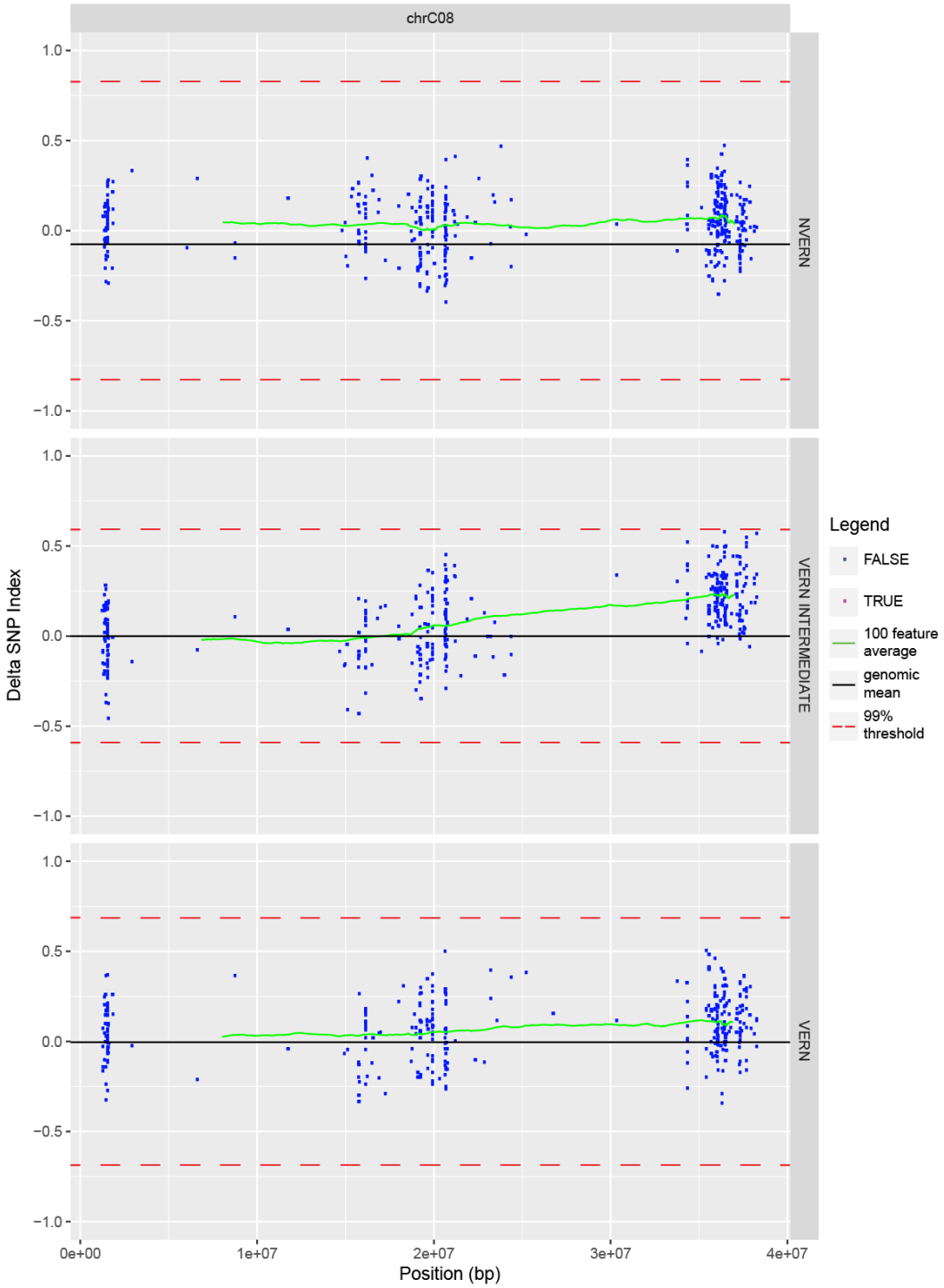


Figure S. 21: Chromosome C08. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in pink, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

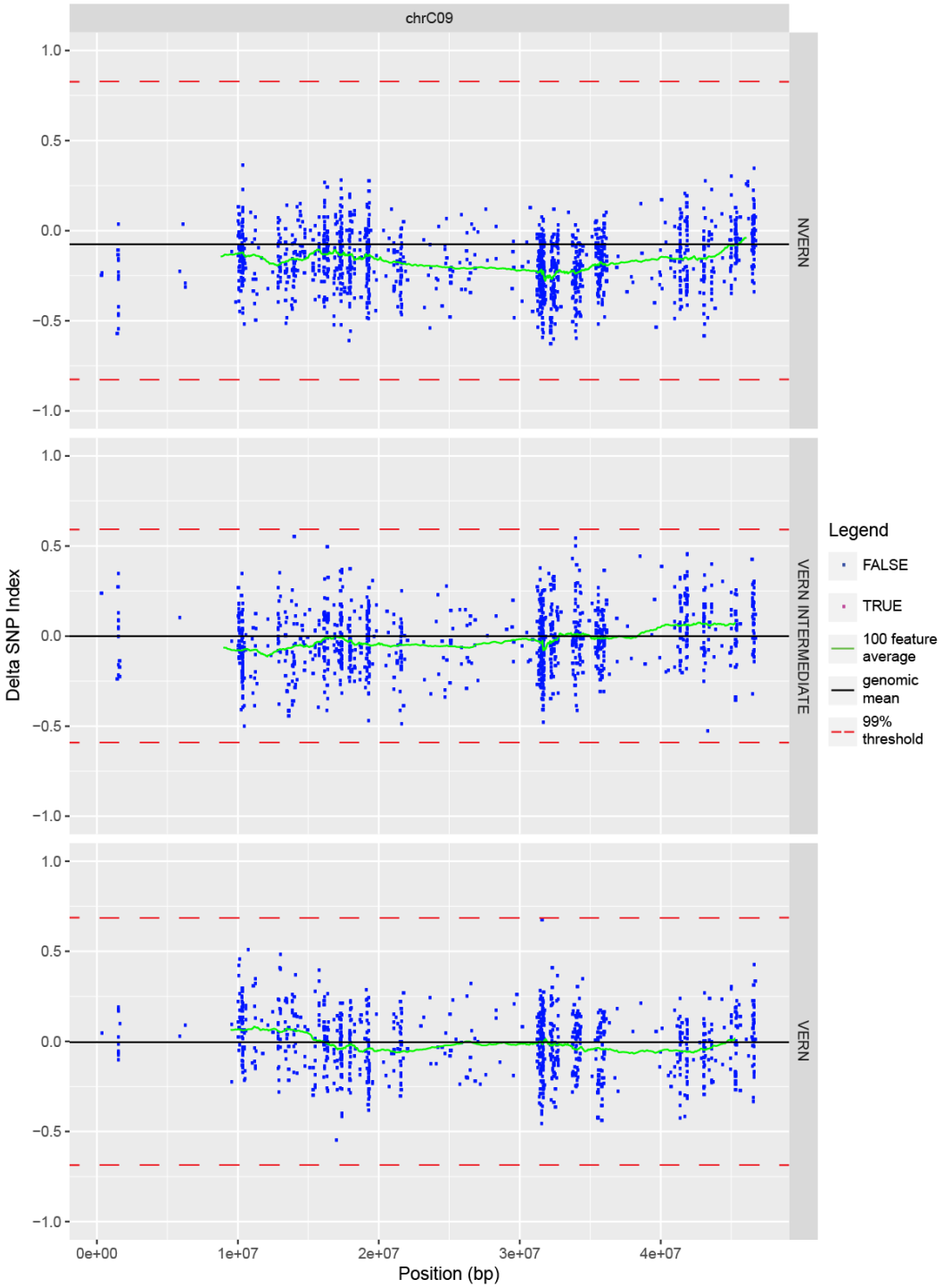


Figure S. 22: Chromosome C09. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in pink, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

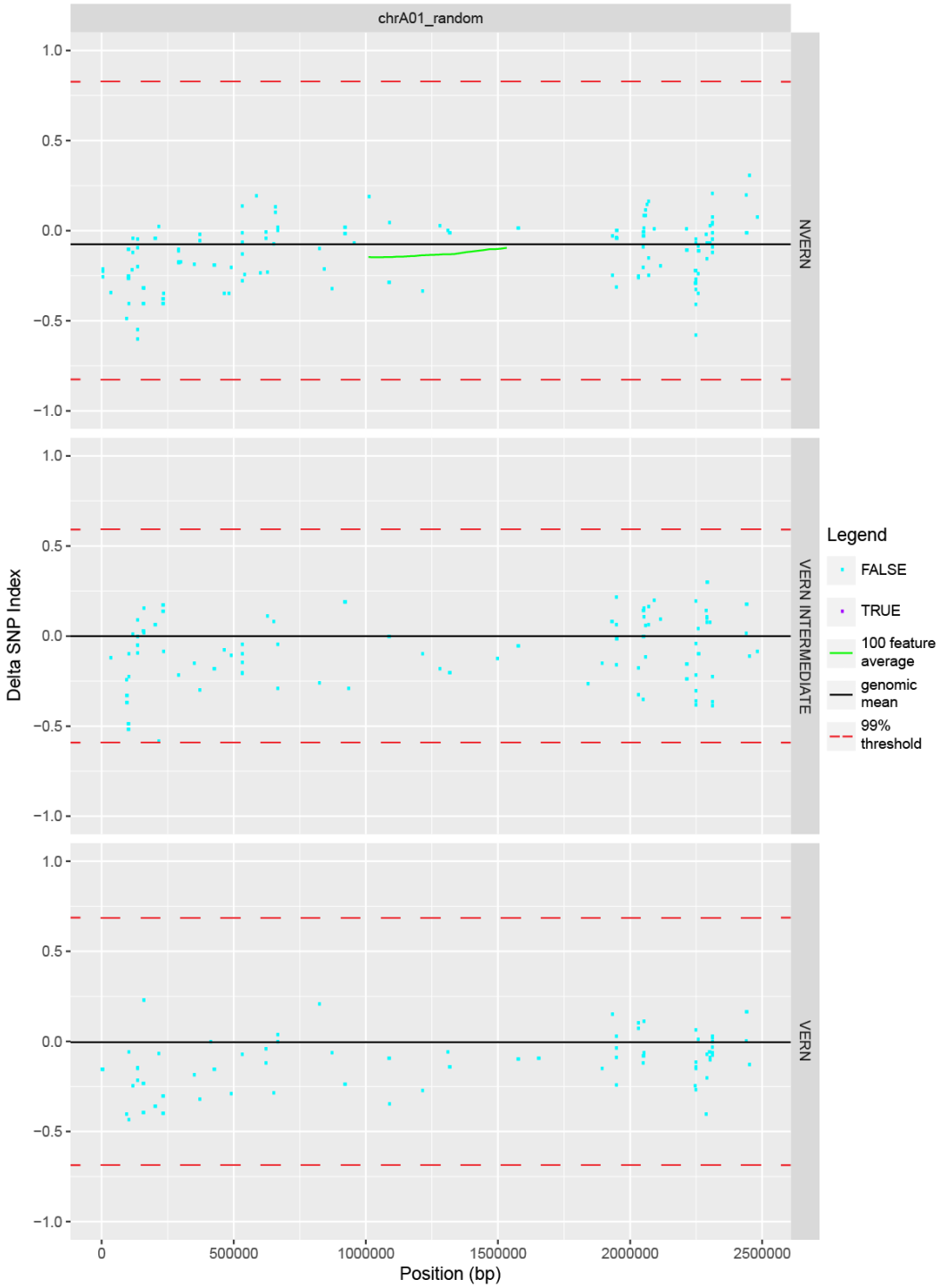


Figure S. 23 Chromosome A01_random. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in purple, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

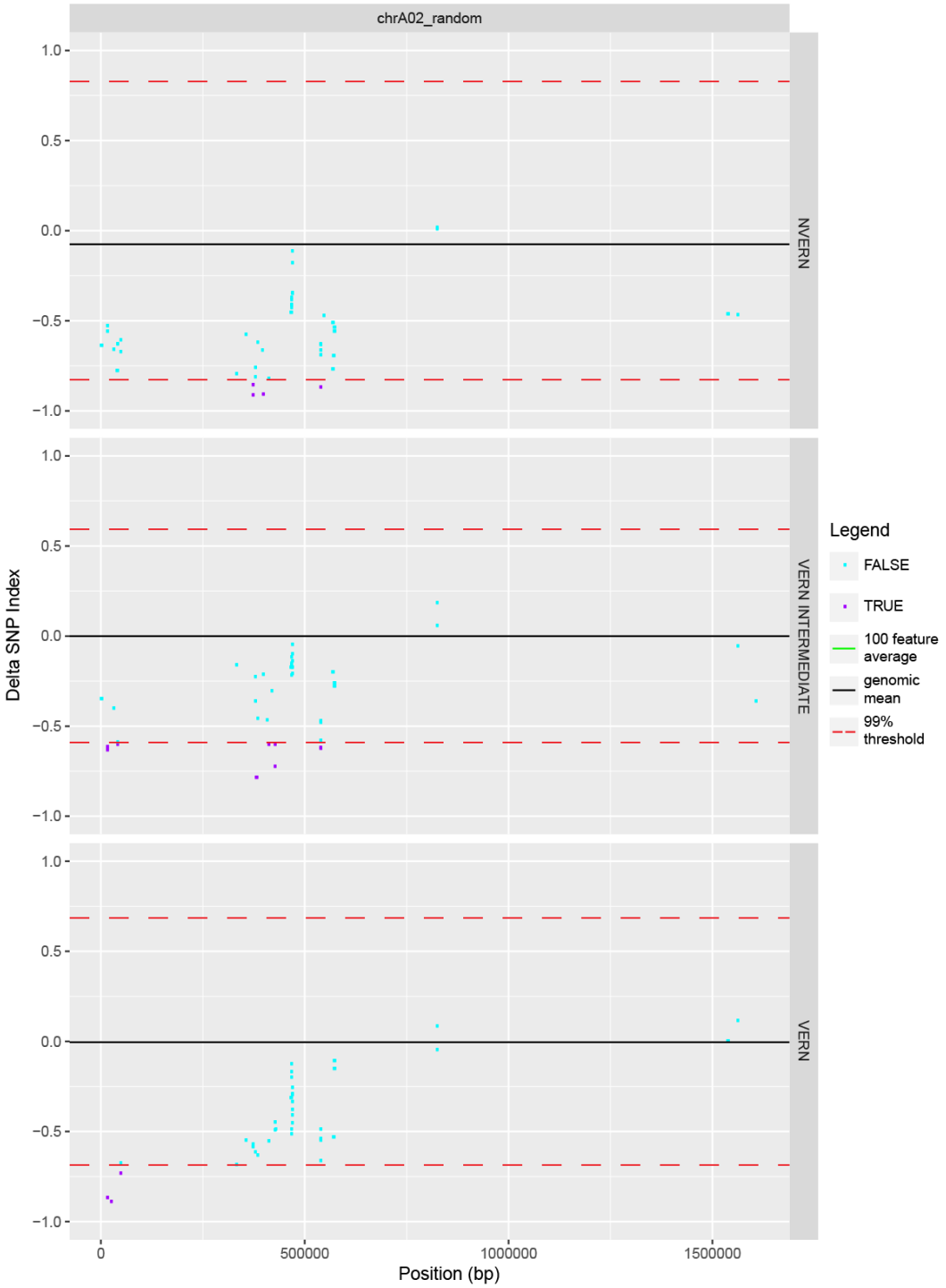


Figure S. 24 Chromosome A02_random. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in purple, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

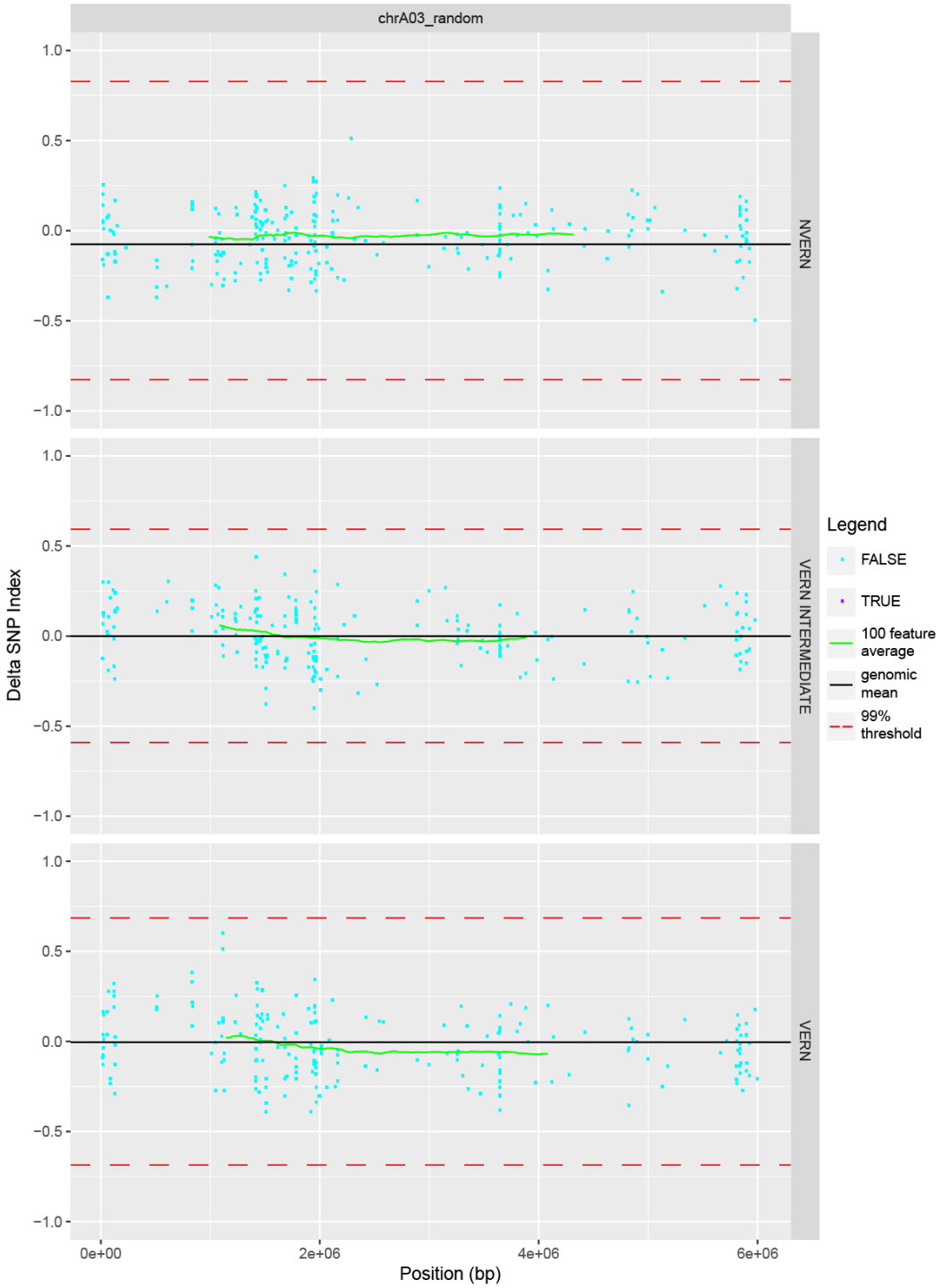


Figure S. 25: Chromosome A03_random. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in purple, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

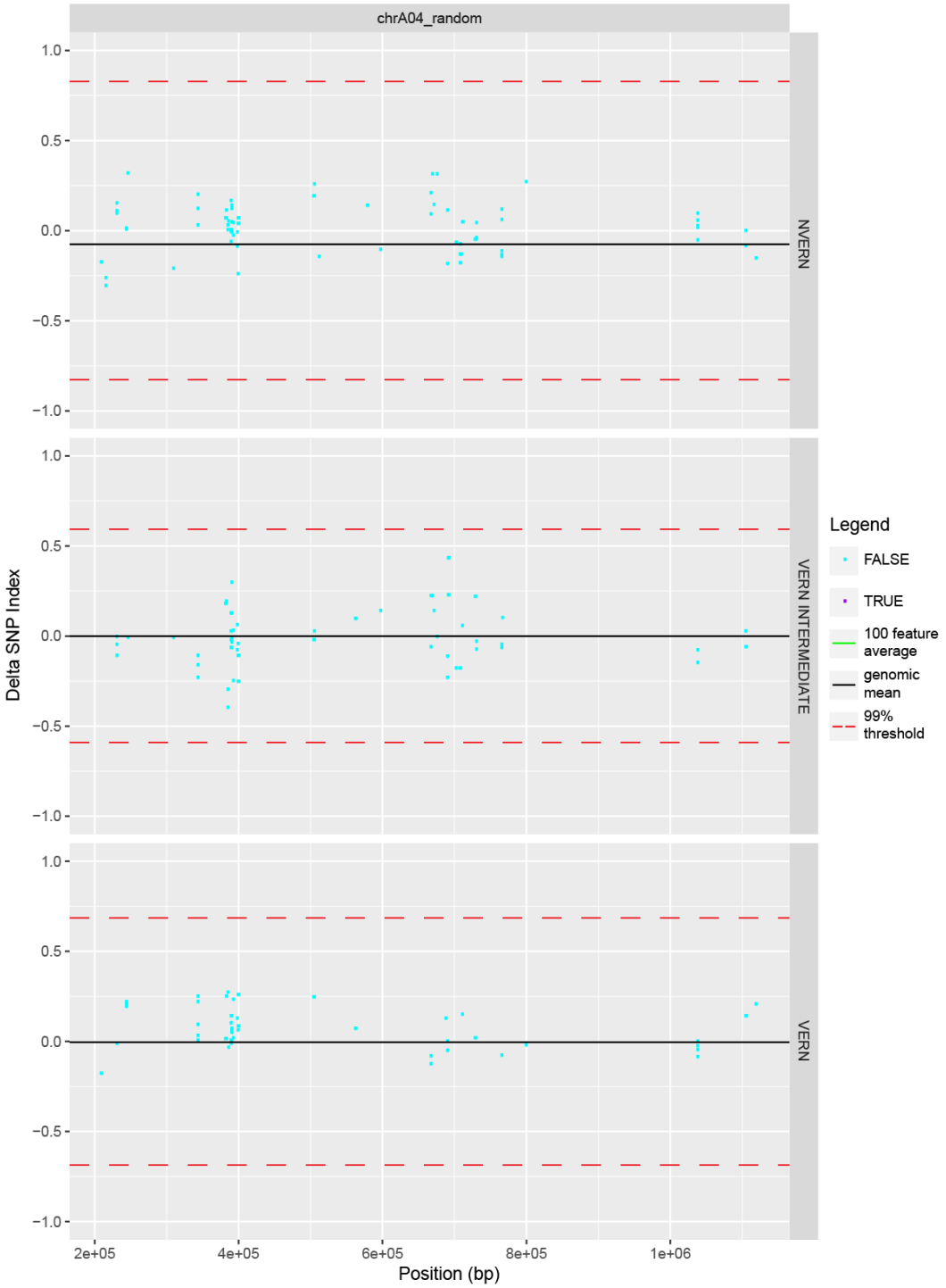


Figure S. 26: Chromosome A04_random. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in purple, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

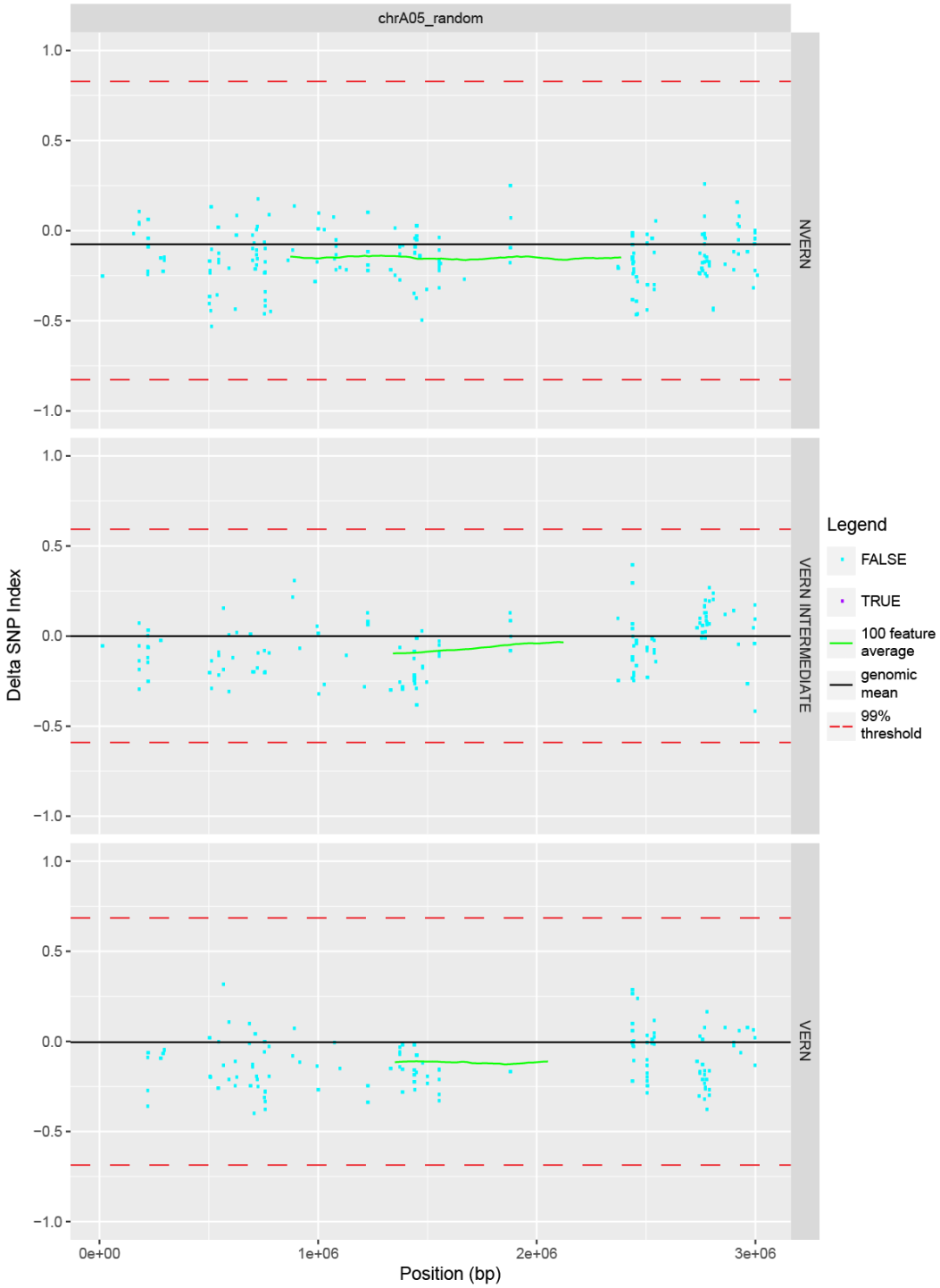


Figure S. 27*:* Chromosome A05_random. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in purple, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

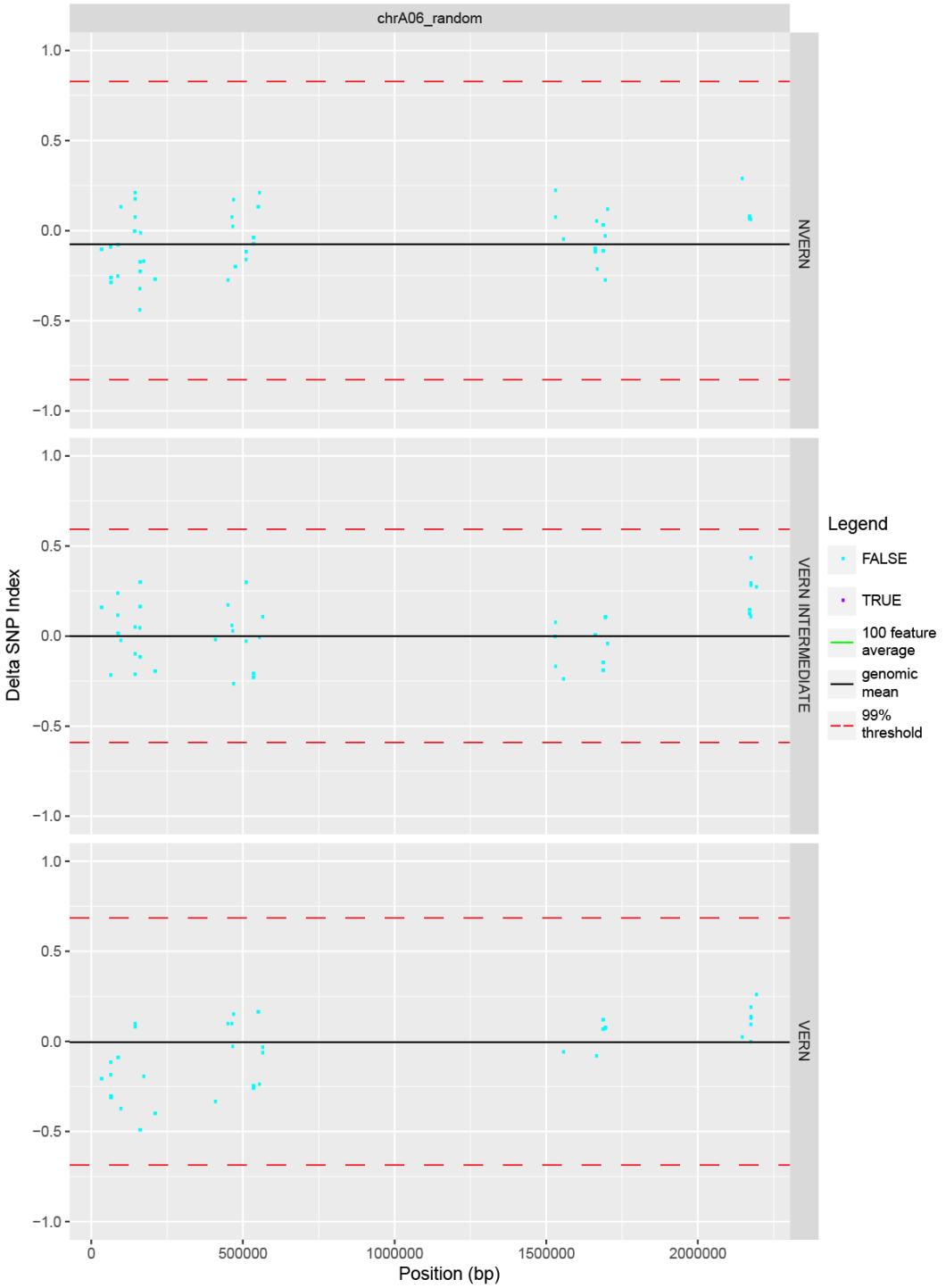


Figure S. 28: Chromosome A06_random. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in purple, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

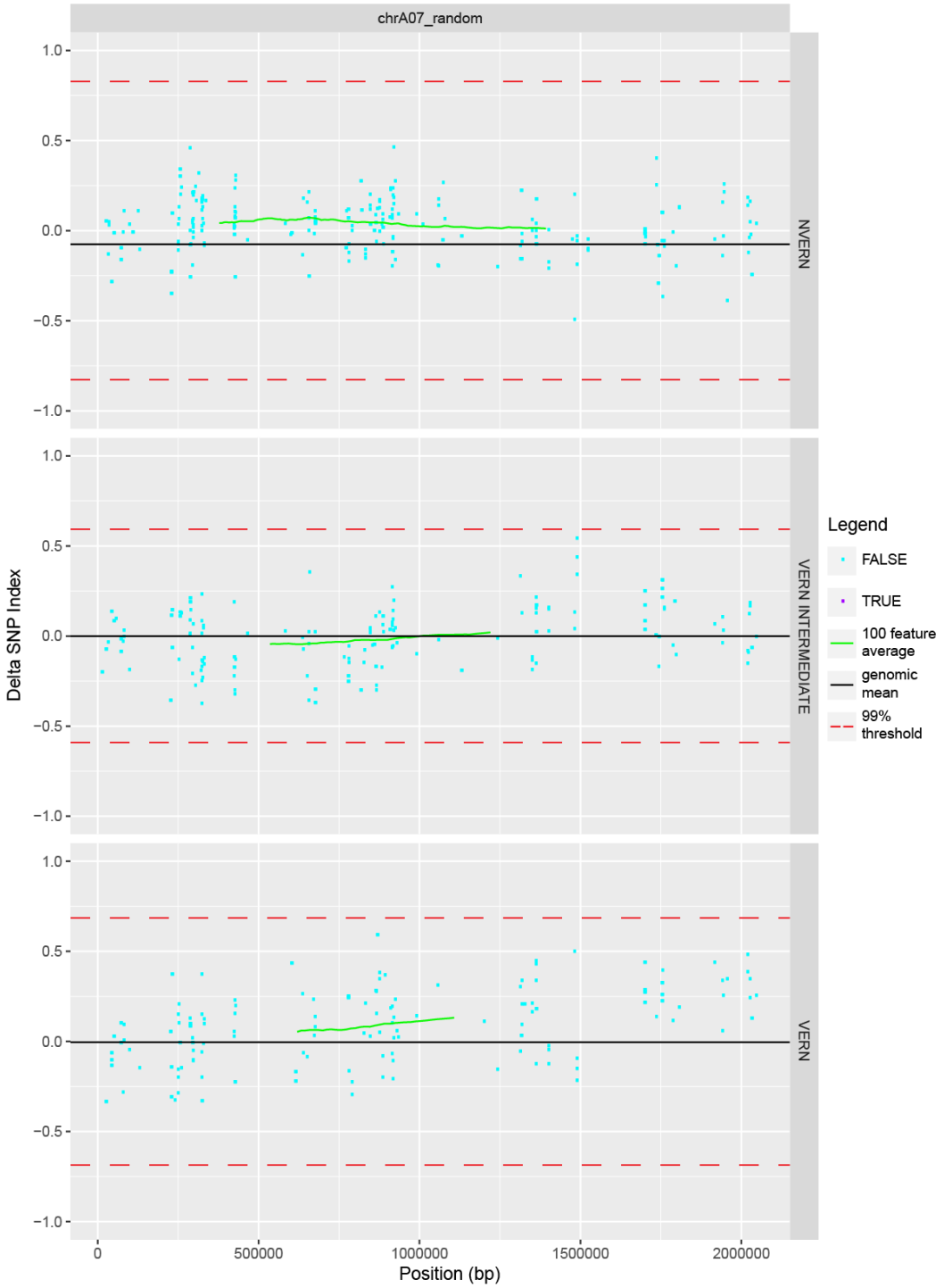


Figure S. 29: Chromosome A07_random. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in purple, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

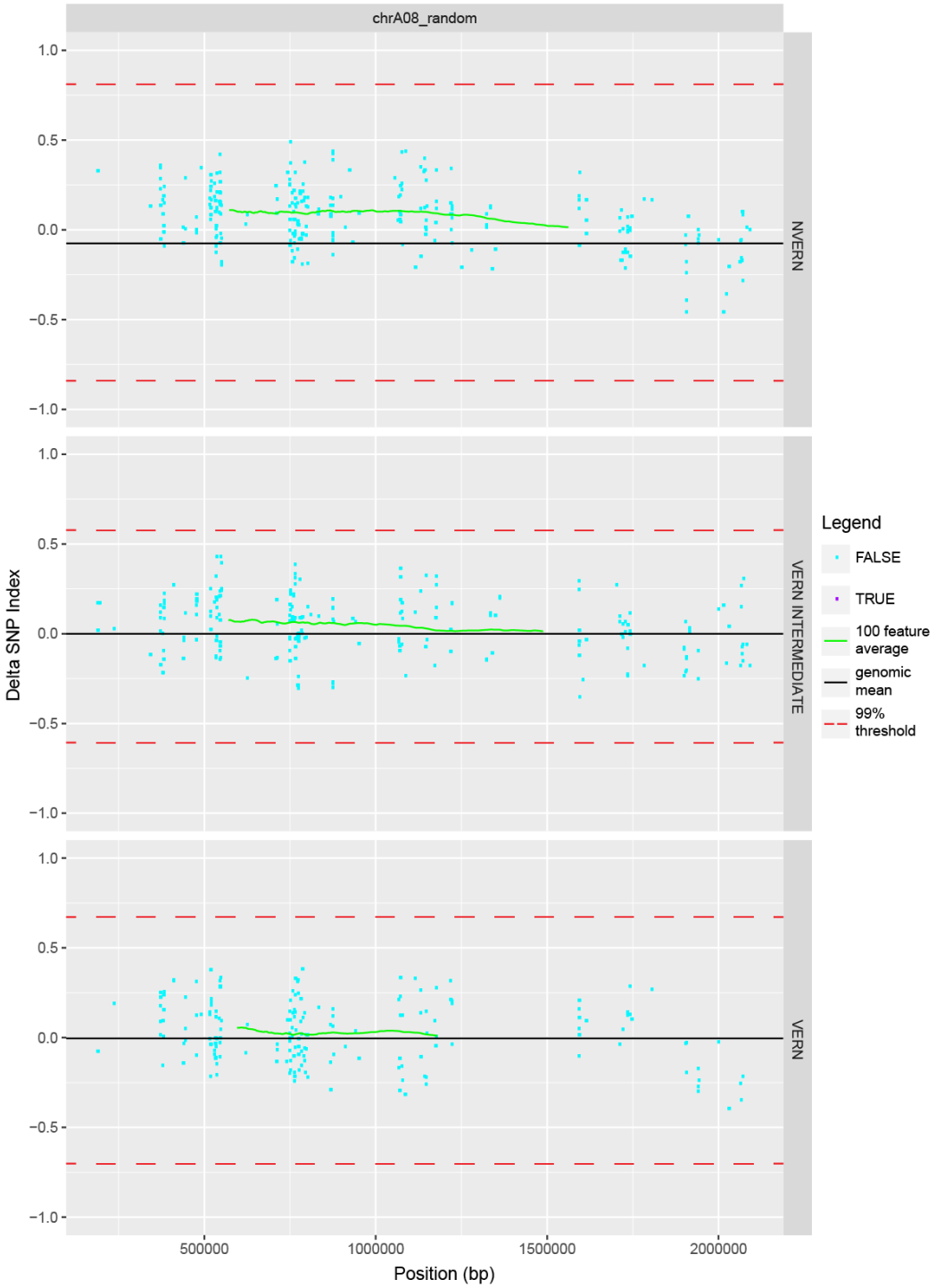


Figure S. 30: Chromosome A08_random. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in purple, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

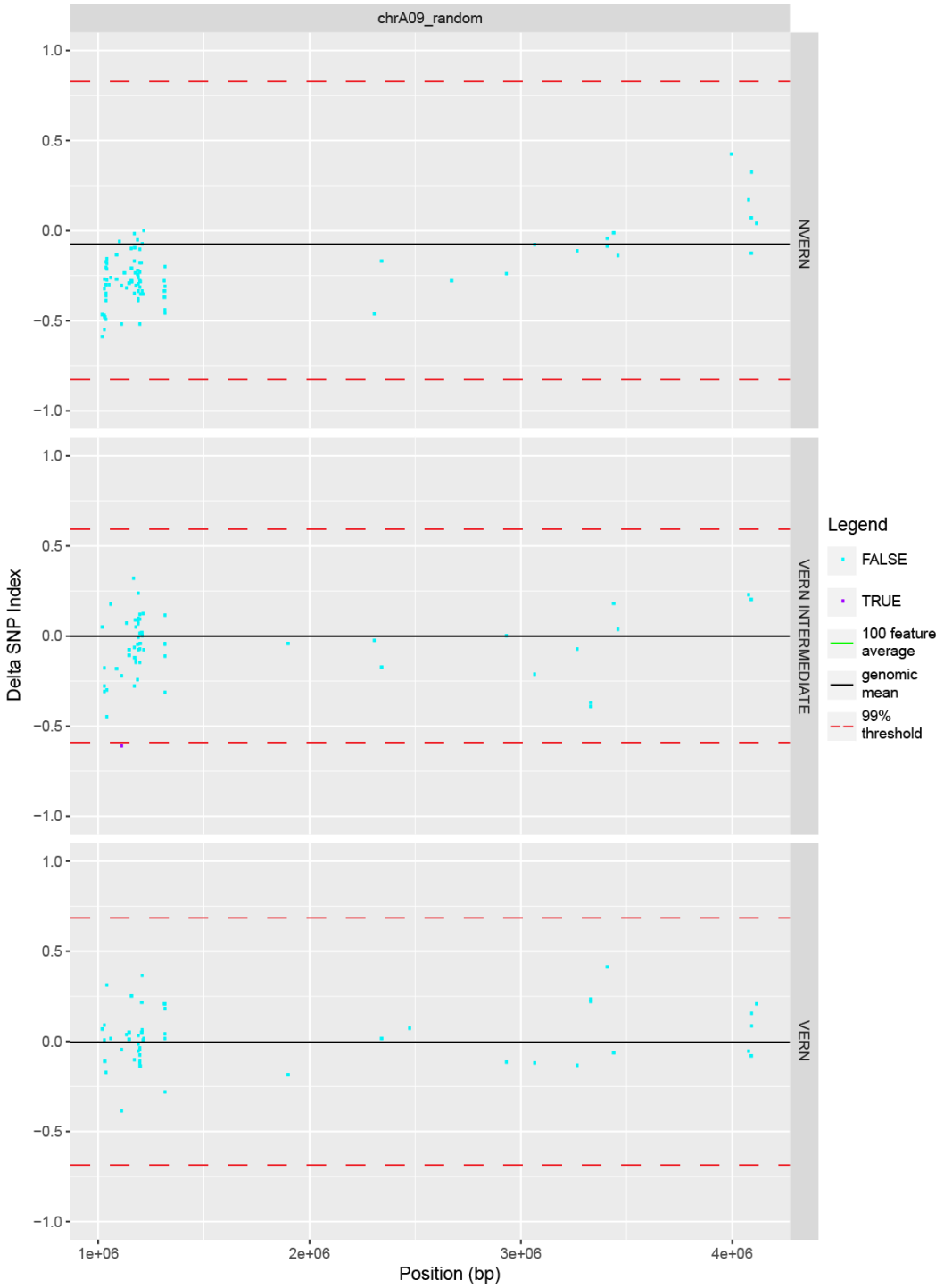


Figure S. 31: Chromosome A09_random. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in purple, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

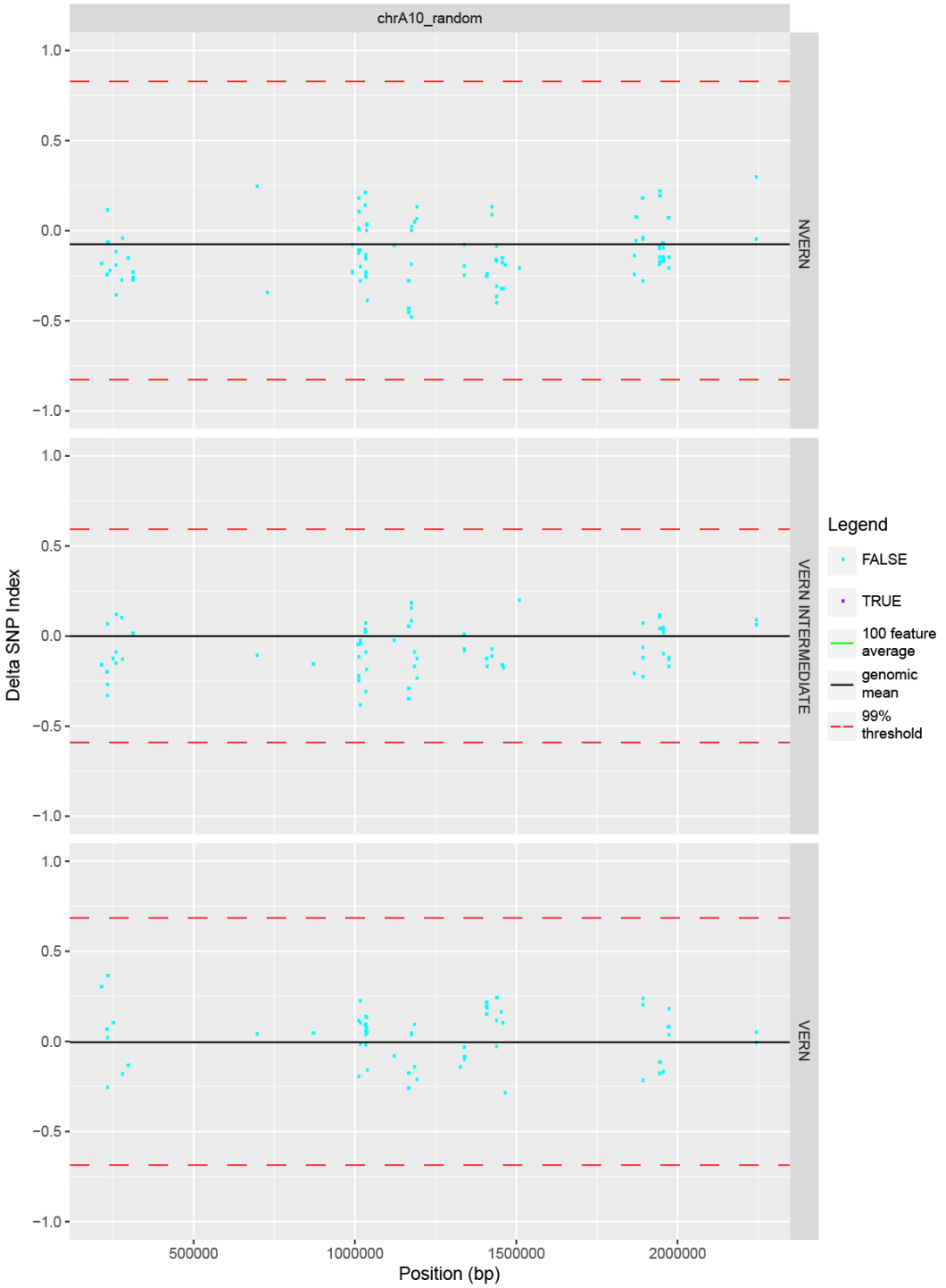


Figure S. 32: Chromosome A10_random. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in purple, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

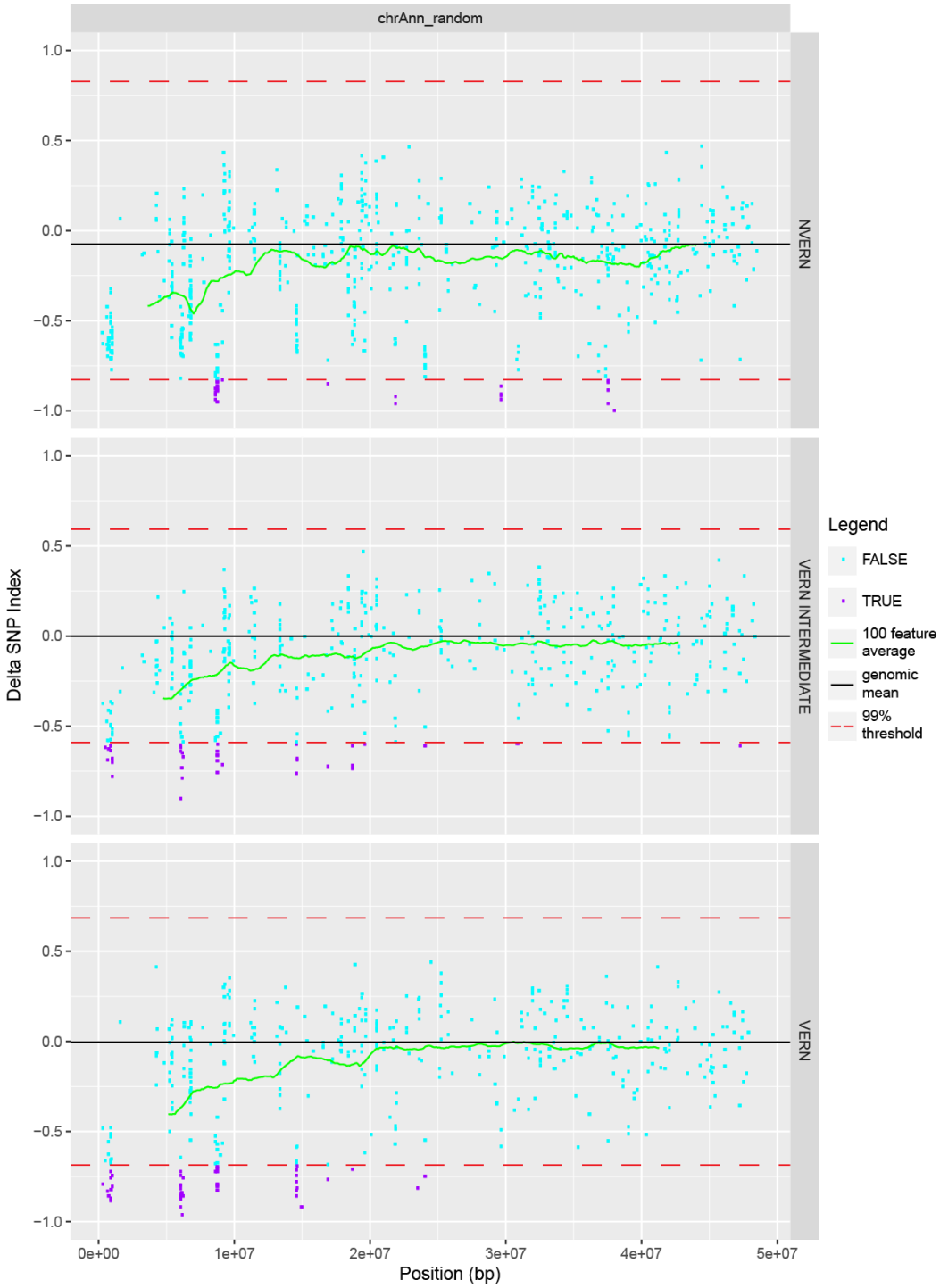


Figure S. 33: Chromosome Ann_random. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in purple, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

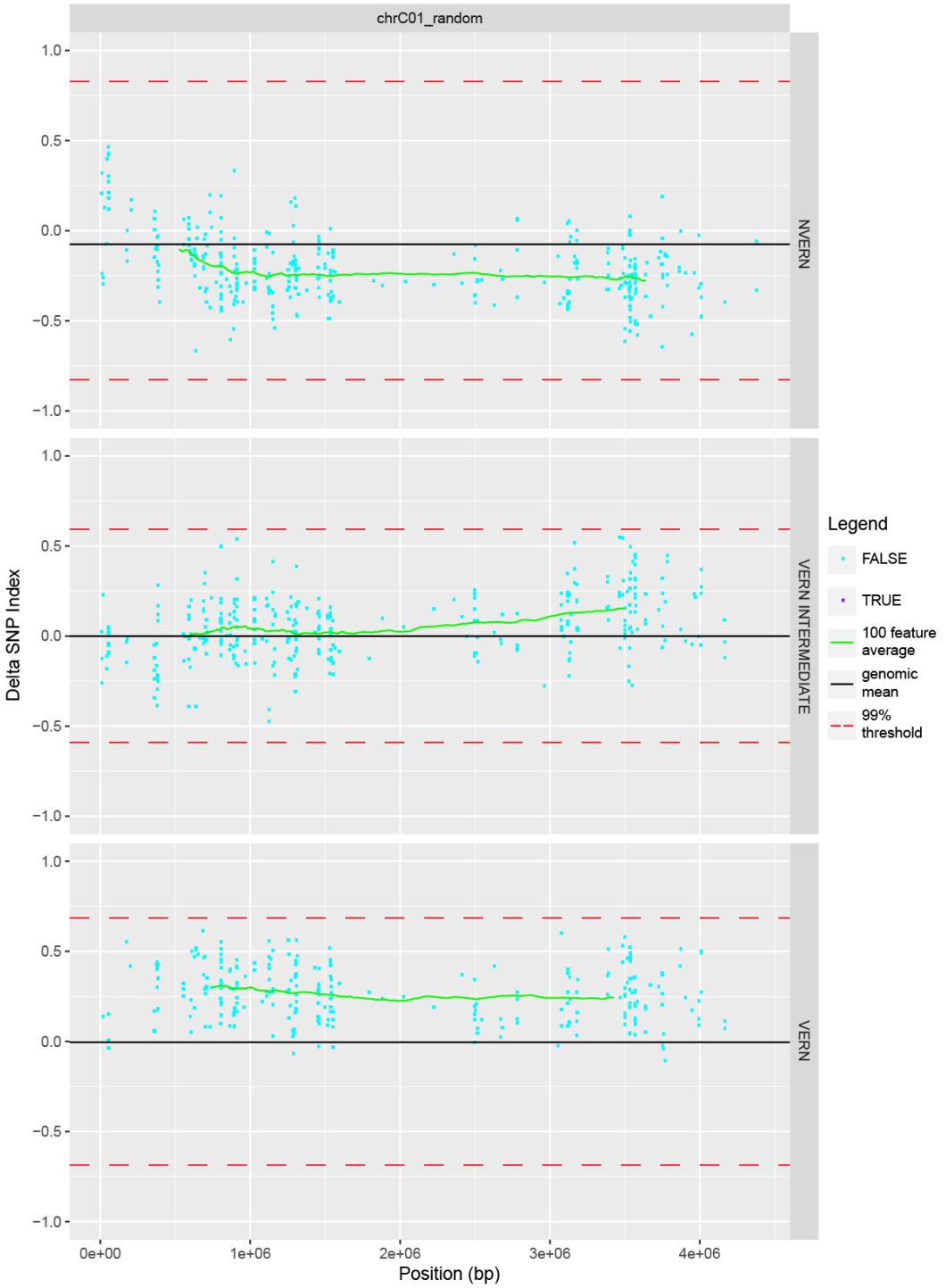


Figure S. 34: Chromosome C01_random. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in purple, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

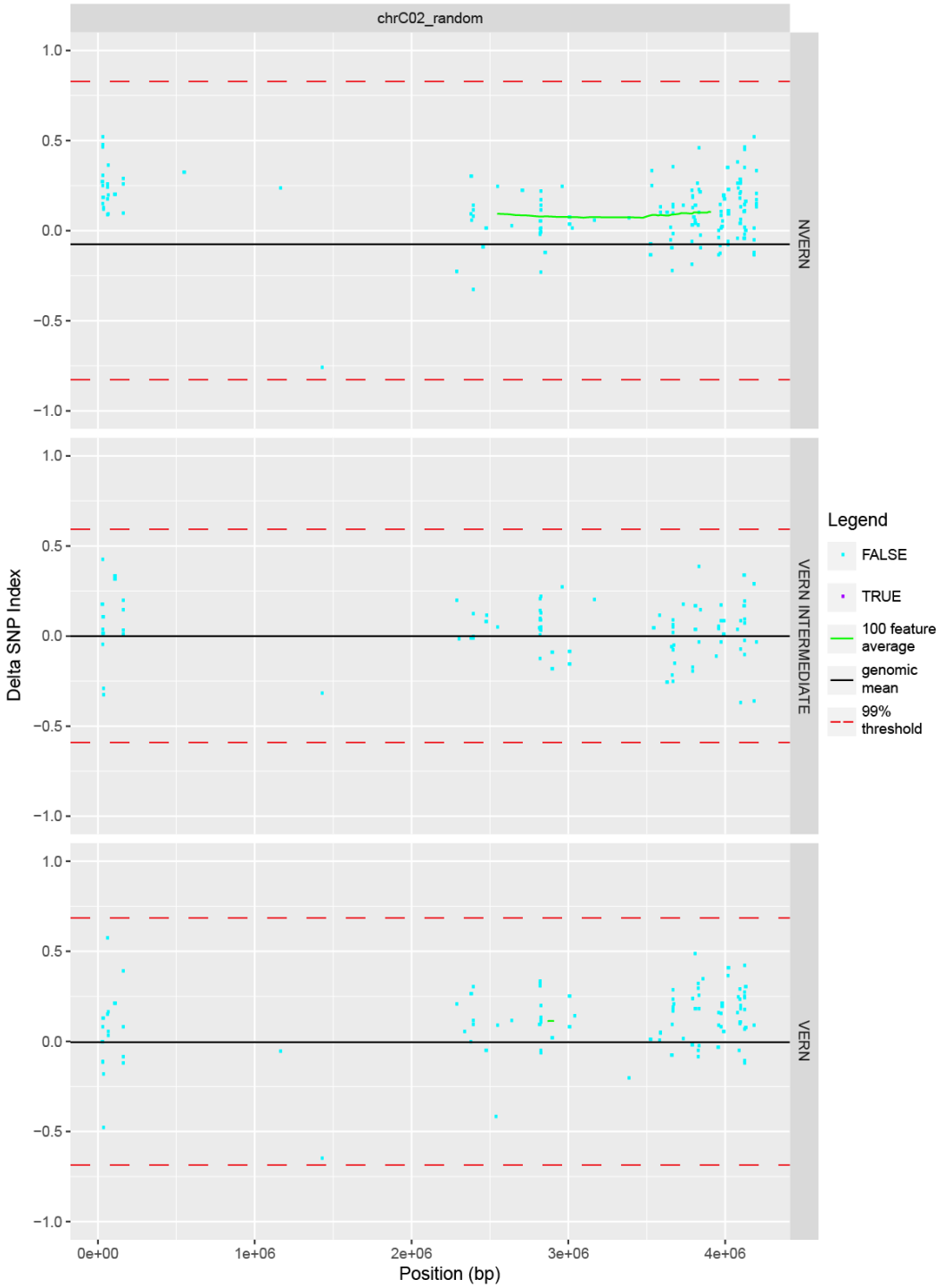


Figure S. 35: Chromosome C02_random. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in purple, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

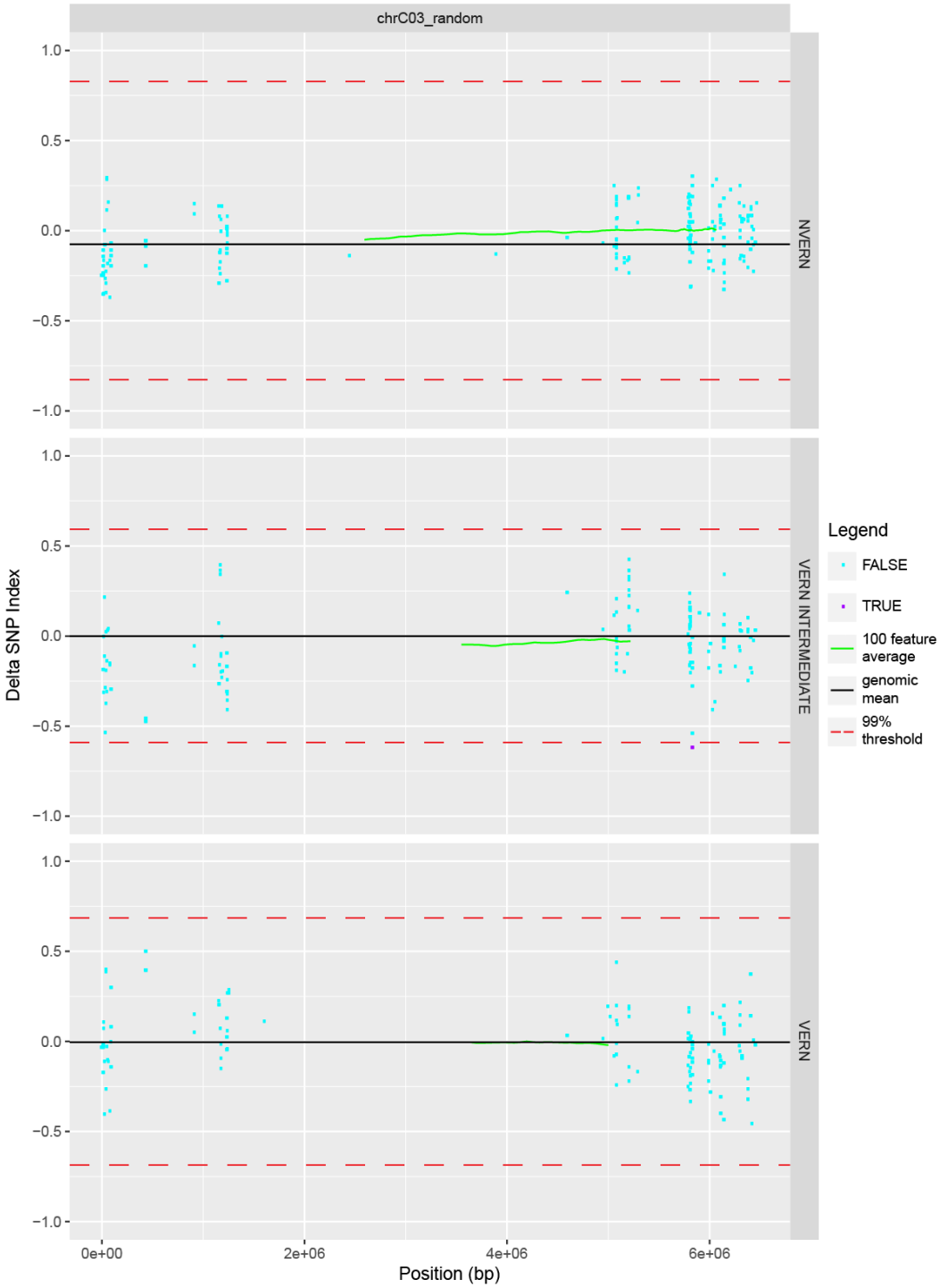


Figure S. 36: Chromosome C03_random. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in purple, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

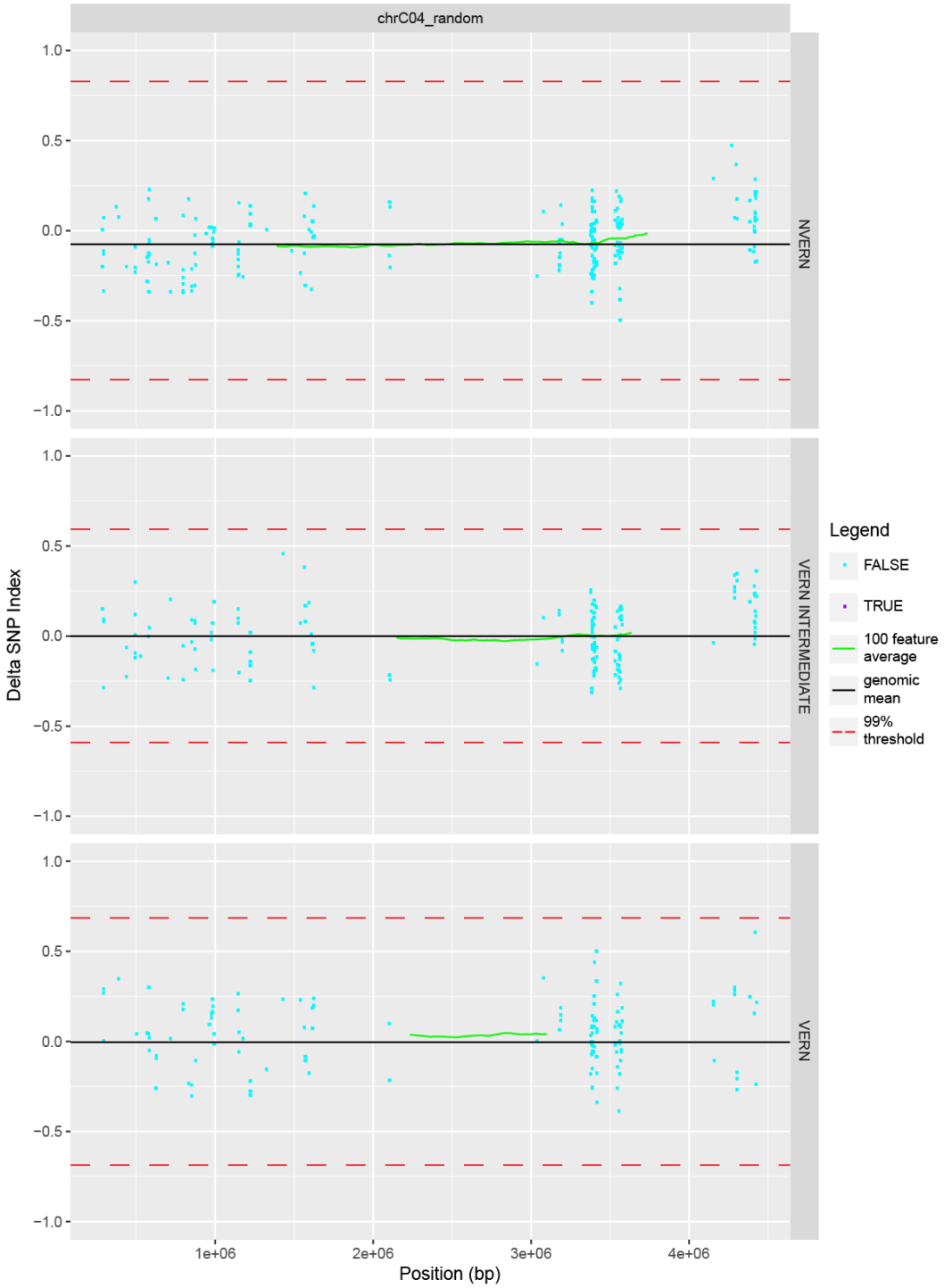


Figure S. 37: Chromosome C04_random. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in purple, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

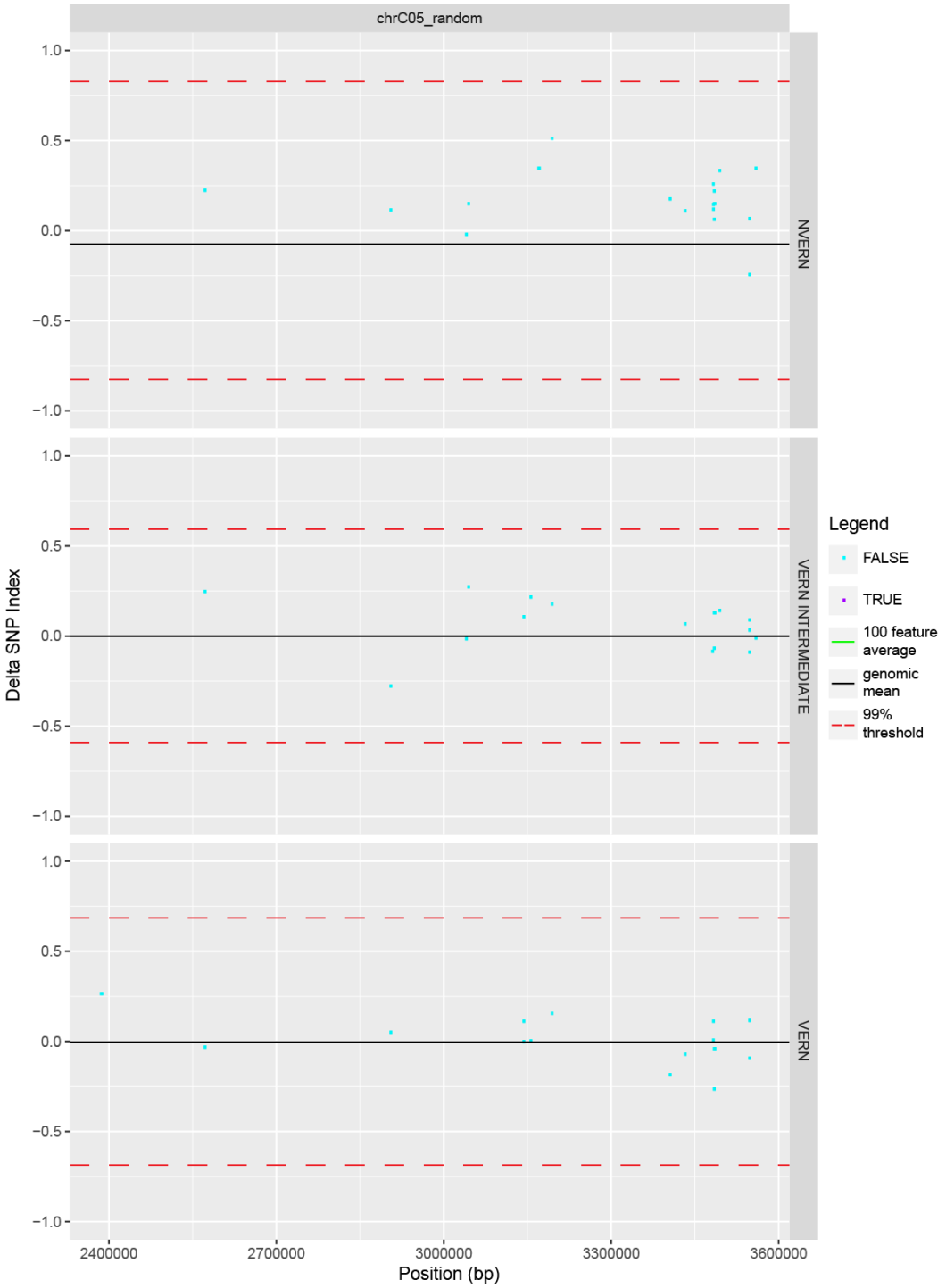


Figure S. 38: Chromosome C05_random. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in purple, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

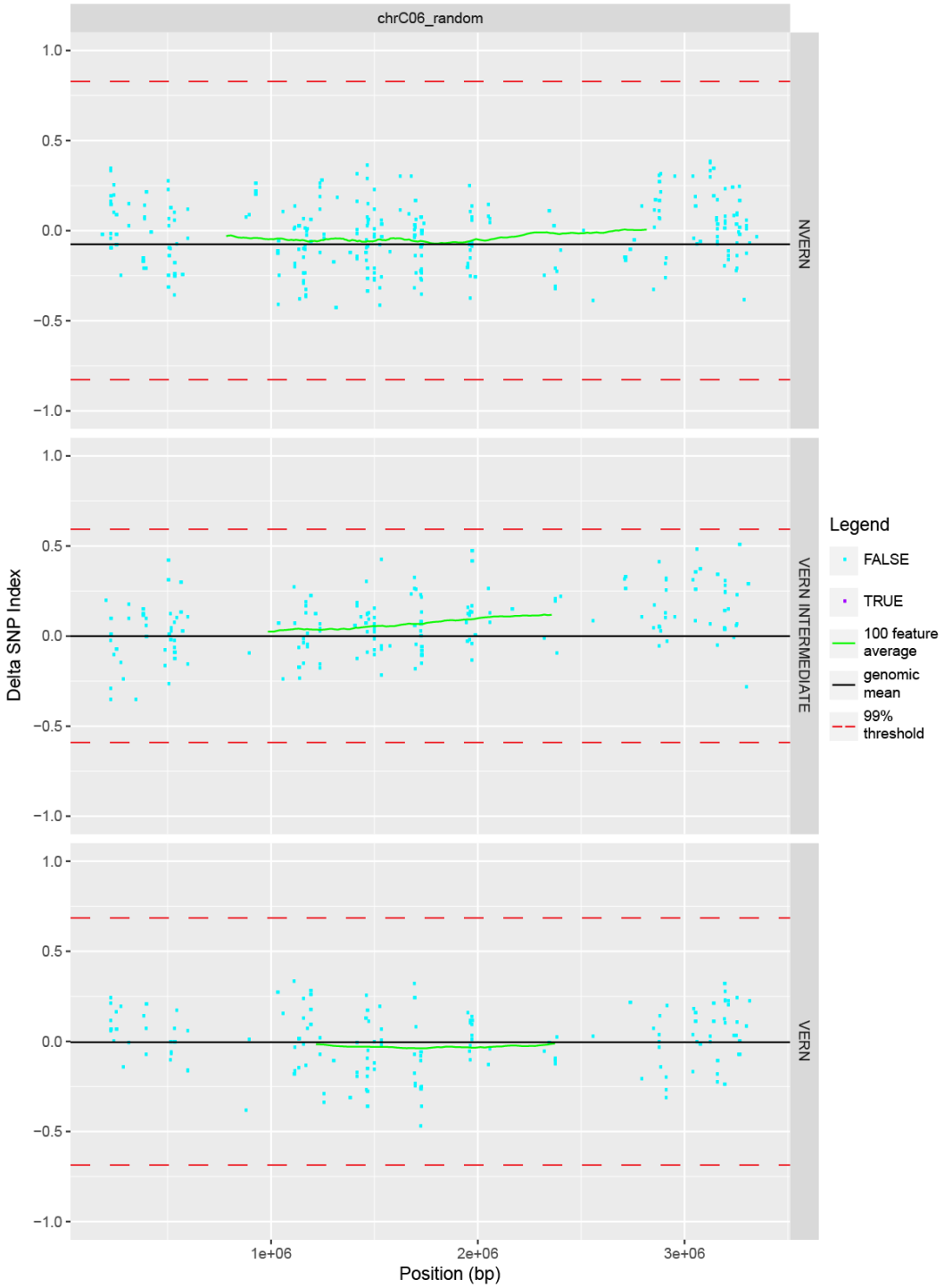


Figure S. 39: Chromosome C06_random. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in purple, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

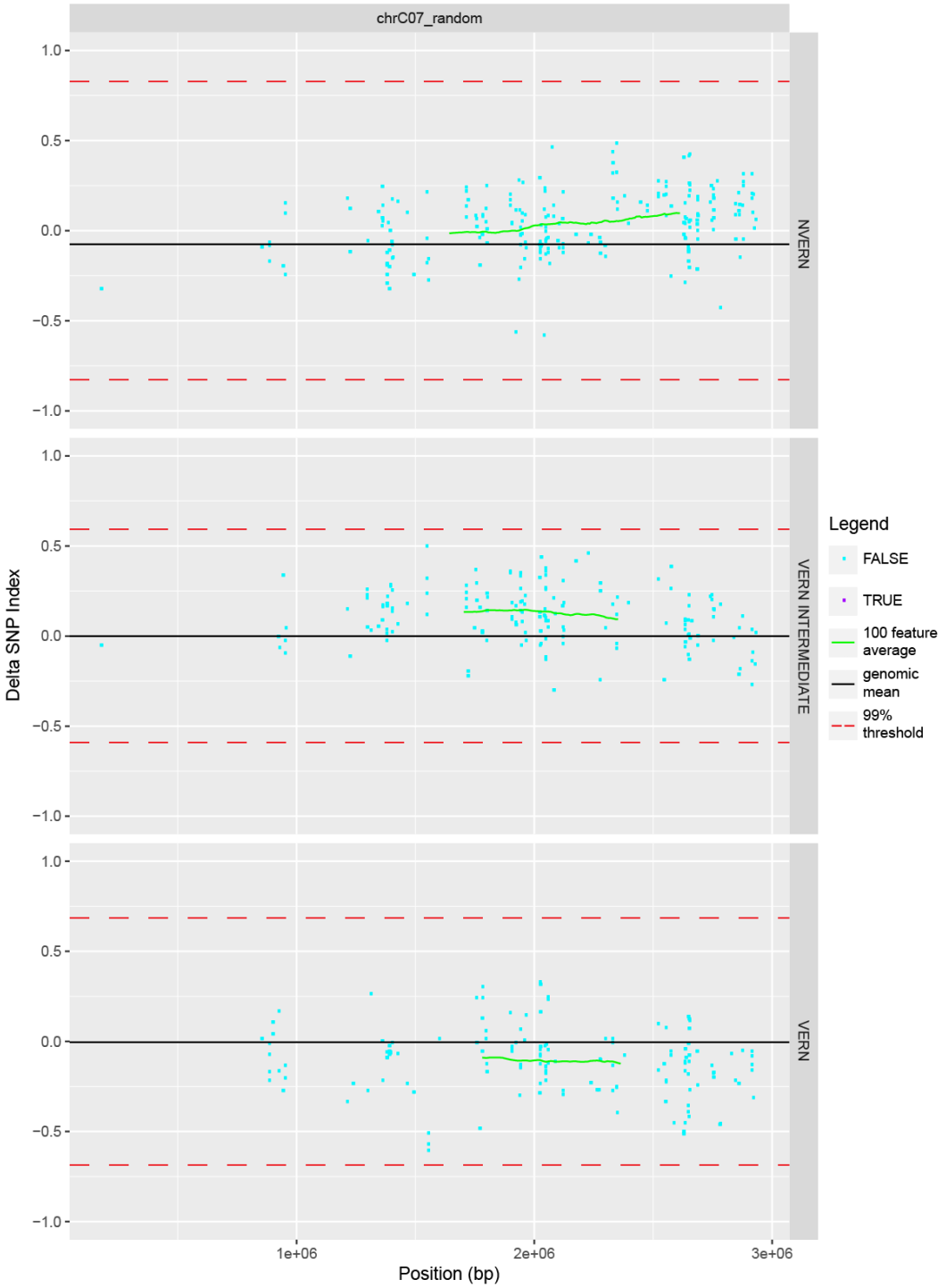


Figure S. 40: Chromosome C07_random. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in purple, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

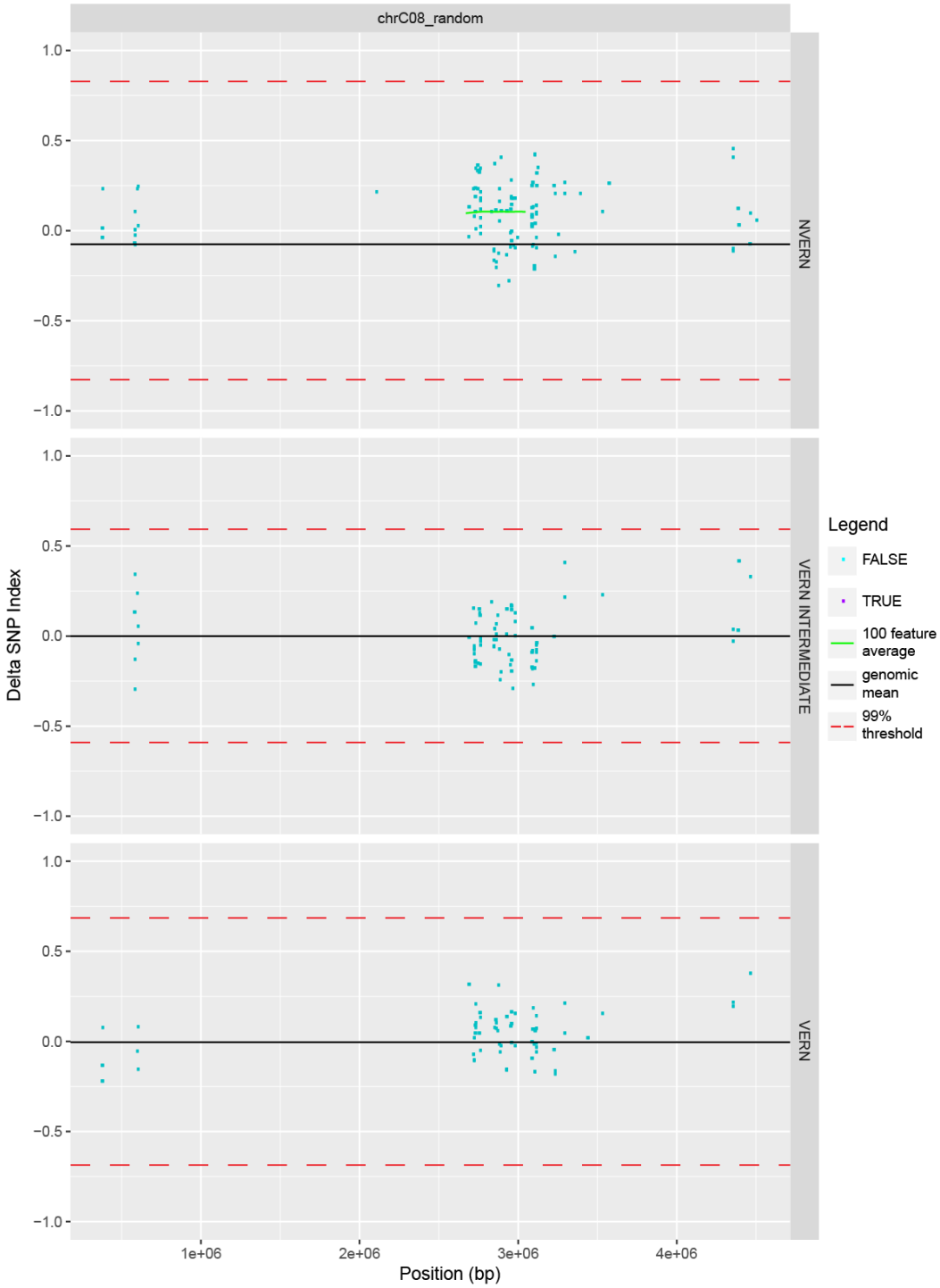


Figure S. 41: Chromosome C08_random. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in purple, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

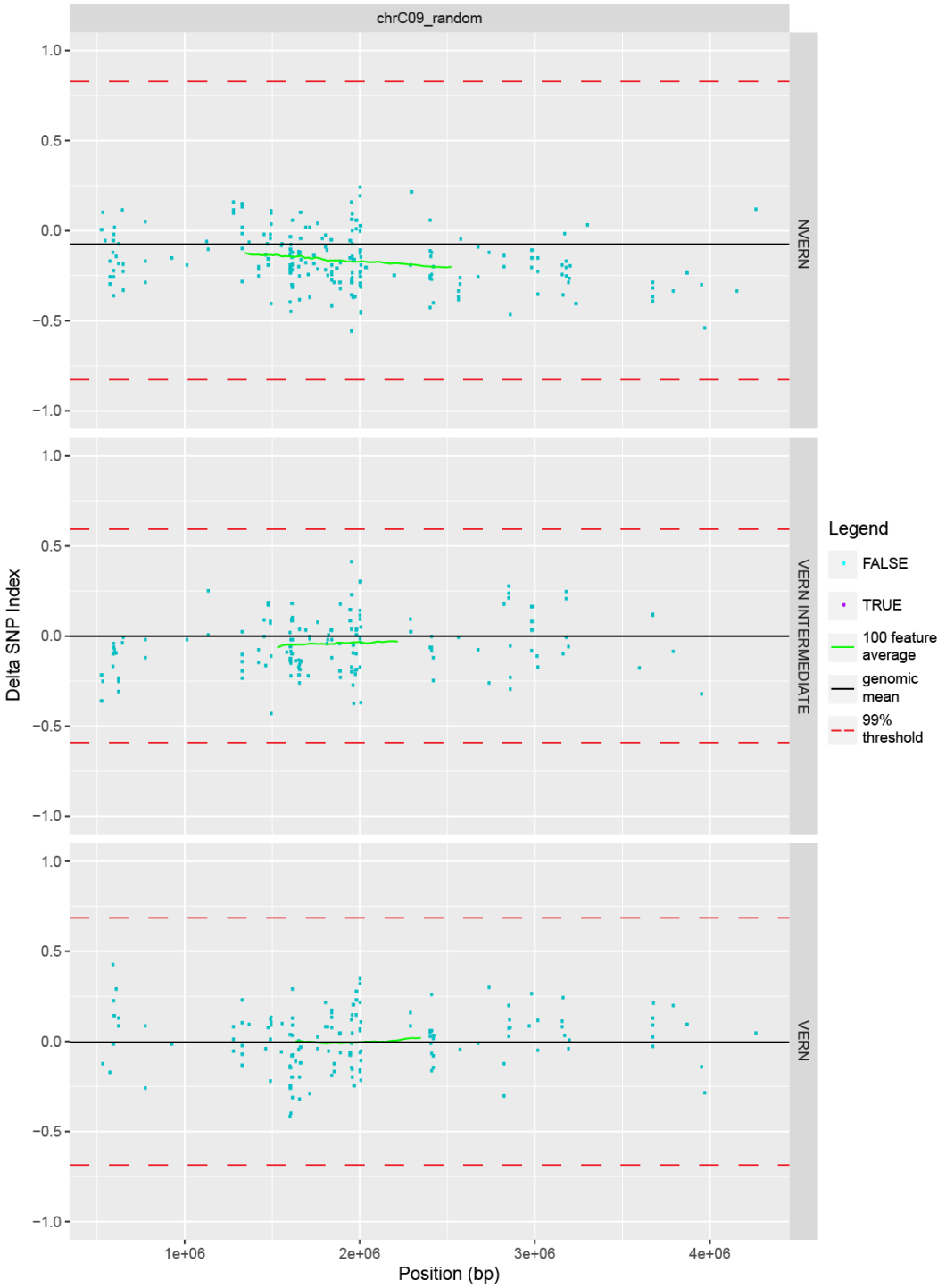


Figure S. 42: Chromosome C09_random. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in purple, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

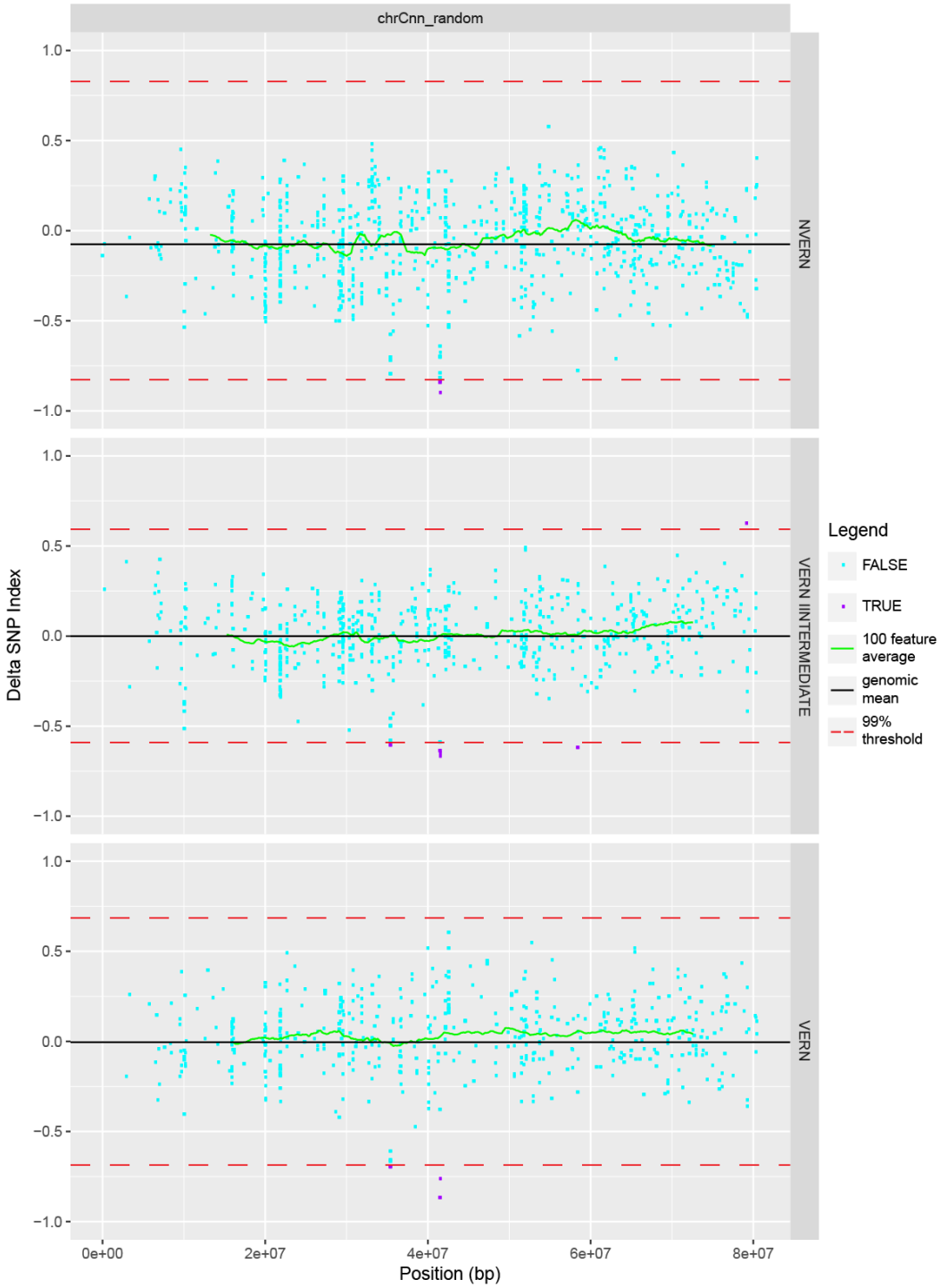


Figure S. 43: Chromosome Cnn_random. ΔSNP index values plotted against chromosomal position for NVERN (upper panel) VERN INTERMEDIATE (middle panel) and VERN (lower panel). ΔSNP index values showing no association are plotted in blue, ΔSNP index values that fall above or below the 99% threshold (red dashed line) are plotted in purple, genome average ΔSNP index value is represented by the black line, 100 ΔSNP index average value is plotted in green

8.2. Supplementary Tables

Table S. 1: The most significant GEMs that fall above the Bonferroni threshold of association for flowering time NVERN dataset. Listed in the table are the Unigene code indicating whether it is located on the A or C genome, the chromosomal location, and the *A. thaliana* gene code (AGI) that exhibits the highest homology. Significance levels are listed as -log10P and P-values.

Trait	Unigene code	Chromosome	AGI	-log10P	Pvalue
NVERN	A_JCVI_22584	A10	AT5G10140.1	12.83580071	1.46E-13
NVERN	C_JCVI_28286	C7	AT4G27710.1	10.49747745	3.18E-11
NVERN	C_JCVI_30211	C9	AT2G40550.1	10.40618668	3.92E-11
NVERN	A_JCVI_28286	A6	AT4G27710.1	10.26053947	5.49E-11
NVERN	C_EX086023	C1	AT1G18335.1	10.20766354	6.20E-11
NVERN	A_EX086023	A6	AT1G18335.1	10.11028319	7.76E-11
NVERN	A_JCVI_36536	A8	AT1G25490.1	10.07910168	8.33E-11
NVERN	C_JCVI_26040	C9	AT4G01910.1	9.88253035	1.31E-10
NVERN	C_JCVI_7295	C5	AT3G07310.1	9.791114193	1.62E-10
NVERN	A_JCVI_30211	A6	AT2G40550.1	9.682808298	2.08E-10
NVERN	C_JCVI_26403	C7	AT4G17810.1	9.674709337	2.11E-10
NVERN	A_JCVI_6943	A9	AT1G23860.1	9.578058539	2.64E-10
NVERN	A_JCVI_42366	A7	AT2G17410.1	9.551504945	2.81E-10
NVERN	A_JCVI_26040	A9	AT4G01910.1	9.471647273	3.38E-10
NVERN	C_JCVI_42366	C7	AT2G17410.1	9.356256193	4.40E-10
NVERN	A_ES929828	A10	AT1G03290.1	9.331257291	4.66E-10
NVERN	C_BQ704376	C1	AT1G18335.1	9.161671604	6.89E-10
NVERN	A_BQ704376	A6	AT1G18335.1	9.061127538	8.69E-10
NVERN	C_JCVI_28769	C5	NA	9.007546793	9.83E-10
NVERN	A_JCVI_24319	A2	NA	8.89926026	1.26E-09
NVERN	A_JCVI_5764	A3	AT5G10140.1	8.8941006	1.28E-09
NVERN	A_EV122333	A9	AT3G26040.1	8.886742381	1.30E-09
NVERN	C_JCVI_7980	C5	NA	8.745120031	1.80E-09
NVERN	C_JCVI_39585	C6	NA	8.733748636	1.85E-09
NVERN	C_EV207684	C8	NA	8.629124342	2.35E-09
NVERN	A_JCVI_19727	NA	NA	8.592651838	2.55E-09
NVERN	A_JCVI_28769	NA	NA	8.585359566	2.60E-09
NVERN	C_JCVI_21412	C3	AT3G09840.1	8.579517995	2.63E-09
NVERN	A_JCVI_7980	NA	NA	8.49423502	3.20E-09
NVERN	A_JCVI_39987	A2	AT5G19010.1	8.481892125	3.30E-09
NVERN	A_EX114883	A9	AT2G31530.1	8.404458983	3.94E-09
NVERN	C_EV124119	C9	AT3G26040.1	8.377916862	4.19E-09
NVERN	C_JCVI_14984	C2	AT1G12350.1	8.334071753	4.63E-09
NVERN	A_JCVI_39585	A2	NA	8.330331629	4.67E-09

Trait	Unigene code	Chromosome	AGI	-log10P	Pvalue
NVERN	A_EV093392	A9	NA	8.29533716	5.07E-09
NVERN	C_EV093392	C3	NA	8.293828361	5.08E-09
NVERN	C_JCVI_15052	C5	AT3G21300.1	8.21387608	6.11E-09
NVERN	A_JCVI_14001	A3	AT5G59050.1	8.19781831	6.34E-09
NVERN	C_EV143804	C9	AT4G00860.1	8.172760111	6.72E-09
NVERN	A_ES266451	A1	NA	8.134879262	7.33E-09
NVERN	C_JCVI_16315	C9	AT5G62000.1	8.095996268	8.02E-09
NVERN	C_JCVI_36536	C3	AT1G25490.1	8.049846918	8.92E-09
NVERN	A_JCVI_26231	A3	AT5G55990.1	8.02006297	9.55E-09
NVERN	A_EV101010	A5	AT3G04160.1	8.01470974	9.67E-09
NVERN	A_JCVI_31384	A2	AT5G60360.1	8.001380243	9.97E-09
NVERN	A_JCVI_38947	A1	NA	7.964883762	1.08E-08
NVERN	A_JCVI_20033	A5	AT3G06483.1	7.960023678	1.10E-08
NVERN	C_JCVI_38947	C6	NA	7.937708074	1.15E-08
NVERN	C_JCVI_31384	C2	AT5G60360.1	7.905656207	1.24E-08
NVERN	C_JCVI_14001	C2	AT5G59050.1	7.866713198	1.36E-08
NVERN	C_CD827468	C8	NA	7.817893005	1.52E-08
NVERN	A_JCVI_26403	A1	AT4G17810.1	7.768231152	1.71E-08
NVERN	C_EV122333	C9	AT3G26040.1	7.695561952	2.02E-08
NVERN	C_JCVI_1710	C4	AT2G47210.1	7.67075794	2.13E-08
NVERN	C_AM395988	C7	AT4G27710.1	7.59039267	2.57E-08
NVERN	A_EX063324	A1	AT1G60500.1	7.579504763	2.63E-08
NVERN	C_JCVI_39987	C3	AT5G19010.1	7.572897719	2.67E-08
NVERN	C_JCVI_29467	C4	AT2G44230.1	7.563839546	2.73E-08
NVERN	A_JCVI_14258	A10	AT5G05830.1	7.558934835	2.76E-08
NVERN	A_JCVI_42335	NA	NA	7.558580103	2.76E-08
NVERN	C_JCVI_15948	C2	AT5G10630.1	7.539419868	2.89E-08
NVERN	C_JCVI_20033	C5	AT3G06483.1	7.523870261	2.99E-08
NVERN	A_CD839266	NA	NA	7.520415827	3.02E-08
NVERN	A_JCVI_13341	A10	AT4G01480.1	7.508784249	3.10E-08
NVERN	A_CD827468	NA	NA	7.499512062	3.17E-08
NVERN	A_JCVI_16315	A6	AT5G62000.1	7.4697641	3.39E-08
NVERN	C_JCVI_28976	C9	AT5G05610.1	7.406476707	3.92E-08
NVERN	C_JCVI_24319	NA	NA	7.398328208	4.00E-08
NVERN	C_JCVI_8575	C2	AT5G16290.1	7.38807897	4.09E-08
NVERN	A_EV207684	A7	NA	7.363270479	4.33E-08
NVERN	C_EV091771	C6	AT4G02440.1	7.326033534	4.72E-08
NVERN	C_ES898453	C3	AT5G56240.1	7.271322451	5.35E-08
NVERN	A_JCVI_11358	A10	AT5G10010.1	7.203973209	6.25E-08
NVERN	C_JCVI_18754	C5	AT1G18660.2	7.195432751	6.38E-08
NVERN	A_JCVI_31529	A5	AT3G01470.1	7.183825231	6.55E-08
NVERN	A_DY016401	A5	NA	7.11817097	7.62E-08
NVERN	A_EV164090	A9	AT2G18160.1	6.994825794	1.01E-07

Trait	Unigene code	Chromosome	AGI	-log10P	Pvalue
NVERN	A_JCVI_17091	A5	AT3G01470.1	6.980948189	1.04E-07
NVERN	A_JCVI_11946	A8	AT1G25490.1	6.955991807	1.11E-07
NVERN	C_JCVI_12972	C3	AT4G19830.1	6.929169565	1.18E-07
NVERN	C_EE558140	C6	AT1G79570.1	6.909915076	1.23E-07
NVERN	C_JCVI_22041	C9	NA	6.880016559	1.32E-07
NVERN	C_EE418092	C1	NA	6.851973641	1.41E-07
NVERN	A_JCVI_7586	A5	AT5G16110.1	6.835339093	1.46E-07
NVERN	C_JCVI_2518	C4	AT2G45660.1	6.824528785	1.50E-07
NVERN	C_JCVI_3473	C2	AT5G18110.1	6.795205102	1.60E-07
NVERN	C_JCVI_16128	C3	AT2G40080.1	6.787092747	1.63E-07
NVERN	C_JCVI_23438	C7	AT2G05440.3	6.763190198	1.73E-07
NVERN	A_AM387806	NA	NA	6.752504813	1.77E-07
NVERN	C_CX188778	C9	AT5G05550.1	6.735637055	1.84E-07
NVERN	C_ES989775	C1	AT4G19350.1	6.719741416	1.91E-07
NVERN	C_EV194620	C8	NA	6.70680066	1.96E-07
NVERN	A_JCVI_10591	A10	AT5G05920.1	6.692251234	2.03E-07
NVERN	A_EE433946	A4	NA	6.684285948	2.07E-07
NVERN	A_JCVI_11694	A10	AT5G03900.2	6.626265905	2.36E-07
NVERN	A_JCVI_18970	A3	AT5G53440.1	6.606090073	2.48E-07
NVERN	C_EV099565	C1	AT1G64190.1	6.587949737	2.58E-07
NVERN	A_ES898453	A10	AT5G56240.1	6.577545455	2.65E-07
NVERN	A_EV100942	A5	AT2G14080.1	6.576104668	2.65E-07
NVERN	C_EX114883	C6	AT2G31530.1	6.565182459	2.72E-07
NVERN	C_EV147935	C2	AT5G17070.1	6.552177994	2.80E-07
NVERN	C_JCVI_23485	C7	NA	6.551154648	2.81E-07
NVERN	C_JCVI_7911	C1	AT3G51430.1	6.551088829	2.81E-07
NVERN	A_JCVI_23438	NA	AT2G05440.3	6.468805708	3.40E-07
NVERN	C_JCVI_18289	C9	AT5G24430.1	6.456636063	3.49E-07
NVERN	C_JCVI_27623	C6	AT1G75390.1	6.448812603	3.56E-07
NVERN	A_JCVI_1108	A10	AT1G07080.1	6.436468252	3.66E-07
NVERN	A_EV136958	A5	AT3G09840.1	6.415146887	3.84E-07
NVERN	A_EE452610	NA	NA	6.406905416	3.92E-07
NVERN	A_EE418092	A8	NA	6.388548279	4.09E-07
NVERN	C_JCVI_7013	C4	AT1G48400.1	6.38843577	4.09E-07
NVERN	C_JCVI_11216	C4	AT2G37130.1	6.378453752	4.18E-07
NVERN	C_JCVI_760	C3	NA	6.372389236	4.24E-07
NVERN	C_JCVI_20513	C5	AT1G05850.1	6.366883529	4.30E-07
NVERN	C_JCVI_11946	C3	AT1G25490.1	6.364445069	4.32E-07
NVERN	C_JCVI_11095	C5	AT1G53240.1	6.360309398	4.36E-07
NVERN	A_JCVI_23485	NA	NA	6.334299662	4.63E-07
NVERN	A_JCVI_16128	A3	AT2G40080.1	6.289600511	5.13E-07
NVERN	C_EE474234	C6	AT4G30220.1	6.276785899	5.29E-07
NVERN	C_EV101010	C3	AT3G04160.1	6.239032381	5.77E-07

Trait	Unigene code	Chromosome	AGI	-log10P	Pvalue
NVERN	A_JCVI_12972	A8	AT4G19830.1	6.218818359	6.04E-07
NVERN	A_JCVI_30297	NA	NA	6.218350429	6.05E-07
NVERN	A_JCVI_15273	A6	AT1G51950.1	6.192145744	6.42E-07
NVERN	A_JCVI_940	A7	AT1G70670.1	6.1752004	6.68E-07
NVERN	C_EL592177	C8	AT2G19740.1	6.162336691	6.88E-07
NVERN	C_JCVI_30423	C2	AT1G78810.2	6.151245347	7.06E-07
NVERN	A_JCVI_31284	A2	AT5G17190.1	6.147124996	7.13E-07
NVERN	A_JCVI_1144	A5	AT3G01440.1	6.137961298	7.28E-07
NVERN	A_JCVI_196	A10	AT5G03880.1	6.131340182	7.39E-07
NVERN	C_EV008705	C2	AT5G13470.1	6.115401586	7.67E-07
NVERN	A_JCVI_12363	A9	AT1G30810.1	6.11210427	7.72E-07
NVERN	C_JCVI_9369	C3	NA	6.108114438	7.80E-07
NVERN	A_JCVI_19103	A2	AT5G49360.1	6.103089881	7.89E-07
NVERN	A_JCVI_41736	A3	NA	6.097301115	7.99E-07
NVERN	C_EH418918	C1	AT1G63460.1	6.092127318	8.09E-07
NVERN	C_JCVI_4780	C6	NA	6.089811027	8.13E-07
NVERN	A_JCVI_29900	A9	AT1G18680.1	6.08661942	8.19E-07
NVERN	A_EX096507	A5	AT2G34300.1	6.085105387	8.22E-07
NVERN	C_JCVI_1876	C2	AT5G53570.1	6.07476993	8.42E-07
NVERN	C_JCVI_13458	C8	AT1G04400.2	6.073212856	8.45E-07
NVERN	A_DN962837	A5	AT3G03305.1	6.069353327	8.52E-07
NVERN	C_JCVI_3005	C2	AT5G03290.1	6.060339957	8.70E-07
NVERN	C_AM395895	C9	AT5G02230.1	6.041981854	9.08E-07
NVERN	A_EL592177	A6	AT2G19740.1	6.039891617	9.12E-07
NVERN	A_JCVI_4780	A5	NA	6.025431015	9.43E-07
NVERN	C_EE452610	C2	NA	6.018270106	9.59E-07
NVERN	C_EE532472	C8	NA	6.006672415	9.85E-07
NVERN	C_JCVI_9610	C9	AT5G27820.1	5.997614729	1.01E-06
NVERN	A_CN729422	A3	AT5G22420.1	5.99113368	1.02E-06
NVERN	C_JCVI_12425	C5	AT1G24030.1	5.988426125	1.03E-06
NVERN	A_EV091356	A5	AT2G45660.1	5.980970835	1.04E-06
NVERN	A_JCVI_19610	A10	AT5G18660.1	5.979297413	1.05E-06
NVERN	A_JCVI_24463	A10	AT5G23140.1	5.971208432	1.07E-06
NVERN	A_EV179210	A4	AT3G52370.1	5.963920864	1.09E-06
NVERN	A_JCVI_18937	A1	AT3G10525.1	5.963023632	1.09E-06
NVERN	C_JCVI_23306	C9	AT5G03760.1	5.954818316	1.11E-06
NVERN	C_EV216648	C4	AT1G48400.1	5.949970447	1.12E-06
NVERN	A_JCVI_27655	A2	NA	5.93734067	1.16E-06
NVERN	A_JCVI_9369	A3	NA	5.905106378	1.24E-06
NVERN	A_JCVI_34086	A10	AT5G02270.1	5.897273427	1.27E-06
NVERN	C_EE433946	C3	NA	5.893865444	1.28E-06
NVERN	C_JCVI_40018	C1	NA	5.891683633	1.28E-06
NVERN	A_EV141032	A6	AT1G18600.1	5.88013342	1.32E-06

Trait	Unigene code	Chromosome	AGI	-log10P	Pvalue
NVERN	A_JCVI_5649	A1	AT3G15510.1	5.875714432	1.33E-06
NVERN	C_CO750600	C7	AT4G11000.1	5.874373935	1.34E-06
NVERN	A_JCVI_41605	A5	AT3G19340.1	5.873955973	1.34E-06
NVERN	A_EX126383	A7	AT2G20120.1	5.854537287	1.40E-06
NVERN	C_ES944788	C3	AT1G17160.1	5.854239685	1.40E-06
NVERN	C_JCVI_39250	C4	AT5G36880.2	5.852545106	1.40E-06
NVERN	A_EX035993	A9	AT1G23800.1	5.849075996	1.42E-06
NVERN	A_EV133152	A3	AT2G28605.1	5.843825632	1.43E-06
NVERN	A_JCVI_33168	A9	AT3G28180.1	5.837977595	1.45E-06
NVERN	A_JCVI_25277	A3	AT3G09670.1	5.837658651	1.45E-06
NVERN	A_EX069636	A9	AT1G59840.1	5.835036404	1.46E-06
NVERN	A_EV099565	NA	AT1G64190.1	5.834676384	1.46E-06
NVERN	A_JCVI_8144	A9	AT5G24400.1	5.812958217	1.54E-06
NVERN	A_AM395988	A6	AT4G27710.1	5.805983443	1.56E-06
NVERN	A_JCVI_521	A9	AT2G18160.1	5.795086985	1.60E-06
NVERN	A_JCVI_14620	NA	AT1G64190.1	5.793030561	1.61E-06
NVERN	A_JCVI_35542	NA	NA	5.792643408	1.61E-06
NVERN	A_EE532472	A1	NA	5.792455033	1.61E-06
NVERN	A_JCVI_31755	A6	AT1G18800.1	5.7818768	1.65E-06
NVERN	A_JCVI_25167	A10	AT1G02890.1	5.781308209	1.65E-06
NVERN	A_JCVI_36868	A3	AT5G10140.1	5.770643793	1.70E-06
NVERN	C_JCVI_11909	C9	AT5G02230.1	5.748671601	1.78E-06
NVERN	C_EX075124	C5	AT1G26550.1	5.733241878	1.85E-06
NVERN	A_EV100560	A2	AT5G17170.1	5.730978151	1.86E-06
NVERN	C_AM390598	C6	AT1G76880.1	5.726849801	1.88E-06
NVERN	C_JCVI_9563	C2	AT2G46580.1	5.723236196	1.89E-06
NVERN	A_JCVI_35896	A10	AT5G03910.1	5.721909421	1.90E-06
NVERN	A_JCVI_19612	A7	AT1G79090.1	5.721670918	1.90E-06
NVERN	C_JCVI_41485	C8	AT4G11050.1	5.718350307	1.91E-06
NVERN	A_JCVI_11404	A2	AT5G58670.1	5.716998696	1.92E-06
NVERN	A_JCVI_24601	A4	NA	5.714948786	1.93E-06
NVERN	C_JCVI_35542	C2	NA	5.710570925	1.95E-06
NVERN	C_JCVI_36265	C3	AT5G07240.1	5.705948234	1.97E-06
NVERN	C_EV187956	C3	AT4G03280.1	5.7041417	1.98E-06
NVERN	C_JCVI_30740	C5	AT3G07560.1	5.701781144	1.99E-06
NVERN	A_JCVI_2275	A10	AT5G11280.1	5.700601338	1.99E-06
NVERN	C_JCVI_21198	C3	NA	5.692699431	2.03E-06
NVERN	A_JCVI_34660	A5	AT3G01460.1	5.686336921	2.06E-06
NVERN	A_ES936825	A9	NA	5.683460858	2.07E-06
NVERN	C_JCVI_39896	C1	AT4G18370.1	5.677145091	2.10E-06
NVERN	C_JCVI_4593	C7	NA	5.673773396	2.12E-06
NVERN	A_DY002257	A4	AT2G36390.1	5.673221802	2.12E-06
NVERN	C_ES906796	C3	NA	5.667744525	2.15E-06

Trait	Unigene code	Chromosome	AGI	-log10P	Pvalue
NVERN	A_JCVI_29571	A2	NA	5.667349584	2.15E-06
NVERN	C_EV133152	C3	AT2G28605.1	5.662035935	2.18E-06
NVERN	A_JCVI_19007	NA	NA	5.657898337	2.20E-06
NVERN	A_JCVI_11953	A10	AT5G56240.1	5.657242419	2.20E-06
NVERN	A_JCVI_2252	A9	AT1G09640.1	5.653960159	2.22E-06
NVERN	C_EV193289	C2	AT5G59670.1	5.644962311	2.26E-06
NVERN	A_JCVI_30596	A7	NA	5.63727895	2.31E-06
NVERN	C_JCVI_42335	C3	NA	5.63460305	2.32E-06
NVERN	C_JCVI_7816	C1	AT3G12560.1	5.63193035	2.33E-06
NVERN	A_CX281482	A1	AT4G32130.1	5.624394906	2.37E-06
NVERN	C_JCVI_28974	C8	NA	5.615441581	2.42E-06
NVERN	C_CX189980	C7	NA	5.608664269	2.46E-06
NVERN	A_EV208166	A10	AT1G05200.1	5.597320111	2.53E-06
NVERN	C_JCVI_11158	C3	AT4G33550.2	5.588150329	2.58E-06
NVERN	A_JCVI_24569	A8	AT1G28440.1	5.579697222	2.63E-06
NVERN	A_EE445303	A7	AT1G79570.1	5.576465378	2.65E-06
NVERN	C_ES911373	C2	AT1G65070.1	5.573548088	2.67E-06
NVERN	A_JCVI_37875	A2	NA	5.568959576	2.70E-06
NVERN	A_EV193289	A3	AT5G59670.1	5.564671748	2.72E-06
NVERN	A_EV002610	NA	AT5G55960.1	5.564184336	2.73E-06
NVERN	A_JCVI_37153	A10	AT5G02270.1	5.563980594	2.73E-06
NVERN	A_JCVI_35326	A2	NA	5.560345485	2.75E-06
NVERN	A_EV048293	NA	NA	5.553929281	2.79E-06
NVERN	A_JCVI_6160	A9	AT1G33265.1	5.552751977	2.80E-06
NVERN	A_DY030278	A3	AT2G32970.1	5.547061606	2.84E-06
NVERN	C_JCVI_23665	C4	AT3G54260.1	5.545709004	2.85E-06
NVERN	A_JCVI_14757	A5	AT3G08890.1	5.545179869	2.85E-06
NVERN	A_EH418918	A9	AT1G63460.1	5.542634586	2.87E-06
NVERN	C_JCVI_1598	C3	AT4G20130.1	5.53939278	2.89E-06
NVERN	C_EV081616	C3	AT3G15880.1	5.537004195	2.90E-06
NVERN	C_JCVI_4270	C9	AT5G06430.1	5.534418495	2.92E-06
NVERN	A_JCVI_25403	A10	AT5G54280.1	5.530838907	2.95E-06
NVERN	C_JCVI_24601	C5	NA	5.523843271	2.99E-06
NVERN	A_JCVI_7286	A9	AT4G08790.1	5.515590559	3.05E-06
NVERN	C_JCVI_33693	C4	AT2G46490.1	5.51444344	3.06E-06
NVERN	C_CD836475	C9	AT5G03770.1	5.510931566	3.08E-06
NVERN	C_EV100560	C2	AT5G17170.1	5.503186331	3.14E-06
NVERN	C_EE448355	C7	AT4G33670.1	5.501956141	3.15E-06
NVERN	A_JCVI_22041	NA	NA	5.501954876	3.15E-06
NVERN	A_JCVI_31236	A7	AT1G78670.1	5.499500363	3.17E-06
NVERN	C_CD823858	C2	AT1G76700.1	5.497543119	3.18E-06
NVERN	A_JCVI_22178	A2	AT5G10140.1	5.494885652	3.20E-06
NVERN	C_JCVI_19007	C4	NA	5.491215955	3.23E-06

Trait	Unigene code	Chromosome	AGI	-log10P	Pvalue
NVERN	A_JCVI_13316	A3	AT5G08560.1	5.489126761	3.24E-06
NVERN	A_AM390716	A2	NA	5.485348527	3.27E-06
NVERN	C_JCVI_23331	C5	NA	5.482648169	3.29E-06
NVERN	C_JCVI_18385	C9	NA	5.479838901	3.31E-06
NVERN	A_EX075124	A9	AT1G26550.1	5.474316848	3.35E-06
NVERN	A_ES906796	NA	NA	5.462989105	3.44E-06
NVERN	A_JCVI_16988	A2	AT3G13180.1	5.456855413	3.49E-06
NVERN	C_EX027565	C8	AT2G25580.1	5.4567242	3.49E-06
NVERN	A_JCVI_42017	A3	NA	5.454889473	3.51E-06
NVERN	A_EE557072	A9	AT5G47840.1	5.454500038	3.51E-06
NVERN	C_JCVI_19103	C2	AT5G49360.1	5.45358545	3.52E-06
NVERN	C_EE476288	C9	NA	5.449468986	3.55E-06
NVERN	C_JCVI_19727	C3	NA	5.447758254	3.57E-06
NVERN	C_EV100942	C5	AT2G14080.1	5.445873198	3.58E-06
NVERN	A_JCVI_34076	A5	AT3G52800.1	5.444065191	3.60E-06
NVERN	A_DW998159	A2	AT5G60370.1	5.44206994	3.61E-06
NVERN	A_EV125413	A5	AT5G53480.1	5.441251287	3.62E-06
NVERN	A_EV030996	A8	NA	5.431103776	3.71E-06
NVERN	C_JCVI_29396	C2	AT5G57580.1	5.42873121	3.73E-06
NVERN	C_JCVI_27187	NA	NA	5.427290389	3.74E-06
NVERN	C_JCVI_39368	C2	AT5G56240.1	5.417814969	3.82E-06
NVERN	C_JCVI_9072	C4	AT2G35120.1	5.416570876	3.83E-06
NVERN	C_JCVI_26932	C7	AT4G28920.1	5.415521751	3.84E-06
NVERN	A_JCVI_23327	NA	NA	5.408862751	3.90E-06
NVERN	A_JCVI_29918	A9	NA	5.407582327	3.91E-06
NVERN	C_EV195955	C2	AT5G03940.1	5.40757401	3.91E-06
NVERN	C_EV091962	NA	NA	5.399995542	3.98E-06
NVERN	C_JCVI_13341	C9	AT4G01480.1	5.399902471	3.98E-06
NVERN	A_CD813695	A8	AT4G38630.1	5.393137596	4.04E-06
NVERN	C_EE451261	C2	AT5G21482.1	5.390132034	4.07E-06
NVERN	C_EX123200	C9	AT5G04360.1	5.389212663	4.08E-06
NVERN	C_CX188349	C2	AT5G21482.1	5.384315125	4.13E-06
NVERN	A_EV201379	A5	AT3G19340.1	5.381501288	4.15E-06
NVERN	C_JCVI_3294	C5	AT3G21300.1	5.380332312	4.17E-06
NVERN	C_JCVI_2381	C1	AT1G58210.1	5.3776738	4.19E-06
NVERN	C_JCVI_37875	C2	NA	5.377061786	4.20E-06
NVERN	C_JCVI_37448	C2	AT5G17050.1	5.376812761	4.20E-06
NVERN	C_JCVI_18136	C2	AT5G49360.1	5.357743086	4.39E-06
NVERN	C_JCVI_34076	C5	AT3G52800.1	5.35429632	4.42E-06
NVERN	A_ES918206	A2	AT1G66970.1	5.351691413	4.45E-06
NVERN	C_JCVI_4004	C2	AT5G02780.1	5.349946061	4.47E-06
NVERN	A_DY020311	A10	AT5G58900.1	5.347404643	4.49E-06
NVERN	A_JCVI_28108	A5	AT2G39840.1	5.34620266	4.51E-06

Trait	Unigene code	Chromosome	AGI	-log10P	Pvalue
NVERN	C_JCVI_10879	C2	NA	5.345353542	4.51E-06
NVERN	C_EE531832	C3	AT3G17900.1	5.341269051	4.56E-06
NVERN	C_JCVI_7847	C1	AT4G18370.1	5.335655394	4.62E-06
NVERN	A_JCVI_7840	A10	AT5G15800.1	5.33180573	4.66E-06
NVERN	C_EV091356	C4	AT2G45660.1	5.328904677	4.69E-06
NVERN	C_JCVI_29571	NA	NA	5.326904748	4.71E-06
NVERN	C_EV002610	C3	AT5G55960.1	5.321008561	4.78E-06
NVERN	C_JCVI_10381	C4	AT2G36390.1	5.320789327	4.78E-06
NVERN	A_JCVI_41136	A9	AT1G12760.1	5.311193333	4.88E-06
NVERN	A_JCVI_18134	A10	AT5G04240.1	5.309104296	4.91E-06
NVERN	A_JCVI_20114	A10	AT5G10320.2	5.30907686	4.91E-06
NVERN	A_EX044090	A3	AT5G10140.1	5.306532733	4.94E-06
NVERN	A_JCVI_38537	A1	NA	5.305484418	4.95E-06
NVERN	A_JCVI_6502	A7	AT1G24260.2	5.304953358	4.96E-06
NVERN	C_JCVI_25403	C2	AT5G54280.1	5.291531257	5.11E-06
NVERN	C_EE432271	C3	NA	5.289537821	5.13E-06
NVERN	C_JCVI_24898	C2	AT5G14790.1	5.285775825	5.18E-06
NVERN	A_JCVI_30412	A1	AT4G19380.1	5.28080495	5.24E-06
NVERN	A_JCVI_34628	A9	NA	5.280259161	5.24E-06
NVERN	A_JCVI_8689	A3	AT1G60140.1	5.278345956	5.27E-06
NVERN	A_JCVI_40108	A2	AT1G65480.1	5.264500532	5.44E-06
NVERN	A_EV130088	A5	AT3G07090.1	5.263676334	5.45E-06
NVERN	A_JCVI_826	NA	NA	5.259933716	5.50E-06
NVERN	A_EV162138	A10	AT5G02310.1	5.258398042	5.52E-06
NVERN	C_ES987566	C5	AT1G09815.1	5.253107489	5.58E-06
NVERN	A_JCVI_28419	A10	AT5G02250.1	5.252175094	5.60E-06
NVERN	C_JCVI_23269	C9	AT3G54610.1	5.247410475	5.66E-06
NVERN	A_EV109260	NA	AT4G27435.1	5.236118726	5.81E-06
NVERN	C_JCVI_35731	C4	NA	5.230907379	5.88E-06
NVERN	C_JCVI_29918	NA	NA	5.21875317	6.04E-06
NVERN	A_JCVI_36043	A9	AT2G03270.1	5.206034403	6.22E-06
NVERN	C_JCVI_6502	C5	AT1G24260.2	5.204801473	6.24E-06
NVERN	C_EV004601	C4	AT2G39800.1	5.204653516	6.24E-06
NVERN	C_JCVI_30596	NA	NA	5.202303706	6.28E-06
NVERN	C_EV035510	C3	NA	5.201974795	6.28E-06
NVERN	A_EE516189	A2	AT5G63370.1	5.200937898	6.30E-06
NVERN	C_JCVI_19898	C8	AT4G14615.1	5.199010866	6.32E-06
NVERN	A_ES969010	NA	NA	5.184812087	6.53E-06
NVERN	C_JCVI_22508	C9	AT5G04560.1	5.181194783	6.59E-06
NVERN	A_BQ704637	NA	AT3G05100.1	5.179716777	6.61E-06
NVERN	A_EV090858	A5	AT1G53350.1	5.177788651	6.64E-06
NVERN	A_JCVI_18832	A2	AT5G05450.1	5.17720532	6.65E-06
NVERN	A_CN729543	A2	AT5G03800.1	5.167751529	6.80E-06

Trait	Unigene code	Chromosome	AGI	-log10P	Pvalue
NVERN	A_JCVI_7320	A2	AT1G74150.1	5.166175395	6.82E-06
NVERN	A_JCVI_35108	A9	AT5G25820.1	5.165988139	6.82E-06
NVERN	C_JCVI_12231	C3	AT4G20130.1	5.159712273	6.92E-06
NVERN	C_EV194419	C5	AT3G16857.2	5.15617827	6.98E-06
NVERN	C_JCVI_826	C5	NA	5.155856792	6.98E-06
NVERN	A_EV201465	A5	AT3G19340.1	5.153170047	7.03E-06
NVERN	A_JCVI_13136	A9	AT5G62810.1	5.150879886	7.07E-06
NVERN	A_JCVI_21360	A9	AT3G55030.1	5.149351294	7.09E-06
NVERN	C_CN729543	C2	AT5G03800.1	5.139085643	7.26E-06
NVERN	C_EV078699	C7	AT5G18460.1	5.138454912	7.27E-06
NVERN	A_EE409569	A5	AT3G07310.1	5.133138136	7.36E-06
NVERN	A_JCVI_4508	A3	AT5G17840.1	5.12924708	7.43E-06
NVERN	A_EE428243	A2	AT5G52410.2	5.125854546	7.48E-06
NVERN	C_JCVI_13109	C9	AT1G02000.1	5.122997555	7.53E-06
NVERN	C_JCVI_7286	C9	AT4G08790.1	5.122888216	7.54E-06
NVERN	A_JCVI_13791	A2	AT5G62140.1	5.116953251	7.64E-06
NVERN	C_ES983643	C2	NA	5.111757751	7.73E-06
NVERN	C_JCVI_20789	C4	AT5G41685.1	5.109657556	7.77E-06
NVERN	C_JCVI_12068	C9	AT4G01850.1	5.108667302	7.79E-06
NVERN	A_JCVI_41378	A6	AT1G17340.1	5.103570967	7.88E-06
NVERN	A_JCVI_31263	A3	NA	5.100554427	7.93E-06
NVERN	C_JCVI_41736	C7	NA	5.096801887	8.00E-06
NVERN	A_JCVI_9111	A5	AT3G06860.1	5.093585214	8.06E-06
NVERN	A_JCVI_6028	A9	AT1G26550.1	5.092792589	8.08E-06
NVERN	C_DY001867	C2	AT5G57580.1	5.092397678	8.08E-06
NVERN	C_EL589622	C2	AT4G02425.1	5.085683547	8.21E-06
NVERN	C_JCVI_9905	C3	AT3G09390.1	5.08253362	8.27E-06
NVERN	A_CX189980	NA	NA	5.078159882	8.35E-06
NVERN	A_ES944788	NA	AT1G17160.1	5.060251251	8.70E-06
NVERN	C_EV091400	C3	AT3G21350.1	5.059725179	8.72E-06
NVERN	A_JCVI_40018	NA	NA	5.058577622	8.74E-06
NVERN	C_EV158802	C4	AT2G26730.1	5.058059603	8.75E-06
NVERN	A_EV226338	A9	AT2G20610.1	5.048833639	8.94E-06
NVERN	A_JCVI_16174	A4	AT3G53630.1	5.045728196	9.00E-06
NVERN	C_EV030996	C5	NA	5.04493914	9.02E-06
NVERN	C_JCVI_6942	C2	AT5G57800.1	5.039063925	9.14E-06
NVERN	C_JCVI_37761	C3	AT4G39230.1	5.038963582	9.14E-06
NVERN	C_JCVI_7448	C8	AT1G14140.1	5.030107356	9.33E-06
NVERN	C_JCVI_11953	C9	AT5G56240.1	5.026918613	9.40E-06
NVERN	C_JCVI_28988	C6	AT1G72880.2	5.017504197	9.60E-06
NVERN	C_JCVI_9903	C5	AT1G14450.1	5.013998852	9.68E-06
NVERN	A_JCVI_20720	A8	AT4G37210.1	5.013741276	9.69E-06
NVERN	C_JCVI_1939	C5	AT1G18210.1	5.013670127	9.69E-06

Trait	Unigene code	Chromosome	AGI	-log10P	Pvalue
NVERN	A_JCVI_31731	A1	AT1G03340.1	5.01241683	9.72E-06
NVERN	A_EV028504	NA	AT4G26630.1	5.011313791	9.74E-06
NVERN	C_EX090988	C9	AT1G60900.1	5.008899694	9.80E-06
NVERN	C_JCVI_31263	C4	NA	5.008063334	9.82E-06
NVERN	A_JCVI_34141	A7	AT2G17410.1	5.000895957	9.98E-06
NVERN	C_EX095959	C8	AT2G25280.1	5.000280828	9.99E-06

Table S. 2: The most significant GEMs that fall above the Bonferroni threshold of association for flowering time VERN dataset. Listed in the table are the Unigene code indicating whether it is located on the A or C genome, the chromosomal location, and the *A. thaliana* gene code (AGI) that exhibits the highest homology. Significance levels are listed as -log10P and P-values.

Trait	Unigene code	Chromosome	AGI	log10P	Pvalue
VERN	A_EV202339	NA	AT1G63140.2	10.6098	2.46E-11
VERN	C_JCVI_17192	C1	AT4G21900.1	10.3112	4.88E-11
VERN	A_JCVI_17192	A1	AT4G21900.1	9.82504	1.50E-10
VERN	C_JCVI_37887	C1	AT4G21900.1	9.24005	5.75E-10
VERN	C_JCVI_4830	C2	AT1G67090.1	8.87659	1.33E-09
VERN	C_JCVI_4004	C2	AT5G02780.1	8.72384	1.89E-09
VERN	A_JCVI_27166	A4	AT3G59790.1	8.70706	1.96E-09
VERN	A_DY009335	A7	NA	8.70016	1.99E-09
VERN	A_JCVI_23190	A2	AT5G01530.1	8.57268	2.68E-09
VERN	A_JCVI_37887	A1	AT4G21900.1	8.49967	3.16E-09
VERN	A_JCVI_26687	A8	AT1G54115.1	8.34372	4.53E-09
VERN	C_DY009335	C5	NA	8.12679	7.47E-09
VERN	A_EX114883	A9	AT2G31530.1	7.89869	1.26E-08
VERN	C_JCVI_12983	C6	AT1G45145.1	7.79175	1.62E-08
VERN	A_EX087749	A2	AT5G45820.1	7.69815	2.00E-08
VERN	A_JCVI_14857	A5	AT3G12345.1	7.59588	2.54E-08
VERN	C_JCVI_16168	C8	AT1G13930.1	7.51552	3.05E-08
VERN	A_EV197623	A3	AT5G60900.1	7.4966	3.19E-08
VERN	C_JCVI_484	NA	NA	7.41218	3.87E-08
VERN	A_JCVI_25616	A9	AT3G25110.1	7.33881	4.58E-08
VERN	A_JCVI_30940	NA	AT1G14720.1	6.9767	1.06E-07
VERN	A_JCVI_26660	A9	AT5G61510.1	6.93955	1.15E-07
VERN	A_JCVI_2275	A10	AT5G11280.1	6.82901	1.48E-07
VERN	C_EV197623	C7	AT5G60900.1	6.73507	1.84E-07
VERN	A_ES967922	A4	NA	6.64689	2.25E-07
VERN	A_JCVI_22584	A10	AT5G10140.1	6.59217	2.56E-07
VERN	A_JCVI_35247	A4	NA	6.58995	2.57E-07
VERN	A_JCVI_11158	A8	AT4G33550.2	6.56011	2.75E-07
VERN	A_EX123292	A3	AT4G15470.1	6.55022	2.82E-07
VERN	C_EV196305	C6	NA	6.54187	2.87E-07
VERN	A_JCVI_23428	A6	AT3G49110.1	6.52404	2.99E-07
VERN	C_EV101010	C3	AT3G04160.1	6.4533	3.52E-07
VERN	C_JCVI_15948	C2	AT5G10630.1	6.4117	3.88E-07
VERN	A_JCVI_30618	A7	NA	6.25246	5.59E-07
VERN	C_JCVI_25968	C5	AT3G01670.1	6.17925	6.62E-07
VERN	A_JCVI_33168	A9	AT3G28180.1	6.17781	6.64E-07
VERN	C_JCVI_30618	C3	NA	6.1299	7.41E-07
VERN	C_DY009791	C2	NA	6.12678	7.47E-07

Trait	Unigene code	Chromosome	AGI	log10P	Pvalue
VERN	C_JCVI_1083	C6	NA	6.10276	7.89E-07
VERN	C_JCVI_2165	C5	AT3G01480.1	6.04146	9.09E-07
VERN	C_JCVI_12972	C3	AT4G19830.1	6.01297	9.71E-07
VERN	A_JCVI_3527	A9	AT1G14030.1	5.98634	1.03E-06
VERN	C_EV196428	C2	NA	5.98495	1.04E-06
VERN	C_JCVI_27327	C1	NA	5.95469	1.11E-06
VERN	A_JCVI_11914	A2	AT1G75380.1	5.94307	1.14E-06
VERN	A_JCVI_18088	A8	AT4G34460.1	5.92023	1.20E-06
VERN	A_JCVI_26544	A7	AT1G74650.1	5.8875	1.30E-06
VERN	A_JCVI_30849	A8	NA	5.86924	1.35E-06
VERN	C_EV201465	C5	AT3G19340.1	5.81887	1.52E-06
VERN	A_EV201595	A2	AT5G45380.1	5.8023	1.58E-06
VERN	A_JCVI_27327	A2	NA	5.77816	1.67E-06
VERN	A_JCVI_17622	A7	AT1G22450.1	5.76892	1.70E-06
VERN	A_JCVI_121	A5	AT2G30490.1	5.73255	1.85E-06
VERN	C_ES937988	C6	AT1G71870.1	5.70625	1.97E-06
VERN	C_EV016874	C4	NA	5.63601	2.31E-06
VERN	A_JCVI_19315	A1	AT4G24930.1	5.59551	2.54E-06
VERN	C_EV078699	C7	AT5G18460.1	5.57447	2.66E-06
VERN	C_EV098831	C1	AT3G16140.1	5.56866	2.70E-06
VERN	A_JCVI_20587	A9	AT1G14710.1	5.54134	2.88E-06
VERN	C_JCVI_8129	C4	AT2G30670.1	5.50974	3.09E-06
VERN	A_JCVI_7462	A10	AT5G02280.1	5.50685	3.11E-06
VERN	A_JCVI_3863	A7	AT1G58270.1	5.49339	3.21E-06
VERN	C_JCVI_27547	C5	NA	5.45573	3.50E-06
VERN	A_JCVI_34086	A10	AT5G02270.1	5.41273	3.87E-06
VERN	A_EV100942	A5	AT2G14080.1	5.40649	3.92E-06
VERN	C_EV197707	C7	AT5G60900.1	5.39732	4.01E-06
VERN	A_JCVI_12400	A9	AT1G12820.1	5.37579	4.21E-06
VERN	C_JCVI_33620	C3	AT5G01910.1	5.37513	4.22E-06
VERN	A_JCVI_10972	A9	AT4G04870.1	5.36563	4.31E-06
VERN	A_EV184965	A8	NA	5.34555	4.51E-06
VERN	C_JCVI_13109	C9	AT1G02000.1	5.34069	4.56E-06
VERN	C_JCVI_41760	C7	NA	5.32013	4.78E-06
VERN	A_EV179210	A4	AT3G52370.1	5.3186	4.80E-06
VERN	A_EE478173	A5	AT3G04970.1	5.30808	4.92E-06
VERN	C_JCVI_29532	C3	AT2G14720.1	5.27261	5.34E-06
VERN	A_JCVI_12972	A8	AT4G19830.1	5.26907	5.38E-06
VERN	C_JCVI_11158	C3	AT4G33550.2	5.25526	5.56E-06
VERN	C_JCVI_23564	C3	AT1G27700.1	5.2542	5.57E-06
VERN	C_JCVI_30351	C6	AT1G69260.1	5.2462	5.67E-06
VERN	C_EE432271	C3	NA	5.22669	5.93E-06
VERN	C_JCVI_14857	C5	AT3G12345.1	5.2235	5.98E-06

Trait	Unigene code	Chromosome	AGI	log10P	Pvalue
VERN	A_JCVI_23438	NA	AT2G05440.3	5.2132	6.12E-06
VERN	C_JCVI_29396	C2	AT5G57580.1	5.20528	6.23E-06
VERN	A_EV204109	A7	AT1G58270.1	5.20336	6.26E-06
VERN	A_AM387806	NA	NA	5.19863	6.33E-06
VERN	C_EE527736	C2	AT4G02500.1	5.19195	6.43E-06
VERN	A_EV196428	A4	NA	5.18773	6.49E-06
VERN	C_JCVI_15549	C9	AT4G00030.1	5.18487	6.53E-06
VERN	C_EV038543	C9	NA	5.18335	6.56E-06
VERN	C_JCVI_9610	C9	AT5G27820.1	5.18241	6.57E-06
VERN	A_JCVI_18520	A3	AT5G20070.1	5.17992	6.61E-06
VERN	A_ES938497	A9	AT5G61390.1	5.1503	7.07E-06
VERN	A_JCVI_28787	A6	AT1G48260.1	5.14357	7.19E-06
VERN	C_JCVI_20620	C4	AT2G38790.1	5.12214	7.55E-06
VERN	A_JCVI_39024	A9	AT1G01630.1	5.10669	7.82E-06
VERN	C_JCVI_16640	C6	AT1G74000.1	5.0979	7.98E-06
VERN	A_JCVI_4141	A8	AT4G37660.1	5.08604	8.20E-06
VERN	C_JCVI_23438	C7	AT2G05440.3	5.08264	8.27E-06
VERN	A_EX018776	NA	AT5G19150.1	5.07676	8.38E-06
VERN	C_JCVI_26041	C3	AT1G54115.1	5.06295	8.65E-06
VERN	A_EX137858	A5	AT2G30490.1	5.0618	8.67E-06
VERN	C_CX189980	C7	NA	5.05247	8.86E-06
VERN	A_JCVI_11734	A2	AT5G60360.1	5.03673	9.19E-06
VERN	A_EV196305	A2	NA	5.0232	9.48E-06
VERN	C_JCVI_26464	C5	AT1G17830.1	5.00906	9.79E-06
VERN	C_EE538308	C3	AT4G21192.1	5.00136	9.97E-06

.

Table S. 3: Summary of sequencing quality data information provided by Novogene Co., Ltd., HK for eight DNA samples sequenced by Illumina ®

Sample	Library	Flowcell/Lane	Raw reads	Raw data (G)	Effective (%)	Error (%)	Q20 (%)	Q30 (%)	GC (%)
Cabriolet (ETNV203)	DSW57009	HFL2JCCXY_L6	121756072	36.5	99.53	0.01	96.65	92.34	37.42
Darmor (ETNV208)	DSW57010	HFL2JCCXY_L7	106033425	31.8	99.62	0.01	96.14	91.35	37.8
VERN EARLY_1 (TD170810373)	DSW50197	H7JCVCCXY_L2	106467543	31.9	99.36	0.01	95.7	91.19	37.97
VERN EARLY_2 (TD170810374)	DSW50198	H7JCVCCXY_L2	113252999	34	99.38	0.01	95.74	91.23	37.98
VERN LATE_3 (TD170810375)	DSW50199	H7JCVCCXY_L3	110030775	33	99.33	0.02	94.95	90.02	38
VERN LATE_4 (TD170810376)	DSW50200	H7JCVCCXY_L3	110166633	33	99.32	0.02	95.07	90.26	37.89
NVERN EARLY_5 (ETNVEP5)	DSW57007	HFL2JCCXY_L7	129129964	38.7	99.53	0.01	96.47	91.93	37.31
NVERN LATE_6 (ETNVLP6)	DSW57008	HFL2JCCXY_L6	132335109	39.7	99.52	0.01	96.91	92.83	37.3

Table S. 4: KASP assay results for F_2 lines for the NVERN treatment. Lines were screened for five SNPs and are grouped in the table according to their flowering time phenotype (early or late flowering) and genotype at each SNP (red=homozygous for Cabriolet, black= homozygous for Darmor, grey = heterozygous). The number of lines per genotype detected are listed.

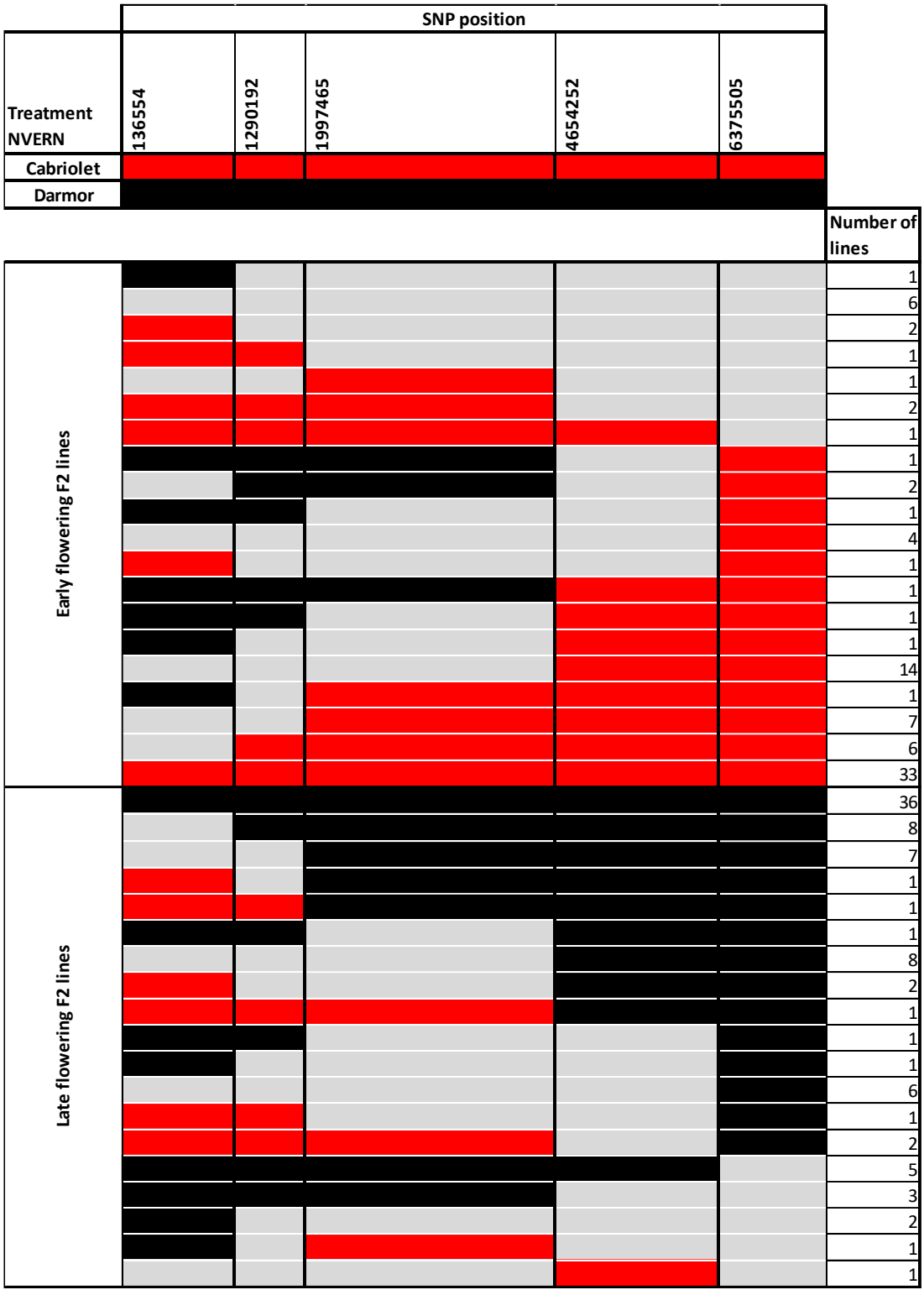
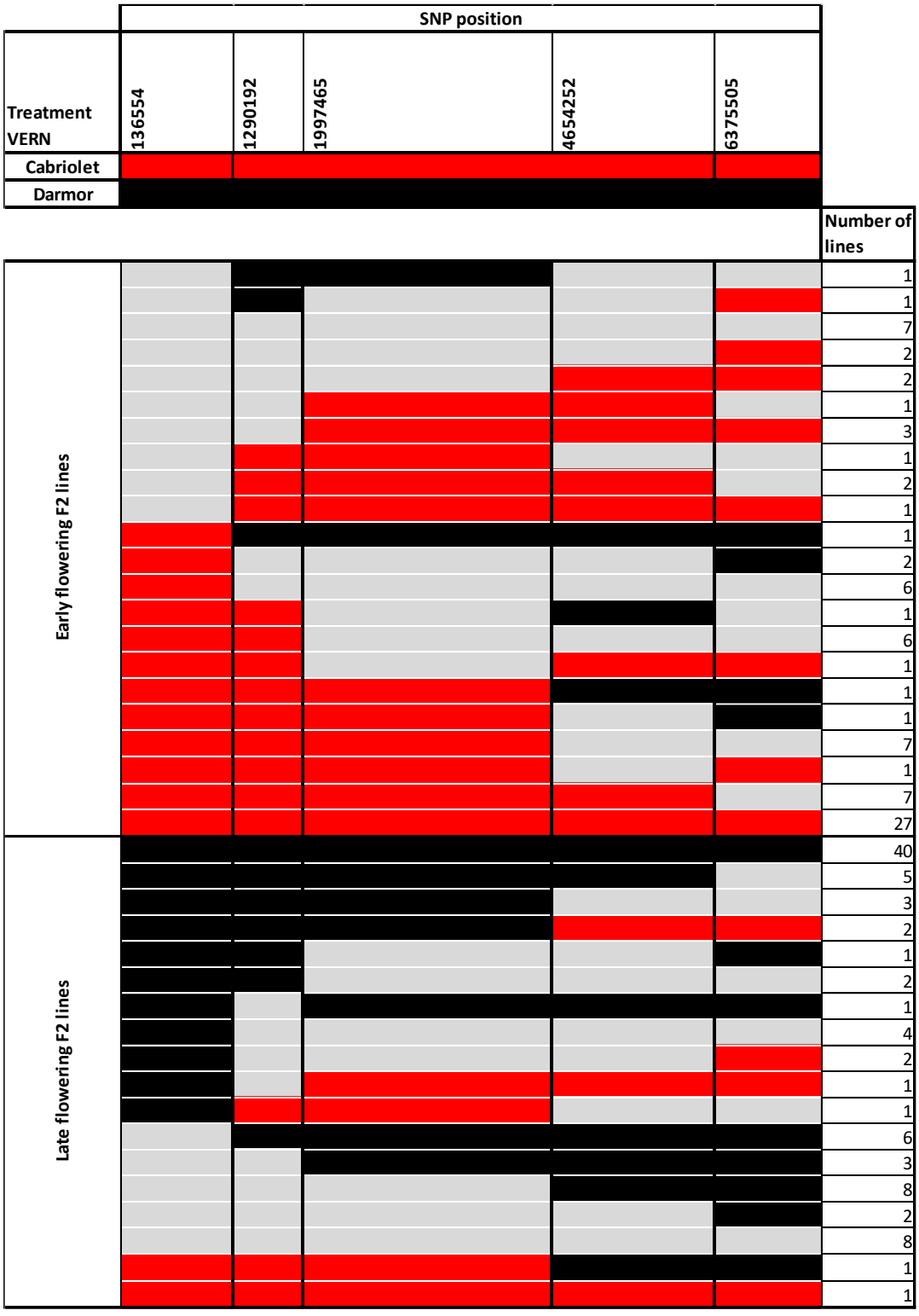


Table S. 5: KASP assay results for F_2 lines for the VERN treatment. Lines were screened for five SNPs and are grouped in the table according to their flowering time phenotype (early or late flowering) and genotype at each SNP (red=homozygous for Cabriolet, black= homozygous for Darmor, grey = heterozygous). The number of lines per genotype detected are listed.



8.3. Supplementary Methods

Edwards DNA Extraction (adapted from Edwards *et al*. 1991)

Reagents

Edwards Extraction Buffer:

200 mM Tris-HCl pH7.5

250 mM NaCl

25 mM EDTA pH8.0

0.5% SDS

Iso-propanol

1. Harvest plant material (whole plant or leaves, can be fresh or frozen) into 1.7ml Eppendorf tube
2. Grind tissue by hand with a blue pestle for approximately 10 seconds
3. Add 400µl of Edwards extraction buffer and grind briefly again with the blue pestle
4. Vortex for 5 seconds
5. Centrifuge at maximum speed for 2 minutes at room temperature
6. Transfer 300µl of the supernatant to a new 1.7ml Eppendorf tube
7. Add 300µl of iso-propanol, invert tubes 5-10 times and leave at room temperature for 2 minutes
8. Centrifuge at maximum speed for 5 minutes at room temperature
9. Remove supernatant and wash the pellet with 70% ethanol
10. Centrifuge at maximum speed for 5 minutes at room temperature
11. Remove the supernatant and air-dry the pellet for 15-30 minutes
12. Resuspend the pellet in 50µl of dH_20

Golden Gate Assembly Protocol: adapted from Engler *et al*., 2009

1. Add all components of the reaction into a 0.2ml PCR tube:
   150ng vector backbone EC00206 (a gift from Dr R. Bloomer)
   50ng of each assembly piece (promoter, ORF1, ORF2, terminator)
   3.0µl 5x T4 buffer (Invitrogen)
   1.5µl 10x CutSmart Buffer (NEB)
   1.0µl BSA-I HF (NEB)
   1.0µl T4 ligase (Invitrogen)
   + DH_20 to a final volume of 15µl
2. Perform an assembly reaction in a thermocycler as follows:
   25 cycles of 37°C for 3 minutes, 16°C for 4 minutes
   1 cycle of 50°C for 5 minutes, 80°C for 5 minutes

Ligation of Level 1 modules to pSLJ-DEST

1. Add all components of the reaction to a 0.2ml PCR tube:
   120ng PSLJ-DEST
   240ng of entry vector assembled by Golden Gate
   1.0µl LR clonase (Invitrogen)
   + TE to a total volume of 10µl
2. Incubate at room temperature overnight
3. Add 1.0µl proteinase K (Sigma Chemical) and incubate at 37°C for 10 minutes
4. Transform into subcloning efficiency DH5α cells (Invitrogen) according to the manufacturer's instructions
5. Grow under selection on LB plates with 10µg/ml tetracycline
6. Positive transformants were confirmed by colony PCR using *BnaFRI.A03* primers N036, A385, N053 and N020

DNA preparation from *Brassica* for Illumina® sequencing

Reagents

Nuclei extraction buffer:
10mM TrisHCL pH9.5
10mM EDTA pH8.0
100mM KCl
500mM Sucrose
4mM Spermidine
1mM Spermine
0.1% 2-mecaptoethanol

Lysis buffer:
10% Triton-X in nuclei extraction buffer (10% by volume)

CTAB extraction buffer:
100mM TrisHCL pH7.5
0.7M NaCl
10mM EDTA pH8.0
0.1% CTAB
1%2-mercaptoethanol

Chloroform/Isoamyl alcohol (24:1)

Phenol/Chloroform/Isoamyl alcohol (24:23:1)

RNase T1 (1000U/µl)

RNase A (10mg/ml)

Proteinase K (15mg/ml)

Steps in Procedure:

Leaf tissue preparation

1. Leaf tissue (1cm leaf discs) were harvested on wet ice and stored at -20°C
2. Using a pestle and mortar, grind approximately 36 leaf discs (~0.1g) to a fine powder in liquid N_2.
3. Turn on bench-top centrifuge to cool to 4°C and water baths to 60°C and 37°C.

Nuclei extraction

4. In a 50ml falcon tube add ground leaf material to 40ml of ice-cold nuclei extraction buffer.
5. Vortex until ground disuse is evenly distributed through the buffer.
6. Filter homogenised tissue through two layers of Miracloth (Calbiochem) in a funnel to a new 50ml tube.

Lysis

7. Add 8ml of lysis buffer to filtered homogenate and keep on ice for 2 minutes.
8. Centrifuge at 1,000x*g* for 20 minutes at 4°C to pellet the nuclei.
9. Remove and discard the supernatant.
10. Resuspend the pellet in 3ml CTAB extraction buffer.
11. Incubate in water bath for 30 minutes at 60°C then transfer all solution to a 15ml falcon tube.

Chloroform/Isoamyl alcohol (24:1) extraction

12. Add equal volumes, roughly 3ml, of chloroform/isoamyl alcohol to the lysed sample.
13. Rotate for 10 minutes in a cold room at 5°C.
14. Centrifuge for 10 minutes at 1,000x*g*.

RNase/Proteinase treatment

15. Remove aqueous phase (usually 2-3ml) and add RNase T1 and RNase A to the sample to final concentrations of 50u/ml and 50µg/ml respectively.
16. Mix sample by inverting the tube 10-20 times.

17. Incubate in a water bath for 45 minutes at 37°C.
18. Add Proteinase K to a final concentration of 150ug/ml.
19. Incubate in a water bath for 45 minutes at 37°C.

Phenol/Chloroform/Isoamyl alcohol (24:23:1) extraction

20. Add equal volume of phenol/chloroform/isoamyl alcohol (24:23:1) to the solution.
21. Invert 20 times to mix.
22. Centrifuge for 10 minutes at 1,000xg.
23. Retain supernatant and repeat steps 20-22.

Chloroform/Isoamyl alcohol (24:1) extraction

24. Add equal volume of chloroform/isoamyl alcohol (24:1) to the solution.
25. Rotate for 10 minutes in a cold room at 5°C.
26. Centrifuge for 10 minutes at 1,000xg.
27. Retain supernatant.

DNA precipitation

28. Add 0.1x volume (usually 200-300ul) of 3M NaAc and 3x volume (usually 6-9ml) of 100% EtOH.
29. Store at -20°C overnight.
30. Centrifuge at 3,000xg for 30 minutes at 4°C.
31. Remove supernatant and wash with 3-5ml of 75% EtOH.
32. Gently centrifuge to collect the DNA.
33. Remove the ethanol and air-dry the DNA.
34. Gently resuspend the DNA in 30µl H_20.

Triparental Mating

Waste disposal

All bio-hazard material must be autoclaved before disposal

Day 1: Start LB liquid culture of *Agrobacterium tumefaciens* strain C58. Take C58 from glycerol stock and incubate in LB with 200µg/ml Rifampicin at 28°C with shaking.

Day 3:

1. Start liquid cultures of the helper plasmid HB101 (pRK2013) taken from glycerol stocks and incubate in LB with 100µg/ml kanamycin at 37°C with shaking. At the same time start liquid cultures of all constructs to be transformed into the Agrobacterium, taking these also from glycerol stocks and incubate in LB with 10µg/ml tetracycline at 37°C shaking. HB101 and constructs must reach mid-log phase (OD600 0.2-0.6) and the C58 culture should reach stationary phase (OD600 1-2) before continuing to the next step
2. To 1.7ml Eppendorf tubes add the following:
   800µl C58
   200µl HB101
   200µl Construct
3. Prepare a negative control. To 1.7ml Eppendorf tubes add the following:
   800µl C58
   200µl HB101
   200µl LB
4. Briefly vortex
5. Centrifuge at 6,000rpm for 2 minutes, then remove all but 100µl of the supernatant
6. Re-suspend cells by pipetting and plate onto LB without selection.
7. Incubate at 28°C overnight

Day 4: Each plate should have an even growth of cells

1. For each construct, plus the negative control, prepare 3 Eppendorf tubes with 1ml MgSO4 (10mM). Label each with 100, 10^{-2}, 10^{-4}
2. Scrape cells from the LB plates with a sterile loop and add to the tube 100 of MgSO4. Vortex.
3. Add 100µl of the cell-MgSO4 mixture from tube 100 to the tube labelled 10-2 and vortex.
4. Add 100µl of the cell-MgSO4 mixture from tube 10-2 to the tube labelled 10-4 and vortex.
5. Pipette 20µl from each tube onto 1/3 sectors of LB plates with 200µg rifampicin and 10µg/ml tetracycline.
6. Allow to dry and incubate at 28°C for 2-3 days. Single colonies will be visible on the 10^{-4} sector on each plate.

Agrobacterium-mediated floral spraying

Waste disposal:

All bio-hazard material must be autoclaved before disposal

Clear lab bench and place autoclave bags over the surface to protect the bench from contamination

Materials:

250ml centrifuge tubes (Corning)

1L tripour pots

1L spray bottles

Autoclavable Biohazard bags (780x600mm)

40x flowering *A. thaliana* plants

Spray solution:

1L sterile water

50g sucrose

200µl SILWET

1. Inoculate 10ml LB with selective antibiotics (200µg/ml rifampicin, 10µg/ml tetracycline) with an *Agrobacterium* strain carrying the construct to be transformed into *A. thaliana*. Take cells from a glycerol stock and incubate in LB at 28°C for 36-48 hours.
2. Use 5ml of this liquid culture to inoculate 500ml of LB with selective antibiotics (200µg/ml rifampicin, 10µg/ml tetracycline) in a 2L flask and incubate at 28°C until the culture reaches OD600 ~2.0-4.0

3. Divide the 500ml between two centrifuge tubes (Corning) and centrifuge at 3,500rpm for 15 minutes at 15°C in a GSA rotor. Be careful to balance the centrifuge tubes.
4. Discard the supernatant and add a small volume of spray solution. Swirl to rinse the pellet and discard.
5. Re-suspend the pellet in ~100ml of spray solution. Pour this into a 1L tripour pot and dilute to OD600 ~1.0. Transfer this to 1L spray bottles.
6. In containment glasshouse, place the tray of *A. thaliana* plants that are to be sprayed in a biohazard bag and spray the flowers liberally with spray solution.
7. Close the biohazard bag to cover the plants, secure with tape, and leave overnight. Remove plants from the bag the next day.
8. Repeat steps 1-7 so that plants are sprayed once more in seven days.

References

Abe M, Kobayashi Y, Yamamoto S, Daimon Y, Yamaguchi A, Ikeda Y, Ichinoki H, Notaguchi M, Goto K, Araki T, 2005. *FD*, a bZIP protein mediating signals from the floral pathway integrator *FT* at the shoot apex. *Science* **309**, 1052-6.

Adams KL, Wendel JF, 2005. Polyploidy and genome evolution in plants. *Current Opinion in Plant Biology* **8**, 135-41.

Adrian J, Farrona S, Reimer JJ, Albani MC, Coupland G, Turck F, 2010. cis-Regulatory elements and chromatin state coordinately control temporal and spatial expression of *FLOWERING LOCUS T* in *Arabidopsis*. *The Plant Cell* **22**, 1425-40.

Ågren J, Oakley CG, Mckay JK, Lovell JT, Schemske DW, 2013. Genetic mapping of adaptation reveals fitness tradeoffs in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences* **110**, 21077-82.

Aguilar-Martínez JA, Poza-Carrión C, Cubas P, 2007. *Arabidopsis BRANCHED1* acts as an integrator of branching signals within axillary buds. *The Plant Cell* **19**, 458-72.

Ahn JH, Miller D, Winter VJ, Banfield MJ, Lee JH, Yoo SY, Henz SR, Brady RL, Weigel D, 2006. A divergent external loop confers antagonistic activity on floral regulators *FT* and *TFL1*. *The EMBO journal* **25**, 605-14.

Aikawa S, Kobayashi MJ, Satake A, Shimizu KK, Kudoh H, 2010. Robust control of the seasonal expression of the *Arabidopsis FLC* gene in a fluctuating environment. *Proceedings of the National Academy of Sciences* **107**, 11632-7.

Ajisaka H, Kuginuki Y, Yui S, Enomoto S, Hirai M, 2001. Identification and mapping of a quantitative trait locus controlling extreme late bolting in Chinese

cabbage (*Brassica rapa* L. ssp. *pekinensis* syn. *campestris* L.) using bulked segregant analysis. *Euphytica* **118**, 75.

Albani MC, Castaings L, Wötzel S, Mateos JL, Wunder J, Wang R, Reymond M, Coupland G, 2012. *PEP1* of *Arabis alpina* is encoded by two overlapping genes that contribute to natural genetic variation in perennial flowering. *PLoS Genetics* **8**, e1003130.

Angel A, Song J, Dean C, Howard M, 2011. A Polycomb-based switch underlying quantitative epigenetic memory. *Nature* **476**, 105-8.

Angel A, Song J, Yang H, Questa JI, Dean C, Howard M, 2015. Vernalizing cold is registered digitally at *FLC*. *Proc Natl Acad Sci U S A* **112**, 4146-51.

Arabidopsis Genome Initiative, 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796.

Arias T, Beilstein MA, Tang M, Mckain MR, Pires JC, 2014. Diversification times among *Brassica* (Brassicaceae) crops suggest hybrid formation after 20 million years of divergence. *American Journal of Botany* **101**, 86-91.

Atwell S, Huang YS, Vilhjálmsson BJ, Willems G, Horton M, Li Y, Meng D, Platt A, Tarone AM, Hu TT, Jiang R, 2010. Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature* **465**, 627-31.

Aukerman MJ, Lee I, Weigel D, Amasino RM, 1999. The *Arabidopsis* flowering-time gene *LUMINIDEPENDENS* is expressed primarily in regions of cell proliferation and encodes a nuclear protein that regulates *LEAFY* expression. *The Plant Journal* **18**, 195-203.

Ausín I, Alonso-Blanco C, Jarillo JA, Ruiz-García L, Martínez-Zapater JM, 2004. Regulation of flowering time by *FVE*, a retinoblastoma-associated protein. *Nature Genetics* **36**, 162-6.

Baduel P, Arnold B, Weisman CM, Hunter B, Bomblies K, 2016. Habitat-Associated Life History and Stress-Tolerance Variation in *Arabidopsis arenosa*. *Plant Physiology* **171**, 437-51.

Bancroft I, Morgan C, Fraser F, Higgins J, Wells R, Clissold L, Baker D, Long Y, Meng J, Wang X, Liu S, 2011. Dissecting the genome of the polyploid crop oilseed rape by transcriptome sequencing. *Nature Biotechnology* **29**, 762-6.

Beilstein MA, Nagalingum NS, Clements MD, Manchester SR, Mathews S, 2010. Dated molecular phylogenies indicate a Miocene origin for *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences* **107**, 18724-8.

Ben-Naim O, Eshed R, Parnis A, Teper-Bamnolker P, Shalit A, Coupland G, Samach A, Lifschitz E, 2006. The CCAAT binding factor can mediate interactions between CONSTANS-like proteins and DNA. *The Plant Journal* **46**, 462-76.

Berry S, Hartley M, Olsson TS, Dean C, Howard M, 2015. Local chromatin environment of a Polycomb target gene instructs its own epigenetic inheritance. *Elife* **4**, e07205.

Blümel M, Dally N, Jung C, 2015. Flowering time regulation in crops-what did we learn from *Arabidopsis*? *Current Opinion in Biotechnology* **32**, 121-9.

Bowman JL, Alvarez J, Weigel D, Meyerowitz EM, Smyth DR, 1993. Control of flower development in *Arabidopsis thaliana* by *APETALA1* and interacting genes. *Development* **119**, 721-43.

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES, 2007. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**, 2633-5.

Brown PT, Caldeira K, 2017. Greater future global warming inferred from Earth's recent energy budget. *Nature* **552**, 45.

Burn J, Bagnall D, Metzger J, Dennis E, Peacock W, 1993. DNA methylation, vernalization, and the initiation of flowering. *Proceedings of the National Academy of Sciences* **90**, 287-91.

Camargo LE, Osborn TC, 1996. Mapping loci controlling flowering time in *Brassica oleracea*. *Theoretical and Applied Genetics* **92**, 610-6.

Carré P, Pouzet A, 2014. Rapeseed market, worldwide and in Europe. *OCL* **21**, D102.

Chalhoub B, Denoeud F, Liu S, Parkin IA, Tang H, Wang X, Chiquet J, Belcram H, Tong C, Samans B, Corréa M, *et al*., 2014. Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science* **345**, 950-3.

Chen B, Xu K, Li J, Li F, Qiao J, Li H, Gao G, Yan G, Wu X, 2014. Evaluation of yield and agronomic traits and their genetic variation in 488 global collections of *Brassica napus* L. *Genetic Resources and Crop Evolution* **61**, 979-99.

Cheng F, Wu J, Fang L, Sun S, Liu B, Lin K, Bonnema G, Wang X, 2012. Biased gene fractionation and dominant gene expression among the subgenomes of *Brassica rapa*. *PLoS One* **7**, e36442.

Chiang GC, Barua D, Kramer EM, Amasino RM, Donohue K, 2009. Major flowering time gene, *FLOWERING LOCUS C*, regulates seed germination in *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A* **106**, 11661-6.

Choi K, Kim J, Hwang HJ, Kim S, Park C, Kim SY, Lee I, 2011. The FRIGIDA complex activates transcription of *FLC*, a strong flowering repressor in *Arabidopsis*, by recruiting chromatin modification factors. *The Plant Cell Online* **23**, 289-303.

Chouard P, 1960. Vernalization and its relations to dormancy. *Annual Review of Plant Physiology* **11**, 191-238.

Clarke JH, Dean C, 1994. Mapping *FRI*, a locus controlling flowering time and vernalization response in *Arabidopsis thaliana*. *Molecular and General Genetics MGG* **242**, 81-9.

Clough SJ, Bent AF, 1998. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal* **16**, 735-43.

Conant GC, Birchler JA, Pires JC, 2014. Dosage, duplication, and diploidization: clarifying the interplay of multiple models for duplicate gene evolution over time. *Current Opinion in Plant Biology* **19**, 91-8.

Corbesier L, Vincent C, Jang S, Fornara F, Fan Q, Searle I, Giakountis A, Farrona S, Gissot L, Turnbull C, Coupland G, 2007. FT protein movement contributes to long-distance signaling in floral induction of *Arabidopsis*. *Science* **316**, 1030-3.

Coustham V, Li P, Strange A, Lister C, Song J, Dean C, 2012. Quantitative modulation of polycomb silencing underlies natural variation in vernalization. *Science* **337**, 584-7.

Craufurd PQ, Wheeler TR, 2009. Climate change and the flowering time of annual crops. *Journal of Experimental Botany* **60**, 2529-39.

Csorba T, Questa JI, Sun Q, Dean C, 2014. Antisense *COOLAIR* mediates the coordinated switching of chromatin states at *FLC* during vernalization. *Proc Natl Acad Sci U S A* **111**, 16160-5.

De Lucia F, Crevillen P, Jones AM, Greb T, Dean C, 2008. A PHD-polycomb repressive complex 2 triggers the epigenetic silencing of *FLC* during vernalization. *Proc Natl Acad Sci U S A* **105**, 16831-6.

Deng W, Ying H, Helliwell CA, Taylor JM, Peacock WJ, Dennis ES, 2011. *FLOWERING LOCUS C* (*FLC*) regulates development pathways throughout the life cycle of *Arabidopsis*. *Proceedings of the National Academy of Sciences* **108**, 6680-5.

Drouaud J, Camilleri C, Bourguignon PY, Canaguier A, Bérard A, Vezon D, Giancola S, Brunel D, Colot V, Prum B, Quesneville H, 2006. Variation in crossing-over rates across chromosome 4 of *Arabidopsis thaliana* reveals the presence of meiotic recombination "hot spots". *Genome Research* **16**, 106-14.

Duncan S, Holm S, Questa J, Irwin J, Grant A, Dean C, 2015. Seasonal shift in timing of vernalization as an adaptation to extreme winter. *Elife* **4**, e09920.

Edwards K, Johnstone C, Thompson C, 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Research* **19**, 1349.

Engler C, Gruetzner R, Kandzia R, Marillonnet S, 2009. Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. *PLoS One* **4**, e5553.

Evanno G, Regnaut S, Goudet J, 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611-20.

Fadina O, Khavkin E, 2014. The vernalization gene *FRIGIDA* in cultivated *Brassica* species. *Russian Journal of Plant Physiology* **61**, 309-17.

Fadina O, Pankin A, Khavkin E, 2013. Molecular characterization of the flowering time gene *FRIGIDA* in *Brassica* genomes A and C. *Russian Journal of Plant Physiology* **60**, 279-89.

Ferguson AA, Jiang N, 2011. Pack-MULEs: recycling and reshaping genes through GC-biased acquisition. Mobile genetic elements **1**, 135-138.

Ferrándiz C, Gu Q, Martienssen R, Yanofsky MF, 2000. Redundant regulation of meristem identity and plant architecture by *FRUITFULL*, *APETALA1* and *CAULIFLOWER*. *Development* **127**, 725-34.

Ferreira ME, Satagopan J, Yandell BS, Williams PH, Osborn TC, 1995. Mapping loci controlling vernalization requirement and flowering time in *Brassica napus*. *Theoretical and Applied Genetics* **90**, 727-32.

Finnegan EJ, Dennis ES, 2007. Vernalization-induced trimethylation of histone H3 lysine 27 at *FLC* is not maintained in mitotically quiescent cells. *Current Biology* **17**, 1978-83.

Fletcher RS, Mullen JL, Heiliger A, Mckay JK, 2014. QTL analysis of root morphology, flowering time, and yield reveals trade-offs in response to drought in *Brassica napus*. *Journal of Experimental Botany* **66**, 245-56.

Foisset N, Delourme R, Barret P, Renard M, 1995. Molecular tagging of the dwarf *BREIZH* (*Bzh*) gene in *Brassica napus*. *Theoretical and Applied Genetics* **91**, 756-61.

Fransz PF, Armstrong S, de Jong JH, Parnell LD, van Drunen C, Dean C, Zabel P, Bisseling T, Jones, GH, 2000. Integrated cytogenetic map of chromosome arm 4S of *A. thaliana*: structural organization of heterochromatic knob and centromere region. *Cell* **100**, 367-76.

Geraldo N, Bäurle I, Kidou S-I, Hu X, Dean C, 2009. *FRIGIDA* delays flowering in *Arabidopsis* via a cotranscriptional mechanism involving direct interaction with the nuclear cap-binding complex. *Plant Physiology* **150**, 1611-8.

Ginzinger DG, 2002. Gene quantification using real-time quantitative PCR: an emerging technology hits the mainstream. *Experimental hematology* **30**, 503-12.

Giovannoni JJ, Wing RA, Ganal MW, Tanksley SD, 1991. Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping populations. *Nucleic Acids Research* **19**, 6553-68.

Grbic V, Bleecker AB, 1996. An altered body plan is conferred on *Arabidopsis* plants carrying dominant alleles of two genes. *Development* **122**, 2395-403.

Grbic V, Bleecker AB, 2000. Axillary meristem development in *Arabidopsis thaliana*. *The Plant Journal* **21**, 215-23.

Greb T, Clarenz O, Schäfer E, Müller D, Herrero R, Schmitz G, Theres K, 2003. Molecular analysis of the *LATERAL SUPPRESSOR* gene in *Arabidopsis* reveals a conserved control mechanism for axillary meristem formation. *Genes & Development* **17**, 1175-87.

Greb T, Mylne JS, Crevillen P, Geraldo N, An H, Gendall AR, Dean C, 2007. The PHD finger protein VRN5 functions in the epigenetic silencing of *Arabidopsis FLC*. *Current Biology* **17**, 73-8.

Grillo MA, Li C, Hammond M, Wang L, Schemske DW, 2013. Genetic architecture of flowering time differentiation between locally adapted populations of *Arabidopsis thaliana*. *New Phytologist* **197**, 1321-31.

Guo Y, Hans H, Christian J, Molina C, 2014. Mutations in single *FT*- and *TFL1*-paralogs of rapeseed (*Brassica napus* L.) and their impact on flowering time and yield components. *Front Plant Sci* **5**, 282.

Guo Y-L, Todesco M, Hagmann J, Das S, Weigel D, 2012. Independent *FLC* mutations as causes of flowering-time variation in *Arabidopsis thaliana* and *Capsella rubella*. *Genetics* **192**, 729-39.

Hanzawa Y, Money T, Bradley D, 2005. A single amino acid converts a repressor to an activator of flowering. *Proceedings of the National Academy of Sciences* **102**, 7748-53.

Harper AL, McKinney LV, Nielsen LR, Havlickova L, Li Y, Trick M, Fraser F, Wang L, Fellgett A, Sollars ES, Janacek SH, 2016. Molecular markers for tolerance of European ash (Fraxinus excelsior) to dieback disease identified using Associative Transcriptomics. *Scientific reports* **6**, 19335.

Havlickova L, He Z, Wang L, Langer S, Harper AL, Kaur H, Broadley MR, Gegas V, Bancroft I, 2018. Validation of an updated Associative Transcriptomics platform for the polyploid crop species *Brassica napus* by dissection of the genetic architecture of erucic acid and tocopherol isoform variation in seeds. *The Plant Journal* **93**, 181-92.

Hawkes EJ, 2017. Conservation and function of *COOLAIR* long non-coding RNAs in *Brassica* flowering time control. *University of East Anglia PhD Thesis*

Hawkes E, Hennelly S, Novikova I, Irwin J, Dean C, Sanbonmatsu K, 2016. *COOLAIR* Antisense RNAs Form Evolutionarily Conserved Elaborate Secondary Structures. *Cell Rep* **16**, 3087-96.

He Y, Doyle MR, Amasino RM, 2004. PAF1-complex-mediated histone methylation of *FLOWERING LOCUS C* chromatin is required for the vernalization-responsive, winter-annual habit in *Arabidopsis*. *Genes & Development* **18**, 2774-84.

He Z, Wang L, Harper AL, Havlickova L, Pradhan AK, Parkin IA, Bancroft I, 2017. Extensive homoeologous genome exchanges in allopolyploid crops revealed by mRNAseq-based visualization. *Plant Biotechnology Journal* **15**, 594-604.

Helliwell CA, Wood CC, Robertson M, James Peacock W, Dennis ES, 2006. The *Arabidopsis* FLC protein interacts directly in vivo with *SOC1* and *FT* chromatin and is part of a high-molecular-weight protein complex. *Plant Journal* **46**, 183-92.

Hempel FD, Feldman LJ, 1994. Bi-directional inflorescence development in *Arabidopsis thaliana*: Acropetal initiation of flowers and basipetal initiation of paraclades. *Planta* **192**, 276-86.

Hempel FD, Zambryski PC, Feldman LJ, 1998. Photoinduction of flower identity in vegetatively biased primordia. *The Plant Cell* **10**, 1663-75.

Hepworth J, Dean C, 2015. Flowering Locus C's lessons: conserved chromatin switches underpinning developmental timing and adaptation. *Plant Physiology* **168**, 1237-45.

Hepworth SR, Valverde F, Ravenscroft D, Mouradov A, Coupland G, 2002. Antagonistic regulation of flowering-time gene *SOC1* by *CONSTANS* and *FLC* via separate promoter motifs. *The EMBO journal* **21**, 4327-37.

Hong JK, Kim S-Y, Kim JS, Kim JA, Park B-S, Lee Y-H, 2011. Promoters of three *Brassica rapa FLOWERING LOCUS C* differentially regulate gene expression during growth and development in *Arabidopsis*. *Genes & Genomics* **33**, 75-82.

Hou J, Long Y, Raman H, Zou X, Wang J, Dai S, Xiao Q, Li C, Fan L, Liu B, Meng J, 2012. A Tourist-like MITE insertion in the upstream region of the *BnFLC. A10* gene is associated with vernalization requirement in rapeseed (*Brassica napus* L.). *BMC Plant Biology* **12**, 238.

Huang J, Pray C, Rozelle S, 2002. Enhancing the crops to feed the poor. *Nature* **418**, 678.

Huang X, Ding J, Effgen S, Turck F, Koornneef M, 2013. Multiple loci and genetic interactions involving flowering time genes regulate stem branching among natural variants of *Arabidopsis*. *New Phytologist* **199**, 843-57.

Huang X, Paulo MJ, Boer M, Effgen S, Keizer P, Koornneef M, van Eeuwijk FA, 2011. Analysis of natural allelic variation in *Arabidopsis* using a multiparent recombinant inbred line population. *Proceedings of the National Academy of Sciences* **108**, 4488-93.

Illa-Berenguer E, Van Houten J, Huang Z, Van Der Knaap E, 2015. Rapid and reliable identification of tomato fruit weight and locule number loci by QTL-seq. *Theoretical and Applied Genetics* **128**, 1329-42.

Irwin JA, Lister C, Soumpourou E, Zhang Y, Howell EC, Teakle G, Dean C, 2012. Functional alleles of the flowering time regulator *FRIGIDA* in the *Brassica oleracea* genome. *BMC Plant Biology* **12**, 21.

Irwin JA, Soumpourou E, Lister C, Ligthart JD, Kennedy S, Dean C, 2016. Nucleotide polymorphism affecting *FLC* expression underpins heading date variation in horticultural brassicas. *Plant Journal* **87**, 597-605.

Jaeger KE, Graf A, Wigge PA, 2006. The control of flowering in time and space. *Journal of Experimental Botany* **57**, 3415-8.

Jaudal M, Yeoh CC, Zhang L, Stockum C, Mysore KS, Ratet P, Putterill J, 2013. Retroelement insertions at the *Medicago FTa1* locus in spring mutants eliminate vernalisation but not long-day requirements for early flowering. *The Plant Journal* **76**, 580-91.

Johanson U, West J, Lister C, Michaels S, Amasino R, Dean C, 2000. Molecular analysis of *FRIGIDA*, a major determinant of natural variation in *Arabidopsis* flowering time. *Science* **290**, 344-7.

Jones DM, Wells R, Pullen N, Trick M, Irwin JA, Morris RJ, 2018. Spatio-temporal expression dynamics differ between homologues of flowering time genes in the allopolyploid *Brassica napus*. *The Plant Journal* **96**, 103-18.

Jung C, Müller AE, 2009. Flowering time control and applications in plant breeding. *Trends in Plant Science* **14**, 563-73.

Kagale S, Robinson SJ, Nixon J, Xiao R, Huebert T, Condie J, Kessler D, Clarke WE, Edger PP, Links MG, Sharpe AG, 2014. Polyploid evolution of the Brassicaceae during the Cenozoic era. *The Plant Cell* **26**, 2777-91.

Kalinina A, Mihajlović N, Grbić V, 2002. Axillary meristem development in the branchless Zu-0 ecotype of *Arabidopsis thaliana*. *Planta* **215**, 699-707.

Kemi U, Niittyvuopio A, Toivainen T, Pasanen A, Quilot-Turion B, Holm K, Lagercrantz U, Savolainen O, Kuittinen H, 2013. Role of vernalization and of duplicated *FLOWERING LOCUS C* in the perennial *Arabidopsis lyrata*. *New Phytologist* **197**, 323-35.

Kim SY, Park BS, Kwon SJ, Kim J, Lim MH, Park YD, Kim DY, Suh SC, Jin YM, Ahn JH, Lee YH, 2007. Delayed flowering time in *Arabidopsis* and *Brassica rapa* by the overexpression of *FLOWERING LOCUS C* (*FLC*) homologs isolated from Chinese cabbage (*Brassica rapa* L.: ssp. *pekinensis*). *Plant Cell Reports* **26**, 327-36.

Kole C, Quijada P, Michaels S, Amasino R, Osborn T, 2001. Evidence for homology of flowering-time genes *VFR2* from *Brassica rapa* and *FLC* from *Arabidopsis thaliana*. *Theoretical and Applied Genetics* **102**, 425-30.

Koornneef M, Alonso-Blanco C, Blankestijn-De Vries H, Hanhart C, Peeters A, 1998. Genetic interactions among late-flowering mutants of *Arabidopsis*. *Genetics* **148**, 885-92.

Koornneef M, Blankestijn-De Vries H, Hanhart C, Soppe W, Peeters T, 1994. The phenotype of some late-flowering mutants is enhanced by a locus on chromosome 5 that is not effective in the Landsberg *erecta* wild-type. *The Plant Journal* **6**, 911-9.

Koornneef M, Hanhart C, Van Der Veen J, 1991. A genetic and physiological analysis of late flowering mutants in *Arabidopsis thaliana*. *Molecular and General Genetics MGG* **229**, 57-66.

Koprivova A, Harper AL, Trick M, Bancroft I, Kopriva S, 2014. Dissection of control of anion homeostasis by associative transcriptomics in *Brassica napus*. *Plant Physiology* **166**, 442-50.

Korves TM, Schmid KJ, Caicedo AL, Mays C, Stinchcombe JR, Purugganan MD, Schmitt J, 2007. Fitness effects associated with the major flowering time gene *FRIGIDA* in *Arabidopsis thaliana* in the field. *The American Naturalist* **169**, E141-E57.

Kuittinen H, Niittyvuopio A, Rinne P, Savolainen O, 2008. Natural variation in *Arabidopsis lyrata* vernalization requirement conferred by a *FRIGIDA* indel polymorphism. *Molecular Biology and Evolution* **25**, 319-29.

Kumar G, Arya P, Gupta K, Randhawa V, Acharya V, Singh AK, 2016. Comparative phylogenetic analysis and transcriptional profiling of MADS-box gene family identified *DAM* and *FLC*-like genes in apple (*Malusx domestica*). *Scientific reports* **6**, 20695.

Langmead B, Trapnell C, Pop M, Salzberg SL, 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* **10**, R25.

Langridge J, 1957. Effect of day-length and gibberellic acid on the flowering of *Arabidopsis*. *Nature* **180**, 36.

Lazaro A, Obeng-Hinneh E, Albani MC, 2018. Extended vernalization regulates inflorescence fate in *Arabis alpina* by stably silencing *PERPETUAL FLOWERING 1*. *Plant Physiology* **176**, 2819-33.

Le Corre V, Roux F, Reboud X, 2002. DNA polymorphism at the *FRIGIDA* gene in *Arabidopsis thaliana*: extensive nonsynonymous variation is consistent with local selection for flowering time. *Molecular Biology and Evolution* **19**, 1261-71.

Lee I, Amasino RM, 1995. Effect of vernalization, photoperiod, and light quality on the flowering phenotype of *Arabidopsis* plants containing the *FRIGIDA* gene. *Plant Physiology* **108**, 157-62.

Lee I, Aukerman MJ, Gore SL, Lohman KN, Michaels SD, Weaver LM, John MC, Feldmann KA, Amasino RM, 1994a. Isolation of *LUMINIDEPENDENS*: a gene involved in the control of flowering time in *Arabidopsis*. *The Plant Cell* **6**, 75-83.

Lee I, Michaels SD, Masshardt AS, Amasino RM, 1994b. The late-flowering phenotype of *FRIGIDA* and mutations in *LUMINIDEPENDENS* is suppressed in the Landsberg *erecta* strain of *Arabidopsis*. *The Plant Journal* **6**, 903-9.

Lee J, Oh M, Park H, Lee I, 2008. SOC1 translocated to the nucleus by interaction with AGL24 directly regulates *LEAFY*. *The Plant Journal* **55**, 832-43.

Li B, Carey M, Workman JL, 2007. The role of chromatin during transcription. *Cell* **128**, 707-19.

Li D, Liu C, Shen L, Wu Y, Chen H, Robertson M, Helliwell CA, Ito T, Meyerowitz E, Yu H, 2008. A repressor complex governs the integration of flowering signals in *Arabidopsis*. *Developmental Cell* **15**, 110-20.

Li F, Chen B, Xu K, Gao G, Yan G, Qiao J, Li J, Li H, Li L, Xiao X, Zhang T, 2016. A genome-wide association study of plant height and primary branch number in rapeseed (*Brassica napus*). *Plant science* **242**, 169-77.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078-9.

Li P, Filiault D, Box MS, Kerdaffrec E, van Oosterhout C, Wilczek AM, Schmitt J, McMullan M, Bergelson J, Nordborg M, Dean C, Multiple *FLC* haplotypes defined by independent cis-regulatory variation underpin life history diversity in *Arabidopsis thaliana*. *Genes & Development* **28**, 1635-40.

Li P, Tao Z, Dean C, 2015. Phenotypic evolution through variation in splicing of the noncoding RNA *COOLAIR*. *Genes & Development* **29**, 696-701.

Li X, Zhang S, Bai J, He Y, 2016. Tuning growth cycles of *Brassica* crops via natural antisense transcripts of *BrFLC*. *Plant Biotechnology Journal* **14**, 905-14.

Lin S-I, Wang J-G, Poon S-Y, Su C-L, Wang S-S, Chiou T-J, 2005. Differential regulation of *FLOWERING LOCUS C* expression by vernalization in cabbage and *Arabidopsis*. *Plant Physiology* **137**, 1037-48.

Liu C, Chen H, Er HL, Soo HM, Kumar PP, Han JH, Liou YC, Yu H, 2008. Direct interaction of AGL24 and SOC1 integrates flowering signals in *Arabidopsis*. *Development* **135**, 1481-91.

Liu F, Marquardt S, Lister C, Swiezewski S, Dean C, 2010. Targeted 3' processing of antisense transcripts triggers *Arabidopsis FLC* chromatin silencing. *Science* **327**, 94-7.

Liu S, Liu Y, Yang X, Tong C, Edwards D, Parkin IA, Zhao M, Ma J, Yu J, Huang S, Wang X, *et al*., 2014. The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nature communications* **5**, 3930.

Livak KJ, Schmittgen TD, 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. *Methods* **25**, 402-8.

Long Y, Shi J, Qiu D, Li R, Zhang C, Wang J, Hou J, Zhao J, Shi L, Park BS, Choi SR, 2007. Flowering time quantitative trait Loci analysis of oilseed *Brassica* in multiple environments and genomewide alignment with *Arabidopsis*. *Genetics* **177**, 2433-44.

Lou P, Zhao J, Kim JS, Shen S, Del Carpio DP, Song X, Jin M, Vreugdenhil D, Wang X, Koornneef M, Bonnema G, 2007. Quantitative trait loci for flowering time and morphological traits in multiple populations of *Brassica rapa*. *Journal of Experimental Botany* **58**, 4005-16.

Lu GU, Harper AL, Trick MA, Morgan CO, Fraser FI, O'Neill C, Bancroft IA, 2014b. Associative transcriptomics study dissects the genetic architecture of seed glucosinolate content in *Brassica napus*. *DNA Research* **21**, 613-25.

Lu H, Lin T, Klein J, Wang S, Qi J, Zhou Q, Sun J, Zhang Z, Weng Y, Huang S, 2014a. QTL-seq identifies an early flowering QTL located near *Flowering Locus T* in cucumber. *Theoretical and Applied Genetics* **127**, 1491-9.

Luterbacher J, Dietrich D, Xoplaki E, Grosjean M, Wanner H, 2004. European seasonal and annual temperature variability, trends, and extremes since 1500. *Science* **303**, 1499-503.

Lysak MA, Koch MA, Pecinka A, Schubert I, 2005. Chromosome triplication found across the tribe Brassiceae. *Genome Research* **15**, 516-25.

Macknight R, Bancroft I, Page T, Lister C, Schmidt R, Love K, Westphal L, Murphy G, Sherson S, Cobbett C, Dean C. 1997. *FCA*, a gene controlling flowering time in *Arabidopsis*, encodes a protein containing RNA-binding domains. *Cell* **89**, 737-45.

Margueron R, Reinberg D, 2011. The Polycomb complex PRC2 and its mark in life. *Nature* **469**, 343.

Marquardt S, Raitskin O, Wu Z, Liu F, Sun Q, Dean C, 2014. Functional consequences of splicing of the antisense transcript *COOLAIR* on *FLC* transcription. *Molecular Cell* **54**, 156-65.

Mason AS, Rousseau-Gueutin M, Morice J, Bayer PE, Besharat N, Cousin A, Pradhan A, Parkin IA, Chèvre AM, Batley J, Nelson MN, 2015. Centromere locations in *Brassica* A and C genomes revealed through half-tetrad analysis. *Genetics* **202**, 513-23.

Mei D, Wang H, Hu Q, Li Y, Xu Y, Li Y, 2009. QTL analysis on plant height and flowering time in *Brassica napus*. *Plant Breeding* **128**, 458-65.

Meier, U, 2001. Growth Stages of Mono- and Dictoyledonous Plants, 2 edn. *Federal Biological Research Centre for Agrciulture and Forestry, Braunshweig, Germany.*

Michaels SD, Amasino RM, 1999. *FLOWERING LOCUS C* encodes a novel MADS domain protein that acts as a repressor of flowering. *The Plant Cell Online* **11**, 949-56.

Michaels SD, Amasino RM, 2001. Loss of *FLOWERING LOCUS C* activity eliminates the late-flowering phenotype of *FRIGIDA* and autonomous pathway mutations but not responsiveness to vernalization. *The Plant Cell Online* **13**, 935-41.

Michaels SD, Bezerra IC, Amasino RM, 2004. *FRIGIDA*-related genes are required for the winter-annual habit in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 3281-5.

Michaels SD, He Y, Scortecci KC, Amasino RM, 2003. 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**, 10102-7.

Michelmore RW, Paran I, Kesseli R, 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proceedings of the National Academy of Sciences* **88**, 9828-32.

Miller CN, Harper AL, Trick M, Werner P, Waldron K, Bancroft I, 2016. Elucidation of the genetic basis of variation for stem strength characteristics in bread wheat by Associative Transcriptomics. *BMC Genomics* **17**, 500.

Mockler TC, Yu X, Shalitin D, Parikh D, Michael TP, Liou J, Huang J, Smith Z, Alonso JM, Ecker JR, Chory J, 2004. Regulation of flowering time in *Arabidopsis* by K homology domain proteins. *Proceedings of the National Academy of Sciences* **101**, 12759-64.

Moon J, Suh SS, Lee H, Choi KR, Hong CB, Paek NC, Kim SG, Lee I, 2003. The *SOC1* MADS-box gene integrates vernalization and gibberellin signals for flowering in *Arabidopsis*. *The Plant Journal* **35**, 613-23.

Moore RC, Purugganan MD, 2005. The evolutionary dynamics of plant duplicate genes. *Current Opinion in Plant Biology* **8**, 122-8.

Müller D, Leyser O, 2011. Auxin, cytokinin and the control of shoot branching. *Annals of Botany* **107**, 1203-12.

Murphy L, Scarth R, 1994. Vernalization response in spring oilseed rape (*Brassica napus* L.) cultivars. *Canadian journal of plant science* **74**, 275-7.

Mylne JS, Barrett L, Tessadori F, Mesnage S, Johnson L, Bernatavichute YV, Jacobsen SE, Fransz P, Dean C, 2006. *LHP1*, the *Arabidopsis* homologue of *HETEROCHROMATIN PROTEIN1*, is required for epigenetic silencing of *FLC*. *Proceedings of the National Academy of Sciences* **103**, 5012-7.

Nagahara U, 1935. Genomic analysis in Brassica with special reference to the experimental formation of B. napus and peculiar mode of fertilisation. *Jpn J Bot* **7,** 389–452

Napp-Zinn K, 1957. Untersuchungen zur Genetik des Kältebedürfnisses bei *Arabidopsis thaliana*. *Zeitschrift für induktive Abstammungs-und Vererbungslehre* **88**, 253-285.

Nelson MN, Książkiewicz M, Rychel S, Besharat N, Taylor CM, Wyrwa K, Jost R, Erskine W, Cowling WA, Berger JD, Batley J, 2017. The loss of vernalization requirement in narrow-leafed lupin is associated with a deletion in the promoter and de-repressed expression of a *Flowering Locus T* (*FT*) homologue. *New Phytologist* **213**, 220-32.

Nelson MN, Rajasekaran R, Smith A, Chen S, Beeck CP, Siddique KH, & Cowling WA, 2014. Quantitative trait loci for thermal time to flowering and photoperiod responsiveness discovered in summer annual-type *Brassica napus* L. PLoS One, **9**, e102611.

Niwa M, Daimon Y, Kurotani KI, Higo A, Pruneda-Paz JL, Breton G, Mitsuda N, Kay SA, Ohme-Takagi M, Endo M, Araki, T, 2013. BRANCHED1 interacts with FLOWERING LOCUS T to repress the floral transition of the axillary meristems in *Arabidopsis*. *The Plant Cell* **25**, 1228-42.

Okazaki K, Sakamoto K, Kikuchi R, Saito A, Togashi E, Kuginuki Y, Matsumoto S, Hirai M, 2007. Mapping and characterization of *FLC* homologs and QTL analysis of flowering time in *Brassica oleracea*. *Theoretical and Applied Genetics* **114**, 595-608.

Orsel M, Moison M, Clouet V, Thomas J, Leprince F, Canoy AS, Just J, Chalhoub B, Masclaux-Daubresse C, 2014. Sixteen cytosolic glutamine synthetase genes identified in the *Brassica napus* L. genome are differentially regulated depending on nitrogen regimes and leaf senescence. *Journal of Experimental Botany* **65**, 3927-47.

Osborn TC, Butrulle DV, Sharpe AG, Pickering KJ, Parkin IA, Parker JS, Lydiate DJ, 2003. Detection and effects of a homeologous reciprocal transposition in *Brassica napus*. *Genetics* **165**, 1569-77.

Parkin IA, Sharpe A, Keith D, Lydiate D, 1995. Identification of the A and C genomes of amphidiploid *Brassica napus* (oilseed rape). *Genome* **38**, 1122-31.

Parkin IA, Gulden SM, Sharpe AG, Lukens L, Trick M, Osborn TC, Lydiate DJ, 2005. Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. *Genetics* **171**, 765-81.

Parkin IA, Koh C, Tang H, Robinson SJ, Kagale S, Clarke WE, Town CD, Nixon J, Krishnakumar V, Bidwell SL, Denoeud F, *et al*., 2014. Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid Brassica oleracea. *Genome biology* **15**, R77.

Poduska B, Humphrey T, Redweik A, Grbić V, 2003. The synergistic activation of *FLOWERING LOCUS C* by *FRIGIDA* and a new flowering gene *AERIAL ROSETTE 1* underlies a novel morphology in *Arabidopsis*. *Genetics* **163**, 1457-65.

Putterill J, Robson F, Lee K, Simon R, Coupland G, 1995. The *CONSTANS* gene of *Arabidopsis* promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. *Cell* **80**, 847-57.

Qaderi MM, Kurepin LV, Reid DM, 2006. Growth and physiological responses of canola (*Brassica napus*) to three components of global climate change: temperature, carbon dioxide and drought. *Physiologia Plantarum* **128**, 710-21.

Questa JI, Song J, Geraldo N, An H, Dean C, 2016. *Arabidopsis* transcriptional repressor *VAL1* triggers Polycomb silencing at *FLC* during vernalization. *Science* **353**, 485-8.

Raman H, Raman R, Coombes N, Song J, Prangnell R, Bandaranayake C, Tahira R, Sundaramoorthi V, Killian A, Meng J, Dennis ES, 2016. Genome-wide association analyses reveal complex genetic architecture underlying natural variation for flowering time in canola. *Plant, Cell & Environment* **39**, 1228-39.

Raman H, Raman R, Eckermann P, Coombes N, Manoli S, Zou X, Edwards D, Meng J, Prangnell R, Stiller J, Batley J, 2013. Genetic and physical mapping of flowering time loci in canola (*Brassica napus* L.). *Theoretical and Applied Genetics* **126**, 119-32.

Razi H, Howell E, Newbury H, Kearsey M, 2008. Does sequence polymorphism of *FLC* paralogues underlie flowering time QTL in *Brassica oleracea*? *Theoretical and Applied Genetics* **116**, 179-92.

Reeves PA, He Y, Schmitz RJ, Amasino RM, Panella LW, Richards C, 2006. Evolutionary conservation of the *FLC* mediated vernalization response: evidence from the sugar beet (*Beta vulgaris*). *Genetics*.

Ridge S, Brown PH, Hecht V, Driessen RG, Weller JL, 2014. The role of *BoFLC2* in cauliflower (*Brassica oleracea* var. *botrytis* L.) reproductive development. *Journal of Experimental Botany* **66**, 125-35.

Riechmann JL, Meyerowitz EM, 1997. MADS domain proteins in plant development. *Biological Chemistry* **378**, 1079-102.

Risk JM, Laurie RE, Macknight RC, Day CL, 2010. FRIGIDA and related proteins have a conserved central domain and family specific N- and C- terminal regions that are functionally important. *Plant Molecular Biology* **73**, 493-505.

Robertson DS, 1978. Characterization of a mutator system in maize. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **51**, 21-28.

Rosa S, Duncan S, Dean C, 2016. Mutually exclusive sense-antisense transcription at *FLC* facilitates environmentally induced gene repression. *Nat Commun* **7**, 13031.

Rosegrant MW, Leach N, Gerpacio RV, 1999. Alternative futures for world cereal and meat consumption. *Proceedings of the Nutrition Society* **58**, 219-34.

Ruelens P, De Maagd RA, Proost S, Theißen G, Geuten K, Kaufmann K, 2013. *FLOWERING LOCUS C* in monocots and the tandem origin of angiosperm-specific MADS-box genes. *Nature communications* **4**, 2280.

Schiessl S, Huettel B, Kuehn D, Reinhardt R, Snowdon R, 2017a. Post-polyploidisation morphotype diversification associates with gene copy number variation. *Sci Rep* **7**, 41845.

Schiessl S, Iniguez-Luy F, Qian W, Snowdon RJ, 2015. Diverse regulatory factors associate with flowering time and yield responses in winter-type *Brassica napus*. *BMC Genomics* **16**, 737.

Schiessl S, Samans B, Hüttel B, Reinhard R, Snowdon RJ, 2014. Capturing sequence variation among flowering-time regulatory gene homologs in the allopolyploid crop species *Brassica napus*. *Frontiers in plant science* **5**.

Schiessl SV, Huettel B, Kuehn D, Reinhardt R, Snowdon RJ, 2017b. Flowering Time Gene Variation in Brassica Species Shows Evolutionary Principles. *Frontiers in Plant Science* **8**, 1742.

Schmalenbach I, Zhang L, Ryngajllo M, Jiménez-Gómez JM, 2014. Functional analysis of the Landsberg *erecta* allele of *FRIGIDA*. *BMC Plant Biology* **14**, 218.

Schranz ME, Lysak MA, Mitchell-Olds T, 2006. The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. *Trends in Plant Science* **11**, 535-42.

Schranz ME, Quijada P, Sung S-B, Lukens L, Amasino R, Osborn TC, 2002. Characterization and effects of the replicated flowering time gene *FLC* in *Brassica rapa*. *Genetics* **162**, 1457-68.

Shannon S, Meeks-Wagner DR, 1991. A mutation in the *Arabidopsis TFL1* gene affects inflorescence meristem development. *The Plant Cell* **3**, 877-92.

Sharpe A, Parkin I, Keith D, Lydiate D, 1995. Frequent nonreciprocal translocations in the amphidiploid genome of oilseed rape (*Brassica napus*). *Genome* **38**, 1112-21.

Sheldon CC, Conn AB, Dennis ES, Peacock WJ, 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**, 2527-37.

Sheldon CC, Hills MJ, Lister C, Dean C, Dennis ES, Peacock WJ, 2008. Resetting of *FLOWERING LOCUS C* expression after epigenetic repression by vernalization. *Proc Natl Acad Sci U S A* **105**, 2214-9.

Sheldon CC, Rouse DT, Finnegan EJ, Peacock WJ, Dennis ES, 2000. The molecular basis of vernalization: the central role of *FLOWERING LOCUS C* (*FLC*). *Proc Natl Acad Sci USA* **97**, 3753-8.

Shindo C, Aranzana MJ, Lister C, Baxter C, Nicholls C, Nordborg M, Dean C, 2005. Role of *FRIGIDA* and *FLOWERING LOCUS C* in determining variation in flowering time of *Arabidopsis*. *Plant Physiology* **138**, 1163-73.

Shindo C, Lister C, Crevillen P, Nordborg M, Dean C, 2006. Variation in the epigenetic silencing of *FLC* contributes to natural variation in *Arabidopsis* vernalization response. *Genes & Development* **20**, 3079-83.

Simpson GG, Dijkwel PP, Quesada V, Henderson I, Dean C, 2003. *FY* is an RNA 3′ end-processing factor that interacts with *FCA* to control the *Arabidopsis* floral transition. *Cell* **113**, 777-87.

Song YH, Song NY, Shin SY, Kim HJ, Yun DJ, Lim CO, Lee SY, Kang KY, Hong JC, 2008. Isolation of CONSTANS as a TGA4/OBF4 interacting protein. *Molecules & Cells (Springer Science & Business Media BV)* **25**.

Spencer CC, Su Z, Donnelly P, Marchini J, 2009. Designing genome-wide association studies: sample size, power, imputation, and the choice of genotyping chip. *PLoS Genetics* **5**, e1000477.

Stinchcombe JR, Weinig C, Ungerer M, Olsen KM, Mays C, Halldorsdottir SS, Purugganan MD, Schmitt J, 2004. A latitudinal cline in flowering time in *Arabidopsis thaliana* modulated by the flowering time gene *FRIGIDA*. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 4712-7.

Strange A, Li P, Lister C, Anderson J, Warthmann N, Shindo C, Irwin J, Nordborg M, Dean C, 2011. Major-effect alleles at relatively few loci underlie distinct vernalization and flowering variation in *Arabidopsis* accessions. *PLoS One* **6**, e19949.

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

Sung S, He Y, Eshoo TW, Tamada Y, Johnson L, Nakahigashi K, Goto K, Jacobsen SE, Amasino RM, 2006. Epigenetic maintenance of the vernalized state in *Arabidopsis thaliana* requires *LIKE HETEROCHROMATIN PROTEIN 1*. *Nature Genetics* **38**, 706.

Swiezewski S, Liu F, Magusin A, Dean C, 2009. Cold-induced silencing by long antisense transcripts of an *Arabidopsis* Polycomb target. *Nature* **462**, 799-802.

Tadege M, Sheldon CC, Helliwell CA, Stoutjesdijk P, Dennis ES, Peacock WJ, 2001. Control of flowering time by *FLC* orthologues in *Brassica napus*. *The Plant Journal* **28**, 545-53.

Takagi H, Abe A, Yoshida K, Kosugi S, Natsume S, Mitsuoka C, Uemura A, Utsushi H, Tamiru M, Takuno S, Innan H, 2013. QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *The Plant Journal* **74**, 174-83.

Tang H, Woodhouse MR, Cheng F, Schnable JC, Pedersen BS, Conant GC, Wang X, Freeling M, Pires JC, 2012. Altered patterns of fractionation and exon deletions in *Brassica rapa* support a two-step model of paleohexaploidy. *Genetics* **190**, 1563-74.

Taoka KI, Ohki I, Tsuji H, Furuita K, Hayashi K, Yanase T, Yamaguchi M, Nakashima C, Purwestri YA, Tamaki S, Ogaki Y, 2011. 14-3-3 proteins act as intracellular receptors for rice Hd3a florigen. *Nature* **476**, 332.

Taylor JL, Massiah A, Kennedy S, Hong Y, Jackson SD, 2017. *FLC* expression is down-regulated by cold treatment in *Diplotaxis tenuifolia* (wild rocket), but flowering time is unaffected. *Journal of Plant Physiology* **214**, 7-15.

Teutonico R, Osborn T, 1995. Mapping loci controlling vernalization requirement in *Brassica rapa*. *Theoretical and Applied Genetics* **91**, 1279-83.

Tian F, Bradbury PJ, Brown PJ, Hung H, Sun Q, Flint-Garcia S, Rocheford TR, McMullen MD, Holland JB, Buckler ES, 2011. Genome-wide association study of

leaf architecture in the maize nested association mapping population. *Nature Genetics* **43**, 159.

Tiwari SB, Shen Y, Chang HC, Hou Y, Harris A, Ma SF, McPartland M, Hymus GJ, Adam L, Marion C, Belachew A, 2010. The flowering time regulator CONSTANS is recruited to the *FLOWERING LOCUS T* promoter via a unique cis-element. *New Phytologist* **187**, 57-66.

Trick M, Long Y, Meng J, Bancroft I, 2009. Single nucleotide polymorphism (SNP) discovery in the polyploid *Brassica napus* using Solexa transcriptome sequencing. *Plant Biotechnology Journal* **7**, 334-46.

Turck F, Fornara F, Coupland G, 2008. Regulation and identity of florigen: *FLOWERING LOCUS T* moves center stage. *Annual Review of Plant Biology* **59**, 573-94.

Udall JA, Quijada PA, Osborn TC, 2005. Detection of chromosomal rearrangements derived from homologous recombination in four mapping populations of *Brassica napus* L. *Genetics* **169**, 967-79.

Voorman A, Lumley T, Mcknight B, Rice K, 2011. Behavior of QQ-plots and genomic control in studies of gene-environment interaction. *PLoS One* **6**, e19416.

Wang H, Cheng H, Wang W, Liu J, Hao M, Mei D, Zhou R, Fu L, Hu Q, 2016b. Identification of *BnaYUCCA6* as a candidate gene for branch angle in *Brassica napus* by QTL-seq. *Scientific reports* **6**, 38493.

Wang J, Long Y, Wu B, Liu J, Jiang C, Shi L, Zhao J, King GJ, Meng J, 2009b. The evolution of *Brassica napus FLOWERING LOCUS T* paralogues in the context of inverted chromosomal duplication blocks. *BMC Evolutionary Biology* **9**, 271.

Wang N, Chen B, Xu K, Gao G, Li F, Qiao J, Yan G, Li J, Li H, Wu X, 2016a. Association mapping of flowering time QTLs and insight into their contributions to rapeseed growth habits. *Frontiers in plant science* **7**, 338.

Wang N, Qian W, Suppanz I, Wei L, Mao B, Long Y, Meng J, Müller AE, Jung C, Flowering time variation in oilseed rape (*Brassica napus* L.) is associated with allelic variation in the *FRIGIDA* homologue *BnaA. FRI. a*. *Journal of Experimental Botany*, err249.

Wang Q, Sajja U, Rosloski S, Humphrey T, Kim MC, Bomblies K, Weigel D, Grbic V, 2007. *HUA2* caused natural variation in shoot morphology of *A. thaliana*. *Current Biology* **17**, 1513-9.

Wang R, Farrona S, Vincent C, Joecker A, Schoof H, Turck F, Alonso-Blanco C, Coupland G, Albani MC, 2009a. *PEP1* regulates perennial flowering in *Arabis alpina*. *Nature* **459**, 423-7.

Wang X, Wang H, Wang J, Sun R, Wu J, Liu S, Bai Y, Mun JH, Bancroft I, Cheng F, Huang S, *et al*., 2011a. The genome of the mesopolyploid crop species *Brassica rapa*. *Nature Genetics* **43**, 1035-9.

Weber E, Gruetzner R, Werner S, Engler C, Marillonnet S, 2011. Assembly of designer TAL effectors by Golden Gate cloning. *PLoS One* **6**, e19722.

Weigel D, Alvarez J, Smyth DR, Yanofsky MF, Meyerowitz EM, 1992. *LEAFY* controls floral meristem identity in *Arabidopsis*. *Cell* **69**, 843-59.

Wenkel S, Turck F, Singer K, Gissot L, Le Gourrierec J, Samach A, Coupland G, 2006. CONSTANS and the CCAAT box binding complex share a functionally important domain and interact to regulate flowering of *Arabidopsis*. *The Plant Cell* **18**, 2971-84.

Werner JD, Borevitz JO, Uhlenhaut NH, Ecker JR, Chory J, Weigel D, 2005. *FRIGIDA*-independent variation in flowering time of natural *Arabidopsis thaliana* accessions. *Genetics* **170**, 1197-207.

Whittaker C, Dean C, 2017. The *FLC* Locus: A Platform for Discoveries in Epigenetics and Adaptation. *Annual Review of Cell and Developmental Biology* **33**, 555-75.

Wigge PA, 2011. *FT*, a mobile developmental signal in plants. *Current Biology* **21**, R374-R8.

Wigge PA, Kim MC, Jaeger KE, Busch W, Schmid M, Lohmann JU, Weigel D, 2005. Integration of spatial and temporal information during floral induction in *Arabidopsis*. *Science* **309**, 1056-9.

Willmann MR, Poethig RS, 2011. The effect of the floral repressor *FLC* on the timing and progression of vegetative phase change in *Arabidopsis*. *Development* **138**, 677-85.

Wilson RN, Heckman JW, Somerville CR, 1992. Gibberellin is required for flowering in *Arabidopsis thaliana* under short days. *Plant Physiology* **100**, 403-8.

Wood CC, Robertson M, Tanner G, Peacock WJ, Dennis ES, Helliwell CA, 2006. The *Arabidopsis thaliana* vernalization response requires a polycomb-like protein complex that also includes *VERNALIZATION INSENSITIVE 3*. *Proceedings of the National Academy of Sciences* **103**, 14631-6.

Woodhouse S, 2017. Analysis of temporal and spatial expression of a key floral repressor in *Brassica*. *Univeristy of East Anglia Msc Thesis*

Wu J, Wei K, Cheng F, Li S, Wang Q, Zhao J, Bonnema G, Wang X, 2012. A naturally occurring InDel variation in *BraA. FLC. b* (*BrFLC2*) associated with flowering time variation in *Brassica rapa*. *BMC Plant Biology* **12**, 151.

Wu Z, Ietswaart R, Liu F, Yang H, Howard M, Dean C, 2016. Quantitative regulation of *FLC* via coordinated transcriptional initiation and elongation. *Proceedings of the National Academy of Sciences* **113**, 218-23.

Xiao D, Zhao JJ, Hou XL, Basnet RK, Carpio DP, Zhang NW, Bucher J, Lin K, Cheng F, Wang XW, Bonnema G, 2013. The *Brassica rapa FLC* homologue *FLC2* is a key regulator of flowering time, identified through transcriptional co-expression networks. *Journal of Experimental Botany* **64**, 4503-16.

Xu L, Hu K, Zhang Z, Guan C, Chen S, Hua W, Li J, Wen J, Yi B, Shen J, Ma C, 2015. Genome-wide association study reveals the genetic architecture of flowering time in rapeseed (*Brassica napus* L.). *DNA Research* **23**, 43-52.

Yamaguchi A, Kobayashi Y, Goto K, Abe M, Araki T, 2005. *TWIN SISTER OF FT* (*TSF*) acts as a floral pathway integrator redundantly with *FT*. *Plant and cell physiology* **46**, 1175-89.

Yanovsky MJ, Kay SA, 2002. Molecular basis of seasonal time measurement in *Arabidopsis*. *Nature* **419**, 308.

Yi G, Park H, Kim J-S, Chae WB, Park S, Huh JH, 2014. Identification of three *FLOWERING LOCUS C* genes responsible for vernalization response in radish (*Raphanus sativus* L.). *Horticulture, Environment, and Biotechnology* **55**, 548-56.

Yi L, Chen C, Yin S, Li H, Li Z, Wang B, King GJ, Wang J, Liu K, 2018. Sequence variation and functional analysis of a *FRIGIDA* orthologue (*BnaA3. FRI*) in *Brassica napus*. *BMC Plant Biology* **18**, 32.

Yoo SK, Chung KS, Kim J, Lee JH, Hong SM, Yoo SJ, Yoo SY, Lee JS, Ahn JH, 2005. *CONSTANS* activates *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* through *FLOWERING LOCUS T* to promote flowering in *Arabidopsis*. *Plant Physiology* **139**, 770-8.

Young MD, Willson TA, Wakefield MJ, Trounson E, Hilton DJ, Blewitt ME, Oshlack A, Majewski IJ, 2011. ChIP-seq analysis reveals distinct H3K27me3 profiles that correlate with transcriptional activity. *Nucleic Acids Research* **39**, 7415-27.

Yu J, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S, 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics* **38**, 203.

Yuan YX, Wu J, Sun RF, Zhang XW, Xu DH, Bonnema G, Wang XW, 2009. A naturally occurring splicing site mutation in the *Brassica rapa FLC1* gene is associated with variation in flowering time. *Journal of Experimental Botany* **60**, 1299-308.

Zhang X, Meng L, Liu B, Hu Y, Cheng F, Liang J, Aarts MG, Wang X, Wu J, 2015. A transposon insertion in *FLOWERING LOCUS T* is associated with delayed flowering in *Brassica rapa*. *Plant Science* **241**, 211-20.

Zhang Z, Ersoz E, Lai CQ, Todhunter RJ, Tiwari HK, Gore MA, Bradbury PJ, Yu J, Arnett DK, Ordovas JM, Buckler ES, 2010. Mixed linear model approach adapted for genome-wide association studies. *Nature Genetics* **42**, 355.

Zhao J, Kulkarni V, Liu N, Del Carpio DP, Bucher J, Bonnema G, 2010. *BrFLC2* (*FLOWERING LOCUS C*) as a candidate gene for a vernalization response QTL in *Brassica rapa*. *Journal of Experimental Botany* **61**, 1817-25.

Zhao K, Aranzana MJ, Kim S, Lister C, Shindo C, Tang C, Toomajian C, Zheng H, Dean C, Marjoram P, Nordborg M, 2007. An *Arabidopsis* example of association mapping in structured samples. *PLoS Genetics* **3**, e4.

Zhao K, Tung CW, Eizenga GC, Wright MH, Ali ML, Price AH, Norton GJ, Islam MR, Reynolds A, Mezey J, McClung AM, 2011. Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nature communications* **2**, 467.

Zhou Q, Han D, Mason AS, Zhou C, Zheng W, Li Y, Wu C, Fu D, Huang Y, 2017. Earliness traits in rapeseed (*Brassica napus*): SNP loci and candidate genes identified by genome-wide association analysis. *DNA Research* **25**, 229-44.

Zou X, Suppanz I, Raman H, Hou J, Wang J, Long Y, Jung C, Meng J, 2012. Comparative analysis of *FLC* homologues in Brassicaceae provides insight into their role in the evolution of oilseed rape. *PLoS One* **7**, e45751.