Deciphering the Variation in Flowering Behaviour in a Norwegian Population of Perennial *Arabis alpina*

INAUGURAL-DISSERTATION

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Yi-Chen Lin

aus Taipeh, Taiwan

Köln, 2022

Die vorliegende Arbeit wurde am Max-Planck-Institut für Pflanzenzüchtungsforschung in Köln in der Abteilung für Entwicklungsbiologie der Pflanzen (Direktor Prof. Dr. G. Coupland), Arbeitsgruppe Prof. Dr. George Coupland durchgeführt.

The work described in this thesis was conducted at the Max Planck Institute for Plant Breeding Research in the Department of Plant Developmental Biology (Director: Prof. Dr. George Coupland) under the supervision of Prof. Dr. George Coupland.

Berichterstatter (Gutachter):
Prof. Dr. George Coupland
Prof. Dr. Martin Hülskamp

Vorsitzender der Komission:
Prof. Dr. Ute Höcker

Tag der mündlichen Prüfung:
14. Juni 2022

Acknowledgements

First and foremost, I would like to thank my supervisor Prof. Dr. George Coupland for his great support during my PhD. George, it has been a true pleasure to work in your group as well as with this comprehensive project. I had enjoyed every scientific discussion with you and they were truly inspiring. Also, many thanks for joining and helping me in Norway for the field work in August 2021 and it was indeed an unforgettable memory.

Secondly, I would like to thank my examination committee: Prof. Dr. Martin Hülskamp, Prof. Dr. Ute Höcker and Dr. Andrea Fulgione, for spending time evaluating my thesis and assessing my PhD defence. Also, I want to thank Prof. Dr. Juliette de Meaux for her valuable feedback during every thesis advisory committee meeting.

Special appreciation to Dr. Ulla Kemi. Ulla, many thanks for teaching me everything about the Riasten population and how to perform the field work in Stugudalen. I truly miss the time that we spent on scoring flowering together in the greenhouse as well as in Andy's plot at the field site and it has always been very rewarding working with you! I would also like to thank Dr. Jörg Wunder, Dr. Stefan Wötzel and Dr. Luise Hecker for the guidance during the early stage of my PhD.

Furthermore, I would like to particularly thank Dr. Edouard Severing for all the nice discussions on the analysis of RNA-sequecning as well as endless support in Bioinformatics. The completion of QTL mapping experiment will not be possible without the help and support from Dr. Eva Madrid, Dr. Niharika Sashidhar, Kerstin Luxa and Andrea Lossow. Many thanks to Dr. Andrea Fulgione, Basitaan Tjeng, and Birgit Walkemeier for performing field work in Norway with me during

the pandemic. Birgit, I have enjoyed every coffee chat that we had and thank you very much for being an amazing friend during my PhD time in MPIPZ!

Michaela Lehnen and Brigitte Koop, many thanks for the great help and make the implementation of my experiment in the lab much easier. Ute, thank you so much for all your help in visa extension, scheduling meetings and other personal requests.

Plenty of gratefulness for all present and former members as well as student helpers in this group during my PhD. Annabel and Chloe, many thanks for teaching the lab skills and giving me brilliant ideas to improve the experiments. Oliver, Samantha, Sarah, Denitsa, Alex and Isabella, thanks for the great help in the greenhouse with the massive number of plants. Thank you Dr. John Chandler for your help with scientific English in this thesis. All other members in the group create a nice working environment and are willing to help, explain and share their own knowledge with me.

I am extremely grateful for the MPIPZ greenhouse, genome centre and technical team to make the implemetation of my large experiments possible and smooth. Aristeidis, Jose, Frank, Andreas, John and Renate, thank you for spending time on taking care of my experiments in the greehnouse and Halle 14. I would like to thank Væktarstua Hotel for always providing a nice accommodation for our field trip, thank you Heidi! I want to also take time to thank our PhD coordinator Dr. Stephan Wagner and I appreciate all your help and support during my PhD. Many thanks for making our PhD programme a welcoming environment for international students.

I would like to also thank my family for all their support during my years in Germany and I certainly can not accomplish my PhD without their love. Specially, I am grateful for having my little brother and sister, Peter and Chrissie, many thanks for taking care of mom and dad when I am far away from home.

Last but not least, Seamus, I am really grateful for having you in my life. Many thanks for all your unconditional love that makes me a better person. Without your support, the completion of this thesis is not even possible and many thanks for always being patient, considerate and understanding.

Abstract

The perennial model species *Arabis alpina*, a relative of *Arabidopsis thaliana* within the Brassicaceae, exhibits a wide variety of flowering behavior amongst natural populations that range geographically from northern Spain to Norway. The Spanish population has previously been studied and it was shown that *PERPETUAL FLOWERING 1* (*PEP1*), the ortholog of the *A. thaliana* gene *FLOWERING LOCUS C* (*FLC*), controls floral induction in response to winter cold (vernalization). However, the Norwegian population is less well studied and exhibits flowering behaviours that are not present in the Spanish population. This thesis describes the flowering phenotypes of Norwegian plants when grown in controlled or field conditions and analyzes the molecular and genetic basis of differences in their flowering behaviour to define the flowering mechanisms of the Norwegian Riasten population of *Arabis alpina*.

In the absence of vernalization, the Norwegian Riasten population exhibits early- and late-flowering phenotypes, which are not due to variation in the timing of floral induction, but to differences in bud outgrowth. However, in the presence of vernalization, the flowering mechanisms of the early- and late-flowering accessions still differed. In the growth chamber, the early-flowering accession began flowering at the primary shoot and the I zone axillary branches followed by flowering on the V1 zone branches; however, the late-flowering accession only flowered at the primary shoot and the I zone branches. The activity of axillary meristems within the V1 zone in the early-flowering accession enables a staggered flowering behaviour, which distinguishes the underlying flowering mechanisms of the early- and late-flowering accessions.

Genome-wide transcriptome analysis indicated that *PEP1* expression increased monotonically in both early and late-flowering accessions during floral induction in the absence of vernalization. Co-expression network analysis revealed that the majority of floral organ identity and floral repressor genes in the early-flowering accession belonged to two distinct clusters that were expressed in opposite patterns, whereas in the late-flowering accession, the floral repressors were split between multiple clusters that showed similar qualitative expression patterns, although these were asynchronous. The genetic basis of the variation in bud outgrowth was narrowed down to 759 Kb ($\sim$120 genes) by analysing near-isogenic lines in the late-flowering genetic background that carried early flowering alleles on chromosome 8 that conferred early flowering.

An *in situ* transplantation experiment in the natural environment also revealed different growth strategies between the two accessions. The early-flowering accession invests resources into growing numerous rosettes, ensuring that sufficient rosettes flower the subsequent year, whereas the late-flowering accession invests resources into producing a greater number of flowers, leading to more siliques, to help to maximise offspring production.

The results of my RNA-sequencing and QTL fine-mapping experiments increase our understanding of the control of flowering behaviour due to the regulation of bud arrest on molecular and genetic levels. In a complementary approach, the results of the phenotypic analysis and fieldwork shed light on the local adaptation of plants to account for short growing season in the natural environment, which is manifested by certain biological traits. Taken together, the four research approaches in this study have facilitated a better understanding of the variation in the flowering behaviour within the Norwegian Riasten *A. alpina* population.

Contents

List of Tables **6**

List of Figures **9**

List of Abbreviations **15**

1 Introduction **19**

1.1 Local adaptation 19

1.2 Local adaptation studies in natural plant populations 20

1.2.1 Transplant experiments and genomic studies on local adaptation 20

1.2.2 Gene expression studies of local adaptation 22

1.3 Variation in flowering behaviours among plants with different life histories 23

1.3.1 Trade-offs between flowering and vegetative growth 24

1.3.2 Flowering behaviour in annuals and perennials 25

1.4 Cycling between vegetative growth and flowering is determined by meristem fate in perennials 29

1.5 *Arabis alpina* as a perennial model species for evolutionary studies . 30

1.6 Genetic interactions regulate flowering in *Arabis alpina* 31

1.7 Aims of this thesis 32

2 Material and methods **35**

2.1 Study population 35

2.2 Plant cultivation in greenhouse and growth chamber 35

2.3 Generation of plant cuttings and crosses 36

2.4 Phenotypic description of plants in the greenhouse and growth chamber . . . 36
2.5 Phenotypic description at the Norwegian field site . . . 37
2.6 Recording of environmental variables in greenhouse, growth chamber and the natural environment . . . 38
2.7 Gibberellin application . . . 38
2.8 Plant DNA extraction, Polymerse Chain Reaction and genotyping . . . 38
2.9 Genetic markers used in QTL mapping . . . 39
2.10 RNA extraction . . . 39
2.11 Processing RNA-sequencing raw reads . . . 40
2.12 Co-expression network and Gene Ontology enrichment analysis . . . 41
2.13 Synteny analysis . . . 41
2.14 Statistical analysis and data visualisation . . . 42

3 Phenotypic variation in flowering behaviour of the Norwegian Riasten population 43
3.1 Background . . . 43
3.2 Phenotypic analysis under greenhouse conditions . . . 46
3.2.1 Experimental design . . . 46
3.2.2 Extended vernalization rescues primary shoot arrest phenotype in the late-flowering accession . . . 46
3.2.3 Longer vernalization decreases the flowering time of axillary branches . . . 48
3.2.4 Longer vernalization promotes growth in the late-flowering accession . . . 49
3.2.5 No vegetative zone (V3) was observed in either accession after vernalization . . . 50
3.3 Phenotypic analysis under growth chamber conditions . . . 53
3.3.1 Experimental Design . . . 53
3.3.2 Differences in flowering time still occur if plants are exposed to vernalization after floral transition . . . 55

3.3.3 Flowering time does not always decrease as the vernalization period is extended 55
3.3.4 A vernalization period of 20 weeks does not cause differences in primary shoot elongation between the accessions 57
3.3.5 Primary shoot elongation of the late-flowering accession is more sensitive to vernalization and pre-growth periods than that of the early-flowering accession 60
3.3.6 The longest vernalization period generates differences in silique production on the primary shoot 61
3.3.7 Extended vernalization increases silique production in both accessions 63
3.3.8 Correlation matrices reveal the dynamics among flowering time, plant height and flower production 65
3.3.9 The relationships between primary shoot and axillary branch traits reveal unique features in each accession 70
3.3.10 The early-flowering accession continues to flower from the V1 zone after the completion of flowering from the I zone . 74
3.3.11 Shoot elongation in the V1 zone of the early-flowering accession is controlled by the pre-growth period 77
3.4 Discussion 79

4 Genome-wide transcriptome analysis of early- and late-flowering accessions **83**
4.1 Experimental design under growth chamber conditions 83
4.2 The bud-arrest phenotype is also exhibited under growth chamber conditions 86
4.3 Principal component analysis is consistent with the timing of the appearance of the bud-arrest phenotype 88
4.4 Expression of *PEP1* in Norwegian *A. alpina* 89
4.5 The molecular basis for the bud-arrest phenotype in the late accession is independent from floral organ specification 91
4.6 Hierarchical clustering of key floral repressors and organ identity genes 93

4.7 Genome-wide co-expression network analysis reveals different cluster structures between accessions . . . 98
4.8 Discussion . . . 104

5 Genetic analysis of variation in bud outgrowth of early- and late-flowering accessions 109
5.1 Background . . . 109
5.2 Fine-mapping using the BC3F2 generation . . . 113
5.3 Fine-mapping using the BC3F3 generation . . . 118
5.4 Differentially expressed candidate genes . . . 121
5.5 *ERL2* is coexpressed with the floral repressor *SVP* and *TFL1* in the early-flowering accession . . . 126
5.6 Discussion . . . 131

6 *In situ* transplantation experiment at the Norwegian Riasten field site 135
6.1 Introduction to transplantation experiment . . . 135
6.1.1 Survival rate . . . 137
6.1.2 Rosette number . . . 137
6.1.3 Flowering episodes . . . 142
6.1.4 Number of flowering shoots and siliques . . . 146
6.2 Phenology of labelled plants in the natural environment . . . 149
6.2.1 Survival and flowering episodes . . . 149
6.2.2 Number of rosettes and inflorescences . . . 151
6.3 Discussion . . . 151
6.3.1 Early and late-flowering accessions show different growth strategies in the natural environment . . . 151
6.3.2 The importance of phenology data to understand the perennial species . . . 155

7 Conclusions and perspectives 159

A Appendix - Section 3.2 Data 163

B Appendix - Section 3.3 Data **189**

C Appendix - *In situ* Transplantation Data **193**

Bibliography **193**

Erklärung zur Dissertation **212**

List of Tables

2.1 PCR mixture for GoTaq PCR 39
2.2 PCR programme for GoTaq PCR 39
2.3 InDel primers used for fine-mapping 40
2.4 SNP primers used for fine-mapping 41

3.1 The height at primary shoot (cm) for each accession one month after 8, 12, 16 or 20 weeks of vernalization 48
3.2 The number of flowering axillary branches for each accession one month after 8, 12, 16 or 20 weeks of vernalization 49
3.3 The 12 different combinations of pre-growth and vernalization periods used for the growth chamber experiment 54
3.4 The number of flowering axillary branches per plant for treatments 1, 2 and 3 for each accession 79

6.1 Model parameters for survival probability in plot 1 139
6.2 Model parameters for survival probability in plot 2 139
6.3 Model parameters for total number of rosette in plot 1. 141
6.4 Model parameters for the flowering probability in plots 1, 2 and 4 for the 2019 flowering data 144
6.5 Model parameters for the expected number of flowering events of plants in plot 1 from 2019 to 2021 145
6.6 Model parameters for mean number of siliques per flowering shoot in plot 1 147

A.1 Data of the 8 weeks of vernalization 163

A.2 Data of the 12 weeks of vernalization 169
A.3 Data of the 16 weeks of vernalization 175
A.4 Data of the for the 20 weeks of vernalization 181

List of Figures

1.1 Conceptual diagram of the life cycle for annual, biennial, and iteroparous perennial plants 26

3.1 Identification of early- and late-flowering individuals from the Norwegian Riasten population 44

3.2 Phenotypic differences in the early- and late-flowering accessions under greenhouse conditions 45

3.3 The height of plants from the early- and late-flowering accessions vernalized for 8, 12, 16 or 20 weeks under greenhouse conditions . . 47

3.4 Flowering time of axillary branches of plants of early- and late-flowering accessions exposed to 8, 12, 16 or 20 weeks of vernalization under greenhouse conditions 49

3.5 The length of axillary branches from each node for 7-week-old plant from early- and late-flowering accession vernalized for 8, 12, 16 or 20 weeks under greenhouse conditions 50

3.6 The developmental zonations of early- and late-flowering accessions vernalized for 8, 12, 16 or 20 weeks under greenhouse conditions . . 51

3.7 Set up of the growth chamber experiment 54

3.8 Microscopic examination of the apices of plants of the early- and late-flowering accessions before vernalization and the flowering time data distribution of the primary shoot from both accessions after vernalization under growth chamber conditions 56

3.9 Flowering time of the primary shoot under growth chamber conditions according to accession and treatment 58

3.10 Distributions of the height of the primary shoot of each accession measured one month after vernalization under growth chamber conditions for each treatment 59

3.11 Height of the primary shoot measured one month after vernalization under growth chamber conditions according to accession and treatment 61

3.12 The number of siliques on the primary shoot of plants of each accession for each treatment, measured one month after vernalization under growth chamber conditions 62

3.13 Number of siliques on the primary shoot for each accession measured one month after vernalization under growth chamber conditions . . 64

3.14 Correlation matrices for primary shoot traits for each accession . . . 67

3.15 Correlation matrices for axillary branch traits for each accession . . 69

3.16 Correlation matrix for primary shoot and axillary branches traits for the early-flowering accession 71

3.17 Correlation matrix for the primary shoot and axillary branches traits for the late-flowering accession 72

3.18 Correlation matrices showing the interaction between the primary shoot and the axillary branches traits for the early- and late-flowering accessions 73

3.19 Analysis of the length, flowering time and developmental fate on each axillary branches of the early- and late-flowering accessions, for treatments including 5 weeks of pre-growth and 12, 16 and 20 weeks of vernalization 76

3.20 Analysis of the length, flowering time and developmental fate on each axillary branch of the early- and late-flowering accessions for treatments including 3 weeks of pre-growth and 12, 16 and 20 weeks of vernalization 78

4.1 Experimental design for RNA-seq analysis of the Norwegian *A. alpina* early- and late-flowering accessions under growth chamber conditions 85

4.2 Phenotypic differences between early- and late-flowering accessions under growth chamber conditions 87

4.3 Principal component analysis (PCA) for three biological replicate samples for each time point of the RNA-seq sample series under non-vernalization and vernalization conditions 89

4.4 The expression of *PEP1* in Norwegian *A. alpina* accessions throughout development 90

4.5 The normalised expression of four floral organ identity genes in Norwegian *A. alpina* accessions in the absence of vernalization over time. 92

4.6 The normalised expression of four floral organ identity genes in vernalized Norwegian *A. alpina* accessions over time 92

4.7 Hierarchical clustering of the expression of genes encoding selected key regulators of plant reproductive development in Norweigian *A. alpina* accessions without vernalization 96

4.8 Hierarchical clustering of the expression of genes encoding selected key regulators of plant reproductive development in Norweigian *A. alpina* accessions with vernalization 97

4.9 Genome-wide co-expression clustering analysis for Norweigian *A. alpina* accessions without vernalization 100

4.10 Genome-wide structure of gene clusters in Norweigian *A. alpina* accessions 100

4.11 Expression pattern of the two largest gene clusters in the Norweigian *A. alpina* early-flowering accession without vernalization 101

4.12 The expression patterns of genes in six selected gene clusters for the Norweigian *A. alpina* late-flowering accession without vernalization 103

4.13 A conceptualised model for flowering in the Norweigian *A. alpina* early- and late-flowering accessions compared with Spanish reference accession Pajares without vernalization 106

5.1 Distribution of flowering time data for F1 and F2 populations of Norwegian *A. alpina* in the greenhouse 110

5.2 Limit of detection (LOD) curve of the simple interval mapping of Norwegian *A. alpina* for the number of days until the first flower opens in the greenhouse in the absence of vernalization 110

5.3 A schematic representation of the fine-mapping strategy for linkage group 8 that confers early flowering to *Arabis alpina* 112

5.4 Flowering-time distribution for individuals from three different selected *Arabis alpina* families and the corresponding genotype for markers in LG1 and 8 115

5.5 The flowering-time distribution of 31 individuals according to their allelic genotype for the genetic markers AL925 and SNP168 within the QTL interval. 116

5.6 The R-squared distributions resulting from 1,000 bootstrap iterations of linear regression modelling of flowering time with respect to allelic combinations of the AL925 and SNP-168 genetic markers. . . 117

5.7 The flowering-time distribution for 1,584 BC3F3 *Arabis alpina* plants derived from the 425(s)(s) mother plant, with corresponding genotypes, and a summary of the genetic markers within the QTL interval.119

5.8 Flowering time distributions separated by allelic combinations for two selected markers in the QTL interval 121

5.9 The expression level of *ERL2* normalised as log2 fragments per kilobase of transcript per million mapped reads (FPKM) in *Arabis alpina* plants of the early- and late-flowering accession that were either not vernalized or were vernalized for 16 weeks 122

5.10 The expression level of *BAG3* normalised as log2 fragments per kilobase of transcript per million mapped reads (FPKM) in *Arabis alpina* plants of the early- and late-flowering accession that were either not vernalized or were vernalized for 16 weeks 123

5.11 The expression level of *AT5G07230* normalised as log2 fragments per kilobase of transcript per million mapped reads (FPKM) in *Arabis alpina* plants of the early- and late-flowering accession that were either not vernalized or were vernalized for 16 weeks 123

5.12 The expression level of *IMD2* normalised as log2 fragments per kilobase of transcript per million mapped reads (FPKM) in *Arabis alpina* plants of the early- and late-flowering accession that were either not vernalized or were vernalized for 16 weeks 125

5.13 The expression levels of *GRP16* and *GRP17* normalised as log2 fragments per kilobase of transcript per million mapped reads (FPKM) in *Arabis alpina* plants of the early- and late-flowering accession that were either not vernalized or were vernalized for 16 weeks 125

5.14 The expression levels of *GRP19* and *GRP20* normalised as log2 fragments per kilobase of transcript per million mapped reads (FPKM) in *Arabis alpina* plants of the early- and late-flowering accession that were either not vernalized or were vernalized for 16 weeks 126

5.15 Temporal expression dynamics for all the genes in the gene cluster containing *ERL2* for both the early- and late-flowering accessions of Norweigian *A. alpina* without vernalization. 127

5.16 A TreeMap view of GO-term clusters for the gene cluster containing the *ERL2* gene in the Norweigian *A. alpina* early-flowering accession without vernalization 129

5.17 A TreeMap view of GO-term clusters for the gene cluster containing the *ERL2* gene in the Norweigian *A. alpina* late-flowering accession without vernalization 130

5.18 Synteny analysis of the QTL interval located on chromosome 8 between the Norweigian *A. alpina* early and late accession 132

5.19 Application of gibberellin acid to plants from the Norweigian *A. alpina* late-flowering accession 134

6.1 The Predicted probability of survival for plants in plot 1 according to family 138

6.2 Predicted probability of survival in plot 2 according to family . . . 138

6.3 The expected number of rosettes in plot 1 according to family using equation 6.2 140

6.4 The total number of rosettes of the early- and late-flowering genotypes in plot 1 . . . 142
6.5 The predicted probability of flowering in plot 1, 2 and 4 for the 2019 data according to family. . . . 143
6.6 The predicted number of flowering events for plants in plot 1 from 2019 to 2021 according to family . . . 145
6.7 The predicted mean number of siliques per flowering shoot in plot 1 according to family . . . 147
6.8 Mean number of siliques per flowering shoot of the early- and late-flowering genotypes in plot 1 . . . 148
6.9 Survival and flowering heatmap from 2015 to 2021 . . . 150
6.10 A heatmap depicting the number of rosettes and inflorescences from 2015 to 2021. . . . 152
6.11 Aggregated rosette phenotype in the natural environment . . . 154
6.12 The length of axillary branches from plants of the early- and l ate-flowering accessions grown in the growth chamber. . . . 156

B.1 The flowering of the primary shoot and developmental fate for the axillary branche on each node from early- and late-flowering accession without vernalization under growth-chamber conditions . . . 189
B.2 Microscopic examination of the apices of plants of the early-flowering accession when vernalization ceases under growth-chamber conditions 190
B.3 Microscopic examination of the apices of plants of the late-flowering accession when vernalization ceases under growth-chamber conditions 191

C.1 Photograph of labelled plant in the natural environment and *in situ* transplantation experiment . . . 193
C.2 Climate data of temperature and light intensity for plot 1 from 2018 to 2021 at Norwegian field site . . . 194
C.3 Plant height in plot 1 for the 2021 data according to family . . . 195
C.4 Number of flowering shoots in plot 1 for the 2021 data according to family . . . 195

List of Abbreviations

A	Adenin
AG	AGAMOUS
AP1	APETALA 1
AP3	APETALA 3
At	Arabidopsis thaliana
bp	Base pair(s)
bZIP	basic LEUCINE ZIPPER
BAG3	BCL-2-ASSOCIATED ATHANOGENE 3
BC	back-cross
C	Cytosin
DEGs	Differentially Expressed Genes
dH_2O	distilled, deionized water
DNA	Deoxyribonucleic acid
DTF	days to the opening of first flower
dNTP	deoxyribonuclease
ERL2	ERECTA-LIKE 2
F1	Filial 1
F2	Filial 2

FLC FLOWERING LOCUS C

FLM FLOWERING LOCUS M

FPKM Fragments Per Kilobase Million

FT FLOWERING LOCUS T

FW Forward

G Guanin

GA Gibberellin (Gibberellic acid)

GA20OX3 GIBBERELLIN 20-OXIDASE 3

GLMs generalized linear models

GNR gene regulatory network

GO Gene Ontology

GRPs glycine-rich proteins

GRP16 GLYCINE-RICH PROTEIN 16

GRP17 GLYCINE-RICH PROTEIN 17

GRP19 GLYCINE-RICH PROTEIN 19

GRP20 GLYCINE-RICH PROTEIN 20

h hours

IMD2 ISOPROPYLMALATE DEHYDROGENASE 2

IQR interquartile range

kb Kilo base(s)

lncRNA long non-coding RNA

LD Long days

LG linkage group

LOD limit of detection

mg milligram

miRNA microRNA

mRNA Messenger Ribonucleic acid

M molar

Mb Megabase

MADS-box MCM1 AGAMOUNS DEFICIENS SRF-box

MAF1-5 MADS AFFECTING FLOWERING 1-5

MIR156 MICRORNA 156

MWU Mann-Whitney U

ng nanogram

PCA Principal Component Analysis

PCR Polymerase chain reaction

PEP1 PERPETUAL FLOWERING 1

PI PISTILLATA

QTL quantitative trait locus

REV Reverse

RNA Ribonucleic acid

RNA-seq RNA-sequencing

sd standard deviation

SAM Shoot apical meristem

SD Short days

SEP3 SEPALLATA 3

SNP single-nucleotide polymorphism

SOC1 SUPPRESSOR OF OVEREXPRESSION OF CONSTANS a

SPL15 SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 15

SV structural variant

SVP SHORT VEGETATIVE PHASE

T Thymin

TFL1 TERMINAL FLOWER 1

TSf TWINSISTER OF FT

w week(s)

ZT Zeitgeber

Chapter 1
Introduction

Life exhibits an extraordinary degree of diversity. It was this diversity for which naturalist Charles Darwin sought an explanation during his voyage on the *Beagle*. Phenotypic variation within and between species does not merely shape biodiversity, but also generates it. Geographical patterns of natural variation in phenotypes reflect the selection pressure by the local environment. Within the framework of evolutionary biology, several key and fascinating questions can be raised about natural variation within a species: How does a species adapt to its environment? Which biological traits are selected during the adaptation process and are manifested as different phenotypes? What are the molecular and genetic bases for the natural variation in these biological traits?

1.1 Local adaptation

The phenotypic traits of populations of the same species can diverge greatly (Turesson, 1922), as exemplified by studies into human height in Europe (Turchin et al., 2012), the colouration patterns of cichlid fishes (Maan and Sefc, 2013), the flowering time of monkeyflowers, *Mimulus guttatus* (Friedman and Willis, 2013), and trichome variation that results in resistance to herbivory in *Arabidopsis lyrata* (Kivimäki et al., 2007). The divergent phenotypes within a species, which lead to population differentiation, often arise from local adaption, i.e., the genotypes of local populations are better adapted to their local environments than those of other populations (Hereford, 2009). The strict criterion for local adaptation is that a population must have a higher fitness at its native site than any other

population introduced to that site (Kawecki and Ebert, 2004; Savolainen et al., 2013). Importantly, local adaptation is not uniquely responsible for population differentiation, which can also be caused by genetic drift and gene flow (Wright, 1931; Sork, 2018).

1.2 Local adaptation studies in natural plant populations

Due to their sessile nature, plants cannot move to avoid adverse environmental conditions. How then do plants adapt and accommodate to such changes? This fundamental and intriguing question leads to the conclusion that plants must be subject to the local selection of traits, which provides a promising avenue to study local adaptation (Leimu and Fischer, 2008; Anderson et al., 2011). Moreover, due to the temporally and spatially heterogeneous environments where conditions vary within the lifetime of plants, adaptive phenotypic plasticity and genetic differentiation within a species can occur. Phenotypic plasticity means the same genotype can lead to phenotypic changes, as in flowering in response to vernalization and snow melt (Anderson et al., 2012). Adaptive genetic differentiation means the selection pressure alters genotype frequencies, which may change the phenotype. These two concepts allow plants to maximise their fitness (the relative number of offspring in the next generation) in response to heterogeneous landscapes (Baythavong and Stanton, 2010; Anderson et al., 2010).

1.2.1 Transplant experiments and genomic studies on local adaptation

Transplantation experiments, including reciprocal transplantation and common garden experiments, represent powerful research approaches with which to decipher the genetic basis of phenotypic differentiation within and among different environments (Sork, 2018). Reciprocal transplantation experiments grow individuals from at least two populations in their respective native and non-native environments whereas common garden experiments grow individuals from different populations in a common environment (Savolainen et al., 2013; de Villemereuil et al., 2016). Clausen et al. (1940) performed the first study that demonstrated local adaption via reciprocal transplantation experiments and they found that plants

typically grow better where they occur naturally. Some of the earliest studies on local adaptation using common garden experiments occurred in forestry: Foresters compared the growth and other traits of different populations to determine the seed resources that would maximise timber production in plantations (Matyas, 1996; Langlet, 1971; Savolainen et al., 2007; Grattapaglia et al., 2009; Sork et al., 2013; Anders et al., 2015). The two approaches of reciprocal transplantation and common garden experiments are complementary and allow the study of two distinct, but related questions. Reciprocal transplantation experiments test whether plants survive best in their natural environment and perform better than plants in non-native environments. By contrast, common garden experiments test whether phenotypic differentiation occurs among individuals of different populations grown in the same environment and is therefore due to genetic differences between the populations (Kawecki and Ebert, 2004).

The integration of data from transplantation experiments with genomic data is crucial to unravel the genetic basis of specific phenotypes that underlie local adaptation. For example, Ågren and Schemske (2012), Ågren et al. (2013) and Postma and Ågren (2016) determined the genetic basis underlying freezing tolerance, as well as its adaptive significance, in the annual plant model species *Arabidopsis thaliana*, by performing reciprocal transplantation experiments together with quantitative trait loci (QTL) mapping using populations from Italy and Sweden. They identified the co-localization between freezing tolerance and survival QTLs in the Swedish genotypes at the Swedish site, whereas the Swedish genotypes at the Italian site showed a reduced fitness at these loci, supporting that the adaptive trait of freezing tolerance was not favoured in Italy (Ågren et al., 2013; Oakley et al., 2014). Notably, QTL mapping conducted with parental populations of the perennial species*Arabidopsis lyrata* from Norway and North Carolina revealed that QTL regions associated with high fitness and early flowering were linked to local alleles (Leinonen et al., 2013).

In addition to QTL mapping, genome-wide association studies (GWAS) have been widely performed to investigate local adaptation via the association of phenotypic traits and allelic variation (Ingvarsson and Street, 2011; Anderson et al., 2011). McKown et al. (2014) performed a common garden experiment together with GWAS, which included phenotyping 40 different traits, as well as genotyping 34,000 single nucleotide polymorphisms (SNPs) across more than 400 accessions of *Populus tricharpoca*. This study successfully identified that most of

the significant associations were linked to phenology traits.

It is important to note that the technique used to study local adaption is dependent on the length of the plant life cycle. For plant species with shorter generation times, QTL mapping using recombinant inbred lines (RILs) and near-isogenic lines (NILs) is more frequently employed (Anderson et al., 2011). For longer-lived, out-crossing species, it is more feasible to directly use individuals bred from populations originating from different geographic ranges to identify phenotypic and genetic variations (Sork, 2018).

1.2.2 GENE EXPRESSION STUDIES OF LOCAL ADAPTATION

Genome-wide transcriptome analysis via RNA-sequencing (RNA-seq) provides a powerful tool with which to investigate how the expression level of genes varies in response to environmental cues, which can reveal evidence of local adaptation on a molecular level (Lasky et al., 2014; Akman et al., 2016; Gugger et al., 2017). One of the greatest advantages of RNA-seq is that it allows a *de novo* transcriptome to be created even without a reference genome (Wang et al., 2009b; Grabherr et al., 2011); thus, the molecular mechanisms of local adaptation in non-model species can be studied. Furthermore, genotype-by-environment (G×E) interactions can be detected by studying gene expression levels, which highlights phenotypic plasticity (Des Marais et al., 2013). An additional and attractive feature of RNA-seq is that gene regulatory networks (GRNs) can be identified, because genes or gene clusters can be sought that are co-expressed in response to environmental cues (Langfelder and Horvath, 2008; Akman et al., 2016). Most importantly, performing the time resolved RNA-seq analysis enables us to study gene expression levels responding to environment during the developmental processes.

Lasky et al. (2014) analysed two published data sets on natural variation in *A. thaliana* and these natural variations included genomic sequence, climate of origin, field fitness, and expression response to abiotic stress from a total of 80 different accessions. By using whole genome expression profiling from different accessions, they classified genes into two categories: eSR (expression stress responsive) and eGEI (expression genotype-by-environment interaction). They further performed enrichment of these two gene categories with associations to climate variables for early- and late-flowering accessions. They showed that eSR genes are more rarely associated with cold-related climate variables compare to eGEI genes

in early-flowering accessions; while the mean temperature of the growing season was the only significant enrichment for late-flowering accessions. For eGEI genes, significant enrichment at most climate variables are only present in early-flowering accessions. Furthermore, considering three populations of California oak species *Quercus lobata Née*, Gugger et al. (2017) investigated the water-stress responses of seedlings grown under greenhouse conditions and demonstrated a strong transcriptional response to drought: 35,347 out of 68,434 contigs (52%) were differentially expressed before and after drought treatment. Genome-wide transcriptome analysis can even be performed for species with large, undocumented genomes. For example, Yeaman et al. (2016) used this approach to show that two distantly related conifers, lodgepole pine and interior spruce, displayed convergent local adaptation, and they also found that the same genes led to the same phenotypes in response to low temperatures.

Taken together these studies demonstrate that genome-wide transcriptome analysis via RNA-seq facilitates a direct investigation of the molecular mechanisms that underlie a phenotype in response to environmental cues. Thus, studying the transcriptomes of natural plant populations can identify genes that are involved in local adaptation.

1.3 Variation in flowering behaviours among plants with different life histories

The extraordinary variation observed in the flowering behaviour of plants highlights the essential question concerning how plants optimise reproduction and survival. The diversity in plant life histories can range from long-lived tree species such as *Adansonia*, to the ephemeral desert annual plant *Boerhavia* (Friedman, 2020). The earliest attempts to classify plant life history/form studied the morphology and position of meristems at the end of the growing season (Raunkiaer et al., 1934). Since then, additional terminologies have emerged to describe the seasonal life expectancy (annual, biennial, perennial), and the number of reproductive events during a plant's life-span (monocarpy/semelparity, polycarpy/iteroparity), which can better characterise a plant's life history (Kirkendall and Chr, 1985; Silvertown, 1986, 1989; Hughes, 2017).

In general, annual plants can be further divided into two types. The first type consists of truly semelparous annuals, i.e., plants that reproduce once and die,

resulting in a single reproduction event where the apical meristem on the primary shoot forms a determinate inflorescence and produces fruits. The second type contains annual plants that can flower continuously due to indeterminate growth, but die in response to adverse environmental conditions. Biennial plants germinate and remain vegetative for the first year before flowering and die the subsequent year. By contrast, perennial plants, due to their iteroparous nature, can alternate between periods of vegetative growth and flowering over the course of many years (Friedman, 2020). Intriguingly, a great degree of variation has been observed in the timing of the transition to the reproductive phase in perennial species. Several herbaceous perennial plants can undergo floral transition in the first year, whereas many tree species need to undergo long juvenile periods prior to floral transition. Some perennial species use clonal reproduction to generate new ramets to maintain perenniality, whereas others continue growing for many years and then die after a single but massive reproductive episode (Young and Augspurger, 1991; Friedman, 2020). These examples illustrate the diversity in life histories and the variation in flowering behaviour that result from the function of complex genetic networks in response to environment cues.

1.3.1 TRADE-OFFS BETWEEN FLOWERING AND VEGETATIVE GROWTH

The allocation of resources between vegetative and reproductive growth is extremely important to ensure plant fitness. Studies have shown that delayed reproduction correlates with increased vegetative growth in plants (Delph and Meagher, 1995; Rubin et al., 2019), supporting the notion that reproduction is generally negatively correlated with traits related to life history, to account for seasonal trade-offs (Obeso, 2002). Nonetheless, environmental factors may modify these trade-offs, as shown by a study of *Lychnis flos-cuculi*, which reported that early flowering is linked to a decrease in subsequent fecundity; however, this link is site-specific (Biere, 1995). Furthermore, a three year common garden experiment from transplants of the sub-alpine perennial *Boechera stricta* also indicates that the probability of flowering opposes fruit length amongst plants that flowered under certain environmental conditions (Wadgymar et al., 2017). Collectively, delaying flowering can lead to higher seed production, but the plant may die before forming any seeds if the environment becomes unfavourable due to drought or onset of winter. By contrast, early flowering ensures reproduction but result in low number

of seeds.

The mechanistic trade-off between the function of primary shoot apical meristems and axillary meristems highlights the association between life history and plant architecture. The flexible nature of meristems enables the development of either vegetative or reproductive organs at a specific timepoint. In alpine perennial *Erysimum capitatum*, additional vegetative growth in the pre-reproductive stage impacts survival and subsequent reproduction (Kim and Donohue, 2011). Notably, in *Mimulus guttatus*, differences in plant shoot architecture between annual and perennial accessions is linked to the fate of axillary meristems on different nodes (Baker and Diggle, 2011). To investigate the trade-offs between vegetative and reproductive growth in annuals and perennials, the use of QTL mapping is effective, and QTLs associated with both the flowering time and vegetative growth of *Mimulus guttatus* have been identified as being pleiotropic and epistatic (Friedman, 2014; Friedman et al., 2015). Also, in *Arabidopsis lyrata*, QTLs have been shown to modulate resource allocation by affecting early vegetative growth patterns (Remington et al., 2013). In summary, plants coordinate the fate of different meristems as a mechanism to regulate the trade-offs between vegetative growth and reproduction.

1.3.2 Flowering behaviour in annuals and perennials

Trade-offs between between vegetative growth and reproduction can easily be observed in the flowering behaviour of annuals and perennials. Although annuals and perennials both transition from vegetative growth to flowering, annuals only flower once before undergoing senescence and completing their whole life cycle, whereas perennials can flower multiple times by alternating between vegetative and reproductive stages (Figure 1.1) (Friedman, 2020). Thus, examining the genetic and molecular basis that is responsible for the transition between vegetative growth and flowering in the context of plant life histories is extremely important in evolutionary biology (Taylor et al., 2019). Using examples of annuals and perennials, I will first underline the genetic and molecular mechanisms of how seasonal flowering behaviour responds to photoperiod; secondly, I will highlight the important role of competence to flower; finally, I will summarise the genetic and molecular mechanisms that underlie how seasonal flowering behaviour responds to vernalization.

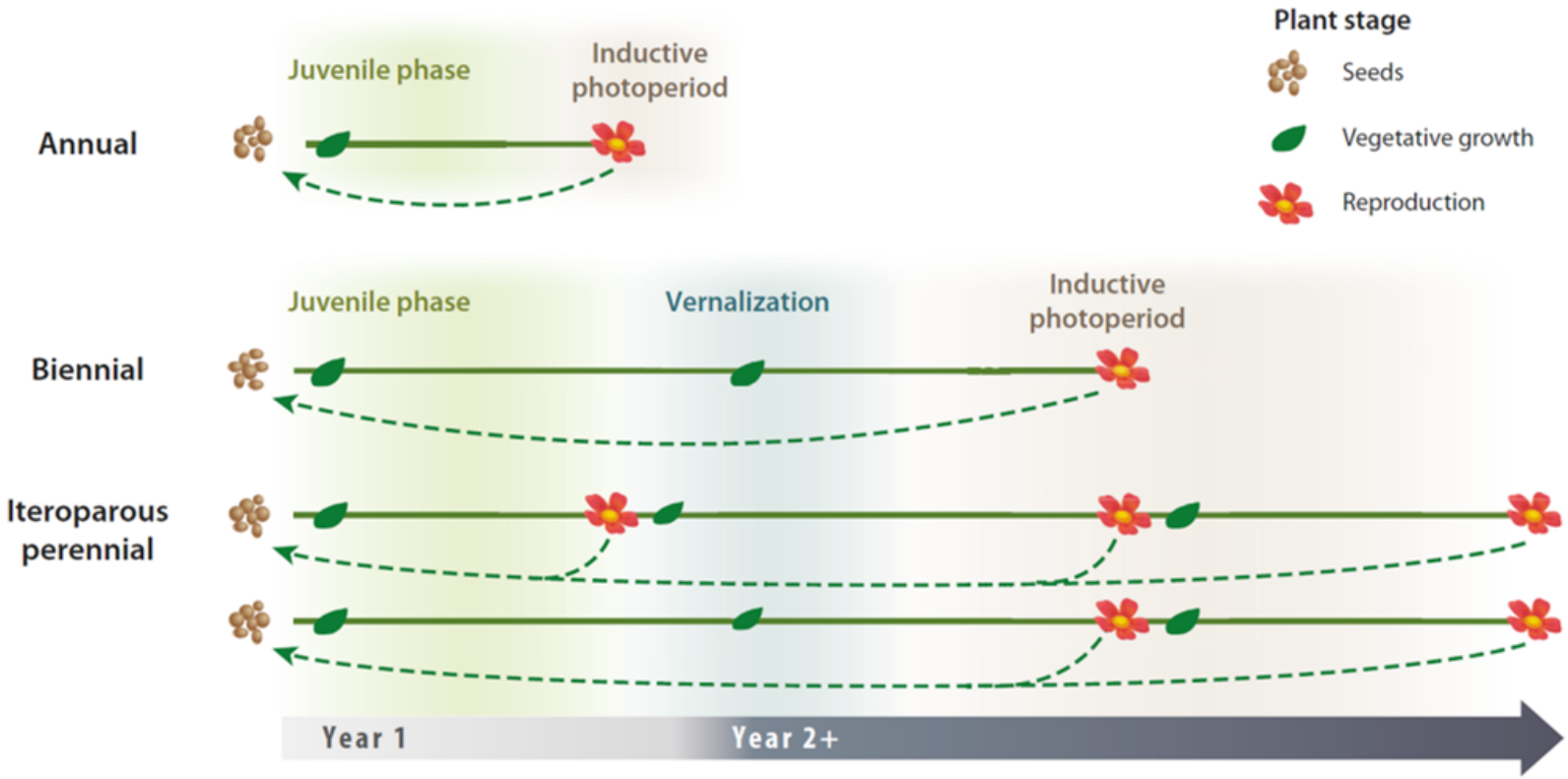

Figure 1.1. Conceptual diagram of the difference in life cycle for annual, biennial, and iteroparous perennial plants. Plants in the vegetative stage are shown with green lines, with a leaf indicating vegetative growth; flowering is shown with orange flowers; and seeds are shown on the left in brown. The dashed green lines indicate the production of seeds that return to start the life cycle anew. Figure is from Friedman (2020).

Photoperiod

Many plant species, especially in temperate zones, perceive differences in the daylength as an environmental cue to flower at the optimal time of year (Andrés and Coupland, 2012).The molecular mechanism for how photoperiod affects flowering time has been well studied in annual model species such as *Arabidopsis thaliana*. In *Arabidopsis thaliana*, the photoperiodic pathway responds to differences in the daylength and promotes flowering during long days. Daylength is first sensed in the leaves by photoreceptors and conveyed through the circadian clock and under inductive long days, CONSTANS (CO) activates transcription of *FLOWERING LOCUS T* (*FT*) (Suárez-López et al., 2001; Valverde et al., 2004; An et al., 2004). Following this activation, FT protein (often referred to as florigen) translocates from the leaves through phloem sieve elements to the shoot apical meristem. Here, the FT protein interacts with the basic leucine zipper (bZIP) transcription factor FD, leading to the formation of a protein complex and the promotion of flowering (Abe et al., 2005; Wigge et al., 2005; Corbesier et al., 2007; Jaeger and Wigge, 2007; Mathieu et al., 2007; Turck et al., 2008; Romera-Branchat et al., 2020). In addition to studies on *Arabidopsis thaliana*, homologs of the *CO* and *FT* genes have been studied in a variety of annual, perennial, and woody dicots and monocots species, and the findings imply that the genetic photoperiodic flowering pathway is conserved and ancient (Andrés and Coupland, 2012; Kemi et al., 2019). Despite the widespread dependence on photoperiod as an environmental cue to promote flowering amongst annual and perennial species, flowering behaviour is further regulated by a plant's competence to flower (age) and/or exposure to the cold (Friedman, 2020).

Competence to flower

Some plants acquire a competence to flower in response to environmental cues as they age. Young juvenile plants do not flower even if exposed to inductive environmental cues, whereas adult plants acquire competence to respond to these cues (Hyun et al., 2017). The length of the juvenile phase differs among annuals and perennials: annuals generally have shorter juvenile phases so that flowering occurs rapidly in response to environmental cues, whereas perennials display delayed flowering, which allows for the sufficient accumulation of biomass and the production of numerous axillary meristems via branching prior to the reproductive stage. This

delayed flowering ensures better survival rates during the perennial life cycle stage (Bergonzi et al., 2013).

The genetic pathways that regulate when plants transition from the juvenile to adult stage are associated with SQUAMOSA PROMOTER BINDING PROTEINLIKE (SPL) transcription factors and regulatory microRNAs (miRNAs) in both annuals and perennials (Hyun et al., 2019). In young plants, miR156 is highly expressed in leaves and in the shoot apical meristem. However, with increasing age, the expression of miR156 gradually decreases, removing the suppression of expression of SPL transcription factors and thus allowing floral transition (Bergonzi and Albani, 2011; Bergonzi et al., 2013; Hyun et al., 2017; Madrid et al., 2021a). Notably, Morea et al. (2016) demonstrated that this miR156/SPL system is conserved across multiple land plant families and appears to be ancient. Furthermore, the long juvenile phases and delayed flowering commonly observed in perennials are associated with the gene *TERMINAL FLOWER 1* (*TFL1*): Kotoda et al. (2006) and Mohamed et al. (2010) showed that flowering of transgenic *Malus domestica* and *Populus trichocarpa* plants with reduced TFL1 function was accelerated due to a shortening of the juvenile phase.

Vernalization

Vernalization is the process whereby flowering in plants is induced by exposure to a prolonged period of low temperatures and it plays an essential role in ensuring the optimal flowering time in natural plant populations. Moreover, the vernalization pathway was proposed to have evolved several times during the diversification of angiosperms (Amasino, 2010). The earliest physiological study on vernalization was performed with the biennial *Hyoscyamus niger* and showed that vernalization was required to ensure that flowering occurred in the second year (Melchers, 1937).

Several cereal crops species in the Poaceae, such as wheat and barley, only flower after overwintering. In these species, flowering in response to winter cold is mainly associated with the CCT protein VERNALIZATION 2 (VRN2), which inhibits the induction of *FT1* under long day conditions and thereafter suppresses the flowering before winter (Yan et al., 2004). However, exposure to cold activates the expression of *VERNALIZATION 1* (*VRN1*) and thus promotes inflorescence initiation at the shoot as well as represses transcription of *VRN2* (Andrés and Coupland, 2012; Fjellheim et al., 2014). In addition, Eagles et al. (2009) has

demonstrated that different vernalization requirements for wheat cultivars in different natural environments are related to allelic variation in *VRN1* and *VRN2*, which implies local adaption can occur in vernalization genes.

Likewise in *A. thaliana*, studies reported that different vernalization requirements to regulate the floral transition are linked to allelic variation at vernalization genes in different populations (Stinchcombe et al., 2004; Méndez-Vigo et al., 2011). Flowering is suppressed by up-regulating expression of *FLOWERING LOCUS C* (*FLC*) mediated by FRIGIDA in vernalization-sensitive ecotypes of *A. thaliana* (Amasino, 2010). Winter cold epigenetically silences *FLC* through histone modification (Bastow et al., 2004; Sung and Amasino, 2004). This promotes the formation of FRIGIDA nuclear condensates, which no longer co-localises with an active *FLC* locus and reduces *FLC* repression. The transcriptional shutdown of *FLC* in the cold temperatures enables the occurrence of floral transition at the shoot apical meristem (Johanson et al., 2000; Zhu et al., 2021). Nonetheless, in many perennial relatives of *A. thaliana* such as *Arabidopsis halleri*, *Arabidopsis lyrata* as well as other species in the tribe Arabideae, the expression of orthologs of *FLC* are only temporally reduced by the cold and rise again when warm temperatures arrive, underlying the perennial cycling of vegetative and flowering stages (Wang et al., 2009a; Aikawa et al., 2010; Kemi et al., 2013; Castaings et al., 2014; Kiefer et al., 2017).

1.4 Cycling between vegetative growth and flowering is determined by meristem fate in perennials

An essential strategy of the perennial life cycle ensures that not all meristems on the plant commit to flowering, such that cycling between vegetative growth and flowering can occur multiple times during the life cycle. Therefore, the reprogramming of meristems in perennials is required for subsequent flowering (Rohde and Bhalerao, 2007; Albani and Coupland, 2010). Furthermore, the location of axillary meristems determines their response to environment cues; for example, some meristems remain dormant during vernalization, whereas others develop into axillary branches (Lazaro et al., 2018; Vayssières et al., 2020).

The underlying mechanism that restricts flowering from specific meristems

of perennials may involve genes with pleiotropic functions. For example, in addition to regulating the juvenile phase, *TFL1* is associated with flowering from axillary branches, as shown by studies on perennial species such as *Lolium perenne*, *Populus trichocarpa*, *Malus domestica* as well as *Fragaria vesca* (Jensen et al., 2001; Kotoda et al., 2006; Mohamed et al., 2010; Tenreira et al., 2017). For forest tree *Populus tremula*, Bohlenius et al. (2006) showed that the flowering-time genes *CO2* and *FT1* are also involved in growth cessation and bud development in short day in autumn, and Hsu et al. (2006) reported that for *Populus deltoides*, *FT2* is not only associated with shortening the juvenile phase but also with regulating seasonal flower initiation.

In perennial species such as strawberry (*Fragaria* spp.), the trade-off between sexual reproduction and vegetative propagation is controlled via meristems on underground stolons (elongated stems). Tenreira et al. (2017) demonstrated that exogenous application of gibberellin affects this trade-off and further showed that a deletion in *gibberellin 20-oxidase* (*GA20ox*) is associated with an alteration in the fate of axillary meristems and the balance between flower and runner production. Moreover, Gould et al. (2017) conducted genomic analyses to compare annual and perennial monkeyflowers *M. guttatus* and found that *GA20ox2* serves as a major allele frequency outlier between different ecotypes. Notably, Lowry et al. (2019) applied exogenous gibberellin to perennial *M. guttatus* ecotypes, which then showed an annual-like morphology. In summary, many genes, including those involved in hormone biosynthesis, may affect the balance between vegetative growth and flowering in perennials.

1.5 *Arabis alpina* AS A PERENNIAL MODEL SPECIES FOR EVOLUTIONARY STUDIES

Arabis alpina, a related species to the annual *Arabidopsis thaliana* within the family Brassicaceae, has been studied extensively in the laboratory and has become a perennial model species, due to its characteristics of being diploid, self-fertile and tractable to transformation by *Agrobacterium tumefaciens*. The regulation of seasonal flowering in *A. alpina* is responsive to environmental signals and it has been shown that the Spanish *A. alpina* accession Pajaras only flowers following exposure to vernalization (Wang et al., 2009a; Wötzel et al., 2022).

A. alpina is broadly distributed throughout the alpine ecosystem of the

Northern Hemisphere and numerous adaptive differences amongst the *A. alpina* populations have been studied. The results from reciprocal transplantation experiments indicated that local adaptation of Spanish and Scandinavian accessions occurs (Toräng et al., 2015). Furthermore, variation in flowering time exists amongst populations from the Iberian Peninsula, the European Alps and Scandinavia (Albani et al., 2012; Toräng et al., 2015). Seeds were collected from different families originating in the various *A. alpina* populations in Europe and the progenies were phenotyped in greenhouse conditions. Spanish populations exhibited an absolute requirement for vernalization to flower, whereas individuals showing early or late onset of flowering without exposure to vernalization were identified in other populations (Wunder et al. in preparation).

1.6 Genetic interactions regulate flowering in *Arabis alpina*

The genetic flowering regulation has been well studied in the Spanish Pajares accession. Previous studies have shown that the *PERPETUAL FLOWERING 1* (*PEP1*) gene in *A. alpina*, which is the orthologue of the *A. thaliana* gene *FLOWERING LOCUS C* (*FLC*), plays a key role in linking seasonal flowering and polycarpy. During vernalization, the transcription of *PEP1* decreases dramatically, but it is reactivated when plants return to normal growth conditions. This reactivation is due to the non-persistent feature of the histone modifications of *PEP1* subsequent to vernalization; in contrast to the annual *A. thaliana*, where chromatin modifications on *FLC* are mitotically stable (Wang et al., 2009a; Friedman, 2020; Wötzel et al., 2022). The reactivation of *PEP1* after vernalization ensures continued vegetative development in *A. alpina* and confers on to it a polycarpic life history. The fates of the inflorescence and axillary branches are also regulated by *PEP1*, dependent upon their position on the shoot and the length of exposure to vernalization. Documentation of the number and the flowering behaviour of axillary branches has shown the existence of different zones along the primary shoot of the Spanish accession Pajaras; in ascending order these are: (V1) axillary branches that flower and partially senesce; (V2) dormant axillary buds; and (V3) axillary vegetative branches. Above the V3 zone is the inflorescence zone, which consists of inflorescence branches and solitary flowers (Lazaro et al., 2018; Hyun et al., 2019; Vayssières et al., 2020).

The repression of *PEP1* during vernalization does not fully describe how reproduction in *A. alpina* is limited, both spatially and temporally. This also requires the involvement of a second regulatory system centred around miR156 that controls the competence of meristems to flower, as described above. When young meristems are vernalized, the expression of *miR156* is high, which represses floral initiation and prevents flowering even though PEP1 mRNA levels are reduced. Conversely, when older meristems are vernalized, the lower level of miR156 allows them to flower when *PEP1* levels fall. Recent studies have shown that the flowering promoting factor SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 15 (SPL15) is the key to coupling the regulation of *PEP1* and miR156. SPL15 is required for flowering after vernalization, and is directly transcriptionally repressed by PEP1 as well as being a direct target for miR156. Therefore, SPL15 integrates competence to flower and vernalization response in *A. alpina* (Hyun et al., 2019; Madrid et al., 2021b).

1.7 AIMS OF THIS THESIS

The aim of this thesis is to decipher the variation in the flowering behaviour within a Norwegian *A. alpina* population, where early- and late-flowering phenotypes were identified without vernalization and to test if these differences in flowering behaviour in controlled conditions are relevant in the natural population. This variation in flowering behaviour without vernalization differs greatly from the Spanish accession Pajares, which shows an obligate requirement for vernalization. While genetic pathways of flowering have been well studied in the Spanish accession Pajares, the underlying flowering mechanisms in Norwegian accessions remain unclear, particularly since the differences in flowering time seem to depend on the rate of floral bud development rather than the timing of floral induction.

Here, I use four different approaches to investigate the underlying mechanisms leading to the expression of the early- and late-flowering phenotypes within a Norwegian population of *A. alpina*. Greenhouse/growth chamber experiments with an accompanying detailed phenotypic analysis is used to examine how multiple reproductive traits and plant architecture respond to different vernalization and pre-growth periods by simulating short growing seasons and long snow coverage. Genome-wide transcriptome analysis via RNA-sequencing is performed to examine whether different gene regulatory networks exist between the early- and

late-flowering phenotypes. A QTL mapping analysis using different near-isogenic lines is conducted to examine the genetic basis of the variation in floral bud development between early- and late-flowering phenotypes. An *in situ* transplantation experiment at the Norwegian Riasten field site is used to examine whether differences in performance are detected via analysing different biological traits in the plants' natural environment.

Chapter 2
Material and methods

2.1 Study population

Arabis alpina (L., Brassicaceae) is considered to be one of the most important model organisms with which to study the perennial life cycle, including flowering behaviour in response to winter cold (Wang et al., 2009a). Because of its wide geographical distribution, *A. alpina* provides additional opportunities for studying ecological genomics as well as life-history evolution. Its natural habitats range from scree slopes, rock ledges to small streams in alpine ecosystem (Wötzel et al., 2022).

The Riasten study population is situated in Norway (62°49´48 to 62°50´09 N and 11°44´41 to 11°44´42 E). The population is distributed between two main valleys and many plants grow on the scree slope adjacent to the valley. The growing season is usually fewer than 4 months (June to September), which means that seedlings germinate, flower and produce siliques within a relatively short period. Several families from this population were collected from the field and phenotyped in the greenhouse, where two inbred families were chosen to represent early and late-flowering accessions.

2.2 Plant cultivation in greenhouse and growth chamber

Seeds were stratified on filter paper at 4°C in darkness for 4 days. Plants were grown in a Bronson growth chamber or greenhouse at a light intensity of

about 150 µmol m^{-2} s^{-1} under LD conditions (16 h light/8 h dark) at 21°C/18°C and 65% humidity. Additional far-red light (100% in the setting) was added to the spectrum in the growth chamber. For vernalization treatment, adult plants were transferred to 4°C under SD conditions (8 h light /16 h dark) in Chamber 1-E16 of Greenhouse 1. After exposure to vernalization, plants were returned to LD conditions in the growth chamber or greenhouse.

2.3 Generation of plant cuttings and crosses

Plant cuttings were taken every 5 months form the field-collected early (S4-41-2) and late (S4-62-8) accessions in the greenhouse to maintain the material of parental plants used at the start of QTL mapping. Five vegetative branches were always selected from a mother plant and the majority of the large leaves were removed to leave one or two small leaves attached to the branch. The base of each branch was dipped in rooting powder and planted in soil.

Crosses were conducted using young floral buds in which pollen had not dehisced from the pollen sac. Sepals, petals and stamen were removed from recipient flowers with tweezer. The stamen from the youngest open flowers were selected as pollen donors and pollen was transferred from these flowers by brushing the anthers over the stigma of recipient plants several times. The identification (ID) numbers of the corresponding donor and recipient plants were recorded in a scoring sheet and the pollinated flowers were individually tagged with this number. The crosses were examined daily until siliques started to develop and all remaining floral buds were removed when silique development was completed on each branch. The branches were covered with a plastic bag labelled with the branch ID and the seeds were harvested according to branch.

2.4 Phenotypic description of plants in the greenhouse and growth chamber

The phenotypic description of plants in the greenhouse and growth chamber included the flowering time of the primary shoot and each axillary branch, the height from the base of plant to the tip of the primary shoot, the number of flowers and siliques on the primary shoot, the length of each axillary branch, the number

of flowers and siliques on each axillary branch, and the developmental stage of each axillary branch. The flowering time was defined as when the first flower opened; the developmental stage of each axillary branch was performed by examining the individual axillary meristem under a magnifying lens in Halle 14; the plant height and length of the axillary branches were measured using a ruler.

2.5 Phenotypic description at the Norwegian field site

The phenotypic description of labelled natural plants and plants in the *in situ* transplantation experiment was performed separately. Individual natural plants in the field were labelled by placing a unique plant ID number on an orange plastic pin next to the corresponding plants (left panel of Figure C.1). The phenotypic descriptions of these labelled plants included survival, the number of rosettes, the developmental stage of each rosette (vegetative or with visible floral bud), the number of elongated flowering shoots, and the number of flowers and siliques on each flowering shoot. All labelled natural plants were photographed at each visit.

The phenotypic description of plants in the *in situ* transplantation experiment included rosette size, the number of rosettes, the developmental stage (vegetative or with visible floral bud) of each rosette, the number of elongated flowering shoots, the number of flowers and siliques on each flowering shoot, and the length of each flowering shoot and the plant height. The scoring of flowering was conducted separately for the primary shoot and axillary branches to distinguish the new flowering shoot produced in the current year. Notably, the first flowering event scored for each plant was always from the primary shoot. Individual plants in the *in situ* transplantation experiment were labelled by placing a unique plant ID number on the green flag attached to the metal pin next to the corresponding plants (right panel of Figure C.1).

2.6 RECORDING OF ENVIRONMENTAL VARIABLES IN GREENHOUSE, GROWTH CHAMBER AND THE NATURAL ENVIRONMENT

Temperature and light intensity were measured for every experiment using HOBO Pendant Temperature/Light 64K Data Logger (Part # UA-002-64). Probes for aerial temperature measurement were protected from direct sunlight in the greenhouse and growth chamber by well-ventilated plastic cases; loggers were protected by a plastic shield in the natural environment to avoid damage from snow. Data from the loggers were recorded three times from the beginning to the end of each experiment to ensure continuous data collection; loggers in the natural environment were read at the first and last visits to the field site with the HOBO waterproof shuttle.

2.7 GIBBERELLIN APPLICATION

Plants were sprayed weekly with 20 mM GA_3, GA_{4+7} (Sigma-Aldrich; stock solution:100 mM in ethanol and 0.1% [v/v] Silwet) or mock (0.1% [v/v] ethanol and 0.1% [v/v] Silwet) after plant germinate (Tilmes et al., 2019). Each treatment included six plants that were fully randomised. Solutions were sprayed onto the SAM of each plant on Monday afternoon between 16:00 and 17:00 to ensure consistency. Plant height was measured and plant architecture was photographed on a weekly basis..

2.8 PLANT DNA EXTRACTION, POLYMERSE CHAIN REACTION AND GENOTYPING

Small portions of two young leaves were harvested and used as two replicates to ensure a sufficient quantity of extracted DNA for genotyping. Samples were frozen at -80°C until use. DNA was extracted from up to 96 samples simultaneously using the Qiagen BioSprint robot and the BioSprint 96 DNA Plant Kit (Qiagen, No.941557), following the manufacturer's instructions. The DNA was elutedin 100 L milli-Q water.

Reagent	µL/sample
dH_20	3.95 µL
5×Green GoTaq Flexi Buffer	2 µL
$MgCl_2$ (25mM)	1 µL
dNTPs (10µM)	0.4 µL
FW1 primer	0.2 µL
Rev1 primer	0.2 µL
Rev2 primer	0.2 µL
GoTaq Flexi DNA Polymerase	0.05 µL
Template DNA	2 µL

Table 2.1. PCR mixture for GoTaq PCR.

PCR step	Temperature	Time	Cycle
Initiation	98°C	2 minutes	1
Denaturation	98°C	30 seconds	35
Annealing	55°C	30 seconds	35
Extension	72°C	30 seconds	35
Hold	4°C	∞	∞

Table 2.2. PCR programme for GoTaq PCR.

Plants were genotyped with all InDel markers (see Table 2.3) using GoTaq DNA Polymerase (Promega, M8291) in 10 µL of reaction mixture. The reagents of reaction mixture are listed in Table 2.1 and the PCR programme is described in Table 2.2. The resulting amplified DNA was visualised on a 1.5% agarose gel.

2.9 Genetic markers used in QTL mapping

All designed genetic markers used for fine-mapping experiments are listed in Table 2.3 and Table 2.4.

2.10 RNA extraction

Total RNA was extracted from apices harvested at Zeitgeber Time 5 (ZT5) using the RNAeasy extraction kit (QIAGEN), treated with Turbo DNase (Invitrogen) and RNA Clean & Concentrator kit (Zymo Research). Apices were harvested in four batches of 12 plants from both early- and late-flowering accessions. The RNA pooled from the apices of each batch was considered as a biological

InDelID	Primer ID	Sequence (5' to 3')
AL925	Fw1	CCAAGTTTACACTGCGGGAG
	Rev1	CGTTTTGAGTTTTGGAGTTTC
	Rev2	GCGAAACCCTCCTAGATCCG
AL922	Fw1	AATAACCCGAGGAGGCCAAG
	Rev1	CTTGTTGTCTACCGCCATGC
	Rev2	TGAGCCATGCTTGTGACCTC
AL921	Fw1	GAGATCCCTTGCATGAGACCA
	Rev1	GCCATATCTAAGTCGCGTGGA
	Rev2	ACGAAGTAGTGACAGACGATC
AL920	Fw1	GGTTATTTTTCCTATCCTAGGTTTGT
	Rev1	ACTTTCCCGGAAAACAGAGGA
	Fw2	AGATACGACGAACCGAACGT

Table 2.3. InDel primers used for fine-mapping.

replicate and three out of four batches that had the highest final RNA concentrations were used for RNA-seq library preparation. The library preparation and sequencing were performed at the Max Planck Genome Center, Cologne, Germany (https://mpgc.mpipz.mpg.de/home/). The libraries were prepared from 1 mg total RNA using the TruSeq RNA kit (Illumina) and 150 bp single-end reads were sequenced on a HiSeq2500 (Illumina) system.

2.11 PROCESSING RNA-SEQUENCING RAW READS

Raw RNA-seq single-end reads were pre-processed by removing potential adaptor sequences using cutadapt (Martin, 2011) and trimming low-quality ends using Trimmomatic (Bolger et al., 2014). Processed RNA-seq reads were mapped onto the Norwegian *A. alpina* early accession reference genome (Kemi et al., in prep) using TopHat2 with additional parameters (Kim et al., 2013). The following additional parameter settings were used for the mapping: maximum edit distance 5; maximum mismatches 5; minimum intron length 50; maximum intron length 10,000; read gap length 2. The read counts per gene were summarized into an initial count table using HTseq-count (Anders et al., 2015). All genes in the count table with a rowSums of less than 10 were removed. The count table was used for further analysis with DESeq2 (Love et al., 2014). The processing of RNA-sequencing raw reads was performed in collaboration with Dr. Edouard Severing.

SNP ID	Primer ID	Sequence (5' to 3')
SNP-174	YL9	AACGGACGCCTTCTTCTCTG
	YL10	ACAAAGATGTTGGTTGTGATCC
SNP-175	NS35	GGCTTGTAAACTGCTAGTTCATCC
	NS36	CGCCTTGTGGTACATGTAGCTC
SNP-176	NS37	GTTCAACGCATATTAATTGACTTTACTG
	NS38	CACCAGACTGGATAAATCTCAAATG
SNP-177	NS39	CCACTGGTCAAATAACACCTAAAGC
	NS40	GGTCTAAGCCTTTACGTTGAGG
SNP-178	NS41	CTAGAAGACGTCAACGTGATTGAC
	NS42	CAAGTGTTCATGTGATACGCATG
SNP-179	NS41	CTAGAAGACGTCAACGTGATTGAC
	NS42	CAAGTGTTCATGTGATACGCATG

Table 2.4. SNP primers used for fine-mapping.

2.12 Co-expression network and Gene Ontology enrichment analysis

Co-expression network analysis and validation of cluster number were originally coded in R by Dr. Edouard Severing and were then modified for plotting by me. Plotting for genome-wide cluster structue in Figure 4.10 was performed with the *WGCNA* package. The validation of cluster number was assessed by the separability metric defined by Horvath (2011). Gene ontology enrichment analysis was performed by *topGO* package; redundant GO terms were removed and visualise as treemaps by Revidgo (Supek et al., 2011).

2.13 Synteny analysis

For synteny analysis, nucleotide alignments between each sequence pair compared were generated using LAST (Kiełbasa et al., 2011). The initial alignment blocks were further processed using SyRI (Goel et al., 2019), which resulted in a minimum set of alignments that reflected colinearity or rearrangements between the sequences. Synteny analysis was performed by Dr. Edouard Severing.

2.14 STATISTICAL ANALYSIS AND DATA VISUALISATION

Most data visualisation and statistical calculations were conducted in R-studio (version 1.1.463; https://rstudio.com/) using packages: *ggplot2*, *plyr*, *dplyr*, *tidyverse*, and *FSA*. Only non-parametric tests were performed in this study and the purposes of applying different non-parametric tests are mentioned in corresponding sections.

Generalised linear modelling was performed using the *lme4* package; retrieving model parameters with the *parameters* package; understanding the fitted effects with the *effects* package; visualising model with the *visreg* and *sjplot* package; model checking with the *performance* package; analysing differences among factor levels with the *modelbased* package; residual diagnostics with the *DHARMa* package.

Correlation matrices in section 3.3.8 and heatmaps for the phenology data in section 6.2 were conducted in Python 3.7.9 by using *numpy*, *scipy.stats*, *matplotlib.pyplot*, *pandas* and *seaborn* packages.

Chapter 3
Phenotypic variation in flowering behaviour of the Norwegian Riasten population

3.1 Background

Arabis alpina harbours a wide geographical distribution which has resulted in a large variation of phenotypes due to genetic variation amongst the different populations. A variation in flowering time exists within and among *A. alpina* populations from the Iberian Peninsula, the French Alps and Scandinavia. While Spanish populations exhibit an absolute requirement for vernalization to flower, individuals showing early and late onset of flowering without vernalization have been identified in French Alps and Scandinavian populations under greenhouse conditions (Wunder et al., in preparation).

Figure 3.1 shows a histogram of the flowering time amongst 383 individuals germinated from field-collected seeds sampled from 53 mother plants in the Norwegian Riasten population (Wunder et al., in preparation). The distribution of flowering time is bimodal: the first mode includes $\sim$ 44% of the data points with a mode between 60 and 90 days after germination; the second mode includes $\sim$ 17% of the data points with a mode above 200 days after germination (this also includes some individuals which did not flower before the end of experiment). One individual was chosen from both the first and second mode to represent the early-

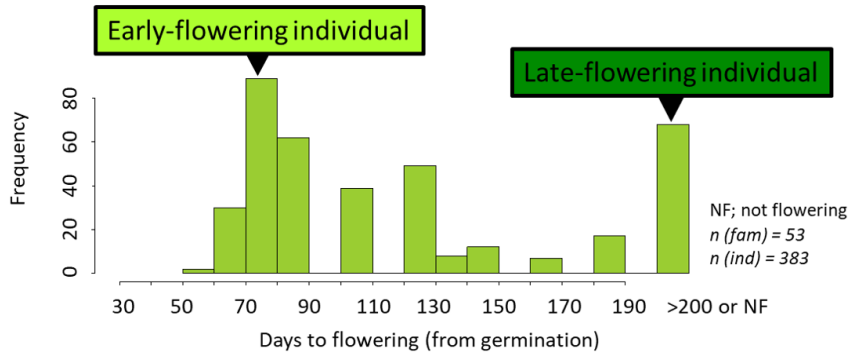

Figure 3.1. Identification of early- and late-flowering individuals from the Norwegian Riasten population. The histogram illustrates the distribution of days to the first open flower without vernalization from field-collected seeds under greenhouse conditions and the frequency represents the number of individuals. The seeds were collected from 53 mother plants in the field site and 383 individuals were scored in this experiment. Data from Wunder et al. (in preparation). The plot was originally provided by Dr. Ulla Kemi and modified by Yi-Chen Lin.

and late-flowering behaviour within this population, respectively. These two individuals were then self-fertilised for three generations where we categorised them as early- and late-flowering accessions (Kemi et al., in preparation).

The early- and late-flowering accessions, grown under greenhouse conditions, were further examined microscopically by Dr. Ulla Kemi as shown in Figure 3.2. This approach identified the bud-arrest phenotype in the late-flowering accession without vernalizaton. In the upper panel of Figure 3.2, bud-arrest phenotype was apparent in the progressive microscopic examination of the shoot apex which not only shows the slow growth of the floral bud after the transition (between 49 and 56 days after sowing) but also the lack of elongation of the primary shoot until the end of experiment. In contrast, the early-flowering accession produced an elongated inflorescence after the transition, showing a completed process from floral induction to inflorescence and flower formation (Kemi et al., in preparation).

To determine how the early- and late-flowering accessions respond to the winter cold, Dr. Ulla Kemi applied a 12-week vernalization treatment to 7-week old plants from both accessions (the lower panel of Figure 3.2). Both accessions produced open flowers after this treatment; however, flowering of the primary shoot of the late-flowering accession occurred approximately 20 days later than that of the early-flowering accession and several axillary branches from the late-flowering accession flowered prior to its primary shoot (154 days after sowing). By contrast,

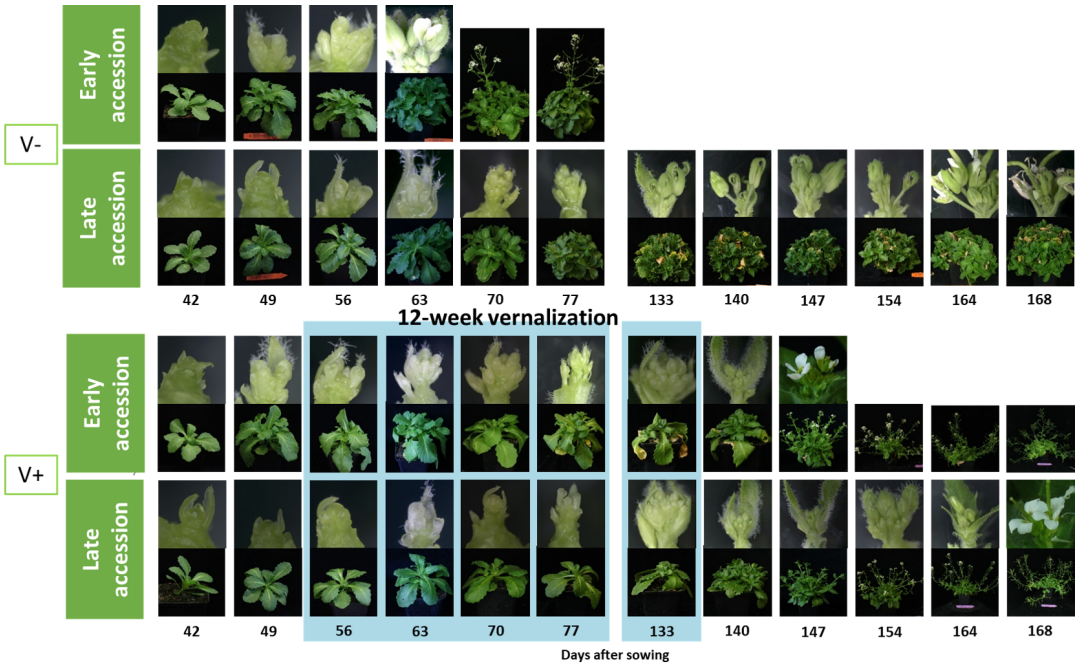

Figure 3.2. Phenotypic differences in early- and late-flowering accessions under greenhouse conditions. Microscopic examination of shoot apex and photographs of whole plant architecture were performed in the early- and late-flowering accessions under greenhouse conditions. The upper panel illustrates phenotypes of both accessions without vernalization; the lower panel illustrates phenotypes of both accession exposed to vernalization for 12 weeks (the blue background represents the vernalization period). The upper row in each panel indicates the early-flowering accession and the lower one indicates the late-flowering accession. The numbers beneath the images represent the days from sowing. Data from Kemi et al. (in preparation).

after a 12-week vernalization, flowering of the primary shoot and axillary branches in the early-flowering accession was nearly synchronous (Figure 3.2). Such variation in flowering behaviour on the primary shoot and axillary branches between early and late-flowering accessions suggests different regulatory mechanisms control flowering of these accessions after vernalization. To further explore these different mechanisms, I performed phenotypic analysis of both accessions under greenhouse and growth chamber conditions with and without vernalization. In this chapter, firstly, I will present the results of my phenotypic analysis of both accessions under greenhouse conditions in section 3.2, and, secondly, the phenotypic analysis of plants under growth chamber conditions in section 3.3.

3.2 PHENOTYPIC ANALYSIS UNDER GREENHOUSE CONDITIONS

3.2.1 EXPERIMENTAL DESIGN

As mentioned in section 3.1, differences exist in the flowering time between the primary shoot and axillary branches after a 12-week vernalization treatment between early- and late-flowering accessions. To investigate how vernalization (4°C in 8h light: 16h dark) affects the dynamics between the primary shoot and axillary branches of both the early- and late-flowering accessions, I applied 4 different vernalization periods (8, 12, 16 and 20 weeks) to three, 7-week old individuals from the early- and late-flowering accessions grown under LD (16h light: 8h dark) in the greenhouse. After vernalization, plants were tranferred back to LD condition in the greenhouse, and shoot architecture was characterised by measuring the height of the primary shoot from base to the shoot apical meristem (SAM) and the time until the opening of the first flower two-times per week for one month after vernalization. All of the data shown in the following subsections represent three plants per treatment.

3.2.2 EXTENDED VERNALIZATION RESCUES PRIMARY SHOOT ARREST PHENOTYPE IN THE LATE-FLOWERING ACCESSION

Figure 3.3 illustrates the height of the primary shoot of plants of the early- and late-flowering accessions after different vernalization periods. Also shown is

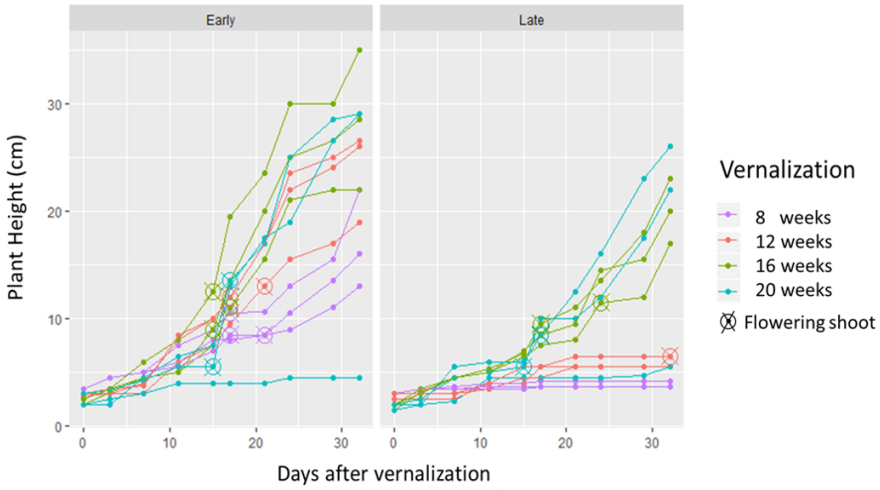

Figure 3.3. The height of plants from the early- and late-flowering accessions vernalized for 8, 12, 16 or 20 weeks under greenhouse conditions. The left and right panels represent plants from the early- and late-flowering accessions, respectively. Each line represents an individual plant and the colour of the line denotes the length of vernalization. Circles with crosses indicate the time at which the first flower on the primary shoot opened.

the time at which each plant flowers from the primary shoot. After 8 weeks of vernalization (purple lines), only the early-flowering accession produced elongated inflorescences from the primary shoot , resulting in a final mean height of 17.0 cm (column 1, Table 3.1). By contrast, the primary shoot of the late-flowering accession did not elongate and did not flower. After 12 weeks of vernalization (red lines), the late-flowering accession still presented short primary shoots which were only slightly elongated, as seen in column 2 of Table 3.1. Flowering did occur in plants of late-flowering accession, but was confined to the axillary branches; only one individual plant produced non-fertile flowers from the primary shoot at the end of experiment. After 16 weeks of vernalization (green lines), primary shoot of the late-flowering accession elongated and flowered, demonstrating that the bud-arrest phenotype was rescued by extended vernalization under greenhouse conditions. The mean final heights of both accessions were similar as indicated in column 3 of Table 3.1. Intriguingly, when both accessions were exposed to 20 weeks of vernalization (blue lines), one individual from each accession displayed the short primary shoot phenotype and did not produce open flowers. Therefore, although extended vernalization (16 weeks) rescued the bud-arrest phenotype in the late-flowering accession, excessively long periods of vernalization (20+ weeks)

may cause the (re-)emergence of the short shoot, non-flowering phenotype in both accessions. Nevertheless, generally longer vernalization treatments promoted shoot elongation and flowering, particularly in the late-flowering accession.

	8-wk vern	12-wk vern	16-wk vern	20-wk vern
Early accession	17.0±3.7	23.8±3.4	28.5±5.3	20.8±11.5
Late accession	3.8±0.2	5.8±0.4	20.0±2.4	17.8±8.8

Table 3.1. The height at primary shoot (cm) for each accession one month after 8, 12, 16 or 20 weeks of vernalization. Values are mean±SD.

3.2.3 LONGER VERNALIZATION DECREASES THE FLOWERING TIME OF AXILLARY BRANCHES

As mentioned in section 3.1, differences were observed between the flowering of the primary shoot and axillary branches in the late-flowering accession after 12 weeks of vernalization so that many axillary branches flowered earlier than the primary shoot. To capture the influence of different vernalization periods on the flowering time of axillary branches, I labelled all axillary branches and scored the time to the first open flower on each individual labelled branch. The results of this are shown as a series of box plots in Figure 3.4. Mean number of flowering axillary branches for each accession after 8, 12, 16 and 20 weeks of vernalization are summarised in Table 3.2.

A decrease in the median time until flowering was observed for axillary branches of both accessions as the vernalization period increased. Although this was seen in both accessions, the late-flowering accession showed a stronger decrease: the median of flowering time reduced by 14 days when vernalization was increased from 8 weeks to 20 weeks for the late-flowering accession whereas the reduction was only 7 days for the early-flowering accession. Therefore, from the current sample size, I observed a linear relationship between median flowering time in axillary branches and vernalization length in both accessions under greenhouse conditions.

	8-wk vern	12-wk vern	16-wk vern	20-wk vern
Early accession	11.0±2.0	11.0±3.0	11.3±1.5	9.3±1.5
Late accession	6.3±2.3	12.6±1.5	12.6±1.5	13.0±1.0

Table 3.2. The number of flowering axillary branches for each accession one month after 8, 12, 16 or 20 weeks of vernalization. Values are mean±SD.

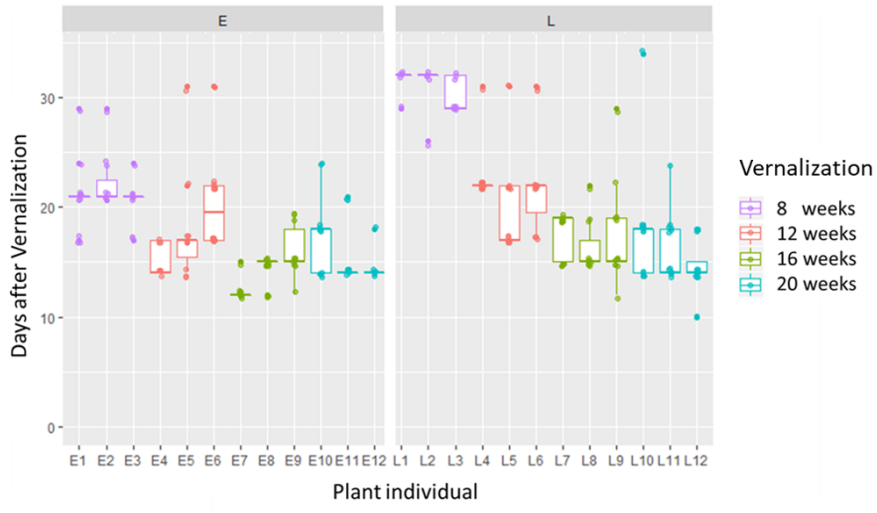

Figure 3.4. Flowering time of axillary branches of plants of early- and late-flowering accessions exposed to 8, 12, 16 or 20 weeks of vernalization under greenhouse conditions. The box plots show the distribution of the time to the first open flower on each axillary branch for plants from the early- (left) and late-flowering (right) accessions. Each box plot represents an individual plant and the colour denotes the length of vernalization.

3.2.4 Longer vernalization promotes growth in the late-flowering accession

The outgrowth of axillary branches in *A. alpina* was proposed to contribute to its perennial life-history, ensuring that flowering occurs from the axillary meristem in the subsequent year (Lazaro et al., 2018; Hyun et al., 2019). To investigate how the outgrowth of axillary branches responds to various vernalization periods, I dissected each plant, and measured the length of the axillary branch at each node. The dissection was performed on three, 7-week old plants from each accession followed by four different vernalization periods and then the measurement of axillary branches was conducted one month after vernalization. The results are shown in Figure 3.5 and the raw data are shown in Appendix Table A.1 to A.4.

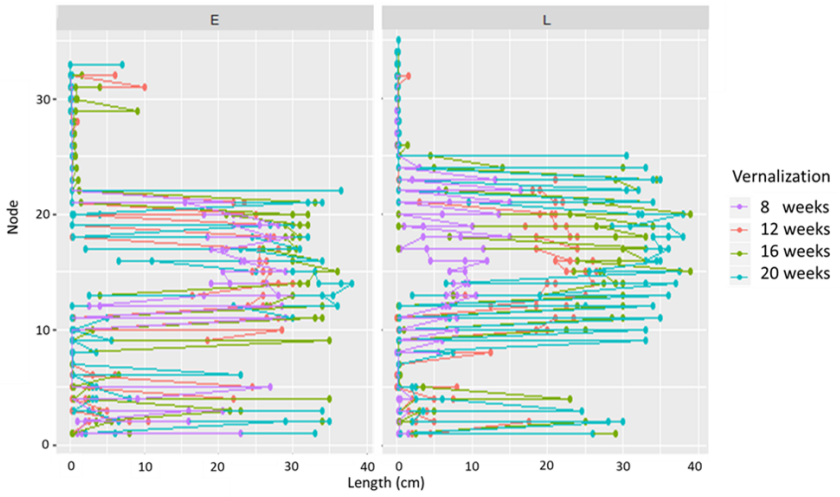

Figure 3.5. The length of axillary branches from each node for 7-week-old plant from early- and late-flowering accession vernalized for 8, 12, 16 or 20 weeks under greenhouse conditions. The four colours represent the four different vernalization periods and each line denotes an individual plant.

The application of four different vernalization periods to early-flowering accession plants did not cause any obvious effects (the left panel of Figure 3.5); rather, the outgrowth of axillary branches in the early accession appeared insensitive to vernalizaiton period. However, a clear effect was observed on the late-flowering accession plants (the right panel of Figure 3.5). Out of the four vernalizaiton periods, the distribution of axillary branch lengths was shorter for the 8 weeks vernalization set (purple lines). The longest axillary branches lay between nodes 10 and 25, similar to the results of the other vernalization periods, but the branch lengths were only between 10 and 15 cm, whereas longer vernalization periods resulted in axillary branches with lengths between 20 to 35 cm. Thus, although the growth of axillary branches in early-flowering accession plants appeared insensitive to vernalization period, a longer vernalization period promoted the growth of axillary branches in plants from the late-flowering accession.

3.2.5 No vegetative zone (V3) was observed in either accession after vernalization

In addition to the outgrowth mentioned above in section 3.2.4, a major feature of axillary branches is the developmental fate of the axillary meristems, which serves as a reservoir for continuous flowering during the perennial life-history. To

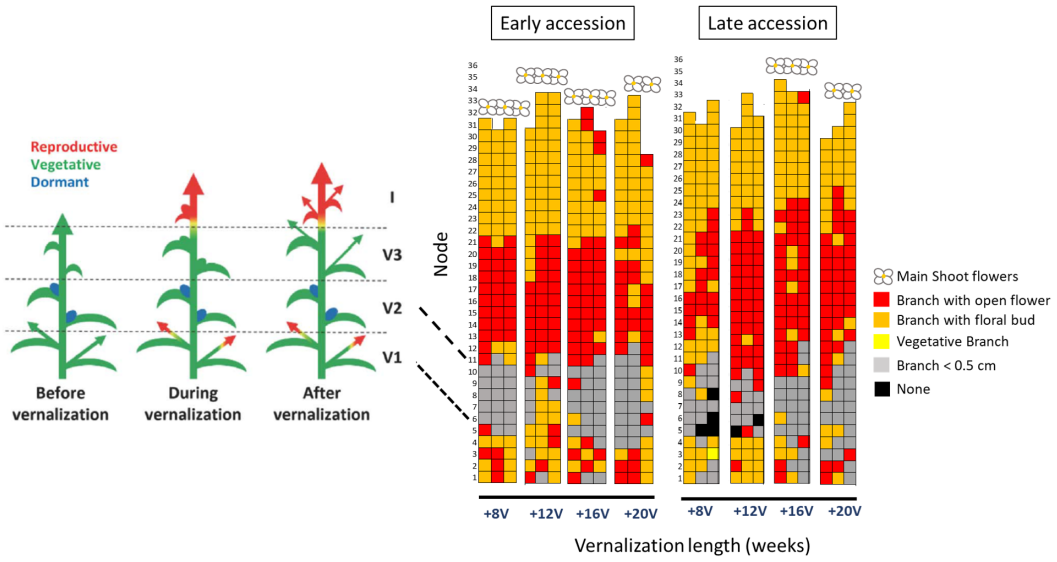

Figure 3.6. The developmental zonations of early- and late-flowering accessions vernalized for 8, 12, 16 or 20 weeks under greenhouse conditions. (Left) The zonations in plants of the Spanish *A. alpina* accession by Lazaro et al. (2018) and Hyun et al. (2019), which consists of V1, V2, V3 and I zones. (Right) The developmental fate of axillary meristems at each node of plants from the early- and late-flowering accessions after four different vernalization periods (shown on the x-axis). Each column represents an individual plant and each pixel shows the developmental stage of the axillary meristem at that node. The axillary meristems were assigned to one of five developmental categories: branch with open flower (red), branch with visible floral bud (orange), branch with vegetative meristem (yellow), branch smaller than 0.5 cm (grey), no axillary meristem present (black). The flower symbols above the columns indicate the primary shoot has flowered.

understand the developmental fate of axillary meristems in response to various vernalization periods during the life-history of Norwegian *A. alpina* accessions, I dissected all axillary branches from three, 7-week-old individual plants of both accessions after exposure to 8, 12, 16 or 20 weeks of vernalization. These three individuals per treatment were the same plants analysed by the above subsections. The dissection was performed one month after vernalization treatment and five categories of developmental fate were assigned: inflorescence branch (branch with open flowers), reproductive branch (branch with visible floral buds), vegetative branch (no visible floral bud), dormant branch (branch smaller than 0.5 cm), and no detectable axillary meristem. The right panel of Figure 3.6 illustrates the results of this dissection. To visualise the developmental fate of the axillary meristems, the above-mentioned five categories were colour-coded. The axillary meristems of early and late-flowering accessions displayed similar developmental fates. Moving basipetally down the plant, several reproductive axillary meristems were observed immediately below the main primary shoot (above node 21), followed by a few inflorescence branches beneath these reproductive axillary meristems (between node 10 and 21), small branches (between node 5 and 10), and flowering branches (below node 5).

In contrast to the zonations proposed by Lazaro et al. (2018) and Hyun et al. (2019) for the Spanish *A. alpina* Pajares accession (left panel of Figure 3.6), no vegetative zone (V3) was observed in either the early- or late-flowering accession after vernalization. Despite the absence of the V3 zone, a zone with small branches (shown in grey) was observed, which may be analogous to the V2 zone in the Spanish accession. Moreover, several flowering branches were observed beneath this V2-analogous zone (below node 5), which could be compared to the V1 zone in the Spanish accession. Thus, the zonations of the Norwegian and Spanish accessions are related except for the absence of a vegetative V3 zone in the Norwegian accessions.

3.3 Phenotypic analysis under growth chamber conditions

3.3.1 Experimental Design

The phenotypic analysis of plants grown under greenhouse conditions, discussed in the above section, identified several important differences between the early- and late-flowering accessions. To better characterise these differences, a detailed large-scale phenotypic analysis was performed under growth chamber conditions. This experiment allowed a large sample of plants to be grown in several different combinations of conditions by varying their age when exposed to vernalization and the duration of vernalization. The purpose of using different combinations of age and vernalization lengths was to simulate how a short growing season might affect the link between vernalization and flowering behaviour in the early- and late-flowering accessions, and also to investigate how both accessions respond to different times of snowmelt as a proxy for the vernalization periods under growth chamber conditions. The experimental set up is shown in Figure 3.7 and the 12 different treatments are summarised in Table 3.3. The growth conditions in the Bronson growth chamber were programmed to simulate greenhouse LD condition of 16h light (21°C) and 8h dark (18°C) with 65% humidity, and treatments of 12, 16 or 20 weeks of vernalization were performed in the vernalization chamber (4°C in 8h light:16h dark) in greenhouse 1. Additional far-red light (100% in the setting) was added to the spectrum in the growth chamber. 12 early- and 12 late-flowering accessions plants were exposed to each of the treatments so that a total of 288 plants were scored. Light microscopic images of the SAM were taken before (Figure 3.8) and after vernalization (Figure B.2 and B.3) to access the developmental stage of the SAM. The scoring of different traits began immediately after vernalization and was performed three times per week for one month. These traits included the time to flowering of the primary shoot, the time to flowering of the axillary branches, plant height, the number of siliques on the primary shoot, and the length and developmental fate of the axillary branches. Ten out of the twelve plants in each treatment and for each accession were dissected node by node approximately 30 days after vernalization to examine the length and developmental fate of each axillary branch. The exception was treatment 1, where all twelve plants of both accessions were dissected. The control experiment, without vernalization

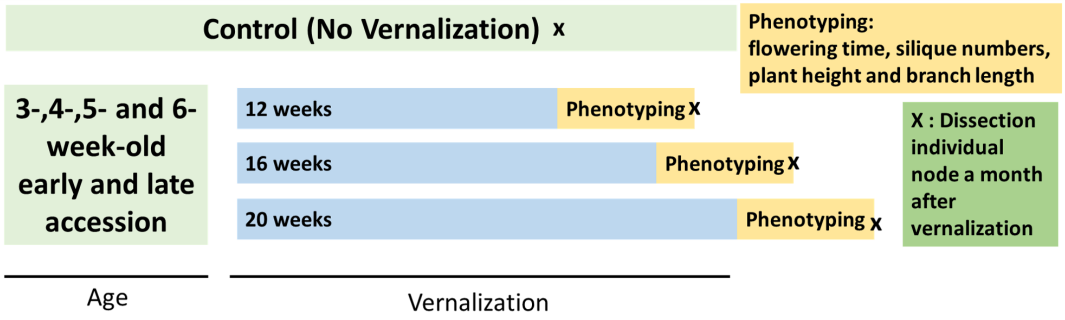

Figure 3.7. Set up of the growth chamber experiment. Plants of the early- and late-flowering accessions were grown for 3, 4, 5 and 6 weeks under LD conditions in the growth chamber, and were then vernalized for 12, 16 or 20 weeks in vernalization chamber. After vernalization, plants were transferred back to LD conditions in growth chamber, scored for flowering time, silique number, height and length of axillary branches for one month. About 30 days after vernalization, each plant was dissected node by node to examine the developmental fate of individual axillary branch around 30 days after vernalization. The controlled plants were dissected when all the early-flowering accession flowered without vernalization treatment. Light green denotes the pre-vernalization growth period, and light blue represents the vernalization period.

treatment, also included 12 plants per accession and all 24 plants were dissected when all the early-flowering accession flowered (Figure B.1).

	Vernalization period		
Pre-growth period	12 weeks	16 weeks	20 weeks
3 weeks	Treatment 1	Treatment 2	Treatment 3
4 weeks	Treatment 4	Treatment 5	Treatment 6
5 weeks	Treatment 7	Treatment 8	Treatment 9
6 weeks	Treatment 10	Treatment 11	Treatment 12

Table 3.3. The 12 different combinations of pre-growth and vernalization periods used for the growth chamber experiment. Each combination is assigned a treatment number which is used throughout section 3.3.

3.3.2 Differences in flowering time still occur if plants are exposed to vernalization after floral transition

The flowering time of the primary shoot of the plants exposed to all 12 treatments listed in Table 3.3 was scored, and the results were shown in Figure 3.8. The left panel of Figure 3.8 shows the weekly microscopic images of the SAM after 3, 4, 5 and 6 weeks of pre-growth. Three plants from each accession were dissected and imaged each week and one plant was selected as being representative for each accession. As demonstrated by the appearance of floral primordia, floral transition had occurred after 6 weeks of pre-growth in both accessions. To address the significance of the difference in flowering time between the early- and late-flowering accessions, I performed a non-parametric Mann–Whitney U test (MWU test) (Mann and Whitney, 1947). The underlying flowering-time distributions for each accession is presented as box plot in the right panel of Figure 3.8. The genotypes showed a p-value less than 0.05 in flowering time in seven of the treatments (1, 3, 5, 7, 10, 11 and 12). Considering a p-value threshold of 0.05 to indicate significance, I may reject the null hypothesis for these seven treatments. Amongst the treatments that showed a statistical difference in flowering time between the early- and late-flowering accessions, treatments 1, 3, 5 and 7 resulted in a median difference of less than 2 days, whereas treatments 10, 11 and 12 showed a median difference greater than 2.5 days. Therefore, because the largest difference in median value was for treatments 10, 11 and 12, long pre-growth periods appeared to accentuate the difference between early- and late-flowering accessions.

3.3.3 Flowering time does not always decrease as the vernalization period is extended

To determine how different treatments affect the flowering time of the primary shoot of the plants of the early- and late-flowering accessions, the data in Figure 3.8 may be visualised by grouping data for each accession (Figure 3.9) rather according to treatment. Different treatments within each accession were tested by Kruskal-Wallis test ($p < 0.05$) followed by the pairwise multiple comparisons post hoc Dunn test (Kruskal and Wallis, 1952; Dunn, 1961). Letters indicate groups with statistically significant differences. Plants of the early-flowering accession in

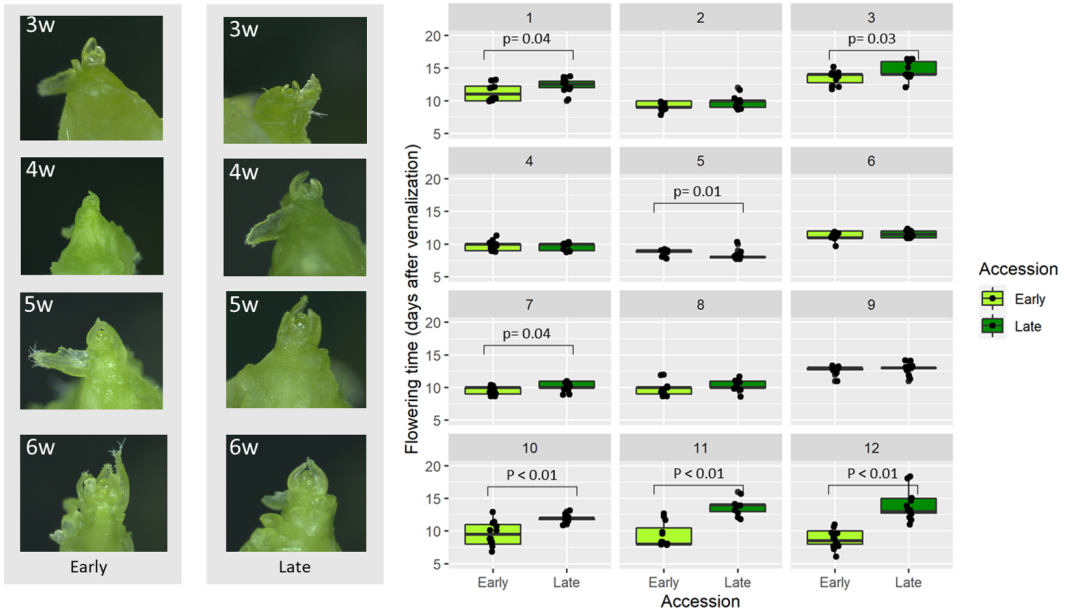

Figure 3.8. Microscopic examination of the apices of plants of the early- and late-flowering accessions before vernalization and the flowering time data distribution of the primary shoot from both accessions after vernalization under growth chamber conditions. (Left) The apices of early- and late-flowering plants grown for different lengths of time prior to vernalization. (Right) The flowering time of the primary shoot after vernalization for the early-flowering (light green) and late-flowering accessions (dark green) shown as individual box plot (n=12). 12 different treatments including 3, 4, 5 or 6 weeks of pre-growth periods followed by 12, 16 or 20 weeks of vernalization (Table 3.3) were indicated by the corresponding number of each sub-panel. The *p*-values of Mann–Whitney U test are indicated if the values are less than 0.05.

all treatments flowered within 15 days of the end of vernalization, whereas individuals of the late-flowering accession in three treatments (3, 11 and 12) flowered later than 15 days after the end of vernalization.

A similar trend was present in both accessions for treatments that included 3 and 4 weeks of pre-growth irrespective of vernalization period (i.e. treatments 1 to 6); the median flowering time decreased as the duration of vernalization extended from 12 to 16 weeks, whereas 20 weeks of vernalization treatments delayed flowering time compared with that after a shorter vernalization period. Intriguingly, a different effect was seen when the duration of vernalization was increased following a pre-growth period of 6 weeks (i.e. treatments 10 to 12). Here, extending the vernalization period resulted in no statistically significant difference in flowering time for the early-flowering accession but a monotonic increase for the late-flowering accession. Thus, although extending vernalization from 12 to 16 weeks reduces the flowering time in most treatments for both accessions, the final results are complicated by the results from treatments with 20 weeks of vernalization.

3.3.4 A vernalization period of 20 weeks does not cause differences in primary shoot elongation between the accessions

As described in section 3.2.2, a 16-week vernalization period rescued the bud-arrest phenotype of the late-flowering accession under greenhouse conditions, indicating a connection between vernalization and primary shoot elongation. To investigate how primary shoot elongation responds to different periods of vernalization and pre-growth, I measured the height of the primary shoot from the base to the apex of each plant once per week after vernalization. Figure 3.10 shows the distribution of heights measured one month after vernalization, when the primary shoot had finished flowering and all flowers had formed siliques.

To determine whether the final height differs between the early- and late-flowering accessions in each treatment, I performed a MWU test to analyse whether the underlying distributions of plant height from one accession stochastically dominates over the other. The results shown in Figure 3.10 demonstrate that the longest vernalization period of 20 weeks, which mostly resembles the natural condition, did not lead to a difference in height except when 4-week old plants were vernalized (treatment 6), which caused the late-flowering accession to produce a

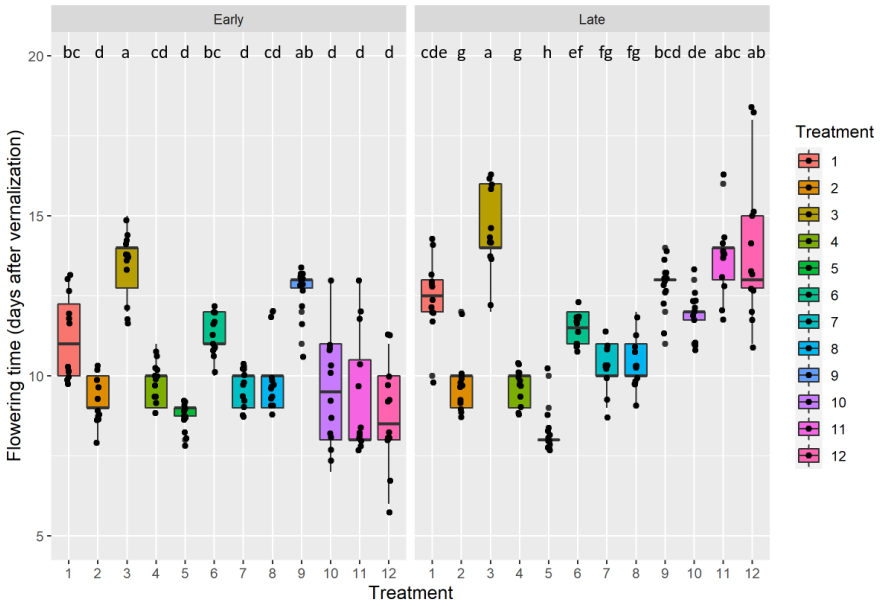

Figure 3.9. Flowering time of the primary shoot under growth chamber conditions according to accession and treatment. The box plots show the distribution of the time after vernalization at which the first flower opened on the primary shoot for plants of the early-flowering (left) and the late-flowering (right) accession. Each box plot represents 12 plants. 12 different treatments including 3, 4, 5 or 6 weeks of pre-growth periods followed by 12, 16 or 20 weeks of vernalization (Table 3.3) are labelled on the x-axis. Letters indicate groups with statistically significant differences ($p < 0.05$, Kruskal-Wallis test, post hoc Dunn test).

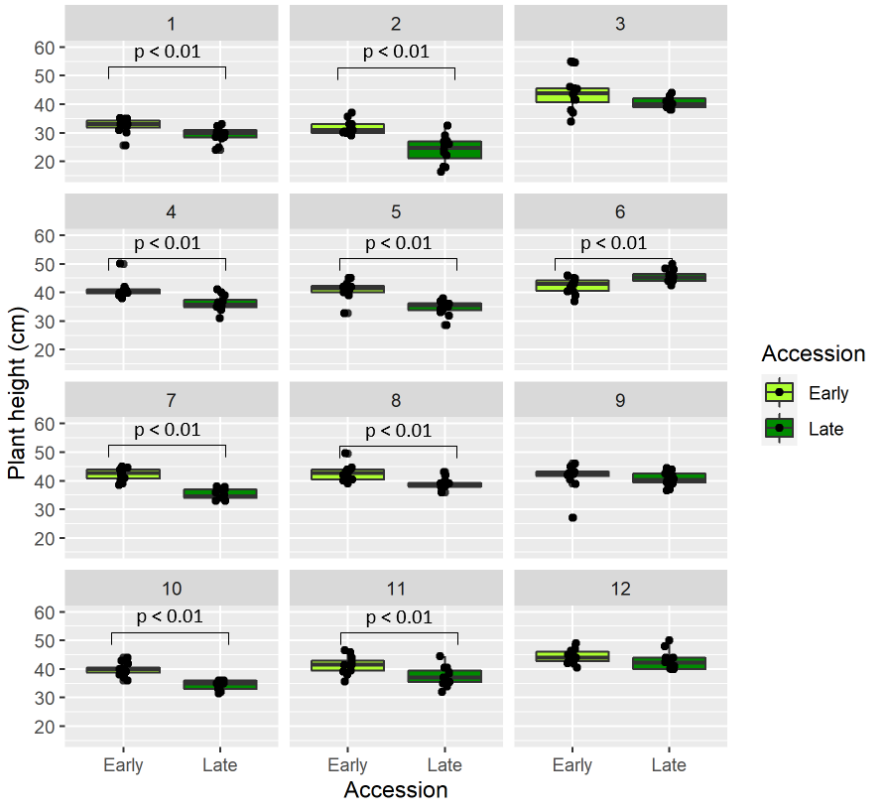

Figure 3.10. Distributions of the height of the primary shoot of each accession measured one month after vernalization under growth chamber conditions for each treatment. The distributions from the early-flowering (light green) and late-flowering accessions (dark green) are shown as individual box plots (n=12) and separated into 12 panels for each treatment. 12 different treatments including 3, 4, 5 or 6 weeks of pre-growth periods followed by 12, 16 or 20 weeks of vernalization (Table 3.3) are indicated by the corresponding number of each sub-panel. The p-values of Mann–Whitney U test less than 0.05 are indicated in each sub-panel.

longer primary shoot that the early-flowering accession. In response to vernalization periods of 12 or 16 weeks (treatments 1, 2, 4, 5, 7, 8, 10 and 11), the early-flowering accession produced a longer primary shoot than the late-flowering accession ($p<0.05$).

3.3.5 PRIMARY SHOOT ELONGATION OF THE LATE-FLOWERING ACCESSION IS MORE SENSITIVE TO VERNALIZATION AND PRE-GROWTH PERIODS THAN THAT OF THE EARLY-FLOWERING ACCESSION

To determine how different treatments affected the elongation of the primary shoot in the early- and late-flowering accessions, the data in Figure 3.10 were visualised according to accession (Figure 3.11). Different treatments within each accession were tested by Kruskal-Wallis test ($p < 0.05$) followed by post hoc Dunn test (Kruskal and Wallis, 1952; Dunn, 1961). Letters indicate groups with statistically significant differences. One notable feature is that the final height of the primary shoot for the early-flowering accession was mostly insensitive to different pre-growth and vernalization periods except when young plants were vernalized for short periods (treatments 1, 2 and 3); however, the late-flowering accession showed a considerably more variable response.

When the late-flowering accession experienced a pre-growth period of 3 weeks, the median final plant height decreases as vernalization increased from 12 to 16 weeks (treatments 1 and 2), but subsequently increased when plants were exposed to 20 weeks of vernalization (treatments 3). This is similar to the trend observed in the flowering time distributions in Figure 3.9, suggesting a possible link between these two traits. By contrast, when the late-flowering accession was vernalization after 5 or 6 weeks of pre-growth, a linear relationship existed between vernalization duration and final plant height, with a longer vernalization period resulting in taller plants. In general, it seemed that the late-flowering accession only reached full height after the longest vernalization irrespective of different pre-growth periods (treatments 3, 6, 9 and 12). For the early-flowering accession, only treatments 1 and 2 caused a significant difference in plant height distribution, which had a median of ~10 cm below that of the other treatments.

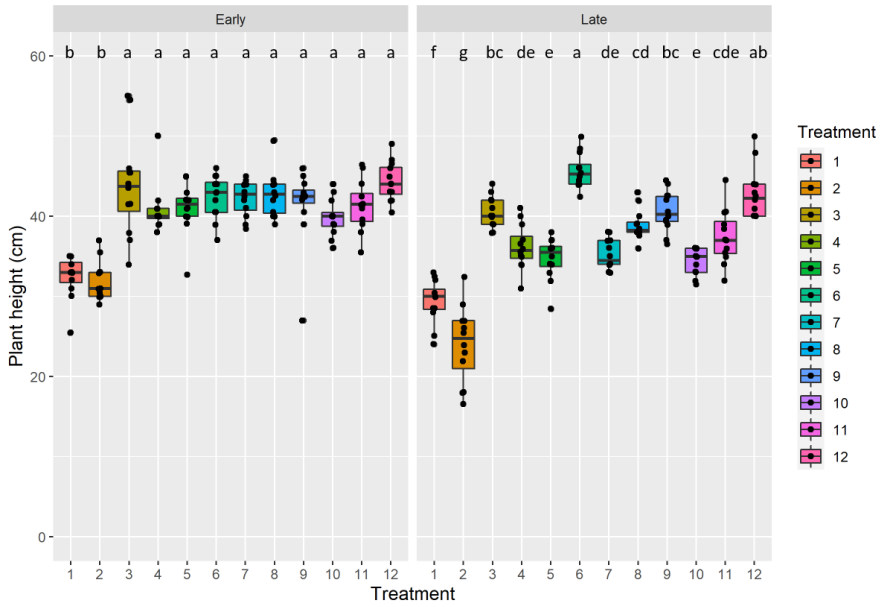

Figure 3.11. Height of the primary shoot measured one month after vernalization under growth chamber conditions according to accession and treatment. The box plots show the final height distributions for plants of early-flowering (left) and late-flowering (right) accessions. Each box plot represents 12 plants. 12 different treatments including 3, 4, 5 or 6 weeks of pre-growth periods followed by 12, 16 or 20 weeks of vernalization (Table 3.3) are labelled on the x-axis. Letters indicate groups with statistically significant differences ($p < 0.05$, Kruskal-Wallis test, post hoc Dunn test).

Collectively, the variation of the median plant height from different treatments in the late-flowering accession is more visible than the early-flowering accession. It appeared that the key factor of regulating plant height in the late-flowering accession is the vernalization length and the pre-growth period has little effect.

3.3.6 The longest vernalization period generates differences in silique production on the primary shoot

To assess whether the early- and late-flowering accessions have different reproductive capacities, I measured the total number of siliques on the primary shoot after all 12 treatments one month after vernalization. At this point all the flowers had fully developed into siliques. The results are shown in Figure 3.12 separated by individual treatments. I performed the MWU test on each pair of

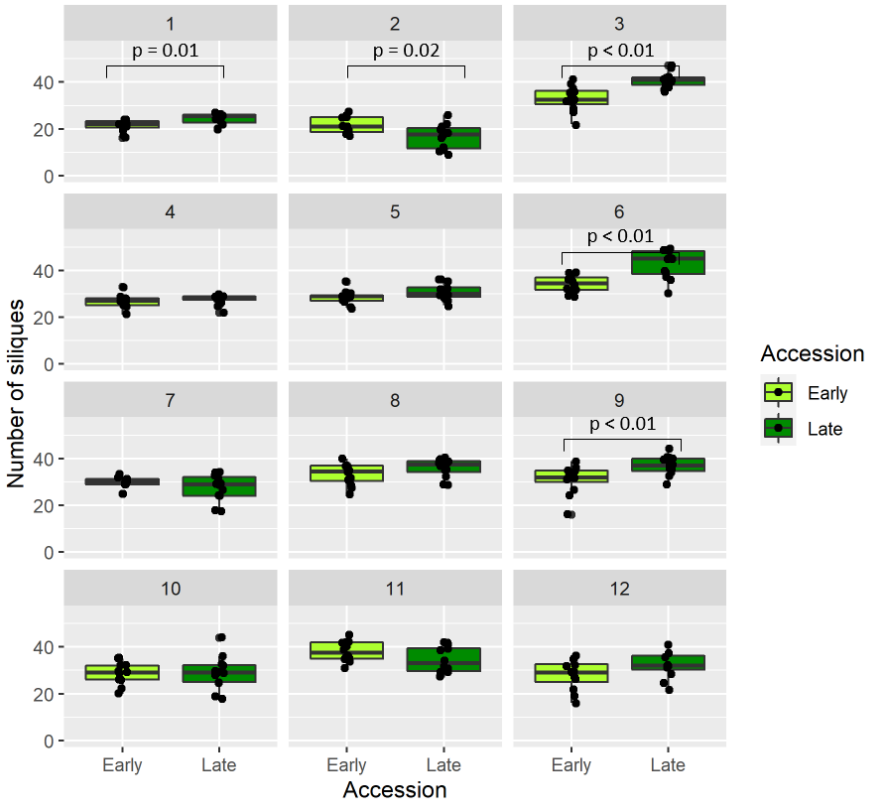

Figure 3.12. The number of siliques on the primary shoot of plants of each accession for each treatment, measured one month after vernalization under growth chamber conditions. The box plots (n=12) show the distributions of the early-flowering (light green) and late-flowering accessions (dark green). Each panel shows the 12 different treatments indicated by the number above the panel. 12 different treatments including 3, 4, 5 or 6 weeks of pre-growth periods followed by 12, 16 or 20 weeks of vernalization (Table 3.3) are indicated by the corresponding number of each sub-panel. The *p*-values of Mann–Whitney U test less than 0.05 are indicated in each sub-panel.

distributions and found that 5 out of the 12 treatments (1, 2, 3, 6 and 9) resulted in a p-value less than 0.05, confirming the statistically significant difference in silique production between both accessions under these environmental circumstances.

Intriguingly, three of these five treatments (3, 6 and 9) included a vernalization period of 20 weeks. These three treatments resulted in increased silique production in the late-flowering accession compared with that in the early-flowering accession and the median difference in silique number between accessions for these treatments was greater than 5. However, the treatment 12 did not lead to a significant difference between the accessions. Thus, although the longest vernalization period resulted in different reproductive output as indicated by the silique numbers between early- and late-flowering accessions, this difference was affected by pre-growth periods. In general, a statistical detectable difference in silique production between both accessions was only observed for 3, 4 or 5 weeks of pre-growth periods and was absent for all three treatments included 6 weeks of pre-growth. Also, it appeared fully vernalized late-flowering accession produced more siliques than the early-flowering accession except for the longest pre-growth period.

3.3.7 Extended vernalization increases silique production in both accessions

To further investigate how each treatment affected silique production of the primary shoot in each accession, the data from Figure 3.12 were plotted according to accession (Figure 3.13). Different treatments within each accession were tested by Kruskal-Wallis test ($p < 0.05$) followed by the pairwise multiple comparisons post hoc Dunn test (Kruskal and Wallis, 1952; Dunn, 1961). Letters indicate groups with statistically significant differences. For both accessions, the effect of different vernalization periods on the production of siliques was most visible when plants experienced 3 or 4 weeks of pre-growth, i.e. a pre-growth period in which plants did not fully undergo the floral transition. In response to these treatments (1 to 6), the early-flowering accession showed an increase of 8-12 siliques as the vernalization period extended from 12 to 20 weeks and the late-flowering accession showed an increase of 12-15 siliques. Among these six treatments, the longest vernalization period (20 weeks) resulted in the greatest silique production for both accessions, indicating that the plants optimised their reproductive production when the vernalization period was long and closest to that in their natural habitat.

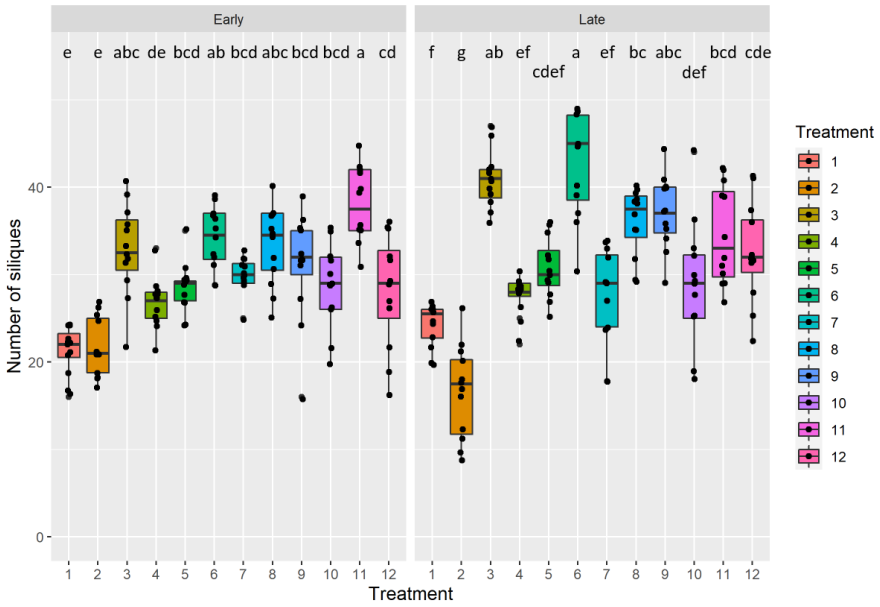

Figure 3.13. Number of siliques on the primary shoot for each accession measured one month after vernalization under growth chamber conditions. The box plots show the distributions for early-flowering (left) and the late-flowering (right) accession plants. Each box plot represents 12 plants. 12 different treatments including 3, 4, 5 or 6 weeks of pre-growth periods followed by 12, 16 or 20 weeks of vernalization (Table 3.3) are labelled on the x-axis. Letters indicate groups with statistically significant differences ($p < 0.05$, Kruskal-Wallis test, post hoc Dunn test).

During and after the floral transition (5 and 6 weeks of pre-growth), the effect of different vernalization periods on the number of siliques was weak for both accessions. In response to these 6 treatments (7 to 12) the median number of siliques increased by only 5-7 between 12 and 20 weeks of vernalization. Moreover, the increase was mainly observed for 12- and 16-week vernalization treatments, whereas 20 weeks of vernalization led to the production of a comparable number of siliques to that for 16 weeks of vernalization. Collectively, regardless of pre-growth period, an increased silique production was observed for both accessions when the vernalization period was extended.

3.3.8 Correlation matrices reveal the dynamics among flowering time, plant height and flower production

To investigate whether the traits scored were correlated, I applied the Kendall rank correlation test to the traits for each accession separately. This is quantified as Kendall's τ, a robust measure correlation due to its non-parametric nature (Kendall, 1938). The analysis included all 12 treatments for each accession and three main features of the primary shoot mentioned from section 3.3.2 to section 3.3.7: days to the opening of the first flower (DTF), final plant height measured one month after vernalization (h) and the final number of siliques measured one month after vernalization (f). Moreover, to capture the temporal changes of height and number of flowers each week, I normalised the weekly rate of elongation as well as flower/silique production from the primary shoot. The normalised rate of elongation, dh, is given by:

$$dh_{\mathrm{n,n+1}} = \frac{h_{\mathrm{n+1}} - h_{\mathrm{n}}}{h}, \tag{3.1}$$

where h_{n} is the height on primary shoot measure on the n^{th} week after vernalization. Thus, h_0 is the height of the primary shoot measured on the day when the vernalization treatment ceased and h_4 is the final height of plant measured one month after vernalization. It is apparent from equation 3.1 that $dh_{\mathrm{n,n+1}}$ is the fractional change in height of the plant's primary shoot between week n and week $n+1$ (e.g. $dh_{1,2} = 0.5$ means that the plant reached 50% of its final height between weeks 1 and 2).

Similarly, the normalised rate of flower/silique production is given by:

$$df_{\mathrm{n,n+1}} = \frac{f_{\mathrm{n+1}} - f_{\mathrm{n}}}{f}, \tag{3.2}$$

where f_{n} is the number of flowers/siliques on the primary shoot measured on the n^{th} week after vernalization. In reality, f_0 does not exist because the plants do not posses flowers/siliques when they are removed from vernalization; and f_4 is the final number of flowers/siliques measured one month after vernalization. Thus, $df_{\mathrm{n,n+1}}$ is the fractional change in flower number on the plant's primary shoot between week n and week $n+1$.

In total, I considered ten features of the primary shoot for each accession by combining the three main traits (DTF, h and f), the weekly elongation rate ($dh_{0,1}$, $dh_{1,2}$, $dh_{2,3}$, $dh_{3,4}$) as well as flower/silique production ($df_{1,2}$, $df_{2,3}$, $df_{3,4}$). Figure 3.14 shows the two correlation matrices when all 12 treatments are combined into a single sample, considering the early- and late-flowering accession separately; significant τ values (i.e. where $p < 0.05$) are shown as black text and non-significant values (i.e. where $p > 0.05$) are shown as white text.

The correlation matrices showed 36 and 35 significant correlations, out of a possible total of 45, for the early- and late-flowering accessions, respectively (Figure 3.14). This high proportion of significant correlations between the features indicates that a high level of interdependence between them and that changes in one will result in changes in others. There were no significant sign-switches of τ (i.e. a correlation becoming an anti-correlation) between the two accessions, but the strength of the correlations between features was mildly stronger for the late-flowering accession.

The significant correlation coefficients in the top row of both matrices (i.e. correlations between DTF and other features) revealed a relationship between the flowering time and the rate of elongation of the primary shoot. This is exemplified best by the strong negative correlation between $dh_{0,1}$ and $dh_{1,2}$, and DTF for both accessions. This strong negative correlation shows that the greater the elongation of the primary shoot is within the first two weeks after vernalization, the plant flowers earlier. This emphasises that shoot elongation is an essential component during the whole flowering process after vernalization, consistent with findings of Kemi et al. (2019). Intriguingly, the only significant correlations between DTF and h and f were observed for the late-flowering accession. Thus, this strong coordination between flowering time, shoot elongation and flower/silique production is unique to the late accession, implying a different flowering mechanism for the primary shoot exist between the early- and late-flowering accessions after vernalization.

Similar to the primary shoot data, I constructed a correlation matrix using traits from the axillary branches. The data from the individual axillary branches were averaged by zone to produce a single data point for the V1, V2 and I zones described in section 3.2.5. Only non-dormant axillary branches were considered when averaging to avoid biases. Similar to the primary shoot data, the axillary branch data for all 12 treatments were combined to produce a single data set for each accession to increase the size of the data sample. Three different traits of

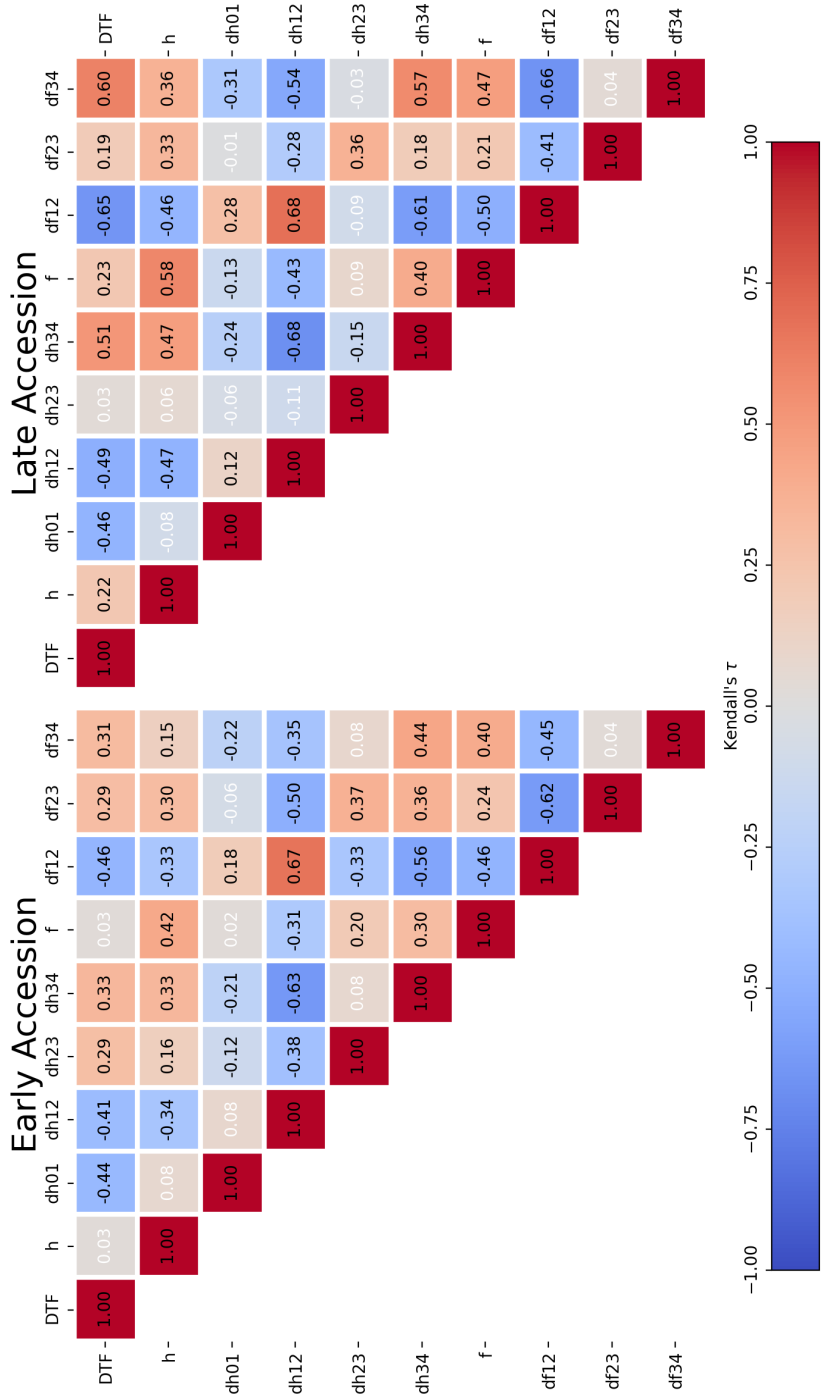

Figure 3.14. Correlation matrices for primary shoot traits for each accession. Positive correlations are in red, negative correlations are in blue. N/A indicates no correlation coefficient due to low numbers. Correlation coefficients of Kendall's τ are reported in each box where significant correlations (p-value < 0.05) are shown in black and non-significant correlations (p-value > 0.05) are shown in white. Ten features of the primary shoot are indicated: days to the opening of first flower (DTF), final plant height measured one month after vernalization (h), the final number of siliques measured one month after vernalization (f), the weekly elongation rate ($dh_{0,1}$, $dh_{1,2}$, $dh_{2,3}$, $dh_{3,4}$) as well as flower/silique production ($df_{1,2}$, $df_{2,3}$, $df_{3,4}$).

the axillary branch were considered: length (L), the number of siliques (S) and the flowering time (F). The length and number of siliques were measured one month after the vernalzation, whereas the flowering time of each axillary branch was measured as the time taken for the first flower to open following vernalization. The results are shown in Figure 3.15, with individual correlation matrix for each accession. Significant τ values are shown as black text and non-significant values as white text. When there were fewer than three plants that had non-zero values for both traits in a pair, the entry for that pair of traits in the correlation matrix was given as N/A to avoid spurious results.

The correlation matrices reported 16 and 7 significant correlation coefficients, out of a possible total of 36, for the early- and late-flowering accessions, respectively (Figure 3.15). Such a low proportion suggests that most traits are not correlated within or across zones, unlike the strong correlations observed among the primary shoot traits.

For the V1 zone in the early-flowering accession, all three traits were significantly correlated: the flowering time (V1F) was negatively correlated with both the length and the number of siliques, whereas the length was positively correlated with the number of siliques. Thus, axillary branches in the V1 zone of the early-flowering accession which flower earlier are longer and produce more siliques. A similar relationship was observed for axillary branches in the I zone, though no significant correlation was observed between flower time and branch length. Considering cross-zone correlations, one sees multiple relationships between the traits of I zone branches and V1/V2 branches in the early-flowering accession. For example, positive correlations existed between the flowering time of I zone branches and the flowering time of V1 and V2 branches. Thus, flowering time appears to be co-ordinated across zones and if flowering occurs late in one zone it also does so in the other zones. However, the correlation was significantly stronger between I and V2 branches than between I and V1 branches showing that this co-ordination is not present across the plant zones.

For the late-flowering accession, significant correlations were restricted to traits of the I zone branches. For the I zone, similar to the early-flowering accession, the flowering time was negatively correlated with both the length and the number of siliques, whereas the length was positively correlated with the number of siliques. However, unlike the early-flowering accession, this pattern was not observed in the V1 zone. The strength of the correlation between IL and IS

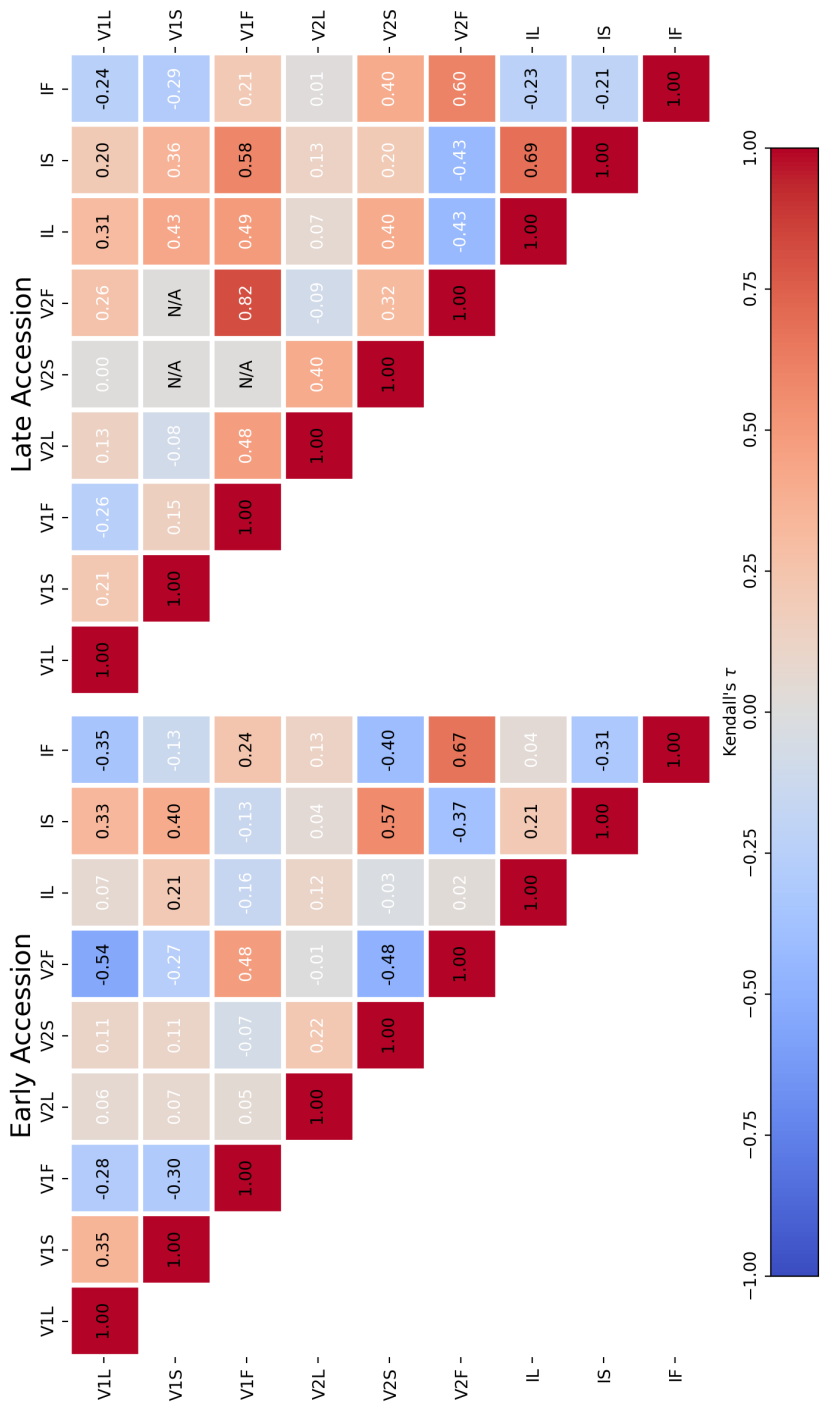

Figure 3.15. Correlation matrices for axillary branch traits for each accession. Positive correlations are in red, negative correlations are in blue. N/A indicates no correlation coefficient due to low numbers. Correlation coefficients of Kendall's τ are reported in each box where significant correlations (p-value < 0.05) are shown in black and non-significant correlations (p-value > 0.05) are shown in white. Three different traits of the axillary branch were considered in the V1, V2 and I zones: length (L), the number of siliques (S) and the flowering time (F).

was stronger in the late-flowering than the early-flowering accession (0.69 vs. 0.21). Similarly to the early-flowering accession, some significant correlations between the length of the V1 branches and the I zone traits were observed. However, a striking difference between the early- and the late-flowering accessions is the lack of significant correlations between the flowering times of different zones in the late-flowering accession; flowering of the axillary branches in the late-flowering accession was not co-ordinated across zones, as it was for the early-flowering accession. Another intriguing difference is that the flowering time of the V1 branches was positively correlated with the number of siliques in the I zone for the late-flowering accession. This positive relationship shows that although *early* flowering of the I zone branches was observed with a higher number of siliques in the I zone, *late* flowering of the V1 zone branches was related to a high number of siliques in the I zone. No other positive correlation between a zone's flowering time and length/silique number was observed in either matrix.

Considering the matrices collectively, it is clear the flowering mechanisms of the axillary branches differ between the early- and the late-flowering accessions.

3.3.9 THE RELATIONSHIPS BETWEEN PRIMARY SHOOT AND AXILLARY BRANCH TRAITS REVEAL UNIQUE FEATURES IN EACH ACCESSION

Due to the perennial growth habit of *A. alpina*, the underlying flowering mechanism that governs flowering of both the primary shoot and axillary branch is extremely important. The primary shoot only flowers in the first year in response to the first winter cold and then undergoes senescence, meaning that flowering in subsequent years occurs solely from the axillary branches. To understand this mechanism in more details, I constructed correlation matrices using all 19 traits from both the primary shoot and the axillary branches for each accession (Figures 3.16 and 3.17). Furthermore, to examine the interaction between the primary shoot and the axillary branch traits more closely, the 90 different correlations indicated by black squares in Figures 3.16 and 3.17 are shown in Figure 3.18.

Figure 3.18 shows several significant correlations for each accession, revealing interactions between primary shoot and axillary branch traits. Out of the 90 total correlations, 27 and 31 significant correlations were observed for the early-

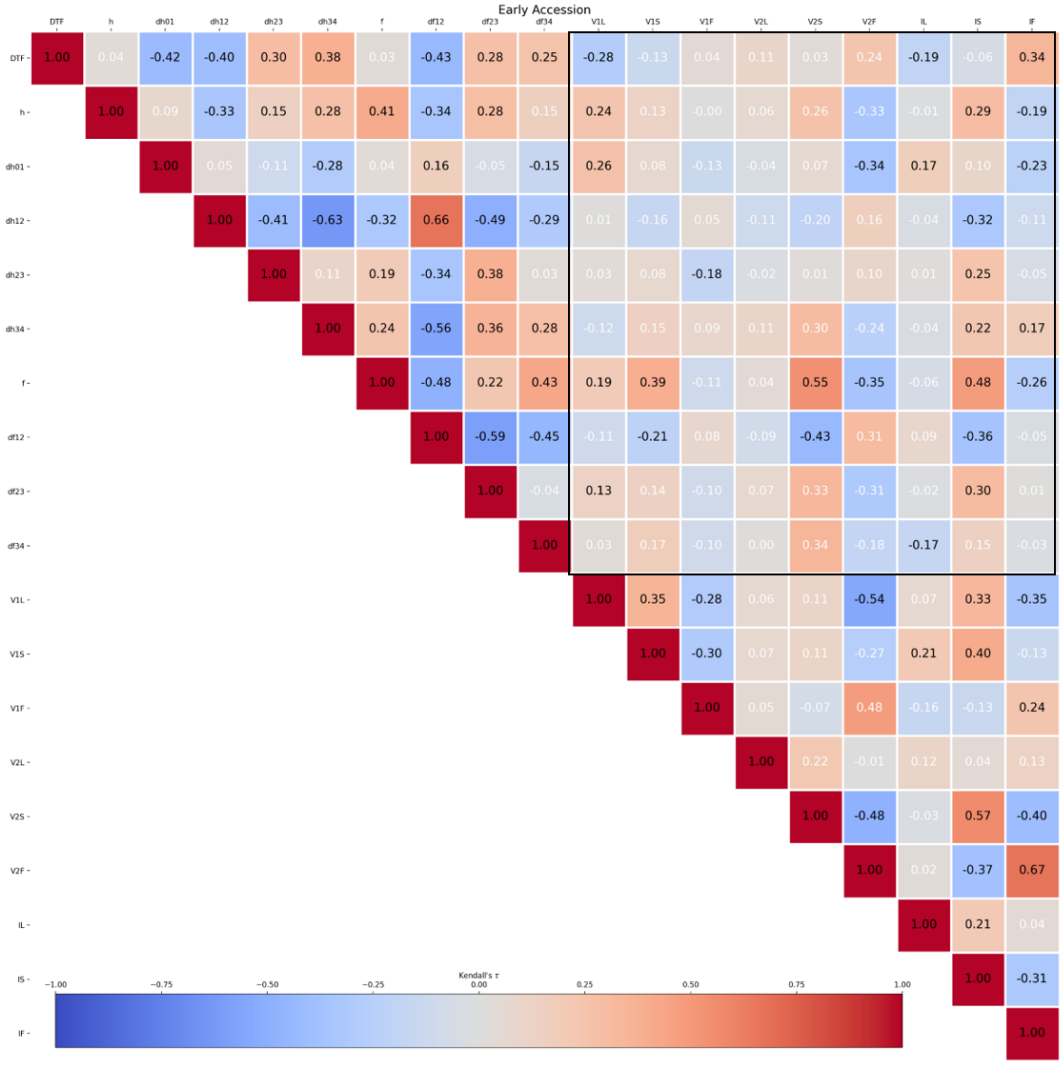

Figure 3.16. Correlation matrix for primary shoot and axillary branches traits for the early-flowering accession. Positive correlations are shown in red, negative correlations are shown in blue. N/A indicates no correlation coefficient due to low numbers. Correlation coefficients of Kendall's τ are shown in each box where significant correlations (p-value < 0.05) are shown in black and non-significant correlations (p-value > 0.05) are shown in white. The black box highlights the area of the matrix containing the correlations between the primary shoot traits and the axillary branch traits.

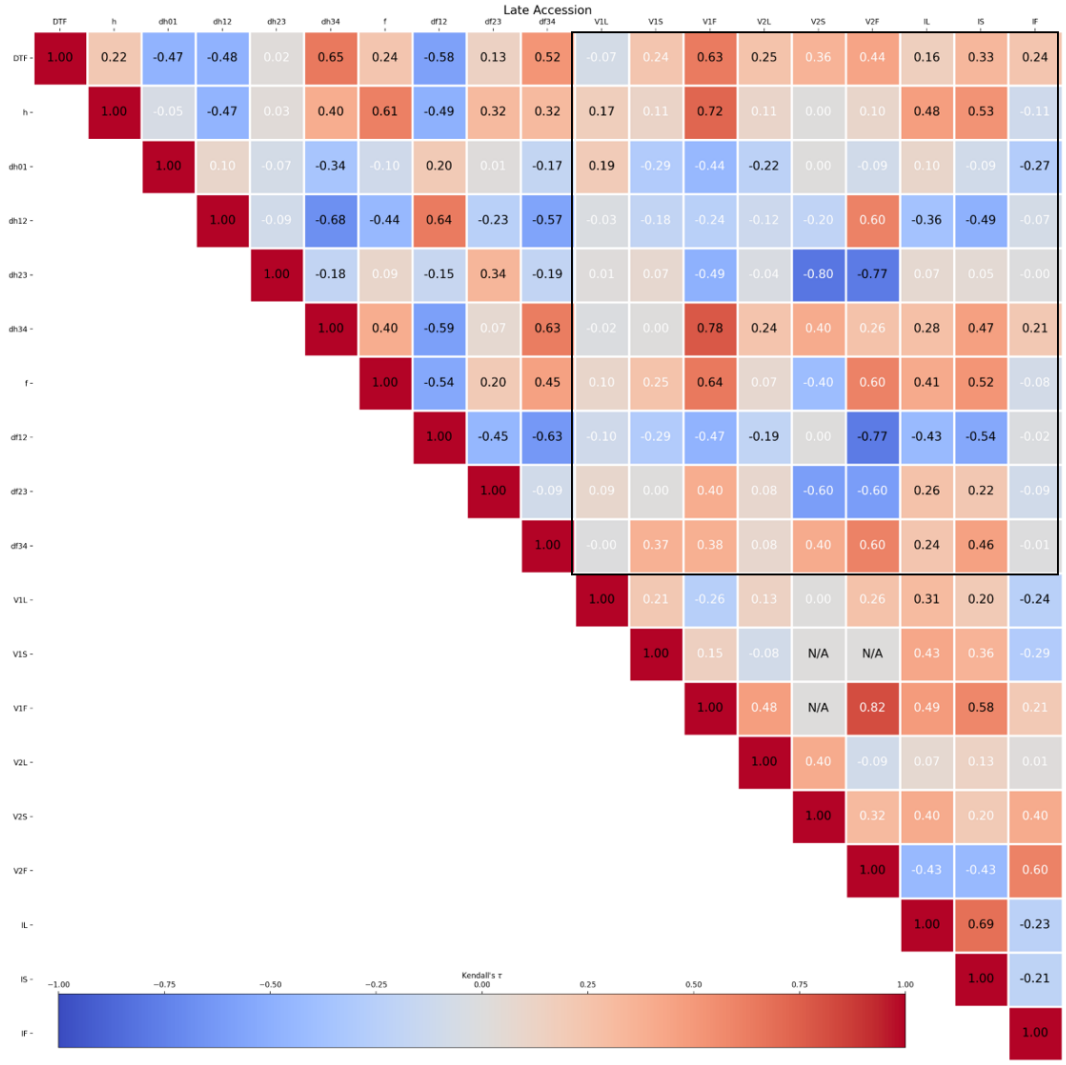

Figure 3.17. Correlation matrix for the primary shoot and axillary branches traits for the late-flowering accession. Positive correlations are shown in red, negative correlations are shown in blue. N/A indicates no correlation coefficient due to low numbers. Correlation coefficients of Kendall's τ are shown in each box where significant correlations (p-value < 0.05) are shown in black and non-significant correlations (p-value > 0.05) are shown in white. The black box highlights the area of the matrix containing the correlations between the primary shoot traits and the axillary branch traits.

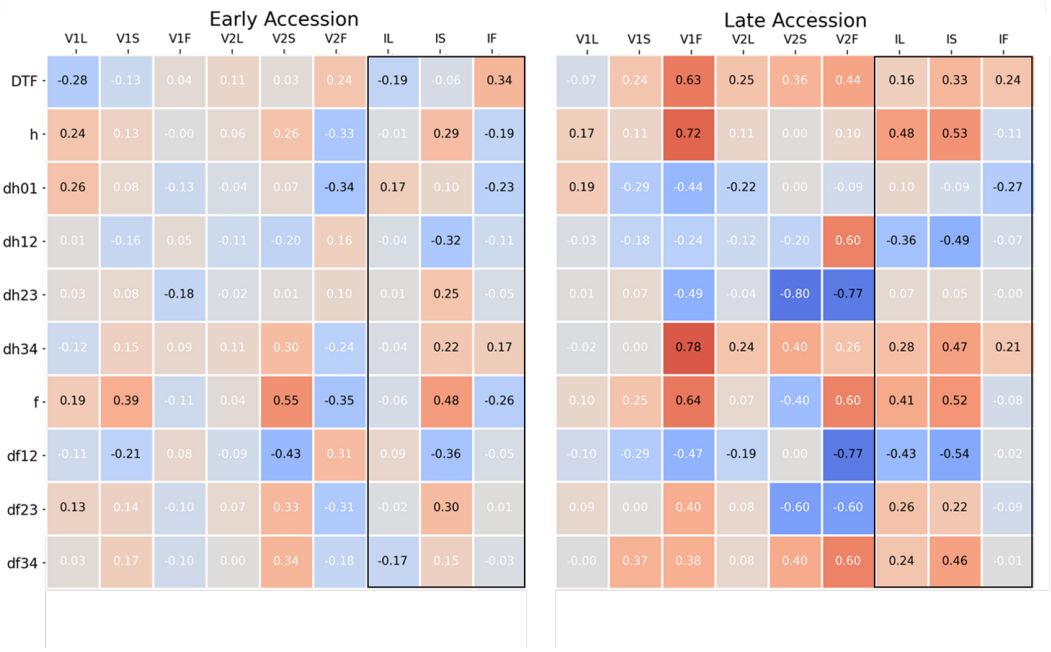

Figure 3.18. Correlation matrices showing the interaction between the primary shoot and the axillary branches traits for the early- and late-flowering accessions. Positive correlations are shown in red, negative correlations are shown in blue. N/A indicates no correlation coefficient due to low numbers. Correlation coefficients of Kendall's τ are shown in each box where significant correlations (p-value < 0.05) are shown in black and non-significant correlations (p-value > 0.05) are shown in white. The black boxes highlight the area of the matrices that contain the correlations between the primary shoot traits and the I zone axillary branch traits.

and late-flowering accessions, respectively. More than half of these significant correlations represented interactions between the primary shoot and I zone axillary branch traits in both the early-flowering (55%) and late-flowering accession (61%).

One example of these relationships was the positive correlation for both accessions between the flowering time of the primary shoot and that of the I zone branches, demonstrating that the flowering of both is coordinated. Intriguingly, the relationship between the flowering time of the primary shoot and the length of I zone branches differed between the accessions; the early-flowering accession showed a negative correlation and the late-flowering accession showed a positive correlation. Another example of this sign-switching is the relationship between $df_{3,4}$ and the length of the I zone branches; in the early-flowering accession it was negative and in the late-flowering accession it was positive. Thus, when the two above-mentioned sign-switching features were combined, early-flowering accession plants which can flower earlier and produce many flowers on their primary shoot earlier after vernalization will have longer I zone branches. By contrast, late-flowering accession plants which can flower earlier but only produce more of their primary shoot flowers later after vernalization will have longer I zone branches.

When the interactions between the primary shoot traits and other zones were considered, a negative correlation was observed between the flowering time of the primary shoot and the length of the V1 zone branches, but only for the early-flowering accession. Thus, if an early-flowering accession plant first flowered earlier, the growth of the axillary branches in the V1 zone was greater.

Collectively, several interactions were observed between the traits of the primary shoot and the axillary branches, mainly the I zone branches. However, some differences in these interactions were observed between the early- and the late-flowering accessions, mostly relating to axillary branch length.

3.3.10 The early-flowering accession continues to flower from the V1 zone after the completion of flowering from the I zone

As described in section 3.3.9, a negative correlation between the flowering time of the primary shoot and the length of the axillary branches in the V1 zone is unique to the early-flowering accession. To further determine how axillary branches grow in the V1 zone, and whether this growth is affected by different vernalization

periods, I examined the length, flowering time and developmental fate of each axillary branch for each accession about 30 days after vernalization, as described in section 3.3.1.

The results of treatments that included 5 weeks of pre-growth (treatments 7, 8, 9) are shown in Figure 3.19, where rows A and B show the data from the early- and late-flowering accession, respectively. The right-most panel of Figure 3.19 shows the developmental fate of each axillary meristem, organised by node for individual plants and treatments. It shows that the zonations under growth chamber conditions are consistent with the results for plants grown under greenhouse conditions described in section 3.2.5, reconfirming that nodes 1 to 5 corresponds to the V1 zone, nodes 6 to 10 correspond to the V2 zone and the nodes above 11 correspond to the I zone.

The left-most panels in Figure 3.19 illustrate the individual box plots for the length of each axillary branch ordered by node; the background is colour-coded based on the location of the V1 (green), V2 (blue) and I (red) zones. Notably, a sharp contrast was observed between the V1 zones of the two accessions; more active axillary meristems resulted in shoot elongation in the early-flowering accession, independently of the duration of vernalization, whereas little shoot elongation was observed for the late-flowering accession, i.e. in the presence of inactive axillary meristems. Comparison of the V2 and I zones showed that the elongation of the axillary branches was similar for both accessions and all three vernalization periods.

The central panels of the two rows show the probability density functions of the flowering time for each zone, on the basis of the time until the first open flower on each individual axillary branch. The flowering-time distributions for the zones in the early-flowering accession were complex and were often bi-modal. However, a clear pattern emerged, in which flowering first occurred in the majority of the I zone axillary branches and then in the V1 branches. Some overlap and concurrent flowering existed, specifically in treatment 7. In treatments 8 and 9, the bi-modality of flowering time was strengthened as the duration of vernalization increased, and during the first flowering episode, the majority of I branches flowered concurrently with a few V1 branches, whereas the inverse scenario occurred in the second flowering episode. For the late-flowering accession, only the I zone branches flowered; however, evidence for slight bi-modality in the distributions still existed, regardless of the duration of vernalization.

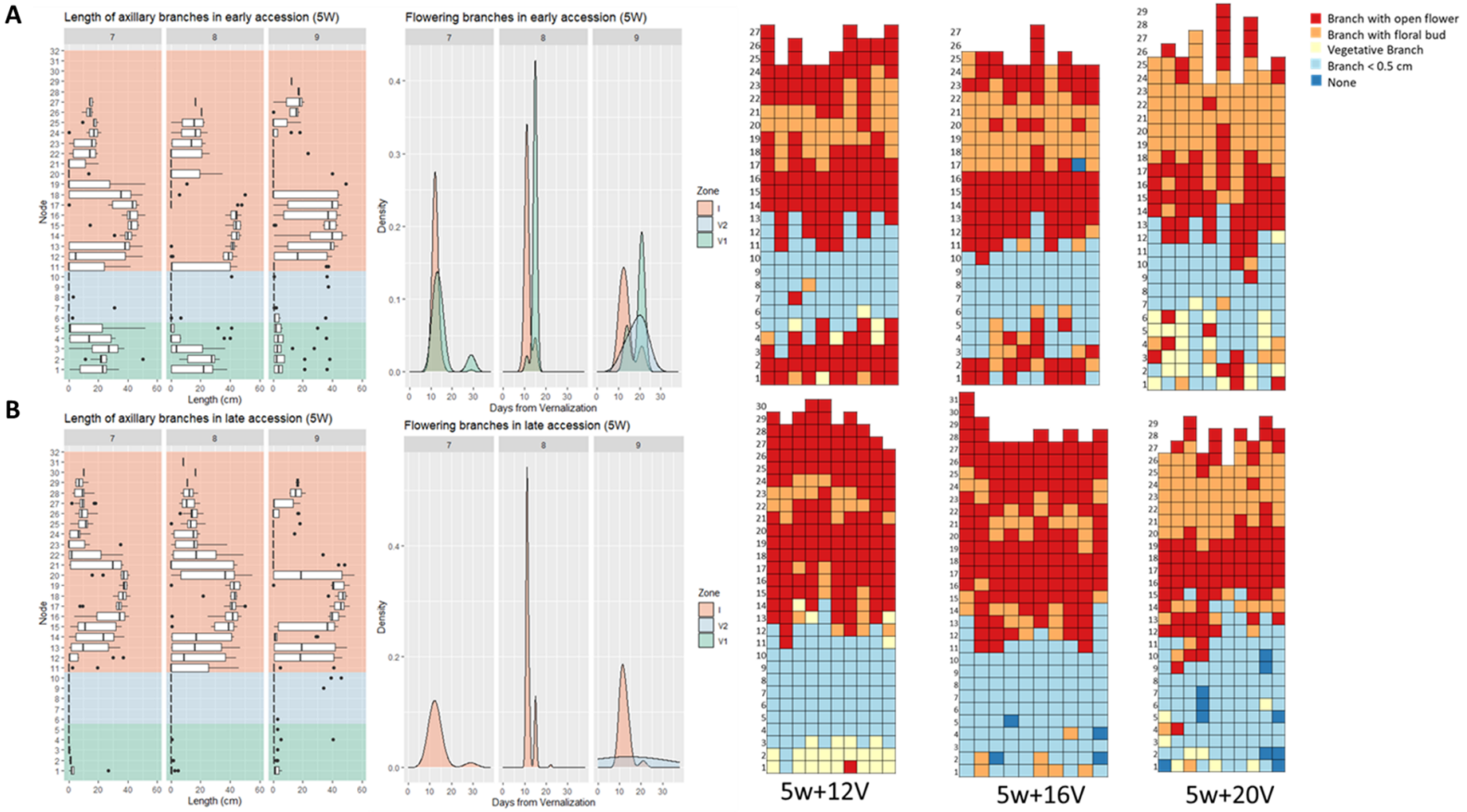

Figure 3.19. Analysis of the length, flowering time and developmental fate on each axillary branches of the early- and late-flowering accessions, for treatments including 5 weeks of pre-growth and 12, 16 and 20 weeks of vernalization. The top (A) and bottom (B) rows show the data for the early- and late-flowering accessions respectively. The left-most panels show the box plots of the length of each axillary branches ordered by node and the background colour indicates the V1 (green), V2 (blue) and I (red) zones. The middle panels show the probability density function of the flowering time of the axillary branches separated by zone, using the same colour scheme. The right-most panels show the developmental fate of each axillary branch, organised by node and separated by plant and treatment. The five categories are: branches with open flower (red), branches with visible floral bud (orange), branches with vegetative meristem (yellow), branches smaller than 0.5 cm (light blue), and no axillary meristem present (dark blue). The left-most and middle panels use data from all 12 plants per treatment and the right-most panel uses 10 out of the 12 plant per treatment. The titles of each sub-panel indicate the corresponding treatment listed in Table 3.3

These results describe a scenario in which flowering of the early-flowering accession is not synchronous across the plant, whereas flowering of the late-flowering accession begins throughout the plant simultaneously. Flowering in plants of the early-flowering accession began at the primary shoot and the I zone axillary branches (compare Figures 3.8 and 3.19) and occurred from branches of the V1 zone 5-10 days later. Flowering in plants of the late-flowering accession plants occurred more or less simultaneously at the primary shoot and the I zone branches. Collectively, the data show that although the location of the V1, V2 and I zones are identical in both accessions, the active axillary meristems within the V1 zone in the early-flowering accession results in a staggered flowering behaviour across the plant, and this represents a difference in the underlying flowering mechanisms of the early- and late-flowering accessions.

3.3.11 Shoot elongation in the V1 zone of the early-flowering accession is controlled by the pre-growth period

As described in section 3.3.10, the elongation of the V1 zone axillary branches is unique to the early-flowering accession. Moreover, flowering of these branches occurs after that of the I zone and the primary shoot for the early-flowering accession but not the late-flowering accession. Figure 3.20 presents the same data as Figure 3.19 except for the 3 week pre-growth treatments (1, 2, 3). Notably, elongation of the V1 zone axillary branches was absent for both the early- and late-flowering accessions, unlike the results of the 5 week pre-growth treatments. A minority of the V1 zone axillary branches of the early-flowering accession still elongated and subsequently produced flowers, but only for treatment 1. Thus, the length of the pre-growth period has a strong regulatory effect on the elongation of the V1 zone branches, but only for the early-flowering accession.

Despite the similarity of plant architecture between both accessions after a 3 week pre-growth period, a few plants of the late-flowering accession did not produce any flowering axillary branched one month after vernalization, independent of the vernalization period (plants without a red node in the right-most panels of Figure 3.20). However, these plants still produced flowers on the primary shoot (Figure 3.8). Furthermore, a decreasing number of flowering axillary branches was observed in both accessions as the vernalization period was extended (quantified in Table

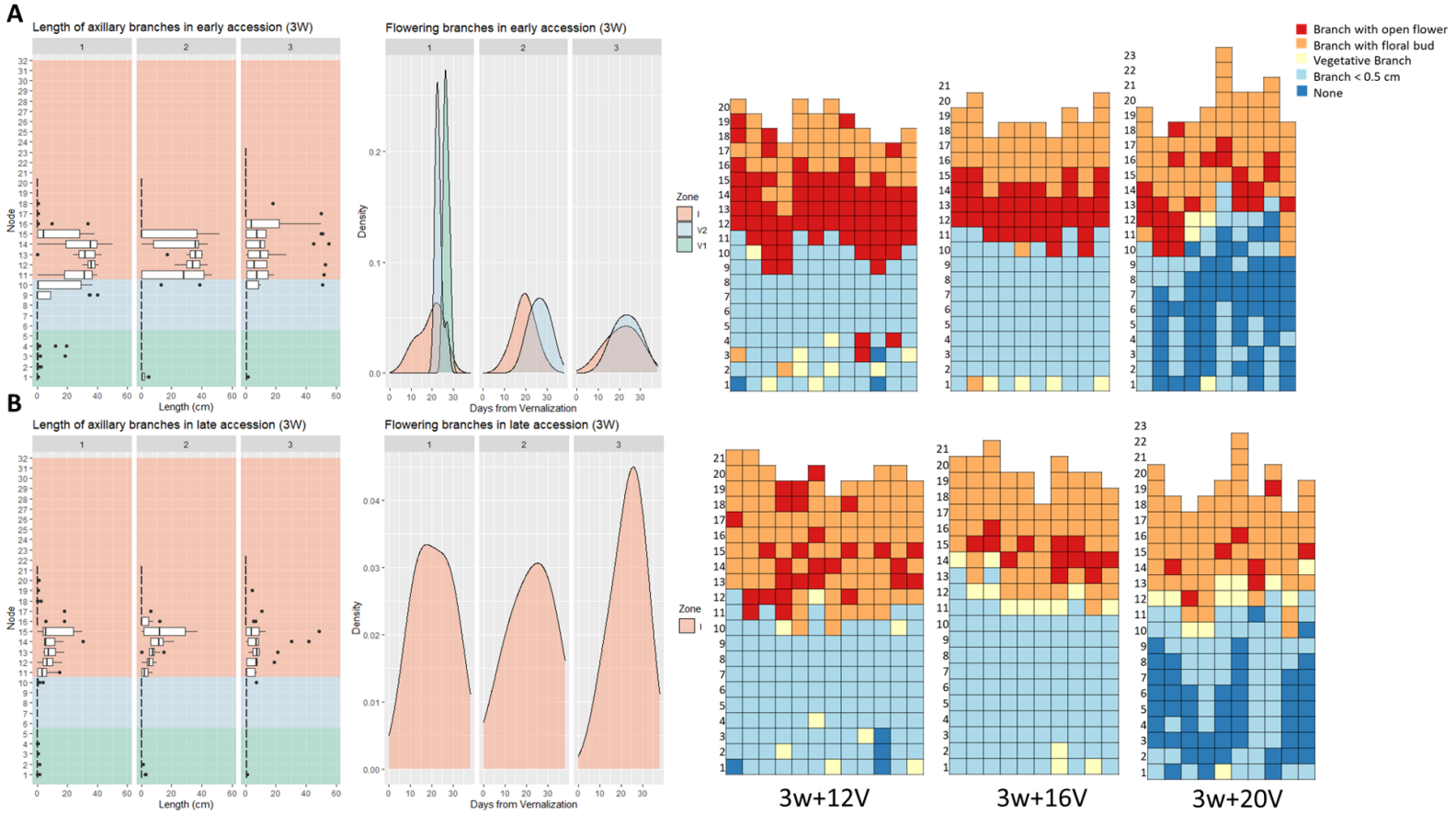

Figure 3.20. Analysis of the length, flowering time and developmental fate on each axillary branch of the early- and late-flowering accessions for treatments including 3 weeks of pre-growth and 12, 16 and 20 weeks of vernalization. The top (A) and bottom (B) rows show the data for the early and late accessions respectively. The left-most panels show the box plots of the length of each axillary branches ordered by node and the background colour indicates the V1 (green), V2 (blue) and I (red) zones. The middle panels show the probability density function of the flowering time of the axillary branches separated by zone, using the same colour scheme. The right-most panels show the developmental fate of each axillary branch, organised by node and separated by plant and treatment. The five categories are: branches with open flower (red), branches with visible floral bud (orange), branches with vegetative meristem (yellow), branches smaller than 0.5 cm (light blue), and no axillary meristem present (dark blue). The left-most and middle panels use data from all 12 plants per treatment and the right-most panel uses 10 out of the 12 plant per treatment. The titles of each sub-panel indicate the corresponding treatment listed in Table 3.3

3.4). Nonetheless, the mean number of flowering axillary branches was greater for the early-flowering accession than the late-flowering accession independent of the duration of vernalization. Thus, whereas the reduced pre-growth period has impeded the elongation and flowering of the V1 branches in the early-flowering accession, causing them to appear similar to the V1 branches in the late-flowering accession, it did not erase the differences between the I zone branches, as evidenced by the greater number of flowering I zone branches per plant of the early-flowering accession. It is clear that the two accessions demonstrate different underlying flowering mechanisms, even when the pre-growth period is short.

	Vernalization period		
	12 weeks	16 weeks	20 weeks
Early accession	5.6±1.4	3.9±0.31	2.4±1.3
Late accession	2.3±1.6	1.3±0.94	0.8±0.6

Table 3.4. The number of flowering axillary branches per plant for treatments 1, 2 and 3 for each accession (n=12). Values are means ± SD.

3.4 Discussion

The comprehensive phenotypic analysis under both greenhouse (section 3.2) and growth chamber (section 3.3) conditions demonstrate different flowering mechanisms between these two Norwegian *A. alpina* accessions after vernalization. Firstly, the largest difference in median flowering time were observed for the treatment in which the oldest plants (6 weeks old) of the early- and late-flowering accessions were vernalized. This implies that the stage at which these accessions were vernalized is extremely important, because the resulting differences in flowering time might strongly impact reproductive success in response to a short growing season in the natural environment. Secondly, I have conclusively shown that the longest vernalization period of 20 weeks, which is closest to natural environmental conditions, leads to similar plant height in both accessions, but to a greater production of flowers and siliques on the primary shoot only in the late-flowering accession. This suggests that only the longest vernalization can overcome the vernalization requirement in plants of the late-flowering accession. Furthermore, despite the number of reports that suggest that maximum plant height is positively

correlated with flowering time (Bolmgren and D. Cowan, 2008; Du and Qi, 2010; Vile et al., 2006), I found that two out of four treatments that included 20 weeks of vernalization did not lead to a statistically significant difference in plant height between the early- and late-flowering accessions, although these two treatments led to a difference in flowering time. An intriguing possibility is that the selection pressure promotes silique production on the primary shoot of the late-flowering accession, but does not affect the traits of the primary shoot of the early-flowering accession.

Unlike the studies of zonations in the Spanish Pajares accession (Lazaro et al., 2018; Hyun et al., 2019; Vayssières et al., 2020), I could not identify a V3 zone in the Norwegian accessions under either greenhouse or growth chamber conditions. This discrepancy between accessions may be due to the different selection pressures in different natural environments: for example, the major selection pressure for the Spanish accession is thought to be a long drought season, whereas for the Norwegian accession it is likely to be a short growing season (Wunder et al., in preparation). To analyse how different zones develop after experiencing a second winter, a long-term phenotypic experiment could be conducted to document the developmental fate of each axillary meristem under growth chamber conditions defined using two-year field-collected climate data.

Lastly, I provide the first direct evidence for a difference in the correlation between primary shoot and axillary branch traits in the early- and late-flowering accessions. A negative and significant correlation between the flowering time of the primary shoot and the elongation of axillary branches in the V1 zone is unique to the early-flowering accession. This result leads me to speculate that if this correlation is only present in the early-flowering accession, the increased activity of meristems within the V1 Zone might be a trait modified by selection pressure that accounts for a reduced silique production on the primary shoot, but additional flowering from the V1 zone, leading to a higher reproductive output. This supports the suggestion that axillary branches that initiate at a distance from the shoot apical meristem are crucial for the perennial lifestyle (Ponraj and Theres, 2020). Moreover, this active behaviour of axillary meristems, which leads to elongation within the V1 zone in the early-flowering accession is absent when plants are vernalized when they are young. The activity of the meristems within the V1 zone is probably regulated by age, to enable the plant to flower properly after vernalization.

To the best of my knowledge, this is the first report of the correlation between primary shoot and axillary branch traits in *A. alpina*. A correlation matrix approach has allowed the phenotypic differences between the early- and late-flowering accessions after vernalization to be identified, and provides a link between phenotype and adaptive significance. Furthermore, the simulation of different pre-growth and vernalization periods suggests that these two accessions respond to both the duration of the growth season and snow cover to optimise reproduction in their natural environment, which is further addressed in Chapter 6 via the *in situ* transplantation experiment. This detailed and systematic phenotypic analysis of primary shoot and axillary branch traits could potentially be applied to other plant species to uncover adaptive flowering mechanisms.

Chapter 4

Genome-wide transcriptome analysis of early- and late-flowering accessions

4.1 Experimental design under growth chamber conditions

To decipher the underlying molecular mechanisms that result in the phenotypic differences described in section 3.1, a 24-week time-course RNA-sequencing (RNA-seq) experiment was performed for the early- and late-flowering accessions under growth chamber conditions with or without vernalization. The LD and vernalization conditions were those used in treatment 8 in section 3.3.1, where a 5-week pre-growth was followed by a 16-week vernalization period. In total, 1728 plants (the full capacity of the Bronson growth chamber) were used in this experiment: 816 plants from the early-flowering accession; 912 plants from the late-flowering accession. Sixteen time points were analysed in the experiment: eight time points for each non-vernalization or vernalization time-series. The first and second harvesting began 4 and 5 weeks, repectively, after sowing but prior to vernalization for those plants that underwent vernalization. Half of the plants that were vernalized, were sampled six times during the vernalization treatment, at roughly 3-week intervals, and twice after vernalization. The non-vernalized plants were sampled seven times between 4 and 15 weeks after germination. The experimental design

and sampling times are summarised in Figure 4.1.

To ensure that sufficient RNA was available, 12 samples were pooled from different primary shoot apices to create a single biological replicate. At each time point, four biological replicates were collected per accession. The three biological replicates with the highest RNA concentrations out of the four were submitted for sequencing. For 2 time points (12w and 15w) within the non-vernalization time-series, biological replicates were only harvested from the late-flowering accession because the early-flowering accession had completed flowering by this time. The plants selected for sampling were labelled one day before sampling to ensure efficient harvesting. This resulted in the collection of the four biological replicates per accession (96 samples in total), which took less than 1.5 hours. To avoid diurnal effects on gene expression, samples were always harvested at Zeitgeber 5 (ZT5). Furthermore, all apices were harvested by me for consistency.

For plants that were not vernalized, the SAM and whole-plant architecture of three different individuals from each accession were examined and photographed at six time points (4w to 9w). One plant per accession was selected to represent that specific time point (depicted in the upper row of Figure 4.2). In addition to documenting the phenotype in parallel to sampling, 12 plants per accession were scored twice weekly to record the time until the emergence of a visible floral bud and until the opening of the first flower. The flowering-time distributions are presented as box plots in the upper right panel of Figure 4.2. Scoring continued until 120 days after sowing, because after this time the late-flowering accession plants were heavily infected with powdery mildew, which led to the early termination of the nonvernalization time-series. Therefore, harvesting scheduled for the 18^{th} and 22^{nd} weeks after sowing was not performed.

For the vernalization treatment, 5-week-old plants were veranalized for 16 weeks. Only two time points (9 and 12 weeks) were selected at which to microscopically examine the SAMs, due to limited space in the Bronson growth chamber. Images of these SAMs are shown in the lower two rows of Figure 4.2. Harvesting plants in the vernalization chamber was impractical; therefore, this was performed outside the chamber. To reduce temperature effects on the plants, they were harvested in batches of four to minimise the time for which they were exposed to warm temperatures. Furthermore, at the two time points of 5 and 10 days after vernalization, harvesting was performed before the first flower opened for either accession. To document the flowering behaviour after vernalization, the time until

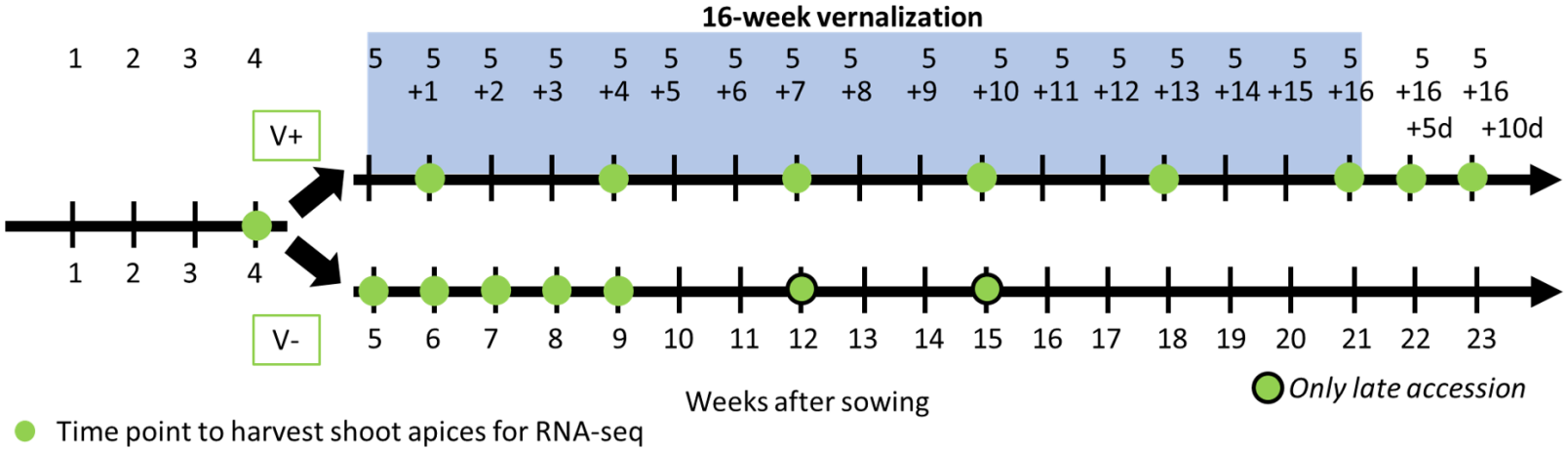

Figure 4.1. Experimental design for RNA-seq analysis of the Norwegian *A. alpina* early- and late-flowering accessions under growth chamber conditions. The 24-week timeline shows the detailed schedule for harvesting shoot apices of both accessions (green circle) or only of the late-flowering accession (green circle with black outline). Sampling was performed for non-vernalized plants and plants vernalized 16 weeks. The timeline represents the number of weeks after sowing. The growth conditions in the Bronson growth chamber were programmed to simulate greenhouse LD condition of 16h light (21°C) and 8h dark (18°C) with 65% humidity, and 16 weeks of vernalization were performed in the vernalization chamber (4°C in 8h light:16h dark) in greenhouse 1. Additional far-red light (100% in the setting) was added to the spectrum in the growth chamber.

the opening of the first flower was recorded for eight plants per accession and is presented as box plots in the lower right panel of Figure 4.2.

4.2 THE BUD-ARREST PHENOTYPE IS ALSO EXHIBITED UNDER GROWTH CHAMBER CONDITIONS

To confirm the phenotypic difference observed between the early- and late-flowering accessions without vernalization under growth chamber conditions, I examined the SAM microscopically at six selected time points to document the development of floral buds (upper panels of Figure 4.2). For the early-flowering accession, the SAM started to dome at 5 weeks, produced visible floral buds at 6 weeks, and subsequently formed the whole primary shoot inflorescence. This resulted in a flowering-time distribution with a median of 51 days after germination and an interquartile range (IQR) of 5 days (upper right panel of Figure 4.2). Only 11 plants flowered, because one plant died before flowering. Although the SAMs of plants of the late-flowering accession started to dome at 6 weeks and produced visible floral buds at 7 weeks , only 1 week later than the early-flowering accession, the floral buds developed relatively slowly after this, which was consistent with the bud-arrest phenotype documented by Dr. Ulla Kemi under greenhouse conditions. Notably, under growth chamber conditions without vernalization, a large variation in flowering time was observed for the late-flowering accession: six plants produced flowers from the primary shoot between 98 to 103 days after germination, but six plants did not flower before the end of experiment (upper right panel of Figure 4.2).

Floral buds on the primary shoot of both accessions continued to develop during the 16-week vernalization period. This growth was verified by the microscopic examination of the SAM performed 4 and 7 weeks after the start of vernalization (lower two panels of Figure 4.2) and led to no statistically significant difference in flowering time between both accessions (eight plants per accession tested by non-parametric MWU test, p-value = 0.11). Only eight plants were scored for flowering due to space limitations in the growth chamber. The flowering-time distribution of the early-flowering accession had a median of 10 days and an IQR of 0 days, showing a narrower range of flowering behaviour on the primary shoot compared with the early-flowering accession plants without vernalization (IQR = 5 days). This narrow flowering-time distribution was also observed for the late-flowering

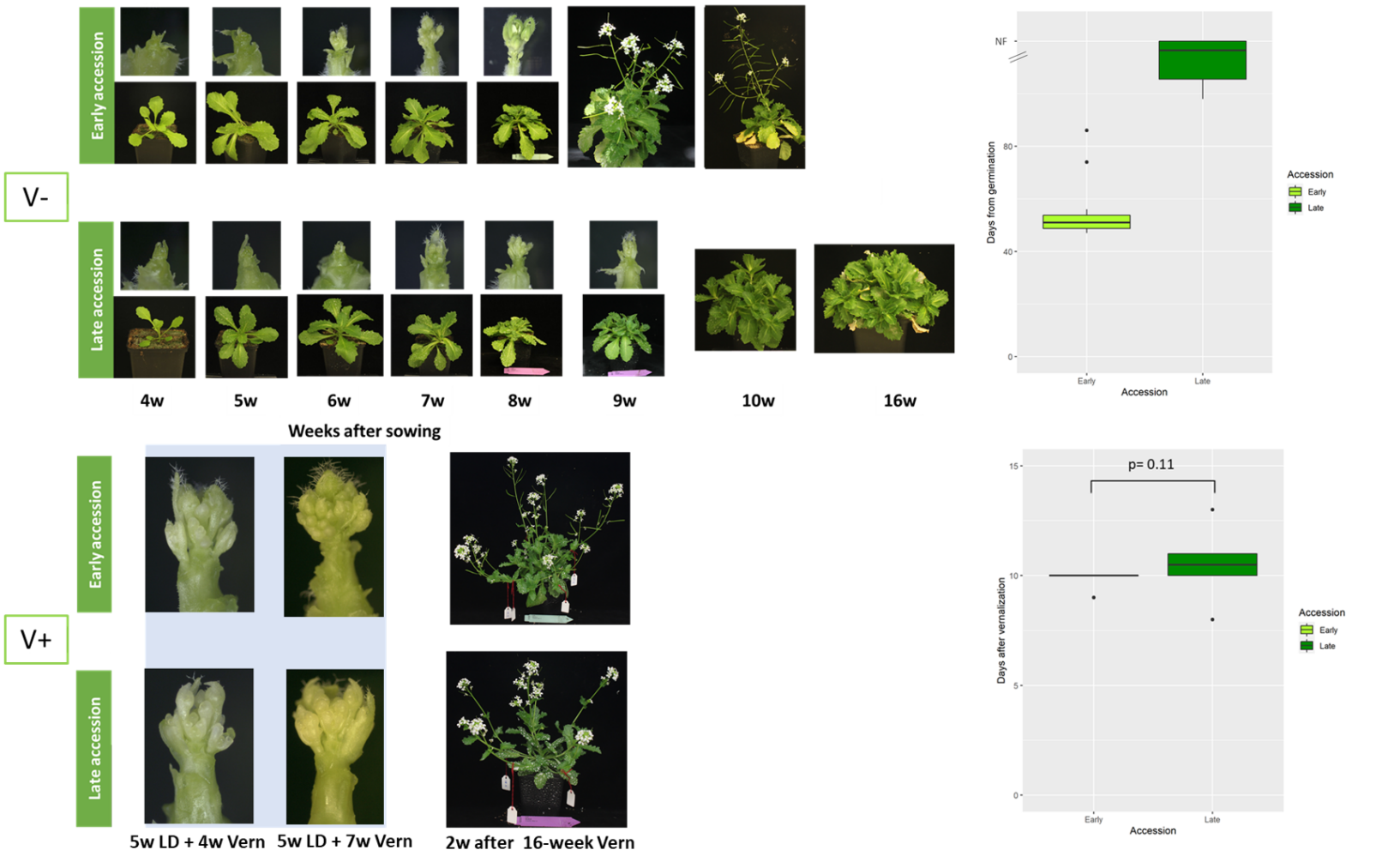

Figure 4.2. Phenotypic differences between early- and late-flowering accessions under growth chamber conditions. The left panels show the microscopic images of the shoot apical meristem and photographs of the whole plant architecture. The upper two rows represent both accessions without vernalization; the lower two rows show both accessions with a 16-week vernalization period. The blue background represents time points during vernalization. The right panels show the flowering-time distributions for the non-vernalization (upper) and vernalization (lower) treatments. The non-parametric Mann-Whitney U test was applied only to the flowering data for vernalized plants. The growth conditions in the Bronson growth chamber were programmed to simulate greenhouse LD condition of 16h light (21°C) and 8h dark (18°C) with 65% humidity, and 16 weeks of vernalization were performed in the vernalization chamber (4°C in 8h light:16h dark) in greenhouse 1. Additional far-red light (100% in the setting) was added to the spectrum in the growth chamber.

accession, where vernalization resulted in an IQR of only 1 day compared with 8 days without vernalization. Thus, although the bud-arrest phenotype was observed for the late-flowering accession without vernalization under growth chamber conditions, a 16-week vernalization treatment abolished the difference in flowering time observed between both accessions without vernalization and reduced the variation in flowering time for each accession after vernalization.

4.3 PRINCIPAL COMPONENT ANALYSIS IS CONSISTENT WITH THE TIMING OF THE APPEARANCE OF THE BUD-ARREST PHENOTYPE

To investigate whether the transcriptomes of the early- and late-flowering accessions differed and also whether differences between time points were present, I performed a principal component analysis (PCA) for the non-vernalization and vernalization data sets. The input for the PCA analysis was a 90 (the total number of biological replicates at all time points) by 36,898 (all genes annotated in Norwegian *A. alpina* accessions) matrix and the results are shown in Figure 4.3. For the non-vernalization time-series, PCA analysis suggested that the major factor that affected the variance (PCl, explaining 79% of the variation) was the developmental stage of the shoot apical meristem, because PCl separated samples in approximate chronological order (time points). Samples from both accessions clustered closely for the two earliest time points (4w and 5w). From the 6w time point onwards, samples from the early-flowering accession were distinguishable from those of the late-flowering accession (marked with an arrow in the left panel of Figure 4.3). This is consistent with the morphological analysis described in section 4.2, which showed that the bud-arrest phenotype in the late-flowering accession became apparent at the 6w time point. After this point, PC2 was maintained at a low value for the early-flowering accession, whereas higher values were observed for the late-flowering accession, which distinguished the two accessions.

For the vernalization data set, the PCA analysis suggested that the major factor that affected the variance (PCl, explaining 72% of the variation) was vernalization: component 1 strongly distinguished time points during and after vernalization, irrespective of accession. Additionally, variation among the three biological replicates during and after vernalization was less than that for the non-vernalized

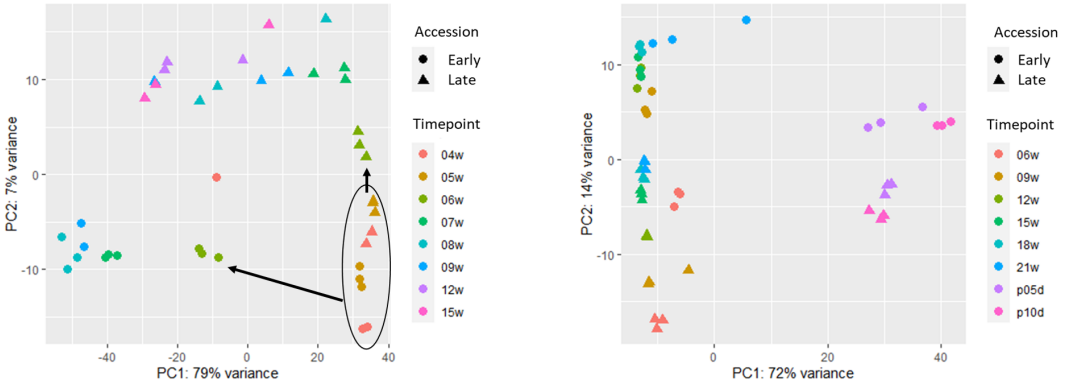

Figure 4.3. Principal component analysis (PCA) for three biological replicate samples for each time point of the RNA-seq sample series under non-vernalization and vernalization conditions. (Left) PCA analysis for the non-vernalization time-series; (Right) PCA analysis for the vernalization time-series. Different symbols indicate different accessions (circles, early-flowering accession; triangles, late-flowering accesssion). Time points are colour coded.

samples, showing that vernalization affected both accessions. Furthermore, the second principle component distinguished the accessions at all time points, with the late-flowering accession consistently having lower values than the early-flowering accession. Intriguingly, although the flowering time after vernalization did not differ significantly between accessions, the PCA analysis separated the early- and late-flowering accessions. Collectively, the results of the PCA analysis provide evidence that changes in gene expression are consistent with the results of the phenotypic analysis in the previous sections; namely, the timing of the appearance of the bud-arrest phenotype in the late-flowering accession without vernalization, as well as the reduced variation between individual plants within each accession after vernalization.

4.4 Expression of *PEP1* in Norwegian *A. alpina*

In the Spanish Pajares accession of *A. alpina*, high expression of *PEP1* inhibits floral initiation and floral development in the absence of vernalization, such that no flowering can occur without exposure to winter cold (Wang et al., 2009a). However, floral initiation was observed in both Norwegian accessions without vernalization, suggesting that floral initiation and floral development are distinct processes and that sequential events in the flowering process are not co-regulated. To investigate how the floral repressor PEPl regulates flowering of the Norwegian

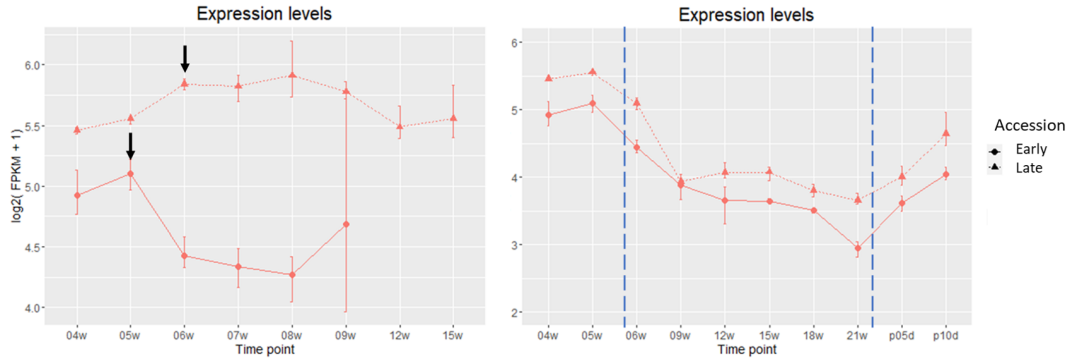

Figure 4.4. The expression of *PEP1* in Norwegian *A. alpina* accessions throughout development. (Left) The normalised expression of *PEP1* in FPKM (fragments per kilobase of transcript per million mapped reads) in unvernalized plants. The black arrows indicates the time at which the floral bud formed. (Right) The normalised expression of *PEP1* in FPKM (fragments per kilobase of transcript per million mapped reads) in vernalized plants. The two blue lines indicates the beginning and end of the 16-week vernalization period. The different symbols indicate the different accessions (circles, early-flowering accession; triangles, late-flowering accession). Expression data are the mean of three biological replicates, and error bars represent the SD.

early- and late-flowering accessions, *PEP1* gene expression was quantified over time for both accessions and was normalised using a scale of log_2 fragments per kilobase of transcript per million mapped reads (FPKM) +1 (Figure 4.4).

In the absence of vernalization, *PEP1* expression in the late-flowering accession was slightly higher than that in the early-flowering accession at all time points. Notably, *PEP1* expression increased in both accessions before floral buds formed (black arrows). After floral bud formation, *PEP1* expression remained high in the late-flowering accession, but decreased in the early-flowering accession as the inflorescence continued to develop. This reduction resulted in a 2-fold difference in *PEP1* expression between both accessions from 6w onwards when the bud-arrest phenotype became apparent in the late-flowering accession.

During vernalization of the Norwegian accessions (the interval between the blue lines in Figure 4.4), the level of *PEP1* mRNA decreased by ~3-fold in both accessions. This decrease is consistent with the continued growth of the floral buds in both accessions during vernalization, (lower row of Figure 4.4). After the plants were transferred to warm conditions following vernalization, *PEP1* transcription was reactivated in both accessions and the difference in *PEP1* expression between both accessions was less than 2-fold. This similar *PEP1* expression level after

vernalization was consistent with the similar flowering time of both accessions (section 4.2).

These data show that *PEP1* is differentially expressed in the Norwegian accessions in the absence of vernalization, but is expressed at a similar level in both accessions during and after vernalization. This is consistent with the phenotype of the accessions.

4.5 The molecular basis for the bud-arrest phenotype in the late accession is independent from floral organ specification

To determine how floral organ development is related to the bud-arrest phenotype and to characterise the molecular status of floral development in Norwegian *A. alpina* accessions, I selected four floral identity genes and quantified their expression levels normalised as fragment counts per kilobase of exon per million fragments mapped (FPKM) over a time course. The molecular functions of these genes in floral organ specification have been well characterised in *A. thaliana* (Chen et al., 2018; Kinoshita and Richter, 2020). The floral meristem identity gene *APETALA 1* (*AP1*) was chosen as a marker of floral initiation, *PISTILLATA* (*PI*) and *APETALA 3* (*AP3*) were selected as markers of petal identity, and *AGAMOUS* (*AG*) was chosen as a marker of carpel and stamen identity. The expression levels of these genes in both accessions are depicted in Figure 4.5 (without vernalization) and Figure 4.6 (with vernalization).

The expression of *AP1* and *PI* increased greatly between 4w and 5w in the early-flowering accession in the absence of vernalization, which correlated with the time at which floral transition occurred according to the morphological examination (Section 4.2). Subsequently, the expression of *AP3* and *AG* increased greatly between 5w and 6w. After 6w, which represents the time at which the early-flowering accession produceed the first open flower on the primary shoot and an elongated inflorescence, the expression of all four genes remained constant. By contrast, in the late-flowering accession, a steep increase of *AP1* and *PI* was observed later than in the early-flowering accession, between the 5w and 6w time points. This correlated with the later floral transition in the late-flowering accession than in the early-flowering accession. Notably, unlike the delay of 1 week

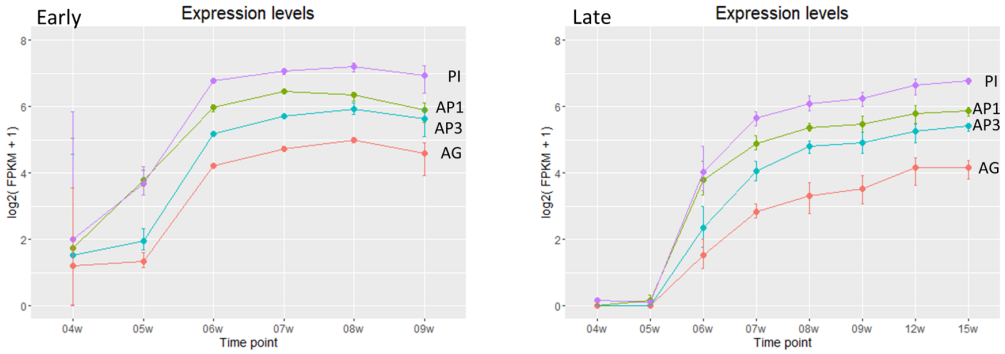

Figure 4.5. The normalised expression of four floral organ identity genes in Norwegian *A. alpina* accessions in the absence of vernalization over time. (Left) The normalised gene expression of *PISTILLATA* (*PI*), *APETALA 1* (*AP1*), *APETALA 3* (*AP3*) and *AGAMOUS* (*AG*) in fragments per kilobase of transcript per million mapped reads (FPKM) for the early-flowering accession without vernalization. (Right) The normalised gene expression of *PI*, *AP1*, *AP3* and *AG* in FPKM for the late-flowering accession without vernalization. Each colour-coded line indicates expression of a single gene: purple, *PI*; green, *AP1*; blue, *AP3*; red, *AG*. Expression data are the mean of three biological replicates, and error bars represent the SD.

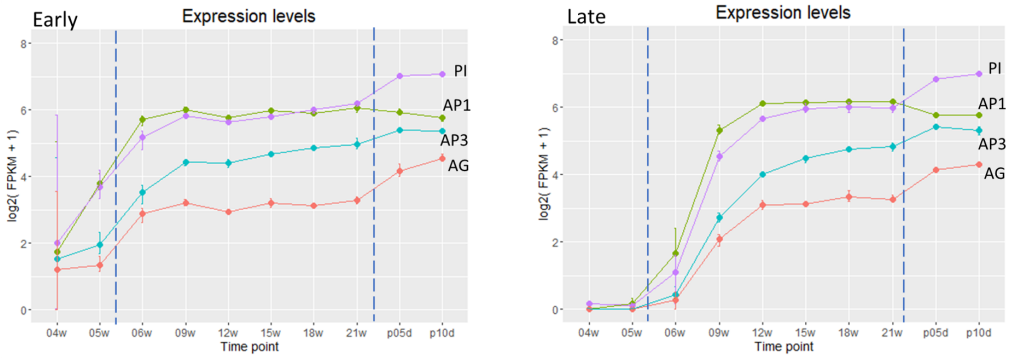

Figure 4.6. The normalised expression of four floral organ identity genes in vernalized Norwegian *A. alpina* accessions over time. (Left) The normalised gene expression of *PISTILLATA* (*PI*), *APETALA 1* (*AP1*), *APETALA 3* (*AP3*) and *AGAMOUS* (*AG*) in fragments per kilobase of transcript per million mapped reads (FPKM) for the early-flowering accession with vernalization. (Right) The normalised gene expression of *PI*, *AP1*, *AP3* and *AG* in FPKM for the late-flowering accession with vernalization. Each colour-coded line indicates expression of a single gene: purple, *PI*; green, *AP1*; blue, *AP3*; red, *AG*. The two vertical blue dashed lines indicate the start and end of the 16-week vernalization period. Expression data are the mean of three biological replicates, and error bars represent the SD.

between the increase in *AP1* and *PI* expression and that of *AP3* and *AG* in the early-flowering accession, the expression of all four genes increased simultaneously in the late-flowering accession and gradually increased further to reach expression levels at the 15w time point similar to those in the early-flowering accession. The gradual increase in expression of the four floral organ identify genes was consistent with the slow growth of floral buds described in section 4.2.

Similar levels of expression for all four genes were observed for both accessions after vernalization for 10 weeks (15w time point). Although the expression levels of all four genes reached a plateau one week after vernalization in the early-flowering accession (6w time point) (Figure 4.6 left), their expression in the late-flowering accession (Figure 4.6 right) also reached a similar level slightly later during vernalization to that in the early-flowering accession. The slow increase of expression in the late-flowering accession might be resulted from the lower pre-vernalization expression level than the early-flowering accession. Thus, these genes had achieved similar expression level from half-way through the vernalization treatment onwards and even after vernalization, consistent with the observed similar flowering time of both accessions after vernalization. Intriguingly, the final expression level of these four genes was relatively similar for each accession either in the presence or absence of vernalization.

Collectively, the data from Figure 4.5 and Figure 4.6 show that similar to its effect on *PEP1* expression, vernalization leads to similar expression levels of *PI*, *AP1*, *AP3* and *AG* in both accessions. In the absence of vernalization, the expression levels of these genes in the late-flowering accession increased slightly later than in the early-flowering accession, but reached similar final levels to those in the early-flowering accession. The presence of phenotypic differences between the accessions in the absence of vernalization, without differences in the final expression levels of four key floral identity genes indicates that the molecular basis for the bud-arrest phenotype is independent from floral organ specification.

4.6 Hierarchical clustering of key floral repressors and organ identity genes

The previous two sections focused on the expression levels of the floral repressor *PEP1* and floral organ identity genes separately. However, this did not clearly indicate the interactions and relationships between these two gene groups.

Therefore, I applied a hierarchical clustering approach to analyse the expression of several selected key floral repressor and floral organ identity genes simultaneously from samples without or with vernalization, and this is summarised as heatmaps as Figures 4.7 and 4.8. Hierarchical clustering was performed via a bottom-up approach using the normalised counts for expression as euclidean distance. The counts were taken from three biological replicates. I normalised the counts of each gene separately, ensuring that they had a mean of zero and a standard deviation of 1, thus producing a Z-score. These Z-scores are visualised with a colour bar, where blue indicates that the normalised counts are below the mean (negative) and red indicates that the normalised counts are above the mean (positive). The naming of the genes in the heatmaps are derived from their orthologues in *A. thaliana*.

In this hierarchical clustering analysis, I selected seven key floral repressor genes and five floral organ identity genes on the basis of their previously reported functions in floral transition. The seven key floral repressor genes included *FLOWERING LOCUS C* (*FLC*), *MADS AFFECTING FLOWERING 1* (*MAF1*) to *MADS AFFECTING FLOWERING 5* (*MAF5*) and *SHORT VEGETATIVE PHASE* (*SVP*). The orthologue of *FLC* in *A. alpina*, *PEP1*, was characterised as the major floral repressor in Spanish accession Pajares of *A. alpina* (Wang et al., 2009a). *MAF1* to *MAF5* was reported to modulate adaptive flowering responses in different Brassicaceae lineages (Madrid et al., 2021b) and *SVP* is a key regulator for floral transition in annual *A. thaliana* (Hartmann et al., 2000) and biennial/perennial species (Wang et al., 2020; Yamagishi et al., 2016). The five floral organ identity genes included the four selected floral homeotic genes from the previous section (*AP1*, *PI*, *AP3* and *AG*) and *SEPALLATA 3* (*SEP3*), which mediates the interactions among floral homeotic proteins in organ-specific protein complexes in *A. thaliana* (Chen et al., 2018).

For the early-flowering accession without vernalization, the analysis revealed a clear separation of floral repressor and floral organ identity genes into two distinct clusters (left panel Figure 4.7). Cluster 1 included six out of seven floral repressor genes (except for *MAF5*) and was characterised by a high expression at the beginning that decreased sharply after 5 weeks. Cluster 2 included all five floral organ identity genes, whose expression was initially low and increased rapidly after 5 weeks. Although the expression of nearly all the floral repressor genes in cluster 1 decreased after 5 weeks, that of *MAF2* and *MAF3* decreased greatly one week earlier (at the 5w time point). However, the expression of *MAF5* fluctuated

throughout the time period and differed from the expression patterns of cluster 1 and 2 (although *MAF5* grouped within cluster 2).

For the late-flowering accession (right panel Figure 4.7), three distinct clusters were observed: the floral repressor genes were divided into two clusters and the third consisted of floral organ identity genes. The expression of the first cluster (*MAF2*, *MAF3* and *SVP*) was high early in development and then decreased, similar to cluster 1, which included six out of seven floral repressor genes, in the early-flowering accession. The second cluster contained all the floral organ identity genes, whose expression showed a similar pattern to that of cluster 2 in the early-flowering accession, although their expression increased one week later (after 6 weeks instead of 5 weeks for the early-flowering accession). The third cluster showed increased expression between 6 and 9 weeks, when the bud-arrest phenotype became apparent in the late-flowering accession. The switch from the expression of floral repressor genes to floral organ identity genes ensures that flowering in the early-flowering accession proceeds rapidly, whereas a different timing of expression of floral repressor and floral organ genes results in slow bud outgrowth in the late-flowering accession.

The hierarchical clustering results for the vernalized plants revealed clear differences in gene expression between the two accessions, even though the flowering time of the the early- and the late-flowering accessions after vernalization was similar. For the early-flowering accession (left panel of Figure 4.8), three large clusters were identified. All five floral organ identity genes belonged to the same cluster, similar to for the non-vernalized dataset, but the floral repressor genes were present in two clusters, which separated *SVP* and *FLC*. Notably, this separation was also observed for the late-flowering accession without vernalization. In the early-flowering accession, the *FLC* cluster was highly expressed prior to vernalization, then expression decreased at the start of vernalization (6w to 21w time point) and slightly increased following vernalization (p5d and plOd time point). However, the *SVP* cluster was similarly expressed before and during vernalization, but expression decreased after vernalization. By contrast, only two large clusters were identified for the late-flowering accession (right panel of Figure 4.8). The first cluster included six out of seven floral repressor genes, which were highly expressed prior to vernalization and then expression decreased. This cluster could be split into two smaller groups: one that included *MAF5*, *SVP* and *MAF1*, whose expression decreased slowly but monotonically, and a second group that included *MAF2*,

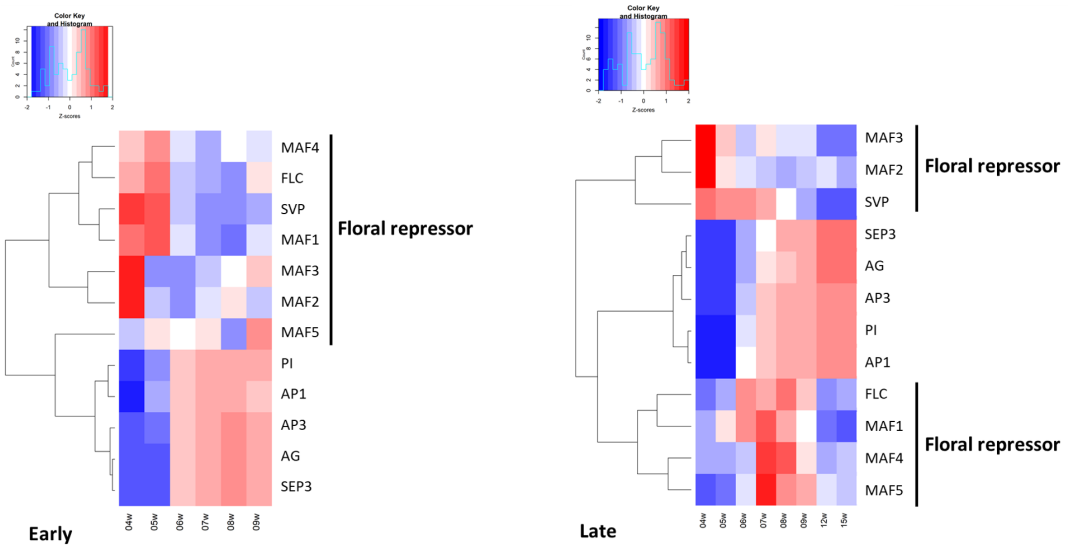

Figure 4.7. Hierarchical clustering of the expression of genes encoding selected key regulators of plant reproductive development in Norweigian *A. alpina* accessions without vernalization. The left and right heatmaps demonstrate the expression dynamic of floral repressor and floral organ identity genes in the early- and late-flowering accessions, respectively. The colour indicates the normalised Z-score, where blue indicates that the normalised counts are below the mean and red indicates that the normalised counts are above the mean. The floral repressor genes are indicated by black lines.

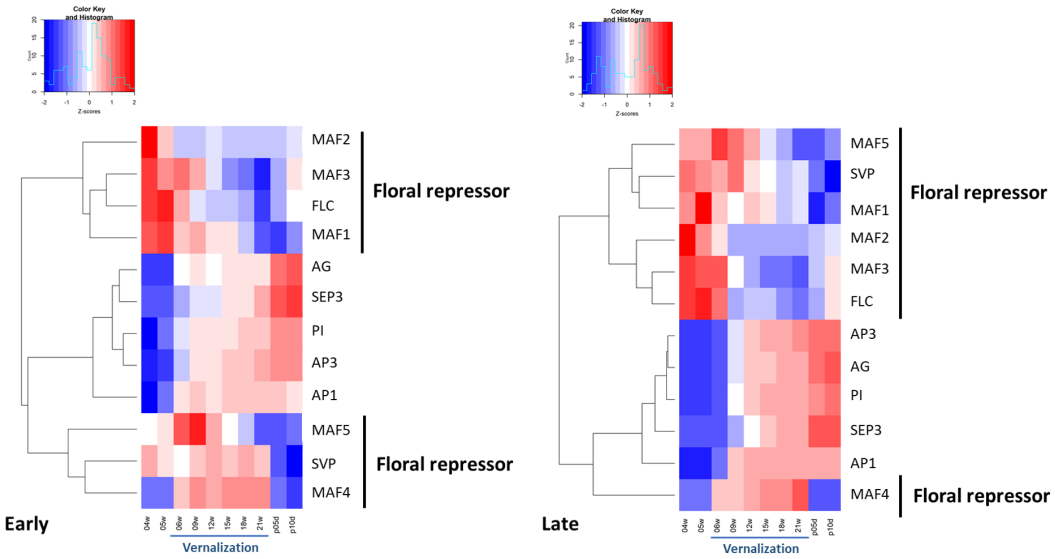

Figure 4.8. Hierarchical clustering of the expression of genes encoding selected key regulators of plant reproductive development in Norweigian *A. alpina* accessions with vernalizationn. The left and right heatmaps demonstrate the expression dynamic of floral repressor and floral organ identity genes in the early- and late-flowering accessions, respectively. The colour indicates the normalised Z-score, where blue indicates that the normalised counts are below the mean and red indicates that the normalised counts are above the mean. The floral repressor genes are indicated by black lines.

MAF3 and *FLC*, whose expression decreased rapidly after the start of vernalization and increased slightly after exposure to warmer temperatures after vernalization. The second cluster included all seven floral organ identity genes, whose expression began to increase when that of the first cluster started to decrease during vernalization, a scenario that represents a switch between the expression of floral repressor genes to floral organ identity genes. Notably, the expression of *MAF4* was strongly induced during vernalization in both accessions, in contrast to the expression of the other gene clusters.

In summary, a transition from the expression of floral repressor genes to floral organ identity genes occurred in the vernalized late-flowering accession and in the early-flowering accession without vernalization. Moreover, although the expression of distinct clusters of floral repressors was observed in the vernalized early-flowering accession, this separation of repressors into multiple clusters was consistent with the two small groups (*SVP* and *FLC*) observed within the large floral repressor cluster in the vernalized late-flowering accession. Intriguingly, despite the presence of distinct FLC and SVP clusters in the vernalized early-flowering accession, the behaviour of these two clusters was similar in both vernalized accessions: the expression of genes in the *FLC* cluster was rapidly downregulated at the start of vernalization, whereas the expression of genes in the *SVP* cluster slowly decreased during vernalization. Therefore, unlike the rapid switch in the unvernalized early-flowering accession, the reduced expression of genes in the FLC and SVP clusters produced an extended transition from the expression of floral repressor genes to floral organ identity genes in both vernalized accessions.

4.7 GENOME-WIDE CO-EXPRESSION NETWORK ANALYSIS REVEALS DIFFERENT CLUSTER STRUCTURES BETWEEN ACCESSIONS

To better characterise different gene clusters on a global level rather than selecting specific genes, and to determine how different floral repressor clusters were correlated with the bud-arrest phenotype, I performed a genome-wide co-expression analysis. Because genes that are involved in the same biological process should be coexpressed (Williams and Bowles, 2004), I used a co-expression network analysis for each accession using the complete RNA-seq data set without vernalization. The

co-expression analysis was constructed using normalised FPKM values, which used robust scaling factors, rather than the total fragment depth. To avoid noise from lowly expressed genes, only the 20,000 genes with the highest mean expression levels were selected. Furthermore, to determine the level of co-expression between genes, the Pearson correlation co-efficient was calculated for each pair of genes; negative correlations were excluded, to ensure that only co-expression was considered rather than anti-expression. The resulting positive correlation co-efficients were transformed into a distance matrix to create the dendrogram.

Once the dendrogram was constructed, I used a distance cut to determine clusters that contained similarly expressed genes that were distinct from genes in other clusters. A small-distance cut resulted in numerous small clusters, because even small differences between gene expression are not permissible within a cluster. Because the distance cut is a free parameter, the optimal value that resulted in a robust selection of clusters had to be determined. To do this, I used a separability metric, S, which compared *inter*-cluster distances, D, with *intra*-cluster distances, d. Here, the separability statistics is defined as:

$$S = 1 - d/D. \tag{4.1}$$

Thus the higher the value of S, the smaller the ratio d/D and the more separated (distinct) are the clusters (Horvath, 2011). The separability metric was calculated for 19 different distance cuts, from 0.05 to 0.5 in steps of 0.05 (two left panels of Figure 4.9). The separability metric closest to 1 (indicated with a black arrow) resulted from a distance cut of 0.2 for the early-flowering accession and a distance cut of 0.225 for the late-flowering accession, both without vernalization. The genome-wide dendrogram and clustering results are shown in the right panels of Figure 4.9, with the genes belonging to the same cluster represented by the same colour. Clusters with fewer than 30 genes were filtered out to better identify only the most meaningful clusters. Clustering in this way resulted in 44 and 52 clusters for the early- and late-flowering accessions, respectively. The 44 clusters in the early-flowering accession include 8,871 genes in total; the 52 clusters in the late-flowering accession include 9,301 genes in total.

Comparison of the two dendrograms in Figure 4.9 revealed a unique clustering of gene expression for each accession: the early-flowering accession possessed

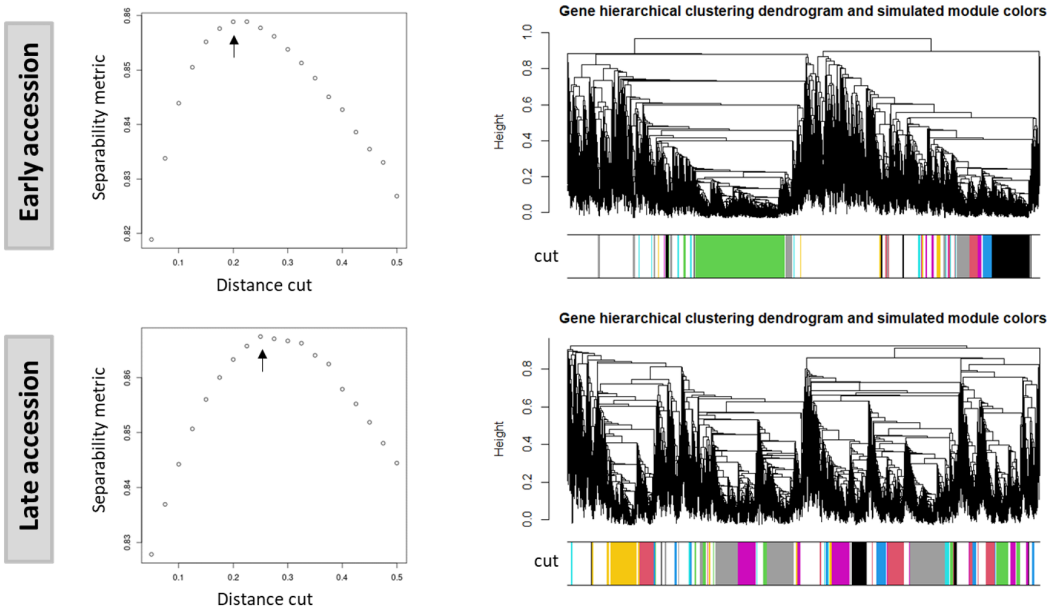

Figure 4.9. Genome-wide co-expression clustering analysis for Norweigian *A. alpina* accessions without vernalization. (Left) The separability metric as a function of dendrogram distance was used to define the optimal clustering result. The black arrows indicate the value where the distance cut results in the highest value of separability metric (S) closest to 1 for each accession. (Right) The dendrograms for expression of the 20,000 genes, with the optimal clusters shown by colour. The top row shows the results for the early-flowering accession, and the bottom row shows the results for the late-flowering accession.

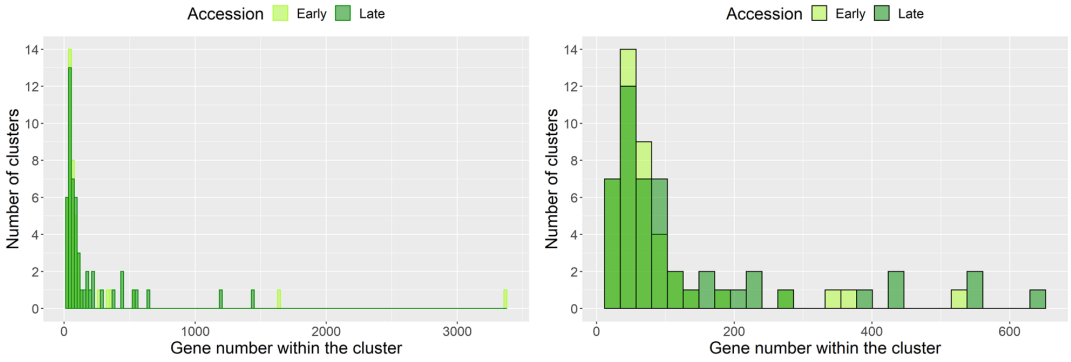

Figure 4.10. Genome-wide structure of gene clusters in Norweigian *A. alpina* accessions. (Left) Histograms showing the distribution of the number of genes within individual clusters for the early- and late-flowering accessions without vernalization. (Right) The same data as in the left panel but only clusters containing fewer than 1,000 genes are shown. The early- and late-flowering accessions are colour coded in yellow and green, respectively.

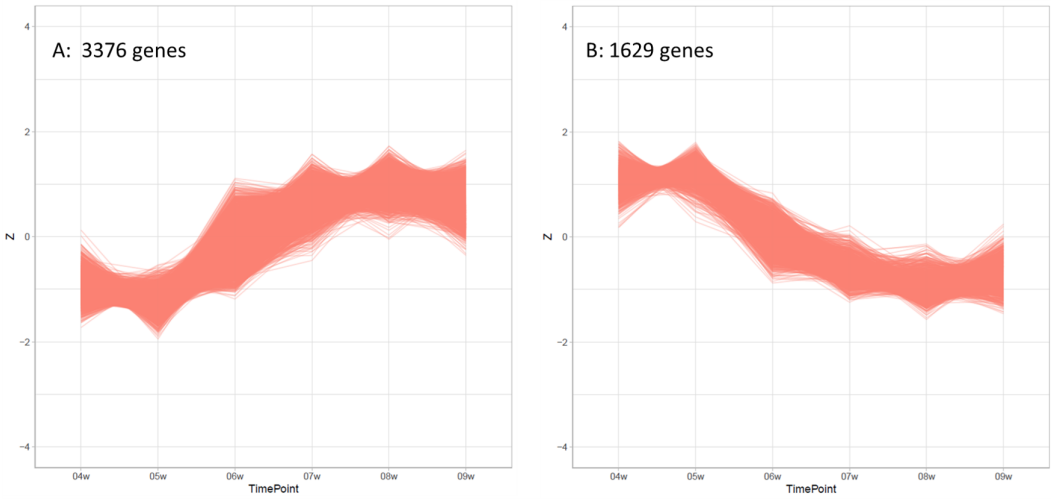

Figure 4.11. Expression pattern of the two largest gene clusters in the Norweigian *A. alpina* early-flowering accession without vernalization. The expression level is shown as normalised Z-scores. The number in the top left of each panel shows the number of genes in each cluster.

two large gene clusters (labelled green and black), whereas the late-flowering accession contained a more even distribution of cluster sizes. To better characterise the cluster structure in each accession, I analysed the number and identities of the genes in the clusters for both accessions. The distribution of the sizes of the clusters for both accessions was visualised using the histograms in Figure 4.10. The two distinct large gene clusters in the early-flowering accession dendrogram contained 3,376 and 1,629 genes, which comprised 56% of the total number of genes in all clusters. Notably, the expression pattern of the largest gene cluster (left panel of Figure 4.11) was similar to that of the floral organ identity gene cluster in the previous section; expression was initially low and then increased monotonically after 5 weeks. The second-largest gene cluster (right panel of Figure 4.11) showed the opposite pattern of expression, which decreased after 5 weeks and was very similar to the expression of the floral repressor gene cluster in the section 4.6. This second-largest cluster included the floral repressor genes *SVP*, *MAF1* as well as *TERMINAL FLOWER1* (*TFL1*), reported to be a floral repressor in the age-dependent response to vernalization in *A. alpina* (Wang et al., 2011). The genes in these two large clusters with opposing expression patterns represent a gene regulatory network that might allow the early-flowering accession to flower without vernalization.

Similar to the early-flowering accession, the late-flowering accession contained two large clusters, which contained 1,188 and 1,434 genes (bottom right panel of Figure 4.9); however, these represented only 28% of the total number of genes in all clusters, whereas the two large clusters in the early-flowering accession represented 56% of the total. The majority of clustered genes in the late-flowering accession belonged to 17 clusters containing from 100 to 1,000 genes and representing 53% of the total number of clustered genes. This difference is seen in Figure 4.10, where the early-flowering accession had only nine clusters containing between 100 and 1,000 genes. Intriguingly, although most of the floral repressor genes were present within a single large cluster in the early-flowering accession, those in the late-flowering accession were distributed among several intermediate-sized clusters.

In section 4.6, the expression patterns of the floral repressor genes such as *FLC*, *MAF1* and *MAF5* caused them to be clustered together; here, the genome-wide co-expression network split them into separate clusters. Figure 4.12 shows the expression patterns of six selected gene clusters from the late-flowering accession. The cluster containing *FLC* (top left panel of Figure 4.12, cluster A) consisted of 84 genes and the cluster containing *MAF1* (top middle panel of Figure 4.12, cluster B) contained 81 genes. The expression of genes in both clusters increased from 4 to 6 weeks, when floral buds formed, but expression of the *MAF1* cluster peaked at 7 weeks and that of genes in the *FLC* cluster peaked at 8 weeks. This differential timing of maximum expression was sufficient to lead to the formation of different clusters, showing the high degree of resolution obtained from the co-expression network analysis. In contrast to these two clusters, the cluster containing *MAF5* (top right panel of Figure 4.12, cluster C) showed a sharper increase in expression before peaking at 7 weeks; this cluster was also over 4-fold larger than those containing *FLC* or *MAF1*. Moreover, although *SVP* and *TFL1* were present in the same cluster in the early-flowering accession, they were separated into different clusters in the late-flowering accession (bottom left and middle panels of Figure 4.12, clusters D and E). This is because expression of genes in the cluster containing *SVP* (D) reached a plateau/slightly increased between 4 and 6 weeks (during floral bud formation) and then decreased, whereas expression of genes in the cluster containing *TFL1* (E) showed a monotonic decrease. Notably, the cluster containing *TFL1* was the second-largest cluster (1,188 genes) and showed an expression pattern analogous to the cluster of floral repressor genes in the early-flowering accession (right panel of Figure 4.11). The expression of

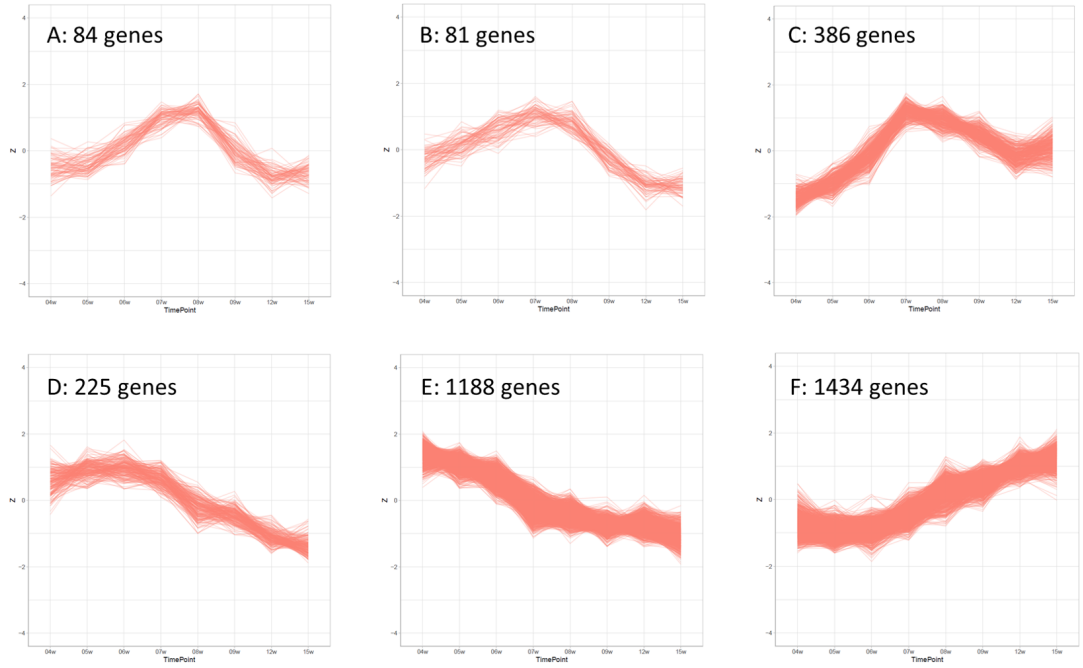

Figure 4.12. The expression patterns of genes in six selected gene clusters for the Norweigian *A. alpina* late-flowering accession without vernalization. The expression patterns were plotted using Z-scores. The text in each panel indicates the name of each cluster and the number of genes it contains.

genes in the largest cluster for the late-flowering accession (bottom right panel of Figure 4.12, cluster F) was opposite to that of the second-largest cluster, cluster E, and increased monotonically with time, similarly to the largest cluster of the early-flowering accession, which contained the floral organ identity genes.

For the early-flowering accession, the majority of floral organ identity and floral repressor genes belonged to two clusters that were expressed in opposite manners; however, for the late-flowering accession, the floral repressors were split among multiple clusters that showed similar expressions patterns, but asynchrony in the timing of the peak expression level. This splitting of floral repressors into multiple, asynchronous clusters suggests that multiple gene regulatory networks might lead to a protracted development of floral buds, resulting in the bud-arrest phenotype observed for the late-flowering accession.

4.8 DISCUSSION

The RNA-seq experiment provided data with which to decipher the molecular mechanisms and gene regulatory networks (GRNs) responsible for variation in flowering-time in the Norwegian *A. alpina* early- and late-flowering accessions. The difference in flowering behaviour between accessions under growth chamber conditions without vernalization validates the robustness of the bud-arrest phenotype of the late-flowering accession. Specifically, I have conclusively shown that without vernalization in the growth chamber, flowering occurs via two pathways. First, the early-flowering accession shows coupled floral transition (i.e. floral bud formation) and inflorescence development, which allows the primary shoot to elongate, flower and form a mature inflorescence immediately after floral initiation. Second, the late-flowering accession undergoes floral transition, but the primary shoot neither elongates nor produces open flowers. These findings support the notion that flowering is regulated at steps subsequent to floral transition, consistent with findings from another perennial species, *Arabidopsis lyrata* (Kemi et al., 2019).

Notably, unlike the Spanish Pajares accession of *A. alpina*, the reduction in *PEP1* expression in the Norwegian early-flowering accession occurs independently of vernalization. This suggests that the GRN that regulates flowering differs between the Norwegian and Spanish accessions. I also showed that *PEP1* expression increased in both accessions without vernalization when floral transition occurred. However, there is an alternative explanation for this observation: *PEP1* regulates the the flowering process subsequent to floral initiation without vernalization, because *PEP1* expression is maintained at a high level in the late-flowering accession compared with the early-flowering accession.

Here, I propose a conceptual model for the genetic regulation of flowering in the Norwegian accessions of *A. alpina* compared with that in the Spanish reference accession Pajares (Figure 4.13). The bar-headed lines in the model do not necessarily indicate direct inhibition of target genes and an additional transcription factor(s) might be involved in a longer transduction cascade. The letters X, Y and Z represent genes that promote floral transition, whereas letters A, B and C represent genes that promote inflorescence development. The high expression of *PEP1* in the Spanish reference accession in the absence of vernalization inhibits floral transition (X, Y and Z genes) and subsequent inflorescence development (A,

B and C genes), making it difficult to study the relationship between these processes. However, I propose that specific genes that are inhibited by PEP1 or other MADS-box proteins in Pajares, function independently of PEP1 or MADS-box proteins in the Norwegian early- and late-flowering accessions and promote floral transition in both accessions without vernalization (shown as the dashed line, gene Z). Moreover, in the early-flowering accession, specific genes can bypass regulation by PEP1 or other MADS-box proteins to promote inflorescence development (one such gene is shown as the dashed line, gene C), whereas in the late-flowering accession, these genes are inhibited, which results in bud-arrest. Although the expression of *PEP1* is always higher in the late-flowering accession than in the early-flowering accession, the function of *PEP1* in floral transition is unclear in both accessions. Furthermore, I showed that the final expression of key floral organ identity genes was similar in both accessions, independent from vernalization (Figure 4.5). I speculate that subsequent to floral initiation, the formation of an elongated inflorescence in the Norwegian accessions is inhibited, but not flower development.

Consistent with the studies from Michaels and Amasino (2000) and Wunder et al., (in preparation), I also showed that vernalization could overcome the late-flowering phenotype as well as variation in flowering time or the Norwegian *A. alpina* accessions. Specifically, I have conclusively shown that *PEP1* expression decreases during vernalization and is reactivated after vernalization, resulting in no difference in flowering time between both accessions after vernalization. This observation further supports the notion that this decrease in *PEP1* expression allows the continual growth of floral bud during vernalization, whereas the reactivation of *PEP1* transcription ensures continued vegetative development in *A. alpina* as part of its polycarpic life history. Although no statistically significant difference in flowering time was observed between the accessions when they were vernalized, the early-flowering and late-flowering accessions were still distinct from each other in the PCA analysis (Figure 4.3). An intriguing possibility is that differences in other features instead of flowering time exist between the accessions, even after vernalization, which is in agreement with section 3.4 in the previous chapter.

In this chapter, I provide the first direct evidence for the relationships between key floral repressor and floral organ identity genes in Norwegian *A. alpina* accessions. A transition from the expression of floral repressor to floral organ identity genes appears to occur in non-vernalized early-flowering accession plants and

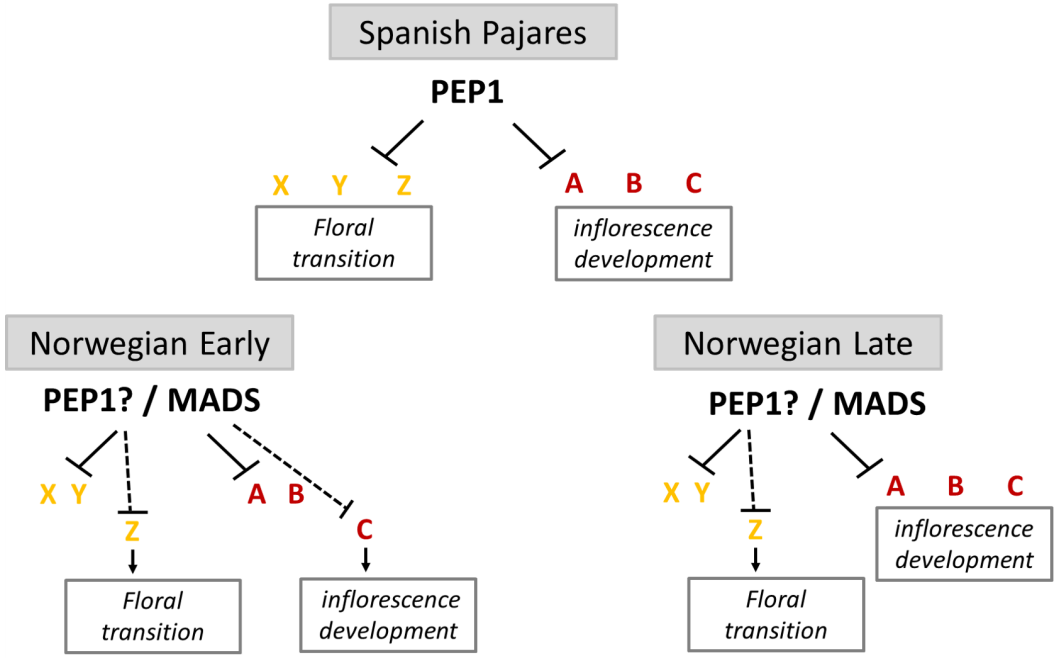

Figure 4.13. A conceptualised model for flowering in the Norweigian *A. alpina* early- and late-flowering accessions compared with Spanish reference accession Pajares without vernalization. The bar-headed lines in the model do not necessarily indicate direct inhibition of target genes and an additional transcription factor(s) might be involved in a longer transduction cascade. The letters X, Y and Z represent genes that promote floral transition, whereas letters A, B and C represent genes that promote inflorescence development. The dashed lines indicate the bypass of the inhibition.

vernalized late-flowering accession plants, all of which produce an elongated primary shoot inflorescence. This leads to the speculation that if floral transition is coupled with the subsequent immediate formation of an elongated inflorescence, a switch from the expression of floral repressor genes to floral organ identity genes is involved.

Additionally, analysis of the vernalization dataset by hierarchical clustering demonstrated the similar membership of genes within each *FLC* and *SVP* cluster in both accessions. Notably, the *FLC* cluster always showed a sharp decrease in expression at the start of vernalization, whereas the *SVP* cluster showed an extended decrease in expression during vernalization. This difference in the timing of the decrease in expression of the *FLC* and *SVP* clusters is a shared feature of both accessions during vernalization. I also show that in the late-flowering accession, *FLC* and *SVP* clustered with specific *MAF* genes, irrespective of vernalization. These results support a previous suggestion that the regulatory networks controlled by *SVP* and *FLC* contribute to variation and robustness in flowering time (Mateos et al., 2015). Although evidence suggests that the effect of *MAF* genes depends on *FLC* in the Spanish Pajares accession (Madrid et al., 2021b), the expression pattern of *MAF* genes in the Norwegian late-flowering accession but not in the early-flowering accession was more similar to that in the *pep1* mutant of the Spanish Pajares accession. Comparison of the expression of *MAF* genes in the Norwegian and Spanish Pajares accessions during vernalization also showed that only *MAF3* and *MAF5* were similarly expressed. Probably, the GRN that controls *MAF* gene expression varies among different *A. alpina* populations.

Here, I provide the first direct evidence for the molecular mechanism underlying the variation in flowering behaviour in Norwegian *A. alpina* accessions. The application of genome-wide co-expression network analysis has allowed different cluster structures and GRNs to be identified between two accessions, which might partially explain the bud-arrest phenotype or flowering process without vernalization on a molecular level. Mutagenesis of the floral repressors described in this chapter might provide deeper insight into identifying the key regulators of flowering in *A. alpina* under non-vernalization conditions.

Chapter 5
Genetic analysis of variation in bud outgrowth of early- and late-flowering accessions

5.1 Background

To unravel the genetic mechanism underlying the variation in bud outgrowth on the primary shoot without vernalization in Norwegian *A. alpina*, Dr. Ulla Kemi previously performed quantitative trait locus (QTL) mapping on progeny produced by crosses between the early- and late-flowering accessions under greenhouse conditions. The first population, F1, consisted of 28 individuals and was generated by crossing the early- and late-flowering accessions. In general, the F1 population displayed a flowering-time phenotype that was intermediate between that of the early- and late-flowering accessions; however, the flowering time of three individuals of the F1 population was similar to that of the early-flowering accession. These individuals were self-pollinated to produce the F2 population, which consisted of 390 individuals. The flowering time of the F1 and F2 populations is presented as histograms in Figure 5.1; the three F1 individuals used to generate the F2 population are also indicated. 19 and 18 plants from the early- and late-flowering parental accessions (after one self-pollination) were also included in the experiment as phenotypic controls for flowering time (Kemi et al., in preparation).

From the 390 individuals of the F2 population, 353 were used for QTL

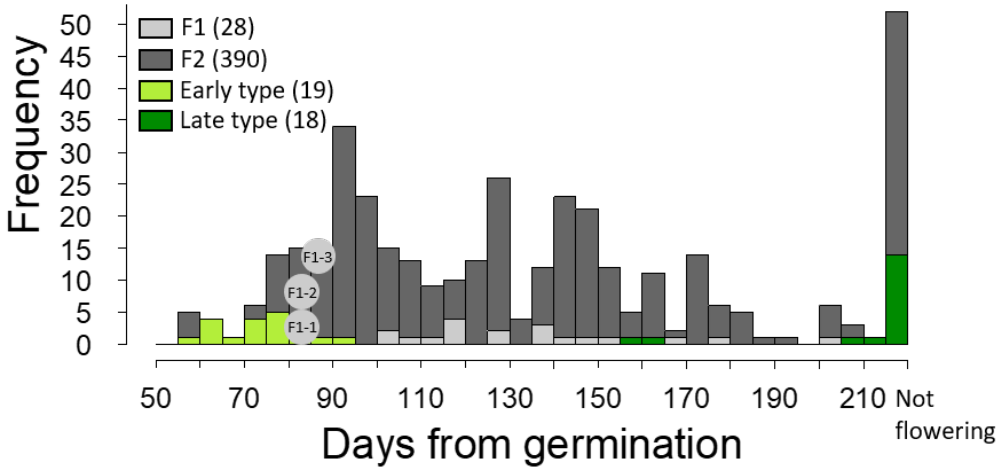

Figure 5.1. Distribution of flowering time data for F1 and F2 populations of Norwegian *A. alpina* in the greenhouse. Each column represents the number of days until the first flower opened on the primary shoot of individual plants grown in the greenhouse without vernalization. Colours represent the different populations: yellow, early-flowering accession; green, late-flowering accession; grey, the F1 population; and dark grey, the F2 population. The number of individuals in each generation is shown in brackets. The plot was originally provided by Dr. Ulla Kemi and modified by Yi-Chen Lin for this thesis.

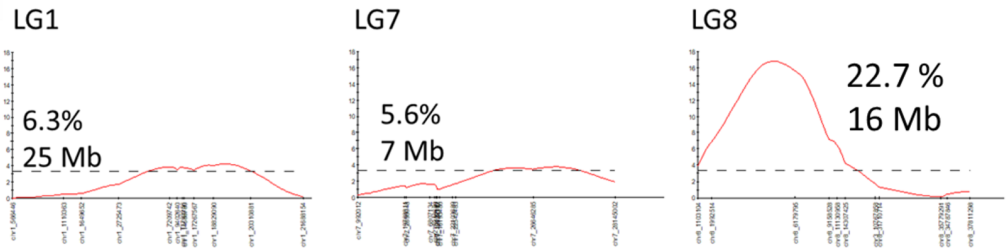

Figure 5.2. Limit of detection (LOD) curve of the simple interval mapping of Norwegian *A. alpina* for the number of days until the first flower opens in the greenhouse in the absence of vernalization. Three significant QTL regions were identified using a 5% LOD significant level. This plot was originally produced by Dr. Ulla Kemi and modified by Yi-Chen Lin for this thesis.

mapping. Each F2 plant individual was genotyped for 128 single-nucleotide polymorphisms (SNPs) by Sequenom technology (Sequenom, San Diego, CA) at the Spanish National Genotyping Centre (CeGen). The SNPs were designed based on the whole-genome sequencing of the early- and late-flowering accession and were designed to be homozygous in each accession. The eight linkage groups corresponding to the eight chromosomes of *A. alpina* were constructed with JoinMap version 4 (Van Ooijen, 2006). For the QTL analysis of a single segregating population, a LOD score of 5% was used to define the genome-wide significance, i.e., the probability of obtaining a LOD score above the threshold somewhere within the whole genome by chance was 5% (Van Ooijen, 1999). A 5% LOD significant level revealed QTL regions in linkage group 1, 7 and 8 (LG1, 7 and 8) with 10,000 permutations identified using parametric QTL mapping analyses (interval mapping, MapQTL version 6) (Van Ooijen, 2009). Figure 5.2 depicts the results of the interval mapping: the major QTL region in LG8 (16Mb) explained 22.7% of the variation in flowering time; the minor QTL regions in LG1 (25Mb) and LG7 (7Mb) explained 6.3% and 5.6% of the variation, respectively.

A strategy of continuously backcrossing an early-flowering plant from the F2 population to the parental late-flowering accession was implemented to enable the sole investigation of linkage group 8. The scheme for fine mapping is shown in Figure 5.3 and the ultimate goal was to produce near-isogenic lines in the late-flowering background that carried a chromosomal segment(s) from the early-flowering accession that promoted early flowering (Kemi et al., in preparation). Therefore, early-flowering plants were identified and selected in each backcrossed generation. The number of selected early-flowering plants based on the desired genotype in each generation is indicated in Figure 5.3. After three backcrossed generations (BC3), LG1 and LG7 were fixed in the late-flowering homozygous background. To avoid genes outside the linkage groups being transferred to each new generation produced via continuous backcrossing, investigation of the self-pollinated lines containing the fixed linkage groups was preferred. Therefore, in the BC3 generation, six early-flowering plants were selected to represent six reference families (31, 67, 110, 425, 431 and 467) and these were used to produce the self-pollinated lines. In the BC3F1 generation, only nine early-flowering plants from three reference families (67(s), 425(s) and 467(s)) were further chosen on the basis of the desired genotype. The expanded square on the right of Figure 5.3 illustrates the detailed fine-mapping scheme adopted after the BC3 generation. The

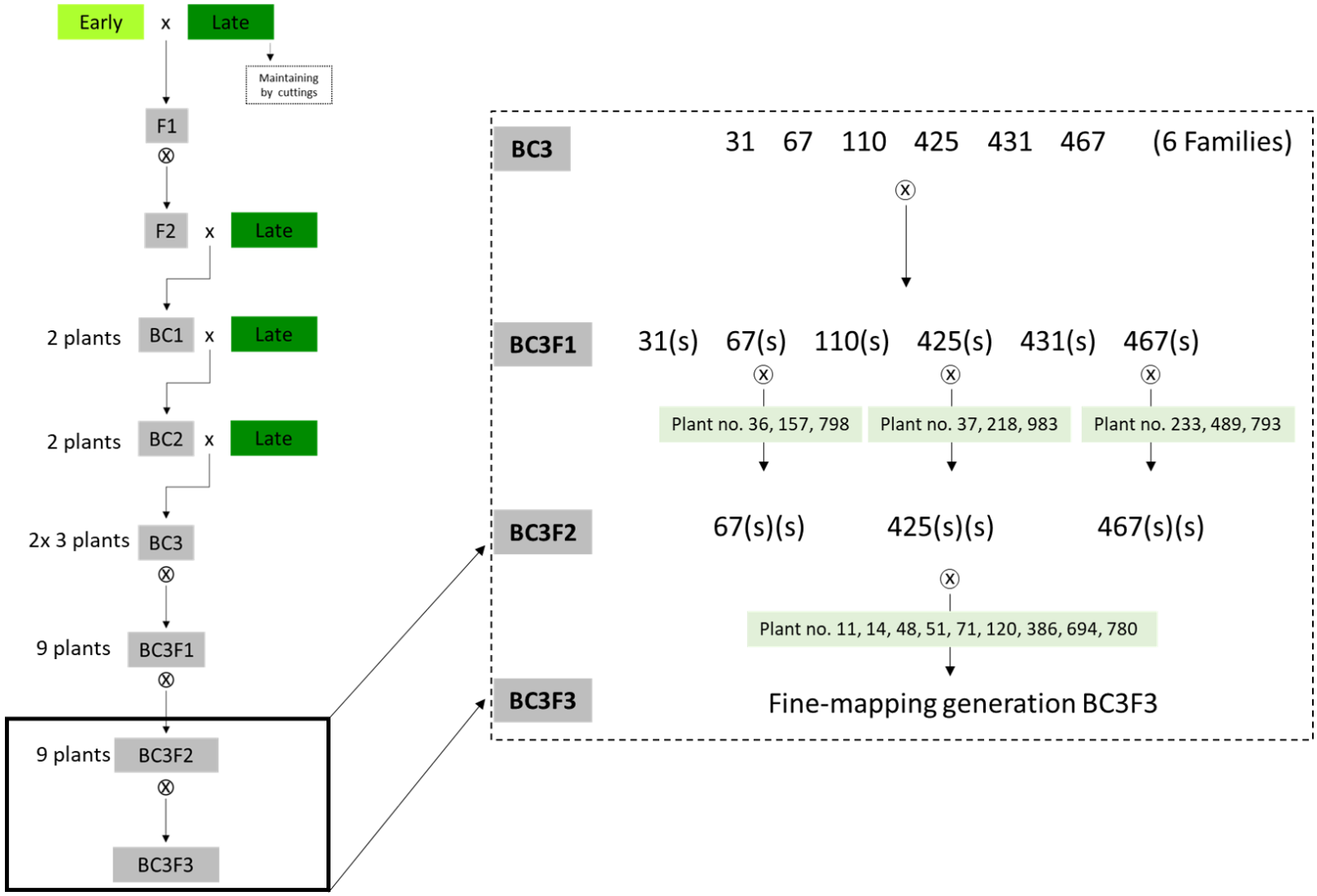

Figure 5.3. A schematic representation of the fine-mapping strategy for linkage group 8 that confers early flowering to *Arabis alpina*. The cross symbol indicates cross-pollination to produce the subsequent generation, whereas a cross within a circle indicates self-pollination. The black square highlights the two generations grown by me. A detailed scheme of the fine-mapping strategy after the BC3 generation is represented by a square with a dashed line. The plant ID numbers (abbreviated as Plant no.) selected from BC3F1 to BC3F3 are highlighted in green. This plot was originally provided by Dr. Ulla Kemi and modified by Yi-Chen Lin for this thesis.

self-pollinated lines following the three backcrossed generations allowed the QTL region in linkage group 8 to be narrowed down from the original 16 Mb to 772 kb. In this chapter, I will only focus on the two fine-mapping generations (BC3F2 and BC3F3) grown by me, marked by the black square on the left of Figure 5.3. The additional genetic markers used in the BC3F2 generation were designed by Dr. Eva Madrid and plants were genotyped by Andrea Lossow. For the BC3F3 generation, the additional genetic markers were designed by Dr. Niharika Sashidhar and plants were genotyped by Kerstin Luxa and Andrea Lossow, together with Dr. Niharika Sashidhar. All the genetic markers used in these two generations are summarised in Table 2.3 and 2.4. The scoring of plants in the greenhouse for the number of days until the first flower opened was always conducted by me to ensure consistency in defining the flowering time of the two generations. I also performed the final statistical analysis on the phenotypes and genotypes of these two generations.

5.2 Fine-mapping using the BC3F2 generation

To further narrow down the QTL region on LG8 that contained the causal gene(s) for early flowering, I performed fine-mapping by analysing the progeny from families 67(s), 425(s) and 467(s) selected in the BC3F1 generation (see the square with the dashed line in Figure 5.3). All plants were grown without vernalization in the greenhouse from May to October 2020. The plant material grown in this experiment included 255 progeny of family 67(s), 170 progeny of family 425(s) and 255 progeny of family 467(s). The 255 progeny of family 67(s) were produced by self-pollinating individuals with ID numbers 36, 157 and 798 in the BC3F1 generation; the 170 progeny of family 425(s) were produced by self-pollinating individuals with ID numbers 37, 218 and 983 in the BC3F1 generation, and the 255 progeny of family 467(s) were produced by self-pollinating individuals with ID numbers 233, 489 and 793 in the BC3F1 generation. Twelve plants from both the early- and late-flowering parental accessions (after one self-pollination) were also included in the experiment as phenotypic controls for flowering time. Plants from different families and the control plants were fully randomised to ensure each population was equally distributed on each table in the greenhouse.

The 67(s), 425(s) and 467(s) families were selected because they contained different lengths of the QTL interval (right panel of Figure 5.4). Progeny from

family 425(s) had the shortest QTL interval (763 kb), whereas progeny from families 67(s) and 467(s) had longer QTL intervals of about 2.2 Mb. In families 67(s) and 467(s), LG1 was fixed with the late-flowering genetic background; however, this was not the case for two individuals (plant ID numbers 37 and 218) from family 425(s). In plant ID number 37, LG1 was fixed with the early-flowering genetic background, whereas in plant ID number 218, LG1 was mostly fixed with the late-flowering genetic background except for a single SNP (SNP-129). Despite the major focus on LG8, the material from these two individuals allowed the parallel investigation of LG1. In addition, newly designed InDel primers (see Table 2.3) were used to increase the resolution of the fine mapping.

The time until the opening of the first flower was scored for the progeny of the three representative individuals from families 67(s), 425(s) and 467(s) and these flowering-time data are displayed as boxplots in the left panel of Figure 5.4. The three selected families in Figure 5.4 are abbreviated as 67(s)(s), 425(s)(s) and 467(s)(s) to indicate the data points from the BC3F2 generation with the corresponding colours. Notably, all of these flowering distributions were relatively broad, indicating that variation in flowering time was a common feature of all three families. Intriguingly, several plants that flowered earlier than the early control were observed in all nine flowering-time distributions, indicating that by fixing one genomic region with the late-flowering genetic background produced extremely early-flowering plants. Furthermore, the flowering-time distribution of progeny from plant ID number 37 in family 425(s), in which LG1 was fixed with the early-flowering homozygous background, showed the lowest median flowering time out of all 10 groups considered here, suggesting that LG1 influences LG8. The data points plotted with a value of 180 days represent plants that did not flower before the end of the experiment.

Progeny from family 425(s)(s) in the BC3F2 generation were chosen for the next fine-mapping experiment, because this family had the shortest QTL region on LG8, but also allowed the effect of LG1 on flowering to be tested. However, prior to the next fine-mapping experiment, I used multiple statistical approaches to determine how genetic markers in LG8 were associated with the phenotype in this family. For this, I analysed 33 plant progeny from plant ID numbers 218 and 983, in which the whole of LG1 was fixed with the late-flowering genetic background. To avoid biases resulting from the inclusion of non-flowering plants, two such individuals were removed, resulting in a sample of 31 plants. The flowering-time distributions

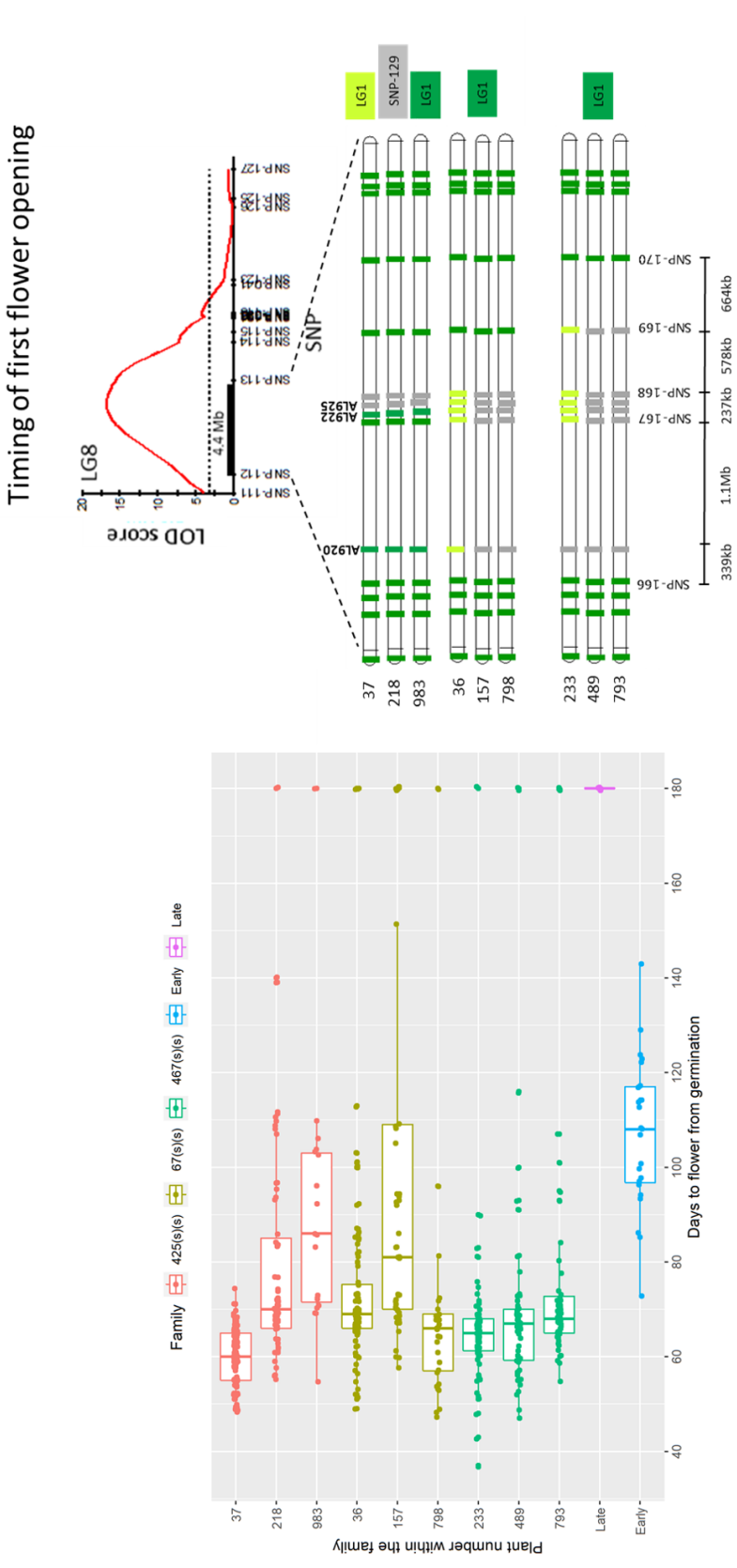

Figure 5.4. Flowering-time distribution for individuals from three different selected *Arabis alpina* families and the corresponding genotype for markers in LG1 and 8. (Left) The flowering-time distribution is represented as boxplots with the colour code of each family. (Right) The genotype of three selected individuals with unique ID number in each family from generation BC3F1 with the genotype of genetic markers in the QTL interval: (green) late homozygous background, (yellow) early homozygous background, (grey) heterozygous background. The right plot was originally provided by Dr. Ulla Kemi and modified by Yi-Chen Lin for this thesis.

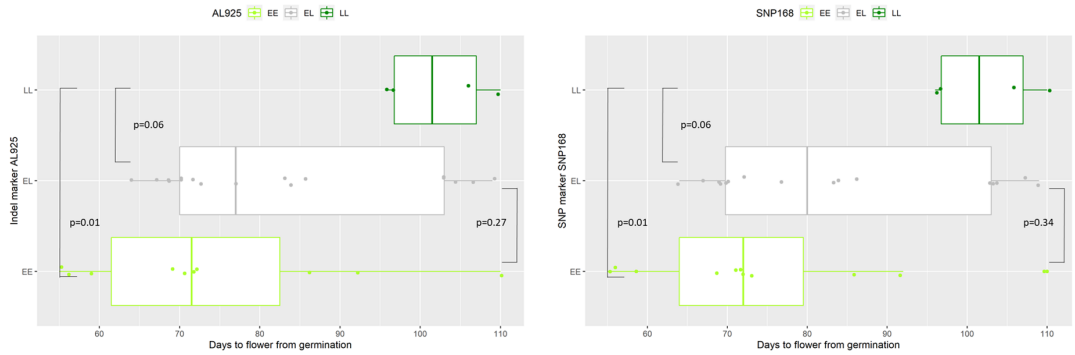

Figure 5.5. The flowering-time distribution of 31 individuals according to their allelic genotype for the genetic markers AL925 and SNP-168 within the QTL interval. (Left) The flowering-time data from different allelic combinations for InDel marker AL925 in the QTL interval. (Right) The flowering-time data from different allelic combinations for SNP marker SNP-168 within the QTL interval. The Anderson-Darling test was performed to statistically compare the distributions.

of these 31 plants are shown in Figure 5.5, according to their allelic combinations at the genetic markers AL925 and SNP168 within the QTL interval. Notably, the number of individuals with the allelic combination LL (late-type homozygous) was much lower than that with early-type homozygous and heterozygous combinations (EE and EL, respectively).

Due to the large variation in flowering time, which resulted in a long tail for each flowering-time distribution, I used the non-parametric two-sample Anderson-Darling test to determine whether the flowering-time distributions of each allelic combination were statistically significantly distinct from each other (Pettitt, 1976). The *p*-values for each pairwise comparison are shown in Figure 5.5. Comparison of the late homozygous and heterozygous distributions resulted in a *p*-value of 0.06 for both markers, which suggests that a difference can still exist, but that is difficult to detect due to the low sample size of individuals fixed as late homozygous. When the difference between the early and the late homozygous combinations was considered, a low *p*-value of 0.01 was found for both markers, showing that a strong statistically detectable difference in flowering time existed when both alleles were derived from the late- and early-flowering accessions. This indicates that the allele from the early-flowering accession contributes to the early-flowering phenotype.

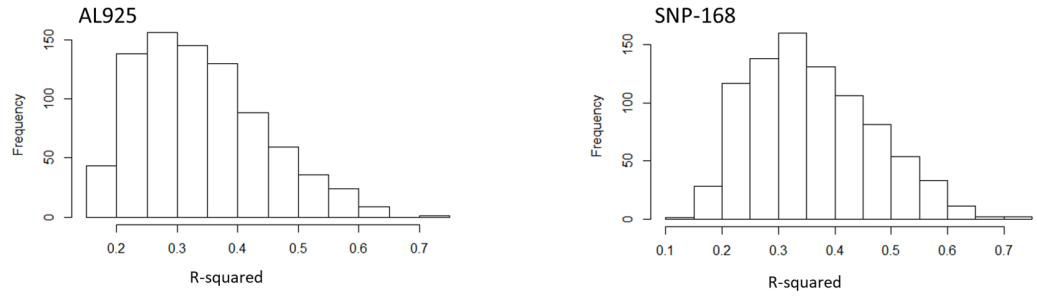

Figure 5.6. The R-squared distributions resulting from 1,000 bootstrap iterations of linear regression modelling of flowering time with respect to allelic combinations of the AL925 and SNP-168 genetic markers. Each distribution only includes the values of R-squared when the corresponding p-value was less than 0.05. On the left is the distribution for InDel marker AL925, and on the right is the distribution for the SNP marker SNP-168 within the QTL interval.

Linear regression was used to investigate whether differences in flowering-time can be explained by the three allelic combinations of either marker. Bootstrapping was performed to quantify the variation in the results due to the size of the data sample. After each bootstrapping iteration (1,000 in total), the linear model was fitted and the coefficient of determination (R-squared) was calculated. Figure 5.6 shows the distribution of the resulting R-squared values when the p-value of the fit was less than 0.05. In total, about 880 out of 1,000 bootstrap iterations produced a significant p-value. The most frequent R-squared value for both markers was approximately 0.3, which indicates that the allelic combinations can predict flowering time, but only with partial accuracy, due to the large variation in the flowering-time distributions.

Taken together, the statistical results for both markers not only show that these two segregating markers partially explained the variation in flowering time, but confirmed that the early allele caused early flowering after many fine-mapping generations. However, further confirmation of the effect of the late homozygous combination of these markers was required by increasing the sample size, which led to the next fine-mapping experiment with a larger sample size from selected family 425(s)(s).

5.3 FINE-MAPPING USING THE BC3F3 GENERATION

To confirm the fine mapping of the QTL region using the BC3F2 generation and to identify more recombinants within the QTL interval (only two were identified), a larger fine-mapping experiment using 1,584 plants was conducted from March to August 2021. Moreover, additional SNP markers were designed by Dr. Niharika Sashidhar to further narrow down the QTL interval. The progeny of nine plants from family 425(s)(s) were selected for this fine-mapping experiment (see the square with the dashed line in Figure 5.3); the genotypes of these nine selected plants can be seen in the centre of Figure 5.7. In total, 165 plants grown from the seeds of three individuals (plant ID numbers 14, 48 and 71) were used to test the effect of LG1; 40 plants grown from the seeds of the two recombinants in 425(s)(s) (plant ID numbers 120 and 386), as well as 40 plants grown from the seeds of plant ID number 11, were used to test whether the effect of the early homozygous segment depended on its location within the QTL interval in LG8; and 1,219 plants grown from the seeds of a further three individuals (plant ID numbers 51, 694 and 780) were used to identify more recombinants. As a control group, 30 individuals from the early-flowering parental accession and 50 from the late-flowering parental accession were included.

This experiment was performed in two adjoining cabins (cabins 1.15 and 1.16) in the greenhouse to ensure comparable light conditions. Each cabin contained 792 plants and the progeny of the nine above-mentioned mother plants were equally distributed between the two cabins. Within each cabin, plants were fully randomised and plants from the early- and late-flowering parental accessions were equally distributed among the three tables in the cabins. The analysis and plotting of flowering time was performed when all the individuals from the early-flowering parental accession flowered. This led to a simple method to distinguish between early- and late-flowering phenotypes in the BC3F3 generation plants; those that flowered before all of the early-flowering parental accession plants in the control group were early flowering, and those that did not flower by this time point were considered to be late-flowering flowering. The flowering-time distributions of the BC3F3 plants according to each mother plant from which they derived are visualised as box plots in Figure 5.7, where panels A and B refer to data for plants grown in cabin 1.15 and 1.16, respectively. The following analysis only focuses on

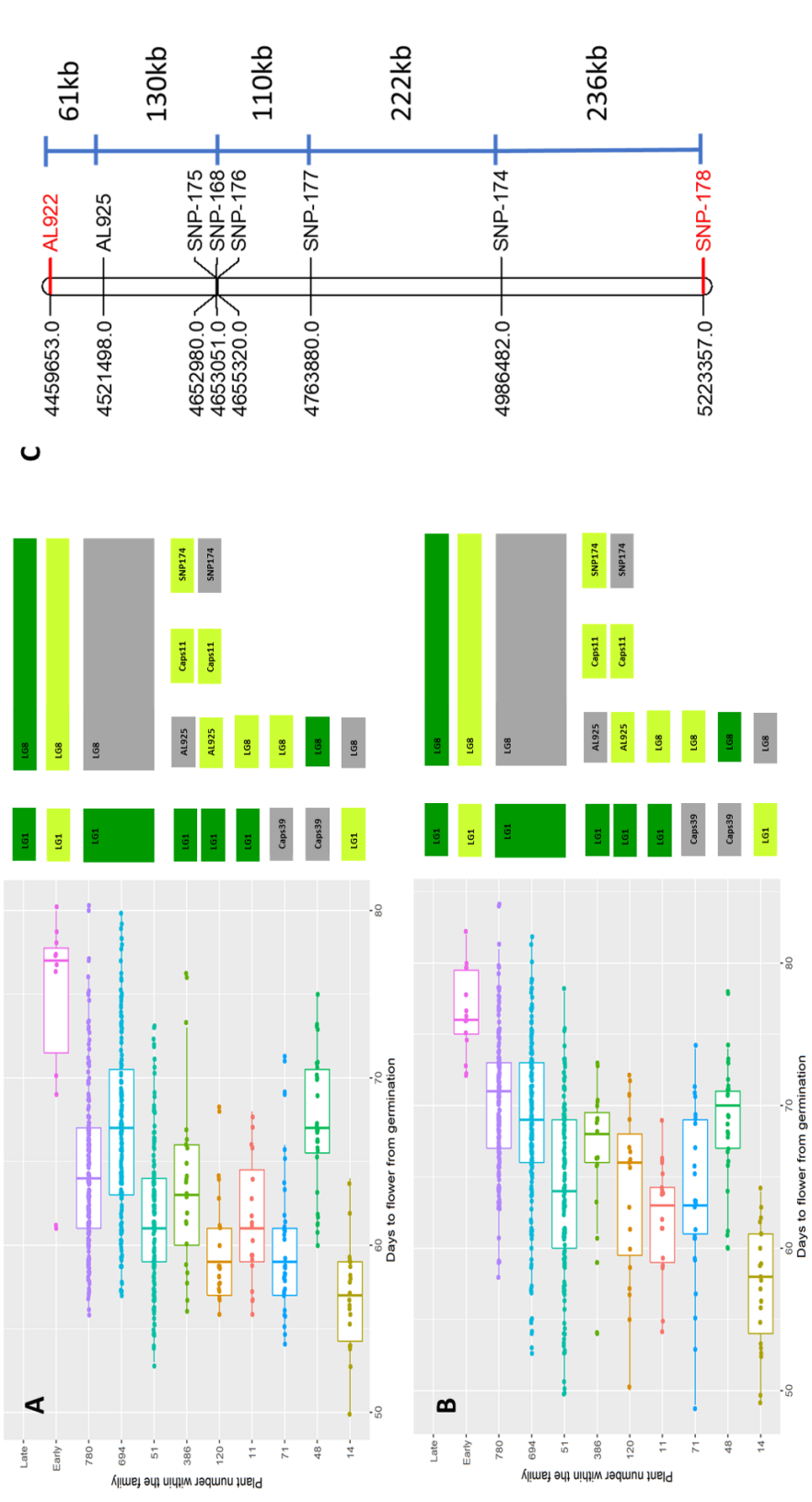

Figure 5.7. The flowering-time distribution for 1,584 BC3F3 *Arabis alpina* plants derived from the 425(s)(s) mother plant, with corresponding genotypes, and a summary of the genetic markers within the QTL interval. (A) The distribution of flowering time for plants in cabin 1.15. (B) The distribution of flowering time for plants in cabin 1.16. The middle panels show the genotype of the mother plants aligned with the corresponding box plots in (A) and (B). For some plants, LG8 has been expanded to show the genotype of multiple markers. (C) A schematic representation of the current QTL interval in LG8, including all genetic markers.

the flowering data from cabin 1.15, because plants in cabin 1.16 experienced a 5-day period without fertilization during the experiment, which might have affected the flowering time.

Similar to data for the previous generation (Figure 5.4), the variation in flowering time shown for BC3F3 in Figure 5.7 led to long tails for each distribution. This variation is consistent with the observed flowering of these accessions without vernalization, a feature described in section 4.2 in Chapter 4. Notably, when LG1 was fixed with the genetic background of the early-flowering accession but LG8 still segregated, plants flowered extremely early (for example, the progeny of plant ID number 14, which showed the lowest median flowering time), which is consistent with results from the previous generation. Additionally, comparison of the flowering-time distributions of the progeny of the two previously identified recombinants (plant ID numbers 120 and 386) showed that fixing the beginning of the QTL region of LG8 with the genetic background of the early-flowering accession (e.g. the progeny of plant ID number 120), resulted in a shorter median flowering time than when the end of the QTL region of LG8 was fixed with the genetic background of the early-flowering accession (e.g. the progeny of plant ID number 386). This suggests that the gene that causes early flowering is located at the beginning of the QTL interval.

As discussed in the previous section, the BC3F2 generation did not include a large number of plants with the late homozygous allelic combination for the InDel and SNP markers AL925 and SNP-168. This resulted in a non-significant statistical difference in flowering-time distributions between the late homozygous and heterozygous combinations. Here, a larger sample size allowed a more robust analysis of this difference. Using the same two-sample Anderson-Darling test, a vanishingly small p-value of $\sim 10^{-7}$ was found when comparing the late homozygous and heterozygous distributions, as well as when comparing the late and early homozygous distributions. Thus, the causal gene(s) that result(s) in the early-flowering phenotype is unequivocally dominant. Intriguingly, although these two markers were segregating, the number of plants that carried the late homozygous combination of these two genetic markers was much lower than expected according to Mendelian segregation when considering the number of plants with the other allelic combinations. The late homozygous combination of these two markers probably resulted in many non-viable seeds that did not germinate, although a similar number of seeds from each mother plant were stratified in this experiment.

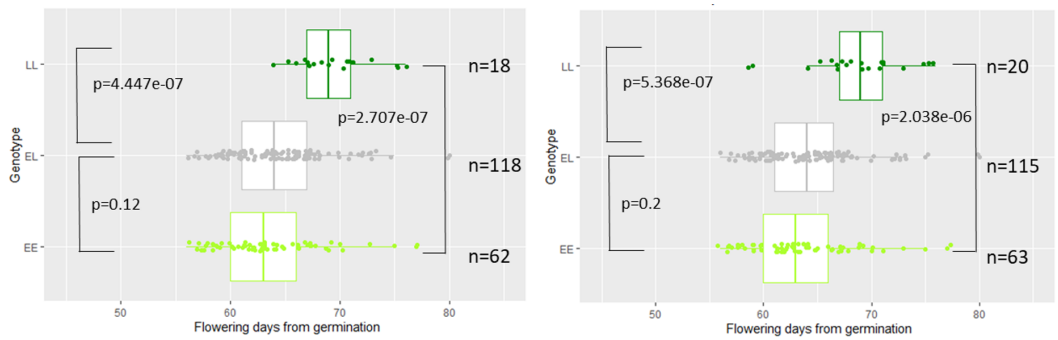

Figure 5.8. Flowering time distributions separated by allelic combinations for two selected markers in the QTL interval (Left) Distributions separated for the indel marker AL925. (Right) Distributions separated for the SNP marker SNP168. The sample size for each distribution is shown as well as the *p*-value resulting from the Anderson-Darling test considering each pair-wise comparison.

Collectively, this large-scale, fine-mapping experiment has better defined the QTL interval and narrowed it down to 759 kb (Figure 5.7C). This interval contains about 120 genes, none of which is a known regulator of flowering time.

5.4 Differentially expressed candidate genes

The combination of data from the RNA-seq and QTL mapping experiments allows the expression levels of candidate genes within the QTL to be quantified during the period in which bud growth is arrested. Quantification of the expression levels of all 120 genes located within the current QTL region identified several genes that were differentially expressed between the early- and late-flowering accessions. I defined differentially expressed genes as those that had a 1.5-fold difference in expression level between the two accessions without vernalization at more than one time point. A 1.5-fold change in log_2 fragments per kilobase of transcript per million mapped reads (FPKM) will result in a difference of 0.5. Out of 120 genes in the current QTL interval, only eight genes were differentially expressed and their individual expression levels over time in early- and late-flowering accessions with or without vernalization are shown in Figure 5.9 to Figure 5.14.

The first differentially expressed gene was the orthologue of the *A. thaliana* gene *ERECTA-LIKE2 (ERL2)*; the change in its expression level over time is shown in Figure 5.9. In the absence of vernalization, *ERL2* was differentially expressed from about 7 weeks onwards, when the bud-arrest phenotype became

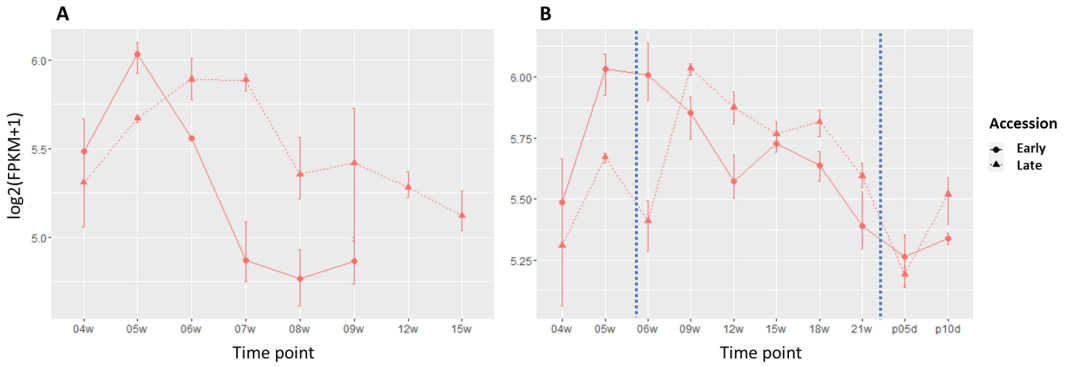

Figure 5.9. The expression level of *ERL2* normalised as log2 fragments per kilobase of transcript per million mapped reads (FPKM) in *Arabis alpina* plants of the early- and late-flowering accession that were either not vernalized (A) or were vernalized for 16 weeks (B). The two blue dashed lines indicate the beginning and end of the veranlization period. Expression data are the mean of three biological replicates, and error bars represent the SD.

apparent. This was due to a large decrease in *ERL2* expression in the early-flowering accession between 5 and 7 weeks, which was absent in the late-flowering accession. This decrease in *ERL2* expression in early-flowering accession plants coincided with the timing of primary shoot elongation and the onset of flower production. When plants were vernalized, the expression level of *ERL2* in both accessions was similar during and after vernalization, which is consistent with the absence of a statistically significant difference in flowering time observed after vernalization described in Figure 4.2 in Chapter 4.

The second differentially expressed gene was the orthologue of the *A. thaliana* gene *BCL-2-ASSOCIATED ATHANOGENE 3* (*BAG3*). The BAG protein family has been proposed to recruit molecular chaperones to target proteins and to modulate their functions by altering protein conformation under stress conditions (Doukhanina et al., 2006). In contrast to *ERL2*, *BAG3* was already differentially expressed at the 5-week time point in the absence of vernalization, and its expression increased between 5 and 8 weeks only in the early-flowering accession (Figure 5.10), which coincided with primary shoot elongation and the start of flower production. When plants were vernalized, the expression level of *BAG3* in both accessions was similar from 4 weeks of vernalization onwards and after vernalization.

The third differentially expressed gene was the orthologue of the *A. thaliana*

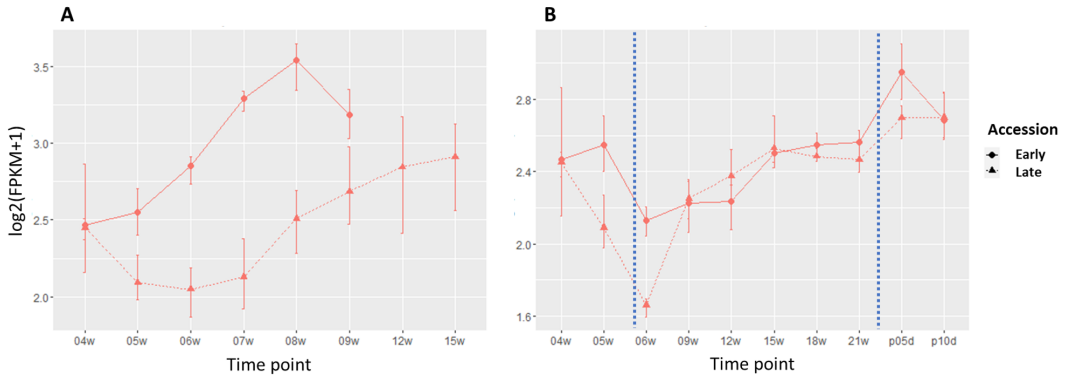

Figure 5.10. The expression level of *BAG3* normalised as log2 fragments per kilobase of transcript per million mapped reads (FPKM) in *Arabis alpina* plants of the early- and late-flowering accession that were either not vernalized (A) or were vernalized for 16 weeks (B). The two blue dashed lines indicate the beginning and end of the veranlization period. Expression data are the mean of three biological replicates, and error bars represent the SD.

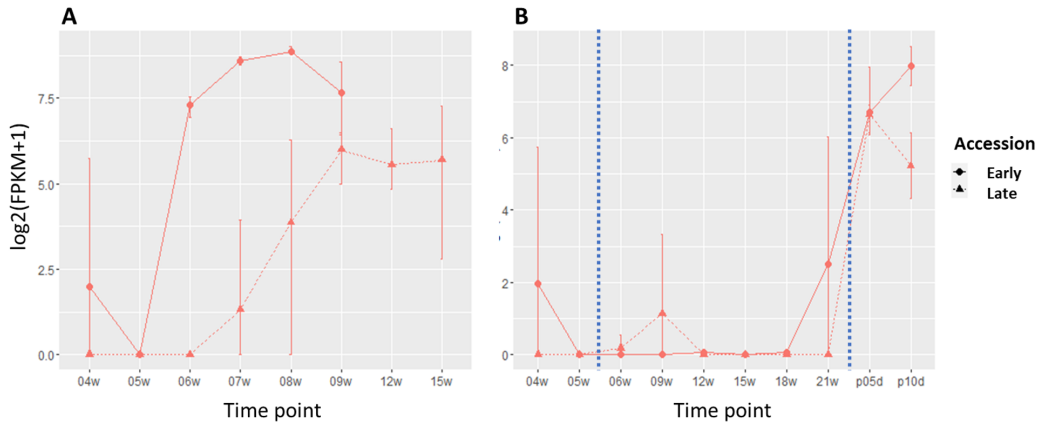

Figure 5.11. The expression level of *AT5G07230* normalised as log2 fragments per kilobase of transcript per million mapped reads (FPKM) in *Arabis alpina* plants of the early- and late-flowering accession that were either not vernalized (A) or were vernalized for 16 weeks (B). The two blue dashed lines indicate the beginning and end of the veranlization period. Expression data are the mean of three biological replicates, and error bars represent the SD.

gene *AT5G07230*, which is a lipid-transfer protein (Huang et al., 2013). In *A. thaliana*, *AT5G07230* is specifically expressed in anthers and expression is highest in young floral buds about stage 7 and thereafter decreases (Wellmer et al., 2004). The expression level of *AT5G07230* in the early- and late-flowering accessions is shown in Figure 5.11. Without vernalization, the initial large increase in the expression of *AT5G07230* in the early-flowering accession correlated with the early-flowering phenotype described in section 4.2 and Figure 4.2. By contrast, the delayed and gradual increase in *AT5G07230* expression in the late-flowering accession was associated with the slow development of the floral buds. Notably, *AT5G07230* was hardly expressed in both accessions during vernalization, but was highly expressed in both accessions after vernalization.

The fourth differentially expressed gene was the orthologue of the *A. thaliana* gene *ISOPROPYLMALATE DEHYDROGENASE 2* (*IMD2*). *IMD2* is one of three genes that encode the enzyme 3-isopropylmalate dehydrogenase, which is involved in leucine biosynthesis in *A. thaliana* (Less and Galili, 2008). The expression pattern of *IMD2* in the early- and late-flowering accessions is shown in Figure 5.12. Without vernalization, *IMD2* was differentially expressed at the 6-week timepoint, because its expression increased sharply between 5 and 8 weeks only in the early-flowering accession, a time point that coincided with flowering on the primary shoot. For the late-flowering accession, expression of *IMD2* only increased gradually from 6 weeks. Notably, the expression level of *IMD2* increased nearly 6-fold in the late-flowering accession after 7 weeks of vernalization and reached a higher expression level than that in the early-flowering accession. Following vernalization, the expression level of *IMD2* in both accessions was similar.

Notably, several genes encoding glycine-rich proteins (GRPs), such as GRP16, GFP17, GRP19 and GRP20, were all differentially expressed between the early- and late-flowering accessions and were located within the QTL region. In *A. thaliana*, these genes are specifically expressed in the tapetum and pollen coat (Wellmer et al., 2004). Moreover, previous study has shown that GRPs, which are also categorised as oleopollenins, play an important role in species with different breeding systems (Schein et al., 2004). The expression levels of *GRP16*, *GFP17*, *GRP19* and *GRP20* are shown in Figures 5.13 and 5.14. A rapid increase in the expression of *GRP16* and *GRP17* occurred in both accessions following floral transition; however, this increase occurred after 6 weeks in the early-flowering accession, whereas it occurred one week later in the late-flowering accession, suggesting that

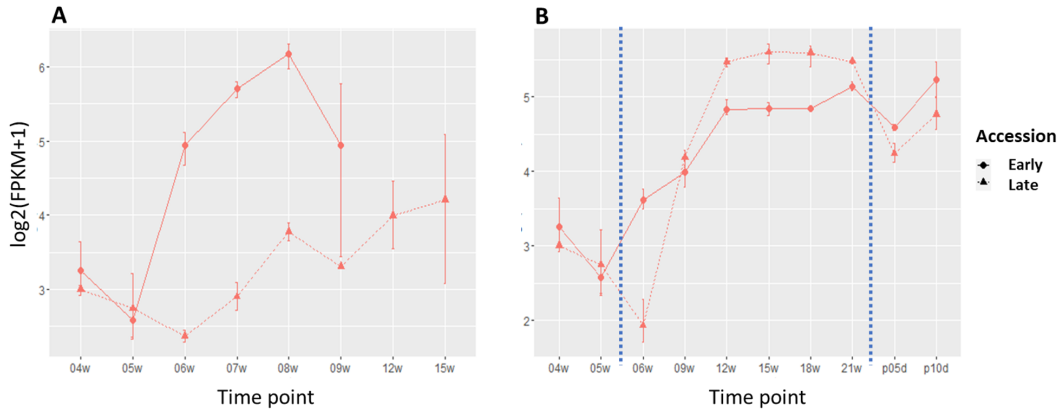

Figure 5.12. The expression level of *IMD2* normalised as log2 fragments per kilobase of transcript per million mapped reads (FPKM) in *Arabis alpina* plants of the early- and late-flowering accession that were either not vernalized (A) or were vernalized for 16 weeks (B). The two blue dashed lines indicate the beginning and end of the veranlization period. Expression data are the mean of three biological replicates, and error bars represent the SD.

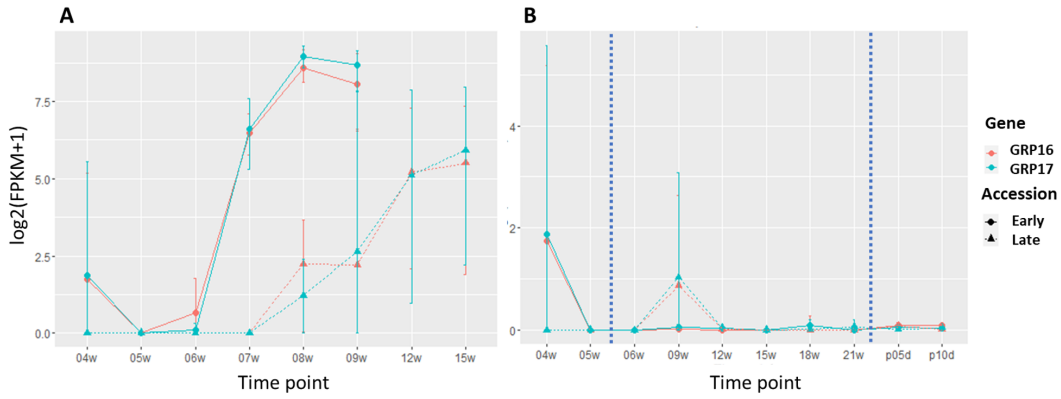

Figure 5.13. The expression levels of *GRP16* and *GRP17* normalised as log2 fragments per kilobase of transcript per million mapped reads (FPKM) in *Arabis alpina* plants of the early- and late-flowering accession that were either not vernalized (A) or were vernalized for 16 weeks (B). The two blue dashed lines indicate the beginning and end of the veranlization period. Expression data are the mean of three biological replicates, and error bars represent the SD.

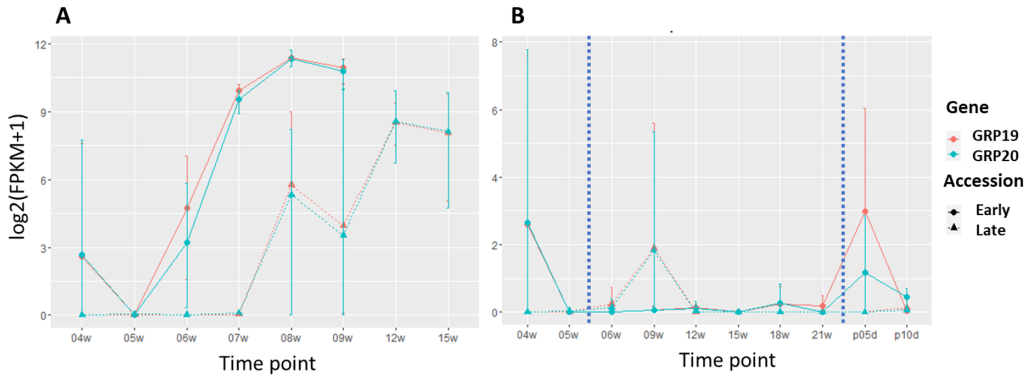

Figure 5.14. The expression levels of *GRP19* and *GRP20* normalised as log2 fragments per kilobase of transcript per million mapped reads (FPKM) in *Arabis alpina* plants of the early- and late-flowering accession that were either not vernalized (A) or were vernalized for 16 weeks (B). The two blue dashed lines indicate the beginning and end of the veranlization period. Expression data are the mean of three biological replicates, and error bars represent the SD.

both genes might be involved in floral development. Vernalization repressed the expression of both genes and a low level of expression was maintained after vernalization, during flowering. An increase in the expression of *GRP19* and *GRP20* also occurred in both accessions following floral transition, but the expression level of *GRP19* and *GRP20* was higher than that of *GRP16* and *GRP17*. The expression of *GRP19* and *GRP20* was also repressed in response to vernalization; however, a slight increase in expression after vernalization was observed in the early-flowering accession. The transcriptional dynamics of *GRP* genes in plants exposed to vernalization has not been previously reported.

5.5 *ERL2* IS COEXPRESSED WITH THE FLORAL REPRESSOR *SVP* AND *TFL1* IN THE EARLY-FLOWERING ACCESSION

To further investigate how these differentially expressed genes within the QTL interval are involved in gene regulatory networks, I used the genome-wide co-expression data described in section 4.7 in Chapter 4. This allowed the expression levels of the gene clusters that contained the deferentially expressed genes in the QTL interval to be examined. Intriguingly, in the early-flowering accession, *ERL2*

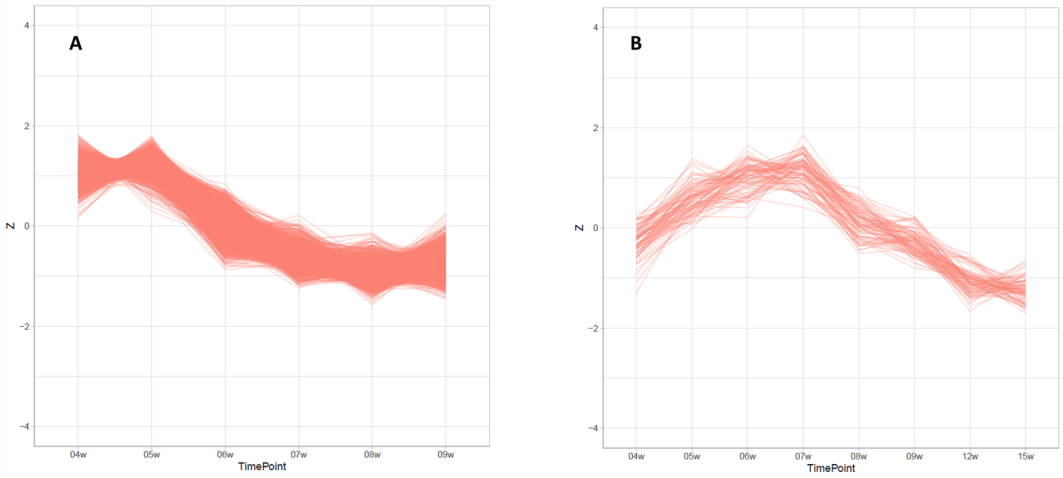

Figure 5.15. Temporal expression dynamics for all the genes in the gene cluster containing *ERL2* for both the early- (A) and late-flowering (B) accessions of Norweigian *A. alpina* without vernalization. The expression levels are plotted as Z-scores.

was identified to be in the same gene cluster as the floral repressor genes *SVP* and *TFL1*, suggesting that *ERL2* is not only coexpressed with, but also participates in, the same biological process as these two floral repressors. This gene cluster is the same cluster that I characterised in section 4.7 in Chapter 4. The expression level of this cluster over time is reproduced here in Figure 5.15A. The continual decrease in the level of gene expression from 6 to 9 weeks correlated with inflorescence development after floral transition in the early-flowering accession without vernalization.

To further decipher the biological processes in which this gene cluster is involved, Gene Ontology (GO) enrichment analysis was performed for all the genes in this cluster, using the R package topGO. This package allowed the testing of GO terms while accounting for the topology of the GO graph. Fisher's exact test was used to eliminate local similarities and dependencies between GO terms. The p-values reported by Fisher's exact test were further adjusted using the method of Benjamini and Hochberg (1995) to control for the false discovery rate. Subsequently, the significant GO terms were selected if the adjusted p-value was smaller than 0.05. These significant GO terms were further summarised as treemaps generated using REVIGO, which removes redundant GO terms (Supek et al., 2011). Figure 5.16 illustrates all of the GO-term clusters resulting from the gene cluster

containing *ERL2* in the early-flowering accession without vernalization. Superclusters were formed by grouping GO-term clusters with related terms, which resulted in six superclusters: non-coding RNA processing, tissue development, regulation of cell cycle, ribosome biogenesis, cell-cycle phase transition and glycosyl compound metabolic process. Intriguingly, the largest supercluster was associated with non-coding RNA processing, suggesting that flowering in the early-flowering accession without vernalization might involve microRNAs (miRNAs) and long non-coding RNAs (lncRNAs). Moreover, the second-largest GO-term supercluster included many GO terms related to tissue development, such as meristem and shoot system development, which is consistent with the formation of an elongated inflorescence in the early-flowering accession without vernalization.

To define the role of the gene cluster containing *ERL2* in the late-flowering accession without vernalization and to compare this role to that in the early-flowering accession, I identified this cluster in the coexpression network analysis; the expression pattern of the *ERL2* cluster in the late-flowering accession without vernalization is shown in Figure 5.15B. As described in the co-expression network analysis in section 4.7 of Chapter 4, by contrast with two large gene clusters identified in the early-flowering accession, scattered and smaller clusters containing floral repressors were present in the late-flowering accession without vernalization. Notably, the gene cluster containing *ERL2* in the late-flowering accession without vernalization did not contain any known floral repressors and was therefore not discussed in section 4.7. Furthermore, the expression of genes in this cluster peaked at 7 weeks, when the bud-arrest phenotype became apparent, implying a potential role for the genes in this cluster in the slow growth of the floral buds.

In addition to performing GO enrichment analysis for the early-flowering accession, I applied the same analysis to the gene cluster in the late-flowering accession that contained *ERL2*. Figure 5.17 illustrates all the GO-term clusters and the representative superclusters of GO terms. The resulting treemap was sparser than the early-flowering accession counterpart, due to the small number of genes in the cluster. Only two superclusters contained more than two GO-term clusters, and these were associated with shoot system development and cellular response to endogenous stimulus. The association of genes in this gene cluster with shoot system development is consistent with the observed absence of primary shoot elongation linked to the bud-arrest phenotype, whereas the association of the genes with endogenous stimulus implies a potential role for phytohormones in

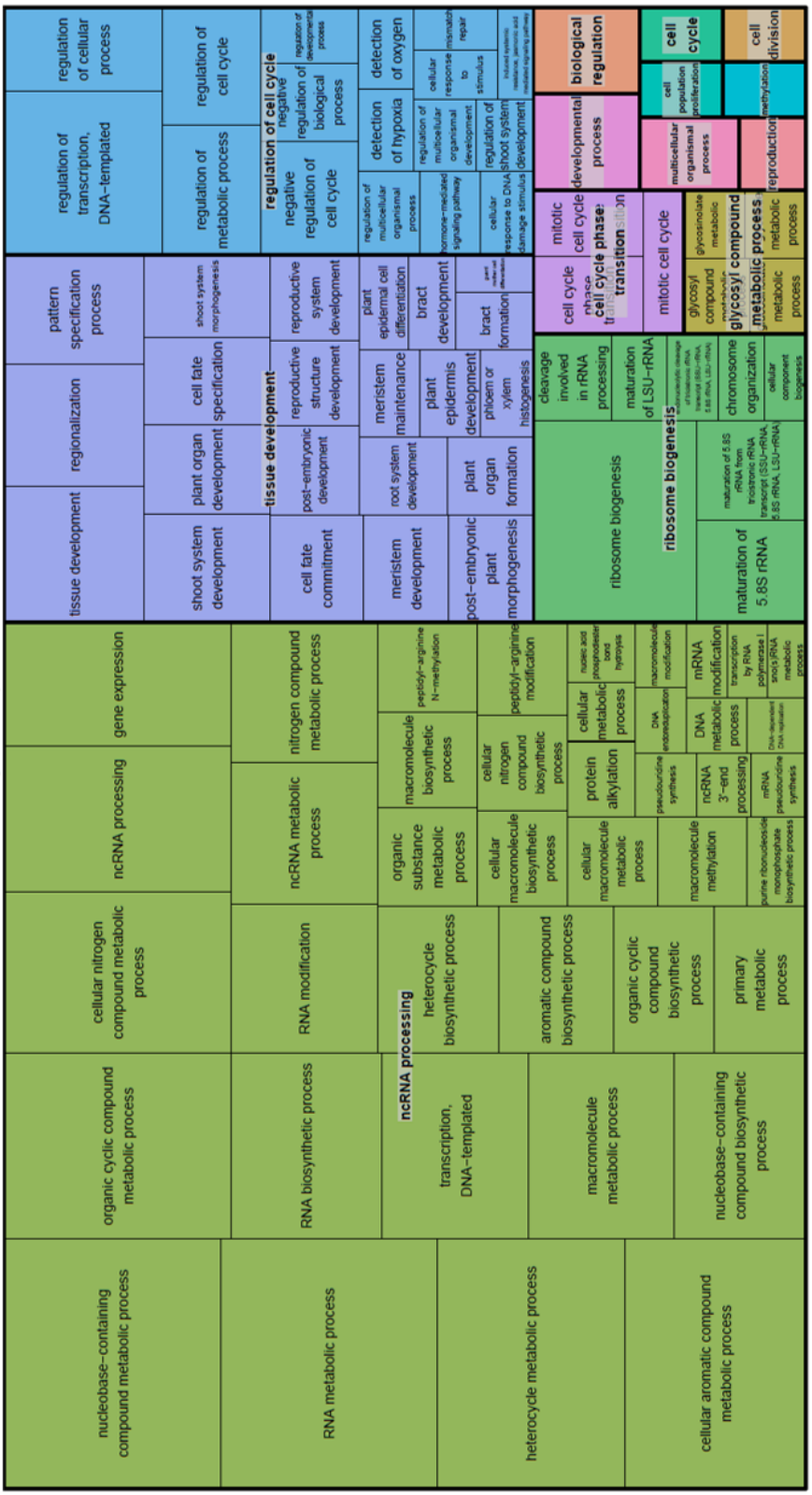

Figure 5.16. A TreeMap view of GO-term clusters for the gene cluster containing the *ERL2* gene in the Norweigian *A. alpina* early-flowering accession without vernalization. The treemap was generated using the default settings of REVIGO. Each rectangle represents a single GO-term cluster. The size of a rectangle is proportional to the frequency of the GO term in the associated GO annotation database. The clusters were grouped into superclusters (highlighted in grey) of loosely related terms, visualized with different colors.

Figure 5.17. A TreeMap view of GO-term clusters for the gene cluster containing the *ERL2* gene in the Norweigian *A. alpina* late-flowering accession without vernalization. The treemap was generated using the default settings of REVIGO. Each rectangle represents a single GO-term cluster. The size of a rectangle is proportional to the frequency of the GO term in the associated GO annotation database. The clusters were grouped into superclusters (highlighted in grey) of loosely related terms, visualized with different colors.

the slow growth of the floral bud.

Taken together, the temporal expression level and the GO enrichment analysis of the gene cluster containing *ERL2* might explain the distinct phenotypes of the early- and late-flowering accessions without vernalization.

5.6 Discussion

This study contributes to understanding the genetic network that underlies variation in floral bud outgrowth without vernalization in Norwegian *A. alpina*. Specifically, I have conclusively shown that the casual gene for this phenotype is dominant. Although the flowering time distributions of the early and late homozygous near-isogenic lines were statistically different, they were smaller than that between the parental phenotypes. This apparent discrepancy may be due to the interaction of other loci. Firstly, when more regions were fixed with the genetic background of the late-flowering accession, the smaller was the difference in flowering time between early and late homozygous backgrounds. Secondly, when LG1 was fixed completely with the early homozygous background, these plants flowered the earliest over two generations. Furthermore, synteny analysis of the QTL region led to the discovery of structural variants (SVs) such as variation in gene copy number at the beginning of the QTL interval. The result of this synteny analysis is shown in Figure 5.18 and in addition to possessing one additional copy of the *bZIP* gene, the late-flowering accession also contained two additional copies of *GIBBERELLIN 20-OXIDASE 3* (*GA20OX3*), both of which copies are highlighted in red in Figure 5.18. This supports the suggestion that SVs can result in the perturbations of cis-regulatory regions of other genes, and that duplicated genes are often involved in the reproductive process (Zhang et al., 2015; Alonge et al., 2020; Yuan et al., 2021). Although the expression levels of all these gene copies were relatively low on the basis of RNA-seq data, I speculate that the location of these copies within the QTL interval plays a role in the flowering-time variation between the two accessions.

Additionally, I have identified eight genes in the current QTL region that are differentially expressed between early- and late-flowering accessions on the basis of RNA-seq data. This leads me to speculate that if the genes in the QTL region that correlate with the difference in the flowering-time phenotype between the two accessions without vernalization are differentially expressed, they are probably

Figure 5.18. Synteny analysis of the QTL interval located on chromosome 8 between the Norweigian *A. alpina* early and late accession. Syntenic regions are colored in yellow and additional copies of *bzip* and *GIBBERELLIN 20-OXIDASE 3* in red.

involved in the bud-arrest phenotype. These eight genes are involved in different functions and none is categorised as a known flowering-time gene. Amongst these eight genes, *ERL2* regulates cell elongation in *A. thaliana* and melon (Qu et al., 2017; Yang et al., 2020), linking it to the main phenotypic difference between the two Norwegian *A. alpina* accessions in which only the early-flowering accession has an elongated primary shoot. Notably, according to the synteny analysis shown in Figure 5.18, the *ERL2* locus is located directly after the *bZIP* gene copies. An intriguing hypopthesis is that variation in the copy number of *bZIP* genes results in different expression levels of *ERL2* in the two accessions without vernalization.

GO enrichment of genes in the *ERL2* cluster has demonstrated their involvement in two major categories. Firstly, it has suggested that they are mainly associated with non-coding RNA processing and tissue development in the early-flowering accession. This supports the notion that reproductive development is regulated by non-coding RNA (Yamaguchi and Abe, 2012). Although the regulation of microRNAs in perennial flowering has been reported in Spanish accession Pajares of *A. alpina*, the relationship between microRNAs and flowering behaviour in Norwegian *A. alpina* accessions remains unclear. Additionally, no nucleotide polymorphisms were identified in the coding sequence or 5'and 3'untranslated regions of *ERL2* in the early- or late-flowering accession. The flowering behaviour of Norwegian *A. alpina* might be epigenetically regulated, because *FLC* and its epigenetic regulation is an important component of flowering in the Spanish accession Pajares of *A. alpina* as well as species of *Arabidopsis lyrata* and *Brassica rapa* L. (Castaings et al., 2014; Shea et al., 2019).

Secondly, GO enrichment analysis of the *ERL2* cluster provides evidence that in the late-flowering accession, these genes are mainly associated with shoot system development and cellular response to endogenous stimulus. The bud-arrest phenotype and an arrested primary shoot are similar to the phenotypes of GA-deficient mutants of *A. thaliana* (Thomas and Sun, 2004). Furthermore, the ERECTA proteins are involved in gibberellin signalling (Sarnowska et al., 2021). Therefore, to analyse the role of GA in the bud-arrest phenotype, I applied GA_{4+7} (20 mM) weekly to apices of plants of the late-flowering accession in the greenhouse. This could overcome the bud-arrest phenotype and the primary shoot of these plants elongated and produced flowers and siliques (Figure 5.19). This suggests that GA is involved in the bud-arrest phenotype, which is consistent with the identified GO terms of cellular response to endogenous stimulus for the *ERL2*

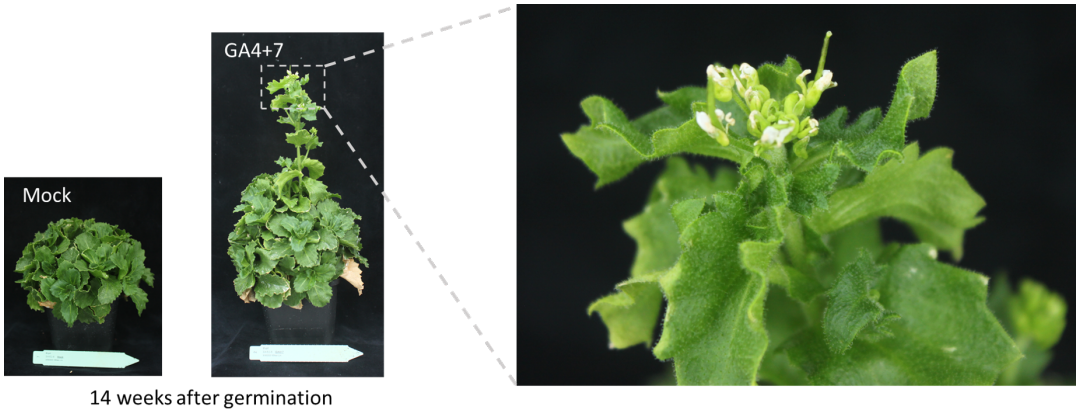

Figure 5.19. Application of gibberellin acid to plants from the Norweigian *A. alpina* late-flowering accession. 20 mM gibberellin acid (GA_{4+7}) was applied to the primary shoot of plants in the late-flowering accession weekly after germination. (Left) Mock and GA_{4+7} treated plants were photographed 14 weeks after germination. These two plant had the same germination date. (Right) A close-up of flowers from the GA-treated plant.

cluster in the late-flowering accession.

This is the first report for QTL mapping in Norwegian *A. alpina* accessions. The combination of QTL mapping, RNA-sequencing and GO enrichment analysis has allowed me to identify the potential casual gene responsible for the variation in flowering time between the early- and late-flowering accessions without vernalization. *ERL2* is the first potential candidate gene that should be analysed further and phenotypic analysis of *erl2* mutant plants will contribute to understanding how flowering is regulated in Norwegian *A. alpina* accessions.

Chapter 6
IN SITU transplantation experiment at the Norwegian Riasten field site

6.1 Introduction to transplantation experiment

To assess the adaptive significance of variation in bud-outgrowth, the early- and late-flowering accessions, together with with six other families, were transplanted to the Norwegian field site by Dr. Ulla Kemi and Dr. Jörg Wunder in the summer of 2018. These eight families were selected because their phenotypes and genotypes represent early- and late-flowering behaviour in the greenhouse. The four early-flowering families were: Nor6-20, Nor6-43, No6-51 and S4-41-2; and the four late-flowering families were: Nor6-06, Nor6-12, Nor6-54 and S4-62-8. The two families S4-41-2 and S4-62-8 are the early- and late-flowering accessions described in all previous chapters, respectively. Each transplantation plot was composed of 20 individuals from each family, resulting in a total sample size of 160 individuals per plot. To minimise the effect of confounding factors, the locations of the 160 plants within the plot were randomised and six transplantation plots were established at different locations in the valley. For each transplantation site, climate (Figure C.2) and light spectrum data were recorded to provide details concerning the growth conditions of the natural environment. The scoring of the transplantation plots

was performed during several visits between June and September annually from 2019 to 2021. The scored traits included rosette surface area, rosette number, the developmental stage of each individual rosette, whether the plant was flowering, the length of the flowering shoots, the number of siliques and the plant height (Kemi et al., in preparation).

In the summer of 2019, only three of the plots (plot 1, 2 and 4) provided sufficient data due to high plant survival rates. Plot 1 was situated in the early snowmelt region (23^{th} week of 2019) whereas plots 2 and 4 were located in the late snowmelt region (30^{th} week of 2019). In the summer of 2020, the time of snowmelt started in week 30 and two out of these three plots (plot 2 and 4) remained covered by snow during our first visit. Only plot 1 was not covered by snow but did not contain any flowering plants yet. Unfortunately, a second visit in September 2020 was not possible due to the coronavirus pandemic. The two visits in the summer of 2021 enabled the previous year's flowering situation at the three remaining plots to be inferred on the basis of the presence of dry shoots. However, it was discovered during the first visit in 2021 that many of the plants that were alive in 2020 had died. This included the whole of plot 4 and the vast majority of plot 2; only the plants in plot 1 survived in sufficient quantities to provide adequate flowering data.

The adaptive significance of the flowering variation can be assessed by comparing the collected field data with the greenhouse experiment data. For this, I constructed generalised linear models (GLMs) to address the differences in familial survival rates, flowering episodes, as well as flowering-related traits from the field data. To account for the effect of genotype derived from the results of genetic markers in current QTL interval on different families, I firstly performed a design using family nested within the genotype as a single predictor. However, due to the collinearity between genotype and family, this caused the inflation of parameter standard errors, generating inaccurate estimation of parameters for the model. Therefore, the application of nested design was not feasible in my current data-set. Instead, the GLMs were done using single predictor models, i.e. the family as a categorical input; including additional features as predictors (e.g. genotype, plot and interaction terms) resulted in no discernable improvement in the models[1].

[1]Model performance was quantified using the Akaike information criteria (Akaike, 1978)

6.1.1 Survival rate

To investigate whether survival rate differed amongst the different families, I fitted a logistic model with a binomial link function to the field data acquired during the second trip of 2021. I considered a plant to have survived if it was still alive in September 2021 from the start of transplantation experiment in 2018. In addition, to understand whether differences in the time of snowmelt affected survival rate, I applied the logistic model to plot 1 and 2 separately. The equation governing the model was the following:

$$\ln\left[\frac{P(\text{Survival}=1)}{1-P(\text{Survival}=1)}\right] = \alpha + \underline{\beta}\cdot\underline{\mathbf{F}}, \tag{6.1}$$

where α is the intercept representing the value for family Nor6-06, $\underline{\beta} = (\beta_{\text{Nor6}-12},\ldots,\beta_{\text{S4}-62-8})$, i.e. a vector containing the model parameters, and $\underline{\mathbf{F}} = (\text{Family}_{\text{Nor6}-12},\ldots,\text{Family}_{\text{S4}-62-8})$, i.e. a vector containing the remaining seven families. One can see from the form of equation 6.1 that the parameter β_i describes differences between family i and the intercept family Nor6-06.

The result of fitting a logistic model to predict the probability of survival with different families in both plot 1 and 2 showed that difference in survival rate amongst these eight families is not statistically clear due to the large error. The results of applying this model to the data from plot 1 are displayed in Figure 6.1 and Table 6.1. No strong difference was observed among the survival rates of the eight families in plot 1, and each family had a survival rate between $\sim 70-90\%$. The same result was observed for the plants from plot 2 (Figure 6.2 and Table 6.2), although the survival rates for these plant were considerably lower, between $\sim 15-50\%$. Thus, although a clear difference in plant survival existed between the two plots, probably due to their different periods of snow cover, no adaptive significance in terms of survival rates between families existed.

6.1.2 Rosette number

The number of axillary branches ensures the perennial flowering behaviour in *A. alpina*, because reproductive success depends on flowering from axillary meristems in the subsequent year (Wang et al., 2009a). To quantify this in the field, the total number of rosettes in plot 1 was scored on the second trip of 2021, because the number of rosettes represents the number of branches in the natural environment.

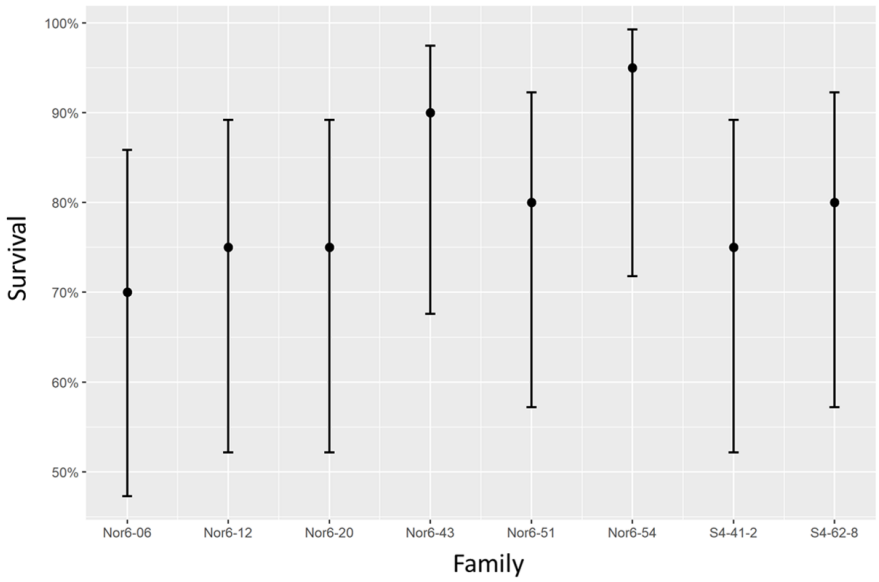

Figure 6.1. The predicted probability of survival for plants in plot 1 according to family. The error bars show the 95% confidence interval of the survival probability.

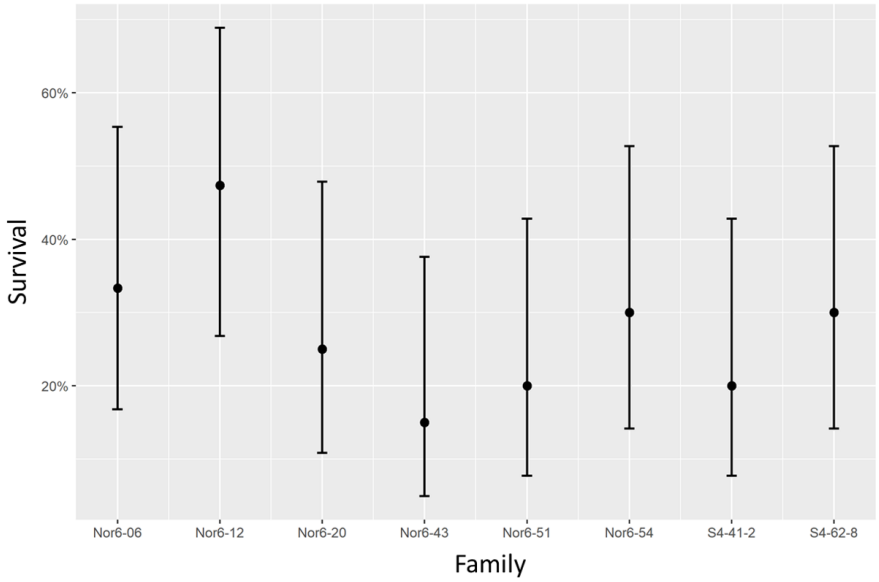

Figure 6.2. The predicted probability of survival for plants in plot 2 according to family. The error bars show the 95% confidence interval of the survival probability.

Family	β_i	σ_i	z-value	Pr(>|z|)
Intercept (Nor6-06)	0.85	0.49	1.74	0.08
Nor6-12	0.25	0.71	0.35	0.72
Nor6-20	0.25	0.71	0.35	0.72
Nor6-43	1.35	0.89	1.52	0.13
Nor6-51	0.54	0.74	0.73	0.47
Nor6-54	2.10	1.14	1.85	0.06
S4-41-2	0.25	0.71	0.35	0.72
S4-62-8	0.54	0.74	0.73	0.47

Table 6.1. Model parameters for survival probability in plot 1. The table shows the best fitting results of the logistics model given by equation 6.1 applied to the data from plot 1. Column 1 shows the family; note that the intercept, α, corresponds to family Nor6-06. Column 2 shows the best fitting result for the parameter β_i. Column 3 shows the estimated error for this parameter, σ_i. Column 4 shows the z-value, i.e. β_i/σ_i. Column 5 shows the p-value resulting from this z-value considering the null hypothesis that $\beta_i = 0$.

Family	β_i	σ_i	z-value	Pr(>|z|)
Intercept (Nor6-06)	-0.69	0.46	-1.50	0.13
Nor6-12	0.59	0.65	0.90	0.37
Nor6-20	-0.41	0.69	-0.58	0.56
Nor6-43	-1.04	0.78	-1.34	0.18
Nor6-51	-0.69	0.73	-0.96	0.34
Nor6-54	-0.15	0.67	-0.23	0.82
S4-41-2	-0.69	0.73	-0.96	0.34
S4-62-8	-0.15	0.67	-0.23	0.82

Table 6.2. Model parameters for survival probability in plot 2. The columns are the same as those in table 6.1.

I first fitted the data with a Poisson model to predict the total number of rosettes using only the family as a predictor; however, the model suffered from over-dispersion, a common failure of a Poisson model when the underlying distribution is not a perfect Poisson distribution. To account for over-dispersion, I fitted the data with a negative-binomial model, allowing for the distribution variance to be independent of the mean. The equation governing the model is the following:

$$\ln(E(\mathrm{N_{ros}})) = \alpha + \underline{\beta} \cdot \underline{\mathbf{F}}, \tag{6.2}$$

where α is the intercept representing the value for family Nor6-06, $\underline{\beta} = (\beta_{\mathrm{Nor6-12}},\ldots,\beta_{\mathrm{S4-62-8}})$, i.e. a vector containing the model parameters, and $\underline{\mathbf{F}} = (\mathrm{Family}_{\mathrm{Nor6-12}},\ldots,\mathrm{Family}_{\mathrm{S4-62-8}})$,

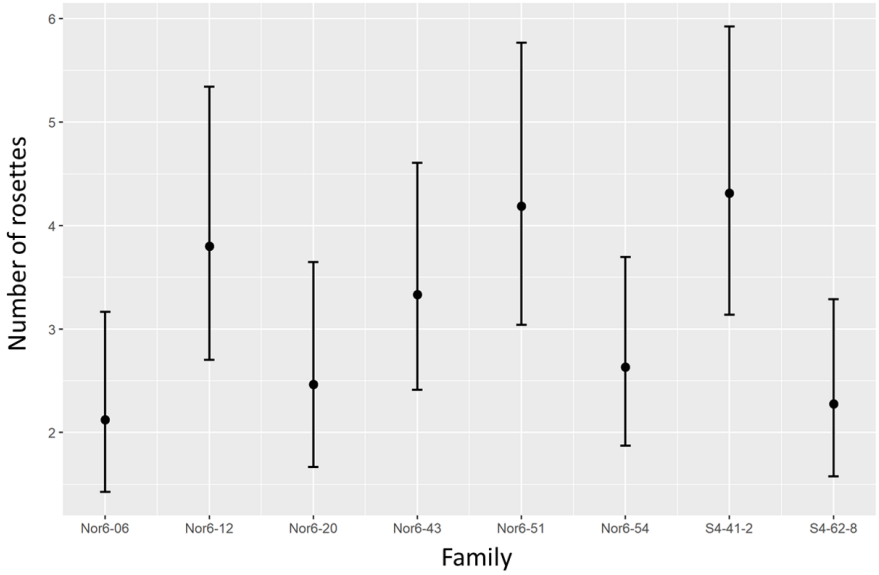

Figure 6.3. The expected number of rosettes in plot 1 according to family using equation 6.2. The error bars show the 95% confidence interval of the expected number.

i.e. a vector containing the remaining seven families.Similar to equation 6.1, this results in parameter β_i, which describes the difference between family i and the intercept family Nor6-06. On the left-hand side of the equation, $E(\mathrm{N_{ros}})$ is the expected number of rosettes given a negative-binomal distribution.

The results of the negative binomial model are shown in Figure 6.3 and Table 6.3. Differences existed between the expected number of rosettes of each families. This can be most easily seen in Table 6.3, where families Nor6-12, Nor6-51 and S4-41-2 showed values of β_i significantly different from 0 (p-values < 0.05), i.e. they had an expected number of rosettes different from that of family Nor6-06. These families had an expected number of rosettes of $\sim 4-5$. The other families are indistinguishable from family Nor6-06 and had an expected number of rosettes of $\sim 2-3$. In addition, the intercept was statistical significant, indicating that the expected number of rosettes of family Nor6-06 was greater than 1.

To examine the connection between genotype in the QTL region and the number of rosettes , I firstly applied non-parametric Kruskal-Wallis test to the 4 families representing either early-flowering (Nor6-20, Nor6-43, No6-51 and S4-41-2) or late-flowering (Nor6-06, Nor6-12, Nor6-54 and S4-62-8) based on the genotyping result by marker SNP-168 in QTL region (Figure 5.4 and 5.7 in Chapter 5). Test

Family	β_i	σ_i	z-value	Pr(>|z|)
Intercept(Nor6-06)	0.75	0.20	3.74	<0.001
Nor6-12	0.58	0.27	2.19	0.03
Nor6-20	0.15	0.28	0.53	0.60
Nor6-43	0.45	0.26	1.73	0.08
Nor6-51	0.68	0.26	2.62	0.01
Nor6-54	0.21	0.27	0.81	0.42
S4-41-2	0.71	0.26	2.74	0.01
S4-62-8	0.07	0.27	0.25	0.80

Table 6.3. Model parameters for total number of rosette in plot 1. The table shows the best fitting results of the negative binomial model given in equation 6.2 when applied to the data from plot 1. Column 1 shows the family; note that the intercept, α, corresponds to family Nor6-06. Column 2 shows the best fitting result for the parameter β_i. Column shows is the estimated error on this parameter, σ_i. Column 4 shows the z-value, i.e. β_i/σ_i. Column 5 shows the p-value resulting from this z-value considering the null hypothesis that $\beta_i = 0$.

statistics from either the early- or late-flowering group reported a p-value greater than 0.1 and could not reject the null-hypothesis, indicating that 4 distributions within each group had been drawn from the same population. Therefore, I grouped the data from all 8 families into two distributions, early- or late-flowering genotype. The resulting two distributions are shown in Figure 6.4. The early-flowering genotype distribution (n=65) has a median of 3 and an interquartile range of 3; the late-flowering genotype distribution (n=68) has median of 2 and interquartile range of 3. To determine whether the difference between these distributions was statistically significant and the early-flowering genotype plants had a greater number of rosettes, I performed the Mann-Whitney U test. This resulted in a p-value of 0.01, allowing one to reject the null hypothesis, and concluding that the difference was significant.

Collectively, these results and those of the negative binomial model showed that the connection between early- or late-flowering genotype and rosette number was non-trivial and complicated by family. Family Nor6-12 belongs to the late homozygous group, as does Nor6-06, but Nor6-12 had more rosettes than Nor6-06. Moreover, the rosette number of only two of the early homozygous families (Nor6-51 and S4-41-2) differed from Nor6-06, whereas the rosette number for the other two families was not significantly different from that of Nor6-06.

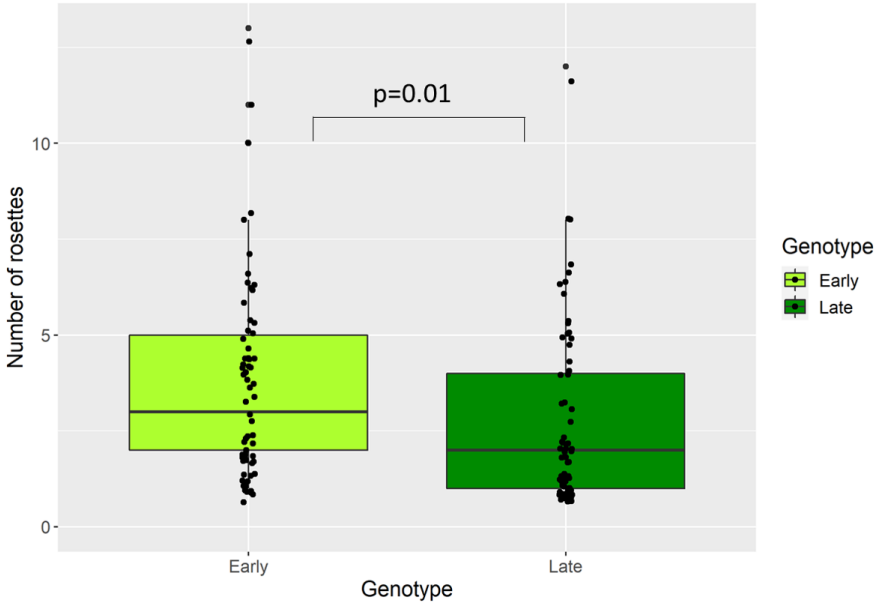

Figure 6.4. The total number of rosettes of the early- and late-flowering genotypes in plot 1. The box plot represents the median and interquartile range of distribution, and the small black dots represent the individual data points. Non-parametric Mann-Whitney U test was applied to the two genotypes data distributions.

6.1.3 FLOWERING EPISODES

To study the flowering behaviour of the plants in the transplantation plots, two variables can be considered: the probability of flowering in a given year and the number of flowering events. Unfortunately, because the plants were only scored two or three times during the summer each year, it was not possible to determine the exact date of flowering, which was, howerver, possible for the greenhouse and growth chamber experiments (section 3.2 and 3.3).

Firstly, I considered the probability of flowering. The data used for this were collected during the last trip in 2019. After the first winter of 2018, plots 1, 2 and 4 provided sufficient flowering data for analysis due to their high survival rate of plants in 2019. To assess whether a difference in flowering probability existed among the eight different families from these three plots in 2019, I fitted a logistic fixed-effect model using family as predictors, similar to the survival probability model in section 6.1.1. In addition, I fitted a logistic mixed-effect model where the plot is included as a random effect. However, this was found to not improve the model's performance, suggesting the strength of the random effect was negligible,

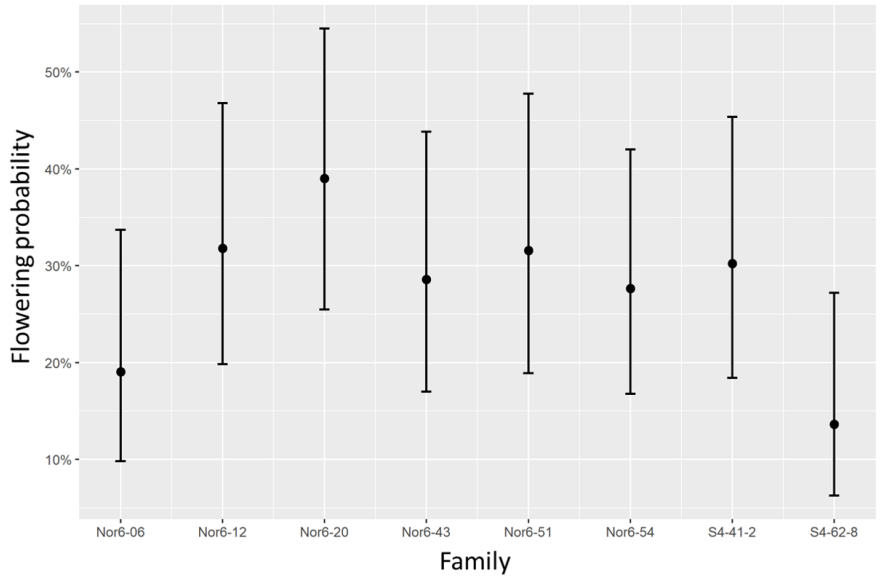

Figure 6.5. The predicted probability of flowering in plot 1, 2 and 4 for the 2019 data according to family. The error bars show the 95% confidence interval of the flowering probability.

and thus I continued to use only the fixed-effect model in this analysis. The equation governing the fixed-effect model was the following:

$$\ln\left[\frac{P(\text{Flowering}=1)}{1-P(\text{Flowering}=1)}\right] = \alpha + \underline{\beta}\cdot\underline{\mathbf{F}}, \tag{6.3}$$

where α is the intercept representing the value for family Nor6-06, $\underline{\beta} = (\beta_{\text{Nor6}-12},\ldots,\beta_{\text{S4}-62-8})$, i.e. a vector containing the model parameters, and $\underline{\mathbf{F}} = (\text{Family}_{\text{Nor6}-12},\ldots,\text{Family}_{\text{S4}-62-8})$, i.e. a vector containing the remaining seven families.

The result of using the fixed-effect model to predict the probability of flowering is shown in Figure 6.5 and Table 6.4. All families had a similar probability of flowering, between $\sim 20-40\%$. Family Nor6-20 had the highest probability of flowering out of the 8, which was marginally significantly different from the reference family Nor6-06 (p=0.05).

When considering the number of flowering events, I used only the data from plot 1 due to the high plant survival rate in this plot, which yielded sufficient data. The number of flowering events was counted from 2019 to 2021, up to a maximum number of events of 3, i.e. a plant flowering every summer. I assumed the number of flowering episodes data in the field followed a Poisson distribution and therefore, the binary response for flowering each year could be summed to test

Family	β_i	σ_i	z-value	Pr(>|z|)
Intercept (Nor6-06)	-1.45	0.39	-3.68	<0.001
Nor6-12	0.68	0.51	1.35	0.18
Nor6-20	1.00	0.51	1.97	0.05
Nor6-43	0.53	0.52	1.02	0.31
Nor6-51	0.67	0.53	1.28	0.20
Nor6-54	0.49	0.51	0.95	0.34
S4-41-2	0.61	0.51	1.19	0.24
S4-62-8	-0.40	0.59	-0.68	0.50

Table 6.4. Model parameters for the flowering probability in plots 1, 2 and 4 for the 2019 flowering data. The table shows the best fitting results of the binomial model given in equation 6.3. Column 1 shows the family; note that the intercept, α, corresponds to family Nor6-06. Column 2 shows the best fitting result for the parameter β_i. Column 3 shows the estimated error on this parameter, σ_i. Column 4 shows the z-value, i.e. β_i/σ_i. Column 5 shows the p-value resulting from this z-value considering the null hypothesis that $\beta_i = 0$.

how many times flowering events occurred among different families from 2019 to 2021. Therefore, I fitted a Poisson model for the number of flowering events during the three winter periods and the model did not indicate any overdispersion. The equation governing the model was the following:

$$\ln(E(\mathrm{N_{flow}})) = \alpha + \underline{\beta} \cdot \underline{\mathbf{F}}, \tag{6.4}$$

where α is the intercept representing the value for family Nor6-06, $\underline{\beta} = (\beta_{\mathrm{Nor6-12}},...,\beta_{\mathrm{S4-62-8}})$, i.e. a vector containing the model parameters, and $\underline{\mathbf{F}} = (\mathrm{Family_{Nor6-12}},...,\mathrm{Family_{S4-62-8}})$, i.e. a vector containing the remaining seven families. On the left-hand side of the equation, $E(\mathrm{N_{flow}})$ is the expected number of flowering events given a Poisson distribution.

The results of fitting a Poisson model to predict the number of flowering events are shown in Figure 6.6 and Table 6.5. Difference existed between the families. This can be easily seen in Table 6.5, where families Nor6-12 and Nor6-51 showed values of β_i significantly different from 0 (p-values < 0.05), i.e. they had an expected occurrence of flowering of ~ 1.75, which was different to family Nor6-06. The expected occurrence of flowering for the other families was $\sim 0.75 - 1.5$, which was indistinguishable from that for family Nor6-06. Notably, in addition to families Nor6-12 and Nor6-51 having a greater predicted number of flowering events than Nor6-06, they also had a greater number of rosettes than Nor6-06 (section 6.1.2).

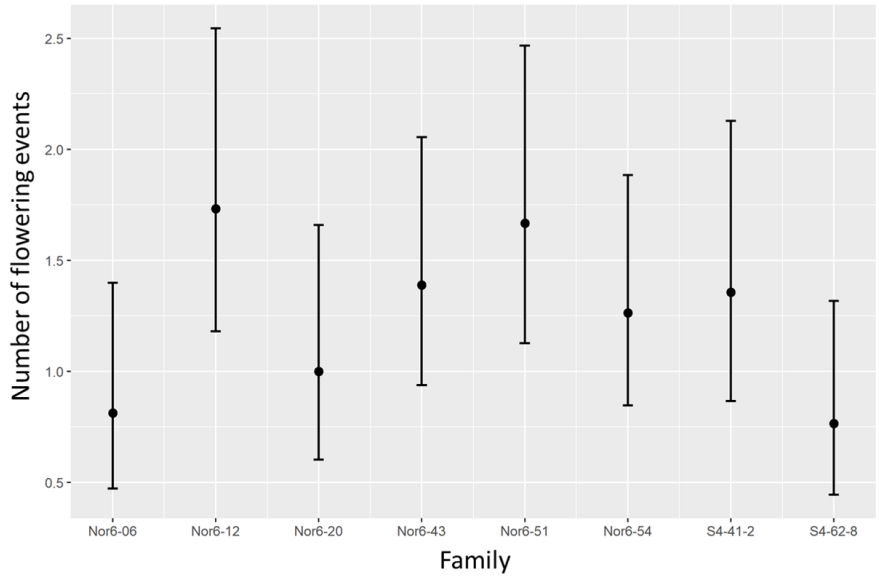

Figure 6.6. The predicted number of flowering events for plants in plot 1 from 2019 to 2021 according to family. The error bars show the 95% confidence interval of the flowering events.

Family	β_i	σ_i	z-value	Pr(>|z|)
Intercept (Nor6-06)	-0.21	0.28	-0.75	0.45
Nor6-12	0.76	0.34	2.23	0.03
Nor6-20	0.21	0.38	0.55	0.58
Nor6-43	0.54	0.34	1.57	0.12
Nor6-51	0.72	0.34	2.10	0.04
Nor6-54	0.44	0.34	1.28	0.20
S4-41-2	0.51	0.36	1.43	0.15
S4-62-8	-0.06	0.39	-0.15	0.88

Table 6.5. Model parameters for the expected number of flowering events of plants in plot 1 from 2019 to 2021. The table shows the best fitting results of the Poisson model given in equation 6.4. Column 1 shows the family; note that the intercept, α, corresponds to family Nor6-06. Column 2 shows the best fitting result for the parameter β_i. Column 3 shows the estimated error on this parameter, σ_i. Column 4 shows the z-value, i.e. β_i/σ_i. Column 5 shows the p-value resulting from this z-value considering the null hypothesis that $\beta_i = 0$.

6.1.4 Number of flowering shoots and siliques

To reveal whether there was a difference in reproductive success amongst the eight families in their natural environment, I scored the number of flowering shoots of each plant and the number of siliques on each flowering shoot. For the following analysis, I used only the data scored in plot 1 during the second trip of 2021 and the data only included flowering individuals.

The relationship between families and the number of flowering shoot is not statistically clear (p-value = 0.25, Kruskal–Wallis test, Figure C.4); however, I further investigated whether differences in silique production occurred amongst these families by dividing the total number of siliques by the total number of flowering shoots for each flowered plant in plot 1 to obtain the mean number of siliques per flowering shoot. As this is a continuous variable, it may be used as a response variable to linear model (i.e. a GLM with a Gaussian linkage function) predicting the mean number of siliques per flowering shoot. The equation governing the model is the following:

$$E(\mathrm{A_{sil}}) = \alpha + \underline{\beta} \cdot \underline{\mathbf{F}}, \tag{6.5}$$

where α is the intercept representing the value for family Nor6-06, $\underline{\beta} = (\beta_{\mathrm{Nor6-12}},\ldots,\beta_{\mathrm{S4-62-8}})$, i.e. a vector containing the model parameters, and $\underline{\mathbf{F}} = (\mathrm{Family}_{\mathrm{Nor6-12}},\ldots,\mathrm{Family}_{\mathrm{S4-62-8}})$, i.e. a vector containing the remaining seven families. On the left-hand side of the equation, $E(\mathrm{A_{sil}})$ is the expected mean number of siliques per flowering shoot given a Gaussian distribution.

The results of the linear model are shown in Figure 6.7 and Table 6.6. Differences existed among the expected mean number of siliques per flowering shoot of each family. This difference was largest for families Nor6-20 and Nor6-43, both of which showed values of β_i significantly different from 0 (p-values = 0.05) (Table 6.6). These families had an expected mean number of siliques per flowering of shoots of ~ 4. The other families had an expected mean number of siliques per flowering shoot of $\sim 5-7$ that was not statistically significantly different from that of family Nor6-06. Notably, families Nor6-06 and S4-62-8 (late-flowering accession) had two of the greatest predicted mean number of siliques per flowering of shoots of ~ 7, whereas they had two of the lowest predicted number of flowering events in section 6.1.3 (Figure 6.6).

To identify the connection between genotype and the mean number of

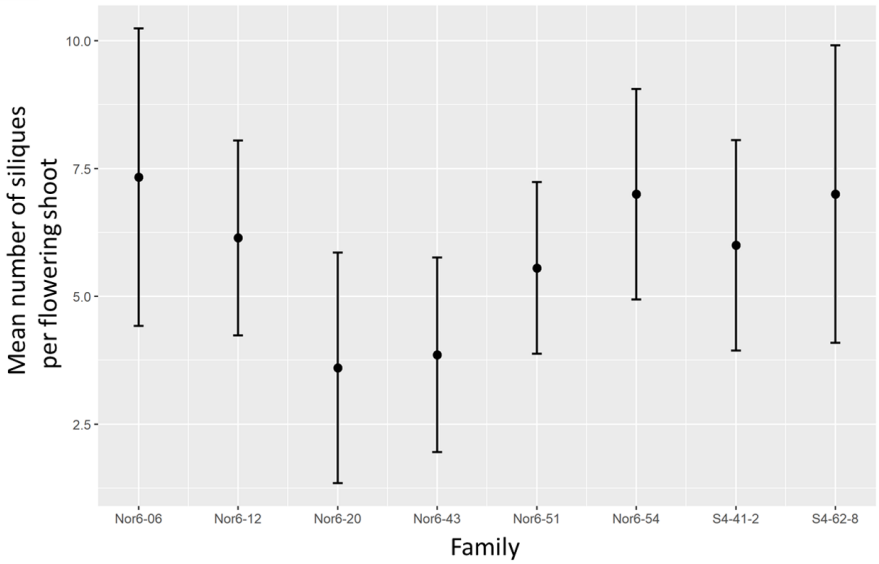

Figure 6.7. The predicted mean number of siliques per flowering shoot in plot 1 according to family. The error bars show the 95% confidence interval of the expected mean number of siliques per flowering shoots.

Family	β_i	σ_i	z-value	Pr(>|z|)
Intercept (Nor6-06)	7.33	1.44	5.10	<0.01
Nor6-12	-1.19	1.72	-0.69	0.49
Nor6-20	-3.73	1.82	-2.05	0.05
Nor6-43	-3.48	1.72	-2.02	0.05
Nor6-51	-1.78	1.66	-1.07	0.29
Nor6-54	-0.33	1.76	-0.19	0.85
S4-41-2	-1.33	1.76	-0.76	0.45
S4-62-8	-0.33	2.03	-0.16	0.87

Table 6.6. Model parameters for mean number of siliques per flowering shoot in plot 1. The table shows the best fitting results of the Gaussian model given in equation 6.4. Column 1 shows the family; note that the intercept, α, corresponds to family Nor6-06. Column 2 shows the best fitting result for the parameter β_i. Column shows the estimated error on this parameter, σ_i. Column 4 shows the z-value, i.e. β_i/σ_i. Column 5 shows the p-value resulting from this z-value considering the null hypothesis that $\beta_i = 0$.

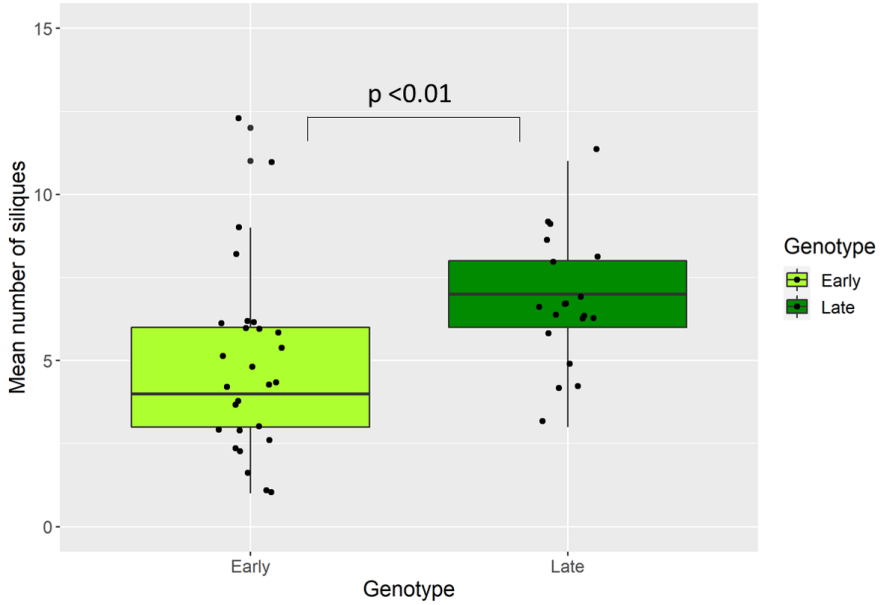

Figure 6.8. Mean number of siliques per flowering shoot of the early- and late-flowering genotypes in plot 1. The boxplot illustrates the data distribution of early and late group based on the genotype result using marker SNP-168 in the QTL region. Nonparametric Mann-Whitney U test was applied to the two genotypes data distributions.

siliques per flowering shoot, I firstly applied non-parametric Kruskal-Wallis test to the 4 families representing either early-flowering (Nor6-20, Nor6-43, No6-51 and S4-41-2) or late-flowering (Nor6-06, Nor6-12, Nor6-54 and S4-62-8) based on the genotyping result using marker SNP-168 in QTL region, similar to section 6.1.2. Test statistics from either the early- or late-flowering group reported a p-value greater than 0.3 and could not reject the null-hypothesis, indicating that 4 distributions within each genotype has been drawn from the same population. Therefore, I grouped the data from all 8 families into two distributions: early- or late-flowering genotype. The resulting distributions are shown in Figure 6.8. The early-flowering genotype distribution (n=27) possesses a median of 4 and an interquartile range of 3; the late-flowering genotype distribution (n=19) possesses median of 7 and interquartile range of 2. The Mann-Whitney U test indicated that the late-flowering genotype plants had a significantly greater number of siliques per flowering shoot (p-value = 0.004) than the early-flowering genotype, allowing the null hypothesis to be rejected.

Collectively, the above results and those of the Gaussian model confirmed

that silique number per flowering shoot differed between early- and late-flowering genotypes and by family. The silique number per flowering shoot for two of the early-flowering homozygous families (Nor6-20 and Nor6-43) was significantly different from that of Nor6-06, which belongs to the late-flowering homozygous group.

6.2 Phenology of labelled plants in the natural environment

6.2.1 Survival and flowering episodes

To document the phenology of *A. alpina* in its natural environment, Dr. Ulla Kemi and Dr. Jörg Wunder initially labelled numerous natural plants *in situ* at different locations in the valley in 2015. A second labelling was conducted in 2017 by Dr. Ulla Kemi and Dr. Stefan Wötzel, which included more individuals from the upper valley. In this analysis, rather than selecting all 114 labelled plants, I selected a subset of the labelled plants that survived for at least two winters, resulting in a total of 30 plants. Six plants (68B, 70, 103, 104, 107 and 108) were situated in the upper valley and experienced late snowmelt (29^{th} week of 2021) whereas the other 24 plants that were distributed in the lower main valley experienced early snowmelt (26^{th} week of 2021).

The labelled plants were scored during each annual visit (left panel of Figure C.1). The scored traits included: survival, the presence of a flowering event, and the number of rosettes and inflorescences. Scoring occurred during the final visit to the site of each year to record the developmental status of each plant at the end of the growing season. This section summarises the scored traits from 2015 to 2021 using heatmaps aligned with a photograph of the valley so that the plants' geographical location is represented.

Figure 6.9 shows one heatmap; the heatmap shows the survival of a labelled natural plant in green, and the solid white circle shows whether the plant flowered in that given year. To examine whether different microhabitats at the site caused differences in survival and flowering, I ordered the labelled plants in the heatmaps to align with their location in the valley. Note that in the heatmap, not all plants were initially scored in 2015 and the scoring of some lines began in 2016 or 2017. Moreover, due to the late snowmelt during the single visit in 2020, it was impossible to score most labelled plants.

Figure 6.9. Survival and flowering heatmap from 2015 to 2021. (Left) An aerial photograph of the valley; the 30 green dots indicate the locations of the individual plants. This photograph was provided by Dr. Ulla Kemi and modified by Yi-Chen Lin for the thesis. (Right) A heatmap showing the years in which a plant survived in green and white solid circles showing the years in which a plant flowered. In the heatmaps, the x-axis indicates the year from 2015 to 2021, and the y-axis indicates the number of individual plants, ordered to align with their location in the valley in the left panel.

The survival data allow the lifespan of *Arabis alpina* in its natural environment to be estimated. Some individuals survived the entire 6-year period (Figure 6.9). These individuals were not localised but were distributed in the upper and lower valley. Although only 26% of the original labelled plants survived for at least three years, 38% survived for 6 years.

Comparison of the survival and flowering heatmap showed that the number of flowering events varied among individuals. Most individuals flowered annually for at least 2–3 years, whereas three individuals (plants 53, 58 and 103) flowered approximately every other year. Two individuals (plants 92 and 94) only flowered once during the whole experiment.

6.2.2 Number of rosettes and inflorescences

The numbers of rosettes and the inflorescences for each plant was scored at each annual visit and these can be used to describe as proxies to indicate the plant's vegetative and reproductive status, respectively. These data are depicted as heatmaps in Figure 6.10. Note that the colour bar is saturated at a value of 10 to highlight differences. Variability in vegetative growth was observed for most individuals and no trends were observed. A similar scenario was observed for reproductive growth; however, many plants flowered for two consecutive years.

6.3 Discussion

6.3.1 Early and late-flowering accessions show different growth strategies in the natural environment

It has been proposed that under certain circumstances, the fitness function in a life-history event may be asymmetric (Austen et al., 2017). Asymmetry in the fitness function indicates that a late reproduction event might lead to a stronger penalty than an early reproduction event. This is because a late event may result in reproductive failure, whereas an early event might only decrease offspring production. Therefore, early-flowering populations are subject to weaker selection pressure than late-flowering populations. However, selection based on fecundity might be balanced by other fitness components (Austen et al., 2017).

Figure 6.10. A heatmap depicting the number of rosettes and inflorescences from 2015 to 2021. (Left) An aerial photograph of the valley. The 30 green dots indicate the locations of the individual plants. This photograph was provided by Dr. Ulla Kemi and modified by Yi-Chen Lin for the thesis. (Centre) A heatmap showing the number of rosettes for each plant in each year, on a coloured scale saturated at a value of 10. (Right) A heatmap showing the number of inflorescence for each plant in each year on a coloured scale saturated at a value of 10. For the heatmaps, the x-axis indicates the year from 2015 to 2021, and the y-axis indicates the plant number, ordered to align with its location in the valley shown in the left panel.

In the transplantation experiment, the early-flowering families produced more rosettes but fewer siliques per flowering shoot than the late-flowering families (sections 6.1.2 and 6.1.4). This lower silique production is consistent with the notion of reduced offspring production; however, the increased production of rosettes is not consistent with the above-mentioned fitness function asymmetry. A similar conflict with the asymmetric fitness function scenario has also been presented by Forrest (2014), who observed that the earlier a plant flowers, the larger is its size. In the field, the additional rosettes on early-flowering plants grew from the basal part of the plant (as shown in Figure 6.11). For plants grown in the growth chamber (section 3.3), where axillary branches are the main indicator of vegetative growth rather than rosettes, those from the early-flowering accessions had an active V1 zone compared with that of the late-flowering accessions. The close-up of treatment 8 in Figure 3.19 of Chapter 3 for plants of the early- and late-flowering accessions is reproduced here in Figure 6.12. This higher activity in the V1 zone, which results in more elongated axillary branches, might be the basis for the greater number of rosettes observed in the field.

In addition to the activity of axillary branches, the increased silique production observed in the natural environment for the late-flowering families is consistent with the results in the growth chamber in Chapter 3 (section 3.3.6); the statistical difference in the number of siliques on the primary shoot was mostly detected for plants that experienced the longest vernalization treatment (Figure 3.12 in Chapter 3), which is closest to the conditions that plants experience in the natural environment. Moreover, the result that the longest vernalization does not cause differences in primary shoot elongation between the two accessions (section 3.3.4) is consistent with the results observed in the natural environment. The difference between the plant height and family is not statistically significant (p-value = 0.42, Kruskal-Wallis test, Figure C.3). Therefore, it appeared fully vernalized, as it is always in the natural environment, the late-flowering accession produced more siliques than the early-flowering accession but remained a similar height to the early-flowering accession.

Collectively, this information suggests a scenario in which the two accessions prioritise different reproductive strategies. The early-flowering accession invests resources into growing numerous rosettes, which ensures that a sufficient number of rosettes flower the subsequent year, whereas the late-flowering accession invests resources into producing a large number of flowers, leading to more siliques, to

Figure 6.11. Aggregated rosette phenotype in the natural environment. An *Arabis alpina* plant growing in the natural environment. The photograph on the right shows a close-up of the region marked by a white square on the left from a different angle. The arrow indicates a vegetative rosette growing from the basal part of the plant. Both photos were taken in September 2019.

maximise offspring production.

6.3.2 The importance of phenology data to understand the perennial species

Various traits that describe the polycarpic life-history of *A. alpina*, such as flowering behaviour and basal bud dormancy, have been well characterised in greenhouse or growth chamber conditions in the last decade (Wang et al., 2009a; Lazaro et al., 2018; Vayssières et al., 2020). However, few studies have focused on polycarpic behaviour in natural populations (Toräng et al., 2015; Hughes et al., 2019; Andrello et al., 2020). The artificial growing environment of greenhouses/growth chambers (due to differences in e.g. soil composition, humidity, nutrient availability) leads to difficulties in extrapolating results for polycarpic traits in these conditions to those that plants experience in their natural environments. For example, the detailed characterisation of the outgrowth of axillary branches in growth chamber conditions (see section 3.3.10) is not applicable to natural plants because they do not produce branched structures, only rosettes. Furthermore, the described zonation in the Spanish accession was impossible to identify in the natural Scandinavian plants due to the phenotype of aggregated rosettes in the basal region, as seen in Figure 6.11.

Despite the above-mentioned difficulties, this study is the first phenology study of this population and documents different polycarpic behaviours in *A. alpina* in its natural environment. Annual visits to labelled natural plants have allowed a 6-year survival period for this Norwegian accession to be recorded. Although *A. alpina* is reported to be a perennial species (Wötzel et al., 2022), we did not record consistent annual flowering behaviour in every labelled plant, but rather observed flowering to be an intermittent event with flowering periods being interspersed with non-flowering years. Similar observations were also reported in six sites spanning most of the elevational range of *A. alpina* in the European Alps where their prediction of the GLM for the probability of flowering among all sites is around 0.5 (Andrello et al., 2020), meaning not every plant in their plots flowered every year. Intriguingly, by scoring the same rosette in a single plant during the two trips in 2021, the formation of floral buds was observed during the second trip in September before the onset of prolonged winter temperatures. In contrast to the notion that floral transition occurs during vernalization in the Spanish Pajares accession,

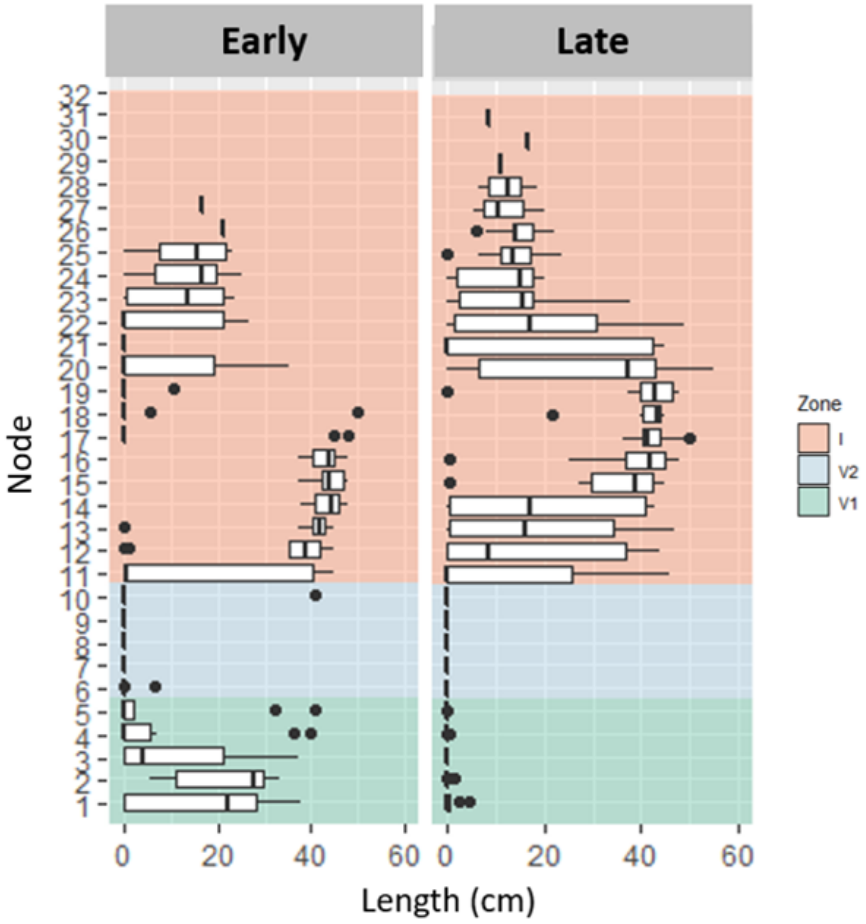

Figure 6.12. The length of axillary branches from plants of the early- and l ate-flowering accessions grown in the growth chamber. The boxplots illustrate the distribution of lengths of axillary branches according to plant node. Twelve plants were scored for each accession and the measurements were performed between 30 to 32 days after a 16-week vernalization treatment. The three different colour backgrounds indicate the V1, V2 dormant and I zones.

we observed that the Norwegian accession underwent floral transition in autumn in the natural environment, which is consistent with the findings of O'Neill et al. (2019).

Different periods of snow cover in this natural environment generate various microhabitats for this population, which increases the complexity of interpreting our collected phenology data. Thus, to better understand the different selection pressures arising from different habitats, more phenology studies of *A. alpina* are needed for plants in diverse environments.

Although local adaption in survival and fruit production has been reported for Spanish and Scandinavian *A. alpina* populations (Toräng et al., 2015), no single phenology study has been performed in any Scandinavian population on the timescale reported here. Thus, the unique combination of phenology and transplantation approaches described in this chapter for the Norwegian population can contribute to strengthen predictions of how natural arctic and alpine plant populations respond to climate change.

.

Chapter 7
Conclusions and perspectives

In this study, I investigated the underlying mechanisms responsible for the variation in flowering behaviour in the Norwegian Riasten population of *A. alpina*, using four different approaches: The phenotypic analysis of plants grown in the greenhouse/growth chamber (Chapter 3), the phenotypic analysis of *in situ* transplantation plants (Chapter 6), genome-wide transcriptome analysis (Chapter 4), and the genetic analysis of a QTL region (Chapter 5).

Firstly, a detailed phenotypic analysis of multiple reproductive traits of early- and late-flowering accessions plants grown in the growth chamber demonstrates different responses to the length of the pre-growth and the vernalization periods between the two accessions. Although the longest vernalization period, closest to that the natural environment, removed the difference in primary shoot elongation, the difference in silique production remained between the two accessions. Additionally, the use of correlation matrices on data covering biological traits from both the primary shoot and the axillary branches not only revealed the dynamics and interconnectedness between flowering time, plant height and silique production but highlighted the unique flowering mechanism in each accession. For the early-flowering accession after vernalization, plants began to flower from the primary shoot and the I zone axillary branches, which was followed by flowering on the V1 zone branches. Conversely, plants in the late-flowering accession only flowered from the primary shoot and the I zone branches. The activity of the axillary meristems within the V1 zone of plants from the early-flowering accession enables flowering to be staggered, which distinguishes the underlying flowering mechanisms of the early- and late-flowering accession plants.

Secondly, *in situ* transplantation at the Norwegian field site also illustrated different growth strategies for the early- and late-flowering accessions in the natural environment. For the early-flowering accession, resources are invested into growing numerous rosettes, ensuring that sufficient rosettes flower the subsequent year, whereas the late-flowering accession invests resources into producing a large number of flowers, leading to more siliques and helping to maximise offspring production. In addition to the difference in silique production under growth chamber and field conditions, an analogy between an active V1 zone observed in the growth chamber condition and the greater number of rosettes of the early-flowering accession in the natural environment confirms the similar findings of these two research approaches. This suggests that a selection pressure exists in the field, which leads to local adaptation in specific traits that reflect the short growing season in the natural environment.

Thirdly, genome-wide transcriptome analysis revealed a link between the phenotypic difference in the growth of floral buds and the gene regulatory networks between the two accessions in the absence of vernalization. One key gene involved in the regulatory network is *PEP1*, which was differentially expressed between the early- and late-flowering accessions without vernalization, consistent with the phenotypic differences. Notably, the decrease in the expression level of *PEP1* in the early-flowering accession was not affected by vernalization, which is strikingly different to the situation for the Spanish Pajares accession of *A. alpina*, which cannot flower without vernalization. This implies that different gene regulatory networks regulate flowering of the Spanish and the Norwegian *A .alpina* accessions. Hierarchical clustering analysis of key floral repressor and organ identity genes identified a clear change in expression pattern between these two groups of genes only in the early-flowering accession without vernalization, suggesting that the timing of expression of different gene regulatory networks is involved in regulating flowering in the absence of vernalization in Norwegian *A. alpina* accessions. Furthermore, genome-wide co-expression network analysis revealed different structures of gene clustering between the two accessions. The majority of floral organ identity and floral repressor genes in the early-flowering accession belonged to two clusters, which were expressed in opposite patterns, whereas in the late-flowering accession, the floral repressors were split between multiple clusters which, while

showing similar patterns, were asynchronous. This separation of genes into multiple, asynchronously expressed clusters suggests that multiple gene regulatory networks exist for the late-flowering accession, leading to a protracted development of floral buds, categorised as the bud-arrest phenotype in this study.

Finally, genetic analysis of the QTL mapping defined a region of 759 kb on chromosome 8, which included casual gene(s) in which structural variants were also present. The use of RNA-sequencing data to examine the expression level of the 120 genes in this region identified that eight genes were differentially expressed between early- and late-flowering accessions. One of these 8 genes, *ERL2*, is a good candidate causal gene for early-flowering. This is because *ERL2* was co-expressed with the gene cluster that includes both *SVP* and *TFL1* (known floral repressors) and the expression pattern of this cluster correlated with the development of an elongated primary shoot with completed inflorescence. Moreover, gene ontology enrichment analysis of this *ERL2* cluster in the two accessions suggests phytohormones play a role in the bud-arrest phenotype. This was confirmed by showing that the application of exogeneous gibberellins rescued the arrested primary shoot.

The work presented in this thesis sheds light on how flowering is regulated in Norwegian *A. alpina* accessions on the molecular and genetic levels. Moreover, the detailed phenotypic analysis of plants grown in growth chamber conditions and at the Norwegian field site as part of the *in situ* transplantation experiment establishes a solid foundation for understanding how local adaptation leads to different flowering mechanisms.

The work in this thesis provides a number of promising avenues for further study of the flowering mechanisms within the Norwegian Riasten *A. alpina* population. Because flowering has been shown here to be highly sensitive to the vernalization and pre-growth periods, further investigation of these environmental cues could be conducted by performing a detailed phenotypic analysis in growth chamber conditions of plants from the near-isogenic lines generated by the fine-mapping experiments. As shown in this thesis, the combination of growth chamber/greenhouse experiments with fieldwork is highly productive; therefore, to further understand how the identified QTL region contributes to local adaptation, *in situ* transplantation experiments should be performed to score the flowering behaviour of the near-isogenic lines at the Norwegian field site. To identify the causal gene(s) involved in bud growth without vernalization, *ERL2* and other differentially expressed genes in the QTL region on chromosome 8 could be mutated

to observe the effect of their loss of function on flowering time. Furthermore, resequencing of the current QTL region would provide better gene annotations and a better characterisation of the region. Finally, to better determine the gene networks that regulate flowering in the absence of vernalization in *A. alpina*, the results of genome-wide transcriptome analysis in this study could be compared with the gene regulatory networks of French *A. alpina* accessions for which early- and late-flowering phenotypes in the absence of vernalization have been identified.

Appendix A
Section 3.2 data

Table A.1. Data for the length and flowering stage (0: no open flower; 1: flowering) of axillary branches on indvidual node from both accessions after 8 weeks of vernalization. In column 1, E and L represents early- and late-flowering accession, respectively.

PlantID	Node	Length (cm)	Flowering stage
E1	1	1.5	0
E1	2	2.0	0
E1	3	16.0	1
E1	4	0.2	0
E1	5	27.0	1
E1	6	0.2	0
E1	7	0.2	0
E1	8	0.2	0
E1	9	0.2	0
E1	10	0.2	0
E1	11	29.0	1
E1	12	28.5	1
E1	13	18.0	1
E1	14	26.5	1
E1	15	20.5	1
E1	16	23.0	1
E1	17	19.0	1

Continued on next page

Table A.1 – *Continued from previous page*

PlantID	Node	Length (cm)	Flowering stage
E1	18	30.0	1
E1	19	22.0	1
E1	20	23.0	1
E1	21	15.5	1
E1	22	0.2	0
E1	23	0.2	0
E1	24	0.2	0
E1	25	0.2	0
E1	26	0.2	0
E1	27	0.2	0
E1	28	0.2	0
E1	29	0.2	0
E1	30	0.2	0
E1	31	0.2	0
E2	1	23.0	1
E2	2	16.0	1
E2	3	20.5	1
E2	4	3.0	0
E2	5	0.2	0
E2	6	0.2	0
E2	7	0.2	0
E2	8	0.2	0
E2	9	0.2	0
E2	10	0.2	0
E2	11	0.2	0
E2	12	4.0	0
E2	13	30.0	1
E2	14	19.0	1
E2	15	20.5	1
E2	16	26.5	1
E2	17	20.5	1
E2	18	18.5	1

Continued on next page

Table A.1 – *Continued from previous page*

PlantID	Node	Length (cm)	Flowering stage
E2	19	25.5	1
E2	20	18.0	1
E2	21	0.2	0
E2	22	0.2	0
E2	23	0.2	0
E2	24	0.2	0
E2	25	0.2	0
E2	26	0.2	0
E2	27	0.2	0
E2	28	0.2	0
E2	29	0.2	0
E2	30	0.2	0
E3	1	1.0	0
E3	2	1.0	0
E3	3	3.0	0
E3	4	9.0	0
E3	5	0.2	0
E3	6	0.2	0
E3	7	0.2	0
E3	8	0.2	0
E3	9	0.2	0
E3	10	0.2	0
E3	11	0.5	0
E3	12	2.5	0
E3	13	28.0	1
E3	14	21.5	1
E3	15	29.0	1
E3	16	23.5	1
E3	17	21.0	1
E3	18	27.0	1
E3	19	28.0	1
E3	20	22.0	1

Continued on next page

Table A.1 – *Continued from previous page*

PlantID	Node	Length (cm)	Flowering stage
E3	21	22.0	1
E3	22	0.2	0
E3	23	0.2	0
E3	24	0.2	0
E3	25	0.2	0
E3	26	0.2	0
E3	27	0.2	0
E3	28	0.2	0
E3	29	0.2	0
E3	30	0.2	0
E3	31	0.2	0
L1	1	1.5	0
L1	2	0.5	0
L1	3	1.5	0
L1	4	0.5	0
L1	5	0.2	0
L1	6	0.2	0
L1	7	0.2	0
L1	8	0.2	0
L1	9	6.0	0
L1	10	8.0	1
L1	11	6.5	0
L1	12	6.5	0
L1	13	10.5	1
L1	14	9.0	1
L1	15	9.0	1
L1	16	12.0	1
L1	17	4.0	0
L1	18	3.5	0
L1	19	1.0	0
L1	20	0.2	0
L1	21	0.2	0

Continued on next page

Table A.1 – *Continued from previous page*

PlantID	Node	Length (cm)	Flowering stage
L1	22	0.2	0
L1	23	0.2	0
L1	24	0.2	0
L1	25	0.2	0
L2	1	0.2	0
L2	2	0.5	0
L2	3	1.5	0
L2	4	0.2	0
L2	5	0.0	0
L2	6	0.2	0
L2	7	0.2	0
L2	8	5.5	0
L2	9	0.2	0
L2	10	0.2	1
L2	11	8.0	0
L2	12	7.0	0
L2	13	8.0	1
L2	14	7.5	1
L2	15	9.0	1
L2	16	9.0	1
L2	17	11.5	0
L2	18	8.5	0
L2	19	3.5	0
L2	20	6.0	0
L2	21	7.0	0
L2	22	5.5	0
L2	23	2.0	0
L2	24	0.2	0
L2	25	0.2	0
L2	26	0.2	0
L2	27	0.2	0
L2	28	0.2	0

Continued on next page

Table A.1 – *Continued from previous page*

PlantID	Node	Length (cm)	Flowering stage
L2	29	0.2	0
L2	30	0.2	0
L2	31	0.2	0
L2	32	0.2	0
L3	1	0.2	0
L3	2	0.2	0
L3	3	0.2	0
L3	4	2.0	0
L3	5	0.0	0
L3	6	0.0	0
L3	7	0.2	0
L3	8	N/A	0
L3	9	0.2	0
L3	10	0.2	0
L3	11	0.2	0
L3	12	6.0	0
L3	13	6.5	0
L3	14	9.5	1
L3	15	7.0	1
L3	16	4.5	1
L3	17	0.5	0
L3	18	15.0	1
L3	19	10.0	1
L3	20	13.5	1
L3	21	15.0	1
L3	22	16.5	1
L3	23	13.0	1
L3	24	3.0	0
L3	25	0.2	0
L3	26	0.2	0
L3	27	0.2	0
L3	28	0.2	0

Continued on next page

Table A.1 – *Continued from previous page*

PlantID	Node	Length (cm)	Flowering stage
L3	29	0.2	0
L3	30	0.2	0
L3	31	0.2	0
L3	32	0.2	0

Table A.2. Data for the length and flowering stage (0: no open flower; 1: flowering) of axillary branches on indvidual node from both accessions after 12 weeks of vernalization. In column 1, E and L represents early- and late-flowering accession, respectively.

PlantID	Node	Length (cm)	Flowering stage
E4	1	8.0	0
E4	2	2.5	0
E4	3	0.2	0
E4	4	0.2	0
E4	5	0.2	0
E4	6	0.2	0
E4	7	0.2	0
E4	8	0.2	0
E4	9	0.2	0
E4	10	28.5	1
E4	11	28.0	1
E4	12	26.0	1
E4	13	26.0	1
E4	14	28.0	1
E4	15	27.0	1
E4	16	26.5	1
E4	17	25.5	1
E4	18	0.2	0
E4	19	0.2	0
E4	20	0.2	0
E4	21	0.2	0

Continued on next page

Table A.2 – *Continued from previous page*

PlantID	Node	Length (cm)	Flowering stage
E4	22	0.2	0
E4	23	0.2	0
E4	24	0.2	0
E4	25	0.2	0
E4	26	0.2	0
E4	27	0.2	0
E4	28	1.0	0
E5	1	0.2	0
E5	2	10.5	0
E5	3	3.0	0
E5	4	2.0	0
E5	5	3.0	0
E5	6	6.0	0
E5	7	0.2	0
E5	8	0.2	0
E5	9	0.2	0
E5	10	0.2	0
E5	11	26.5	1
E5	12	27.0	1
E5	13	26.0	1
E5	14	30.0	1
E5	15	26.0	1
E5	16	30.0	1
E5	17	28.0	1
E5	18	26.5	1
E5	19	27.0	1
E5	20	21.0	1
E5	21	22.0	1
E5	22	0.2	0
E5	23	0.2	0
E5	24	0.2	0
E5	25	0.2	0

Continued on next page

Table A.2 – *Continued from previous page*

PlantID	Node	Length (cm)	Flowering stage
E5	26	0.2	0
E5	27	0.2	0
E5	28	0.2	0
E5	29	0.2	0
E5	30	0.2	0
E5	31	0.2	0
E5	32	6.0	0
E6	1	0.2	0
E6	2	8.0	0
E6	3	5.0	0
E6	4	22.0	1
E6	5	24.5	1
E6	6	3.0	0
E6	7	0.2	0
E6	8	0.2	0
E6	9	18.5	1
E6	10	0.2	1
E6	11	0.2	0
E6	12	24.0	1
E6	13	16.5	1
E6	14	26.0	1
E6	15	24.5	1
E6	16	25.5	1
E6	17	25.5	1
E6	18	27.5	1
E6	19	27.0	1
E6	20	25.0	1
E6	21	23.5	1
E6	22	0.2	0
E6	23	0.2	0
E6	24	0.2	0
E6	25	0.2	0

Continued on next page

Table A.2 – *Continued from previous page*

PlantID	Node	Length (cm)	Flowering stage
E6	26	0.2	0
E6	27	0.2	0
E6	28	0.2	0
E6	29	0.2	0
E6	30	0.2	0
E6	31	10.0	0
L4	1	2.5	0
L4	2	17.5	1
L4	3	1.5	0
L4	4	2.0	0
L4	5	N/A	0
L4	6	0.2	0
L4	7	0.2	0
L4	8	12.5	1
L4	9	0.2	0
L4	10	20.0	1
L4	11	21.0	1
L4	12	18.5	1
L4	13	25.0	1
L4	14	20.0	1
L4	15	25.0	1
L4	16	21.0	1
L4	17	18.5	1
L4	18	18.5	1
L4	19	22.5	1
L4	20	20.5	1
L4	21	3.0	0
L4	22	19.0	1
L4	23	0.4	0
L4	24	0.2	0
L4	25	0.2	0
L4	26	0.2	0

Continued on next page

Table A.2 – *Continued from previous page*

PlantID	Node	Length (cm)	Flowering stage
L4	27	0.2	0
L4	28	0.2	0
L4	29	0.2	0
L4	30	0.2	0
L4	31	N/A	0
L4	32	1.5	0
L4	33	0.2	0
L5	1	2.0	0
L5	2	2.5	0
L5	3	4.0	0
L5	4	7.5	0
L5	5	8.0	1
L5	6	0.2	0
L5	7	0.2	0
L5	8	0.2	0
L5	9	0.2	0
L5	10	0.2	0
L5	11	23.5	1
L5	12	22.5	1
L5	13	9.0	0
L5	14	21.0	1
L5	15	26.5	1
L5	16	26.0	1
L5	17	24.0	1
L5	18	24.0	1
L5	19	17.0	1
L5	20	19.0	1
L5	21	22.0	1
L5	22	18.0	1
L5	23	21.0	1
L5	24	0.2	0
L5	25	0.2	0

Continued on next page

Table A.2 – *Continued from previous page*

PlantID	Node	Length (cm)	Flowering stage
L5	26	0.2	0
L5	27	0.2	0
L5	28	0.2	0
L5	29	0.2	0
L5	30	0.2	0
L5	31	0.2	0
L5	32	0.2	0
L6	1	4.5	0
L6	2	5.0	0
L6	3	0.5	0
L6	4	5.0	0
L6	5	0.2	0
L6	6	N/A	0
L6	7	0.2	0
L6	8	0.2	0
L6	9	0.2	0
L6	10	18.5	1
L6	11	N/A	0
L6	12	12.5	1
L6	13	19.0	1
L6	14	26.0	1
L6	15	22.5	1
L6	16	23.0	1
L6	17	24.0	1
L6	18	23.0	1
L6	19	21.0	1
L6	20	21.0	1
L6	21	21.0	1
L6	22	0.2	0
L6	23	0.2	0
L6	24	0.2	0
L6	25	0.2	0

Continued on next page

Table A.2 – *Continued from previous page*

PlantID	Node	Length (cm)	Flowering stage
L6	26	0.2	0
L6	27	0.2	0
L6	28	0.2	0
L6	29	0.2	0
L6	30	0.2	0

Table A.3. Data for the length and flowering stage (0: no open flower; 1: flowering) of axillary branches on indvidual node from both accessions after 16 weeks of vernalization. In column 1, E and L represents early- and late-flowering accession, respectively.

PlantID	Node	Length (cm)	Flowering stage
E7	1	8.0	0
E7	2	5.5	0
E7	3	21.5	0
E7	4	2.5	0
E7	5	2.5	0
E7	6	6.5	0
E7	7	0.2	0
E7	8	0.2	0
E7	9	35.0	1
E7	10	3.0	0
E7	11	33.0	1
E7	12	34.0	1
E7	13	30.0	1
E7	14	31.0	1
E7	15	33.0	1
E7	16	30.0	1
E7	17	26.0	1
E7	18	29.0	1
E7	19	29.0	1
E7	20	21.0	1

Continued on next page

Table A.3 – *Continued from previous page*

PlantID	Node	Length (cm)	Flowering stage
E7	21	1.4	0
E7	22	1.1	0
E7	23	1.1	0
E7	24	0.8	0
E7	25	0.5	0
E7	26	0.6	0
E7	27	0.6	0
E7	28	0.6	0
E7	29	0.7	0
E7	30	0.7	0
E7	31	0.7	0
E8	1	0.2	0
E8	2	34.0	1
E8	3	4.0	0
E8	4	35.0	1
E8	5	0.5	0
E8	6	0.4	0
E8	7	0.4	0
E8	8	0.4	0
E8	9	0.4	0
E8	10	0.4	0
E8	11	34.0	1
E8	12	34.0	1
E8	13	34.0	1
E8	14	32.0	1
E8	15	30.0	1
E8	16	30.0	1
E8	17	29.5	1
E8	18	30.0	1
E8	19	32.0	1
E8	20	32.0	1
E8	21	32.0	1

Continued on next page

Table A.3 – *Continued from previous page*

PlantID	Node	Length (cm)	Flowering stage
E8	22	0.2	0
E8	23	0.2	0
E8	24	0.2	0
E8	25	0.2	0
E8	26	0.2	0
E8	27	0.2	0
E8	28	0.2	0
E8	29	0.2	0
E8	30	0.2	0
E8	31	4.0	0
E8	32	1.5	0
E9	1	0.4	0
E9	2	3.5	0
E9	3	23.0	1
E9	4	0.5	0
E9	5	0.4	0
E9	6	0.4	0
E9	7	0.4	0
E9	8	0.4	0
E9	9	0.4	0
E9	10	0.4	0
E9	11	0.4	0
E9	12	26.5	1
E9	13	4.0	0
E9	14	32.0	1
E9	15	36.0	1
E9	16	34.0	1
E9	17	30.5	1
E9	18	31.0	1
E9	19	31.0	1
E9	20	30.0	1
E9	21	33.0	1

Continued on next page

Table A.3 – *Continued from previous page*

PlantID	Node	Length (cm)	Flowering stage
E9	22	0.2	0
E9	23	0.2	0
E9	24	0.2	0
E9	25	0.7	0
E9	26	0.2	0
E9	27	0.2	0
E9	28	0.2	0
E9	29	9.0	0
E9	30	1.0	0
L7	1	29.0	0
L7	2	25.0	1
L7	3	3.0	0
L7	4	2.5	0
L7	5	3.5	0
L7	6	0.5	0
L7	7	0.2	0
L7	8	0.2	0
L7	9	0.2	0
L7	10	22.5	1
L7	11	30.0	1
L7	12	30.0	1
L7	13	30.0	1
L7	14	27.5	1
L7	15	27.0	1
L7	16	24.0	1
L7	17	30.0	1
L7	18	7.0	1
L7	19	26.5	1
L7	20	23.0	1
L7	21	25.0	1
L7	22	6.5	0
L7	23	29.0	1

Continued on next page

Table A.3 – *Continued from previous page*

PlantID	Node	Length (cm)	Flowering stage
L7	24	5.0	0
L7	25	4.5	0
L7	26	1.4	0
L7	27	0.2	0
L7	28	0.2	0
L7	29	0.2	0
L7	30	0.2	0
L7	31	0.2	0
L7	32	0.2	0
L7	33	0.2	0
L7	34	0.2	0
L8	1	2.5	0
L8	2	0.4	0
L8	3	5.0	0
L8	4	2.5	0
L8	5	0.4	0
L8	6	0.4	0
L8	7	0.4	0
L8	8	0.4	0
L8	9	0.4	0
L8	10	25.0	1
L8	11	28.5	1
L8	12	24.0	1
L8	13	7.5	0
L8	14	30.0	1
L8	15	23.5	1
L8	16	29.5	1
L8	17	0.2	0
L8	18	29.0	1
L8	19	0.2	0
L8	20	0.2	0
L8	21	28.5	1

Continued on next page

Table A.3 – *Continued from previous page*

PlantID	Node	Length (cm)	Flowering stage
L8	22	32.0	1
L8	23	34.5	1
L8	24	30.0	1
L8	25	0.4	0
L8	26	0.1	0
L8	27	0.2	0
L8	28	0.2	0
L8	29	0.2	0
L8	30	0.2	0
L8	31	0.2	0
L8	32	0.2	0
L8	33	0.2	0
L9	1	0.5	0
L9	2	0.5	0
L9	3	0.5	0
L9	4	23.0	1
L9	5	0.2	0
L9	6	0.2	0
L9	7	0.2	0
L9	8	0.2	0
L9	9	0.2	0
L9	10	1.0	0
L9	11	0.2	0
L9	12	0.2	0
L9	13	19.0	1
L9	14	29.0	1
L9	15	39.0	1
L9	16	33.0	1
L9	17	35.0	1
L9	18	33.0	1
L9	19	34.0	1
L9	20	39.0	1

Continued on next page

Table A.3 – *Continued from previous page*

PlantID	Node	Length (cm)	Flowering stage
L9	21	34.0	1
L9	22	32.0	1
L9	23	34.0	1
L9	24	14.0	1
L9	25	0.2	0
L9	26	0.2	0
L9	27	0.2	0
L9	28	0.2	0
L9	29	0.2	0
L9	30	0.2	0
L9	31	0.2	0
L9	32	0.2	0
L9	33	0.2	0

Table A.4. Data for the length and flowering stage (0: no open flower; 1: flowering) of axillary branches on indvidual node from both accessions after 20 weeks of vernalization. In column 1, E and L represents early- and late-flowering accession, respectively.

PlantID	Node	Length (cm)	Flowering stage
E10	1	6.0	1
E10	2	29.0	1
E10	3	2.5	0
E10	4	3.0	0
E10	5	3.0	0
E10	6	0.4	0
E10	7	0.4	0
E10	8	0.4	0
E10	9	0.4	0
E10	10	0.2	0
E10	11	5.0	0
E10	12	22.0	1

Continued on next page

Table A.4 – *Continued from previous page*

PlantID	Node	Length (cm)	Flowering stage
E10	13	34.0	1
E10	14	38.0	1
E10	15	33.0	1
E10	16	11.0	1
E10	17	23.0	1
E10	18	24.5	1
E10	19	27.0	1
E10	20	0.5	0
E10	21	32.0	1
E10	22	0.2	0
E10	23	0.2	0
E10	24	0.2	0
E10	25	0.2	0
E10	26	0.2	0
E10	27	0.2	0
E10	28	0.2	0
E10	29	0.2	0
E10	30	0.2	0
E10	31	0.2	0
E11	1	33.0	1
E11	2	35.0	1
E11	3	34.0	1
E11	4	3.5	0
E11	5	3.5	0
E11	6	0.4	0
E11	7	0.4	0
E11	8	0.2	0
E11	9	0.2	0
E11	10	0.2	0
E11	11	0.2	0
E11	12	0.2	0
E11	13	2.5	0

Continued on next page

Table A.4 – *Continued from previous page*

PlantID	Node	Length (cm)	Flowering stage
E11	14	36.5	1
E11	15	30.0	1
E11	16	6.5	0
E11	17	2.0	0
E11	18	32.0	1
E11	19	30.0	1
E11	20	4.0	0
E11	21	34.0	1
E11	22	36.5	1
E11	23	0.2	0
E11	24	0.2	0
E11	25	0.2	0
E11	26	0.2	0
E11	27	0.2	0
E11	28	0.2	0
E11	29	0.2	0
E11	30	0.2	0
E11	31	0.2	0
E11	32	0.2	0
E11	33	7.0	0
E12	1	2.0	0
E12	2	6.5	0
E12	3	0.5	0
E12	4	8.0	0
E12	5	3.5	0
E12	6	23.0	1
E12	7	0.4	0
E12	8	3.5	0
E12	9	5.5	0
E12	10	0.5	0
E12	11	30.0	1
E12	12	36.0	1

Continued on next page

Table A.4 – *Continued from previous page*

PlantID	Node	Length (cm)	Flowering stage
E12	13	35.5	1
E12	14	33.5	1
E12	15	25.0	1
E12	16	34.0	1
E12	17	31.0	1
E12	18	0.4	0
E12	19	0.2	0
E12	20	0.2	0
E12	21	0.2	0
E12	22	0.2	0
E12	23	0.2	0
E12	24	0.2	0
E12	25	0.2	0
E12	26	0.2	0
E12	27	0.2	0
E12	28	0.2	0
L10	1	26.0	1
L10	2	30.0	1
L10	3	3.5	0
L10	4	6.0	0
L10	5	2.0	0
L10	6	0.2	0
L10	7	0.2	0
L10	8	7.5	0
L10	9	33.0	1
L10	10	33.0	1
L10	11	35.0	1
L10	12	34.0	1
L10	13	36.0	1
L10	14	37.0	1
L10	15	37.5	1
L10	16	35.0	1

Continued on next page

Table A.4 – *Continued from previous page*

PlantID	Node	Length (cm)	Flowering stage
L10	17	36.0	1
L10	18	34.0	1
L10	19	36.0	1
L10	20	38.0	1
L10	21	9.5	0
L10	22	0.2	0
L10	23	0.4	0
L10	24	0.2	0
L10	25	0.2	0
L10	26	0.2	0
L10	27	0.2	0
L10	28	0.2	0
L10	29	0.2	0
L11	1	0.5	0
L11	2	28.0	1
L11	3	0.5	0
L11	4	2.5	0
L11	5	2.5	0
L11	6	0.2	0
L11	7	0.2	0
L11	8	0.2	0
L11	9	0.2	0
L11	10	0.4	0
L11	11	7.0	0
L11	12	25.0	1
L11	13	2.0	0
L11	14	33.0	1
L11	15	27.5	1
L11	16	34.5	1
L11	17	33.0	1
L11	18	36.0	1
L11	19	28.5	1

Continued on next page

Table A.4 – *Continued from previous page*

PlantID	Node	Length (cm)	Flowering stage
L11	20	32.5	1
L11	21	0.2	0
L11	22	30.5	1
L11	23	35.0	1
L11	24	33.0	1
L11	25	30.5	1
L11	26	0.4	0
L11	27	0.4	0
L11	28	0.4	0
L11	29	0.2	0
L11	30	0.2	0
L11	31	0.2	0
L11	32	0.2	0
L11	33	0.2	0
L11	34	0.2	0
L11	35	0.2	0
L12	1	0.5	0
L12	2	2.0	0
L12	3	24.5	1
L12	4	2.5	0
L12	5	0.2	0
L12	6	0.2	0
L12	7	0.2	0
L12	8	0.2	0
L12	9	0.4	0
L12	10	0.4	0
L12	11	0.4	0
L12	12	0.4	0
L12	13	19.0	1
L12	14	6.5	1
L12	15	25.5	1
L12	16	33.0	1

Continued on next page

Table A.4 – *Continued from previous page*

PlantID	**Node**	**Length (cm)**	**Flowering stage**
L12	17	35.0	1
L12	18	38.0	1
L12	19	31.0	1
L12	20	32.0	1
L12	21	34.0	1
L12	22	32.0	1
L12	23	34.0	1
L12	24	0.2	0
L12	25	0.2	0
L12	26	0.2	0
L12	27	0.2	0
L12	28	0.2	0
L12	29	0.2	0
L12	30	0.2	0
L12	31	0.2	0
L12	32	0.2	0

Appendix B

Section 3.3 data

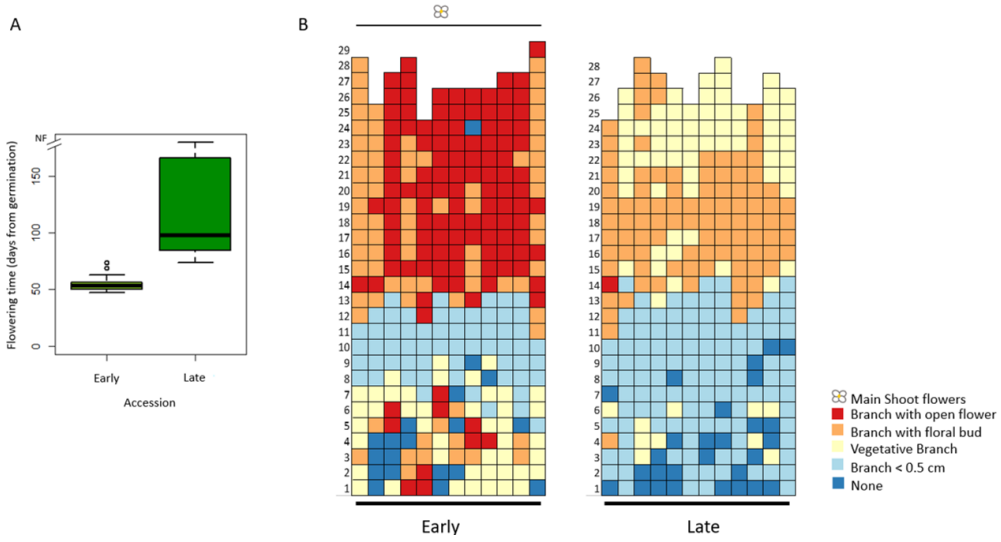

Figure B.1. The flowering of the primary shoot and developmental fate for the axillary branche on each node from early- and late-flowering accession without vernalization under growth-chamber conditions. (A) Flowering time of the primary shoot for each accession (n=12) without vernalization under growth-chamber conditions. (B) The flower symbols above the columns indicate the primary shoot has flowered. Each column represents an individual plant and each pixel shows the developmental stage of the axillary meristem at that node. The axillary meristems were assigned to one of five developmental categories: branches with open flower (red), branches with visible floral bud (orange), branches with vegetative meristem (yellow), branches smaller than 0.5 cm (light blue), and no axillary meristem present (dark blue).

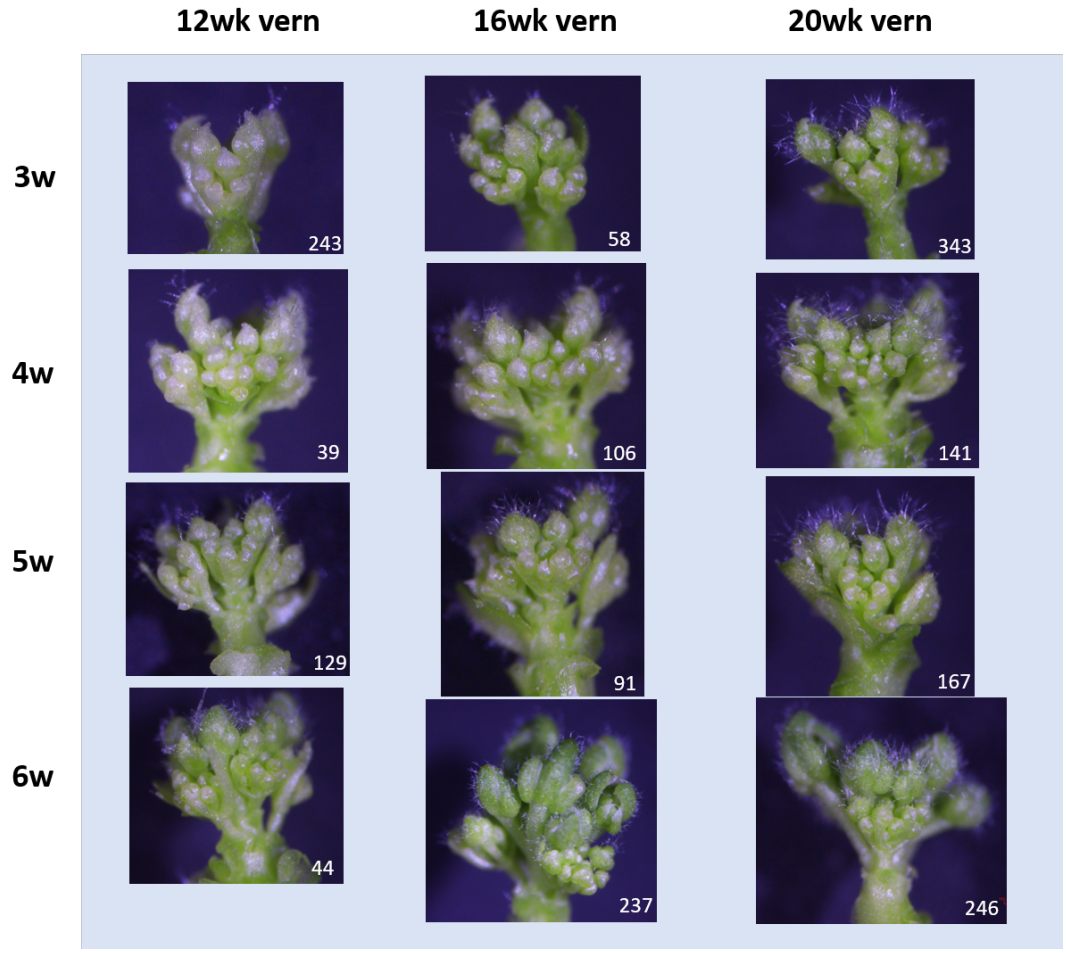

Figure B.2. Microscopic examination of the apices of plants of the early-flowering accession when vernalization ceases under growth-chamber conditions. The apices of early-flowering plants grown for different lengths of time prior to 12, 16 and 20 weeks of vernalization. The plant ID number is shown as white text in each image.

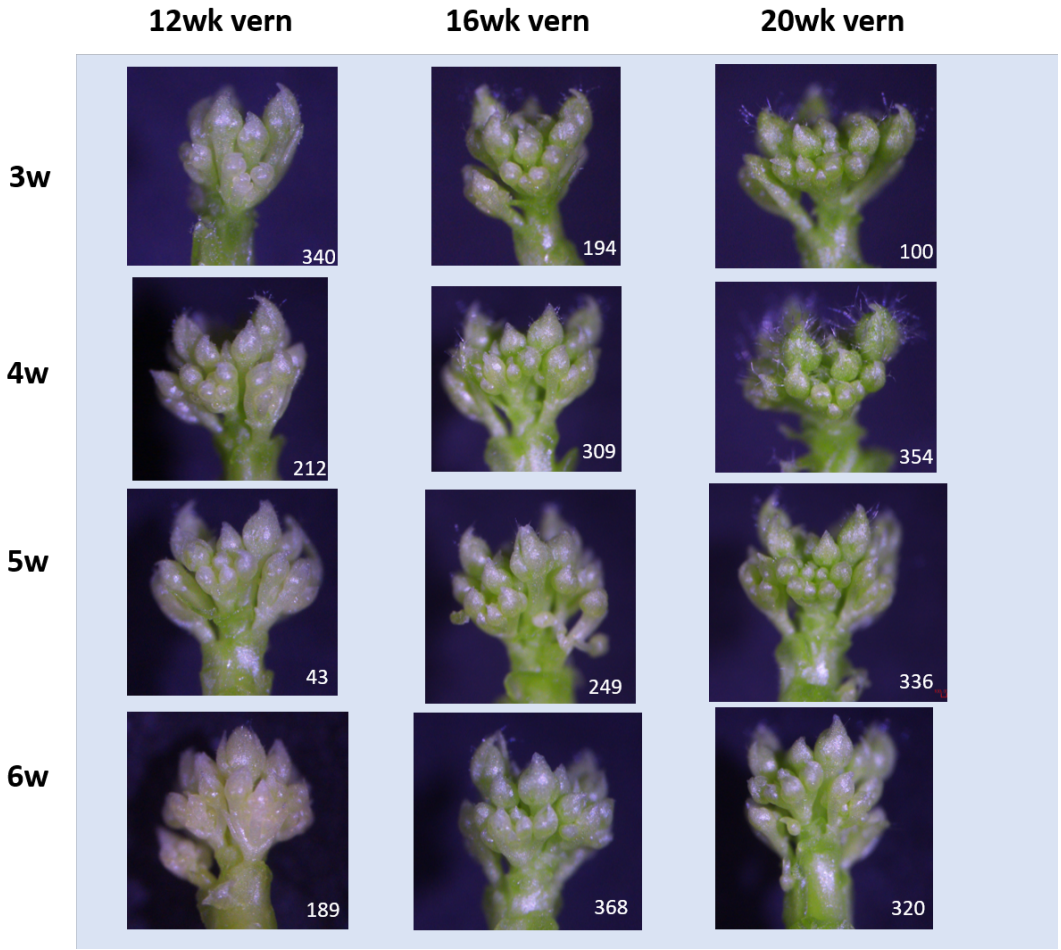

Figure B.3. Microscopic examination of the apices of plants of the late-flowering accession when vernalization ceases under growth-chamber conditions. The apices of early-flowering plants grown for different lengths of time prior to 12, 16 and 20 weeks of vernalization. The plant ID number is shown as white text in each image.

Appendix C
IN SITU transplantation Data

Figure C.1. Photograph of labelled plant in the natural environment (left) and *in situ* transplantation experiment (right). The labelled plant in the natural environment was photographed in September 2021. The labelled plant of the *in situ* transplantation experiment was photographed in August 2019.

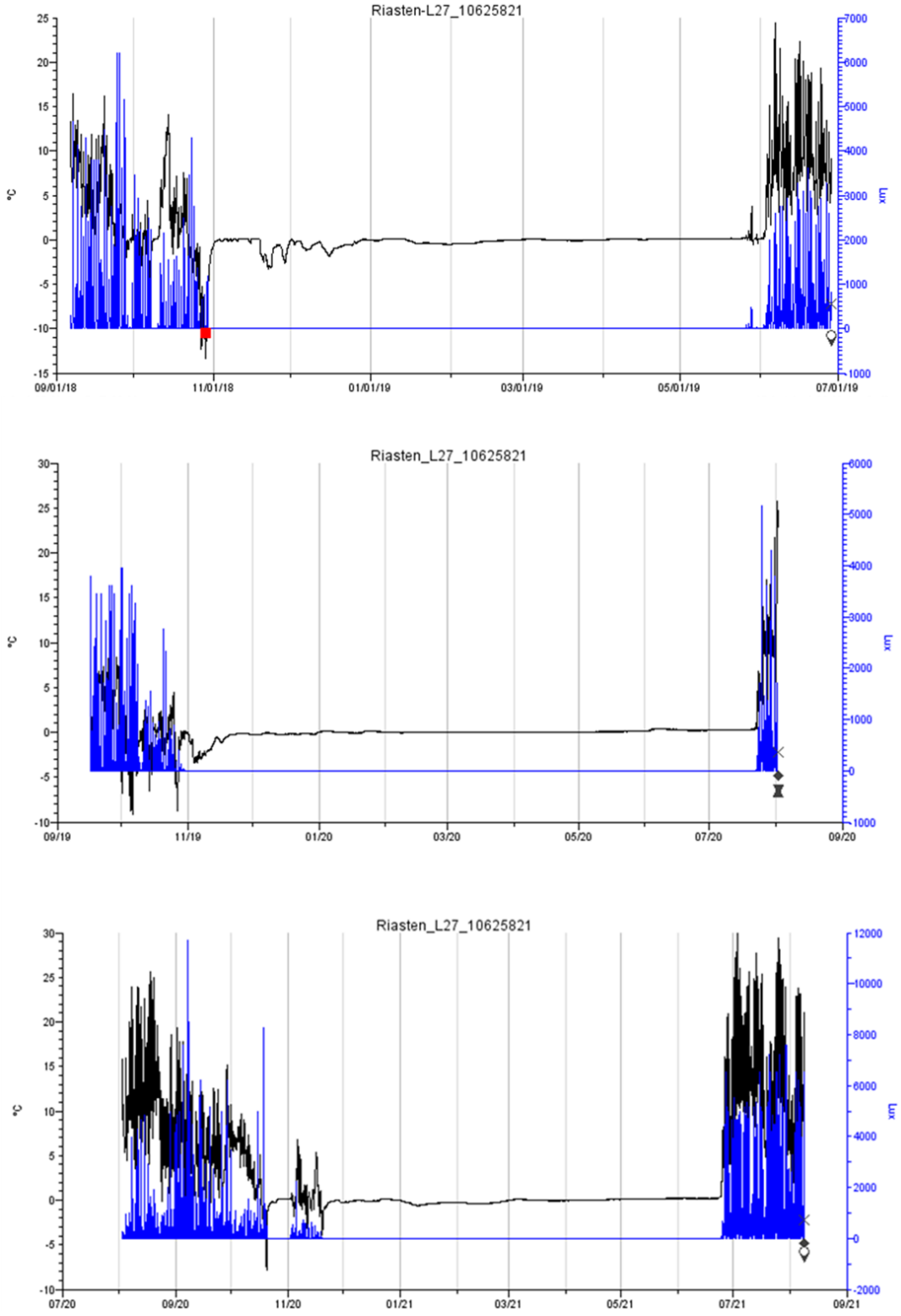

Figure C.2. Climate data of temperature and light intensity for plot 1 from 2018 to 2021 at Norwegian field site. The Celsius temperature and light intensity (lux) are shown on the black and blue line graph, respectively.

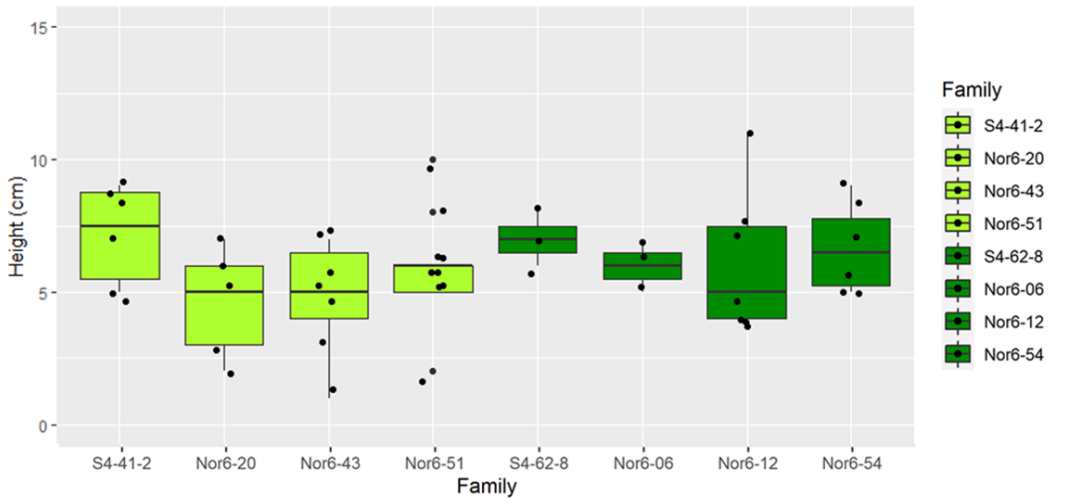

Figure C.3. Plant height in plot 1 for the 2021 data according to family. Plant height was measured for individual flowering rosette from the base of that rosette to the tip of the flowering shoot. If there are more than one flowering rosettes, the mean height was used for a single plant. The relationship between families and the plant height is not statistically clear (p-value = 0.42, Kruskal-Wallis test).

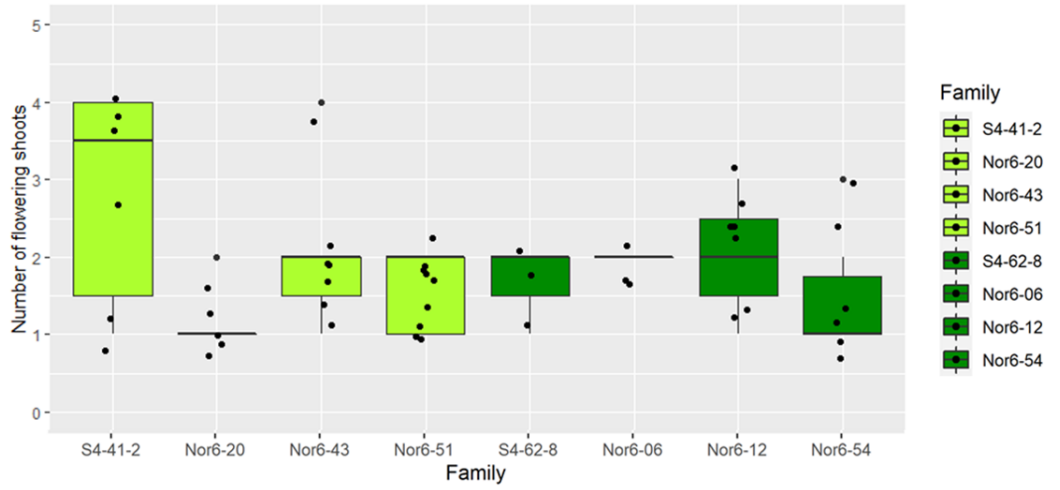

Figure C.4. Number of flowering shoots in plot 1 for the 2021 data according to family. The relationship between families and the number of flowering shoots is not statistically clear (p-value = 0.25, Kruskal-Wallis test).

Bibliography

Abe, M., Kobayashi, Y., Yamamoto, S., Daimon, Y., Yamaguchi, A., Ikeda, Y., Ichinoki, H., Notaguchi, M., Goto, K., and Araki, T. (2005). Fd, a bzip protein mediating signals from the floral pathway integrator ft at the shoot apex. *Science*, 309(5737):1052–1056.

Ågren, J., Oakley, C. G., McKay, J. K., Lovell, J. T., and Schemske, D. W. (2013). Genetic mapping of adaptation reveals fitness tradeoffs in arabidopsis thaliana. *Proceedings of the National Academy of Sciences*, 110(52):21077–21082.

Ågren, J. and Schemske, D. W. (2012). Reciprocal transplants demonstrate strong adaptive differentiation of the model organism arabidopsis thaliana in its native range. *New Phytologist*, 194(4):1112–1122.

Aikawa, S., Kobayashi, M. J., Satake, A., Shimizu, K. K., and 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(25):11632–11637.

Akaike, H. (1978). On the likelihood of a time series model. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 27(3-4):217–235.

Akman, M., Carlson, J. E., Holsinger, K. E., and Latimer, A. M. (2016). Transcriptome sequencing reveals population differentiation in gene expression linked to functional traits and environmental gradients in the south african shrub protea repens. *New Phytologist*, 210(1):295–309.

Albani, M. C., Castaings, L., Wötzel, S., Mateos, J. L., Wunder, J., Wang, R., Reymond, M., and 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(12):e1003130.

Albani, M. C. and Coupland, G. (2010). Comparative analysis of flowering in annual and perennial plants. *Current topics in developmental biology*, 91:323–348.

Alonge, M., Wang, X., Benoit, M., Soyk, S., Pereira, L., Zhang, L., Suresh, H., Ramakrishnan, S., Maumus, F., Ciren, D., et al. (2020). Major impacts of widespread structural variation on gene expression and crop improvement in tomato. *Cell*, 182(1):145–161.

Amasino, R. (2010). Seasonal and developmental timing of flowering. *The Plant Journal*, 61(6):1001–1013.

An, H., Roussot, C., Suárez-López, P., Corbesier, L., Vincent, C., Piñeiro, M., Hepworth, S., Mouradov, A., Justin, S., Turnbull, C., et al. (2004). Constans acts in the phloem to regulate a systemic signal that induces photoperiodic flowering of arabidopsis.

Anders, S., Pyl, P. T., and Huber, W. (2015). Htseq—a python framework to work with high-throughput sequencing data. *bioinformatics*, 31(2):166–169.

Anderson, J. T., Inouye, D. W., McKinney, A. M., Colautti, R. I., and Mitchell-Olds, T. (2012). Phenotypic plasticity and adaptive evolution contribute to advancing flowering phenology in response to climate change. *Proceedings of the Royal Society B: Biological Sciences*, 279(1743):3843–3852.

Anderson, J. T., Sparks, J. P., and Geber, M. A. (2010). Phenotypic plasticity despite source–sink population dynamics in a long-lived perennial plant. *New Phytologist*, 188(3):856–867.

Anderson, J. T., Willis, J. H., and Mitchell-Olds, T. (2011). Evolutionary genetics of plant adaptation. *Trends in Genetics*, 27(7):258–266.

Andrello, M., de Villemereuil, P., Carboni, M., Busson, D., Fortin, M.-J., Gaggiotti, O. E., and Till-Bottraud, I. (2020). Accounting for stochasticity in demographic compensation along the elevational range of an alpine plant. *Ecology letters*, 23(5):870–880.

Andrés, F. and Coupland, G. (2012). The genetic basis of flowering responses to seasonal cues. *Nature Reviews Genetics*, 13(9):627–639.

Austen, E. J., Rowe, L., Stinchcombe, J. R., and Forrest, J. R. (2017). Explaining the apparent paradox of persistent selection for early flowering. *New Phytologist*, 215(3):929–934.

Baker, R. L. and Diggle, P. K. (2011). Node-specific branching and heterochronic changes underlie population-level differences in mimulus guttatus (phrymaceae) shoot architecture. *American Journal of Botany*, 98(12):1924–1934.

Bastow, R., Mylne, J. S., Lister, C., Lippman, Z., Martienssen, R. A., and Dean, C. (2004). Vernalization requires epigenetic silencing of flc by histone methylation. *Nature*, 427(6970):164–167.

Baythavong, B. S. and Stanton, M. L. (2010). Characterizing selection on phenotypic plasticity in response to natural environmental heterogeneity. *Evolution: International Journal of Organic Evolution*, 64(10):2904–2920.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, 57(1):289–300.

Bergonzi, S. and Albani, M. C. (2011). Reproductive competence from an annual and a perennial perspective. *Journal of experimental botany*, 62(13):4415–4422.

Bergonzi, S., Albani, M. C., van Themaat, E. V. L., Nordström, K. J., Wang, R., Schneeberger, K., Moerland, P. D., and Coupland, G. (2013). Mechanisms of age-dependent response to winter temperature in perennial flowering of arabis alpina. *Science*, 340(6136):1094–1097.

Biere, A. (1995). Genotypic and plastic variation in plant size: effects on fecundity and allocation patterns in lychnis flos-cuculi along a gradient of natural soil fertility. *Journal of Ecology*, pages 629–642.

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics*, 30(15):2114–2120.

Bolmgren, K. and D. Cowan, P. (2008). Time–size tradeoffs: A phylogenetic comparative study of flowering time, plant height and seed mass in a north-temperate flora. *Oikos*, 117(3):424–429.

Bohlenius, H., Huang, T., Charbonnel-Campaa, L., Brunner, A. M., Jansson, S., Strauss, S. H., and Nilsson, O. (2006). Co/ft regulatory module controls timing of flowering and seasonal growth cessation in trees. *Science*, 312(5776):1040–1043.

Castaings, L., Bergonzi, S., Albani, M. C., Kemi, U., Savolainen, O., and Coupland, G. (2014). Evolutionary conservation of cold-induced antisense rnas of flowering locus c in arabidopsis thaliana perennial relatives. *Nature Communications*, 5(1):1–9.

Chen, D., Yan, W., Fu, L.-Y., and Kaufmann, K. (2018). Architecture of gene regulatory networks controlling flower development in arabidopsis thaliana. *Nature communications*, 9(1):1–13.

Clausen, J., Keck, D. D., Hiesey, W. M., et al. (1940). Experimental studies on the nature of species. i. effect of varied environments on western north american plants. *Experimental studies on the nature of species. I. Effect of varied environments on western North American plants.*, (520).

Corbesier, L., Vincent, C., Jang, S., Fornara, F., Fan, Q., Searle, I., Giakountis, A., Farrona, S., Gissot, L., Turnbull, C., et al. (2007). Ft protein movement contributes to long-distance signaling in floral induction of arabidopsis. *science*, 316(5827):1030–1033.

de Villemereuil, P., Gaggiotti, O. E., Mouterde, M., and Till-Bottraud, I. (2016). Common garden experiments in the genomic era: new perspectives and opportunities. *Heredity*, 116(3):249–254.

Delph, L. F. and Meagher, T. R. (1995). Sexual dimorphism masks life history trade-offs in the dioecious plant silene latifolia. *Ecology*, 76(3):775–785.

Des Marais, D. L., Hernandez, K. M., and Juenger, T. E. (2013). Genotype-by-environment interaction and plasticity: exploring genomic responses of plants to the abiotic environment. *Annual Review of Ecology, Evolution, and Systematics*, 44:5–29.

Doukhanina, E. V., Chen, S., van der Zalm, E., Godzik, A., Reed, J., and Dickman, M. B. (2006). Identification and functional characterization of the bag protein

family in arabidopsis thaliana. *Journal of Biological Chemistry*, 281(27):18793–18801.

Du, G. and Qi, W. (2010). Trade-offs between flowering time, plant height, and seed size within and across 11 communities of a qinghai-tibetan flora. *Plant Ecology*, 209(2):321–333.

Dunn, O. J. (1961). Multiple comparisons among means. *Journal of the American statistical association*, 56(293):52–64.

Eagles, H. A., Cane, K., and Vallance, N. (2009). The flow of alleles of important photoperiod and vernalisation genes through australian wheat. *Crop and Pasture Science*, 60(7):646–657.

Fjellheim, S., Boden, S., and Trevaskis, B. (2014). The role of seasonal flowering responses in adaptation of grasses to temperate climates. *Frontiers in plant science*, 5:431.

Forrest, J. R. (2014). Plant size, sexual selection, and the evolution of protandry in dioecious plants. *The American Naturalist*, 184(3):338–351.

Friedman, J. (2014). Genetic determinants and epistasis for life history trait differences in the common monkeyflower, mimulus guttatus. *Journal of Heredity*, 105(6):910–921.

Friedman, J. (2020). The evolution of annual and perennial plant life histories: ecological correlates and genetic mechanisms. *Annual Review of Ecology, Evolution, and Systematics*, 51:461–481.

Friedman, J., Twyford, A. D., Willis, J. H., and Blackman, B. K. (2015). The extent and genetic basis of phenotypic divergence in life history traits in mimulus guttatus. *Molecular Ecology*, 24(1):111–122.

Friedman, J. and Willis, J. H. (2013). Major qtl s for critical photoperiod and vernalization underlie extensive variation in flowering in the m imulus guttatus species complex. *New Phytologist*, 199(2):571–583.

Goel, M., Sun, H., Jiao, W.-B., and Schneeberger, K. (2019). Syri: finding genomic rearrangements and local sequence differences from whole-genome assemblies. *Genome biology*, 20(1):1–13.

Gould, B. A., Chen, Y., and Lowry, D. B. (2017). Pooled ecotype sequencing reveals candidate genetic mechanisms for adaptive differentiation and reproductive isolation. *Molecular ecology*, 26(1):163–177.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., et al. (2011). Full-length transcriptome assembly from rna-seq data without a reference genome. *Nature biotechnology*, 29(7):644–652.

Grattapaglia, D., Plomion, C., Kirst, M., and Sederoff, R. R. (2009). Genomics of growth traits in forest trees. *Current opinion in plant biology*, 12(2):148–156.

Gugger, P. F., Peñaloza-Ramírez, J. M., Wright, J. W., and Sork, V. L. (2017). Whole-transcriptome response to water stress in a california endemic oak, quercus lobata. *Tree physiology*, 37(5):632–644.

Hartmann, U., Höhmann, S., Nettesheim, K., Wisman, E., Saedler, H., and Huijser, P. (2000). Molecular cloning of svp: a negative regulator of the floral transition in arabidopsis. *The Plant Journal*, 21(4):351–360.

Hereford, J. (2009). A quantitative survey of local adaptation and fitness trade-offs. *The American Naturalist*, 173(5):579–588.

Horvath, S. (2011). *Weighted network analysis: applications in genomics and systems biology*. Springer Science & Business Media.

Hsu, C.-Y., Liu, Y., Luthe, D. S., and Yuceer, C. (2006). Poplar ft2 shortens the juvenile phase and promotes seasonal flowering. *The Plant Cell*, 18(8):1846–1861.

Huang, M.-D., Chen, T.-L. L., and Huang, A. H. (2013). Abundant type iii lipid transfer proteins in arabidopsis tapetum are secreted to the locule and become a constituent of the pollen exine. *Plant physiology*, 163(3):1218–1229.

Hughes, P. W. (2017). Between semelparity and iteroparity: empirical evidence for a continuum of modes of parity. *Ecology and evolution*, 7(20):8232–8261.

Hughes, P. W., Soppe, W. J., and Albani, M. C. (2019). Seed traits are pleiotropically regulated by the flowering time gene perpetual flowering 1 (pep1) in the perennial arabis alpina. *Molecular ecology*, 28(5):1183–1201.

Hyun, Y., Richter, R., and Coupland, G. (2017). Competence to flower: age-controlled sensitivity to environmental cues. *Plant physiology*, 173(1):36–46.

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

Ingvarsson, P. K. and Street, N. R. (2011). Association genetics of complex traits in plants. *New Phytologist*, 189(4):909–922.

Jaeger, K. E. and Wigge, P. A. (2007). Ft protein acts as a long-range signal in arabidopsis. *Current biology*, 17(12):1050–1054.

Jensen, C. S., Salchert, K., and Nielsen, K. K. (2001). A terminal flower1-like gene from perennial ryegrass involved in floral transition and axillary meristem identity. *Plant Physiology*, 125(3):1517–1528.

Johanson, U., West, J., Lister, C., Michaels, S., Amasino, R., and Dean, C. (2000). Molecular analysis of frigida, a major determinant of natural variation in arabidopsis flowering time. *Science*, 290(5490):344–347.

Kawecki, T. J. and Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology letters*, 7(12):1225–1241.

Kemi, U., Leinonen, P. H., Savolainen, O., and Kuittinen, H. (2019). Inflorescence shoot elongation, but not flower primordia formation, is photoperiodically regulated in arabidopsis lyrata. *Annals of botany*, 124(1):91–102.

Kemi, U., Niittyvuopio, A., Toivainen, T., Pasanen, A., Quilot-Turion, B., Holm, K., Lagercrantz, U., Savolainen, O., and Kuittinen, H. (2013). Role of vernalization and of duplicated flowering locus c in the perennial arabidopsis lyrata. *New Phytologist*, 197(1):323–335.

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, 30(1/2):81–93.

Kiefer, C., Severing, E., Karl, R., Bergonzi, S., Koch, M., Tresch, A., and Coupland, G. (2017). Divergence of annual and perennial species in the brassicaceae

and the contribution of cis-acting variation at flc orthologues. *Molecular Ecology*, 26(13):3437–3457.

Kiełbasa, S. M., Wan, R., Sato, K., Horton, P., and Frith, M. C. (2011). Adaptive seeds tame genomic sequence comparison. *Genome research*, 21(3):487–493.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S. L. (2013). Tophat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome biology*, 14(4):1–13.

Kim, E. and Donohue, K. (2011). Population differentiation and plasticity in vegetative ontogeny: Effects on life-history expression in erysimum capitatum (brassicaceae). *American Journal of Botany*, 98(11):1752–1761.

Kinoshita, A. and Richter, R. (2020). Genetic and molecular basis of floral induction in arabidopsis thaliana. *Journal of Experimental Botany*, 71(9):2490–2504.

Kirkendall, L. R. and Chr, N. (1985). On defining" breeding once". *The American Naturalist*, 125(2):189–204.

Kivimäki, M., Kärkkäinen, K., Gaudeul, M., Løe, G., and Ågren, J. (2007). Gene, phenotype and function: Glabrous1 and resistance to herbivory in natural populations of arabidopsis lyrata. *Molecular Ecology*, 16(2):453–462.

Kotoda, N., Iwanami, H., Takahashi, S., and Abe, K. (2006). Antisense expression of mdtfl1, a tfl1-like gene, reduces the juvenile phase in apple. *Journal of the American Society for Horticultural Science*, 131(1):74–81.

Kruskal, W. H. and Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association*, 47(260):583–621.

Langfelder, P. and Horvath, S. (2008). Wgcna: an r package for weighted correlation network analysis. *BMC bioinformatics*, 9(1):1–13.

Langlet, O. (1971). Two hundred years genecology. *Taxon*, pages 653–721.

Lasky, J. R., Des Marais, D. L., Lowry, D. B., Povolotskaya, I., McKay, J. K., Richards, J. H., Keitt, T. H., and Juenger, T. E. (2014). Natural variation in abiotic stress responsive gene expression and local adaptation to climate in arabidopsis thaliana. *Molecular biology and evolution*, 31(9):2283–2296.

Lazaro, A., Obeng-Hinneh, E., and Albani, M. C. (2018). Extended vernalization regulates inflorescence fate in arabis alpina by stably silencing perpetual flowering1. *Plant physiology*, 176(4):2819–2833.

Leimu, R. and Fischer, M. (2008). A meta-analysis of local adaptation in plants. *PloS one*, 3(12):e4010.

Leinonen, P. H., Remington, D. L., Leppälä, J., and Savolainen, O. (2013). Genetic basis of local adaptation and flowering time variation in arabidopsis lyrata. *Molecular ecology*, 22(3):709–723.

Less, H. and Galili, G. (2008). Principal transcriptional programs regulating plant amino acid metabolism in response to abiotic stresses. *Plant physiology*, 147(1):316–330.

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for rna-seq data with deseq2. *Genome biology*, 15(12):1–21.

Lowry, D. B., Popovic, D., Brennan, D. J., and Holeski, L. M. (2019). Mechanisms of a locally adaptive shift in allocation among growth, reproduction, and herbivore resistance in mimulus guttatus. *Evolution*, 73(6):1168–1181.

Maan, M. E. and Sefc, K. M. (2013). Colour variation in cichlid fish: developmental mechanisms, selective pressures and evolutionary consequences. In *Seminars in Cell & Developmental Biology*, volume 24, pages 516–528. Elsevier.

Madrid, E., Chandler, J. W., and Coupland, G. (2021a). Gene regulatory networks controlled by flowering locus c that confer variation in seasonal flowering and life history. *Journal of Experimental Botany*, 72(1):4–14.

Madrid, E., Severing, E., de Ansorena, E., Kiefer, C., Brand, L., Martinez-Gallegos, R., Woetzel, S., Kemi, U., Jiao, W.-B., Schneeberger, K., et al. (2021b). Transposition and duplication of mads-domain transcription factor genes in annual and perennial arabis species modulates flowering. *Proceedings of the National Academy of Sciences*, 118(39).

Mann, H. B. and Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, pages 50–60.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. journal*, 17(1):10–12.

Mateos, J. L., Madrigal, P., Tsuda, K., Rawat, V., Richter, R., Romera-Branchat, M., Fornara, F., Schneeberger, K., Krajewski, P., and Coupland, G. (2015). Combinatorial activities of short vegetative phase and flowering locus c define distinct modes of flowering regulation in arabidopsis. *Genome biology*, 16(1):1–23.

Mathieu, J., Warthmann, N., Küttner, F., and Schmid, M. (2007). Export of ft protein from phloem companion cells is sufficient for floral induction in arabidopsis. *Current biology*, 17(12):1055–1060.

Matyas, C. (1996). Climatic adaptation of trees: rediscovering provenance tests. *Euphytica*, 92(1):45–54.

McKown, A. D., Guy, R. D., Klápště, J., Geraldes, A., Friedmann, M., Cronk, Q. C., El-Kassaby, Y. A., Mansfield, S. D., and Douglas, C. J. (2014). Geographical and environmental gradients shape phenotypic trait variation and genetic structure in p opulus trichocarpa. *New Phytologist*, 201(4):1263–1276.

Melchers, G. (1937). *Die Wirkung von Genen, tiefen Temperaturen und blühenden Pfropfpartnern auf die Blühreife von Hyoscyamus niger L.*

Méndez-Vigo, B., Picó, F. X., Ramiro, M., Martínez-Zapater, J. M., and Alonso-Blanco, C. (2011). Altitudinal and climatic adaptation is mediated by flowering traits and fri, flc, and phyc genes in arabidopsis. *Plant physiology*, 157(4):1942–1955.

Michaels, S. D. and Amasino, R. M. (2000). Memories of winter: vernalization and the competence to flower. *Plant, Cell & Environment*, 23(11):1145–1153.

Mohamed, R., Wang, C.-T., Ma, C., Shevchenko, O., Dye, S. J., Puzey, J. R., Etherington, E., Sheng, X., Meilan, R., Strauss, S. H., et al. (2010). Populus cen/tfl1 regulates first onset of flowering, axillary meristem identity and dormancy release in populus. *The Plant Journal*, 62(4):674–688.

Morea, E. G. O., da Silva, E. M., Valente, G. T., Barrera Rojas, C. H., Vincentz, M., Nogueira, F. T. S., et al. (2016). Functional and evolutionary analyses of the mir156 and mir529 families in land plants. *BMC plant biology*, 16(1):1–13.

Oakley, C. G., Ågren, J., Atchison, R. A., and Schemske, D. W. (2014). Qtl mapping of freezing tolerance: links to fitness and adaptive trade-offs. *Molecular Ecology*, 23(17):4304–4315.

Obeso, J. R. (2002). The costs of reproduction in plants. *New phytologist*, 155(3):321–348.

O'Neill, C. M., Lu, X., Calderwood, A., Tudor, E. H., Robinson, P., Wells, R., Morris, R., and Penfield, S. (2019). Vernalization and floral transition in autumn drive winter annual life history in oilseed rape. *Current Biology*, 29(24):4300–4306.

Pettitt, A. N. (1976). A two-sample anderson-darling rank statistic. *Biometrika*, 63(1):161–168.

Ponraj, U. and Theres, K. (2020). Keep a distance to be different: axillary buds initiating at a distance from the shoot apical meristem are crucial for the perennial lifestyle of arabis alpina. *New Phytologist*, 227(1):116–131.

Postma, F. M. and Ågren, J. (2016). Early life stages contribute strongly to local adaptation in arabidopsis thaliana. *Proceedings of the National Academy of Sciences*, 113(27):7590–7595.

Qu, X., Zhao, Z., and Tian, Z. (2017). Erecta regulates cell elongation by activating auxin biosynthesis in arabidopsis thaliana. *Frontiers in plant science*, 8:1688.

Raunkiaer, C. et al. (1934). The life forms of plants and statistical plant geography; being the collected papers of c. raunkiaer. *The life forms of plants and statistical plant geography; being the collected papers of C. Raunkiaer.*

Remington, D. L., Leinonen, P. H., Leppälä, J., and Savolainen, O. (2013). Complex genetic effects on early vegetative development shape resource allocation differences between arabidopsis lyrata populations. *Genetics*, 195(3):1087–1102.

Rohde, A. and Bhalerao, R. P. (2007). Plant dormancy in the perennial context. *Trends in plant science*, 12(5):217–223.

Romera-Branchat, M., Severing, E., Pocard, C., Ohr, H., Vincent, C., Née, G.,

Martinez-Gallegos, R., Jang, S., Andrés, F., Madrigal, P., et al. (2020). Functional divergence of the arabidopsis florigen-interacting bzip transcription factors fd and fdp. *Cell Reports*, 31(9):107717.

Rubin, M. J., Schmid, K. M., and Friedman, J. (2019). Assortative mating by flowering time and its effect on correlated traits in variable environments. *Ecology and Evolution*, 9(1):471–481.

Sarnowska, E., Kubala, S., Cwiek, P., Sacharowski, S., Oksinska, P., Steciuk, J., Zaborowska, M., Szurmak, J. M., Dubianski, R., Maassen, A., et al. (2021). A non-canonical function of arabidopsis erecta proteins in gibberellin signaling. *bioRxiv*.

Savolainen, O., Lascoux, M., and Merilä, J. (2013). Ecological genomics of local adaptation. *Nature Reviews Genetics*, 14(11):807–820.

Savolainen, O., Pyhäjärvi, T., and Knürr, T. (2007). Gene flow and local adaptation in trees. *Annu. Rev. Ecol. Evol. Syst.*, 38:595–619.

Schein, M., Yang, Z., Mitchell-Olds, T., and Schmid, K. J. (2004). Rapid evolution of a pollen-specific oleosin-like gene family from arabidopsis thaliana and closely related species. *Molecular Biology and Evolution*, 21(4):659–669.

Shea, D. J., Nishida, N., Takada, S., Itabashi, E., Takahashi, S., Akter, A., Miyaji, N., Osabe, K., Mehraj, H., Shimizu, M., et al. (2019). Long noncoding rnas in brassica rapa l. following vernalization. *Scientific reports*, 9(1):1–11.

Silvertown, J. (1986). " biennials": A reply to kelly. *The American Naturalist*, 127(5):721–724.

Silvertown, J. (1989). binary classification of plant life history and some possibilities for its evolutionary application. *Evolutionary trends in plants*.

Sork, V., Aitken, S., Dyer, R., Eckert, A., Legendre, P., and Neale, D. (2013). Putting the landscape into the genomics of trees: approaches for understanding local adaptation and population responses to changing climate. *Tree Genetics & Genomes*, 9(4):901–911.

Sork, V. L. (2018). Genomic studies of local adaptation in natural plant populations. *Journal of Heredity*, 109(1):3–15.

Stinchcombe, J. R., Weinig, C., Ungerer, M., Olsen, K. M., Mays, C., Halldorsdottir, S. S., Purugganan, M. D., and Schmitt, J. (2004). A latitudinal cline in flowering time in arabidopsis thaliana modulated by the flowering time gene frigida. *Proceedings of the National Academy of Sciences*, 101(13):4712–4717.

Suárez-López, P., Wheatley, K., Robson, F., Onouchi, H., Valverde, F., and Coupland, G. (2001). Constans mediates between the circadian clock and the control of flowering in arabidopsis. *Nature*, 410(6832):1116–1120.

Sung, S. and Amasino, R. M. (2004). Vernalization in arabidopsis thaliana is mediated by the phd finger protein vin3. *Nature*, 427(6970):159–164.

Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011). Revigo summarizes and visualizes long lists of gene ontology terms. *PloS one*, 6(7):e21800.

Taylor, M. A., Wilczek, A. M., Roe, J. L., Welch, S. M., Runcie, D. E., Cooper, M. D., and Schmitt, J. (2019). Large-effect flowering time mutations reveal conditionally adaptive paths through fitness landscapes in arabidopsis thaliana. *Proceedings of the National Academy of Sciences*, 116(36):17890–17899.

Tenreira, T., Lange, M. J. P., Lange, T., Bres, C., Labadie, M., Monfort, A., Hernould, M., Rothan, C., and Denoyes, B. (2017). A specific gibberellin 20-oxidase dictates the flowering-runnering decision in diploid strawberry. *The Plant Cell*, 29(9):2168–2182.

Thomas, S. G. and Sun, T.-p. (2004). Update on gibberellin signaling. a tale of the tall and the short. *Plant physiology*, 135(2):668–676.

Tilmes, V., Mateos, J. L., Madrid, E., Vincent, C., Severing, E., Carrera, E., López-Díaz, I., and Coupland, G. (2019). Gibberellins act downstream of arabis perpetual flowering1 to accelerate floral induction during vernalization. *Plant physiology*, 180(3):1549–1563.

Toräng, P., Wunder, J., Obeso, J. R., Herzog, M., Coupland, G., and Ågren, J. (2015). Large-scale adaptive differentiation in the alpine perennial herb a rabis alpina. *New Phytologist*, 206(1):459–470.

Turchin, M. C., Chiang, C. W., Palmer, C. D., Sankararaman, S., Reich, D., and Hirschhorn, J. N. (2012). Evidence of widespread selection on standing variation in europe at height-associated snps. *Nature genetics*, 44(9):1015–1019.

Turck, F., Fornara, F., and Coupland, G. (2008). Regulation and identity of florigen: Flowering locus t moves center stage. *Annu. Rev. Plant Biol.*, 59:573–594.

Turesson, G. (1922). The species and the variety as ecological units. *Hereditas*, 3(1):100–113.

Valverde, F., Mouradov, A., Soppe, W., Ravenscroft, D., Samach, A., and Coupland, G. (2004). Photoreceptor regulation of constans protein in photoperiodic flowering. *Science*, 303(5660):1003–1006.

Van Ooijen, J. (2006). Joinmap® 4, software for the calculation of genetic linkage maps in experimental populations. *Kyazma BV, Wageningen*, 33(10.1371).

Van Ooijen, J. (2009). Mapqtl® 6, software for the mapping of quantitative trait loci in experimental populations of diploid species. *Kyazma BV, Wageningen, Netherlands*, 64.

Van Ooijen, J. W. (1999). Lod significance thresholds for qtl analysis in experimental populations of diploid species. *Heredity*, 83(5):613–624.

Vayssières, A., Mishra, P., Roggen, A., Neumann, U., Ljung, K., and Albani, M. C. (2020). Vernalization shapes shoot architecture and ensures the maintenance of dormant buds in the perennial arabis alpina. *New Phytologist*, 227(1):99–115.

Vile, D., Shipley, B., and Garnier, E. (2006). A structural equation model to integrate changes in functional strategies during old-field succession. *Ecology*, 87(2):504–517.

Wadgymar, S. M., Daws, S. C., and Anderson, J. T. (2017). Integrating viability and fecundity selection to illuminate the adaptive nature of genetic clines. *Evolution Letters*, 1(1):26–39.

Wang, R., Albani, M. C., Vincent, C., Bergonzi, S., Luan, M., Bai, Y., Kiefer, C., Castillo, R., and Coupland, G. (2011). Aa tfl1 confers an age-dependent response to vernalization in perennial arabis alpina. *The Plant Cell*, 23(4):1307–1321.

Wang, R., Farrona, S., Vincent, C., Joecker, A., Schoof, H., Turck, F., Alonso-Blanco, C., Coupland, G., and Albani, M. C. (2009a). Pep1 regulates perennial flowering in arabis alpina. *Nature*, 459(7245):423–427.

Wang, Y., Severing, E. I., Koornneef, M., and Aarts, M. G. (2020). Flc and svp are key regulators of flowering time in the biennial/perennial species noccaea caerulescens. *Frontiers in plant science*, 11.

Wang, Z., Gerstein, M., and Snyder, M. (2009b). Rna-seq: a revolutionary tool for transcriptomics. *Nature reviews genetics*, 10(1):57–63.

Wellmer, F., Riechmann, J. L., Alves-Ferreira, M., and Meyerowitz, E. M. (2004). Genome-wide analysis of spatial gene expression in arabidopsis flowers. *The Plant Cell*, 16(5):1314–1326.

Wigge, P. A., Kim, M. C., Jaeger, K. E., Busch, W., Schmid, M., Lohmann, J. U., and Weigel, D. (2005). Integration of spatial and temporal information during floral induction in arabidopsis. *Science*, 309(5737):1056–1059.

Williams, E. J. and Bowles, D. J. (2004). Coexpression of neighboring genes in the genome of arabidopsis thaliana. *Genome research*, 14(6):1060–1067.

Wötzel, S., Andrello, M., Albani, M. C., Koch, M. A., Coupland, G., and Gugerli, F. (2022). Arabis alpina: A perennial model plant for ecological genomics and life-history evolution. *Molecular Ecology Resources*, 22(2):468–486.

Wright, S. (1931). Evolution in mendelian populations. *Genetics*, 16(2):97.

Yamagishi, N., Kume, K., Hikage, T., Takahashi, Y., Bidadi, H., Wakameda, K., Saitoh, Y., Yoshikawa, N., and Tsutsumi, K.-i. (2016). Identification and functional analysis of svp ortholog in herbaceous perennial plant gentiana triflora: implication for its multifunctional roles. *Plant Science*, 248:1–7.

Yamaguchi, A. and Abe, M. (2012). Regulation of reproductive development by non-coding rna in arabidopsis: to flower or not to flower. *Journal of plant research*, 125(6):693–704.

Yan, L., Loukoianov, A., Blechl, A., Tranquilli, G., Ramakrishna, W., SanMiguel, P., Bennetzen, J. L., Echenique, V., and Dubcovsky, J. (2004). The wheat vrn2 gene is a flowering repressor down-regulated by vernalization. *Science*, 303(5664):1640–1644.

Yang, S., Zhang, K., Zhu, H., Zhang, X., Yan, W., Xu, N., Liu, D., Hu, J., Wu, Y., Weng, Y., et al. (2020). Melon short internode (cmsi) encodes an erecta-like

receptor kinase regulating stem elongation through auxin signaling. *Horticulture research*, 7(1):1–14.

Yeaman, S., Hodgins, K. A., Lotterhos, K. E., Suren, H., Nadeau, S., Degner, J. C., Nurkowski, K. A., Smets, P., Wang, T., Gray, L. K., et al. (2016). Convergent local adaptation to climate in distantly related conifers. *Science*, 353(6306):1431–1433.

Young, T. P. and Augspurger, C. K. (1991). Ecology and evolution of long-lived semelparous plants. *Trends in Ecology & Evolution*, 6(9):285–289.

Yuan, Y., Bayer, P. E., Batley, J., and Edwards, D. (2021). Current status of structural variation studies in plants. *Plant Biotechnology Journal*, 19(11):2153–2163.

Zhang, Z., Mao, L., Chen, H., Bu, F., Li, G., Sun, J., Li, S., Sun, H., Jiao, C., Blakely, R., et al. (2015). Genome-wide mapping of structural variations reveals a copy number variant that determines reproductive morphology in cucumber. *The Plant Cell*, 27(6):1595–1604.

Zhu, P., Lister, C., and Dean, C. (2021). Cold-induced arabidopsis frigida nuclear condensates for flc repression. *Nature*, 599(7886):657–661.

Erklärung zur Dissertation
gemäß der Promotionsordnung vom 12. März 2020

Diese Erklärung muss in der Dissertation enthalten sein.
(*This version must be included in the doctoral thesis*)

„Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind als solche kenntlich gemacht. Ich versichere an Eides statt, dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen und eingebundenen Artikeln und Manuskripten - noch nicht veröffentlicht worden ist sowie, dass ich eine Veröffentlichung der Dissertation vor Abschluss der Promotion nicht ohne Genehmigung des Promotionsausschusses vornehmen werde. Die Bestimmungen dieser Ordnung sind mir bekannt. Darüber hinaus erkläre ich hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten der Universität zu Köln gelesen und sie bei der Durchführung der Dissertation zugrundeliegenden Arbeiten und der schriftlich verfassten Dissertation beachtet habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht."

Teilpublikationen:

04.04.2022

Datum, Name und Unterschrift

27.02.2022 Yi-Chen Lin yi-chen Lin

Erklärung zum Gesuch um Zulassung zur Promotion

gemäß der Promotionsordnung vom 12. März 2020

1. Zugänglichkeit von Daten und Materialien

Die Dissertation beinhaltet die Gewinnung von Primärdaten oder die Analyse solcher Daten oder die Reproduzierbarkeit der in der Dissertation dargestellten Ergebnisse setzt die Verfügbarkeit von Datenanalysen, Versuchsprotokollen oder Probenmaterial voraus.

☐ Trifft nicht zu

☒ Trifft zu.
In der Dissertation ist dargelegt wie diese Daten und Materialien gesichert und zugänglich sind (entsprechend den Vorgaben des Fachgebiets beziehungsweise der Betreuerin oder des Betreuers).

2. Frühere Promotionsverfahren

Ich habe bereits einen Doktortitel erworben oder ehrenhalber verliehen bekommen.

Oder: Für mich ist an einer anderen Fakultät oder Universität ein Promotionsverfahren eröffnet worden, aber noch nicht abgeschlossen.

Oder: Ich bin in einem Promotionsverfahren gescheitert.

☒ Trifft nicht zu

☐ Zutreffend
Erläuterung:

3. Straftat

Ich bin nicht zu einer vorsätzlichen Straftat verurteilt worden, bei deren Vorbereitung oder Begehung der Status einer Doktorandin oder eines Doktoranden missbraucht wurde.

Ich versichere, alle Angaben wahrheitsgemäß gemacht zu haben.

Datum	Name	Unterschrift
27.02.2022	Yi-Chen Lin	Yi-Chen Lin