Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Characterisation of the Genetic Mechanism Behind Floral Trichomes in *Brassica rapa*

A thesis submitted to the University of Dublin for the degree of

Doctor of Philosophy

Christina Fitzsimons

Thesis Supervisor: Professor Frank Wellmer

Smurfit Institute of Genetics

Trinity College Dublin

Dublin, Ireland

March 2023

Declaration

I declare that his thesis has not been submitted as an exercise for a degree at this or any other University and it is entirely my own work. I agree to deposit this thesis in the University's open-access institutional repository or allow the Library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

Christina Fitzsimons

Summary

In his seminal essay "The Metamorphosis of Plants", published in the 18th century, Johann Wolfgang von Goethe hypothesized that all above-ground organs of a plant, including floral organs, are derived from leaf-like structures. Over 200 years later, his ideas were confirmed when plants containing mutations for A-, B-, and C-class floral homeotic genes, or all functionally redundant E-class floral homeotic genes, were found to develop flowers that consisted of leaf-like organs with trichomes. Work carried out previously in our laboratory provided insight into the genetic mechanism underlying this metamorphosis, finding that the activities of AGAMOUS (a C-class floral homeotic protein) overlap with the gene network controlling trichome formation; AGAMOUS suppresses their formation on floral organs under normal developmental conditions. When a weak *ag* mutant allele was combined with mutant alleles of trichome repressor genes, *Arabidopsis* plants were generated with flowers that developed supernumerary trichomes on the floral organs.

Given that trichomes are known to confer protection from pests, it was thought that this mechanism of inducing trichomes to develop on the gynoecium and subsequent seedpod could provide an increased level of innate pest protection. The advantage of this approach is that the highest level of protection is offered to the fruit of the plant which is often the most agriculturally and economically important part of a plant. The main aim of the work described in this thesis was to translate these findings on the regulation of trichome initiation in *Arabidopsis* flowers, described by Ó'Maoiléidigh et al. (2018), to *Brassica rapa* in order to generate plants with supernumerary trichomes as a potential new avenue for crop protection.

Three key elements were necessary in order for this project to succeed. Firstly, a *B. rapa ag* mutant line that was as similar to the weak *Arabidopsis ag-10* allele used in Ó'Maoiléidigh et al. (2018) need to be obtained. Secondly, *B. rapa* lines containing strong mutations for key trichome repressors *TRIPTYCHON* and *CAPRICE* needed to be identified and crossed into the *ag* mutant line. Lastly, given that the widely used *B. rapa* line R-o-18 almost completely lacks trichomes, the mechanism of trichome development in *B. rapa* needed to be investigated. A way to get the plants to produce trichomes successfully was required in order for there to be a chance of trichomes developing on the gynoecium of the lines containing mutations in *AG*, *CPC* and *TRY*. This work was complicated by the occurrence of a whole genome triplication event in the *Brassica* species after its divergence from *Arabidopsis* - meaning there were likely more than one homolog of these genes present in *B. rapa*

The first part of this thesis focused on identifying the homologs of *AG* in *B. rapa* and finding suitable mutant lines. Two homologs were identified - *BrAG1* and *BrAG2* - using a variety of techniques such as sequence identity comparison, sequence alignment and examination of synteny. The genes were then characterised using numerous methods such as RT-qPCR, 3'RACE, complementation experiments and investigating transcription factor binding sites. Expression data indicated that *BrAG1* was the critical homolog to target the function of in this work as it was the with the highest level of expression in gynoecium tissue. *B. rapa brag* mutant alleles were obtained from a commercial supplier, characterised and it was determined that the best alleles for the purposes of this project were the *BrAG1* mutant alleles *brag1-1* and *brag1-2* as they were the closest to the *ag-10* allele used in Ó'Maoiléidigh et al. (2018).

The second part of this thesis focused on identifying the *B. rapa* homologs of *CPC* and *TRY* and employed similar experimental methods used for *AG* as described above. There were two homologs identified for *CPC* - *BrCPC1* and *BrCPC2* and three were identified for *TRY* - *BrTRY1*, *BrTRY2* and *BrTRY3*. RT-qPCR data revealed that *BrTRY2* and *BrCPC2* were the most highly expressed homologs in gynoecium tissue and thus were the essential homologs to knockout function of in order to allow for the formation of any potential trichomes on the female floral reproductive organs. Again, mutant alleles were obtained from a commercial supplier and characterised with the mutants *brtry2-1* and *brcpc2-1* identified as the best alleles to use.

The third part of this thesis examined why the R-o-18 line of *B. rapa* almost entirely lacks trichomes and it was found that this was due to a mutation in *BrGL1a*, a key component of the trichome activator complex, rendering the gene non-functional. It was discovered that the lack of trichomes could be rescued by crossing with another accession of *B. rapa* C634 which contains a functional copy of *BrGL1a* and develops trichomes. With all three elements present, an *ag-10* like *B. rapa* mutant, a strong mutant alleles for *BrTRY2* and a working trichome activator complex, a line of *B. rapa* that develops with trichomes on the gynoecia and with the potential for pest protection was produced. These results showed that research findings in *Arabidopsis* can be directly translated into the closely related species *B. rapa*.

The final part of this thesis examined the role of the floral meristem identity genes *LEAFY* (*LFY*) and *APETALA1* (*AP1*) in controlling the levels of the plant hormone auxin in the floral meristem of *Arabidopsis*. I showed through a series of experiments using an inducible LFY line, RNA sequencing and quantification of levels of the most common form of auxin, indole-3-acetic acid (IAA), that LFY and AP1 do not have a significant impact on the auxin levels in floral meristem thus challenging current models that suggest that LFY is key in downregulating auxin levels in the floral meristem.

For my grandfather, Christopher Fitzsimons.

Acknowledgments

I would like to thank Prof. Frank Wellmer for the opportunity to complete my Ph.D. in his laboratory and for all the guidance, patience, and advice he provided over the years. Many thanks to Dr Emmanuelle Graciet for all her advice and help with experimental troubleshooting. Thank you to Science Foundation Ireland and the Environmental Protection Agency for funding this work. Thanks should also go to all the technical and support staff in the department especially Brenda Campbell, Philip Campbell, and Dr Dave Noonan.

I would like to extend my sincere thanks to all the members of the Wellmer lab, past and present, who have helped me throughout my Ph.D. You all made the lab such a wonderful environment to work and learn in. Thank you to Kevin Goslin, Joseph Beegan, Andrea Finocchio, and Christos Velanis for their encouragement, advice and help over the years. To lab dad, Bennett Thomson, thank you for your endless patience and advice with the troubleshooting of experiments. To my work wives, Caoilfhionn Breathnach and Kate Horgan, thank you so much for the endless support, laughter, and always encouraging me to continue on.

I am also thankful to everyone else in the Smurfit Institute of Genetics especially Zoe, Emily and Orla who always had time for coffee or lunch to de-stress when work felt overwhelming. Thank you to my friends Beth, Laura, Stewart, and Snow who were always understanding and supportive throughout this journey. Thank you to my brothers, Gregory and Brendan, for being a source of laughter and levity over the last few years. I am eternally grateful to my mother for always believing in me. I would like to thank the most important person in my life, my partner Lucy Walsh - without her endless love, support, and patience none of this would have been possible.

Abbreviations

Abbr.	Term
AG	AGAMOUS
AGL-24	AGAMOUS-LIKE 24
AlcR	Ethanol Inducible Promoter
amiRNA	artificial microRNA
ANT	AINTEGUMENTA
AP1	APETALA 1
AP2	APETALA 2
AP3	APETALA 3
AR	Androgen Receptor
AREs	AU-Rich Elements
ARFs	AUXIN RESPONSE FACTORS
ARF5	AUXIN RESPONSE FACTOR 5
BAP	Cytokinin 6-benzylaminopurine
bp	basepair
CAL	CAULIFLOWER
CDS	Coding DNA Sequence
CO	CONSTANS
CLV1	CLAVATA 1
CLV2	CLAVATA 2
CLV3	CLAVATA 3
CPC	CAPRICE
CPL3	CAPRICE-LIKE MYB 3
CPR5	CONSTITUTIVE PATHOGEN RESPONSE 5
CRC	CRABS CLAW
CRN	CORYNE
dCAPS	Derived Cleaved Amplified Polymorphic Sequences
DAD1	DEFECTIVE IN ANTHER DEHISCENCE 1
DEX	Dexamethasone
DHT	Dihydrotestosterone
DMAS	Double-Mismatch Allele-Specific
ETC1	ENHANCER OF TRY AND CPC 1
ETC2	ENHANCER OF TRY AND CPC 2
ETC3	ENHANCER OF TRY AND CPC 3
EGL3	ENHANCER OF GLABRA 3

EMS	Ethyl Methanesulfonate
FLC	FLOWERING LOCUS C
FT	FLOWERING LOCUS T
FDH	FIDDLEHEAD
FIL	FILAMENTOUS FLOWER
FIS	Floral Induction System
FT	FLOWERING LOCUS T
FM	Floral Meristem
FUL	FRUITFULL
GA	Gibberellic acid
GI	GIGANTEA
GIS2	GLABROUS INFLORESCENCE STEMS 2
GL1	GLABRA1
GL2	GLABRA2
GL3	GLABRA3
GR	Glucocorticoid Receptor
GUS	β-Glucuronidase
H3K27me3	Histone 3 Lysine 27 Trimethylation
HSP	Heat Shock Proteins
HYP6	HYPOCOTYL6
IAA	Indole-3-acetic acid
IAA1	INDOLE-3-ACETIC ACID INDUCIBLE1
IAA17	INDOLE-3-ACETIC ACID INDUCIBLE17
IAA29	INDOLE-3-ACETIC ACID INDUCIBLE29
IM	Inflorescence Meristem
KAK	KAKTUS
kb	kilobase
KNU	KNUCKLES
LUG	LEUNIG
LUH	LEUNIG HOMOLOG
LFY	LEAFY
LP	Leaf Primordia
LMI1	LATE MERISTEM IDENTITY 1
LMI2	LATE MERISTEM IDENTITY 2
MADS	MCM1-AGAMOUS-SRF-DEFICIENS
MBW	MYB/bHLH/WD
MP	MONOPTEROS
NZZ	NOZZLE
OC	Organising Centre
PFMs	Position Frequency Matrices
PWMs	Position Weight Matrices
PIN	PIN-FORMED
PIN1	PIN-FORMED1
PI	PISTILLATA
PRC1	Polycomb Repressive Complex 1

PCR2	Polycomb Repressive Complex 2
QTL	Quantitative Trait Locus
RACE-PCR	Rapid Amplification of cDNA Ends
RBR1	RETINOBLASTOMA RELATED1
RHL2	ROOT HAIRLESS 2
RT-qPCR	Reverse Transcription Quantitative Polymerase Chain Reaction
SAM	Shoot Apical Meristem
SCF	Skp1-Cul1-F-box
SEA	Sequence Enrichment Analysis
SEU	SEUSS
SHP	SHATTERPROOF
SOC1	SUPPRESSOR OF OVEREXPRESSION OF CO1
SEP1-4	SEPALLATA 1-4
SIM	SIAMESE
SNP	Single Nucleotide Polymorphism
SPL	SQUAMOSA-Promoter binding protein-like
SPL	SPOROCYTELESS
SPT	SPATULA
SVP	SHORT VEGETATIVE PHASE
SYD	SPLAYED
TAA	TRYPTOPHAN AMINO-TRANSFERASE OF ARABIDOPSIS
TAIR	The *Arabidopsis* Information Resource
TAR	TRYPTOPHAN AMINO-TRANSFERASE RELATED
TAR1	TRYPTOPHAN AMINO-TRANSFERASE RELATED1
TAR2	TRYPTOPHAN AMINO-TRANSFERASE RELATED2
TCL1	TRICHOMELESS1
TCL2	TRICHOMELESS2
TFBS	Transcription Factor Binding Sites
TFL1	TERMINAL FLOWER 1
TFL2	TERMINAL FLOWER 2
TILLING	Targeting Induced Local Lesions in Genomes
TTG1	TRANSPARENT TESTA GLABRA1
TTG2	TRANSPARENT TESTA GLABRA2
TRY	TRIPTYCHON
UFO	UNUSUAL FLORAL ORGANS
UTR	Untranslated Region
(V/V)	% Volume per Volume (mL/100 mL)
WT	Wild-type
WUS	WUSCHEL
YUC	YUCCA
YUC1	YUCCA 1
YUC4	YUCCA 4
ZFP8	Zinc Finger Protein 8
ZWI	ZWICHEL

List of Figures

1.1 Vegetative and reproductive phase of *Arabidopsis* development. 3

1.2 Flower of *Arabidopsis thaliana* (accession: Landsberg *erecta*). 8

1.3 Domain structure of Type II MADS-domain transcription factors. 9

1.4 The quartet model. 11

1.5 Mechanism of trichome formation. 14

1.6 Trichome phenotype observed upon a knockdown of *AG*. 17

1.7 Trichome development on reproductive floral organs. 18

1.8 Trichome phenotype of a weak *ag* allele. 19

1.9 Model for suppression of trichomes on gynoecia. 21

1.10 Fates of duplicate genes. 24

1.11 Phylogeny of *Brassica*, *Arabidopsis* and their close relatives with polyploidy events indicated. 26

1.12 The triangle of U. 27

1.13 Flower of *Brassica rapa* var. Yellow Sarson. 29

3.1 Expression profile of *AG* in *Arabidopsis*. 60

3.2 Gene structure of *AtAG* and of two homologs from *B. rapa*. 63

3.3 Alignment of coding sequences of *AtAG* and the putative *B. rapa* homologs. 66

3.4 Alignment of the protein sequence of *AtAG* and putative *B. rapa* homologs. 67

3.5 Chromosomal location of *BrAG* genes. 69

3.6 Expression of *BrAG1* and *BrAG2* as determined by RT-qPCR. 72

3.7 Alignment of 3'UTR sequences of *AG* homologs. 73

3.8 Predicted AG binding sites in the upstream regions of *AG* homologs. . . 77
3.9 Predicted LFY binding sites in promoters of *AG* homologs. 78
3.10 Alignment of large intron sequences of *AG* homologs. 79
3.11 Predicted AG binding sites in the intronic region of *AG* homologs. . . . 81
3.12 Predicted LFY binding sites in regulatory introns of *AG* homologs. . . . 82
3.13 Floral phenotype of *ag-1* mutants. 83
3.14 Effects of *AtAG* overexpression. 85
3.15 Effects of overexpression of *BrAG1*. 86
3.16 Effect of *BrAG2* overexpression. 87
3.17 Results for the expression of pAG:*AtAG* constructs. 88
3.18 Results for the expression of pAG:*BrAG1* constructs. 89
3.19 Results for the expression of pAG:*BrAG2* constructs. 90
3.20 Mutant alleles for *BrAG* genes. 93
3.21 Analysis of *BrAG2* splicing in *brag2-1* mutants. 94
3.22 Alignment of cDNA sequences for *BrAG2* obtained by PCR amplification from R-o-18 and *brag2-1* mutant plants, respectively. 95
3.23 Alignment of the amino acid sequence of BrAG2 from R-o-18 and from *brag2-1*. 96
3.24 Floral phenotypes of heterozygous *brag1-1* plants. 98
3.25 Floral phenotypes of homozygous *brag1-1* plants. 99
3.26 Floral phenotype of homozygous *brag1-2* plants. 100
3.27 Floral phenotype of heterozygous *brag2-1* plants. 101
3.28 Floral phenotype of homozygous *brag2-1* mutant plants. 102
3.29 Floral phenotype of *brag2-2* mutant plants. 103
3.30 Floral phenotype of *brag1-2 brag2-2* double-mutant plants. 105
3.31 Floral phenotypes of *brag1-2 brag2-1* double-mutant plants. 106
3.32 Floral phenotype of *brag1-1 brag2-2* double-mutant plants. 107
3.33 Floral phenotypes of *brag1-1 brag2-1* double-mutant plants. 108
3.34 Floral phenotype of *ant* and *seu* mutant plants. 115

3.35 Silique of *brag1-2 brag2-1* double-mutant and *ag-10* and *knu* single mutant plants. 117
3.36 Floral phenotype of *spt crc Arabidopsis* flowers. 118

4.1 Gene structure of *AtTRY* and of three putative homologs from *B. rapa.* . 127
4.2 Alignment of coding sequences of *AtTRY* and the putative *B. rapa* homologs. 129
4.3 Alignment of protein sequences of *AtTRY* and putative *B. rapa* homologs. 130
4.4 Gene structure of *AtCPC* and of two putative homologs from *B. rapa*. . 132
4.5 Alignment of coding sequences of *AtCPC* and the putative *B. rapa* homologs. 134
4.6 Alignment of the protein sequence of *AtCPC* and putative *B. rapa* homologs. 134
4.7 Chromosomal location of *BrCPC* and *BrTRY* genes. 136
4.8 Expression profile of *AtTRY* and *AtCPC* in *Arabidopsis*. 138
4.9 Expression of *BrTRY1*, *BrTRY2* and *BrTRY3* as determined by RT-qPCR. 140
4.10 Expression of *BrCPC1* and *BrCPC2* as determined by RT-qPCR. 143
4.11 Expression of *BrTRY1*, *BrTRY2*, *BrTRY3*, *BrCPC1* and *BrCPC2* in leaf and gynoecium tissue of two *B. rapa* lines. 145
4.12 Alignment of 3'UTR sequences of *TRY* homologs. 146
4.13 Alignment of 3'UTR sequences of *CPC* homologs. 147
4.14 Phenotype of *try-29760* mutant plants. 156
4.15 Results of overexpression of *AtTRY*. 157
4.16 Results of overexpression of *BrTRY1*. 158
4.17 Results of overexpression of *BrTRY2*. 158
4.18 Effects of a pTRY:*AtTRY* construct on trichome formation. 159
4.19 Effects of a pTRY:*BrTRY1* construct on trichome formation. 160
4.20 Effects of a pTRY:*BrTRY2* construct on trichome formation. 161
4.21 Effects of a pTRY:*BrTRY3* construct on trichome formation. 162
4.22 Trichome phenotype of *cpc-1* mutant plants. 163

4.23 Results of overexpression of *AtCPC*. 164
4.24 Results of overexpression of *BrCPC1*. 165
4.25 Results of overexpression of *BrCPC2*. 166
4.26 Effects of a pCPC:*AtCPC* construct on trichome formation. 167
4.27 Effects of a pCPC:*BrCPC1* construct on trichome formation. 168
4.28 Effects of a pCPC:*BrCPC2* construct on trichome formation. 169
4.29 Mutant lines for *BrTRY* genes. 171
4.30 Mutant lines for *BrCPC* genes. 172
4.31 Leaf phenotype of *brtry1-1* plants. 173
4.32 Stem trichome phenotype of *brtry2-1* plants. 174
4.33 Petiole trichome phenotype of *brtry2-1* plants. 174
4.34 Leaf trichome phenotype of *brcpc1-1* plants. 175
4.35 Leaf trichome phenotype of *brcpc2-1* plants. 176
4.36 Leaf trichome phenotype of *brtry1-1 brtry2-1*. 177
4.37 Leaf phenotype of *brcpc1-1 brcpc2-1* double-mutant plant. 178

5.1 Phenotypic differences between R-o-18 line and C634 accession of *B. rapa*. 191
5.2 Results of *BrGL1a* amplification in R-o-18 and C634. 192
5.3 Alignment of sequences obtained for *BrGL1a* by PCR amplification from R-o-18 and C634, respectively, to *BrGL1a* from the Chiifu reference genome. 193
5.4 Alignment of the amino acid sequence of *BrGL1a* from R-o-18 and C634 to Chiifu sequence. 195
5.5 Alignment of sequences obtained for *BrGL1b* by PCR amplification from R-o-18 and C634 to *BrGL1b* from the Chiifu reference genome. . 196
5.6 Leaf trichome phenotype of the F1 progeny of C634 X R-o-18 crosses. . 197
5.7 Leaf trichome phenotype of F2 progeny of C634 R-o-18 crosses. 198
5.8 Leaf phenotype of *brgl1b-1* mutant plants. 200
5.9 Leaf trichome phenotype of the F1 progeny of C634 *brgl1b-1* crosses. . 201

5.10 Leaf trichome phenotype of the F2 progeny of an C634 *brgl1b-1* crosses. 202
5.11 Expression of *B. rapa* homologs of genes involved in the control of trichome initiation in *Arabidopsis* in whole leaf tissue. 203
5.12 Expression in leaf blade tissue of *B. rapa* homologs of genes involved in the control of trichome initiation in *Arabidopsis*. 204
5.13 Expression in leaf margin tissue of *B. rapa* homologs of genes involved in the control of trichome initiation in *Arabidopsis*. 205
5.14 *BrGL2* expression. 206
5.15 *B. rapa* lines to be generated for testing floral trichome formation. . . . 208
5.16 Floral phenotype of *brtry2-1 brag1-1* double-mutant plants. 209

6.1 A floral induction system. 219
6.2 Molecular mechanism underlying the floral induction system and the line used for a specific activation of the LFY transcription factor. 220
6.3 Effect of LFY activation on the expression of auxin-related genes. . . . 223
6.4 RNA sequencing results for auxin-related genes. 225
6.5 Read counts from RNA sequencing for *YUC4*. 226
6.6 Effects of LFY and AP1 activation on IAA levels in meristematic tissue. 227

7.1 Alignment of the genomic sequence of *AtAG*, *Bra012564* and *Bra013364* 242
7.2 Alignment of the genomic sequence of *AtTRY*, *Bra022637*, *Bra026297* and *Bra029089* 243
7.3 Alignment of the genomic sequence of *AtCPC*, *Bra004539* and *Bra039283* 243
7.4 JASPAR binding site matrices used for analysis of sequences of *AtAG*, *BrAG1* and *BrAG2* 244

List of Tables

2.1 *Arabidopsis* lines generated during the course of this study. 32

2.2 *B. rapa* TILLING lines used in this study. 33

2.3 Primer pairs used to genotype TILLING lines. 40

2.4 DMAS primers used for genotyping each TILLING line. 41

3.1 Percentage identity matrix for *AtAG* and potential *B. rapa* homologs. . . 64

3.2 Synteny of *AtAG* and putative *B. rapa* homologs. 68

3.3 Percentage identity matrix for 3' UTRS of *AG* homologs. 74

3.4 ARE binding sites detected in 3'UTR of AG homologs. 74

3.5 Percentage identity matrix for promoter sequences of *AG* homologs. . . 76

3.6 Predicted MADS-domain transcription factor binding sites for upstream promoter regions of *AG* homologs. 76

3.7 Predicted LFY binding sites in upstream regions of *AG* homologs. . . . 77

3.8 Percentage identity matrix for intronic sequences of *AG* homologs. . . . 80

3.9 Predicted MADS-domain transcription factor binding sites for regulatory introns of *AG* homologs. 81

3.10 Predicted LFY binding sites for large intronic sequences of *AG* homologs. 82

4.1 Percentage identity matrix for *AtTRY* and potential *B. rapa* homologs. . 128

4.2 Synteny of *AtTRY* and putative *B.rapa* homologs. 131

4.3 Percentage identity matrix of *AtCPC* and potential *B. rapa* homologs. . 133

4.4 Synteny of *AtCPC* and putative B.rapa homologs. 135

4.5 Percentage identity matrix for *TRY* 3'UTRs. 147

4.6 Percentage identity matrix for *CPC* 3'UTRs. 148
4.7 ARE binding sites detected in 3'UTR of *TRY* homologs. 149
4.8 ARE binding sites detected in 3'UTR of *CPC* homologs. 150
4.9 *JASPAR* binding site analysis of promoter regions of *TRY* homologs for SPL transcription factor binding sites. 152
4.10 *JASPAR* binding site analysis of promoter regions of *TRY* homologs for MADS-domain transcription factor binding sites. 153
4.11 *JASPAR* binding site analysis of introns of *TRY* homologs for SPL transcription factor binding sites. 154
4.12 *JASPAR* binding site analysis of introns of *TRY* homologs for MADS-domain transcription factor binding sites. 155

5.1 *B. rapa* homologs of *Arabidopsis* genes involved in trichome formation. 188
5.2 BLAST search results for the insertion in *BrGL1a* of R-o-18. 194
5.3 Phenotypic analysis of the F1 progeny of crosses between C634 and R-o-18. 197
5.4 Phenotypic analysis of F2 progeny of C634 R-o-18 crosses. 198
5.5 Phenotypic analysis of F1 progeny of C634 *brgl1b-1* crosses. 200

6.1 Genes involved in auxin biosynthesis and response studied by Li et al. (2013). 218
6.2 Comparison of RT-qPCR results for auxin-related genes. 222
6.3 Comparison of expression data for auxin-related genes. 224

7.1 Oligonucleotides used in the course of this study 232
7.2 BLAST results for the search of *AtAG* genomic sequence against the *B. rapa* chiffu genome. 238
7.3 BLAST results for *AtTRY* genomic sequences against the *B. rapa* chifuu genome 239
7.4 BLAST results for *AtCPC* genomic sequences against the *B. rapa* chifuu genome 240

Contents

1 Introduction **1**

1.1 Genetic Control of Flower Development 3

1.1.1 Specification of the Floral Meristem 4

1.1.2 Genetic Models of Flower Development 7

1.1.3 ABCE Model and Quartet Model 9

1.1.4 Floral Organs are Modified Leaves 11

1.2 Trichomes 12

1.2.1 Regulation of Trichome Initiation and Development 13

1.3 Control of Trichome Initiation on Floral Organs 16

1.4 Trichomes as a Strategy for Crop Protection 21

1.5 Polyploidy in Plants 23

1.5.1 Polyploidy in the Genus *Brassica* 25

1.5.2 Translating Research Findings from *Arabidopsis* to *Brassica* Crops 27

1.6 Aims of the Thesis 29

2 Materials and Methods **31**

2.1 Strains and Growth Conditions 31

2.1.1 Bacterial Strains 31

2.1.2 Plant Lines 31

2.1.3 Plant Growth Media 33

2.1.4 Bacterial Growth Media 33

2.2 Cloning 34

2.2.1 RNA Extraction and cDNA Synthesis 34
2.2.2 Preparation of Plasmid DNA 34
2.2.3 PCR Product Purification 34
2.2.4 Agarose Gel Electrophoresis 34
2.2.5 Agarose Gel Extraction of DNA 35
2.2.6 Quantification of Nucleic Acids 35
2.2.7 Preparation and Transformation of Competent *Escherichia coli* . 35
2.2.8 Colony PCR of *Escherichia coli* 35
2.2.9 Preparation and Transformation of Competent *Agrobacterium tumefaciens* 36
2.2.10 Colony PCR of *Agrobacterium tumefaciens* 36
2.2.11 Cetyltrimethyl Ammonium Bromide (CTAB) Genomic DNA Extraction 37
2.2.12 Dephosphorylation of DNA 37
2.2.13 DNA Ligation 37
2.2.14 GreenGate Cloning 38
2.2.15 Sequencing 38
2.2.16 Design of Oligonucleotides 39
2.3 Genotyping 39
2.3.1 "Edwards" DNA Preparation 39
2.3.2 Genotyping of TILLING Lines 39
2.3.3 Double-Mismatch Allele-Specific Genotyping 40
2.3.4 dCAPS genotyping 41
2.4 Rapid Amplification of cDNA Ends-PCR 41
2.5 Quantitative PCR 42
2.6 Plant Growth Conditions 43
2.6.1 Growth Conditions 43
2.6.2 Crossing Plant Lines 43
2.7 Plant Transformation and Selection 43

2.7.1 Transformation of *Arabidopsis* 43
2.7.2 Seed Sterilisation 43
2.7.3 Selection on Soil 44
2.8 Prediction of Binding Sites 44
2.9 Microscopy 45
2.10 Treatments of Lines Encoding Dexamethasone Inducible Promoters 45
2.11 Determination of Free IAA by GC-MS 45
2.12 Image Manipulation and Figure Construction 46
2.13 Tissue Collection 46
2.13.1 Floral Induction System Tissue 46
2.13.2 *Brassica rapa* Tissue 46
2.14 Generation of Transgenic Lines 47
2.14.1 Overexpression Lines 47
2.14.2 Expression Analysis Lines 51
2.14.3 Promoter Swap Experiments 52

3 Chapter 3 - Identification and Characterisation of Homologs of *AGAMOUS* in *Brassica rapa* 56

3.1 Introduction 56
3.1.1 The Role of *AG* in Floral Meristem Determinacy 57
3.1.2 The Role of *AG* in Stamen and Carpel Development 59
3.1.3 Aims of the Chapter 61
3.2 Results 63
3.2.1 Identification of *AG* Homologs in *B. rapa* 63
3.2.2 Characterisation of *BrAG1* and *BrAG2* Expression 69
3.2.3 Characterisation of the 3'UTRs of *BrAG1* and *BrAG2* 73
3.2.4 Comparative Analysis of Upstream Promoter Regions of *AG* Homologs 75
3.2.5 Comparative Analysis of Regulatory Introns of *AG* Homologs . 78

3.2.6 Results of Complementation Experiments 82
3.2.7 Identification of Mutant Alleles for *BrAG1* and *BrAG2* 91
3.2.8 Analysis of *brag* Single-Mutants 97
3.2.9 Analysis of *brag* Double-Mutants 104
3.3 Discussion 110
3.3.1 Identification and Characterisation of *AG* Homologs in *B. rapa* . 110
3.3.2 Characterisation of *BrAG1* and *BrAG2* by Examining Mutant Alleles 113
3.4 Future Work 122

4 Chapter 4 - Identification and Characterisation of Homologs of *CAPRICE* and *TRIPTYCHON* in *Brassica rapa* 124

4.1 Introduction and Aims 124
4.2 Results 126
4.2.1 Identification of *TRY* and *CPC* Homologs in *B. rapa* 126
4.2.2 Characterisation of the Expression of *B. rapa* Homologs of *CPC* and *TRY* 137
4.2.3 Comparing Expression of *BrCPC* and *BrTRY* Genes in Two Lines of *B. rapa* 144
4.2.4 Characterisation of the 3' UTRs of *BrCPCs* and *BrTRYs* 146
4.2.5 Comparative Analysis of Upstream Promoter Regions of *TRY* Homologs 150
4.2.6 Comparative Analysis of Introns of *TRY* Homologs 153
4.2.7 Results of *TRY* Complementation Experiments 156
4.2.8 Results of *CPC* Complementation Experiments 162
4.2.9 Identification of Mutant Alleles for *BrCPC1*, *BrCPC2*, *BrTRY1*, *BrTRY2* and *BrTRY3* 169
4.2.10 Generation and Analysis of *BrCPC* and *BrTRY* Single Mutants . 172
4.2.11 Analysis of *brtry* and *brcpc* Double Mutants 176
4.3 Discussion 179

4.3.1 Identification and Characterisation of *TRY* and *CPC* homologs in *B. rapa* 179
4.3.2 Characterisation of BrTRYs and BrCPCs by Examining Mutant Alleles 181
4.4 Future Work 185

5 Chapter 5 - Investigating Trichome Phenotype Variation in Two Lines of *Brassica rapa* & Generating Plants With Trichomes on Floral Organs 187
5.1 Introduction 187
5.2 Aim of the Chapter 189
5.3 Results 190
5.3.1 Examining the Genetic Mechanism Behind the Phenotypic Difference between R-o-18 and C634 190
5.3.2 Generation of Hybrids between R-o-18 and C634 196
5.3.3 Investigating the Role of *BrGL1b* in the Formation of Trichomes on the Leaves 199
5.3.4 Examining the Expression Patterns of Genes Involved in Trichome Formation in R-o-18, C634 and the Hybrid 202
5.3.5 Double Mutants of *B. rapa* homologs of *AG*, *CPC* and *TRY* . . 207
5.4 Discussion 211
5.5 Future Work 215

6 Chapter 6 - Investigating the Role of LEAFY in Auxin Modulation in Floral Meristems 216
6.1 The Role of Auxin in the Floral Meristems 216
6.1.1 Approaches to Investigating Genetic Processes in the Floral Meristem 219
6.2 Aims of the Chapter 221
6.3 Results 222

6.3.1 LFY-Dependent Regulation of Genes Involved in Auxin Responses . . . 222
6.3.2 Assessing LFY Function in the Regulation of Auxin-Related Genes through RNA Sequencing . . . 224
6.3.3 Quantification of IAA in Floral Meristems . . . 226
6.4 Discussion . . . 229
6.5 Future Work . . . 231

7 Appendix **232**
7.1 List of Oligonucleotides Used in this Study: . . . 232
7.2 Buffers and Solutions . . . 237
7.3 BLAST Results . . . 238
7.4 Genomic Sequence Alignments . . . 241
7.5 JASPAR Binding Matrices . . . 244

Introduction

Over the past 40 years, much has been learned about the genetic mechanisms underlying plant development. One of the processes that has been especially well studied during this time is the formation of flowers. This work led to the identification and characterization of dozens of genes that make up a complex gene network controlling flower development. The focus of the research on flower development is in part driven by the critical role that flowers of angiosperms have for societies, with the fruits and seeds they produce being a key source of nutrition for humans and their livestock. Flower development has been extensively studied in *Arabidopsis thaliana*, a member of the Brassicaceae family, which has proven an ideal model organism for plant research given its small size, compact diploid genome, ability to self-fertilize and available genome sequence (Meyerowitz and Pruitt, 1985; The Arabidopsis Genome Initiative, 2000). A landmark event for the field was the proposal of the ABC model of floral organ identity specification, which describes how the development of different floral organ types is driven by three classes of homeotic genes that all encode transcription factors (Coen and Meyerowitz, 1991). More recently, the advancement of new experimental techniques (e.g., RNA-sequencing and methods for the genome-wide mapping of transcription factor binding sites) has allowed for the identification of many of the downstream targets of key flower regulators (Wuest et al., 2012; Ó'Maoiléidigh et al., 2013; Pajoro et al., 2014; Yan et al., 2019). These genomic approaches have vastly expanded our knowledge of the way in which flower development is driven but also highlighted the incredible level of complexity of the gene network controlling this process.

In the sections that follow, I will describe the current understanding of the genetic network controlling flower development in *Arabidopsis* and how key experiments conducted over the past 40 years confirmed a hypothesis originally formulated by Johann Wolfgang von Goethe in the late 18th century, namely that flowers are composed of modified leaves. One of the features of leaves that was modified during floral evolution is the presence of trichomes, and I will summarize what is known about the genetic mechanism underlying trichome development in *Arabidopsis*. This will be followed by a description of work carried out previously in our laboratory showing that by knocking down the key floral development gene *AGAMOUS* (*AG*) together with trichome repressor genes, one can generate *Arabidopsis* plants that form trichomes on reproductive floral organs. I will then discuss the possibility of this supernumerary trichome phenotype being used as a way to increase the pest resistance of plants and how this could be translated into plants that are closely related to *Arabidopsis* and that are in wide agricultural use such as those in the *Brassica* genus. As there has been at least one polyploidy event in the *Brassica* lineage that is not shared with *Arabidopsis*, I will discuss the difficulties that the resulting increase in gene number has on the efforts to translate research findings from *Arabidopsis* to *Brassica* crops.

The final part of this thesis aims to examine two genes that are key to the proper establishment of the floral meristem - *LEAFY* and *APETALA1* - and if they contribute to the changing of auxin levels in the floral meristem. The movement of auxin in the inflorescence meristem is essential to ensure both the initiation of the floral meristem and the correct pattern of phyllotaxis of flowers emerging from the shoot. This was primarily carried out using previously published plant lines that allow for hormone-dependant activation of the LFY or AP1 proteins and examining the effects that activation has on expression levels of genes involved in auxin production and transport as well as quantifying changes in auxin levels over time.

1.1 Genetic Control of Flower Development

Following the completion of embryogenesis, the life cycle of *Arabidopsis* can be split into two broad stages - the vegetative phase and the reproductive phase (Figure 1.1). The vegetative phase is marked by the production and growth of all aerial parts of the plant from the shoot apical meristem - a highly organised group of stem cells. The transition to the reproductive phase is controlled by a tightly regulated genetic network and is arguably the most important developmental transition that the plant will undergo. It is marked by the transformation of the shoot apical meristem to the inflorescence meristem on the flanks of which floral meristems form and over a period of approximately 14 days, develop into mature flowers (Wellmer and Riechmann, 2010). Floral meristems contain stem cell pools which act as the source of cells for the different organs of the flower.

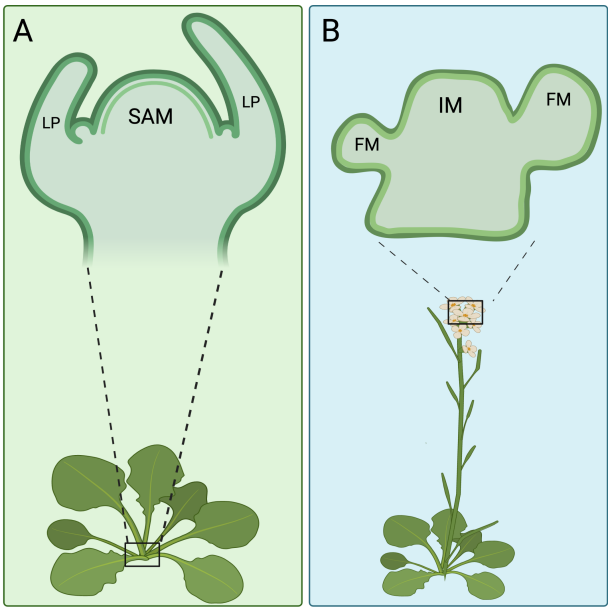

Figure 1.1: Vegetative and reproductive phase of *Arabidopsis* development. (A) Schematic of the vegetative phase where the shoot apical meristem (SAM) of the plant produces all aerial parts of the plant with rosette leaves being produced from leaf primordia (LP). (B) Once the plant transitions to the reproductive phase, the SAM becomes the inflorescence meristem (IM), and floral meristems (FM) form on the flanks of the IM and develop into mature flowers. Figure generated with BioRender.com.

1.1.1 Specification of the Floral Meristem

Flowering in *Arabidopsis* is prevented from occurring prematurely by several pathways that all involve integrating external environmental cues such as temperature and daylight and internal cues such as the age of the plant and the levels of the phytohormone gibberellic acid (GA) (Kinoshita and Richter, 2020). In the case of *Arabidopsis* accessions that must undergo vernalization in order to flower, there is continual expression of the powerful floral repressor gene *FLOWERING LOCUS C* (*FLC*) as well as continual expression of shoot identity genes such as *TERMINAL FLOWER1* (*TFL1*) in the shoot of the plant. (Searle et al., 2006; Hanano and Goto, 2011). Following exposure to a prolonged period of cold, *FLC* is silenced by the recruitment of the Polycomb Repressive Complex 2 (PRC2) to the locus and the deposition of repressive Histone H3 Lysine-27 tri-methylation (H3K27me3) marks. Once *FLC* has been silenced, the gene *FLOWERING LOCUS T* (*FT*) becomes expressed in leaf tissue. The FT protein is synthesised in the phloem and travels *via* the vascular system to the shoot apical meristem (Turck et al., 2008; Chen et al., 2018) where it acts in conjunction with the transcription factor FD to upregulate genes involved in the establishment of the floral meristem such as *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* (*SOC1*), *APETALA1* (*AP1*) and *FRUITFULL* (*FUL*) (Abe et al., 2005; Wigge et al., 2005).

Another important pathway controlling flowering is the photoperiodic pathway, which integrates increased length and strength of daylight and induces flowering in *Arabidopsis*. Under short-day conditions such as winter when daylight averages 8-10 hours a day, peak levels of the mRNA of the transcription factor *CONSTANS* (*CO*) do not coincide with daylight and is degraded by the RING motif containing E3 ubiquitin ligase CONSTITUTIVE PHOTOMORPHOGENIC1 and SUPPRESSOR OF PHYTOCHROME A-105 1 (Laubinger et al., 2006; Jang et al., 2008; Turck et al., 2008). Under long-day conditions (approximately 16 hours of light per day), CO becomes stabilized and accumulates to sufficient levels to activate the expression of FT in leaves. Thus, FT is

a so-called floral pathway integrator where several flowering pathways converge (Turck et al., 2008; Johansson and Staiger, 2015).

SOC1 works in conjunction with *AGAMOUS-LIKE24* (*AGL24*) to directly promote the expression of *LEAFY* (*LFY*) and indirectly promote the expression of *AP1 via* LFY (Gregis et al., 2006; Liu et al., 2008). LFY is a plant-specific transcription factor that is one of the main floral meristem identity genes. The expression of *LFY* in the floral *anlagen* (the emerging floral primordia) marks the beginning of the emergence of the floral meristem, however *LFY* itself is not expressed in the floral meristem due to repression by *TFL1* (Weigel et al., 1992; Wils and Kaufmann, 2017; Goslin et al., 2017). With the activation of *AP1* and *LFY*, the floral meristem becomes fully established and the genetic cascade that drives the development of floral organs begins.

The formation of the floral meristem is also controlled by the plant hormone auxin with its development from the SAM preceded by the establishment of an auxin maxima at the point of emergence (Reinhardt et al., 2003). This high concentration of auxin is driven by polar auxin transport which involves the PIN-FORMED family of auxin efflux carriers PIN1. The protein kinase PINOID phosphorylates PIN proteins thus altering its apical-basal polarity and controlling the direction in which auxin is transported, acting in conjunction with other auxin-influx carriers likely members of the AUX1/LAX family, resulting in an auxin sink. (Friml et al., 2004; Périlleux et al., 2019). The F box protein TIR1 is an auxin receptor and subunit of the ubiquitin ligase complex SCF (Skp1-Cul1-F-box). In the presence of auxin, SCF^{TIR1} triggers the degradation of transcriptional repressors Aux/IAA thus allowing for the transcriptional regulators known as AUXIN RESPONSE FACTORS (ARFs) to become active (Mockaitis and Estelle, 2008).

The ARF MONOPTEROUS (MP/ARF5) directly induces expression of *LFY* by binding to its promoter (Yamaguchi et al., 2013). MP also directly binds to the promoter

regions of AINTEGUMENTA (ANT) and AINTEGUMENTA- LIKE6/PLETHORA3 which have been shown to be key regulators of the outgrowth of the floral meristem (Yamaguchi et al., 2014; Krizek et al., 2021). LFY then activates the expression of *LATE MERISTEM IDENTITY1* (*LMI1*) and both act together to activate the expression of *CAULIFLOWER* (*CAL*) (Saddic et al., 2006). LFY and CAL also activate the expression of *APETALA1* and AP1 promotes the expression of *LFY* as part of a positive feedback loop (Kaufmann et al., 2010). Given the critical role of auxin in the initiation of the floral meristem, the work detailed in Section 6.2 examines the role that AP1 and LFY have in controlling the auxin levels in the floral meristem.

In order to ensure correct flower development, AP1 initially represses genes involved in shoot development as well as the genes that suppress the onset of flowering including *TEMPRANILLO1* and *TEMPRANILLO2*, *TFL1*, *TARGET OF EAT 1* and *3*, *SCHNARCHZAPFEN* and *SVP* (Kaufmann et al., 2010). In a second step, AP1 then causes upregulation of the expression of *LFY* thus completing a positive feedback loop between the two that ensures high expression of both key floral meristem identity genes thus ensuring maintenance of the floral meristem (Wagner et al., 1999; Kaufmann et al., 2010).

In the newly established floral meristem, LFY works in conjunction with UNUSUAL FLORAL ORGANS (UFO) and AP1 to activate the expression of the floral organ identity genes *APETALA3 (AP3)* and *PISTILLATA* (*PI*) in regions of the floral meristem that will give rise to the second and third whorl of the flower (Honma and Goto, 2000; Krizek and Fletcher, 2005; Rieu et al., 2023). *LFY* also works alongside *WUSCHEL* (*WUS*) in order to activate the expression of another floral organ identity gene, *AGAMOUS* (*AG*), in the centre of the floral meristem to drive the development of the reproductive floral organs and to regulate the determinacy of the floral meristem (Shang et al., 2019).

WUS encodes a transcription factor which promotes the maintenance of the stem

cell pool whilst the protein products of three *CLAVATA* genes antagonise its activity in a regulatory feedback loop (Sablowski, 2007; Zhou et al., 2018; Zhang et al., 2021). The termination of the floral meristem occurs as a result of the stem cell pool being depleted and needs to be carefully timed to ensure that there are sufficient stem cells in order to allow for complete and proper development of the flower and all its floral organs but terminate at an appropriate time in order to avoid an over-proliferation of tissue (Chang et al., 2020). *AG* begins to be expressed at stage three of flower development (when the sepal primordia begin to arise) and in addition to its role in the development of the reproductive organs, AG also represses the expression of WUS thus preventing the maintenance of stem cells in the organizing centre (Smyth et al., 1990; Krizek and Fletcher, 2005). The role that AG plays in the development of reproductive floral organs and floral meristem determinacy is discussed in greater detail in Section 3.1.

1.1.2 Genetic Models of Flower Development

A typical flower of *Arabidopsis* is made up of four distinct whorls as shown in Figure 1.2. The first whorl contains four sepals, the second whorl is made up of four petals, the third whorl contains six stamens, and the innermost fourth whorl contains the gynoecium, which is composed of two fused carpels (Ó'Maoiléidigh et al., 2014). Over 30 years ago, Coen and Meyerowitz proposed a model to explain the genetic basis for the development of the floral organs of *Arabidopsis* known as the ABC model. This model classified the genes controlling the developmental process into three groups - A-class genes, B-class genes, and C-class genes (Coen and Meyerowitz, 1991). These genes are known as the floral organ identity or floral homeotic genes and their function was determined *via* the examination and the genetic analysis of mutant lines that developed flowers with whorls containing incorrect organs – similar in nature to the homeotic mutations seen in the fruitfly *Drosophila melanogaster* (Goto et al., 2001).

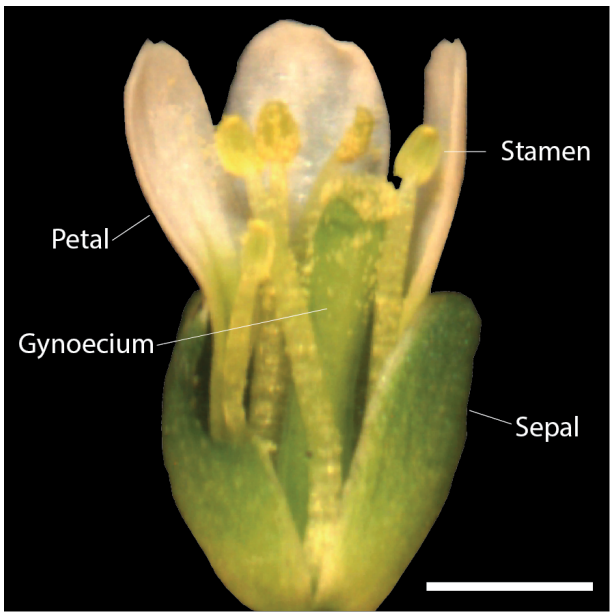

Figure 1.2: Flower of *Arabidopsis thaliana* (accession: Landsberg *erecta*). The flower is made up of four sepals in the outermost whorl, four petals in the second whorl, the third whorl contains six stamens and the fourth whorl contains two fused carpels form the gynoecium in the centre. A petal and a sepal have been removed to aid viewing. Scale bar: 1 mm.

Under this model, the formation of sepals is controlled by A-class genes alone whose members include *AP1* and *APETALA2* (*AP2*). Petal formation is controlled by both A-class genes and the B-class genes *AP3* and *PISTILLATA* (*PI*). The formation of stamens is coordinated by both the B-class genes and the C-class gene *AG*. Lastly, the formation of carpels is driven solely by *AG*. The A-class genes and C-class genes act antagonistically to the other's activity in addition to their roles in driving organ formation. The homeotic transformations observed in mutant lines for these genes show striking phenotypes with organs forming in an inappropriate whorl. Specifically, the A-class mutant *ap2* forms flowers with carpels in the place of sepals in the first whorl and stamens in the place of petals in the second whorl (Bowman et al., 1989); the B-class mutants *ap3* and *pi* form flowers that contain sepals in the second whorl and carpels in the third whorl in place of petals and stamens respectively (Bowman et al., 1989); and the C-class mutant *ag* develops flowers that have petals in the place of stamens and an additional flower of the same phenotype forms in whorl four (Bowman et al., 1989).

All of the floral organ identity genes (with the exception of *AP2*) encode transcription factors of the MADS-domain protein family (MADS for **M**CM1-**A**GAMOUS-**D**EFICIENS-**S**ERUM-RESPONSE-FACTOR). These transcription factors recognize and bind DNA motifs known as CArG-box sequences which have the consensus sequence $CC(A/T)_6GG$ or $GTA(A/T)_4TAG$ (Becker and Ehlers, 2016). These transcription factors belong to the Type II family of the MADS-domain family and are characterized by the possession of four distinct domains – a MADS domain, an Intervening (I) domain, a Keratin-like (K) domain, and a carboxy-terminal (C) domain (Figure 1.3). The MADS domain and I domain are essential for the proteins' ability to bind to DNA whilst the I and K domains mediate protein-protein interactions (Causier et al., 2010).

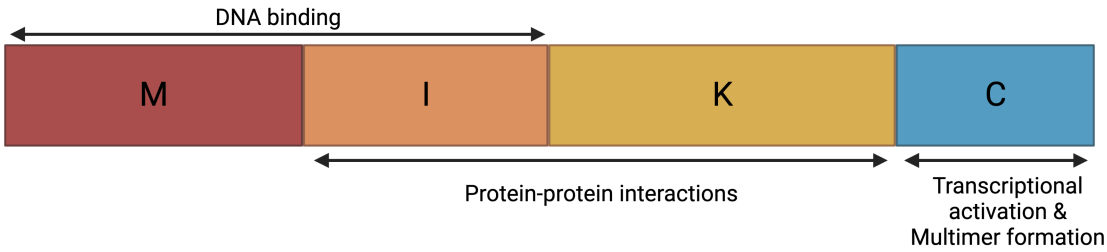

Figure 1.3: Domain structure of Type II MADS-domain transcription factors. Type II MADS-domain transcription factors are made up of four domains: the MADS domain (M) and the intervening domain (I) which function in DNA binding. The I-domain and the keratin (K) domain are involved in protein-protein interactions. The carboxyl domain (C) functions in transcriptional activation and multimer formation. Figure generated with Biorender.com.

1.1.3 ABCE Model and Quartet Model

When all three classes of floral organ identity genes are mutated simultaneously, leaf-like organs are observed in place of floral organs indicating that all three classes are necessary for normal flower development to occur (Bowman et al., 1991). However, when A-, B- and C-class genes together were ectopically expressed, leaves were not converted into floral organs, indicating that the ABC genes alone are not sufficient to drive floral organ development (Mizukami and Ma, 1992, Honma and Goto, 2001). Through re-

verse genetic analysis, additional factors were later identified, which are encoded by four closely related and partially redundant genes termed *SEPALLATA1-4* (*SEP1-4*). These genes were denoted as 'E-class genes' because the term 'D-class gene' had already been assigned to genes that drive the development of ovules within the gynoecium (Pelaz et al., 2000; Favaro et al., 2003)

The expression of *SEP1-4* commences after that of the floral meristem identity genes but before the activation of the A, B, and C-class genes and thus, they seem to lay the base for the other floral organ identity genes to build upon (Wils and Kaufmann, 2017). A mutant affected in all four *SEP* genes forms leaf-like structures in place of floral organs and thus mimics plants with combinations of mutations in A, B and C class genes. Co-expression of the ABCE genes was shown to be sufficient to convert leaves into floral organs with the type of floral organ formed depending on the exact combination of floral homeotic genes used in the experiments (Goto et al., 2001; Pelaz et al., 2001).

As mentioned above, almost all of the floral organ identity genes encode MADS-domain proteins which are able to interact with other proteins *via* their I- and K-domains (Figure 1.3). Experimental evidence obtained in the 1990s showed that homodimers of AP1 and AG as well as AP3/PI heterodimers can form and these dimers are able to form higher order complexes (Riechmann et al., 1996; Mizukami et al., 1996). SEP proteins were later shown *via* yeast two-hybrid and co-immunoprecipitation experiments to also have the ability to interact with other floral organ identity factors (Honma and Goto, 2001). These combined findings led to the proposal of an updated model of how floral organ identity factors act to control the expression of their target genes. This model, which is known as "the quartet model" proposes that floral organ identity genes work in higher-order protein complexes which have the ability to recognise two CArG-boxes simultaneously to regulate the expression of their targets (Figure 1.4). These quartets of proteins are thought to have the ability to cause DNA to bend and loop and this modulates the expression of genes located nearby (Theißen, 2001; Smaczniak et al., 2012;

Mendes et al., 2013).

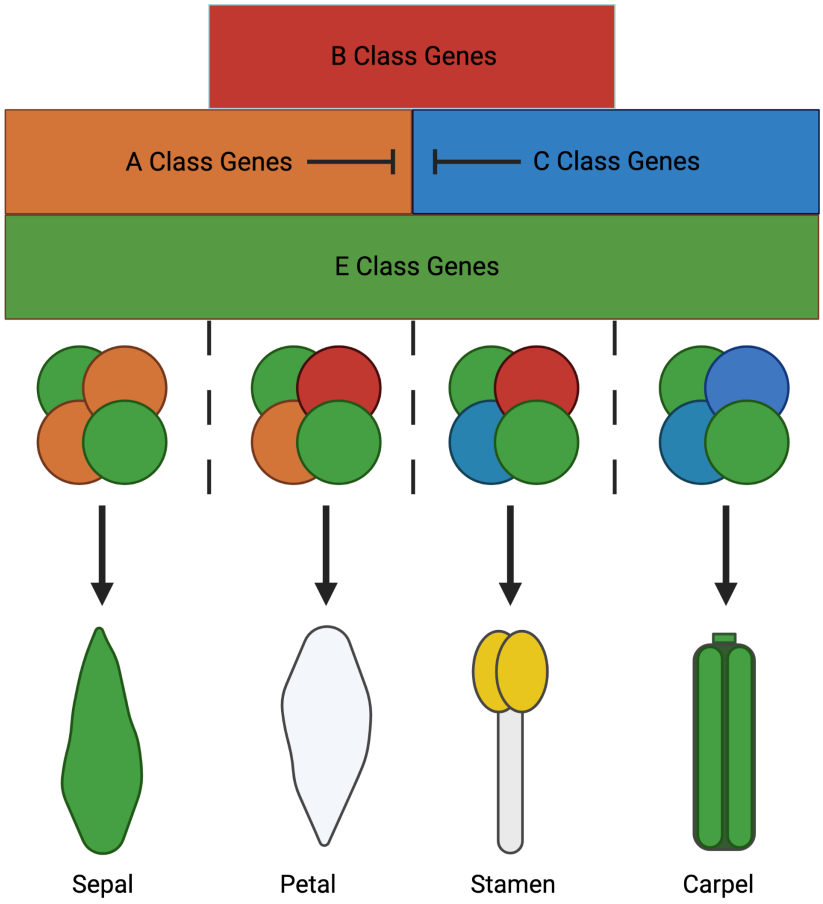

Figure 1.4: The quartet model. A-, B-, C- and E-class genes and their resulting proteins are responsible for the development of floral organs. E-class genes are required for the correct development of floral organs across all four whorls. Sepals require A- and E-class genes, petals require A-, B- and E- class genes, stamens require B-, C- and E-class genes and carpels require C- and E-class genes for proper development. A- and C-class genes act antagonistically to restrict the other's expression domain. The MADS-domain transcription factors form tetramers. The circles represent the different MADS-domain transcription factors involved in the development of each floral organ. These tetramers induce DNA looping and influence gene expression during flower development. Figure generated with Biorender.com.

1.1.4 Floral Organs are Modified Leaves

In his seminal essay "The Metamorphosis of Plants", published in 1790, Johann Wolfgang von Goethe hypothesized that all above-ground organs of a plant, including floral organs, are derived from leaf-like structures (von Goethe, 1790). Over 200 years later,

his ideas were confirmed when plants containing mutations for A-, B-, and C-class genes or for all functionally redundant E-class genes were found to develop flowers that consisted of leaf-like organs (see above). Furthermore, the ectopic expression of certain combinations of floral homeotic factors was found to be sufficient for the transformation of leaves into floral organs providing further support for Goethe's hypothesis. However, the molecular mechanisms underlying this transformation process are currently not well understood in spite of the significant progress that has been made in recent years in elucidating the gene expression programmes that are controlled by the floral homeotic factors (Bowman et al., 1991; Pelaz et al., 2001; Honma and Goto, 2001; Ditta et al., 2004). One aspect of leaf development that is clearly affected by floral organ identity factors is the formation of trichomes.

1.2 Trichomes

Trichomes are aerial projections that grow out from the epidermis of the plant aerial tissues. They are typically found on the stems, petioles, leaves, and sometimes the fruits or seedpods of plants with each singular trichome forming from an epidermal pavement cell (Hülskamp et al., 1994). Trichomes vary widely across plant species – they can be unicellular or multicellular, they can be simple or form branches and they can be glandular or non-glandular (Wagner et al., 2004). The structure of trichomes can vary widely within species depending on the type of tissue on which they form or at which stage of development they come about. For example in *Arabidopsis*, trichomes are unicellular and branched, with the first few rosette leaves forming trichomes on the adaxial surface and lacking trichomes on the abaxial surface. Leaves that subsequently form have trichomes on both the adaxial and abaxial surface (Telfer et al., 1997; Szymanski et al., 2000).

Trichomes have several known biological functions, including temperature regula-

tion and decreasing the amount of water lost *via* transpiration by increasing the leaf-air boundary resistance layer (Wagner et al., 2004). Trichomes also play an important role in the reflection of UV light in order to provide a level of protection from the damaging rays of the sun (Karabourniotis et al., 1993). These trichomes are hollow, filled with air and have a white colour which enables them to reflect sunlight efficiently. Perhaps most importantly, trichomes are involved in pest defences (Kessler and Baldwin, 2002). Simple trichomes on the leaves and stems of plants provide defence by acting as a physical barrier against insects that stops them from feeding, prevents oviposition, and prevents infection with pathogens that insects often carry (Levin, 1973; Obermeier et al., 2022). Glandular trichomes also play a role in plant defences by secreting exudates containing metabolites that repel insects or producing secretions that are extremely viscous thus impeding the movement of any insects that attempt to access the plant (Wagner, 1991; Wheeler and Krimmel, 2015).

1.2.1 Regulation of Trichome Initiation and Development

The process of trichome formation can be broken down into three distinct stages - initiation, endoreplication and branching - and is well studied in *Arabidopsis* which forms primarily branched trichomes on its rosette leaves and less branched or unbranched trichomes on cauline leaves, stems and sepals of flowers (Schnittger and Hülskamp, 2002; Xu et al., 2019). *Arabidopsis* trichomes form as a result of individual cells on the epidermis exiting the mitotic cycle of replication and undergoing endoreplication - a process where the nuclear genome of the cell replicates without undergoing mitosis (Hülskamp et al., 1994; Yang and Ye, 2013). The developing trichome typically completes four cycles of endoreplication, reaching an average DNA content of 32C. The trichomes on the surface of *Arabidopsis* develop with an average of three to four cells separating them and form no clusters under normal developmental conditions (Larkin et al., 1996; Bramsiepe et al., 2010).

Trichome development is mediated by the formation of a trichome initiation complex which is also known as the MYB/bHLH/WD (MBW) activator complex as seen in Figure 1.5.

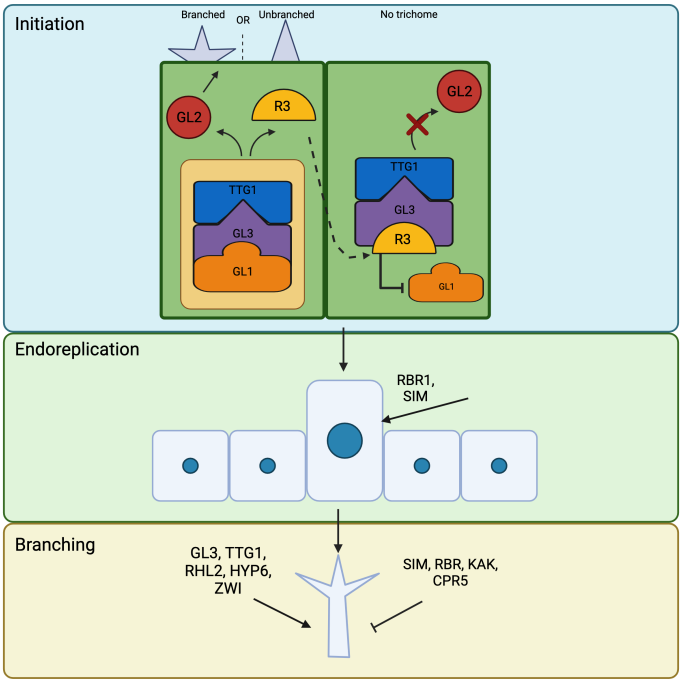

Figure 1.5: Mechanism of trichome formation. In epidermal cells (green rectangles), the formation of trichomes is dependent on the formation of a transcriptional activator complex made up of GLABRA1 (GL1), TRANSPARENT TESTA GLABRA1 (TTG1) and GLABRA3 (GL3). The activator complex directly activates the expression of many genes involved in the control of trichome initiation, branching and spacing including *GLABRA2* (*GL2*). The formation of the activator complex additionally activates the expression of single repeat R3-MYBs which act as inhibitors of trichome formation. These R3-MYBs proteins travel laterally to neighbouring cells and block trichome formation in these nearby epidermal cells by competing with GL1 to bind with TTG1 and GL3. In these neighbouring cells, the activator complex fails to form, there is no activation of *GL2* or *TTG2* and trichome formation is inhibited. The cell that contains the activator complex enters endoreplication which is positively regulated by SIAMESE (SIM) and RETINOBLASTOMA RELATED1 (RBR1). Once the process of endoreplication is complete, branching begins with many positive interactors (GL3, TTG1, RHL2, HYP6, ZWI) and negative interactors (SIM, RBR, KAK, CPR5) directing the process. Arrows indicate positive regulation and blunt lines indicate negative regulation. Figure generated with Biorender.com.

This complex is made up of three proteins – the R2R3-MYB type transcription factor GLABRA1 (GL1), the WD40 repeat protein TRANSPARENT TESTA GLABRA1

(TTG1), and a basic helix-loop-helix transcription factor, either GLABRA3 (GL3) or ENHANCER OF GLABRA3 (EGL3) (Payne et al., 2000; Zhao et al., 2008). The formation of the MBW activator complex directly activates the expression of *GLABRA2* (*GL2*) which marks the process of trichome initiation (Arteaga et al., 2021).

ChIP experiments have led to the identification of a number of targets of the MBW activator complex that are involved in initiating endoreplication such as *SIAMESE* (*SIM*) and *RETINOBLASTOMA RELATED1* (*RBR1*) (Morohashi and Grotewold, 2009). SIM is a putative CDK inhibitor and is thought to enable the cell to enter into endoreplication by inhibiting cytokinesis with *sim* mutants developing multicellular trichomes and clusters of trichomes (Walker et al., 2000; Churchman et al., 2006; K. Wang et al., 2020). RBR1 interacts with members of the E2F transcription factor family and is likely to restrict the number of endocycles during trichome formation (Desvoyes et al., 2006; Bramsiepe et al., 2010).

The final stage of trichome development is the process of branching (Figure 1.5). Some of the known positive regulators of trichome branching include GL3, TTG1, ROOT HAIRLESS 2 (RHL2), HYPOCOTYL6 (HYP6) and ZWICHEL (ZWI) (Wang et al., 2021). Plants containing *zwi*, *rhl2*, or *hyp6* mutant alleles develop trichomes with fewer branches (Esch et al., 2003; Wang et al., 2022). Negative regulators of trichome branching include SIM, RBR, KAKTUS (KAK) and CONSTITUTIVE PATHOGEN RESPONSE 5 (CPR5). SIM, CPR5, and RBR all affect trichome branching by controlling nuclear replication while KAK mainly regulates the expression pattern of GL3 and EGL3 causing reduced trichome branching (Perazza et al., 1999; Kirik et al., 2001; Bramsiepe et al., 2010).

The formation of the MBW activator complex also results in the expression of single repeat R3-MYBs which act to inhibit trichome formation in neighbouring cells (Wang et al., 2022). Seven of these trichome repressors have been identified in *Arabidopsis*

to date: *CAPRICE* (*CPC*), *TRIPTYCHON* (*TRY*), *ENHANCER OF TRY AND CPC1* (*ETC1*), *ENHANCER OF TRY AND CPC 2* (*ETC2*), *ENHANCER OF TRY AND CPC 3* (*ETC3*), *TRICHOMELESS1* (*TCL1*) and *TRICHOMELESS2* (*TCL2*). (Schellmann et al., 2002; Kirik, Simon, Huelskamp, et al., 2004; Wang et al., 2007; Ishida et al., 2008; Wester et al., 2009; Gan et al., 2011). The R3-MYBs travel laterally to neighbouring cells, preventing the binding of GL1 to GL3, blocking the formation of the MBW activator complex and preventing expression of *GL2* (Figure 1.5) (Wang et al., 2021). It has been shown that although the R3-MYBs are closely related, they are not fully functionally redundant (Yang and Ye, 2013). For example, *cpc* mutants have a higher density of trichomes on the stem and leaves compared to wild-type plants whereas in *try* mutants trichomes are observed developing in clusters on the plant surface (Schellmann et al., 2002; Kirik, Simon, Wester, et al., 2004). It has been shown that these repressor proteins move laterally between neighbouring cells to nearby epidermal cells where they compete with GL1 for inclusion in the MBW activator complex (Figure 1.5) (Wang and Chen, 2014). In these neighbouring epidermal cells, the activator complex is unable to form and thus is unable to initiate the expression of downstream targets such as *GL2* or *TTG2*, and trichome formation is blocked (Wang et al., 2021). This is known as the lateral inhibition model of trichome formation and explains the mechanism by which trichomes develop and are distributed evenly across surfaces on which they develop. There has been limited experimental work investigating how far the R3-MYBs can travel once activated but recent mathematical modelling using a weak *ttg1* allele postulates that CPC is working on a long-range basis and TRY is working on a short-range basis (Balkunde et al., 2020).

1.3 Control of Trichome Initiation on Floral Organs

As mentioned above, while experimental work confirmed von Goethe's proposal that floral organs are modified leaves, the genetic mechanisms underlying this metamorpho-

sis are not well understood. Work carried out previously in our laboratory provided some insights into this process by elucidating how trichome initiation is repressed in *Arabidopsis* flowers. In experiments carried out by former group members and described in Ó'Maoiléidigh et al. (2013) and Ó'Maoiléidigh et al. (2018), a transgenic *Arabidopsis* line allowing the ethanol inducible expression of an artificial microRNA (amiRNA) that targets *AG* (amiRNA-*AG*^{AlcR}) was generated and characterised. It was observed that after an amiRNA-mediated perturbation of AG function, trichomes occasionally formed on carpel valves (Figure 1.6).

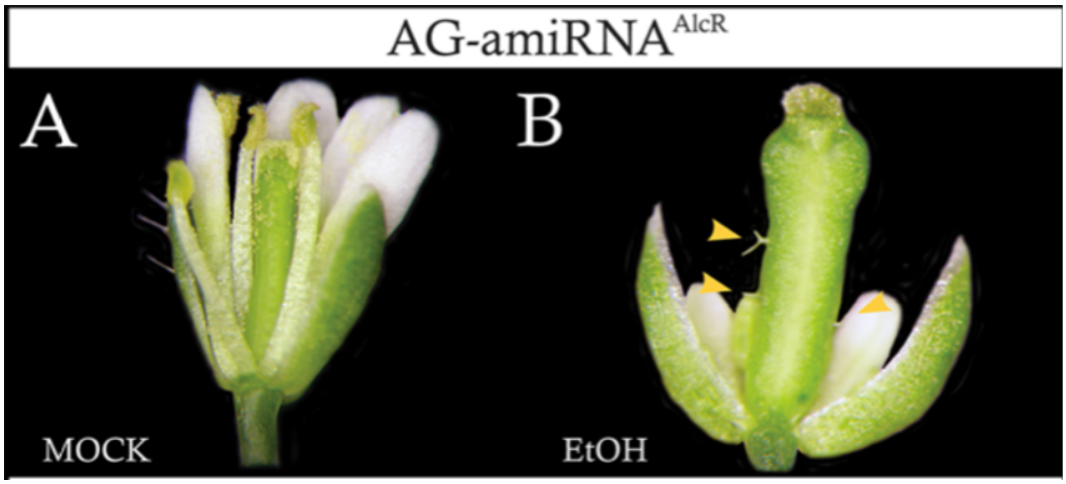

Figure 1.6: Trichome phenotype observed upon a knockdown of *AG*. (A) Light microscope image of *Arabidopsis* flower from the AG-amiRNA^{AlcR} line treated with mock-solution. No trichomes are observed on mock-treated flowers. (B) Light microscope image of a flower from the AG-amiRNA^{AlcR} line treated with ethanol. The yellow arrow indicates a branched trichome present on the carpel value of the flower. Figure taken from from Ó'Maoiléidigh et al., 2018.

Additional experiments showed that the mechanism of trichome formation on gynoecia was dependent on the action of the trichome initiation complex (Section 1.2). When the amiRNA-AG^{AlcR} was introgressed into loss of function mutants for key components of this complex (see above) trichome formation on floral organs was suppressed. It had previously been found that AG inhibits the expression of *GL1* by binding to its regulatory region, indicating that AG acts to inhibit trichome formation on the reproductive floral organs by preventing the trichome initiation complex from forming

(Ó'Maoiléidigh et al., 2013). The amiRNA-AG^{AlcR} line was then crossed with mutant lines for some of the R3-MYB transcription factors which act as trichome repressors to test if this would affect the trichome phenotype observed in the gynoecium (Figure 1.7).

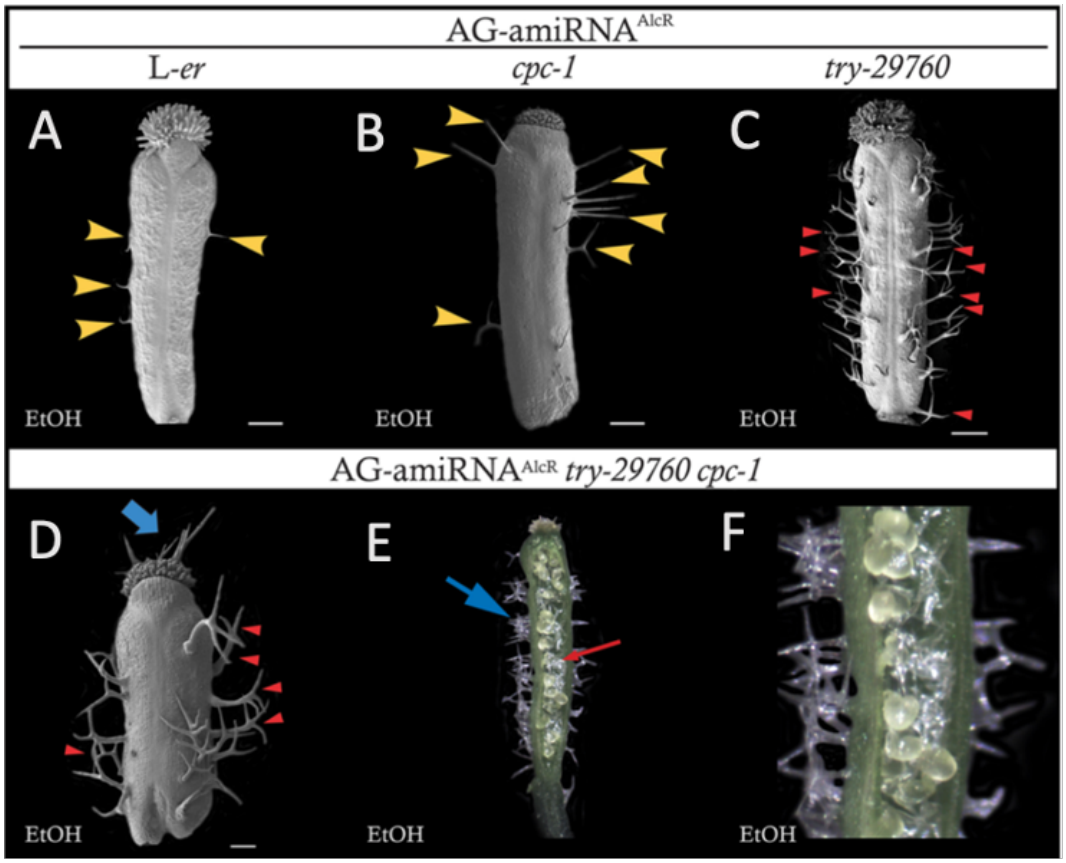

Figure 1.7: Trichome development on reproductive floral organs. Scanning electron micrographs of gynoecia from ethanol-treated AG-amiRNA^{AlcR}. (A) Gynoecium from AG-amiRNA^{AlcR} plant with trichomes present on the carpel valve. (B) Gynoecium from AG-amiRNA^{AlcR} *cpc-1* plant. An increased number of trichomes are observed compared to (A). (C) Gynoecium from AG-amiRNA^{AlcR} *try-29760* plant. A large increase in trichome number is observed compared to (A) and (B) and the trichomes developing tend to be branched in structure. (D) AG-amiRNA^{AlcR} *cpc-1 try-29760* gynoecium. A large number of branched trichomes form on the carpel valve and also on the abaxial surface. (E) and (F) a gynoecium from an AG-amiRNA^{AlcR} *cpc-1 try-29760* plant with trichomes forming from the carpel valves pictured using a light microscope. Yellow arrows indicate unbranched trichomes, red arrows indicate branched trichomes and blue arrows indicate the formation of trichome clusters on the abaxial carpel valves. Scale bars: 200 μm. Figure taken from Ó'Maoiléidigh et al. (2018).

When crossed individually with *cpc* or *try* mutants there was a significant increase in the number of trichomes seen forming on the carpel valves of the gynoecia (Figure

1.7 B-C). When *AG* was knocked down in a *try* mutant background, a strong increase in the numbers of trichomes on the carpel valve was observed and the trichomes that developed were similar to those seen on rosette leaves. An even stronger effect was observed when *AG* was knocked down in a *try cpc* double-mutant background (Figure 1.7D-F). In these plants, a greater number of trichomes formed on the carpel valve than in either *try* or *cpc* alone and all the trichomes appeared to be branched with additional trichomes developing occasionally on the anthers. These results were replicated in plants carrying the weak *ag-10* mutant allele instead of the inducible amiRNA line for *AG*. Plants containing the *ag-10* allele itself never develop trichomes on the gynoecia but when crossed with a strong *try* allele or a strong *cpc* allele, the gynoecia consistently developed trichomes and occasionally trichomes could be seen on stamens (Figure 1.8).

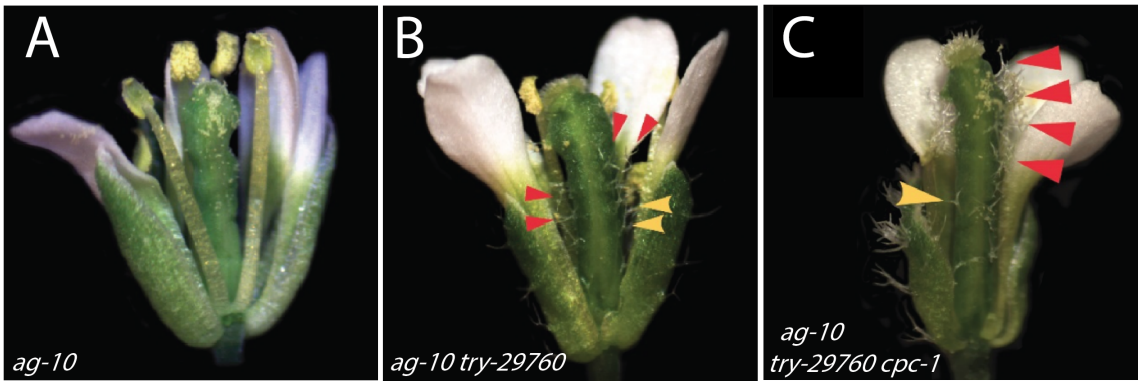

Figure 1.8: Trichome phenotype of a weak *ag* allele. (A) Light microscope image of a flower from an *ag-10* plant. This line alone never develops trichomes on the gynoecium. (B) Light microscope image of a flower from a *ag-10 try-29760* plant. Both unbranched and branched trichomes form on the gynoecium. (C) Light microscope image of *ag-10 try-29760 cpc-1* flower. An increased number of trichomes on the gynoecium is observed compared to (B). Yellow arrows and red arrows indicate unbranched and branched trichomes respectively. Figure taken from from Ó'Maoiléidigh et al. (2018)

Phytohormones such as gibberellins, jasmonic acid and cytokinin are known to promote trichome formation initiation on leaves (Maes and Goossens, 2010). Wild-type plants treated with the cytokinin 6 benzylaminopurine (BAP) never formed trichomes on the gynoecia while a particularly strong increase in the trichomes numbers was observed following treatment of both the amiRNA-AG^{AlcR} and *ag-10* plants. AG-amiRNA^{AlcR}

cpc-1 try-29760 plants treated with BAP developed an even higher number of trichomes than the AG-amiRNA^{AlcR} alone. From these results, it appears that in the gynoecium, *AG* prevents the formation of trichomes by limiting the cytokinin response in the gynoecium.

Previous work had shown that SPL transcription factors encoded by *SQUAMOSA PROMOTER BINDING- LIKE* (*SPL*) genes are involved in the regulation of trichomes by promoting the expression of trichome repressors such as *CPC*, *TRY* and *TCL1* (Yu et al., 2010; Cui et al., 2022). By introducing a transgene into AG-amiRNA^{AlcR} plants that allow for the overexpression of the microRNA miR156 (miR156^{OE}), which targets *SPL* genes, the role of the SPL/miR156 pathway of trichome development in the gynoecia was investigated. AG-amiRNA^{AlcR} miR156^{OE} formed significantly more trichomes on the gynoecia than AG-amiRNA^{AlcR} plants alone. A similar phenotype was observed for *ag-10* ^{OE} plants.

The results described by Ó'Maoiléidigh et al. (2018) suggest the activities of the floral homeotic proteins overlap with the gene network controlling trichome formation (Figure 1.9). *AG* appears to suppress the emergence of trichome formation by several means: by upregulating the expression of R3-MYB trichome repressors, by preventing expression of *GL1*, and by limiting the cytokinin response thus blocking trichome development.

A recent study demonstrated the relevance of the work described above for the biology of *Arabidopsis*. It was shown that trichome formation occurs on fruits of natural populations of *Arabidopsis* found across the Iberian peninsula (Arteaga et al., 2021). One such accession contains a series of mutations causing both a reduction in the function of the trichome repressors *TRY* and *TCL1* as well as a gain of function of the positive regulator *GL1*. Further research showed that the presence of additional trichomes on stems, pedicels and fruits has likely evolved as a response to abiotic stresses such as

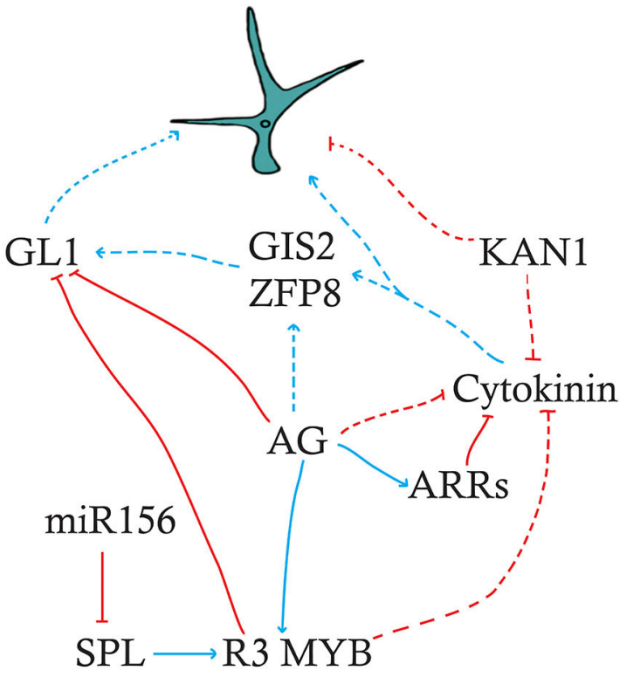

Figure 1.9: Model for suppression of trichomes on gynoecia. AG is central to suppressing the trichomes on gynoecia, acting to promote the expression of genes that repress trichome formation (R3 MYB) and repress genes essential to the formation of trichomes (GL1). AG also controls cytokinin responses directly through type-A ARRs and indirectly *via* regulation of cytokinin response genes *GIS2* and *ZFP8*. *SPL* genes which encode transcription factors positively regulate the repressors of trichome formation (R3 MYBs) in the gynoecium which was shown via overexpressing miR156 which targets *SPLs*. SPL and KAN1 transcription factors which control trichome formation *via* different regulatory mechanisms overlap with this pathway. Red lines indicate negative regulation and blue lines indicate positive regulation. Direct regulation is indicated by solid lines and indirect regulation or an unknown mechanism is indicated by dashed lines. Figure taken from Ó'Maoiléidigh et al. (2018).

---

periods of drought and intense UV exposure (Arteaga et al., 2022).

1.4 Trichomes as a Strategy for Crop Protection

An average of 35% of crops produced globally never make it to harvest due to pest infestation, weeds, or disease (Popp et al., 2013). Many pesticides are currently in wide use to prevent crop loss with approximately 2.5 million tonnes used every year to protect plants (Alavanja, 2009; Orzech and Nichter, 2008; Ma et al., 2021). As early as the

1970s, it was suggested that eventually, pesticide resistance that could not be countered with increased dosage or increased frequency of treatment as well paired with a lack of development of new effective pesticides would result in the treatments becoming ineffective which nearly 50 years later is fast becoming a reality (Van Den Bosch, 1976; Orzech and Nichter, 2008). The rise of resistance levels paired with government organizations such as the European Protection Agency banning or restricting the use of some classes of pesticides to prevent damage to the biosphere show the need for developing new strategies for crop protection (European Commission, 2013; Green, 2018).

As discussed in Section 1.2, trichomes are a key part of the defence system that protects plants from pests by acting as a physical barrier, preventing access by obstructing insects or even secreting repellents. A transgenic line of *Brassica napus* (rapeseed), which overexpresses *GL3* from *Arabidopsis*, was found to drive the formation of supernumerary trichomes on leaves and stems. This increase in trichome formation was then shown to lead to more pest-resistant plants (Gruber et al., 2006; Soroka et al., 2011). However, in addition to being transgenic, this line had considerable drawbacks: the plants showed stunted growth and matured later than non-transgenic *B. napus* (Alahakoon et al., 2016). Also, the increased trichome density was limited to stems and leaves, while fruits and seedpods, which are often the most economically important part of the plant, were not affected. In the case of *B. napus* and *B. rapa* (turnip, field mustard and others), flea beetles cause massive damage to the seedpod by feeding on the epidermal surfaces. Other pests such as cabbage seedpod weevil (*Ceutorhynchus obstrictus*) lay their eggs into the seedpods of *B. napus* and *B. rapa* with the larvae feeding on the growing seeds causing 20-30% loss (Briar et al., 2019; Shpanev and Moseyko, 2021).

In some species however, including several varieties of soybean (*Glycine max*) and white mustard (*Sinapis alba*, a member of the Brassicaceae family), trichomes are found naturally occurring on gynoecia and the seedpods (Bodnaryk, 1996; Lam and Pedigo, 2001). It has been shown that these trichomes confer protection by acting as a physical

barrier for pests as when 'shaved' away the glabrous seed pods were fed on to a greater level (Bodnaryk, 1996).

As the genus *Brassica* is closely related to the genus *Arabidopsis*, it is possible that the genetic modulation found to lead to trichome formation in *Arabidopsis* (see work described in Section 1.3) could be translated directly to *Brassica* crops such as *B. napus* or *B. rapa*. For *B. rapa*, a collection of ethyl methanesulphonate (EMS) generated (non-transgenic) mutant lines is available that could be used to recreate the *ag try cpc* phenotype (Figure 1.8). Critically in the case of the gene perturbations detailed in Section 1.3, the highest level of protection is given to the gynoecium and fruit of the plant protecting the most agriculturally valuable part of the plant. The translation of this work could be complicated by the fact there has been an additional polyploidy event in the ancestry of the *Brassica* lineage after its divergence from the *Arabidopsis* lineage. The effect that polyploidy events have on the gene numbers of plants will be discussed in the section that follows.

1.5 Polyploidy in Plants

Polyploidy, which is a common phenomenon among angiosperm, occurs when two or more genomes undergo fusion in a nucleus (Wendel, 2000; Wang et al., 2012). Polyploids that form as a result of one chromosome set duplicating are known as autopolyploids and those that form when chromosome sets from different species merge are known as allopolyploids (Wolfe, 2001). Many agriculturally important crops are stable polyploids such as wheat, cotton, potato and oat. Even most diploid plants that are extant today are paleopolyploids that have undergone diploidization over time (Blanc and Wolfe, 2004b).

Following an autopolyploidy event, all the genes in the genome of an organism are

duplicated. As summarised in Figure 1.10, there are several outcomes that are possible for the duplicated genes that are also known as paralogs. Because after a duplication event, there is now a redundant copy, it is possible for one of the paralogs to accumulate deleterious mutations to the point where it ceases to function and becomes a pseudogene. This has been shown to be the most likely outcome occurring in approximately 99% of cases (Clark and Donoghue, 2018).

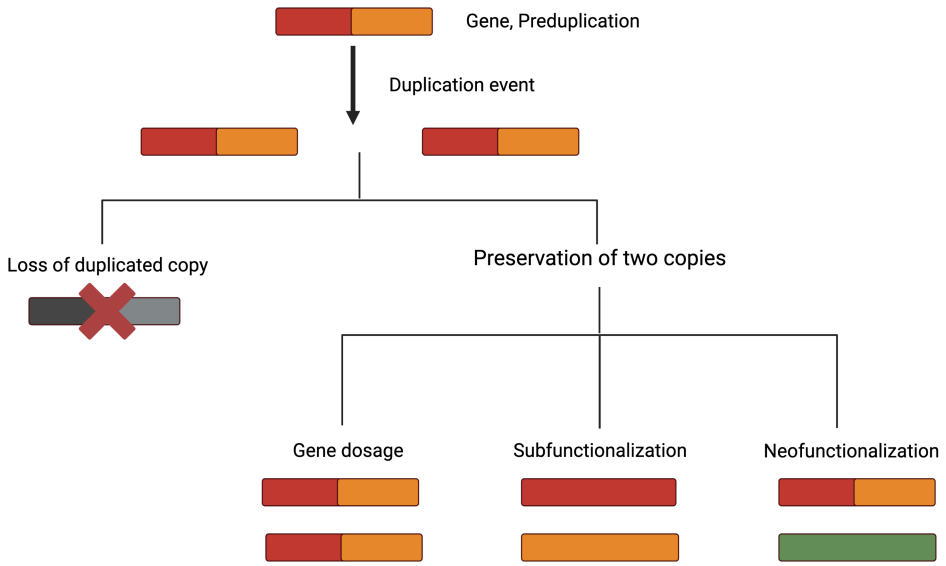

Figure 1.10: Fates of duplicate genes. Schematic summarising the possible outcomes for a gene following a duplication or polyploidy event. A gene represented in red and orange is duplicated, producing two daughter genes. The most likely outcome is the loss of one of the copies. Other outcomes involve the retention of both daughter genes. Both copies could be retained due to gene dosage effects. The gene could undergo subfunctionalization where the original function of the ancestral gene is split between the two daughter genes. A process known as neofunctionalization could also occur which sees one daughter gene retain the function of the ancestral gene whilst the other gains a new function. Figure generated with Biorender.com.

Alternatively, both copies are retained by positive selection due to gene dosage effects. In fact, it has been shown that genes involved in DNA replication and transport processes and genes with regulatory functions are retained at a higher rate compared to other genes in the genome (Blanc and Wolfe, 2004a; Jiang et al., 2013). Gene pairs may also undergo subfunctionalization where both daughter genes are needed to carry out the function of the ancestral gene - either due to any functions being split between

the two paralogs or due to divergence in expression levels or domains (Birchler and Yang, 2022). For example, there is evidence of subfunctionalization of genes involved in flowering time such as *SOC1* and *VERNALIZATION INSENSITIVE 3* in *B. napus* (Schiessl, 2020). Arguably the least likely outcome is neofunctionalization where one of the duplicates retains the function of the ancestral gene whilst the other genes evolve a completely new function (Wendel, 2000; Panchy et al., 2016). An example of neofunctionalization in plants is the evolution of two genes involved in pigment synthesis which are unique to the flowering plant order Caryophyllales - 4,5-dioxygenase which is involved in the production of betalamic acid and CYP76AD1, a cytochrome P450 gene that catalyzes the formation of cyclo-DOPA. These are key precursors to betalain pigments which result in the orders vibrantly coloured petals (Brockington et al., 2015).

1.5.1 Polyploidy in the Genus *Brassica*

As mentioned in Section 1.4, *Brassica* is an agriculturally important genus that like *Arabidopsis* is part of the Brassicaceae family. It is estimated that the two genera diverged from each other approximately 29.5 million years ago (Song et al., 2021). There have been at least three polyploidy events in the ancestral line of the Brassicaceae family - commonly referred to as the α, β and γ events - thought to have occurred 47, 124 and 130 Mya respectively (Figure 1.11) (Bowers et al., 2003; Panchy et al., 2016; Teng et al., 2022). The γ hexaploidy event occurred early in the core eudicot lineage while β and then α tetraploidy events are common to the order of Brassicales (Bowers et al., 2003; Vekemans et al., 2012). A whole-genome triplication event occurred after the divergence of *Arabidopsis* and *Brassica* species approximately 23 Mya (Lysak et al., 2005; Zhang et al., 2019). This is observed in the fact that the *Arabidopsis* genome has been determined to be 157 Mbp whilst the diploid *Brassica* species are much larger ranging from 529 Mbp to 696 Mbps (Johnston et al., 2005). There is evidence for preferential retention of genes involved in stress response, signalling or transcription regulation indicating that genes that confer robustness are retained at a higher level (Fomeju et al.,

2015; Song et al., 2020; Zhang et al., 2021).

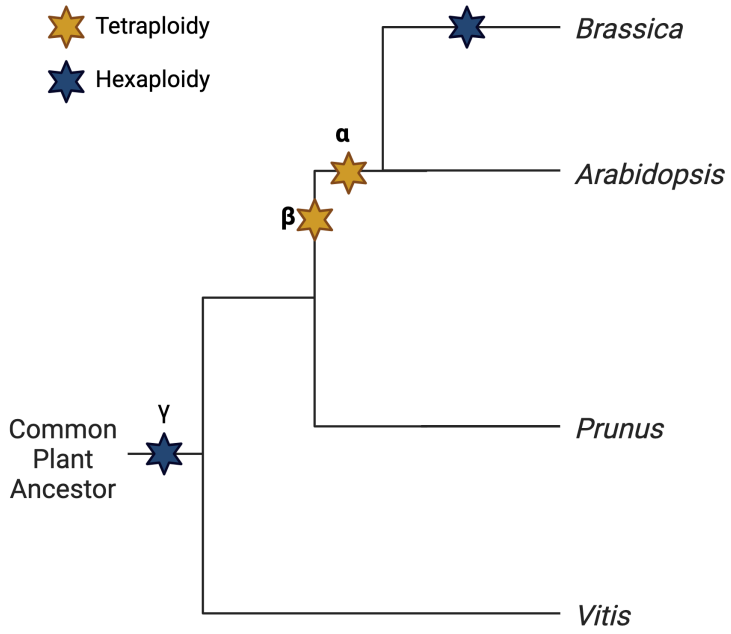

Figure 1.11: Phylogeny of *Brassica*, *Arabidopsis* and their close relatives with polyploidy events indicated. Polyploidy events in the plant lineage are indicated by star shapes on the phylogenetic tree with tetraploidy events indicated in yellow and hexaploidy events indicated in dark blue. The γ event is common to the core eudicot lineage and includes *Vitis* and *Prunus* genera. The β and α polyploidy events are common to the order of Brassicales. There was an additional whole genome triplication event after the divergence of the *Arabidopsis* and *Brassica* species. Figure generated using Biorender.com

There are six major species in the genus *Brassica* and their evolutionary relationship are summarized by the "Triangle of U" (Figure 1.12) (Nagaharu, 1935). According to this model, species in the genus are either diploid or allotetraploid as a consequence of interspecies hybridization between diploids. Specifically, it was proposed that the allotetraploids *B. napus* (AACC, 2n = 4x = 38), *B. juncea* (AABB, 2n = 4x = 36) and *B. carinata* (BBCC, 2n = 4x = 34) are derived from the diploid species *B. rapa* (AA, 2n = 20), *B. oleracea* (CC, 2n = 18) and *B. nigra* (BB, 2n = 16). This model was originally proposed in 1935 and has been confirmed in recent years as the complete genomes of the six major *Brassica* species have become available (Kim et al., 2018; Song et al.,

2021).

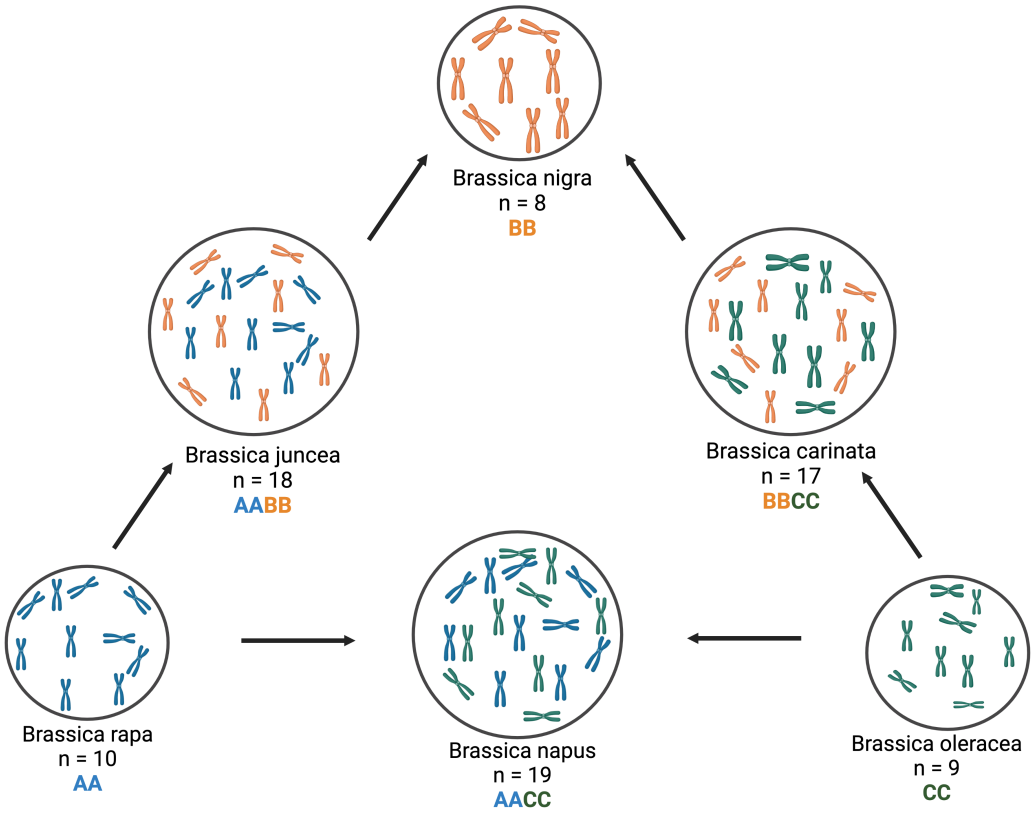

Figure 1.12: The of triangle of U. Schematic of the Triangle of U which explains the evolutionary relationship between the six major species of the genus *Brassica*. The allotetraploid *B. napus* (AACC) is derived from the diploids *B. rapa* (AA) and *B. oleracea* (CC). *B. juncea* (AABB) is derived from *B. rapa* (AA) and *B. nigra* (BB) while *B. carinata* (BBCC) is derived from *B. oleracea* (CC) and *B. nigra* (BB). The number of chromosomes for each species is indicated. Figure generated with Biorender.com.

1.5.2 Translating Research Findings from *Arabidopsis* to *Brassica* Crops

As outlined at the beginning of this chapter, research using *Arabidopsis* as a model organism has led to fundamental new insights into the genetic mechanisms controlling plant development. Because of the many challenges modern agriculture faces in light of problems such as climate change and pesticide resistance, it appears important to use the new knowledge that has been generated for improving crop plants. However, such trans-

lational research can be difficult if the crop species that is being worked on is distantly related to *Arabidopsis* and shows substantial differences in its genetic makeup. In order to translate the work previously conducted in our laboratory on floral trichomes in *Arabidopsis* and given the high degree of relatedness of *Arabidopsis* to *Brassica* species, it was decided to target a crop in the genus *Brassica*. As discussed in Section 1.5.1, three of the six major species of *Brassica* are allotetraploids (*B. napus*, *B. juncea* and *B. carinata*) and were ruled out of use due to potential complications caused by a complex genome structure.

Of the remaining three major *Brassica* genus members, *B. rapa* was chosen for use in the course of this study. *B. rapa* is a crop that is cultivated worldwide in many forms including vegetable crops such as turnip and pak-choi and as an oil crop which alongside other *Brassica* oil crops provides 12% of world vegetable oil (Wang et al., 2011). An extensive and well-characterized collection of non-transgenic mutant lines for *B. rapa* is available whilst no similar catalogue exists for the other diploids *Brassica oleracea* or *Brassica nigra* (Stephenson et al., 2010). As the aim of the overall work would involve generating plants with trichomes on the gynoecia and flowers and testing the levels of pest protection they may provide with the aim of putting the lines into wide agricultural use, it was essential that non-transgenic lines were used in the course of this work due to the restrictions on the growth and use of GM plants within the EU. Flowers of *B. rapa* plants are very similar in structure and composition to those of *Arabidopsis* with four sepals present in the outer whorl, four yellow petals in the second whorl, six stamens - two short and four tall - in the third whorl and two carpels fused to form the gynoecium in the fourth whorl (Figure 1.13).

Despite these advantages, a factor that often complicates the translation of research findings from *Arabidopsis* to *B. rapa* is the additional polyploidy event in the *Brassica* genus (see Section 1.5.1) that led to increased gene numbers and to paralogous genes that may have diverged in function. Thus, re-creating findings from *Arabidopsis* to *B.*

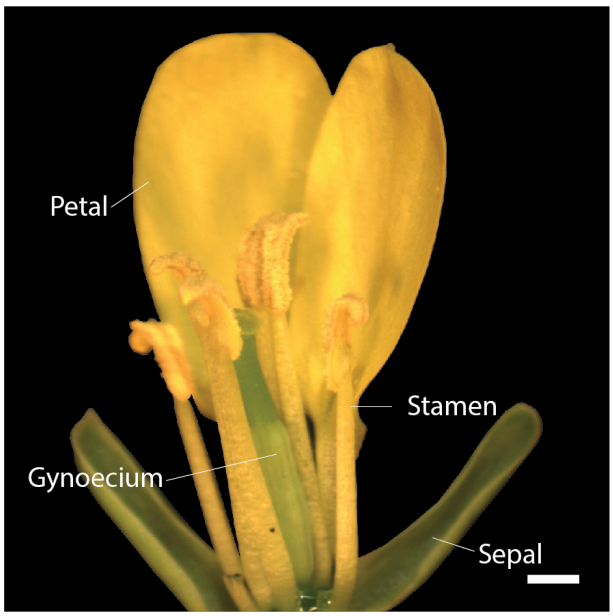

Figure 1.13: Flower of *Brassica rapa* var. Yellow Sarson. Light microscope image of a flower of *B. rapa* R-o-18. The flowers of *B. rapa* are similar in structure to those of *Arabidopsis* consisting of four sepals in the first whorl, four yellow petals in the second whorl, six stamens and two carpels fused to form the gynoecium. Some sepals and petals were removed to visualise the internal floral organs. Scale bar: 1 mm.

---

rapa (and other *Brassicas*) can be difficult if orthologous genes in *Arabidopsis* and *B. rapa* cannot be readily identified. In addition, the higher gene number in *Brassicas* can substantially increase the degree of functional redundancy between genes, often requiring the genetic manipulation of more *B. rapa* genes than in *Arabidopsis*. This, together with the much longer generation time of *Brassica* crops, makes translational research even in the case of these closely related species challenging.

1.6 Aims of the Thesis

As described in the preceding sections, the main aim of the work described in this thesis was to translate the findings on the regulation of trichome initiation in *Arabidopsis* flowers described by Ó'Maoiléidigh et al. (2018) to *B. rapa* as a potential new avenue for crop protection. To this end, a weak *ag* mutant line that retains fertility would have to be isolated. This line would then provide a sensitised background that could be combined

with mutant lines for *try* and *cpc* in order to generate plants that could possibly develop trichomes on seed pods. With the knowledge that a whole genome triplication event has occurred after the divergence of *Arabidopsis* and *Brassica* species, it is probable that there are multiple homologs of genes present in *B. rapa* that are present in single copies in *Arabidopsis*. One aim of my work was therefore to identify the homologs of *AG*, *CPC* and *TRY* and to functionally characterize them. Specifically, their degree of functional redundancy and any evidence for a loss of function or subfunctionalization should be investigated. Ultimately, the aim was to generate the higher-order mutants needed to replicate the floral trichome phenotypes observed in *Arabidopsis*.

In a second project, the role of LFY and AP1 in controlling the levels of the plant hormone auxin in the emerging floral meristem in *Arabidopsis* was investigated using experimental techniques examining how their activation impacts the expression of genes known to be involved in auxin synthesis and transport as well and the quantification of their impact on auxin levels in the floral meristem.

Materials and Methods

2.1 Strains and Growth Conditions

2.1.1 Bacterial Strains

The following bacterial strains were used in the described experiments:

- *Escherichia coli* strain stbl2 (F- endA1 glnV44 thi-1 recA1 gyrA96 relA1 Δ (lac-proAB) mcrA Δ (mcrBC-hsdRMS-mrr) λ - (Invitrogen)
- *Agrobacterium tumefaciens* strain pGV2260 (McBride and Summerfelt, 1990)

Growth conditions for *E. coli* and *Agrobacterium* are described in Section 2.2.7 and Section 2.2.9 respectively.

2.1.2 Plant Lines

2.1.2.1 *Arabidopsis thaliana*

Previously published lines used in this study are:

- *try-29760* (Tominaga et al., 2008)
- DR5rev::GFP (Friml et al., 2003)
- DR5rev::GUS (Benková et al., 2003)
- *ag-1* (Bowman et al., 1989)
- p35S::AP1-GR *ap1-1 cal-1* (Wellmer et al., 2006)
- p35S::LFY-GR *ap1-1 cal-1* (Goslin et al., 2017)

- *cpc-1* (Wada et al., 1997)

Lines generated during the course of this study are as listed in Table 2.1.

Table 2.1: *Arabidopsis* lines generated during the course of this study. Constructs and the mutants/wild-type accessions that were transformed are listed. Lines generated by crossing existing lines are also shown.

Transformed into *ag-1* & L-*er*	
pAG:*AtAG*	p35S:*AtAG*
pAG:*BrAG1*	p35S:*BrAG1*
pAG:*BrAG2*	p35S:*BrAG2*
Transformed into *try-29760* & Col-0	
pTRY:*AtTRY*	p35S:*AtTRY*
pTRY:*BrTRY1*	p35S:*BrTRY1*
pTRY:*BrTRY2*	p35S:*BrTRY2*
pTRY:*BrTRY3*	p35S:*BrTRY3*
Transformed into *cpc-1* & Ws	
pCPC:*AtCPC*	p35S:*AtCPC*
pCPC:*BrCPC1*	p35S:*BrCPC1*
pCPC:*BrCPC2*	p35S:*BrCPC2*
Generated through crosses	
DR5rev:GFP p35S:AP1-GR *ap1-1 cal-1*	
DR5rev:GFP p35S:LFY-GR *ap1-1 cal-1*	
DR5rev:GUS p35S:AP1-GR *ap1-1 cal-1*	
DR5rev:GUS p35S:LFY-GR *ap1-1 cal-1*	

2.1.2.2 *Brassica rapa*

For these experiments the following accessions of *B. rapa* were used:

- R-o-18 (donated by Dr Emmanuelle Graciet, Maynooth University) (Stephenson et al., 2010)
- C634 (obtained from the Tohoku University Brassica Seed Bank) (Li et al., 2009)

The mutant *B. rapa* lines used in these experiments are Targeting Induced Local Lesions in Genomes (TILLING) mutant lines obtained from RevGenUK at the John Innes Centre (Stephenson et al., 2010). The lines used in the course of this study are listed in Table 2.2.

Table 2.2: *B. rapa* TILLING lines used in this study. Summary of the TILLING lines used in this work. This table includes the *Arabidopsis* gene name, the *B. rapa* gene for which the line contains a mutation, the ID used by the supplier RevGenUK, the abbreviation used for each line, and a description of the mutation that defines the line.

At Gene	Gene	ID	Abbr.	Mutation
AG	Bra012564	JI31652	*brag1-1*	stop gained, 4th amino acid converted to stop
		JI30258	*brag1-2*	missense, 418 G to A, Glu to Lys
	Bra013364	JI30013	*brag2-1*	mutation at splice donor site, G to A at base 231
		JI31547	*brag2-2*	missense, 415 G to A, Glu to Lys
CPC	Bra004539	JI32146	*brcpc1-1*	stop gained, 79th amino acid converted to stop
	Bra039283	JI32170	*brcpc2-1*	stop gained, 24th amino acid converted to stop
TRY	Bra022637	JI31767	*brtry1-1*	missense, 100 G to A, Glu to Lys
	Bra026297	JI30284	*brtry2-1*	stop gained, 20th amino acid converted to stop
	Bra039065	JI33214	*brgl1b-1*	stop gained, 83rd amino acid converted to stop

The M3 plants of the lines in Table 2.2, generated as described in McCallum et al., 2000, were used in the work described.

2.1.3 Plant Growth Media

Plants were either grown on soil made up of compost, vermiculite, and perlite (NAD, Dublin, Ireland) mixed in a 3:1:1 ratio, or on MS (Murashige and Skoog) agar consisting of 2.2 g/L MS salts (Sigma-Aldrich), and 6 g/L agar (Fisher Scientific) adjusted to pH5.7 using KOH. Plant growth conditions are described in Section 2.6.

2.1.4 Bacterial Growth Media

All bacteria were grown in LB (Luria-Bertani) medium or on LB agar made up of 10 g/L NaCl (Fisher Scientific), 10 g/L Bacto-tryptone (Fisher Scientific), and 5 g/L Bacto-yeast extract (Fisher Scientific). Fifteen g/L of agar was added to make LB agar.

2.2 Cloning

2.2.1 RNA Extraction and cDNA Synthesis

The Spectrum Plant Total RNA kit (Sigma-Aldrich) was used to perform total RNA extraction from plant tissue according to the manufacturer's instructions. Purified RNA was stored at -80°C. Copy DNA (cDNA) was generated by reverse transcription using RevertAid H Minus Reverse Transcriptase (Thermofisher) and oligo$(dT)_{18}$ primers (Thermofisher) as per the manufacturer's instructions.

2.2.2 Preparation of Plasmid DNA

Cultures of transformed *E. coli* colonies were grown in 5 mL of LB medium with the appropriate selective antibiotic at 37°C with constant agitation overnight. The cultures were then centrifuged at 4500 rpm for 10 minutes to pellet the cells and the plasmids were extracted using Plasmid DNA Mini Kit (O.M.E.G.A) as per the manufacturer's instructions.

2.2.3 PCR Product Purification

PCR products were purified when necessary with the EZNA Cycle Pure Kit (O.M.E.G.A) following the manufacturer's instructions with the following changes. The columns were centrifuged at 15,000 rpm for 5 minutes post-washing steps to ensure any residual ethanol was removed. The columns were transferred to a clean 1.5mL tube, placed in a heat block at 42°C, and incubated for 10 minutes to further aid drying.

2.2.4 Agarose Gel Electrophoresis

All DNA for analysis was run out on 1% or 2% agarose gels except for genotyping using derived cleaved amplified polymorphic sequence (dCAPS)(see Section 2.3.4) where the products were run on a 3% gel. Agarose was resuspended in 1X TAE buffer (see Appendix 7.2) and a 1 in 10,000 dilution of SYBR SAFE (ThermoFisher). Electrophoresis

of gels was carried out in 1X TAE buffer at a current of 120 V. A UV transilluminator/gel imaging system was used to visualize and photograph the gels.

2.2.5 Agarose Gel Extraction of DNA

Digested plasmids and PCR products were extracted from agarose gels using the EZNA Gel extraction kit (O.M.E.G.A) following the manufacturer's instructions with the amendments as described in Section 2.2.3.

2.2.6 Quantification of Nucleic Acids

The concentration of DNA and RNA was determined using a NanoDrop ND-1000 spectrophotometer. The instrument was initialized with ultrapure water and then blanked with the same solution that the sample was suspended in.

2.2.7 Preparation and Transformation of Competent *Escherichia coli*

The protocol as described by Inoue et al., 1990 was used to prepare competent stbl2 cells. Transformations were carried out by incubating 100 μL of competent cells with the plasmid on ice for 10 minutes. The cells were then incubated at 42°C for 45 seconds. They were placed back on ice for 2 minutes and then incubated at 37°C with 1 mL of LB medium for an hour to allow the cells to recover. The cells were then plated on an LB agar plate containing the relevant antibiotic for selection and incubated at 37°C overnight.

2.2.8 Colony PCR of *Escherichia coli*

Colony PCR was used to confirm the presence of plasmids containing the construct of interest in *E. coli* colonies following transformation. A single colony was suspended in 20 μL of LB and 2 μL of this solution was used in subsequent genotyping PCRs, 20 μL total.

2.2.9 Preparation and Transformation of Competent *Agrobacterium tumefaciens*

A 5 mL culture of *Agrobacterium* was grown at 28°C with constant agitation overnight. A 250 mL culture supplemented with 0.2 g/L $MgSO_4$ was inoculated with the 5 mL culture. This culture was grown at 28°C with constant agitation until an OD_{600} of between 1 and 2.5 was reached. The culture was centrifuged at 4,000 *g* for 10 minutes to pellet the cells. The pellet was resuspended in 20 mL ice-cold 10 mM $CaCl_2$ to wash. The cells were then centrifuged at 4,000 *g* for 10 minutes and the pellet was resuspended in 4 mL 10 mM $CaCl_2$. Cells were aliquoted in 200 μL volumes and snap-frozen in liquid N_2. To transform, an aliquot of the cells was thawed and incubated with the plasmid for 5 minutes in liquid N_2. The cells were then thawed, 1 mL of LB medium was added and the cells were allowed to recover at 28°C under constant agitation for 4 hours. After recovering the cells were centrifuged at 5,000 rpm for 2 minutes, 800 μL of the supernatant was removed and the pellet was resuspended in the remaining LB medium. Cells were then plated on LB agar containing rifampicin (50 μg/mL), carbenicillin (100 μg/mL) and the selective antibiotic for the plasmid being transformed. The plates were incubated at 28°C for 2-3 days.

2.2.10 Colony PCR of *Agrobacterium tumefaciens*

Single colonies of *A. tumefaciens* were picked and suspended in 20 μL lysis buffer (see Appendix 7.2). The solution was heated at 95°C for 10 minutes to lyse the cells and then centrifuged at 14,000 rpm for 3 minutes to pellet cellular debris. Two μL of this solution was used in subsequent genotyping PCR using a total reaction volume of 20 μL total.

2.2.11 Cetyltrimethyl Ammonium Bromide (CTAB) Genomic DNA Extraction

Genomic DNA was extracted from *B. rapa* or *Arabidopsis* using this method. Approximately 100 mg of tissue was snap-frozen in liquid N_2 and then ground to a fine powder using a pestle. Five hundred μL of preheated CTAB Buffer (see Appendix 7.2 for composition) was added to the tissue and was incubated at 65°C for 15 minutes. Five hundred μL of phenol/chloroform/isoamyl (48:48:4) (Roth) was then added and vortexed thoroughly. The mixture was transferred to a phase lock tube (VWR International) and spun down in a micro-centrifuge at maximum speed for 3 minutes. The extraction was repeated. Three-hundred fifty μL of isopropanol was added to precipitate DNA, the tube was inverted to mix, and was incubated at room temperature for 10 minutes. The sample was then spun at maximum speed in a microcentrifuge for 15 minutes at room temperature. The pelleted DNA was washed with 70% (v/v) ethanol by centrifuging at 14,000 rpm for 3 minutes. The pellets were dried either on a heat block or in a vacuum centrifuge and then suspended in 50 μL TE buffer (see Appendix 7.2).

2.2.12 Dephosphorylation of DNA

FastAP (ThermoFisher) was added during restriction enzyme digestion in order to 5' dephosphorylate the plasmid DNA. This was carried out to prevent the recircularization of the plasmid during subsequent ligation reactions. One μL of the enzyme was added directly to the digest, incubated at 37°C for 1 hour, and then inactivated by incubating at 75°C for 10 minutes.

2.2.13 DNA Ligation

DNA fragments that had been digested with the same restriction enzymes were ligated using T4 DNA ligase (NEB). Ligation reactions were carried as follows: 100 ng plasmid, three molar digested insert DNA, 1.5 μL 10X T4 Ligase Buffer (NEB), 1 μL T4 Ligase (NEB), 1.5 μL 10 mM ATP (ThermoFisher) and ddH_2O to 15 μL. The reaction

was incubated at 22°C for at least 20 minutes and then heat-inactivated by incubating at 65°C for 10 minutes. Five μL of the ligation reaction was transformed into *E. coli* stbl2 cells as detailed in Section 2.2.7.

2.2.14 GreenGate Cloning

The GreenGate cloning system was used to construct plasmids generated during the course of this study (Lampropoulos et al., 2013). This system is based on Golden Gate cloning and allows for the assembly of six modules - promoter, N-tag, CDS, C-tag, terminator, and resistance cassette into pUC19-based entry vectors. These entry vectors are then used to assemble a pGREEN-IIs-based destination vector. Each module is flanked by overhangs that enable cloning into the entry vector and assembly in the correct order when cloning the final vector. A GreenGate entry reaction consists of 1.5 μL 10X T4 Ligase Buffer(NEB), 1 μL T4 Ligase (NEB), 1 μL *BsaI* (NEB), 1.5 μL 10mM ATP (ThermoFisher), 1.5 μL 10X BSA (NEB), 100 ng of the entry vector, insert DNA and ddH_2O to 15 μL. The reaction conditions were [37°C 2 minutes, 16°C 2 minutes] x 30 followed by heat inactivation at 80°C for 5 minutes. Five μL of the reaction was used to transform 100 μL of stlb2 *E. coli* cells as described in Section 2.2.7. To assemble the destination vector, 100 ng of each entry and the destination vector, 1.5 μL 10X T4 Ligase Buffer(NEB), 1 μL T4 Ligase (NEB), 1 μL *BsaI* (NEB), 1.5 μL 10mM ATP (ThermoFisher), 1.5 μL 10X BSA (NEB) and ddH_2O to 15 μL. The reaction conditions were [37°C 2 minutes, 16°C 2 minutes] x 50 cycles followed by heat inactivation at 80°C for 5 minutes. Five μL of the ligation reaction was transformed into *E. coli* stbl2 cells as detailed in Section 2.2.7.

2.2.15 Sequencing

All sequencing was carried out by Eurofins GATC. All DNA files were constructed with the software ApE (https://www.apesoftware.com/). ApE was used to compare the DNA files to the sequencing files provided by Eurofins GATC.

2.2.16 Design of Oligonucleotides

Oligonucleotides were designed using the "Design Primer" feature in the ApE files of the sequence of interest. All were designed to avoid the occurrence of secondary structures and dimerization, have a GC content of between 40-60%, and a melting temperature between 50-60°C.

2.3 Genotyping

2.3.1 "Edwards" DNA Preparation

Genomic DNA extractions for the purposes of genotyping were carried out using a published protocol (Edwards et al., 1991). A small section of leaf tissue was ground using a pestle in 400 μL extraction buffer (see Appendix 7.2). The samples were centrifuged at 14,000 rpm to pellet the debris and 300 μL of the supernatant was transferred to a new tube. Three hundred μL of isopropanol was added and tubes were centrifuged at 14,000 rpm for 5 minutes in order to pellet the DNA. The supernatant was removed and the pellet was washed with 70% (v/v) ethanol by centrifuging at 14,000 rpm for 3 minutes. The supernatant was decanted and tubes were dried either on a heat block or in a vacuum centrifuge. Once dry, the pellet was suspended in 50 μL TE buffer by briefly vortexing.

2.3.2 Genotyping of TILLING Lines

Primer pairs were designed as described in Section 2.2.16 to create an amplicon that contained the SNP that defined the TILLING line (Stephenson et al., 2010). PCRs using the primer pairs as listed in Table 2.3 were carried out using the Phusion enzyme (ThermoFisher) and the PCR product was cleaned using O.M.E.G.A PCR Cycle Pure kit as per the manufacturer's instructions as described in Section 2.2.3. The products were sent to GATC Eurofins for Sanger sequencing. Primer sequences are listed in Table 7.1 in the Appendix.

Table 2.3: Primer pairs used to genotype TILLING lines. Line identifiers from RevGenUK, allele names, and primers used are shown. Primer sequences are listed in Table 7.1 in the Appendix.

RevGenID	Allele	Primer Pair
JI31652	*brag1-1*	oCF15/31
JI30258	*brag1-2*	oCF30/16
JI30013	*brag2-1*	oCF27/32
JI31547	*brag2-2*	oCF28/18
JI32146	*brcpc1-1*	oCF33/34
JI32170	*brcpc2-1*	oCF35/36
JI31767	*brtry1-1*	oCF37/38
JI30284	*brtry2-1*	oCF39/40
JI33214	*brgl1b-1*	oCF511/512

2.3.3 Double-Mismatch Allele-Specific Genotyping

DMAS genotyping consists of two parallel qPCR reactions as described by Lefever et al. 2019. Each reaction contained a common wild-type (WT) primer and two modified allele-specific (AS) primers. The 3' end of the AS primers overlapped with the SNP of interest. An additional mismatch (double mismatch) was also introduced to the AS primers in order to increase the sensitivity of the assay. Each DMAS primer was referred to as a 'match' for the corresponding allele and a 'mismatch' for the other allele. Primers used in DMAS genotyping are listed in Table 7.1 in the Appendix. Edwards prepped genomic DNA (prepared as described in Section 2.3.1) was used as the template in a qPCR reaction consisting of 1 μL gDNA, 1 μL of oligos 10 μM each, 3 μL H_20 and 5 μL 2x SYBR Green Master 1 (Roche). Primer pairs were designed to have a Tm of between 60°C and 63°C. The qPCR reaction was run using the LightCycler480 (Roche) and the software was used to calculate the Cp value of the reaction. By subtracting the Cp values for the 'match' and the 'mismatch' reaction the genotype of the sample can be concluded.

- Cp Match - Cp Mismatch ≈ 0 the sample is heterozygous for the SNP of interest.
- Cp Match - Cp Mismatch < 0 the sample is homozygous for the 'match' allele
- Cp Match - Cp Mismatch > 0 the sample is homozygous for the 'mismatch' allele.

Table 2.4: DMAS primers used for genotyping each TILLING line. Allele names, the common primer and the allele-specific (wild-type and mutant) primers used for DMAS genotyping are shown. Primer sequences are listed in Table 7.1 in the Appendix.

Allele	Common Primer	WT Primer	Mutant Primer
brag1-1	oCF526	oCF524	oCF525
brag1-2	oCF532	oCF530	oCF531
brag2-1	oCF54	oCF547	oCF548
brag2-2	oCF544	oCF542	oCF543
brtry1-1	oCF64	oCF577	oCF578
brtry2-1	oCF146	oCF551	oCF552
brcpc1-1	oCF555	oCF553	oCF554
brcpc2-1	oCF72	oCF556	oCF557
brgl1b-1	oCF560	oCF558	oCF559

2.3.4 dCAPS genotyping

Genotyping of *ag-1* was carried out using derived Cleaved Amplified Polymorphic Sequence (dCAPS) analysis as described by (Neff et al., 1998). During PCR, a restriction enzyme recognition site which includes the SNP of interest was introduced into the amplicon by a primer that contains at least one mismatch to the template. The PCR product was then digested using the restriction enzyme matching the recognition site. The digest was then analyzed on an agarose gel. In the case of *ag-1*, the primers used are oCF509/510 (see Table 7.1 Appendix) and the amplicon produced was 144bp in length. The wild-type amplicon remained uncut by the restriction enzyme *AflII* and the amplicon containing the mutant SNP was cut producing a fragment of 114bp and 30bp.

2.4 Rapid Amplification of cDNA Ends-PCR

Rapid Amplification of cDNA Ends (RACE)-PCR was performed to determine the 3'UTR sequence of the mRNA of the *B. rapa* homologs of *AG* (Section 3.2.3), *CPC* and *TRY* (Section 4.2.4). To this end, a previously published protocol (Scotto-Lavino et al., 2007) was followed. Briefly, cDNA was generated using a modified oligo(dT) primer which contained restriction digest recognition sites. A PCR was carried out using this cDNA as a template with a gene-specific primer and the modified oligo(dT)

primer to generate an amplicon from the gene of interest which contains the 3'UTR. The 3'RACE product along with pUC19 was digested with HindIII (ThermoFisher) restriction enzyme and the two were ligated together using T4 ligase (Thermofisher). The pUC19 plasmid containing the 3' RACE fragment was sent to Eurofins for sequencing.

2.5 Quantitative PCR

Primer pairs were designed to have a Tm of between 60°C and 63°C and have no more than 1°C difference between the pair. Primer pairs were also designed to span an exon-exon junction. The relative enrichment of cDNA was quantified using the LightCycler 480 (Roche) with SYBR Green Master 1 (Roche). One reaction consisted of 5 μL of 2x SYBR Green Master 1, 2 μL cDNA (diluted 1 in 5), 1 μL of 10 μM primers, and 2 μL of PCR-grade water. An elongation time of 60 seconds per 1 kb of DNA was set. Annealing temperatures were set between 58°C and 60°C depending on the primer pair in use. The enrichment was calculated relative to the reference gene. The relative enrichment for each gene was calculated as follows:

$$2^{\text{CP(REF)-(GENE)}}$$

For *Arabidopsis*, the reference primers used were oCF203/204 (AT1G13320) (see Table 7.1 in the Appendix) and for *B. rapa*, the reference primers oCF100/101 (Bra005395 homolog of AT1G75780) (see Table 7.1 in the Appendix) were used. The Cp is the crossing point as calculated by the LightCycler480. The Cp (REF) is the value for the chosen reference gene and the Cp (GENE) is the value for the gene of interest.

2.6 Plant Growth Conditions

2.6.1 Growth Conditions

All *Brassica* plants were sown and placed at 4°C in the dark for 3 days. Then they were grown at 21°C under white fluorescent LEDs with constant light until 7 days after germination. The plants were then grown at 21°C under white fluorescent LEDs with 20 hours of light and 4 hours of darkness. In the case of *Arabidopsis*, the plants were grown under constant white fluorescent LED lights at 21°C.

2.6.2 Crossing Plant Lines

In the case of *Brassica*, the sepals of the flowers of maternal plants were peeled back and the stamens were removed. A stamen from the paternal plant was used to pollinate the maternal plant. Crosses were reciprocally out both ways where possible. In the case of *Arabidopsis*, a flower had the stamens removed before anthesis. The flower was manually pollinated using a stamen from the other line and the plants of the F1 were genotyped for the genotype or construct of interest.

2.7 Plant Transformation and Selection

2.7.1 Transformation of *Arabidopsis*

The floral dip protocol as described by Clough and Bent, 1998 was used to transform *Arabidopsis* (Clough and Bent, 1998).

2.7.2 Seed Sterilisation

Both *Arabidopsis* and *B. rapa* seeds were vapour sterilised in a sealed chamber for four hours using 100 mL bleach (Domestos) with 3 mL concentrated HCl added in order to release chlorine gas. Once sterilised, seeds were then transferred to 0.5 MS agar plates

with the appropriate selective agent. Plates were sealed with surgical tape, kept at 4°C for 3 days, and then transferred to the cold room.

2.7.3 Selection on Soil

Transgenic seeds containing a BAR resistance cassette were planted on soil. Once the seeds germinated they were sprayed with a 0.15% (v/v) ammonium-glufosinate (BASTA^{TM}(BAYER)). The plants were sprayed three times in total with a two-day interval between sprays. Plants that formed true leaves were genotyped for the construct of interest.

2.8 Prediction of Binding Sites

In the case of predicting transcription factor binding sites, JASPAR was used (Castro-Mondragon et al., 2022). JASPAR is an online database of transcription factor (TF) binding profiles. The database stores transcription factors binding profiles as position frequency matrices (PFMs). PFMs summarise the frequency at which each nucleotide at each position in a given transcription factor binding site occurs. PFMs can be used to construct position weight matrices (PWMs). These PWMs can be used to scan a user-input DNA sequence to predict TF binding sites. JASPAR was used to predict transcription factor binding sites for *AG* and *TRY* and their *B. rapa* homologs in Sections 3.2.3 and 4.2.3 respectively. In order to predict AU-rich elements (ARE) binding sites in the 3'UTRs for *AG*, *TRY*, *CPC* and their *B. rapa* homologs the MEME-Suite tool Simple Enrichment Analysis (SEA) was used (https://meme-suite.org/meme/tools/sea). SEA analyses user-inputted DNA sequences and identifies selected motifs that are relatively enriched compared to either an input control sequence or a control sequence consisting of primary sequences that have been shuffled randomly.

2.9 Microscopy

Tissue collection for *Arabidopsis* was carried out using an Olympus SZX7 zoom stereomicroscope as described in Section 2.14.1. All photographs were taken using an OPTIKAM microscope fitted with an OPTIKAM Pro Plus 3-mounted camera.

2.10 Treatments of Lines Encoding Dexamethasone Inducible Promoters

The cauliflower-like structures of p35:LFY-GR *ap1-1 cal-1* and p35S:AP1-GR *ap1-1 cal-1* lines were drenched using plastic pipettes with a solution containing 10 μM dexamethasone (Sigma-Aldrich), 0.01% (v/v) ethanol and 0.015% (v/v) Silwett L-77 (De Sangosse) or with an identical mock solution lacking the dexamethasone. After treatment, the meristematic tissue was collected under the microscope using a blade on liquid nitrogen and stored at -80°C until further use.

2.11 Determination of Free IAA by GC-MS

The quantification of free indole-3-acetic acid (IAA), the most common form of auxin in plant tissues, was carried out by the Pollmann laboratory at the Polytechnic University of Madrid. The amount of IAA present in tissue samples was measured using gas chromatography-mass spectrometry (GC-MS). Gas chromatography is an analytical technique that works by heating a sample in order to separate it into its constitutive parts. A carrier gas carries the molecules through the GC system to the detector which produces a signal allowing for the molecules to be identified. After separation, the components are analysed using mass spectrometry (MS). The components are ionized, the m/z ratio of the component is measured and then the current of the component is detected. Using GC-MS, the IAA in the samples was separated from other components in the tissue and the levels were quantified. The tissue collection and preparation were

carried out by myself. In the case of examining the effect of LFY and AP1 on the levels of auxin in the floral meristem, tissue was collected as described in Section 2.4.1. Mock and dexamethasone-treated samples were collected at 8-hour and 16-hour time points. All samples were collected in quadruplicate and each sample contained 25mg of tissue. The tissue was prepared as per the protocol provided by the Pollmann lab and then sent to Madrid for analysis via GC-MS.

2.12 Image Manipulation and Figure Construction

Adobe Illustrator was used to edit photographs and agarose gel images. Figures were constructed using Adobe Illustrator, PRISM, or Biorender. All DNA alignment figures were made using Jalview software (https://www.jalview.org/) (Troshin et al., 2011).

2.13 Tissue Collection

2.13.1 Floral Induction System Tissue

In the case of both p35S::AP1-GR *ap1-1 cal-1* and p35S::LFY-GR *ap1-1 cal-1*, meristematic tissue was collected under an Olympus light microscope using a sterile No. 4 jeweller's forceps. The tissue was immediately deposited into Eppendorf tubes that were immersed in liquid N_2 and then stored at -80°C until ready for use.

2.13.2 *Brassica rapa* Tissue

Brassica tissue was collected using sterile forceps and then deposited into Eppendorf tubes that were immersed in liquid N_2. Where necessary, floral or leaf tissue was dissected using either a sterile No. 4 jeweller's forceps or a No. 10 scalpel blade (Swann-Morton). The tissue was then stored at -80°C until ready for use.

2.14 Generation of Transgenic Lines

2.14.1 Overexpression Lines

2.14.1.1 pGREEN-IIs p35S::*AtAG*

The CDS of *AtAG* was amplified from cDNA generated from RNA extracted from *Arabidopsis* flowers (accession: L-*er*) using primers oCF415/444 (see Table 7.1 in the Appendix). The sequence amplified began at the start codon of the gene and encompassed the 3' UTR. For the purpose of this cloning, the start codon of *AtAG* was changed from "ACG" to "ATG" to ensure proper expression in the plant. The amplicon was cloned into the GreenGate CDS module entry vector. This plasmid was used with commercially available GreenGate plasmid components (35S promoter module, N-tag dummy module, C-tag dummy module, BASTA resistance cassette module, and terminator module) in a destination reaction to assemble the final product. *Agrobacterium* containing the helper plasmid pSOUP were transformed with the plasmid and used to transform *Arabidopsis* plants. The transformed seedlings were selected using BASTA^{TM} as described in Section 2.7.3.

2.14.1.2 pGREEN-IIs p35S::*BrAG1*

The CDS of *BrAG1* was amplified from cDNA generated from RNA extracted from *B. rapa* flowers (accession: R-o-18). The sequence amplified began at the start codon of the gene and encompassed the 3' UTR using primers oCF409/441 (see Table 7.1 in the Appendix). Once cloned into the GreenGate CDS module entry vector, the *BrAG1* CDS plasmid was used with commercially available GreenGate plasmids components (35S promoter module, N-tag dummy module, C-tag dummy module, BASTA resistance cassette module, and terminator module) in a destination reaction to assemble the final product. *Agrobacterium* containing the helper plasmid pSOUP were transformed with the plasmid and used to transform *Arabidopsis* plants. The transformed seedlings were selected using BASTA^{TM} as described in Section 2.7.3.

2.14.1.3 pGREEN-IIs p35S::*BrAG2*

The CDS of *BrAG2* was amplified from cDNA generated from RNA extracted from *B. rapa* flowers (accession: R-o-18). The sequence amplified began at the start codon of the gene and encompassed the 3' UTR using primers oCF411/442 (see Table 7.1 in the Appendix). Once cloned into the GreenGate CDS module entry vector, the *BrAG2* CDS plasmid was used with commercially available GreenGate plasmids components (35S promoter module, N-tag dummy module, C-tag dummy module, BASTA resistance cassette module, and terminator module) in a destination reaction to assemble the final product. *Agrobacterium* containing the helper plasmid pSOUP were transformed with the plasmid and used to transform *Arabidopsis* plants. The transformed seedlings were selected using BASTA^{TM} as described in Section 2.7.3.

2.14.1.4 pGREEN-IIs p35S::*AtCPC*

The CDS of *AtCPC* was amplified from cDNA generated from RNA extracted from *Arabidopsis* leaf tissue (accession: Col-0) using primers oCF475/499 (see Table 7.1 in the Appendix). The region amplified began at the start codon of the gene and included the 3'UTR. This amplicon was ligated into a GreenGate CDS module entry vector which was subsequently used alongside other GreenGate plasmids (35S promoter module, N-tag dummy module, C-tag dummy module, BASTA resistance cassette module, and terminator module) to assemble the final destination vector. *Agrobacterium* containing the helper plasmid pSOUP were transformed with the plasmid and used to transform *Arabidopsis* plants. The transformed seedlings were selected using BASTA^{TM} (BAYER) as described in Section 2.7.3.

2.14.1.5 pGREEN-IIs p35S::*BrCPC1*

The CDS of *BrCPC1* was amplified from cDNA generated from RNA extracted from *B. rapa* leaf margin tissue (accession: C634) using primers oCF477/500 (see Table 7.1 in the Appendix) which encompassed from the start codon to the end of the 3'UTR. This amplicon was ligated into a GreenGate CDS module entry vector which was subse-

quently used alongside other GreenGate plasmids (35S promoter module, N-tag dummy module, C-tag dummy module, BASTA resistance cassette module, and terminator module) to assemble the final destination vector. *Agrobacterium* containing the helper plasmid pSOUP were transformed with the plasmid and used to transform *Arabidopsis* plants. The transformed seedlings were selected using BASTA^{TM} (BAYER) as described in Section 2.7.3.

2.14.1.6 pGREEN-IIs p35S::*BrCPC2*

The CDS of *BrCPC2* was amplified from cDNA generated from RNA extracted from *B. rapa* leaf margin tissue (accession: C634) using primers oCF479/480 (see Table 7.1 in the Appendix) which encompassed the start codon to the end of the 3'UTR. This amplicon was ligated into a GreenGate CDS module entry vector which was subsequently used alongside other GreenGate plasmids (35S promoter module, N-tag dummy module, C-tag dummy module, BASTA resistance cassette module, and terminator module) to assemble the final destination vector. *Agrobacterium* containing the helper plasmid pSOUP were transformed with the plasmid and used to transform *Arabidopsis* plants. The transformed seedlings were selected using BASTA^{TM} (BAYER) as described in Section 2.7.3.

2.14.1.7 pGREEN-IIs p35S::*AtTRY*

The CDS of *AtTRY* was amplified from cDNA generated from RNA extracted from *Arabidopsis* flowers (accession: L-*er*) using primers oCF481/482 (see Table 7.1 in the Appendix). The amplicon includes the start codon and the 3'UTR of the gene and was ligated into a GreenGate CDS module entry vector. This entry vector was used alongside other GreenGate plasmids (35S promoter module, N-tag dummy module, C-tag dummy module, BASTA resistance cassette module, and terminator module) to assemble the final destination vector. *Agrobacterium* containing the helper plasmid pSOUP were transformed with the plasmid and used to transform *Arabidopsis* plants. The transformed seedlings were selected using BASTA^{TM} (BAYER) as described in Section 2.7.3.

2.14.1.8 pGREEN-IIs p35S::*BrTRY1*

The CDS of *BrTRY1* was amplified from cDNA generated from RNA extracted from *B. rapa* leaf tissue (accession: C634) using primers oCF483/484 (see Table 7.1 in the Appendix). This amplicon was ligated into a GreenGate CDS module entry vector which was subsequently used alongside other GreenGate plasmids (35S promoter module, N-tag dummy module, C-tag dummy module, BASTA resistance cassette module, and terminator module) to assemble the final destination vector. *Agrobacterium* containing the helper plasmid pSOUP were transformed with the plasmid and used to transform *Arabidopsis* plants. The transformed seedlings were selected using BASTA^{TM} (BAYER) as described in Section 2.7.3.

2.14.1.9 pGREEN-IIs p35S::*BrTRY2*

Primers oCF485/486 (see Table 7.1 in the Appendix) were used to amplify the CDS of *BrTRY2* from cDNA generated from RNA extracted from *B. rapa* flower tissue (accession: R-o-18). The amplicon was ligated into the GreenGate CDS module entry vector which was then used alongside other GreenGate plasmids (35S promoter module, N-tag dummy module, C-tag dummy module, BASTA resistance cassette module, and terminator module) to assemble the final destination vector. *Agrobacterium* containing the helper plasmid pSOUP were transformed with the plasmid and used to transform *Arabidopsis* plants. The transformed seedlings were selected using BASTA^{TM} (BAYER) as described in Section 2.7.3.

2.14.1.10 pGREEN-IIs p35S::*BrTRY3*

The CDS of *BrTRY3* was amplified from cDNA generated from RNA extracted from *B. rapa* gynoecium tissue (accession: C634) using primers oCF487/488 (see Table 7.1 in the Appendix). The amplicon was ligated into the GreenGate CDS module entry vector which was then used alongside other GreenGate plasmids (35S promoter module, N-tag dummy module, C-tag dummy module, BASTA resistance cassette module, and terminator module) to assemble the final destination vector. *Agrobacterium* containing the

helper plasmid pSOUP were transformed with the plasmid and used to transform *Arabidopsis* plants. The transformed seedlings were selected using BASTA^{TM} (BAYER) as described in Section 2.7.3.

2.14.2 Expression Analysis Lines

2.14.2.1 pGREEN-IIs pAG::GUS

The complete second intron of *AtAG* was amplified from *Arabidopsis* genomic DNA (accession: L-*er*)using Phusion DNA polymerase (ThermoFisher) using primers oCF463/451 (see Table 7.1 in the Appendix). This amplicon was ligated into the GreenGate promoter module entry vector which was then used alongside other GreenGate plasmids (N-tag dummy module, GUS CDS, C-tag dummy module, BASTA resistance cassette module, and terminator module) to assemble the final destination vector. *Agrobacterium* containing the helper plasmid pSOUP were transformed with the plasmid and used to transform *Arabidopsis* plants. The transformed seedlings were selected using BASTA^{TM} (BAYER) as described in Section 2.7.3.

2.14.2.2 pGREEN-IIs pCPC::GUS

A Greengate promoter entry vector containing 1.5kb upstream of the start codon of *AtCPC* was donated by Andrea Finocchio (Wellmer Laboratory). This plasmid was used alongside other GreenGate plasmids in a destination reaction (N-tag dummy module, GUS CDS, C-tag dummy module, BASTA resistance cassette module, and terminator module) to assemble the final destination vector. *Agrobacterium* containing the helper plasmid pSOUP were transformed with the plasmid and used to transform *Arabidopsis* plants. The transformed seedlings were selected using BASTA^{TM} (BAYER) as described in Section 2.7.3.

2.14.2.3 pGREEN-IIs pTRY::GUS

A Greengate promoter entry vector containing 1.5kb upstream of the start codon of *At-TRY* was donated by Andrea Finocchio (Wellmer Laboratory). This plasmid was used alongside other GreenGate plasmids in a destination reaction (N-tag dummy module, GUS CDS, C-tag dummy module, BASTA resistance cassette module, and terminator module) to assemble the final destination vector. *Agrobacterium* containing the helper plasmid pSOUP were transformed with the plasmid and used to transform *Arabidopsis* plants. The transformed seedlings were selected using BASTA^{TM} (BAYER) as described in Section 2.7.3.

2.14.3 Promoter Swap Experiments

2.14.3.1 pAG::*AtAG*

The GreenGate promoter entry vector containing the *AG* second intron was used in a GreenGate destination reaction with the GreenGate CDS entry vector containing the *AtAG* CDS and other GreenGate plasmids (N-tag dummy module, C-tag dummy module, BASTA resistance cassette module, and terminator module) to assemble the destination vector. *Agrobacterium* containing the helper plasmid pSOUP were transformed with the plasmid and used to transform *Arabidopsis* plants. The transformed seedlings were selected using BASTA^{TM} (BAYER) as described in Section 2.7.3.

2.14.3.2 pAG::*BrAG1*

The GreenGate promoter entry vector containing the *AtAG* second intron was used in a GreenGate destination reaction with the GreenGate CDS entry vector containing the *BrAG1* CDS and other GreenGate plasmids (N-tag dummy module, C-tag dummy module, BASTA resistance cassette module, and terminator module) to assemble the destination vector. *Agrobacterium* containing the helper plasmid pSOUP were transformed with the plasmid and used to transform *Arabidopsis* plants. The transformed seedlings were selected using BASTA^{TM} (BAYER) as described in Section 2.7.3.

2.14.3.3 pAG::*BrAG2*

The GreenGate promoter entry vector containing the *AG* second intron was used in a GreenGate destination reaction with the GreenGate CDS entry vector containing the *BrAG2* CDS and other GreenGate plasmids (N-tag dummy module, C-tag dummy module, BASTA resistance cassette module, and terminator module) to assemble the destination vector. *Agrobacterium* containing the helper plasmid pSOUP were transformed with the plasmid and used to transform *Arabidopsis* plants. The transformed seedlings were selected using BASTA^{TM} (BAYER) as described in Section 2.7.3.

2.14.3.4 pCPC::*AtCPC*

The GreenGate promoter entry vector containing the 1.5kb upstream sequence of *AtCPC* and the GreenGate CDS entry vector containing the sequence of *AtCPC* was used in a GreenGate destination reaction other GreenGate plasmids (N-tag dummy module, C-tag dummy module, BASTA resistance cassette module, and terminator module) to assemble the destination vector. *Agrobacterium* containing the helper plasmid pSOUP were transformed with the plasmid and used to transform *Arabidopsis* plants. The transformed seedlings were selected using BASTA^{TM} (BAYER) as described in Section 2.7.3.

2.14.3.5 pCPC::*BrCPC1*

The GreenGate promoter entry vector containing the 1.5kb upstream sequence of *AtCPC* and the GreenGate CDS entry vector containing the sequence of *BrCPC1* was used in a GreenGate destination reaction other GreenGate plasmids (N-tag dummy module, C-tag dummy module, BASTA resistance cassette module, and terminator module) to assemble the destination vector. *Agrobacterium* containing the helper plasmid pSOUP were transformed with the plasmid and used to transform *Arabidopsis* plants. The transformed seedlings were selected using BASTA^{TM} (BAYER) as described in Section 2.7.3.

2.14.3.6 pCPC::*BrCPC2*

The GreenGate promoter entry vector containing the 1.5kb upstream sequence of *AtCPC* and the GreenGate CDS entry vector containing the sequence of *BrCPC2* was used in a GreenGate destination reaction other GreenGate plasmids (N-tag dummy module, C-tag dummy module, BASTA resistance cassette module, and terminator module) to assemble the destination vector. *Agrobacterium* containing the helper plasmid pSOUP were transformed with the plasmid and used to transform *Arabidopsis* plants. The transformed seedlings were selected using BASTA^{TM} (BAYER) as described in Section 2.7.3.

2.14.3.7 pTRY::*AtTRY*

The GreenGate promoter entry vector containing the 1.5kb upstream sequence of *AtTRY* and the GreenGate CDS entry vector containing the sequence of *AtTRY* was used in a GreenGate destination reaction other GreenGate plasmids (N-tag dummy module, C-tag dummy module, BASTA resistance cassette module, and terminator module) to assemble the destination vector. *Agrobacterium* containing the helper plasmid pSOUP were transformed with the plasmid and used to transform *Arabidopsis* plants. The transformed seedlings were selected using BASTA^{TM} (BAYER) as described in Section 2.7.3.

2.14.3.8 pTRY::*BrTRY1*

The GreenGate promoter entry vector containing the 1.5kb upstream sequence of AtTRY and the GreenGate CDS entry vector containing the sequence of *BrTRY1* was used in a GreenGate destination reaction other GreenGate plasmids (N-tag dummy module, C-tag dummy module, BASTA resistance cassette module, and terminator module) to assemble the destination vector. *Agrobacterium* containing the helper plasmid pSOUP were transformed with the plasmid and used to transform *Arabidopsis* plants. The transformed seedlings were selected using BASTA^{TM} (BAYER) as described in Section 2.7.3.

2.14.3.9 pTRY::*BrTRY2*

The GreenGate promoter entry vector containing the 1.5kb upstream sequence of *AtTRY* and the GreenGate CDS entry vector containing the sequence of *BrTRY2* was used in a GreenGate destination reaction other GreenGate plasmids (N-tag dummy module, C-tag dummy module, BASTA resistance cassette module, and terminator module) to assemble the destination vector. *Agrobacterium* containing the helper plasmid pSOUP were transformed with the plasmid and used to transform *Arabidopsis* plants. The transformed seedlings were selected using BASTA^{TM} (BAYER) as described in Section 2.7.3.

2.14.3.10 pTRY::*BrTRY3*

The GreenGate promoter entry vector containing the 1.5kb upstream sequence of AtTRY and the GreenGate CDS entry vector containing the sequence of *BrTRY3* was used in a GreenGate destination reaction other GreenGate plasmids (N-tag dummy module, C-tag dummy module, BASTA resistance cassette module, and terminator module) to assemble the destination vector. *Agrobacterium* containing the helper plasmid pSOUP were transformed with the plasmid and used to transform *Arabidopsis* plants. The transformed seedlings were selected using BASTA^{TM} (BAYER) as described in Section 2.7.3.

Chapter 3 - Identification and Characterisation of Homologs of *AGAMOUS* in *Brassica rapa*

3.1 Introduction

As discussed in Section 1.1.4, floral organs are modified leaves and in the case of the floral reproductive organs, *AGAMOUS* (AG) is central in repressing trichome formation on the reproductive organs. *AG* overlaps with the gene network controlling trichome formation in several pathways. It prevents the expression of *GL1* which is essential for trichome formation, upregulates the expression of R3-MYB trichome repressors and represses the plant hormone cytokinin which is known to be a positive regulator of trichome development both directly through type-A ARRs and indirectly via regulation of cytokinin response genes *GLABROUS INFLORESCENCE STEMS 2 GIS2* and *ZINC FINGER PROTEIN 8 ZFP8* (Ó'Maoiléidigh et al., 2018). It could be said that *AG* is a gatekeeper, holding back the leaf-like features such as trichomes and once it is knocked down these qualities are free to emerge.

As discussed in Sections 1.1.1 and 1.1.2, the process of flower development is driven by key floral organ identity genes which encode MADS-domain transcription factors, acting to activate downstream targets in order to drive the development of floral organs. The function of *AG* in flower development can be split broadly into two roles - ensuring the floral meristem ceases proliferation at an appropriate time to allow for the correct development of the reproductive organs of the flower and floral determinacy in addition

to its function in suppressing the formation of trichomes on gynoecia.

3.1.1 The Role of *AG* in Floral Meristem Determinacy

The role of *AG* in the process of floral meristem determinacy was first hinted at when *ag-1* mutant plants showed an indeterminate floral phenotype with an additional flower produced inside the inner whorl of the first flower. These flowers also lacked any reproductive organs (Yanofsky et al., 1990). The stem cells in the shoot apical meristem are located in the central zone and are marked by the expression of *CLAVATA3* (*CLV3*) (Fletcher et al., 1999). The stem cell niche is maintained by the expression of *WUSCHEL* (*WUS*) in the organizing centre located below the stem cell niche (Mayer et al., 1998). *WUS* expression is activated by the binding of SPLAYED (SYD), a chromatin remodelling factor which increases the level of chromatin accessibility (Kwon et al., 2005). WUS travels in a non-cell autonomous manner to the central zone where it promotes stem cell fate by repressing differentiation (Yadav et al., 2011; Daum et al., 2014). Plants containing *wus* mutant alleles produce a limited number of organs with the meristem termination occurring too early in the development processes although they do eventually produce shoots and a small number of abnormal flowers (Laux et al., 1996; Kwon et al., 2005).

In the central zone, WUS binds to the *CLV3* locus and activates its expression (Schoof et al., 2000). CLV3 is a small signalling peptide that through interactions with its receptors such as CLV1, CLV2 and CORYNE (CRN) act to repress *WUS* ensuring that its expression domain is contained to the cells of the organizing centre underneath the stem cell domain (Somssich et al., 2016). Plants lacking CLV3 activity develop enlarged meristems due to lack of repression of *WUS* (Fletcher et al., 1999). This marks the establishment of a feedback loop, with WUS promoting stem cell fates in the central zone by activating CLV signalling and CLV acting in turn to restrict the expression of *WUS*. This feedback loop ensures the maintenance of the stem cell pool from which the

whole structure of the flower will form (Schoof et al., 2000). Stem cell maintenance ceases at stage 6 of floral development (when sepals enclose the floral bud) and this avoids the over-proliferation of tissue and the correct development of carpels (Sun et al., 2009).

AG is induced at the centre of emerging floral buds at stage 3 - the stage in flower development as described in Smyth et al. 1990 when the sepal primordia begin to arise. This occurs by the binding of WUS and LFY to regulatory elements in the second intron of *AG* (Hong et al., 2003). In order to mediate floral meristem determinacy, *AG* represses the expression of *WUS* through direct and indirect means. This is achieved *via* the recruitment of proteins associated with Polycomb Group proteins to the promoter of *WUS* which is silenced due to the deposition of repressive histone markers (Liu et al., 2011). Another pathway by which *AG* represses *WUS* is *via* activation of *KNUCKLES* (*KNU*) at stage 6 (when sepals enclose the developing flower) in the flower development process (Sun et al., 2009). KNU is a C2H2-type zinc finger protein and *knu* mutants display abnormal silique morphology as well as reduced fertility (Payne et al., 2004). These mutants develop short siliques that are often kinked and all contain a singular bulge or 'knuckle' of tissue. This bulge of tissue is a mass of tissue consisting of ectopic stamen and carpel tissue. It has been shown that the ectopic tissue formation is a result of extended WUS expression resulting in the failure of the floral meristem to terminate at the correct time in flower development (Sun et al., 2019).

KNU is thought to bind directly to the *WUS* promoter, causing the eviction of SPLAYED (SYD) which results in reduced accessibility at the *WUS* locus and decreases the levels of active histone markers on *WUS* chromatin (Sun et al., 2019). AG plays a critical role in ensuring that the meristem terminates in a timely manner to ensure correct flower development.

3.1.2 The Role of *AG* in Stamen and Carpel Development

In addition to its role in floral meristem determinacy, *AG* controls the development of the reproductive organs of the flower - stamens in the third whorl and carpels in the fourth whorl. As previously mentioned, *AG* mutants such as *ag-1* develop flowers that completely lack floral organs with additional sepals and petals as well as a secondary flower taking their place (Yanofsky et al., 1990). This strong phenotype paired with data obtained from microarray experiments comparing *ag* and wild-type flowers, as well as from genome-wide localization studies, showed that hundreds of genes are regulated by AG and that in its absence, reproductive floral organs cannot form (Wellmer et al., 2004; Ó'Maoiléidigh et al., 2013).

The expression pattern across a range of tissue in *Arabidopsis* is shown in Figure 3.1, with expression detected in the floral meristem, young flowers, and at its highest level in stamens and gynoecium (Winter et al., 2007; Waese et al., 2017).

As detailed in Section 1.1.3, stamen identity is specified by the combined action of B-, C- and E-class genes - *AP3*, *PI*, *AG* and *SEP*. In situ-hybridisation experiments have revealed that during flower development, AG mRNA is detected from stage 5 (when petal and stamen primordia emerge) in stamen primordia until stage 14 (when the long anthers extend above the stigma) in the endothecium and nectaries, but the expression is not detected in the microspore cell lineage (Smyth et al., 1990; Drews et al., 1991). The expression domain of *AG* begins more broadly in early stamen development with the domain restricting as stamens mature (Drews et al., 1991). These expression patterns indicate the expression of *AG* is necessary throughout the process of stamen development and maturation and is likely targeting different genes during this process in order to drive development.

To date, the downstream regulatory network of stamen development is not well de-

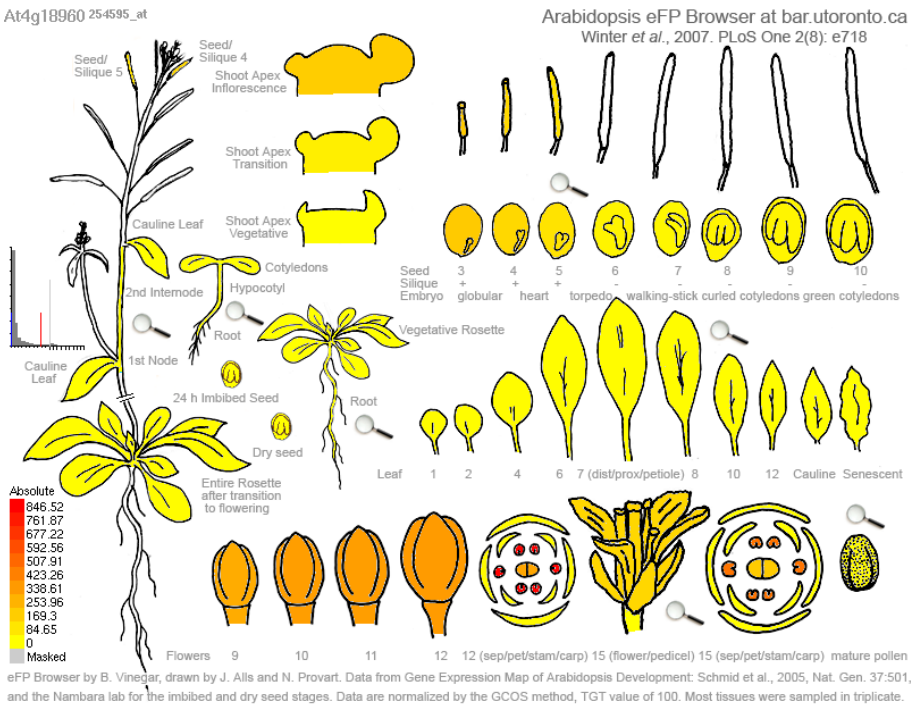

Figure 3.1: Expression profile of *AG* in *Arabidopsis*. Expression domains of *AG* in *Arabidopsis* obtained from the Arabidopsis eFP Browser (Winter et al., 2007; Waese et al., 2017). Yellow indicates no or low expression whilst red indicates higher levels of expression. In the case of *AG*, expression is highest in younger developing flowers and in the reproductive organs of fully developed flowers.

---

fined and how exactly *AG* is driving the entire process is unclear. Several genes and how they function in stamen development have been identified such as *NOZZLE/SPOROCYTELESS* (*NZZ/SPL*) and *DEFECTIVE IN ANTHER DEHISCENCE 1 (DAD1)*. It appears that AG does not directly drive the development of microsporogenesis but induces the formation of pollen by directly activating *NZZ/SPL* a known regulator of sporogenesis (Ito et al., 2004). It was shown that a CArG box present in the 3' promoter region of *NZZ/SPL* had AG binding activity and when it was mutated expression activity was disrupted. Using an inducible expression system it was shown that *AG* was sufficient to induce microsporogenesis. AG has also been shown to regulate the expression of *DAD1*. DAD1 is required for the production of jasmonate, a phytohormone which is required for filament formation and proper elongation as well as anther morphogenesis and dehiscence (Ito et al., 2007).

Carpel identity is specified by C- and E-class genes alone. *AG* is detected in carpel primordia as the ovules develop (stage 8) up until stage 14 when fertilization takes place (Smyth et al., 1990; Drews et al., 1991). It is clear that AG is essential for proper carpel development but to date, the complex downstream network of how exactly AG is driving this process has not been elucidated. LEUNIG (LUG) acting in conjunction with SEUSS (SEU) and LEUNIG-HOMOLOG (LUH) has been shown to have a role in restricting the expression domain of *AG* to the third and fourth whorl (Conner and Liu, 2000). Gynoecia of strong *lug* mutants show abnormal development with only the basal part fused correctly and the apex is completely unfused (Sitaraman et al., 2008). Another gene acting downstream of *AG* is the YABBY family gene *CRABS CLAW* (*CRC*) which is necessary for the development of the carpels (Bowman and Smyth, 1999). AG works in conjunction with SEP to activate the expression of *CRC*. When the *CRC* promoter fused to a GUS reporter module showed expression in the expected domain but this was eliminated upon mutation of CArG-boxes providing evidence that AG and SEP binding is essential in order to ensure correct expression of *CRC* (Lee et al., 2005).

3.1.3 Aims of the Chapter

As described in Section 1.3, it has been shown that in *Arabidopsis* a combination of weak *ag* mutant alleles with strong mutant alleles of trichome repressor genes leads to the formation of trichomes on seed pods without major effects on fertility. Based on these results, one of the main aims of the SFI/EPA-funded project that supported my PhD work was to engineer *Brassica* crops with floral trichomes as a possible new avenue for crop protection. The assumption here was that the results obtained in *Arabidopsis* could be directly translated to plants of the *Brassica* genus. To test whether this is indeed possible, we first needed to identify and characterize the *Brassica* orthologs of the relevant *Arabidopsis* genes. Secondly, we needed to isolate mutant alleles that could be used to test our main working hypothesis, namely that the mechanism for the suppression of trichome formation in flowers is conserved between *Arabidopsis* and

Brassica. In this chapter, I focused on the orthologs of *Arabidopsis AG* in *B. rapa*. The specific aims of this chapter were to identify and functionally characterize these genes and to assess their level of redundancy. To this end, a combination of mutant analyses and transgenic approaches were used. The isolation of mutant alleles for these genes with different phenotypic strengths would further allow us to fine-tune C function and ultimately, to test whether the trichome phenotypes observed in *Arabidopsis* can be re-created in *Brassica* plants.

3.2 Results

3.2.1 Identification of *AG* Homologs in *B. rapa*

In order to identify homologs of *AG* from *Arabidopsis* (AT4G18960; hereafter referred to as *AtAG*) in *B. rapa*, its genomic and coding DNA sequence (CDS) were used to perform BLAST searches on the *B. rapa* genome sequence that was available at the time (from accession Chiifu-401-42 (Wang et al., 2011; Cai et al., 2017)). Two highly significant hits were obtained – gene *Bra012564* located on chromosome 1 (E-value of 3.20E-82) and gene *Bra013364* on chromosome 3 (E-value of 1.20E-75). The full results of the BLAST search are shown in Table 7.2 in the Appendix.

The genomic sequences of *Bra012564* and of *Bra013364* are approximately 3.9 kb and 4 kb, respectively, in length, and are thus shorter than the genomic sequence for *AtAG* (5.3 kb). The reduced length of the *Brassica* genes is in agreement with results from a previous report showing that genes in *Brassica* species are on average 27% shorter than their homologs in *Arabidopsis* (Mun et al., 2009). Hence, the sizes of *Bra012564* and of *Bra013364* are approximately what would be expected of *AG* homologs in *B. rapa*. A schematic of the structures of the genes is shown in Figure 3.2.

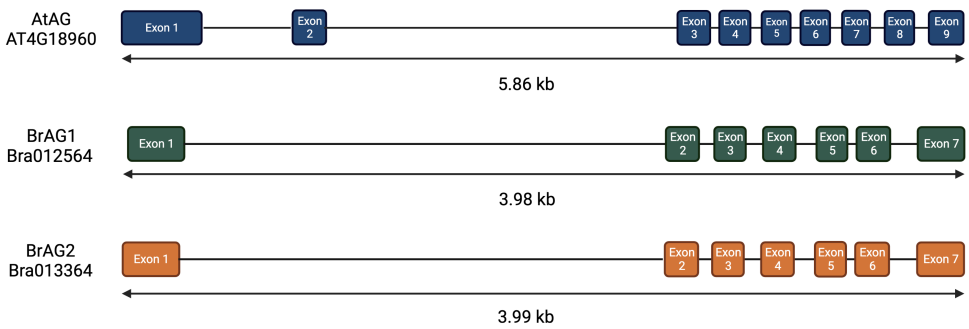

Figure 3.2: Gene structure of *AtAG* and of two homologs from *B. rapa*. Diagram showing the structure of *AtAG* (dark blue), *Bra012564* (green) and *Bra013364* (orange). Exons of each gene are labelled and introns are represented by black lines. The length of each gene is indicated by arrows. Figure generated with BioRender.com.

AtAG, depending on the splice variant, contains either 9 or 10 exons and 8 or 9 introns. *Bra012564* and *Bra013364* both contain 7 exons and 6 introns. *AtAG* has a large second intron that is 2.8 kb in length and is known to play an essential role in the regulation of its expression (Busch et al., 1999; Hong et al., 2003; Wu et al., 2018). For both *Bra012564* and *Bra013364*, the first intron is the largest spanning 2.7 kb and 2.8 kb in length, respectively. Thus, although the gene structures of the *Arabidopsis* and *Brassica* genes are not identical, they all have an unusually long intron in common (the mean length of an intron in *Arabidopsis* is approximately 170bp (The Arabidopsis Genome Initiative, 2000)). This lends support to the idea that *Bra012564* and *Bra013364* are the homologs of *AtAG* in *B. rapa*. When examining the level of conservation of the genomic sequence of the three genes, a high level of similarity - greater than 80% - is seen (see Figure 7.1 in the Appendix for the alignment of the genomic sequence of the three genes).

The genomic sequence, the coding DNA sequence (CDS) and the amino acid sequence of *Bra012564* and *Bra013364* were aligned to the corresponding sequences of *AtAG* using the *EMBL-EBI Clustal Omega* tool (Madeira et al., 2022). A summary of the percentage identity between these sequences is shown in Table 3.1.

Table 3.1: Percentage identity matrix for *AtAG* and potential *B. rapa* homologs. Table comparing the percentage identity of the CDS and protein sequences of *AtAG*, *Bra012564*, and *Bra013364*.

	Genomic			**CDS**		
	1	**2**	**3**	**1**	**2**	**3**
1. AtAG	100	82.58	83.11	100	91.90	93.02
2. Bra012564	82.58	100	83.37	91.90	100	94.02
3. Bra013364	83.11	83.37	100	93.02	94.02	100

	Protein		
	1	**2**	**3**
1. AtAG	100	93.17	94.44
2. BrAG1	93.17	100	94.38
3. BrAG2	94.44	94.38	100

The average level of sequence conservation between the genes and proteins of *Arabidopsis* and *B. rapa* is in the region of 70% (Quiros et al., 2001). The results of the genomic alignment show that there is a higher-than-average level of sequence conservation between the three genes. Both *Bra012564* and *Bra013364* have $\approx$ 83% identity with the genomic sequence of *AtAG*. There is an even higher level of identity between the CDS of the three genes, with *Bra013364* having the higher score at $\approx$ 94% and *Bra012564* at $\approx$ 92%. A similar level of identity is observed when the protein sequences are aligned.

Figure 3.3 shows the alignment between the CDS of *AtAG*, *Bra012564* and *Bra013364*. Both *Bra012564* and *Bra013364* possess the usual start codon of "AUG" unlike the *AtAG* which possesses the unusual start codon of "ACG". This is in line with previous findings that *AG* in *Arabidopsis* is unique in this non-AUG initiation translation. In other plant species such as *B. napus*, *Solanum lycopersicum* (tomato), *Zea mays* (corn) and *(Nicotiana tabacum* (tobacco), the CDS of *AG* begins with "AUG (Riechmann et al., 1999).

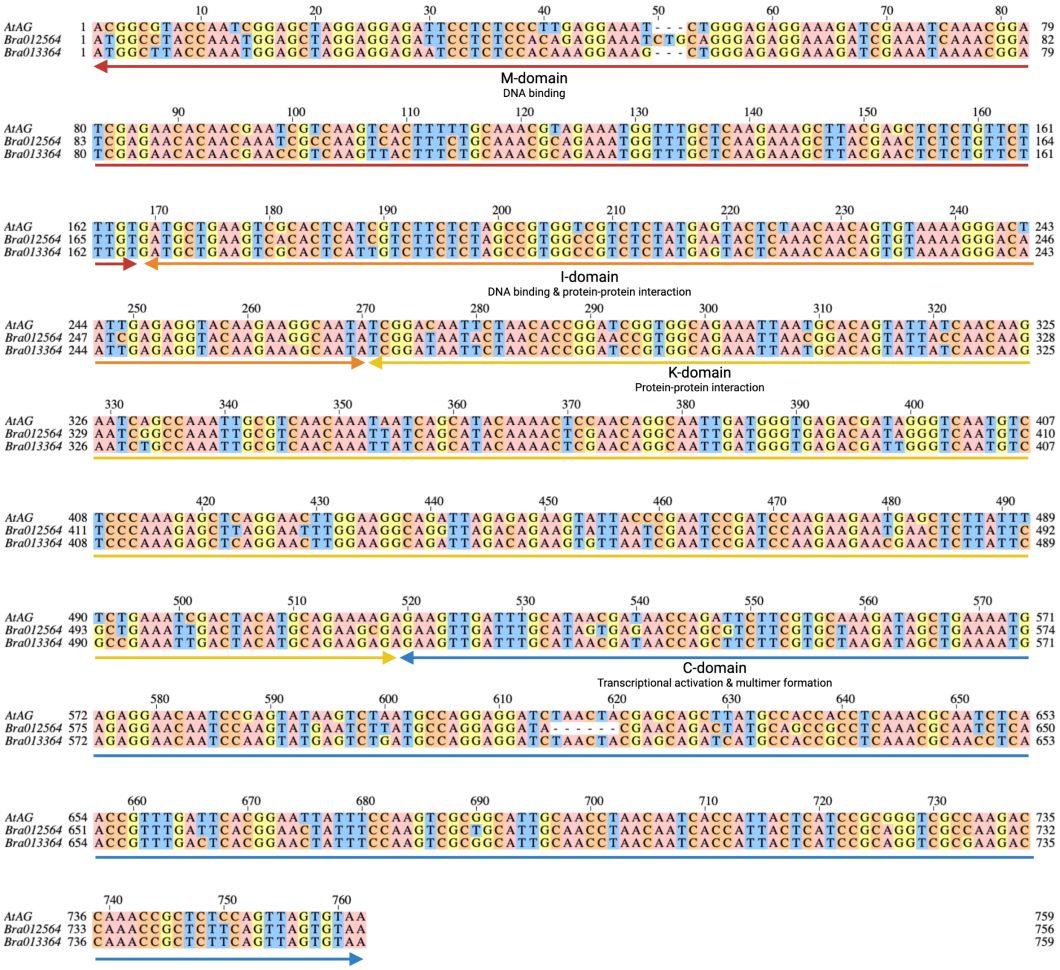

Figure 3.3: Alignment of coding sequences of *AG* and the putative *B. rapa* homologs. The CDS of *AtAG* and the two putative homologs of *AG* in *B. rapa* were aligned using the *Clustal Omega Multiple Sequence Alignment* tool from EMBL-EBI (Madeira et al., 2022). The results were visualised using *Jalview* software. The domains of MADS domain transcription factors and their functions are annotated. Figure generated with BioRender.com

---

Figure 3.4 shows the alignment between the protein sequences of *AtAG*, *Bra012564* and *Bra013364*. The MADS-domain of *AtAG* consists of amino acid residues 19 to 73 (as discussed in Section 1.1.2) which are coded by bases 57 to 219 of the CDS. As can be seen in both Figures 3.3 and 3.4, there is a high level of sequence identity at this region in *Bra012564* and *Bra013364*. The only difference in the MADS-domain occurs at position 61 of *Bra012564* where a threonine is present where an alanine occurs in both *AtAG* and *Bra013364*. Although alanine is classed as a hydrophobic amino acid and threonine is classed as a polar amino acid, they are structurally similar, suggesting

that this amino acid substitution does not affect the function of the proteins.

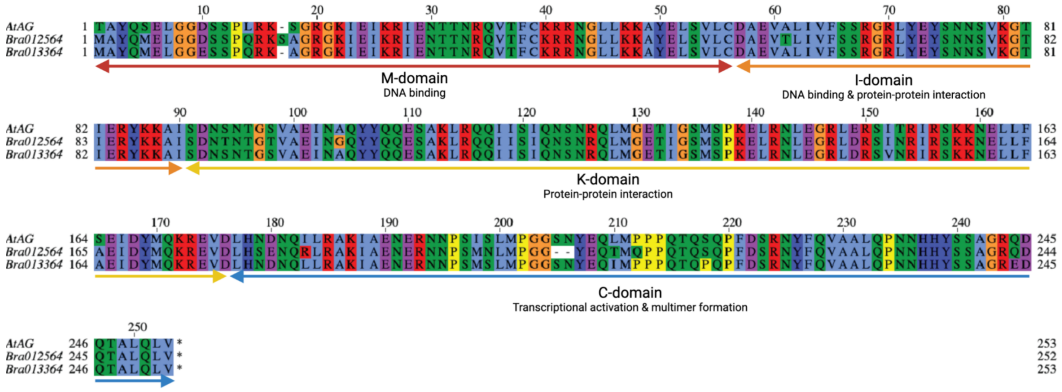

Figure 3.4: Alignment of the protein sequence of *AtAG* and putative *B. rapa* homologs. The protein sequence of *AtAG* aligned to the protein sequence of Bra012564 and Bra013364 using the *Clustal Omega Multiple Sequence Alignment* tool from EMBL-EBI (Madeira et al., 2022). Results were visualised using *Jalview* software. The following colour scheme was used to denote amino acids that belong to the same residue groups: blue: hydrophobic; red: positively charged; magenta: negatively charged; green: polar; pink: cysteines; orange: glycines; yellow: prolines; cyan: aromatic. The domains of MADS domain transcription factors and their functions are annotated. Figure generated with BioRender.com

---

To further test whether *Bra012564* and *Bra013364* are indeed *AtAG* homologs, I compared the genes flanking the different loci. As shown in Table 3.2, *Bra012564* and *Bra013364* share many neighbouring genes with *AtAG*. Of the 20 surrounding genes examined for *Bra012564*, 12 are located on the same chromosome as *AtAG* and 8 of the 20 genes are also direct neighbours of *AtAG*. In the case of *Bra013364*, the level of synteny is higher with 18 of 20 genes examined being located on chromosome 4 and 11 of the 20 genes being direct neighbours of *AtAG*. Thus, the two *B. rapa* genes have a high level of synteny with *AtAG*, suggesting that they are indeed homologous.

On the basis of the similarity of the gene structures, the high level of sequence identity and the level of synteny conservation between the three genes, it was concluded that *Bra012564* and *Bra013364* are homologs of *AG* in *B. rapa*. As such, they were given shortened gene names - *BrAG1* for *Bra012564* and *BrAG2* for *Bra013364*.

Table 3.2: Synteny of *AtAG* and putative *B. rapa* homologs. The table lists the ten neighbouring genes upstream and downstream of *AtAG* (AT4G18960), *Bra012564* and *Bra013364*. The neighbouring genes were identified using the EnsemblPlants 'Region in Detail' view in order to identify the surrounding genes. In the case of the *B. rapa* genes, the BRAD gene IDs were "BLASTed" and the corresponding TAIR IDs were listed for easy comparison. Gene IDs that appear in the *AtAG* column and one of the homolog columns are highlighted in the same colour for ease of comparison. N/A indicates there was no corresponding TAIR ID found for the BRAD gene ID.

Position	AT4G18960	Bra012564	Bra013364
-10	AT4G18870	AT5G50960	AT4G18830
-9	AT4G18880	AT3G62850	AT4G18840
-8	AT4G18890	AT3G62850	N/A
-7	AT4G18900	AT4G18880	AT4G36150
-6	AT4G18905	AT4G18890	AT4G18880
-5	AT4G18910	AT2G19030	AT4G18890
-4	AT4G18920	AT4G18905	AT4G18905
-3	AT4G18930	AT4G18910	AT4G18910
-2	AT4G18940	AT4G18920	AT4G18920
-1	AT4G18950	AT2G07760	AT4G18950
+1	AT4G18970	AT4G18970	AT4G18975
+2	AT4G18975	N/A	AT5G45630
+3	AT4G18980	AT1G14800	AT4G18990
+4	AT4G18990	AT4G18980	AT4G19003
+5	AT4G19000	AT4G18990	AT4G19010
+6	AT4G19003	AT4G19110	AT4G19020
+7	AT4G19010	AT4G19120	AT4G19040
+8	AT4G19020	N/A	AT4G19050
+9	AT4G19030	AT4G19140	AT4G19060
+10	AT4G19035	AT4G19180	AT4G19120

As shown in Figure 3.5, *BrAG1* and *BrAG2* genes are located on different chromosomes and are thus unlinked. This strongly suggested that mutant alleles for both genes, once isolated and characterized, could be crossed to obtain the double-mutants needed to assess the level of redundancy between the genes and to assess them for fine-tuning C function (see Section 3.1 above).

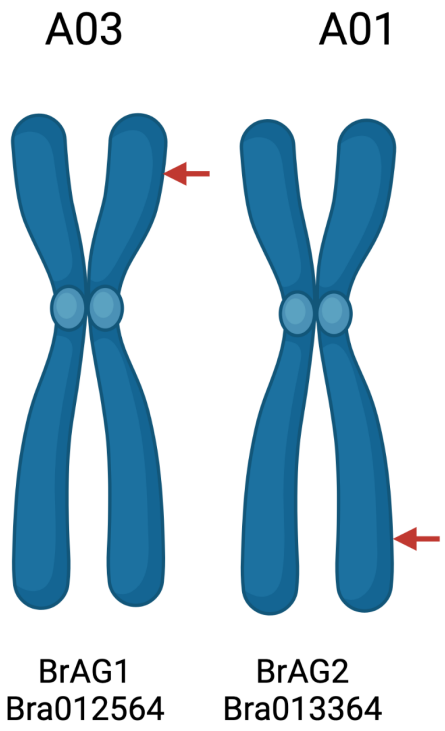

Figure 3.5: Chromosomal location of *BrAG* genes. Arrows indicate the approximate positions of BrAG1 and BrAG2 on chromosomes A03 and A01, respectively. Figure generated with Biorender.com.

3.2.2 Characterisation of *BrAG1* and *BrAG2* Expression

As previously detailed in Section 1.5, genes that undergo duplication can have one of several possible outcomes - loss or pseudogenization of one copy, retention of both copies due to gene dosage effects, subfunctionalization where functional roles of the original gene are split between the two copies or neofunctionalization where one copy retains the original function and the other copy is free of evolutionary constraint to evolve a new function. Sub- and neofunctionalization are thought to be often a consequence of altered expression patterns (Birchler and Yang, 2022). As there appear to be two homologs of *AtAG* identified in *B. rapa*, I examined whether both are expressed and whether their expression patterns would show any evidence for possible subfunctionalization. To this end, reverse transcription-quantitative PCR (RT-qPCR) was used to characterize the level of expression of the two homologs across a wide range of *B. rapa* plant tissue.

The tissue samples used for the expression analysis were as follows: whole 10-day-old seedlings; whole leaf tissue; floral buds of different stages that were sized in 1 mm increments from $<$ 1 mm to 10 mm; and individual types of floral organs from mature flowers (equivalent to stage 13 flowers in *Arabidopsis*; stages according to Smyth et al., 1990). All tissue samples were collected from the R-o-18 accession of *B. rapa*. Three biologically independent sets of samples were collected for analysis.

The results of the RT-qPCR experiments are shown in Figure 3.6. The expression results obtained for *BrAG1* and *BrAG2* broadly match the expression profile that has been described for *AtAG* (Figure 3.1). No expression was detected in 10-day-old seedlings and leaf tissue. In contrast, high levels of expression of both homologs were detected in floral meristems and young floral buds. The level of expression detected decreases as the flower increases in size and matures. In the floral meristem up to flowers 4 mm in size, the expression of both homologs was found to be about equal with no clear difference between the two. In flowers of 5 to 10 mm, *BrAG1* expression was detected at a higher level than *BrAG2*.

When examining the expression data for floral organ types, little to no expression was detected in sepals and petals. As expected for *AtAG* homologs, the highest level of expression was detected in the reproductive floral organs. The expression levels of *BrAG1* were approximately the same in both stamens and the gynoecium whilst *BrAG2* expression appeared higher in stamens than in the gynoecium. In general, the expression of *BrAG1* appeared to be higher than that of *BrAG2* in the reproductive floral organs.

From the results of these experiments, it was concluded that *BrAG1* and *BrAG2* are both expressed in flowers and have similar expression patterns. Thus, no evidence for pseudogenization or sub-/neofunctionalization of *BrAG1* or *BrAG2* could be detected given that the expression patterns of both broadly matched that expected of *AtAG*.

With regards to the aim of knocking down *AG* function in *B. rapa* in order to generate gynoecium with trichomes (as detailed in Section 1.3), this data reveals that *BrAG1* is the most important homolog of the two to target, given that it is more highly expressed. However as *BrAG2* expression is detected in gynoecium tissue, albeit at a much lower level, and redundancy between the two could not be ruled out, mutant alleles for both *BrAG1* and *BrAG2* were obtained and analysed (see Sections 3.2.8 and 3.2.9).

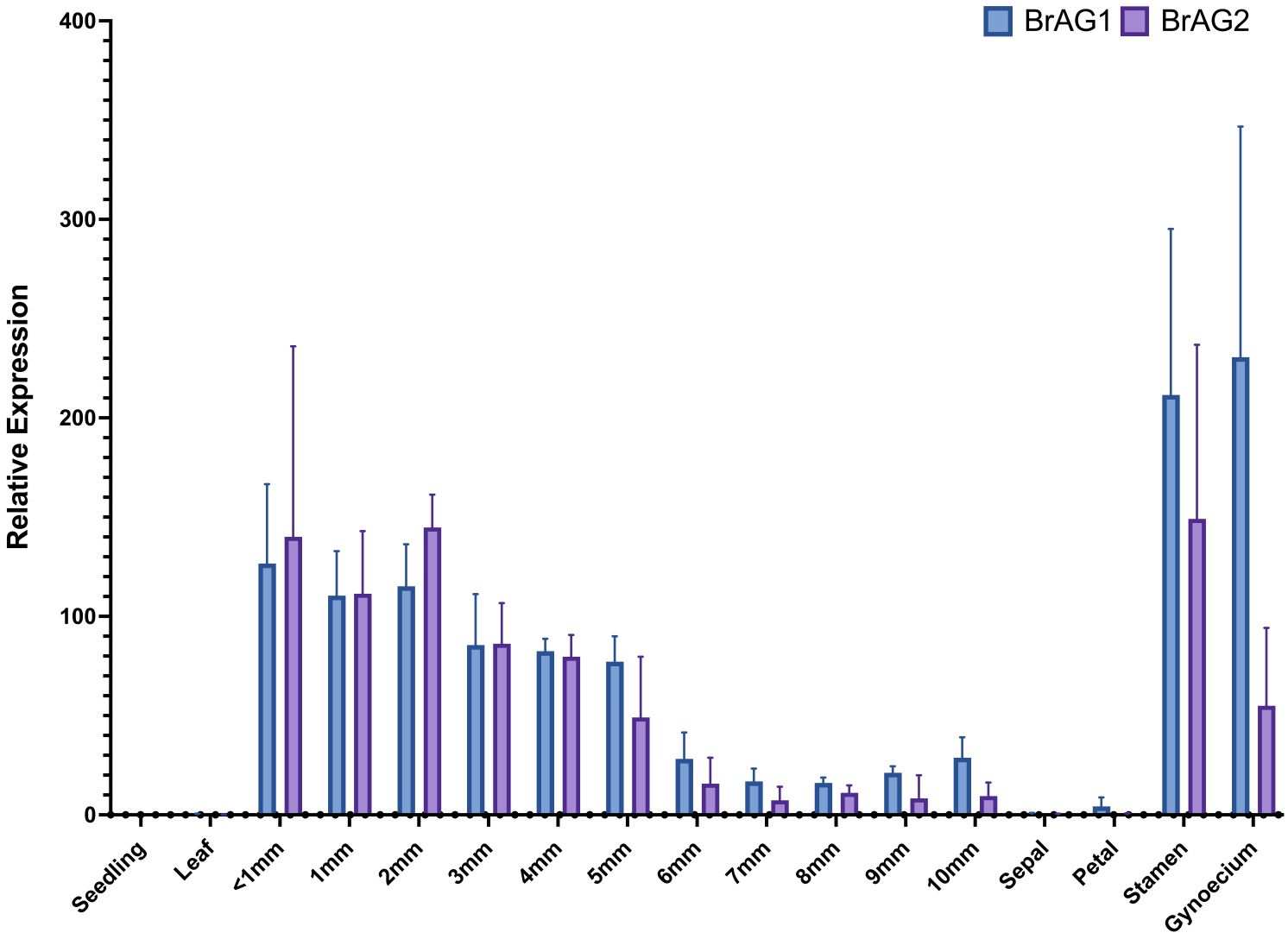

Figure 3.6: Expression of *BrAG1* and *BrAG2* as determined by RT-qPCR. Results for *BrAG1* and *BrAG2* are shown by blue and purple bars, respectively. Values are the mean of three biological replicates. Error bars represent standard deviations. The different tissue types used for the analysis are indicated. Samples from flowers of different stages are sorted by size ($<$ 1 – 10 mm). Expression values were normalized using a reference gene, *TUBULIN1 B* (*Bra015815*). No further statistical analyses were carried out on this data.

3.2.3 Characterisation of the 3'UTRs of *BrAG1* and *BrAG2*

To further characterise *BrAG1* and *BrAG2*, I determined their 3' untranslated regions (UTRs) which had not been previously described. It is known that UTRs can contain *cis*-acting elements which function in post-transcriptional control and also have an effect on the stability, transport and translation efficiency of mRNAs (Mignone et al., 2002; Srivastava et al., 2018). To identify the 3' UTRs, 3' Rapid Amplification of cDNA Ends (RACE)-PCR, as per the protocol described by Scotto-Lavino et al., 2007, was employed. The sequences obtained *via* 3'RACE-PCR were aligned with the known 3'UTR of AtAG (Figure 3.7). The 3' UTR of *AtAG* spans 267 bp, while the 3'UTRs of *BrAG1* and *BrAG2* are slightly shorter at 230 bp and 241 bp, respectively.

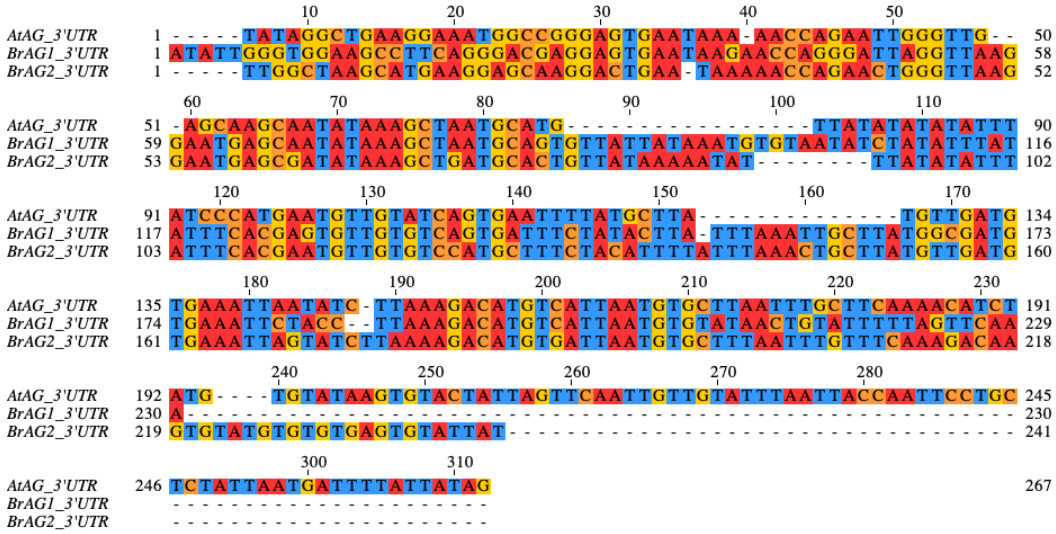

Figure 3.7: Alignment of 3'UTR sequences of *AG* homologs. Alignment of 3'UTR sequences of *AtAG*, *BrAG1* and *BrAG2*. The sequence of the 3'UTR of *AtAG* was obtained from *EnsemblPlants* while the sequence of *BrAG1* and *BrAG2* were determined *via* 3'RACE-PCR. The sequences were aligned using the *Clustal Omega Multiple Sequence Alignment* tool from EMBL-EBI (Madeira et al., 2022). The results were visualised using *Jalview* software.

The percentage identity of the different 3'UTRs sequences is high (≈70%) as summarized in Table 3.3. This result further supports the idea that *BrAG1* and *BrAG2* are *AtAG* homologs.

Table 3.3: Percentage identity matrix for 3' UTRS of *AG* homologs. Table displaying the percentage identity between the 3'UTR sequences of *AtAG*, *BrAG1* and *BrAG2*.

	1	2	3
1. AtAG	100.00	67.54	72.68
2. BrAG1	67.54	100.00	72.22
3. BrAG2	72.68	72.22	100

3'UTRs contain sequences known as AU-rich elements (AREs) that are known to have the ability to increase protein production, destabilize mRNAs and regulate mRNA localisation by the recruitment of RNA binding proteins such as exosome proteins, motor proteins or proteins involved in posttranslational modification (Mayr, 2019). As the sequences of the 3'UTRs of the *AG* homologs lack any notable difference in their sequence in terms of sequence identity or alignment, the potential AREs binding sites were examined. The 3'UTR sequences were scanned using the Meme-Suite Tool Simple Enrichment Analysis (SEA) (Section 2.8) which identifies motifs that are relatively enriched in the sequences under examination compared with shuffled sequences or control sequences. As shown in Table 3.4, a number of these sites were detected in the sequences of the *AG* homologs. While most of them appeared to be shared between all of them, several sites were present only in *AtAG* or only in one of the *B. rapa* homologs.

Table 3.4: ARE binding sites detected in 3'UTR of *AG* homologs. The table shows the number of sites for each ARE motif detected in the 3'UTR sequence using SEA analysis. W in the motif indicates that the nucleotide could be an A or U base.

Motif Sequence	AtAG	BrAG1	BrAG2
AUUUA	1	1	0
WUUUW	1	0	1
WWUUUWW	1	1	1
$(W)_3UUU(W)_3$	1	1	1
$(W)_4UUU(W)_4$	1	1	1
$(W)_5UUU(W)_5$	1	1	1
UUUGUUU	0	0	1
GUUUG	1	0	1
AWUAAA	1	1	1
Total	**8**	**6**	**7**

AtAG and *BrAG1* exclusively share one recognition site - "AUUUA", while *AtAG* and *BrAG2* exclusively share two sites - "WUUUW" and "GUUUG". *BrAG2* also possesses a recognition site that is not detected for the 3'UTR for *AtAG* and *BrAG1* - "UUUGUUU". The differences detected in the numbers and types of ARE binding sites in the 3'UTRs of the *B. rapa* homologs could explain the different expression levels in reproductive floral organs as discussed in Section 3.2.2. However, without further experimental work such as the deletion and swapping of the motifs that are unique to *BrAG1* and *BrAG2*, it is not possible to conclusively state this.

3.2.4 Comparative Analysis of Upstream Promoter Regions of *AG* Homologs

Another possible explanation for the difference in the expression strength of *BrAG1* and *BrAG2* could be differences in transcription factor binding sites in the promoters of the *B. rapa* homologs. In the context of genetic drift after a polyploidy event, deleterious mutations may have not accumulated in the coding DNA sequence but in the regulatory regions of the genes, thus leading to changes in gene expression. To test this hypothesis, I used the software *JASPAR* (Section 2.8) to analyze the upstream promoter regions of *AtAG*, *BrAG1*, and *BrAG2* for transcription factor binding sites. I initially considered only sites for MADS-domain transcription factors because they are known to play important roles in the regulation of the floral homeotic genes. In the case of *AtAG*, this includes auto-regulation, i.e. the binding of *AtAG* to its own promoter (Hong et al., 2003).

The binding site matrices used for different MADS-domain transcription factors are shown in Figure 7.4 in the Appendix. For the purpose of this analysis, the promoter regions were defined as 2kb upstream of the start codon of a gene. The percentage identity of the three upstream regions under examination is summarised in Table 3.5. *BrAG1* and *BrAG2* show a similar level of identity to *AtAG* of $\approx$ 55%, and the *BrAG1*

and *BrAG2* upstream sequences are more similar to each other (≈ 78%) than they are to *AtAG*.

Table 3.5: Percentage identity matrix for promoter sequences of *AG* homologs. The percentage of identity of the promoter regions of *AtAG* compared to *BrAG1* and *BrAG2*.

	1	2	3
1. *AtAG*	100.00	55.81	54.52
2. *BrAG1*	55.81	100	77.57
3. *BrAG2*	54.52	77.57	100.00

The MADS-domain transcription factor binding sites predicted by *JASPAR* are summarized in Table 3.6.

Table 3.6: Predicted MADS-domain transcription factor binding sites for upstream promoter regions of *AG* homologs. The names of MADS-domain transcription factors and the corresponding *JASPAR* matrix IDs are listed. The results shown are the number of sites predicted using the default relative profile score threshold of 90%.

Gene	Matrix ID	AtAG	BrAG1	BrAG2
AG	MA0005.1	4	1	2
	MA0005.2	0	0	0
AGL15	MA0548.1	0	1	0
	MA0548.2	0	0	0
AGL16	MA1199.1	0	1	0
AGL42	MA1201.1	4	2	1
AGL55	MA1202.1	0	2	2
AP1	MA0940.1	1	2	5
AP3	MA0556.1	0	1	0
FLC	MA0558.1	0	1	0
PI	MA0559.1	0	2	3
SEP3	MA0563.1	1	0	3
SEP4	MA0001.1	3	0	2
	MA0001.2	0	0	0
SOC1	MA0554.1	0	1	0
Total		**13**	**14**	**18**

The results shown are those obtained using a stringent 90% relative profile score. There are a similar number of binding sites predicted for *AtAG*, *BrAG1* and *BrAG2*. Interestingly, some of the largest differences between binding sites in between *AtAG*,

BrAG1 and *BrAG2* were observed for the floral homeotic transcription factors AP3, PI, SEP3, SEP4 and for AG itself. The most interesting results are perhaps seen in the case of predicted binding sites for AG (MA0005.1 and MA0005.2). There are no sites predicted for MA0005.2. For MA0005.1, there are four sites predicted for *AtAG*, only one predicted for *BrAG1* and two predicted for *BrAG2*. The location of the putative AG binding sites in the *AG* homologs is shown in Figure 3.8. There is no obvious conservation of an AG binding site between *AtAG*, *BrAG1* and *BrAG2*.

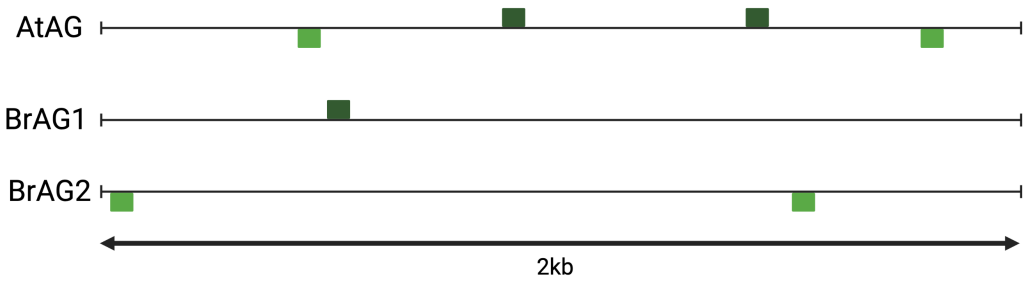

Figure 3.8: Predicted AG binding sites in the upstream regions of *AG* homologs. The location of the predicted AG binding sites in the promoter regions of *AtAG*, *BrAG1* and *BrAG2* are shown. Dark green indicates a binding site on the forward strand and light green indicates a binding site on the reverse strand of the DNA. Figure generated with BioRender.com.

As mentioned in Section 3.1, LFY acts as a transcriptional activator of *AG*. To test whether any differences in LFY binding sites could be detected in the upstream regions of the *AG* homologs, the *JASPAR* software was again used. The results of the analysis are shown in Table 3.7.

Table 3.7: Predicted LFY binding sites in upstream regions of *AG* homologs. The number of LFY binding sites detected in the region 2 kb upstream of the start codon of a gene are shown. A relative profile score threshold of 70% was used for the analysis using *JASPAR*.

Gene	Matrix ID	AtAG	BrAG1	BrAG2
LFY	MA0590.1	10	8	12

Using a relative profile score threshold of 70%, similar numbers of sites were detected for the three homologs. The location of the predicted binding sites for LFY is shown in Figure 3.9.

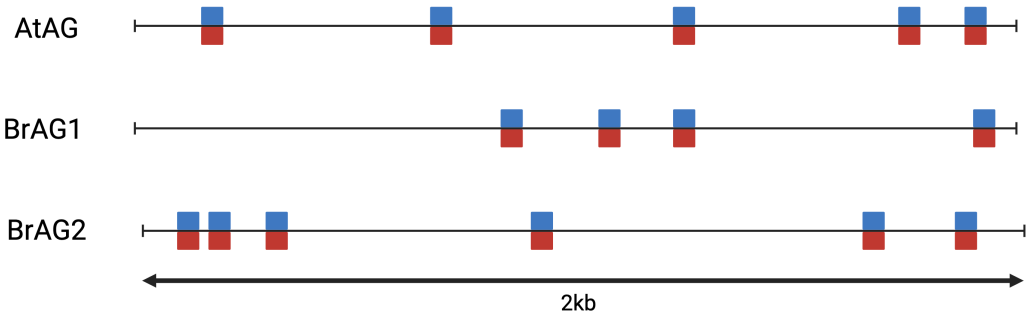

Figure 3.9: Predicted LFY binding sites in promoters of *AG* homologs. The location of the predicted LFY binding sites in the promoter regions of *AtAG*, *BrAG1*, and *BrAG2* are shown. Blue indicates a binding site on the forward strand and red indicates a binding site on the reverse strand of the DNA. Figure generated with BioRender.com.

Again, there is no clear conservation of LFY binding sites among the three homologs. While the analyses outlined above led to the identification of some differences in the number and distribution of putative binding sites for LFY and selected MADS-domain transcription factors, it is currently unclear whether any of these contribute to differences in the expression of the *AG* homologs. Further work will be needed to investigate this, for example by disrupting individual candidate binding sites through genome-editing approaches such as CRISPR-Cas9.

3.2.5 Comparative Analysis of Regulatory Introns of *AG* Homologs

As mentioned in Section 3.1, the large intron of *AtAG* contains essential regulatory elements. In *Arabidopsis*, AG has been shown to bind to this intron, suggesting that it auto-regulates its own expression (Ó'Maoiléidigh et al., 2013). To test whether the difference in expression levels between *BrAG1* and *BrAG2* could be due to differences in their intronic sequences, I conducted a comparative analysis of the large introns of the

AG homologs. An alignment of the intronic sequences is shown in Figure 3.10.

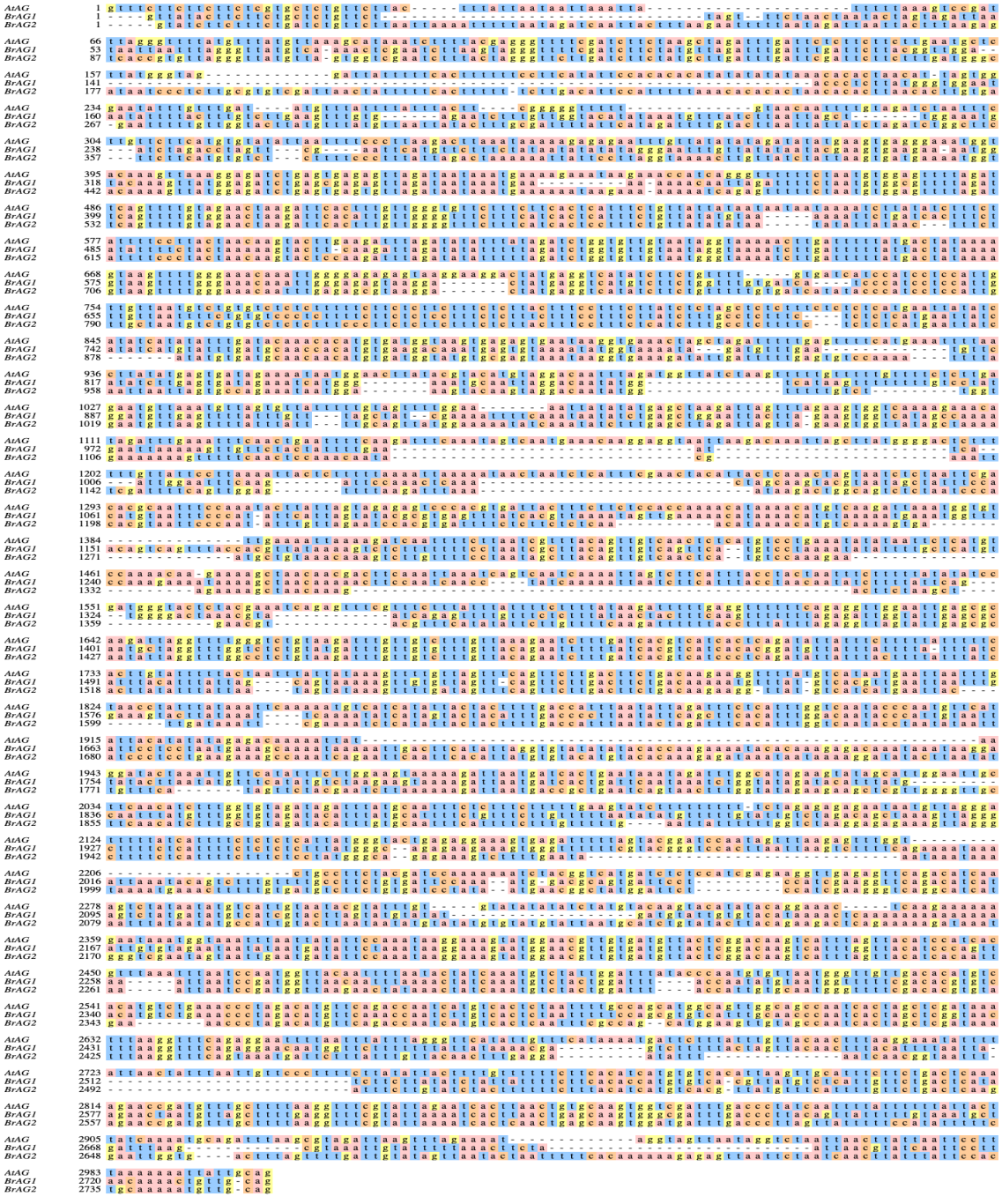

Figure 3.10: Alignment of large intron sequences of *AG* homologs. Alignment of the DNA sequence of the second intron in *AtAG* and both of the first introns in *BrAG1* and *BrAG2*. The sequences were aligned using the *Clustal Omega Multiple Sequence Alignment* tool from EMBL-EBI (Madeira et al., 2022). The results were visualised using *Jalview* software.

The percentage identities of the three large introns are listed in Table 3.8.

Table 3.8: Percentage identity matrix for intronic sequences of *AG* homologs. The percentage of identity of the second intron of *AtAG* with the first introns of *BrAG1* and *BrAG2*.

	1	**2**	**3**
1. *AtAG* intron	100.00	82.21	82.55
2. *BrAG1* intron	82.21	100	82.08
3. *BrAG2* intron	82.55	82.08	100.00

From the alignment, it is clear that there are large regions of the introns that are nearly identical. Given this high degree of conservation in a non-coding DNA sequence where mutations can normally accumulate, it appears that there is functional constraint that has prevented the sequences from diversifying over time. To investigate whether any differences in transcription binding sites could be detected in the large introns of the *AG* homologs, *JASPAR* was used again. For this analysis, I focused initially on MADS-domain transcription factor binding sites and specifically on AG. The results of the analysis are displayed in Table 3.9.

The results shown are those obtained using a stringent 90% relative profile score. In general, a substantially lower number of putative binding sites was detected for *BrAG1* (17 sites) than for *AtAG* or *BrAG2* (35 and 32 sites respectively). There is a similar pattern to these results to those described in Section 3.2.4, with the largest differences between binding sites in *AtAG*, *BrAG1* and *BrAG2* observed for the floral homeotic transcription factors AP1, SEP1, SEP4 and SHP1. What's more, while there are several predicted AG binding sites in the large introns of *AtAG* and *BrAG2* in similar positions (Figure 3.11), these seem to be partially absent in *BrAG1*. These differences may explain, at least in part, the observed differences in expression between the two *Brassica AG* homologs as seen in Section 3.2.2.

Table 3.9: Predicted MADS-domain transcription factor binding sites for regulatory introns of *AG* homologs. The names of MADS-domain transcription factors and the corresponding *JASPAR* matrix IDs are listed. The results shown are the number of sites predicted using the default relative profile score threshold of 90%.

Gene	Matrix ID	AtAG	BrAG1	BrAG2
AG	MA0005.1	3	1	2
	MA0005.2	2	0	2
AGL6	MA1205.1	1	1	1
AGL13	MA1204.1	1	1	1
AGL15	MA0548.1	1	1	1
	MA0548.2	1	1	1
AGL16	MA1199.1	1	1	1
AGL27	MA1012.1	1	0	1
AGL42	MA1201.1	2	2	3
AGL55	MA1202.1	0	0	2
AGL63	MA1203.1	1	1	1
AP1	MA0940.1	7	1	4
AP3	MA0556.1	1	1	1
FLC	MA0558.1	1	1	1
PI	MA0559.1	2	1	2
SEP1	MA0584.1	3	0	1
SEP3	MA0563.1	1	0	1
SEP4	MA0001.1	1	2	1
	MA0001.2	1	0	1
SHP1	MA0585.1	2	0	2
SOC1	MA0554.1	1	1	1
SVP	MA0555.1	1	1	1
Total		**35**	**17**	**32**

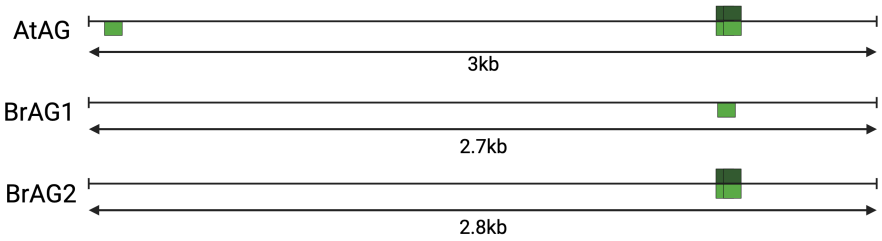

Figure 3.11: Predicted AG binding sites in the intronic region of *AG* homologs. The location of predicted AG binding sites in the intronic regions of *AtAG*, *BrAG1* and *BrAG2*. Dark green indicates a binding site on the forward strand and light green indicates a binding site present on the reverse strand of the DNA. Figure generated with BioRender.com.

The results of the analysis of putative binding sites for LFY are shown in Table 3.10 and in Figure 3.12. Following the trend seen for the upstream promoter region of the

AG homologs, there are a similar number of sites detected for LFY - four for *AtAG* and *BrAG1* and six for *BrAG2*. The location of the putative binding sites across the three sequences is similar, suggesting that perhaps these sites are not likely to contribute to the difference in the expression patterns of BrAG1 and BrAG2 as seen in Section 3.2.2.

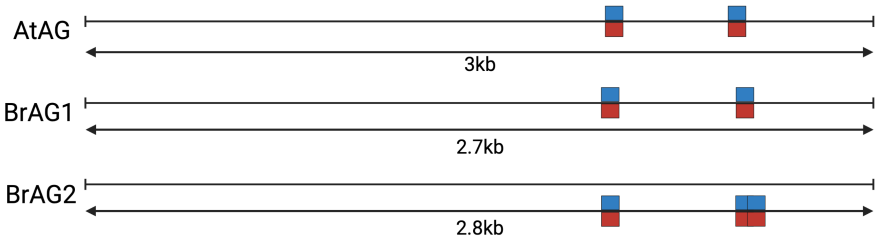

Figure 3.12: Predicted LFY binding sites in regulatory introns of *AG* homologs. The location of LFY binding sites in the regulatory introns of *AtAG*, *BrAG1* and *BrAG2*. Blue indicates a binding site on the forward DNA strand and red indicates a binding site on the reverse strand. Figure generated with BioRender.com.

Table 3.10: Predicted LFY binding sites for large intronic sequences of *AG* homologs. The number of LFY binding sites detected in the first intron of *BrAG1* and *BrAG2* and the second intron of *AtAG*.

Intronic Sequences				
Gene	**Matrix ID**	**AtAG**	**BrAG1**	**BrAG2**
LFY	MA0590.1	4	4	6

Similar to the predicted binding sites in the promoter regions of *AG* homologs, further work will be needed to investigate the function of the predicted binding sites in the introns, for example by disrupting individual candidate binding sites through genome-editing approaches.

3.2.6 Results of Complementation Experiments

To further characterise the function of *BrAG1* and *BrAG2* and to test whether they are true orthologs of *AtAG*, I conducted a second line of experiments in which I tested whether they could rescue the strong *ag-1* mutant allele from *Arabidopsis*. *Ag-1* was

generated by EMS mutagenesis (Bowman et al., 1989) and contains a splice acceptor site mutation at the sixth intron (Yanofsky et al., 1990). The floral phenotype of *ag-1* is shown in Figure 3.13. In the mutant, flowers develop sepals as usual in the first whorl. Petals form in the second whorl as well as in the third whorl in place of stamens. In the fourth whorl, an additional mutant flower develops instead of a gynoecium.

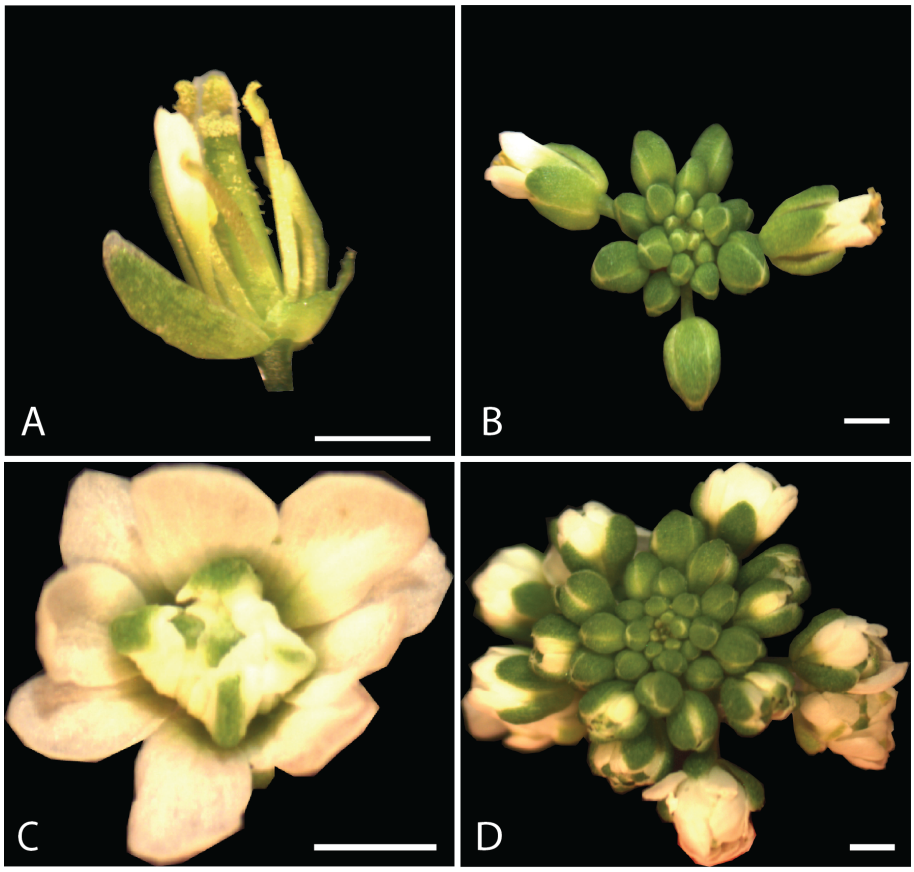

Figure 3.13: Floral phenotype of *ag-1* mutants. (A) A typical flower from a wild-type plant (accession: L-*er*). (B) Top view of the primary inflorescence of a wild-type plant. (C) Flower of an *ag-1* plant. The flowers completely lack reproductive organs whilst there is an increase in the number of sepals and petals developing. In the place of the reproductive organs, there is an additional flower developing, again with sepals developing as normal but with a greater number of petals. (D) A top view of the primary inflorescence of an *ag-1* plant. Scale bars: 1 mm.

For the complementation experiments, the CDS (including the 3'UTRs) of *BrAG1* and *BrAG2* were placed under the control of the CaMV 35S promoter. An equivalent construct was generated for *AtAG* as a control. These overexpression lines were gener-

ated in order to assess the potential of BrAG1 and BrAG2 to functionally compensate for AtAG regardless of expression domain. In order to assess if BrAG1 and BrAG2 could functionally compensate for AtAG when expressed only in the AtAG domain, the CDS of all three genes were fused to the *AtAG* promoter. To this end, I used the complete second intron of *AtAG* which had been shown previously to drive expression in the normal domain of *AtAG* (by Lenhard et al., 2001). The CaMV 35S promoter was used in order to determine if *BrAG1* and *BrAG2* could provide a rescue of *ag-1* when over-expressed and replicate the 35S:AtAG line described in Mizukami and Ma (1992). The *AtAG* promoter was used to determine if *BrAG1* and *BrAG2* could provide a rescue of the *ag-1* mutant phenotype when expressed only in the *AtAG* expression domain. All constructs were transformed into *ag-1/+* as well as into wild-type plants (accession: L-*er*) and transformants were analysed for phenotypic alterations. The results of these experiments are described below.

3.2.6.1 Complementation Experiments Using a p35S:*AtAG* Construct

Mizukami and Ma (1992) previously generated *AtAG* overexpression lines which were described as having flowers where sepals are replaced by fused carpels with stigmatic papillae, and that lacked petals. Furthermore, the leaves of the plants were curled up at the lateral margins. The *AtAG* overexpression lines generated in the course of this study are shown in Figure 3.14.

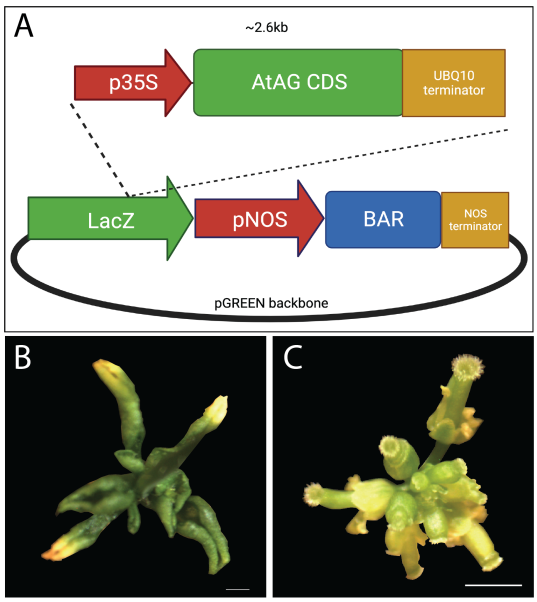

Figure 3.14: Effects of *AtAG* overexpression. (A) Schematic of the genetic construct used to generate the *AtAG* overexpression lines. p35S, CaMV 35S promoter sequence; AtAG CDS, the coding DNA sequence of *AtAG*; Terminator, the *UBQ10* terminator; LacZ, lacZ a modified version of the β- galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA^{TM} resistance gene; NOS terminator, nopaline synthase terminator sequence. Figure generated using Biorender.com (B) Plant carrying the p35S::*AtAG* construct with a curly leaf phenotype and (C), flowers produced by a plant overexpressing *AtAG*. Scale bars: 1 mm.

Similar to what had been described by Mizukami and Ma (1992), all transgenic lines I obtained (n=7 for *ag-1*; n=5 for L-*er*) showed a strong curly leaf phenotype (Figure 3.14B-C). Flowers lacked petals and the outer whorl organs were carpelloid with stigmatic papillae. The construct always provided a rescue of the *ag-1* mutant phenotype in plants that were confirmed to be homozygous mutants *via* dCAPS genotyping (Section 2.3.4).

3.2.6.2 Complementation Experiments Using a p35S:*BrAG1* Construct

All independent lines (n=8 for *ag-1*; n=8 for L-*er*) generated for the *BrAG1* overexpression construct showed phenotypes (Figure 3.15) that were very similar to the ones observed for the p35S:*AtAG* construct described above.

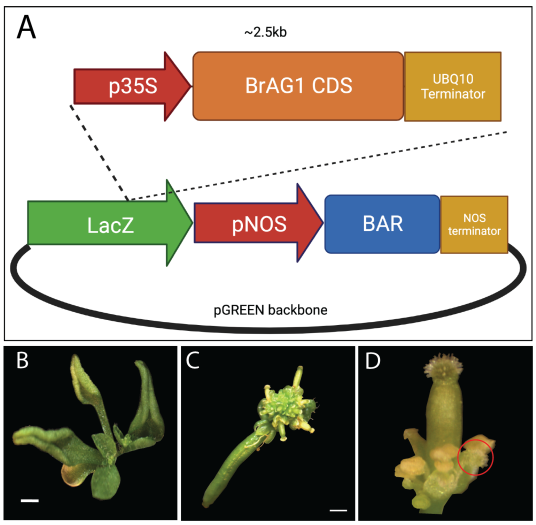

Figure 3.15: Effects of overexpression of *BrAG1*. (A) Schematic of the genetic construct used to generate the *BrAG1* overexpression lines. p35S, CaMV 35S promoter sequence; BrAG1 CDS, the coding DNA sequence of *BrAG1*; Terminator, the *UBQ10* terminator; LacZ, lacZ a modified version of the β- galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA^{TM} resistance gene; NOS terminator, nopaline synthase terminator sequence. Figure generated with Biorender.com. (B) The curly leaf phenotype seen in every p35S:*BrAG1* transgenic plant. (C) Typical primary inflorescence observed in these lines. None of the flowers develop petals. (D) Flower with no petals and the stigmatic papillae highlighted by a red circle. Scale bars: 1 mm.

The construct always provided a rescue of the *ag-1* mutant phenotype in plants that were confirmed to be homozygous mutants *via* dCAPS genotyping (Section 2.3.4). Thus, it appears that *BrAG1* can functionally replace *AtAG* when overexpressed.

3.2.6.3 Complementation Experiments Using a p35S:*BrAG2* Construct

All independent lines (n=6 for *ag-1*; n=4 for L-*er*) generated for the *BrAG2* over-expression construct showed a curly leaf phenotype as described above for the *AtAG* and *BrAG1* over-expression lines (Figure 3.16).

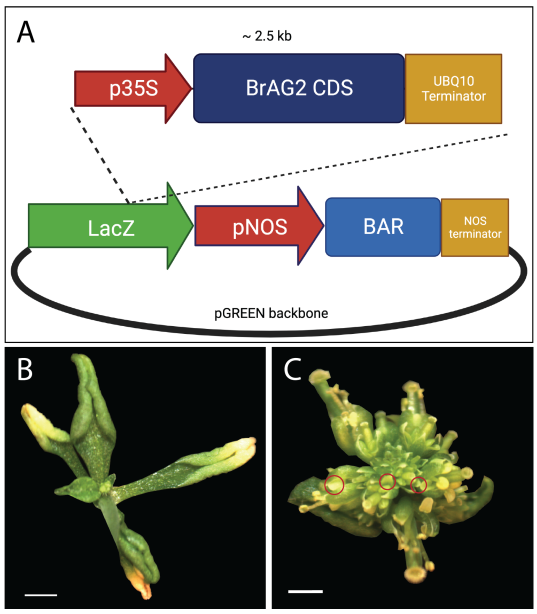

Figure 3.16: Effect of *BrAG2* overexpression. (A) Schematic of the genetic construct used to generate the *BrAG2* overexpression lines. p35S, CaMV 35S promoter sequence; BrAG2 CDS, the coding DNA sequence of *BrAG2*; Terminator, the *UBQ10* terminator; LacZ, lacZ a modified version of the β- galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA^{TM} resistance gene; NOS terminator, nopaline synthase terminator sequence. Figure generated with Biorender.com. (B) The curly leaf phenotype seen in all p35S:*BrAG2* lines generated. (C) Typical primary inflorescence observed in these lines. Flowers developed petals which are indicated by red circles. Scale bars: 1 mm

---

In contrast, unlike *AtAG* and *BrAG1* over-expression lines, the identity of floral organs was less affected in the *BrAG2* over-expression lines and all flowers formed petals. At the same time, the construct always led to a rescue of the *ag-1* mutant phenotype (as confirmed by dCAPS genotyping; Section 2.3.4), suggesting that *BrAG2* can functionally replace *AtAG*.

3.2.6.4 Complementation Experiments Using pAG:*AtAG*

In addition to the constitutive over-expression lines described above, I also drove expression of the *AG* homologs from the regulatory region of *AtAG* to test whether they could rescue *ag-1* mutants when expressed in the normal domain of *AtAG*. As a control, I first generated lines carrying a pAG:*AtAG* construct. A total of 11 independent lines (n=7 for ag-1; n=4 for L-er) were isolated. The construct always led to a rescue of the

ag-1 mutant phenotype (as confirmed by dCAPS genotyping; Section 2.3.4) with stamen and gynoecium development being restored (Figure 3.17).

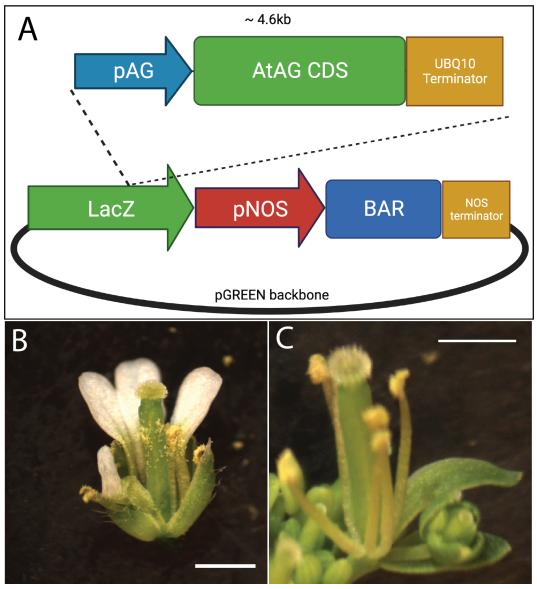

Figure 3.17: Results for the expression of pAG:*AtAG* constructs. (A) Schematic of the genetic construct used to generate the pAG:*AtAG* lines. pAG, *AtAG* promoter sequence as described in Section 3.2.8; AtAG CDS, the coding DNA sequence of *AtAG*; Terminator, the *UBQ10* terminator; LacZ, lacZ a modified version of the β- galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA^{TM} resistance gene; NOS terminator, nopaline synthase terminator sequence. Figure generated with Biorender.com. (B) The floral phenotype of *ag-1* plants containing the transgene. (C) The floral phenotype of L-*er* plants containing the pAG:*AtAG* transgene.

Flowers of wild-type plants carrying the construct lacked petals and produced secondary flowers from the floral axis, a phenotype reminiscent of *ap1-1* mutants (Irish and Sussex, 1990). Because AG is known to repress *AP1* (Gustafson-Brown et al., 1994), this result suggests that in a wild-type background, where endogenous *AG* is active, the addition of another copy of the gene resulted in a dosage effect that led to *AP1* down-regulation.

3.2.6.5 Complementation Experiments Using pAG:*BrAG1*

A total of 13 independent lines (n= 6 for *ag-1*; n=7 for L-*er*) were generated containing *BrAG1* fused to the *AtAG* promoter. The construct always led to a rescue of the *ag-1* mutant phenotype (as confirmed by dCAPS genotyping; Section 2.3.4). All rescues in the *ag-1* background displayed the phenotype shown in Figure 3.18B with a restoration of the stamen number and the gynoecium.

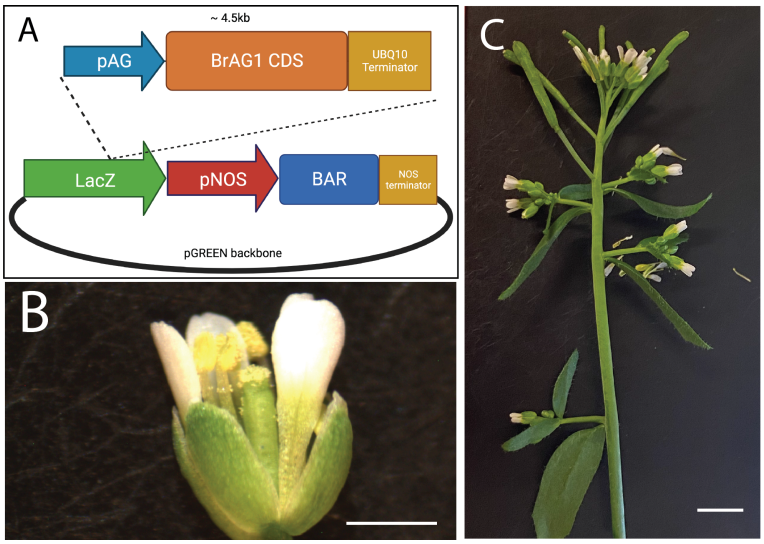

Figure 3.18: Results for the expression of pAG:*BrAG1* constructs. (A) Schematic of the genetic construct used to generate the pAG:*BrAG1* lines. pAG, *AtAG* promoter sequence as described in Section 3.2.8; BrAG1 CDS, the coding DNA sequence of *BrAG1*; Terminator, the *UBQ10* terminator; LacZ, lacZ a modified version of the β-galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA^{TM} resistance gene; NOS terminator, nopaline synthase terminator sequence. Figure generated using Biorender.com (B) The floral phenotype of *ag-1* plants containing the transgene. Scale bar: 1 mm. (C) The inflorescence of a L-*er* plant containing the transgene. Scale bar: 5 mm.

Figure 3.18C shows a phenotype seen in wild-type plants carrying the construct in which there are a greater number of secondary inflorescences produced by plants containing the transgene. This could be due to *BrAG1* repressing *AP1* as C class genes act to repress A class genes but perhaps this is not happening as efficiently as *AtAG*.

3.2.6.6 Complementation Experiments Using pAG:*BrAG2*

A total of 15 independent lines (n=9 for *ag-1*; n=6 for L-*er*) were generating for the construct containing *BrAG2* fused to the *AtAG* promoter (Figure 3.19A).

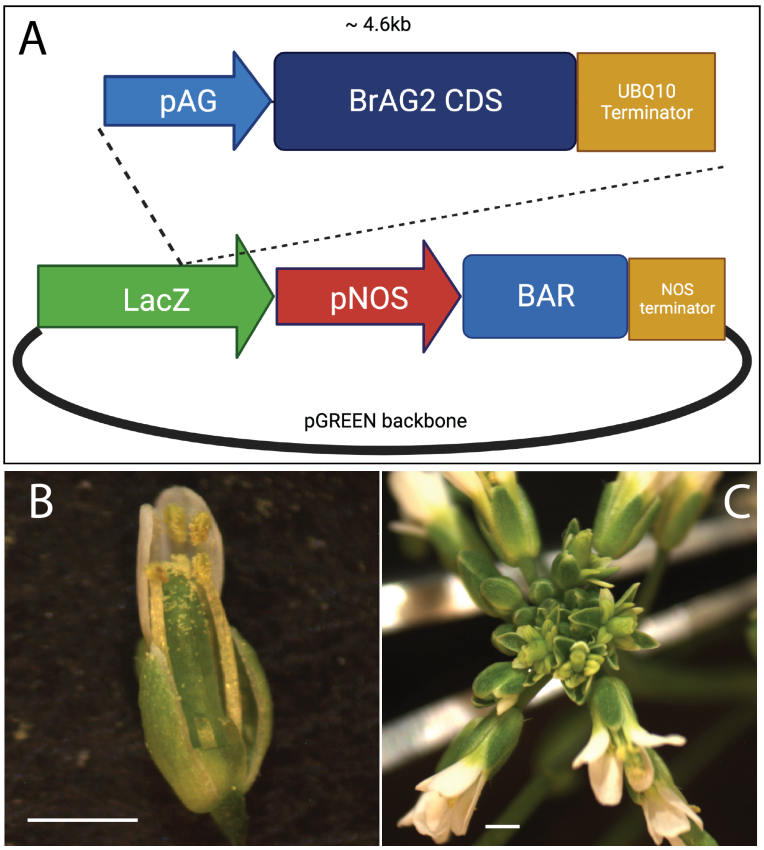

Figure 3.19: Results for the expression of pAG:*BrAG2* constructs. (A) Schematic of the genetic construct used to generate the pAG:*BrAG2* lines. pAG, *AtAG* promoter sequence as described in Section 3.2.8; BrAG2 CDS, the coding DNA sequence of *BrAG2*; Terminator, the *UBQ10* terminator; LacZ, lacZ a modified version of the β-galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA^{TM} resistance gene; NOS terminator, nopaline synthase terminator sequence. Figure generated with Biorender.com. (B) The floral phenotype of *ag-1* plants containing the transgene. (C) The inflorescence phenotype of L-*er* plants containing the pAG:*BrAG2* transgene

---

The construct always led to the complete rescue in the *ag-1* background (as confirmed by dCAPS genotyping; Section 2.3.4)(Figure 3.19B). The phenotype seen in all of the lines generated in a wild-type background is shown in Figure 3.19C. The first ten or so flowers produced by these plants resembled wild-type flowers. After this, the

flowers emerging showed a phenotype similar to the one observed in p35S:*BrAG2* lines (Figure 3.16C) with flowers being smaller in size with small green immature petals.

These experiments showed that under the 35S promoter both *BrAG1* and *BrAG2* are able to reproduce the curly leaf phenotype seen in *AtAG* overexpression. There is a difference in the rescues of the *ag-1* floral phenotypes. As seen in Figure 3.15C-D, the flowers of p35S:*BrAG1* develop with stigmatic papillae tissue in the first whorl and lack petals due to repression of A-class genes, as is seen in *AtAG* overexpression lines. In the case of p35S:*BrAG2* (Figure 3.16C), the flowers develop petals and lack stigmatic papillae in the first whorl, indicating that A-class genes are not being repressed to the same level seen in *AtAG* or *BrAG1* overexpression lines.

The expression of *BrAG1* under the *AtAG* promoter was able to provide rescue of the *ag-1* floral phenotype with a restoration of the stamen number and the gynoecium in all flowers. A similar rescue is seen in pAG:*BrAG2* but only in the first ten or so flowers with flowers produced afterwards containing stamens and gynoecia that appear underdeveloped (Figure 3.19C). This suggests that *BrAG2* can only provide rescue for the *ag-1* floral phenotype early on in the flower development process and later on, it alone cannot drive the complete development of reproductive organs. As shown in Section 3.2.1, the CDS of *BrAG1* and *BrAG2* have a percentage of sequence identity ($\approx$94%) so the difference in rescue is unlikely to be down to differences in the sequences. It is possibly caused by the differences in the sequences of the 3'UTRs of *BrAG1* and *BrAG2* as discussed in Section 3.2.3 resulting in different levels of mRNA stability or translation efficiency.

3.2.7 Identification of Mutant Alleles for *BrAG1* and *BrAG2*

As shown above (Section 3.2.1), *BrAG1* and *BrAG2* encode proteins that are very similar to AG from *Arabidopsis*. Furthermore, their expression patterns, as assessed by

RT-qPCR, mimic that of *AtAG* (see Section 3.2.2). It, therefore, appeared very likely that *BrAG1* and *BrAG2* are orthologs of *AtAG*.

To test whether these genes are indeed C-function genes, I aimed to analyse the effects that their inactivation has on flower development in *B. rapa*. I also aimed to identify a mutant line that was similar in nature to the *Arabidopsis* mutant line *ag-10* in order to only mildly knockdown *AG* function to retain fertility whilst allowing for the potential of trichome formation on the gynoecia. As shown in Section 3.2.2, *BrAG1* is expressed more highly than *BrAG2* in gynoecium tissue but the expression of both is detected. Given the potential for redundancy between *BrAG1* and *BrAG2*, I thought it possible that a *brag1* mutant line alone may not be sufficient to replicate the *ag-10* phenotype and a double mutant may be needed in order to obtain enough of a knockdown of *AG* function. Therefore, the strategy decided upon was to obtain mutant alleles for both *BrAG1* and *BrAG2* and to obtain both a strong mutant allele and a weak mutant allele for each in order to generate double mutants with various levels of loss of *AG* function.

To this end, *B. rapa* mutant lines were selected from the catalogue of commercially available TILLING lines from RevGenUK at the John Innes Centre in Norwich, UK. In the case of *BrAG1*, 35 mutant lines were available whilst for *BrAG2* the RevGenUK database contained 32 mutant lines. The vast majority of these lines available were deemed not suitable for our work because the mutations involved were synonymous mutations, mutations in intronic sequences, or missense mutations outside of conserved regions that are unlikely to affect protein function.

For *BrAG1*, two TILLING lines were chosen for analysis. In the first, JI31652-a, a C to T conversion had occurred at position 10 of the CDS which led to a premature stop codon at position 4 of the protein which was assumed to be a null allele. This mutant line was named *brag1-1*. In the second line that was chosen for *BrAG1*, JI30258-b, a G to A conversion at position 418 of the CDS causes a glutamic acid at position 140 of

the protein to be converted into lysine. Notably, this mutation mirrors that of the weak *ag-10* allele in *Arabidopsis* that was used to induce trichomes on carpel valves (Ji et al., 2011; Ó'Maoiléidigh et al., 2018). This mutant line was named *brag1-2*.

In the case of *BrAG2*, there was no allele available that would introduce an early stop codon in the CDS. The best candidate for a line with a possibly strong effect on gene function was JI30013-a. This line contained a G to A conversion at base 231 in a splice donor site. The guanine at the +1 position of the splice donor site that is mutated in the line is highly conserved in plants, suggesting that the mutation may have severe effects on the splicing of the mRNA (Lorković et al., 2000). This mutant line was named *brag2-1*. The second allele chosen for *BrAG2* was line JI31547-b. This line had a G to A conversion at position 415 of the CDS, which results in the glutamic acid at position 139 of the protein being converted to a lysine. This line is similar in nature to *brag1-2* (see above) and was named *brag2-2*. The mutations described above are summarized in Figure 3.20.

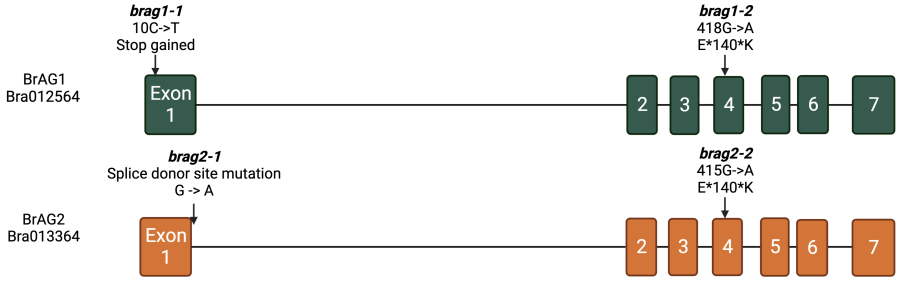

Figure 3.20: Mutant alleles for *BrAG* genes. Schematic showing mutations in *BrAG1* (top) and *BrAG2* (bottom) in the lines obtained from RevGenUK. Exons are numbered and shown as coloured boxes; introns are represented by black lines. The *brag1-1* line contains an early stop codon in exon 1 and *brag1-2* contains a missense mutation that results in a lysine at position 140 instead of a glutamic acid. *brag2-2* contains a missense mutation that results in the same amino acid substitution as *brag1-2*. *Brag2-1* contains a mutation at the highly conserved splice donor site at the first intron of *BrAG2*. Figure generated with BioRender.com.

3.2.7.1 Characterisation of *brag2-1*, a *BrAG2* Allele with a Splice Donor Site Mutation

As mentioned in Section 3.2.7, the *brag2-1* allele contained a mutation in a conserved residue of a splice donor site. To test whether splicing of the *BrAG2* transcript was affected in this line, flowers were collected from homozygous *brag2-1* plants and from wild-type plants (line: R-o-18). RNA was extracted from the tissue and used to synthesise cDNA as described in Section 2.2.1. This cDNA was then used as a PCR template to amplify the CDS of *BrAG2*. The expected length of the wild-type amplicon was 1 kb. The PCR products obtained were analysed by agarose gel electrophoresis (Figure 3.21). Both the homozygous *brag2-1* mutant and R-o-18 produced an amplicon 1 kb in length but the *brag2-1* mutant also produced an additional faint band of approximately 250 bp.

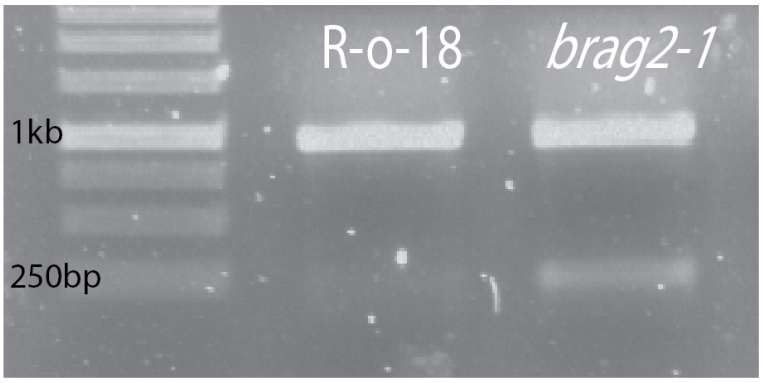

Figure 3.21: Analysis of *BrAG2* splicing in *brag2-1* mutants. The CDS sequence of *BrAG2* was PCR-amplified from cDNA generated from flowers of *brag2-1* mutants and wild-type plants (R-o-18) plants, respectively. PCR fragments were resolved by agarose gel electrophoresis. The expected amplicon size was approximately 1kb. Note the additional band of approximately 250bp in the brag2-1 sample. The leftmost lane contains a 1 kb size marker (ThermoFisher). The bands corresponding to 250 bp and 1 kb are indicated.

The 1 kb bands and the 250 bp band were gel extracted, sequenced and the results analysed. The analysis showed that the mutation in the splice donor site in *brag2-1* is indeed affecting the processing of the *BrAG2* transcript. In the case of the 1 kb transcripts, the cDNA from *brag2-1* flowers contains a 15 nt deletion relative to the cDNA from the wild-type gene, which starts at the affected splice donor site and ends at the

next "GT" site which is contained within the CDS and is apparently functioning as a cryptic splice site in this case. The deletion in the case of the 1 kb transcript leaves the CDS in-frame and is aligned to the wild-type *BrAG2* CDS in Figure 3.22.

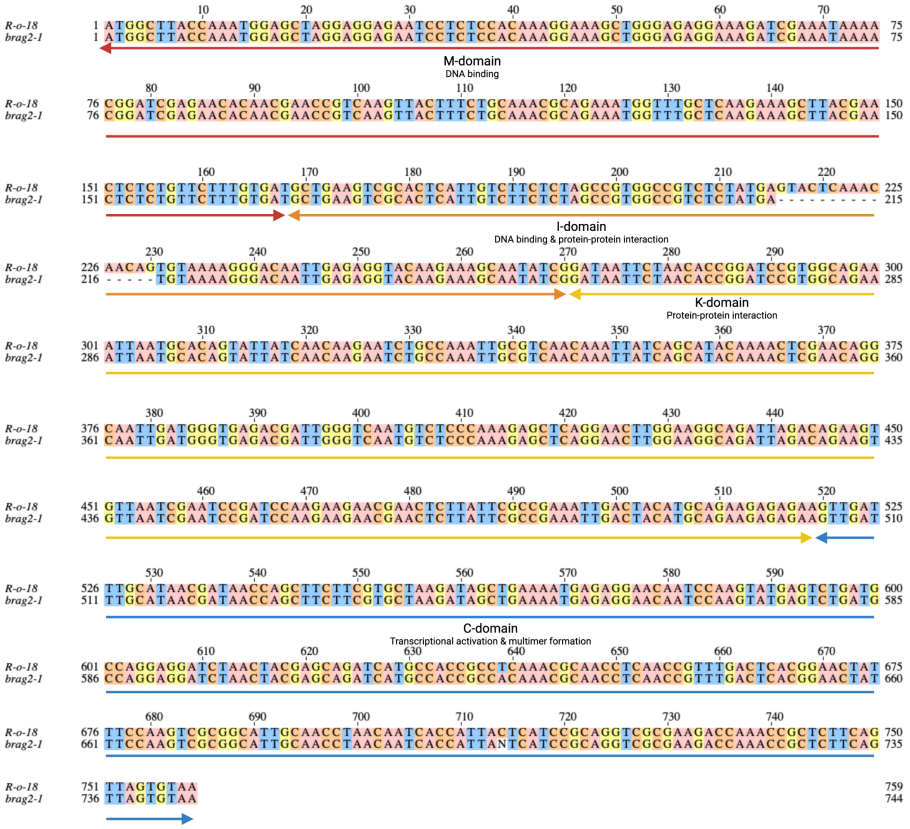

Figure 3.22: Alignment of cDNA sequences for *BrAG2* obtained by PCR amplification from R-o-18 and *brag2-1* mutant plants, respectively. Alignment of cDNA sequences from R-o-18 and *brag2-1*. Note the 15 nt deletion in *brag2-1* starting at position 216. The sequences were aligned using the *Clustal Omega Multiple Sequence Alignment* tool from EMBL-EBI (Madeira et al., 2022). The results were visualised using *Jalview* software. The domains of MADS domain transcription factors and their functions are annotated. Figure generated with BioRender.com

---

The effect that this deletion is predicted to have on the sequence of the *BrAG2* protein is shown in Figure 3.23.

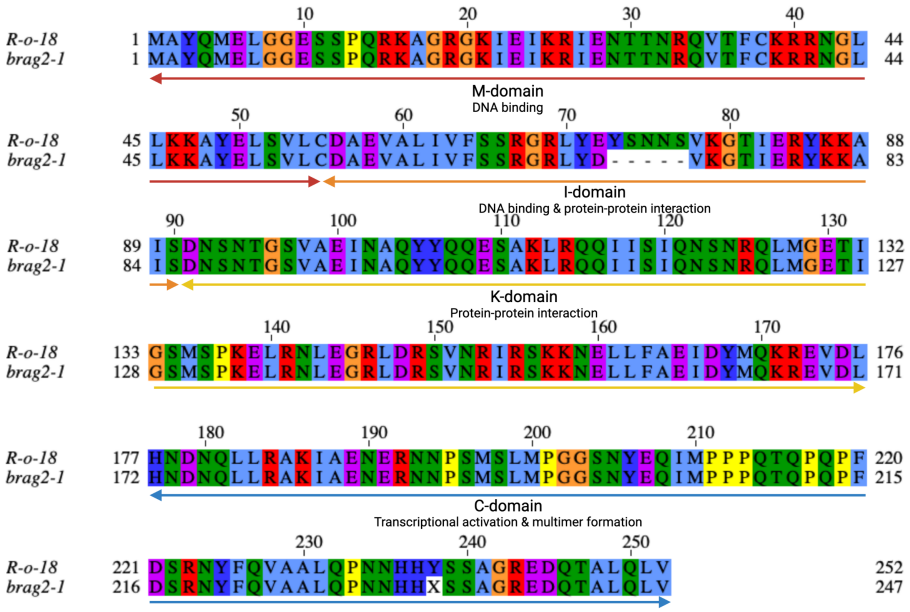

Figure 3.23: Alignment of the amino acid sequence of BrAG2 from R-o-18 and from *brag2-1*. Alignment of the two protein sequences shows that the protein produced in the *brag2-1* line lacks five amino acids which are in the MADS-domain of the structure. The sequences were aligned using the *Clustal Omega Multiple Sequence Alignment* tool from EMBL-EBI (Madeira et al., 2022). The results were visualised using *Jalview* software. The following colour scheme was used to denote amino acids that belong to the same residue groups: blue: hydrophobic; red: positively charged; magenta: negatively charged; green: polar; pink: cysteines; orange: glycines; yellow: prolines; cyan: aromatic. The domains of MADS domain transcription factors and their functions are annotated. Figure generated with BioRender.com

---

Sequencing of the smaller 250bp amplicon isolated from *brag2-1* showed that it was the result of another splicing event using a different cryptic splice site, which resulted in a transcript that was no longer in-frame. The faint band obtained in the PCR reaction suggested that this splicing event occurred less frequently than the one leading to the ≈1 kb fragment.

The *BrAG2* protein produced by *brag2-1* mutants lacks in all probability five amino acids from position 73 to 78 and contains a single amino acid conversion from glutamic acid to aspartic acid at position 72. Notably, these positions are at the C-terminal end of the MADS-domain, which in *AtAG* spans positions 19 to 73 (Mizukami et al., 1996). The deletion also encompasses the beginning of the I domain, which has recently been

shown to be essential for the DNA binding and dimerization specificity of AtAG (Lai et al., 2021). Thus, the deletion, in addition to affecting the folding of the protein, may severely impair its DNA binding activity and dimerization ability. I concluded that the mutation in the *brag2-1* allele likely affected the functionality of the *BrAG2* protein and as a result, may effectively be a strong mutant line. As discussed earlier in this section, I aimed to obtain both strong and weak mutant alleles for both *BrAG1* and *BrAG2* in order to generate double mutants of various strengths of knockout of both with the aim of replicating the *ag-10* line in *Arabidopsis*. The results of these experiments indicate that this mutant allele will be a suitable strong mutant allele for *BrAG2* and this is further investigated in Sections 3.2.8 and 3.2.9.

3.2.8 Analysis of *brag* Single-Mutants

As advised by RevGenUK, all TILLING lines obtained were back-crossed with wild-type plants in order to improve plant health by reducing their mutational load and ensure that any phenotypic differences observed could be attributed to the mutation that defines each line. All *BrAG* mutant lines were back-crossed with R-o-18 or C634 (see Section 5.3.1 for C634) accession at least 4 times.

3.2.8.1 Floral Phenotypes of *brag1-1* Mutant Plants

Plants heterozygous and homozygous for the *brag1-1* mutation were initially isolated by PCR amplification and by sequencing as described in Section 2.3.2. Subsequently, a DMAS genotyping assay was developed (described in Section 2.3.3) to be able to rapidly follow the mutation in a segregating population. To assess the phenotypic consequences of the *brag1-1* mutation both heterozygous and homozygous plants were examined once four backcrosses were performed.

In flowers of heterozygous *brag1-1* plants, sepals did not fully enclose the inner

whorls and petals could be seen (Figure 3.24A). This was likely due to the stigma of the flower being bulged and larger than wild-type flowers at this stage (Figure 3.24B). Stamens of mature flowers appeared slightly shorter than those of R-o-18 flowers. Normally, at the equivalent of stage 13 in *Arabidopsis*, the anther of the long stamens have grown above the stigma of the gynoecium and the short stamens are approximately the same height (Smyth et al., 1990). It appears that this was due to the style being elongated compared to the wild-type flowers. Some of the flowers of these plants had gynoecia and siliques that were slightly misshapen and bent and they generally had fewer stigma papillae (Figure 3.24D-E).

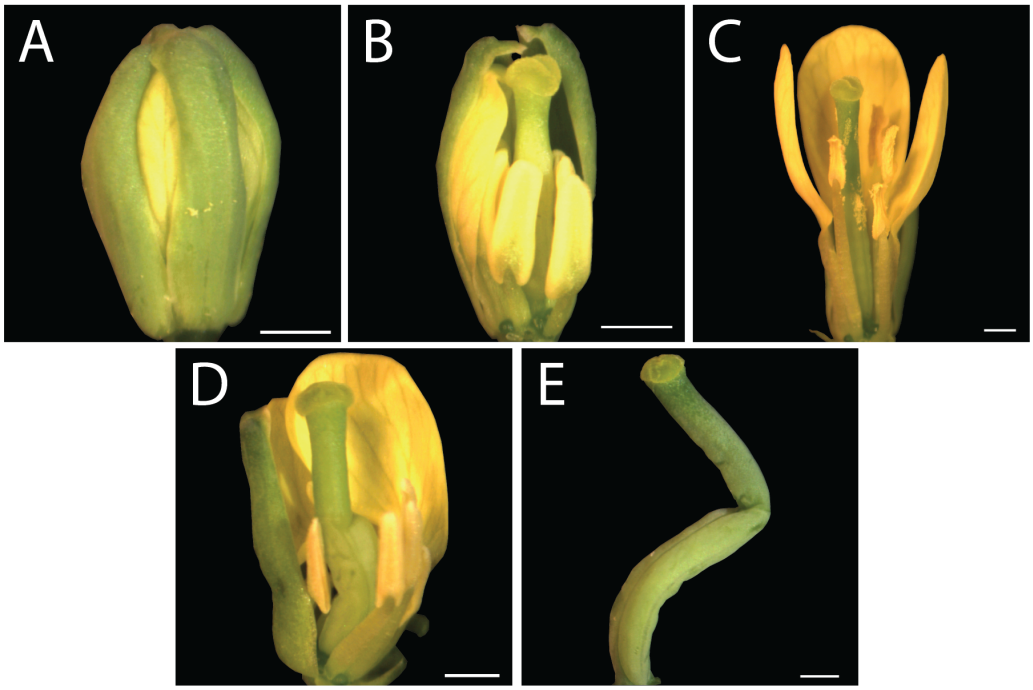

Figure 3.24: Floral phenotypes of heterozygous *brag1-1* plants. (A) The sepals do not enclose the inner whorls in flowers of these plants. (B) A flower with sepals and petals removed to view the reproductive organs. The stigma of the flower is larger and more domed than the stigma of wild-type flowers of the same age (C) A mature flower. The stamens do not reach the top of the gynoecium due to the style being elongated compared to the wild-type flowers. (D) Phenotype that arises in *brag1-1* flowers occasionally where the gynoecium appears kinked and (E), the resulting silique is also bent. Scale bars: 1 mm.

Flowers of plants homozygous for *brag1-1* also had sepals that did not fully engulf the floral organs (Figure 3.25A). Petals appeared normal but the reproductive floral or-

gans were affected (Figure 3.25B).

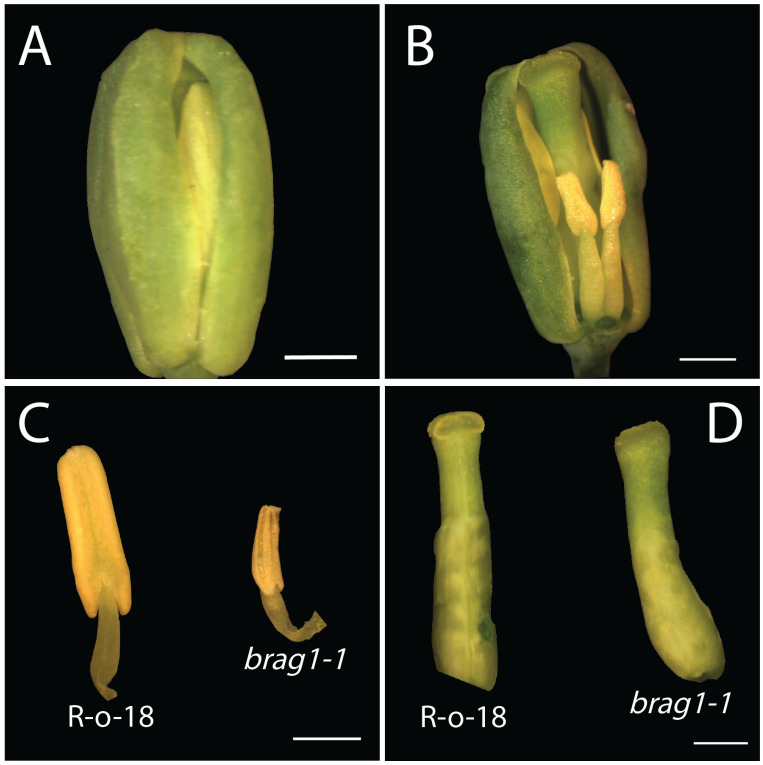

Figure 3.25: Floral phenotypes of homozygous *brag1-1* plants. (A) The sepals of flowers of *brag1-1* mutants do not fully enclose the other floral organs as in R-o-18. (B) The stamens of *brag1-1* flowers are shorter and have much smaller anthers than R-o-18 stamens. Some sepals and petals were removed to show the reproductive organs. (C) Stamens of R-o-18 and *brag1-1* plants are shown in greater detail. (D) Gynoecia of *brag1-1* and R-o-18 flowers. The stigma of the *brag1-1* gynoecium is less defined and less domed when compared to R-o-18. The style of the *brag1-1* gynoecium is also less defined as a separate and distinct region from the ovary. Scale bars: 1 mm.

While their number appeared to be unchanged, stamens were shorter (Figure 3.25C) mainly because their anthers were much smaller than those of the wild type. As a result, the stamens in this line never reached the height of the stigma and flowers needed to be manually pollinated but were fertile when manually pollinated. In the fourth whorl, both the style and stigma of the gynoecium were abnormal (Figure 3.25D) and appeared less defined and the stigma less domed than in the wild-type with fewer stigma papillae again. These mutant phenotypes were observed in all flowers examined.

3.2.8.2 Floral Phenotypes of *brag1-2* Mutant Plants

As described above for *brag1-1*, plants heterozygous and homozygous for the *brag1-2* mutation were initially isolated by PCR amplification and sequencing and subsequently, a DMAS genotyping assay was developed. While plants heterozygous for the *brag1-2* showed no obvious mutant phenotype, the flowers of homozygous plants were mildly affected with siliques being smaller than those seen in R-o-18 (Figure 3.26). This difference in the length appeared to be due to the ovary being shorter in *brag1-2* than in the wild type. As a result, *brag1-2* plants produced fewer seeds than R-o-18 plants.

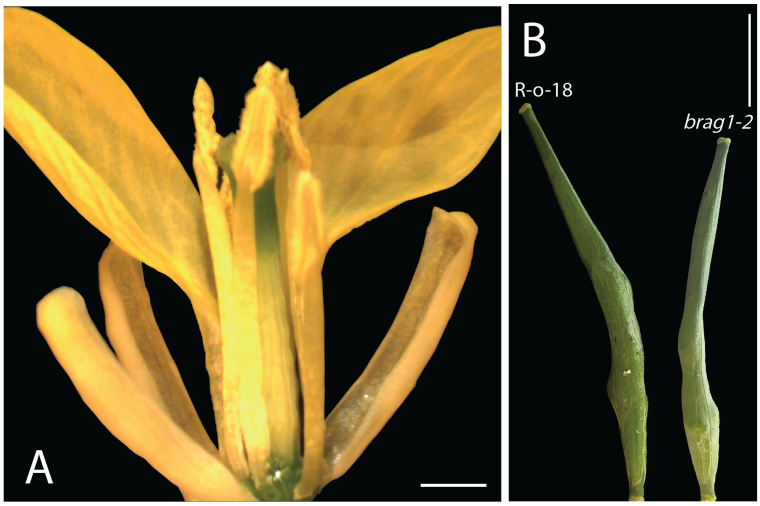

Figure 3.26: Floral phenotype of homozygous *brag1-2* plants. (A) A flower of a *brag1-2* plant showing no obvious floral defects. Scale bar: 1 mm. (B) The siliques of *brag1-2* plants (on the right) are shorter and contain fewer seeds than those produced by R-o-18 plants (on the left). Scale bar: 1 cm.

In general, the mutant defects of *brag1-2* plants appeared much weaker than those of *brag1-1*, which was expected given the different nature of the mutations in the two alleles (see Section 3.2.7 above).

3.2.8.3 Floral Phenotype of *brag2-1* Mutant Plants

As for the other mutant lines described above, plants heterozygous and homozygous for the *brag2-1* mutation were initially isolated by PCR-amplification and sequencing (Section 2.3.2), followed by the development of a DMAS genotyping assay (Section 2.3.3).

Plants that were heterozygous for the *brag2-1* mutation showed a floral phenotype that was almost identical to wild-type flowers (Figure 3.27A). The only floral abnormality observed was that occasionally, flowers formed additional stamens - 7 or 8 – instead of 6 in the wild type (Figure 3.27B-C).

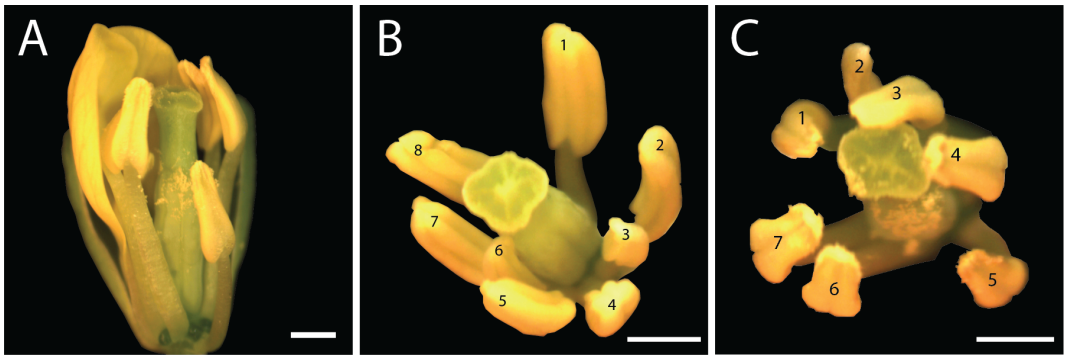

Figure 3.27: Floral phenotype of heterozygous *brag2-1* plants. (A) Mature flower of a *brag2-1*/+ plant. Some sepals and petals were removed to show the reproductive floral organs. (B) and (C) Flowers of a *brag2-1*/+ plant with increased numbers of stamens. Individual stamens are numbered. Scale bars: 1 mm.

Plants homozygous for the *brag2-1* allele showed strong floral defects (Figure 3.28). While the identity of the floral organs was overall unaffected, stamens did not reach the top of the gynoecium. However, in contrast to *brag1-1* flowers where similar defects were observed (Section 3.2.8.1), this appeared to be at least in part due to an extension of the gynophore at the base of the gynoecium (Figure 3.28A). The extension of the gynophore seen in *brag2-1* flowers is similar to the "basalized" gynoecia of *knu-1* flowers and it has been suggested that *KNU* plays a key role in suppressing the overgrowth of basal gynoecial structures including the gynophore and the nectaries thus allowing for the correct development of the ovaries contained within Payne et al., 2004. This suggests that the phenotype seen in Figure 3.28A is a result of inadequate activation of *KNU* and therefore a failure to repress basal overgrowth.

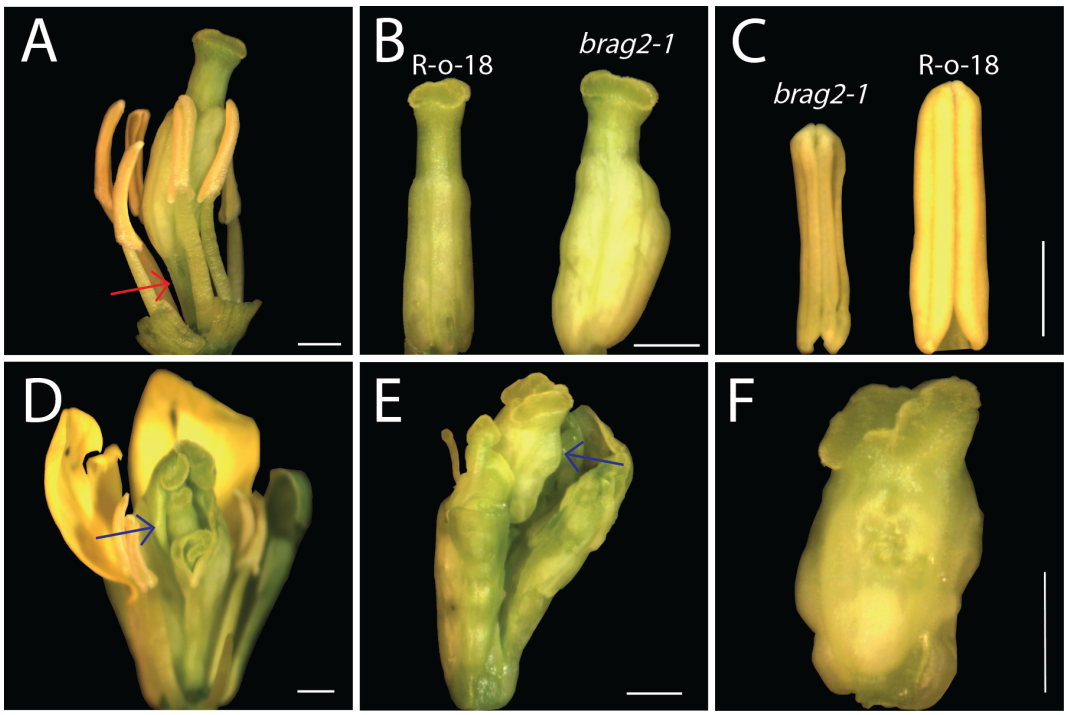

Figure 3.28: Floral phenotype of homozygous *brag2-1* mutant plants. (A) A mature flower with petals and sepals removed to show the reproductive floral organs. An elongated gynophore is marked by a red arrow (B) Gynoecia of R-o-18 and of *brag2-1* mutant plants. Note the wider and bulged ovary in *brag2-1*. (C) Anthers of R-o-18 and of *brag2-1* mutant plants. (D) Flower of a *brag2-1* mutant plant with partially unfused carpels and extra tissue forming in the gynoecium (blue arrow). (E) Close-ups on the gynoecium with the extra tissue (blue arrow) (F) The extra tissue contained within the unfused carpels as shown in panels (D-E). Scale bars: 1 mm.

The anthers of stamens from *brag2-1* mutants were considerably smaller than those of the wild type (Figure 3.29C). Anthers of these plants also produce less pollen than wild-type anthers. Gynoecia of *brag2-1* mutant flowers were much wider than those of the wild type and were often bulged (Figure 3.28B), suggesting the formation of extra tissue in the gynoecium.

Occasionally, more severe gynoecium defects were observed with carpels not properly fusing, revealing the extra tissue forming on the inside of the gynoecium (Figure 3.28D-E). A closer examination of this tissue suggested that it stemmed from the formation of a secondary gynoecium (Figure 3.28F). Thus, the flowers of *brag2-1* mutant plants showed many characteristics of mutants with defects in floral meristem determinacy. The phenotype of *brag2-1* is similar to that of *knu* mutants in *Arabidopsis* which

also produce flowers with extended gynophores and contain ectopic tissue in the carpels (Payne et al., 2004). Because of the severe floral defects, the fertility of the mutant plants was much reduced and only occasionally, flowers produced some seeds. As a result, this mutant line had to be propagated in a segregating population.

3.2.8.4 Floral Phenotypes of *brag2-2* Mutant Plants

Plants heterozygous and homozygous for the *brag2-2* mutation were initially isolated by genotyping as described above and a DMAS genotyping assay was developed. The flowers produced by heterozygous *brag2-2* plants looked identical to those produced by the wild type. Flowers of homozygous *brag2-2* plants are detailed in Figure 3.29.

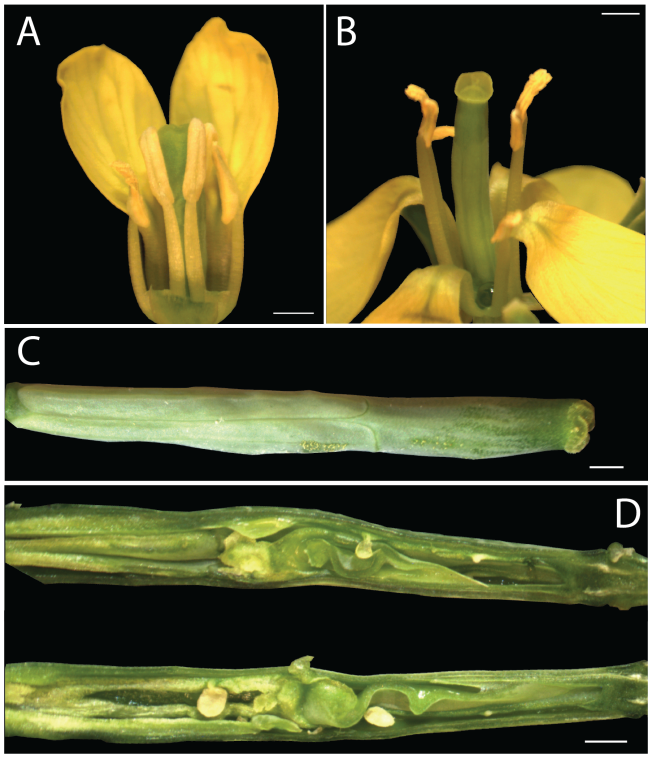

Figure 3.29: Floral phenotype of *brag2-2* mutant plants. (A) Young flower of *brag2-2* plant. Stamens and the gynoecium appear to develop normally. (B) A more mature flower. The style is poorly defined and the ovary is flatter than seen in R-o-18. (C) Silique of a *brag2-2* flower. The siliques of these plants are shorter than those of R-o-18 with the style length reduced and the ovary lacking in volume compared to R-o-18. (D) The same silique as in (C) but dissected. There are fewer ovules than would be seen in a wild-type silique. Scale bar: 1 mm

As shown in panel (A), flowers began to develop similarly to R-o-18. As the flower matured, the gynoecium (Figure 3.29B) continued to grow in size but the style lacked definition from the ovary and the ovary lacked the volume normally seen in R-o-18. Figure 3.29C shows a mature *brag2-2* silique. The style had not extended to the length that is seen in wild-type siliques and the ovary continued to lack volume. Dissection of the siliques showed that the number of ovules was reduced to one or two. The *brag2-2* flowers look phenotypically similar to *Arabidopsis* flowers produced by plants containing mutant alleles for *AINTEGUMENTA* (*ANT*) and *SEUSS* (*SEU*) (Franks et al., 2002). This will be discussed in further detail in Section 3.3.2. As a result of the defects in fruit development, this mutant was infertile and had to be propagated in a segregating population.

3.2.9 Analysis of *brag* Double-Mutants

To test whether *BrAG1* and *BrAG2* are functionally redundant in the control of flower development or have (partially) non-overlapping functions, double-mutant combinations between the different mutant alleles described in the section above were generated. For all mutant combinations discussed below, genotypes were determined by DMAS assays (Section 2.3.3) and through sequencing.

3.2.9.1 Floral Phenotypes of *brag1-2 brag2-2* Double-Mutant Plants

Floral phenotypes of the *brag1-2 brag2-2* double-mutant plants are shown in Figure 3.30. The inflorescence of this mutant developed fewer flowers compared to R-o-18 (Figure 3.30A). The flowers produced showed a mostly wild-type phenotype and were fertile (Figure 3.30B-C).

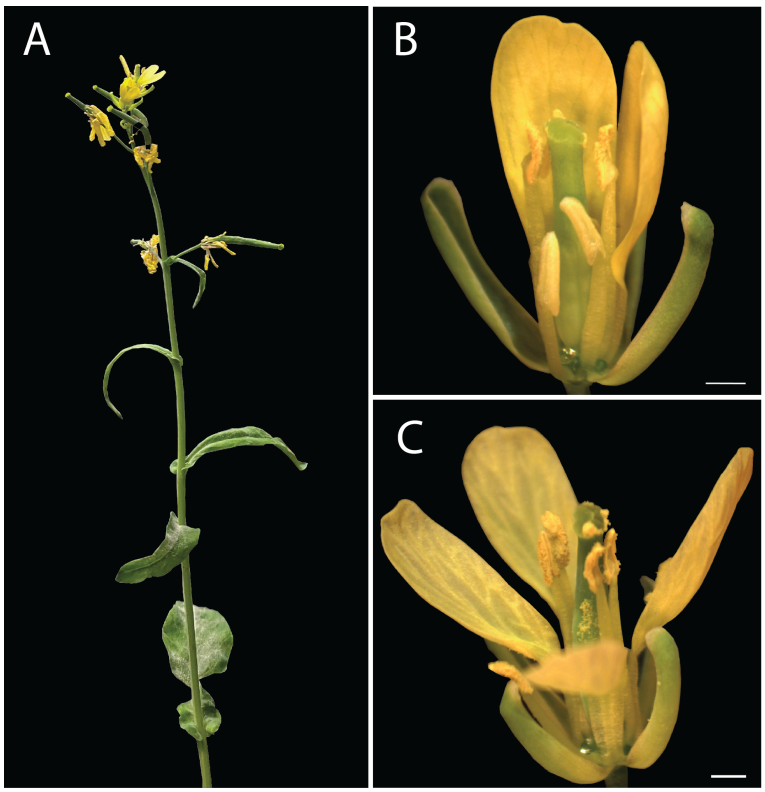

Figure 3.30: Floral phenotype of *brag1-2 brag2-2* double-mutant plants. (A) The inflorescence develops fewer flowers than R-o-18. (B) Flower of a *brag1-2 brag2-2* plant. (C) The flowers that are produced later on show a petal phenotype where they appear more translucent and contain less yellow pigment compared to the wild-type. Scale bar: 1 cm.

The only abnormality that was observed was seen in later-arising flowers that formed petals containing less yellow pigment and that appeared more translucent than wild-type petals. Unlike *brag2-2* flowers (as seen in Figure 3.29), the flowers of the double-mutant produced flowers with gynoecia that had well-defined styles and the ovary appeared to contain ovules. As this plant developed flowers that were close to the wild-type phenotype and the siliques produced were not underdeveloped like those produced by *brag2-2*, this double-mutant was fertile and was able to produce seeds. This double mutant has a milder floral phenotype than the *brag2-2* mutant which was unexpected and it was postulated that perhaps the mutant brag1-2 and brag2-2 proteins are able to interact better with one and other than either mutant protein paired with its wild-type counterpart.

3.2.9.2 Floral Phenotypes of *brag1-2 brag2-1* Double-Mutant Plants

Flowers of *brag1-2 brag2-1* double-mutants formed an extended gynophore (Figure 3.31), similar to what was observed for *brag2-1* (Section 3.2.8.3) single mutants. However, the stamens produced in the flowers of the double-mutant were closer to the stamens of wild-type flowers than those observed in *brag2-1* flowers (Figure 3.28 A, C).

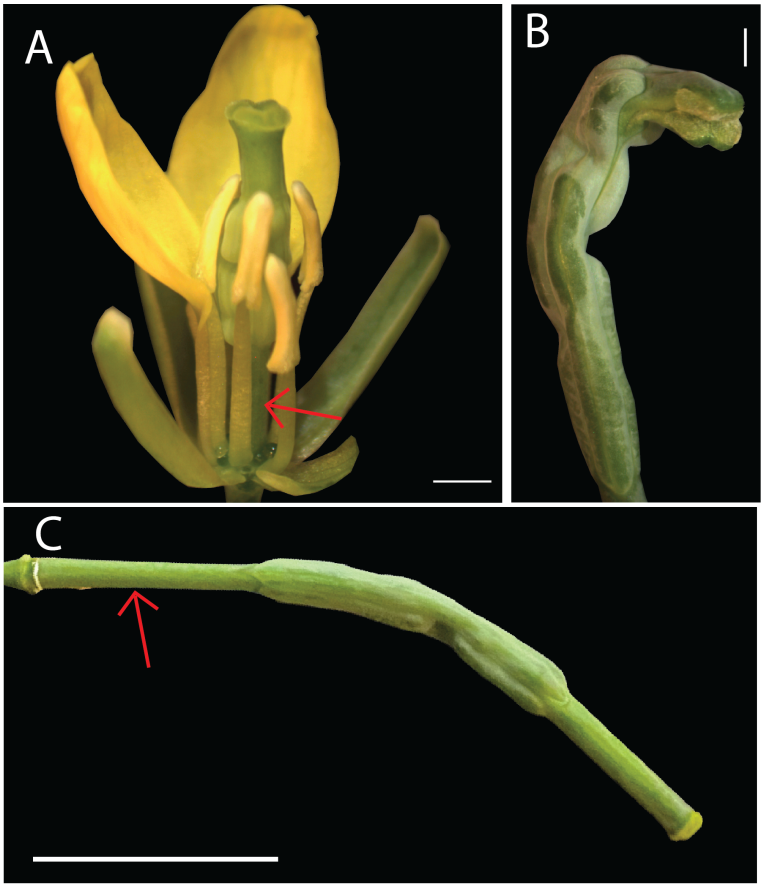

Figure 3.31: Floral phenotypes of *brag1-2 brag2-1* double-mutant plants. (A) Flower of a *brag1-2 brag2-1* plant. An elongated gynophore is marked by an arrow. Scale bar: 1 mm. (B) A silique produced by an early-arising flower. Scale bar: 1 mm. (C) A silique produced by a later-arising flower. An elongated gynophore is marked by an arrow. Scale bar: 1 cm

---

Figure 3.31B shows a typical silique that forms from early-arising flowers with a knobbly appearance and a curved shape. This is a phenotype that is reminiscent of the phenotype seen in the *ag-10* allele or in *knu-1* in *Arabidopsis* (Payne et al., 2004; Ji et al., 2011). Siliques from later-arising flowers looked more similar to an R-o-18 silique.

However the gynophore was hyper-extended compared to R-o-18 siliques and there was still a slight kink in the silique, again similar to the *ag-10* allele.

3.2.9.3 Floral Phenotypes of *brag1-1 brag2-2* Double-Mutant Plants

Similar to the *brag1-2 brag2-1* double-mutant (Section 3.2.9.2), the inflorescences of *brag1-1 brag2-2* double-mutants produced fewer flowers compared to R-o-18 (Figure 3.32A). While sepals, petals and stamens appeared mostly normal, gynoecia were enlarged and unfused at the top (Figure 3.32B-C).

Figure 3.32: Floral phenotype of *brag1-1 brag2-2* double-mutant plants. (A) The inflorescence develops fewer flowers than R-o-18. Scale bar: 1cm. (B) Flower of a *brag1-1 brag2-2* plant. Sepals, petals and stamens appear largely normal but the gynoecium is partially unfused. Scale bar: 1 mm (C) A partially unfused gynoecium with ovules visible. Scale bar: 1 mm.

In later-arising flowers, carpels were even more unfused and internal organs such as nectaries and other ovary tissue could be clearly seen similar to those seen in *crc* mutants (Alvarez and Smyth, 2002). Due to the severe gynoecium defects, this double-mutant was infertile and had to be propagated in a segregating population.

3.2.9.4 Floral Phenotypes of *brag1-1 brag2-1* Double-Mutant Plants

Similarly to the *brag1-2 brag2-1* double-mutant (Section 3.2.9.2) and the *brag1-1 brag2-2* double-mutant (Section 3.2.9.3), the inflorescences of *brag1-1 brag2-1* double-mutants

produced fewer flowers compared to R-o-18 (not pictured). Flowers produced by *brag1-1 brag2-1* showed strong floral defects (Figure 3.33). Sepals appeared mostly normal, but all other floral organs showed defects. There were a normal number of petals but they were smaller than those of R-o-18 (Figure 3.33B). This appeared to be due to the claw of the petal developing in a much more narrow and filamentous manner and the blade area being reduced. The stamens of these flowers produced less pollen than wild-type stamens (Figure 3.33C) and appeared similar to those of *brag2-1* single mutants. In addition to the six stamens, the formation of additional filaments that lacked anthers was observed in these flowers.

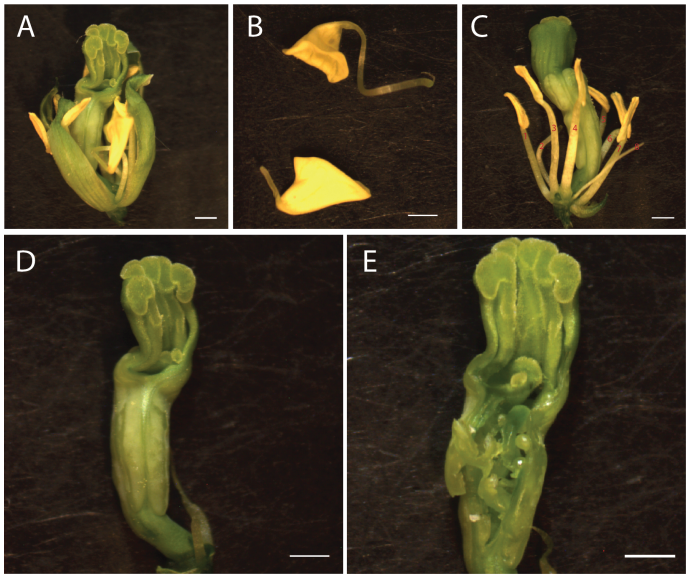

Figure 3.33: Floral phenotypes of *brag1-1 brag2-1* double-mutant plants. (A) Flower of a of *brag1-1 brag2-1* plant. Sepals appeared largely normal but the petals were shaped abnormally. Gynoecia were enlarged and partially unfused, while extra filamentous organs were observed in the stamen whorl. (B) The petals from a *brag1-1 brag2-1* flower. Note the smaller size and misshapen appearance. (C) The mature flower from (A), with sepals and flowers removed to show the reproductive organs. Note the six stamens and two filamentous structures that lack anthers. (D) A partially unfused gynoecium. (E) The same gynoecium as in (D) but dissected. There are fewer ovules than would be seen in a wild-type silique. Scale bars: 1 mm.

The gynoecia were partially unfused (Figure 3.33A, D-E). This was similar to the phenotype seen in the flowers of *brag1-1 brag2-2* mutant plants (Figure 3.32 B-C). However, unlike for *brag1-1 brag2-2*, no ovules were visible at the top of the partially

unfused flowers of *brag1-1 brag2-1* mutants (Figure 3.33D). When the gynoecium was partially dissected (Figure 3.33E), ovules were seen within the gynoecium but fewer than in R-o-18 and fewer than in the late-flowers of *brag1-1 brag2-2* double-mutant flowers. There is an overgrowth of tissue within the gynoecia of this double mutant (Figure 3.33E) which indicates that there is an issue with floral determinancy. Due to the severe gynoecium defects, this double-mutant was infertile and had to be propagated in a segregating population.

3.3 Discussion

3.3.1 Identification and Characterisation of *AG* Homologs in *B. rapa*

The aim of this chapter was to identify homologs of *AtAG* in *B. rapa* and identify mutant alleles that could be used to test our main working hypothesis, namely that the mechanism for the suppression of trichome formation in flowers is conserved between *Arabidopsis* and *Brassica*. During the course of this work, I have successfully identified two homologs of *AtAG* in *B. rapa* - *BrAG1* (*Bra012564*) and *BrAG2* (*Bra013364*) with both genes having been characterised in terms of their expression patterns, and their 3'UTR sequences, and I have shown *via* a series of complementation experiments that both genes can functionally replace *AtAG*.

The RT-qPCR experiments described in Section 3.2.2 revealed that both homologs were expected to be functional as the expression of both was detected across various tissues including the reproductive organs of the flower and broadly matched the expression of *AtAG*. As discussed in Section 1.5, expression of both genes would only be expected if they had both retained functions similar to that of *AtAG*. If the function of either homolog had been lost it would be expected that expression would be lower or at undetectable levels. The expression patterns detected for both *BrAG1* and *BrAG2* reflected the expression patterns for *AG* in *Arabidopsis* which are shown in Figure 3.1. The expression of both were detected at similar levels for the most part, however, the largest difference was seen in the stamen and gynoecium tissue with *BrAG2* expression considerably lower than *BrAG1*. These results also agreed with the phenotypes of mutant alleles as reported in Sections 3.2.8 and 3.2.9.

The 3'UTRs of the *BrAG1* and *BrAG2* were characterised using 3'RACE-PCR. The level of conservation between the sequence was approximately 70% which was higher than that observed for the 3'UTRs of *TRY* and *CPC* (discussed in Section 4.2.4) and

their respective *B. rapa* homologs. There was no clear difference in the sequences that could explain the difference in expression that was observed. Subsequently, the sequences were analysed using the MEME-suite SEA tool to check for common binding motifs in 3'UTRs that are associated with post-translational modification and processing. Again, when compared to *AtAG*, *BrAG1* and *BrAG2* showed no clear difference that could explain the expression patterns detected.

To this end, the upstream and intronic promoter regions of *AtAG*, *BrAG1* and *BrAG2* were analysed using *JASPAR* software to predict potential binding sites for MADS-domain transcription factors that are known to be involved in the regulation and modulation of the expression of *AtAG*. For the upstream regions analysed, there was a similar number of binding sites predicted for *AtAG*, *BrAG1* and *BrAG2*. Particular focus was paid to predicted binding sites of AG given that *AtAG* has been implied to auto-regulate its own expression (Hong et al., 2003). There were four sites predicted for *AtAG*, one for *BrAG1* and two for *BrAG2*. Predicted binding sites for LFY were also examined as it is a known regulator of *AtAG* and a similar number was observed for all three genes. The analyses of the AG and LFY binding sites present in the upstream promoter regions identified some differences in the number of sites and locations but it remains unclear if these differences could be contributing to the differences in expression patterns as described in Section 3.2.2.

The same techniques were also used to analyse the intronic regions of the three genes. There is a high level of sequence conservation across the large intron sequence of the three genes (80%) which indicates the sequence contains functionally important sequences. When the MADS-domain transcription factor binding sites were predicted using *JASPAR* software, a similar number of sites was predicted for *AtAG* and *BrAG2* - 35 and 32 respectively - while there were 17 predicted sites for *BrAG1*. For AG binding sites, five sites were predicted for *AtAG*, four for *BrAG2* and only one for *BrAG1*. For the predicted LFY binding sites, there was a similar number detected across the three

genes. Given that *BrAG1* possesses fewer MADS-domain transcription factor binding sites and specifically AG binding sites perhaps some of the differences in expression that are detected were due to the reduced ability of *BrAG1* or *BrAG2* to bind to the *BrAG1* regulatory intron.

The complementation experiments revealed that for the most part, both *BrAG1* and *BrAG2* were able to functionally complement *AtAG* in an *ag-1* mutant background. Under the CaMV 35S promoter, both *BrAG1* and *BrAG2* show a curly leaf phenotype that is associated with an *AtAG* overexpression phenotype. The floral phenotype of *BrAG1* matched that of *AtAG* overexpression with flowers lacking petals and carpels with stigmatic papillae present in the outer whorl. In contrast, the flowers of *BrAG2* overexpression lines developed petals. As the amino acid sequence of *BrAG1* and *BrAG2* are almost identical ($\approx 95\%$) (Section 3.2.1), the results suggest that the difference in the rescue phenotypes is due to the differences in the 3'UTR of the two. The results of expressing the *B. rapa* homologs under the *AtAG* promoter showed that both *BrAG1* and *BrAG2* are able to rescue the *ag-1* phenotype. A similar rescue of *ag-1* is seen for *AtAG*, *BrAG1* and *BrAG2*. In the case of lines generated using pAtAG:*BrAG2* in an L-*er* background, the later forming flowers (Figure 3.19C) show a phenotype similar to that seen in p35S:*BrAG2* plants (Figure 3.26). It appears that *BrAG2* expression becomes stronger after the development of the first ten flowers and thus the native *AtAG* combined with the pAtAG:*BrAG2* expression results in an overexpression phenotype. Again, given that this is seen across all the independent pAtAG:*BrAG2* lines, the results suggest that the difference in the rescue phenotypes is due to the differences in the 3'UTR sequence.

Taken together, these results indicate that there are two homologs of *AG* present in *B. rapa* - *BrAG1* (*Bra012564*) and *BrAG2* (*Bra013364*).

3.3.2 Characterisation of *BrAG1* and *BrAG2* by Examining Mutant Alleles

In order to further test if *BrAG1* and *BrAG2* were C-function genes, mutant lines were obtained from RevGenUK in order to assess the effects that their inactivation has on flower development in *B. rapa* and allow for investigation of potential subfunctionalization between the two. The isolation of mutant alleles for these genes with different phenotypic strengths would further allow for fine-tuning of C function and ultimately, to test whether the trichome phenotypes as detailed in Section 1.3 in *Arabidopsis* can be re-created in *B. rapa*.

The phenotypic analysis of the mutant lines confirmed the basis of selection for alleles as detailed in Section 3.2.7 - both *brag1-1* and *brag2-1* which were chosen for their suspected potential to eliminate function of *BrAG1* and *BrAG2* and they show the strongest floral defects as shown in Figure 3.25 and Figure 3.28. The other alleles chosen were thought to mimic the weak *ag-10* allele (Ji et al., 2011). Both *brag1-2* and *brag2-2* showed comparatively milder mutant phenotypes.

The flowers produced by *brag1-1* plants develop an unchanged number of stamens (Figure 3.25B-C), however, the anthers are shorter than those produced by R-o-18 plants. The gynoecia of *brag1-1* flowers showed only minor differences to R-o-18 (Figure 3.25D). As the flowers produced by *brag1-1* plants are fertile, it was possible to conclude that *BrAG2* is able to fulfil the role of driving floral organ development in the absence of *BrAG1*. As discussed in Section 3.1.2, stamen identity in *Arabidopsis* is specified by B-, C- and E-class genes. One possible cause of the stamen phenotype of *brag1-1*, is that BrAG1 is the preferential partner of the B- and E-class *B. rapa* proteins driving stamen development with BrAG2 being at least partially functionally redundant. Flowers of *brag1-2* plants were only mildly affected with the ovary of the siliques being shorter than R-o-18 siliques, proving that as predicted the *brag1-2* floral phenotype was

milder than that of *brag1-1*.

The most striking floral mutant phenotype of the single mutants was that of *brag2-1* mutants (Figure 3.28). Like *brag1-1*, the identity of the floral organs was unaffected however defects were seen in both stamens and gynoecia. The stamens produced less pollen and the anthers were smaller compared to R-o-18 stamens (Figure 3.28A, C). Gynoecia of *brag2-1* mutant flowers had elongated gynophores, were much wider than those of the wild type and were often bulged (Figure 3.28B) and occasionally carpels were not properly fused with extra tissue forming inside the gynoecium (Figure 3.28D-E). As discussed in Section 3.2.8.3, the phenotype of *brag2-1* is similar to that of *knu* mutants in *Arabidopsis* which also produce flowers with extended gynophores and contain ectopic tissue in the carpels (Payne et al., 2004). As detailed in Section 3.1.1 in *Arabidopsis*, one of the pathways by which *AG* represses the production of stem cells by indirectly repressing *WUS* by activating *KNU* at stage 6 of flower development (Sun et al., 2009). These results suggest that *BrAG2* is the homolog that fulfils the role that *AtAG* has in floral meristem determinacy in *Arabidopsis*. However, the *brag2-1* mutant allele is a mutation in a conserved splice donor site and this likely results in a 15 nt deletion in the mRNA of *brag2-1* compared to wild-type *BrAG2* mRNA as shown in Section 3.2.7. The resulting *BrAG2* protein transcribed in *brag2-1* plants likely lacks five amino acids that occur at the C-terminal end of the MADS-domain and I-domain which are essential for DNA-binding and dimerization specificity of *AG* in *Arabidopsis* (Lai et al., 2021). It could be that BrAG2 in *brag2-1* plants is acting in a dominant negative manner perhaps having the ability to form higher-order protein complexes but lacking the ability to bind to DNA. Or the mutant protein could lack the dimerization specificity of the wild-type protein and is acting to interfere in the formation of the normal higher-order protein complexes needed to drive the correct development of reproductive floral organs or to drive the termination of the floral meristem in a timely manner.

The flowers of *brag2-2* plants initially developed similarly to flowers produced by

R-o-18 (Figure 3.29 A-B). As the flower matured, the style lacked definition from the ovary, the ovary lacked the volume normally seen in R-o-18 and the resulting silique developed with far fewer ovaries than R-o-18. The *brag2-2* flowers look phenotypically similar to *Arabidopsis* flowers produced by plants containing mutant alleles for *AINTEGUMENTA* (*ANT*) and *SEUSS* (*SEU*) shown in Figure 3.34 (Franks et al., 2002).

Figure 3.34: Floral phenotype of *ant* and *seu* mutant plants. (A) A wild-type *Arabidopsis* from accession Col-0. (B) A flower from a *ant-3* mutant plant. (C) A flower from a *ant-1* plant. (D) A flower from a *seu-3* mutant plant. Note the lack of definition between the style and ovary as well as the elongated ovary of the mutant flowers. Scale bars: 1 mm. Figure adapted from Azhakanandam et al., 2008

ANT encodes a sequence-specific DNA-binding protein required for ovules development and *SEU* encodes a transcriptional co-regulator (Klucher et al., 1996; Franks et al., 2002). The flowers of *ant* and *seu* look similar to *brag2-2* flowers (Figure 3.29B) with poorly defined styles and the ovary lacking volume compared to wild-type flowers. However, *ant* and *seu* single mutants still develops ovules with *ant-3* and *seu-3* producing a similar number to Col-0 and *ant-1* producing half that of Col-0.

It is only in double-mutant combinations of *ANT*, *SEU*, *FILAMENTOUS FLOWER* (*FIL*) and *LEUNIG* (*LUG*) that flowers develop a similar floral mutant phenotype to *brag2-2* flowers. *FIL* is a *YABBY* gene involved in the establishment of abaxial cell type specification and *LUG* is another transcriptional co-regulator and is known to repress *AG* (Sawa et al., 1999; Conner and Liu, 2000). Plants containing *lug ant*, *fil ant* or *seu lug* double-mutants produce flowers that lack ovules but also develop with unfused carpels or more severe defects which *brag2-2* lacks (Liu et al., 2000; Franks et al., 2002;

Nole-Wilson and Krizek, 2006). Currently, there is no evidence that *AG* directly or indirectly controls the expression of *ANT*, *SEU*, *FIL* or *LUG*. It seems that *brag2-2* flowers are an intermediate phenotype where they lack the ovules as seen in the double-mutants described above but phenotypically appear more like *ant* or *seu*. It was concluded that *BrAG1* is likely driving the development of the carpels resulting in the fused carpel phenotype observed in *brag2-2* but lacks the ability to drive ovule development.

The *brag1-2 brag2-2* plants produce flowers that show a mostly wild-type phenotype with later flowers developing with petals that contain less pigment (Figure 3.30B-C). The most dramatic phenotype seen in *brag1-2 brag2-2* double-mutants are the reduced number of flowers produced from the inflorescence which resembles mutants of *TERMINAL FLOWER 1* (*TFL1*) in *Arabidopsis* where *tfl1-1* mutants produce one to five flowers and occasionally one or two lateral branches that terminate in a single flower (Shannon and Meeks-Wagner, 1991). It appears that *brag1-2 brag2-2* mutant alleles in this line are able to function in their roles in driving the development of reproductive floral organs. It is unclear how but *brag1-2 brag2-2* is driving the termination of the shoot meristem at an earlier stage than wild-type plants or either of the single mutants.

The flowers of *brag 1-2 brag2-1* double-mutants show an intermediate phenotype between the weak *brag1-2* and strong *brag2-1* floral defects. Flowers develop an elongated gynophore (Figure 3.31A, C) which is a phenotype only seen in *brag2-1* mutants (Figure 3.28A). As mentioned above this is similar to a *knu* phenotype, however, these plants show a less dramatic ectopic growth phenotype as seen in *brag2-1* single mutant (Figure 3.28B, D, E)(Payne et al., 2004). The siliques produced by early arising flowers have a knobbly appearance and a curved shape (Figure 3.31B) and are similar to those produced by *ag-10* and *knu-1* plants (Ji et al., 2011). Figure 3.35 shows a silique of *brag1-2 brag2-1* plant compared to a silique from *Arabidopsis ag-10* and *knu-1*.

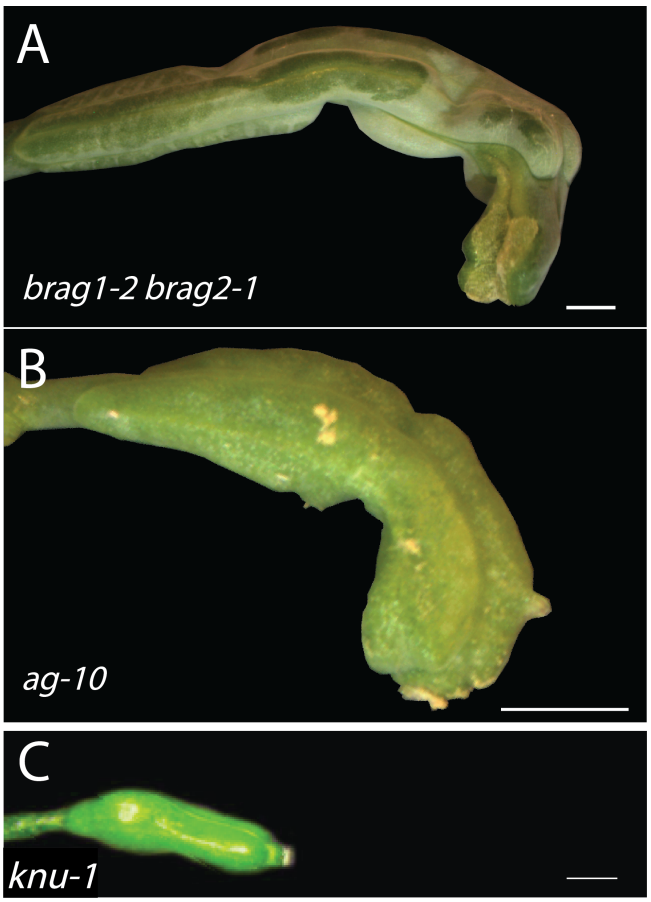

Figure 3.35: Silique of *brag1-2 brag2-1* double-mutant and *ag-10* and *knu* single mutant plants. (A) The silique from an early-arising *brag1-2 brag2-1* flower with increased curvature compared to the wild-type. (B) Silique from an *Arabidopsis ag-10* mutant plant that is curved and slightly bulged. (C) Silique from an *Arabidopsis knu* plant. Scale bars: 1 mm. Panel C adapted from Kwaśniewska et al., 2021.

This silique phenotype in the *brag1-2 brag2-1* could be caused by the *ag-10* like *brag1-2* mutant allele (Section 3.2.8) or that the expression of *KNU* is not being activated correctly but the lack of ectopic growth in the siliques makes this unlikely. It could also be the case that if the *brag2-1* mutant allele is acting in a dominant negative fashion as discussed previously, it is less able to impede the function of the *BrAG1* protein in the *brag1-2 brag2-1* mutant.

The floral phenotype of *brag1-1 brag2-2* double-mutants also show the inflorescence phenotype seen in *brag1-2 brag2-2* with fewer flowers produced by this mutant (Figure 3.32A). However, unlike *brag1-1 brag2-2* flowers, *brag1-2 brag2-2* flowers develop

with large unfused gynoecia with ovules visible (Figure 3.32B-C). In flowers produced later on, the carpels were even more unfused and internal organs could be seen clearly. This phenotype is similar to the phenotypes observed in *Arabidopsis* mutants for the genes *SPATULA* (*SPT*), a transcription factor which is part of the bHLH protein family and *CRABS CLAW* (*CRC*), a YABBY family gene (Bowman and Smyth, 1999). Both genes are known to be essential for the correct development of the gynoecium in *Arabidopsis* with single mutants for each gene showing abnormal development and small areas where the carpels are unfused (Figure 3.36B-C). The double-mutant shows a dramatic phenotype where the two carpels are almost completely unfused (Figure 3.36D) (Alvarez and Smyth, 2002).

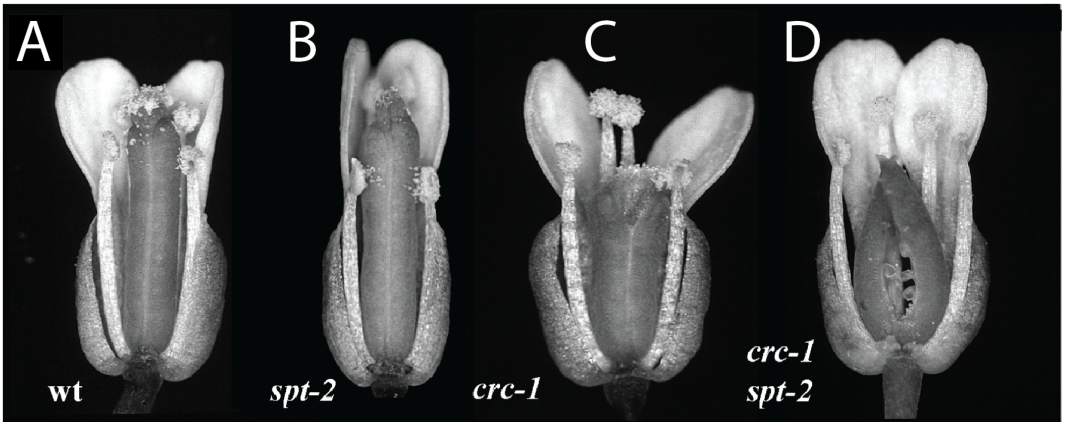

Figure 3.36: Floral phenotype of *spt crc Arabidopsis* flowers. (A) A wild-type flower showing normal gynoecium development (B) A flower from an *spt-2* plant with a style that is shorter, narrower than wild-type and containing an apical split that traverses the style and upper part of the ovary. (C) A flower from an *crc-1* mutant which develops a shorter and wider gynoecium compared to wild-type plants. The two carpels are unfused at the style and the upper ovary. (D) A *cpc-1 spt-2* double-mutant flower with a gynoecium that is unfused except at the base and lacks a septum. Figure adapted from (Alvarez and Smyth, 2002).

It is clear that the *brag1-1 brag2-2* double-mutant is disrupting normal gynoecium development and interestingly seems to not affect stamen development. This line shows the same inflorescence phenotype as the other double *brag1 brag2* mutants with much fewer flowers produced by this mutant (not pictured). Unsurprisingly, the strongest floral defects are seen in the flowers of *brag1-1 brag2-1* double-mutants. The gynoecium

is enlarged and unfused with no ovules visible (Figure 3.33A, C-E). The petals produced by *brag1-1 brag2-1* flowers are smaller than wild-type flowers (Figure 3.33B). The stamens of these flowers do not produce as much pollen as wild-type flowers and contained two additional filament-like structures that lacked anthers (Figure 3.33C). Similarly to *brag1-1 brag2-2* double-mutants, the carpels are unfused at the top. However, there are no ovules visible in the unfused portion and when dissected there are fewer ovules than both *brag1-1 brag2-2* and wild-type flowers. This unfused gynoecium phenotype and reduced ovule number is similar to that seen in *crc spt* in *Arabidopsis* (Figure 3.36D) (Alvarez and Smyth, 2002).

In general, the results of the genetic analysis showed that *BrAG1* and *BrAG2* are not fully functionally redundant as the inactivation of one of these genes was sufficient to induce phenotypic abnormalities. At the same time, the generation of different combinations of *brag1 brag2* double-mutants resulted in much stronger phenotypes than observed in either of the corresponding single mutants (with the exception of *brag1-2 brag2-2*), suggesting that functions of *BrAG1* and *BrAG2* partially overlap. Notably, both *brag1* and *brag2* single-mutants and the different combinations of *brag1 brag2* double-mutants were most severely affected in gynoecium development. As discussed above, the carpel fusion defects and the growth of additional tissue inside of gynoecia are likely linked to the known functions of AG in the control of key regulatory genes for gynoecium development such as *CRC* and *KNU*. However, even in the case of the *brag1-1 brag2-1* double mutant, where both *BrAG* genes should be inactivated, no homeotic transformations of the reproductive floral organ were seen as observed in strong *ag* mutants in *Arabidopsis*. This was an unexpected result especially given that a recent study has shown that a CRISPR-mediated knockout of all four *AG* homologs in *B. napus* (which would include *BrAG1* and *BrAG2*) leads to an *ag*-like mutant phenotype (Shin et al., 2022).

One possible explanation for the lack of homeotic transformations could be that the

brag1-1 and *brag2-1* alleles are not null alleles and that some C function remains active in the double mutant. Given that *brag1-1* contains a very early stop codon in its CDS, any residual activity would have to come from the *brag2-1* allele. As outlined in Section 3.2.7, this allele carries a splice donor site mutation, resulting in a (fairly conservative) change of one amino acid residue at the end of the MADS-domain and a 5 amino acid residue deletion at the beginning of the I-domain. The results of the single-mutant analysis for this allele showed that it is much stronger than the *brag2-2* allele, indicating that the mutated protein has indeed lost function. However, whether it is a true null allele cannot be deduced from this observation alone. To test the functionality of the brag2-1 protein, I conducted experiments, in which I attempted to rescue *ag-1* mutants (similar to what has been described in Section 3.2.6) by over-expressing *BrAG2-1*. The results from this analysis have so far been inconclusive with not enough transgenic lines generated in order to form a conclusion but they are essential for determining whether the brag2-1 protein has residual activity and could be the reason for the lack of homeotic transformations in the *brag1-1 brag2-1* mutant flowers.

An alternative explanation for the absence of homeotic transformations in the *brag1-1 brag2-1* double mutant is that in *B. rapa*, other genes can functionally replace the *BrAG* genes when they become inactive. Strong candidates for such genes are the *SHATTERPROOF* (*SHP*) genes, which are closely related to *AG*. Notably, in *Arabidopsis*, it has been shown that they can functionally replace *AG* in certain mutant backgrounds or when expressed early on in flower development (Colombo et al., 2010; Ó'Maoiléidigh et al., 2018). At the same time, *AG* has been shown to activate these genes specifically in developing gynoecia (Savidge et al., 1995). Thus, in *Arabidopsis*, the *SHP* genes act downstream of *AG* at stages when organ specification has already occurred (Ó'Maoiléidigh et al., 2013). It appears the reason that *AG* and the *SHP* genes are not functionally redundant in the specification of floral organs in *Arabidopsis* is that their expression patterns have diverged with the *SHP* genes taken on a more specialized function during gynoecium development. Interestingly, recent work in our laboratory has

shown that the expression of the homologs of *SHP* genes in *B. rapa* does not seem to be confined to gynoecia (Caoilfhionn Breathnach, unpublished observations). Instead, they also seem to be expressed in stamen primordial. Thus, their expression patterns appear much more similar to those of the *BrAG* genes than what is observed in *Arabidopsis*. From these results, it seems possible that in *B. rapa*, C function is mediated by a combination of *BrAG* and *BrSHP* genes. If true, the floral defects of *brag1-1 brag2-1* double mutants would be due to a partial reduction of C function, which mainly leads to abnormalities in the 4th whorl in agreement with what has been described for weak *ag* mutants in *Arabidopsis* such as *ag-10*. Experiments are currently underway in our laboratory to test the functional relationship between the *BrAG* and *BrSHP* genes through genetic analysis.

3.4 Future Work

For the purpose of this project, the aim of this chapter was to identify the homologs of *AG* in *B. rapa* and identify suitable mutant lines that could be used as the sensitised background for combination with trichome repressor mutant allele and examine for the development of trichomes on the gynoecia. The work carried out in this chapter indicated that *BrAG1* is the homolog that should be focused on and that the mutant lines obtained and characterised - *brag1-1* and *brag1-2* - are close to the *ag-10* alleles that I wish to replicate. Further work with these mutant *B. rapa* lines is detailed in Chapter 5.

In further efforts to characterise the *AG* homologs in *B. rapa*, there is much that can still be carried out. Further work is necessary to determine the exact cause of the differential expression patterns of *BrAG1* and *BrAG2* seen in Section 3.2.2. This could be achieved by interrupting individual candidate binding sites in the upstream and intronic promoter regions identified in Sections 3.2.4 and 3.2.5 *via* site-directed mutagenesis. This would have to be done in *Arabidopsis ag-1* mutants to compare the rescue with binding sites interrupted to the rescue without binding sites interrupted, given the difficulties in performing *Agrobacterium*-mediated transformation in *B. rapa* successfully. The *BrAG1* and *BrAG2* upstream and intronic promoter sequences could be cloned driving the expression of a reporter such as $\beta-$glucuronidase (GUS) in order to compare the expression profiles of the two and have a control to compare to when predicted sites for AG and LFY are mutated.

In order to determine if the floral mutant phenotypes of the double-mutant lines are due to target genes not being regulated properly in the absence of functional *BrAG1* and *BrAG2*, RT-qPCR could be performed. For example, the expression of *B. rapa* homologs of *CRC* in floral tissue from *brag1-1 brag2-2* and *brag1-1 brag2-1* could be examined to determine if the unfused carpel phenotype was due to altered expression in the absence of functional *BrAG1* and *BrAG2*. The expression levels could be compared

to wild-type and single-mutant plants. Experiments such as these would allow for elucidation of how conserved the *AG* downstream targets and characterise how/if the roles of *AG* in modulating the expression of downstream target genes have been split between *BrAG1* and *BrAG2*.

In order to determine the impact of the *brag2-1* mutation (Section 3.2.7) and how well the resulting truncated protein is able to function will be investigated by driving the expression of *brag2-1* CDS under the control of the pAG promoter. Similar to the complementation experiments described in Section 3.2.6, the pAG:*brag2-1* construct will be transformed in *ag-1* in order to assess its potential for rescue. It will also be transformed into L-*er* in the hope that any possible interactions with AtAG will give insight into whether the BrAG2 protein produced by *brag2-1* is acting in a dominant negative manner. The nature of the *brag1-2* and *brag2-2* mutant alleles could be investigated and characterised further by cloning their coding DNA sequence under the control of the pAG promoter and testing their potential to rescue the *ag-1* mutant phenotype. As mentioned in Section 3.2.8, *brag1-2* and *brag2-2* are suspected to be similar to the weak *Arabidopsis ag-10* allele. An *ag-10* rescue allele would be expected if the brag1-2 and brag2-2 proteins are like the AG mutant protein produced by *ag-10*.

Chapter 4 - Identification and Characterisation of Homologs of *CAPRICE* and *TRIPTYCHON* in *Brassica rapa*

4.1 Introduction and Aims

Single repeat R3 MYBs play a key role in repressing the development of trichomes. In stems and leaves, they are activated following the formation of the MBW activator complex in a neighbouring epidermal cell. These R3 MYBs travel to neighbouring cells, blocking the formation of the MBW activator complex, and this ensures even spacing of trichomes on the surface of the plant (see Section 1.2.1 for further detail) (Wang et al., 2021). In contrast, in the gynoecium of flowers, the C-class homeotic protein AGAMOUS acts to positively regulate the expression of R3 MYBs in order to prevent trichome formation from occurring. As described in Section 1.3, it has been shown that in *Arabidopsis*, the combination of weak *ag* mutant alleles with strong mutant alleles of the trichome repressor genes *TRIPTYCHON* (*TRY*) and/or *CAPRICE* (*CPC*) leads to the formation of trichomes on carpel valves (Ó'Maoiléidigh et al., 2018).

As previously discussed (Section 1.6), the main aim of the SFI/EPA-funded project that supported my PhD work was to engineer *Brassica* crops that develop supernumerary and floral trichomes to investigate their potential for pest control. To this end, mutant alleles of the genes known to control trichome formation in flowers had to be isolated

in *B. rapa*. Chapter 3 focused on the identification of such mutant alleles for *B. rapa* homologs of *AG*. In this chapter, I describe work that aimed at the identification and functional characterisation of homologs of *Arabidopsis CPC* and *TRY* in *B. rapa*. For it, I employed approaches very similar to the ones described in Chapter 3. However, unlike in the case of *AG*, there was less concern about a loss of *TRY* and *CPC* activity resulting in a reduction of fertility. Thus, strong mutant alleles were preferred and were thought to give the best chance at recreating in *B. rapa* the trichome phenotypes observed initially in *Arabidopsis*.

4.2 Results

4.2.1 Identification of *TRY* and *CPC* Homologs in *B. rapa*

In order to identify the homologs of *TRY* (AT5G53200; hereafter referred to as *AtTRY*) and *CPC* (AT2G46410; hereafter referred to as *AtCPC*) from *Arabidopsis*, their genomic and coding DNA sequence (CDS) were used to perform BLAST searches on the *B. rapa* genome sequence that was available at the time this work began - accession Chiifu-401-42 (Wang et al., 2011, Cai et al., 2017). The results and subsequent work carried out to confirm the identity of the genes identified are described below.

4.2.1.1 Identification of *TRY* Homologs in *B. rapa*

The genomic and coding DNA sequence (CDS) of *AtTRY* were obtained from The Arabidopsis Information Resource (TAIR) and used to perform BLAST searches on the *B. rapa* genome sequence. Three highly significant hits were obtained - *Bra022637* located on chromosome 2 (E-value of 9.20E-36), *Bra026297* located on chromosome 1 (E-value of 6.6E-09) and *Bra029089* located on chromosome 3 (E-value of 9.20-36). The BLAST results are summarized in Table 7.3 in the Appendix. The gene structures of *AtTRY* and the three *B. rapa* genes are compared in Figure 4.1.

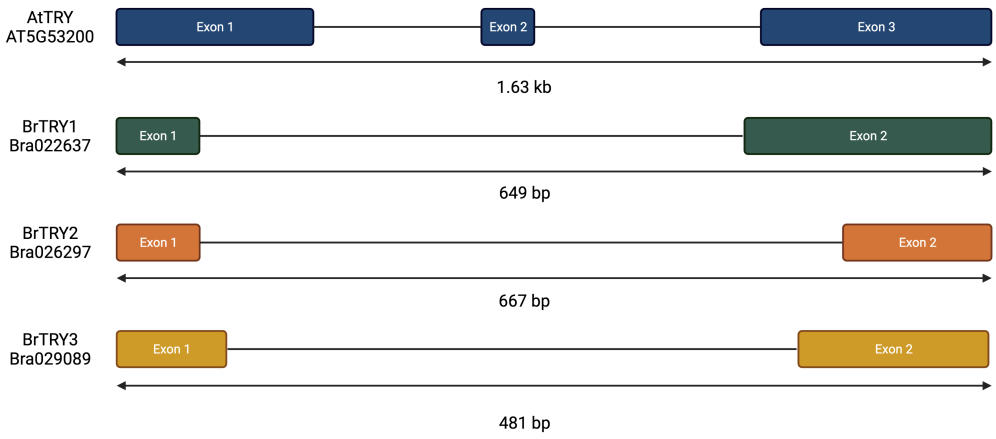

Figure 4.1: Gene structure of *AtTRY* and of three putative homologs from *B. rapa*. Diagram showing the structure of *AtTRY* (dark blue), *Bra022637* (green), *Bra026297* (orange) and *Bra029089* (yellow). Exons of each gene are labelled and black lines represent introns. The lengths of the genes are marked by arrows below. Figure generated using Biorender.com.

In contrast to *AtTRY*, which contains three exons, the three putative homologs in *B. rapa* all have only two exons. Thus, if these genes are true homologs, their gene structures have changed in either *Arabidopsis* or *B. rapa* after the divergence of their last common ancestor. To test this further, I compared the genomic sequence, the coding DNA sequence (CDS) and the amino acid sequence of *AtTRY* to those of the three candidate homologs from *B. rapa* using the EMBL-EBI *Clustal Omega* tool (Madeira et al., 2022).

A summary of the percentage identity between these sequences is shown in Table 4.1. As mentioned in Section 3.2.1, the mean level of sequence conservation between homologs in *Arabidopsis* and *B. rapa* is $\approx$ 70%. The genomic sequences of *AtTRY* and the putative homologs under examination show a wide range of percentage identity - between 55% and 70%. When examining the CDS, I found that the three putative homologs have higher than average identity percentages than would be expected between homologs. *Bra029089* has the highest percentage identity at $\approx$ 92% identity whilst *Bra022637* is the next highest at $\approx$ 84% followed by *Bra026297* at $\approx$ 78% identity. There was a similar trend seen for the protein sequences with the highest being

Bra029089 at ≈ 92%, *Bra022637* the next highest at ≈ 86%, and then *Bra026297* at ≈ 69%. These values for percentage identity support the notion that three *B. rapa* genes are homologs of *AtTRY*.

Table 4.1: Percentage identity matrix for *AtTRY* and potential *B. rapa* homologs. Table comparing the percentage identity of the genomic, coding DNA sequence (CDS) and protein sequences of *AtTRY*, *Bra022637*, *Bra026297*, *Bra029089*.

	Genomic				CDS			
	1	2	3	4	1	2	3	4
1. AtTRY	100	69.67	55.12	60.34	100	83.56	77.58	92.31
2. Bra022637	69.67	100	59.96	56.22	83.56	100	83.94	86.67
3. Bra026297	55.12	59.96	100	47.65	77.58	83.56	100	76.67
4. Bra029089	60.34	56.22	47.65	100	92.31	86.67	76.67	100

	Protein			
	1	2	3	4
1. AtTRY	100	86.49	68.52	92.16
2. Bra022637	86.49	100	68.52	89.47
3. Bra026297	68.52	68.25	100	84.21
4. Bra029089	92.16	89.47	84.21	100

The alignment of the CDS and the protein sequences of *AtAG*, *Bra022637*, *Bra026297* and *Bra029089* are shown in Figure 4.2 and Figure 4.3, respectively. The genomic sequence alignment is shown in Section 7.4 in the Appendix. As can be seen in the figures, there is a high level of sequence conservation between these genes. As mentioned in Section 1.2.1, *AtTRY* encodes a member of the R3-MYB protein family which is characterised by an MYB-related DNA binding domain that is critical to their function in repressing trichome formation (Schellmann et al., 2002).

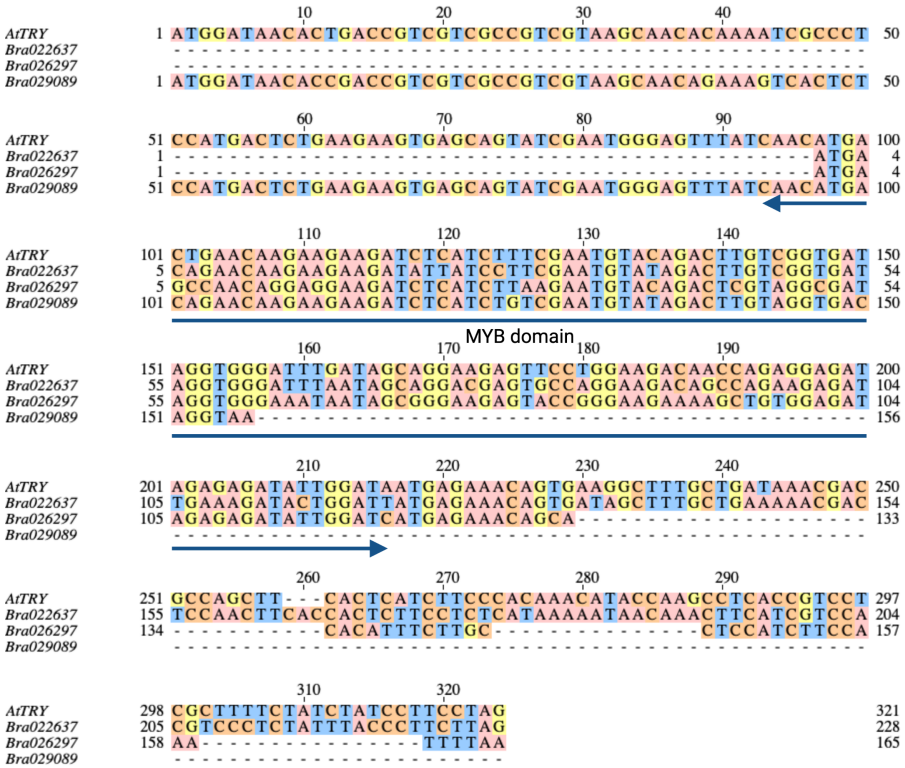

Figure 4.2: Alignment of coding sequences of *AtTRY* and the putative *B. rapa* homologs. The coding DNA sequence of *AtTRY* and the three putative homologs were aligned using the *Clustal Omega Multiple Sequence Alignment* tool from EMBL-EBI (Madeira et al., 2022). The alignment results were visualised using *Jalview* software. The R3 MYB domain which facilitates the interaction with bHLH transcription factors is annotated. Figure generated using Biorender.com

---

As can be seen in Figure 4.3, *Bra022637* and *Bra026297* have a high level of conservation to *AtTRY* in this domain with *Bra022637* possessing 100% sequence identity in this domain and *Bra026297* possessing $\approx$ 85% sequence identity with most amino acids that differ belonging to the same residue group. In contrast, *Bra029089* only possesses the first 20 amino acids of this domain and it was concluded that this may render the protein non-functional.

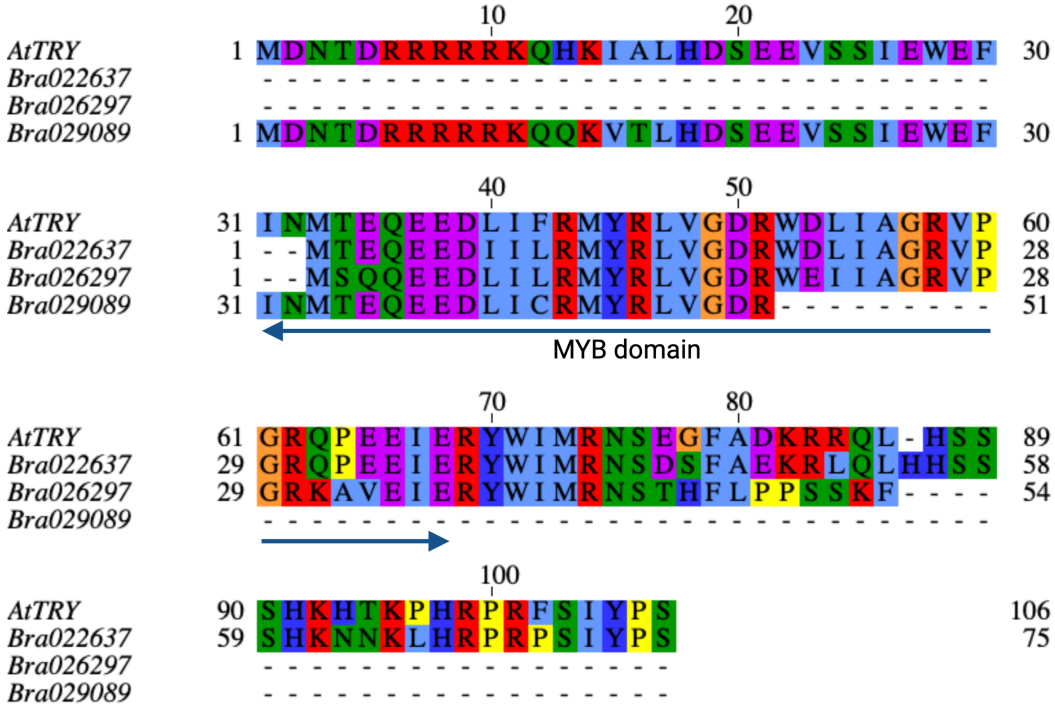

Figure 4.3: Alignment of protein sequences of AtTRY and putative *B. rapa* homologs. Alignment of the protein sequence of AtTRY with the putative homologs identified in *B. rapa* using the *Clustal Omega Multiple Sequence Alignment* tool from EMBL-EBI (Madeira et al., 2022). The following colour scheme was used to denote amino acids that belong to the same residue groups: blue: hydrophobic; red: positively charged; magenta: negatively charged; green: polar; pink: cysteines; orange: glycines; yellow: prolines; cyan: aromatic. The alignment results were visualised using *Jalview* software. The R3 MYB domain which facilitates the interaction with bHLH transcription factors is annotated. Figure generated with BioRender.com

---

To further test whether *Bra022637*, *Bra026297* and *Bra029089* are indeed *AtTRY* homologs, I compared the genes flanking the different loci. As shown in Table 4.2, *Bra022637* and *Bra029089* share many neighbouring genes with *AtTRY*. Of the 20 surrounding genes examined for *Bra022637*, 19 of them are located on the same chromosome and 11 are also direct neighbours of *AtTRY*. In the case of *Bra029089*, 18 of the genes are located on chromosome 5 and 12 share identity with the direct neighbours of *AtTRY*. In the case of *Bra026297*, only one of the 20 genes examined was located on chromosome 5 while none of them shared identity with the direct neighbours of *AtTRY*. However given the relatively high level of sequence identity, it was decided to not exclude this gene from the list of possible homologs.

Table 4.2: Synteny of *AtTRY* and putative *B.rapa* homologs. The table lists the ten neighbouring genes upstream and downstream of *AtTRY* (AT5G53200), *Bra022637*, *Bra026297*, *Bra029089*. The neighbouring genes were identified using the EnsemblPlants Region in Detail view to identify the surrounding genes. In the case of the *B. rapa* genes, the BRAD gene IDs were "BLASTed" and the corresponding TAIR IDs were listed for easy comparison. Gene IDs that appear in the *AtTRY* column and one of the homolog columns are highlighted in the same colour for ease of comparison. N/A indicates there was no corresponding TAIR ID found for the BRAD gene ID.

Position	AT5G53200	Bra022637	Bra026297	Bra029089
-10	AT5G53100	AT5G53340	AT4G27780	AT5G53300
-9	AT5G53110	AT5G53310	AT4G27790	N/A
-8	AT5G53120	AT5G53300	AT4G27810	AT5G53290
-7	AT5G53130	N/A	AT4G27860	AT5G58230
-6	AT5G53140	AT5G53290	AT4G27860	AT5G02050
-5	AT5G53150	AT5G53260	AT4G27870	AT5G53280
-4	AT5G53160	AT5G53250	AT4G27880	AT5G53260
-3	AT5G53170	AT5G54550	AT4G27890	AT5G53250
-2	AT5G53180	AT5G05450	AT4G27920	AT5G53220
-1	AT5G53190	AT5G53210	AT4G27940	AT5G53210
+1	AT5G53210	AT5G53190	AT1G17920	AT5G53190
+2	AT5G53220	AT5G53170	AT4G27950	AT5G53180
+3	AT5G53230	AT5G53160	AT4G27970	AT5G53170
+4	AT5G53240	AT3G18720	AT4G28000	AT5G53160
+5	AT5G53250	AT5G53130	N/A	AT3G21910
+6	AT5G53260	AT5G53120	N/A	AT5G53140
+7	AT5G53270	AT5G53110	AT4G28025	AT5G53100
+8	AT5G53280	AT5G53090	AT4G28040	AT5G03020
+9	AT5G53290	AT5G53080	AT4G28050	AT5G53640
+10	AT5G53300	AT5G53070	AT5G57815	AT5G53050

On the basis of the high level of sequence identity, the presence of a key functional domain and the level of synteny conservation between the four genes, it was concluded that in *Bra022637*, *Bra026297* and *Bra029089* were homologs of *TRY* in *B. rapa*. As such, they were given the names *BrTRY1* for *Bra022637*, *BrTRY2* for *Bra026297*, and *BrTRY3* for *Bra029089*.

4.2.1.2 Identification of *CPC* Homologs in *B. rapa*

The genomic sequence and CDS of *AtCPC* were obtained from TAIR and used to perform BLAST searches on the *B. rapa* genome. Two highly significant hits were obtained - *Bra004539* located on chromosome 5 (with an E-value of 1.1e-31) and *Bra039283* located on chromosome 4 (with an E-value of 4.5e-310). The BLAST results are summarized in Table 7.4 in the Appendix. The structure of the three genes is compared in Figure 4.4 which shows that *Bra004539* and *Bra039283* have similar gene structures to *AtCPC* with all three genes containing three exons.

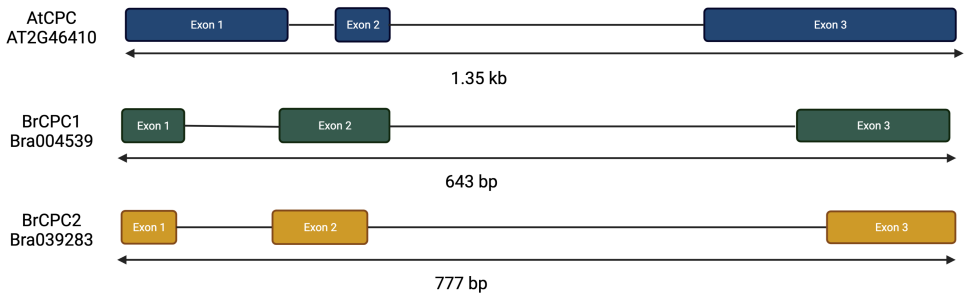

Figure 4.4: Gene structure of *AtCPC* and of two putative homologs from *B. rapa*. Diagram showing the structure of *AtCPC* (dark blue), *Bra004539* (green) and *Bra039283* (yellow). Exons of each gene are labelled and introns are represented by black lines. The length of each gene is marked below by an arrow. Figure generated using Biorender.com.

The genomic sequence, the CDS and the amino acid sequence of *AtCPC*, *Bra004539* and *Bra039293* were aligned using the EMBL-EBI *Clustal Omega* tool (Madeira et al., 2022). A summary of the percentage identity between these sequences is shown in Table 4.3.

The genomic sequences exhibit a high level of sequence conservation with both

Table 4.3: Percentage identity matrix of *AtCPC* and potential *B. rapa* homologs. Table comparing the percentage identity of the genomic sequences, coding DNA sequence (CDS) and protein sequences of *AtCPC*, *Bra004539* and *Bra039283*.

	Genomic			CDS		
	1	**2**	**3**	**1**	**2**	**3**
1. AtCPC	100	76.06	74.54	100	89.53	87.98
2. Bra004539	76.06	100	79.09	89.53	100	89.92
3. Bra039283	74.54	79.09	100	87.98	89.92	100

	Protein		
	1	**2**	**3**
1. AtCPC	100	92.94	90.59
2. Bra004539	92.94	100	91.76
3. Bra039283	90.59	91.76	100

Bra004539 and *Bra039283* possessing $\approx$ 75% identity with the genomic sequence of *AtCPC*. A higher level of identity is observed when examining the CDS with *Bra004539* and *Bra039283* having $\approx$ 90% and $\approx$ 88% identity with *AtCPC*, respectively. A similarly high level of identity was observed for the sequences of the three proteins encoded by the putative CPC homologs.

Figures 4.5 and 4.6 show the alignment of the CDS and the protein sequences of *AtCPC*, *Bra004539* and *Bra039283*. The genomic sequence alignment is shown in Section 7.4 in the Appendix. The figures show the high level of sequence conservation between *AtCPC*, *Bra004539* and *Bra039283*.

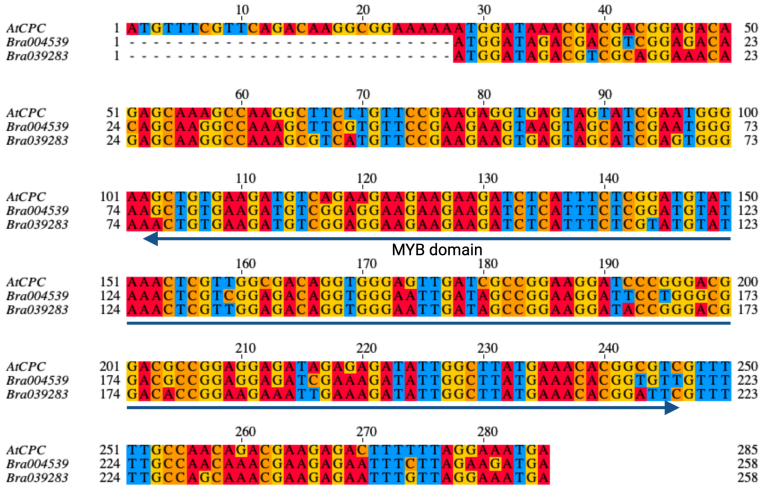

Figure 4.5: Alignment of coding sequences of *AtCPC* and the putative *B. rapa* homologs. The CDS of *AtCPC* and the two putative homologs of *CPC* in *B. rapa* were aligned using the *Clustal Omega Multiple Sequence Alignment* tool from EMBL-EBI (Madeira et al., 2022). The results were visualised using *Jalview* software. The R3 MYB domain which facilitates the interaction with bHLH transcription factors is annotated. Figure generated using Biorender.com

---

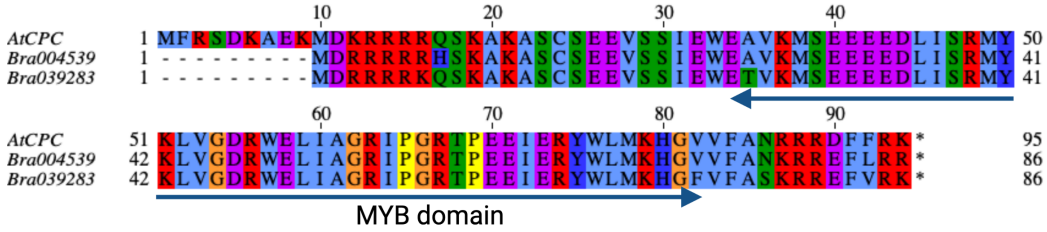

Figure 4.6: Alignment of the protein sequence of *AtCPC* and putative *B. rapa* homologs. The protein sequence of *AtCPC* aligned to the protein sequence of *Bra004539* and *Bra039283* using the *Clustal Omega Multiple Sequence Alignment* tool from EMBL-EBI (Madeira et al., 2022). Results were visualised using *Jalview* software. The following colour scheme was used to denote amino acids that belong to the same residue groups: blue: hydrophobic; red: positively charged; magenta: negatively charged; green: polar; pink: cysteines; orange: glycines; yellow: prolines; cyan: aromatic. The R3 MYB domain which facilitates the interaction with bHLH transcription factors is annotated. Figure generated using Biorender.com

---

As discussed in Section 1.2.1, *AtCPC* is a member of the R3-MYB family which function in repressing trichome formation and all contain an MYB-related DNA binding domain which is critical to their function (Schellmann et al., 2002). The MYB domain of *AtCPC* consists of amino acids 33 to 83 which, as can be seen in Figure 4.6, are almost perfectly conserved between *AtCPC* and the two putative *B. rapa* homologs (Wada et al., 2002).

In order to further test if *Bra004539* and *Bra039283* are indeed *AtCPC* homologs, I compared the genes flanking the different loci, and the results are presented in Table 4.4. Of the 20 genes examined for *Bra004539*, 18 are located on the same chromosome as *AtCPC* and 13 are direct neighbours. *Bra039283* has 18 of its closest genes located on the same chromosome as *AtCPC* and 3 of these are also direct neighbours of *AtCPC*.

Table 4.4: Synteny of *AtCPC* and putative B.rapa homologs. The table lists the ten neighbouring genes upstream and downstream of *AtCPC* (AT2G46410), *Bra004539* and *Bra039283*. The neighbouring genes were identified using the EnsemblPlants Region in Detail view to identify the surrounding genes. In the case of the *B. rapa* genes, the BRAD gene IDs were "BLASTed" and the corresponding TAIR IDs were listed for easy comparison. Gene IDs that appear in the *AtCPC* column and one of the homolog columns are highlighted in the same colour for ease of comparison.

Position	AT2G46410	Bra004539	Bra039283
-10	AT2G46310	AT2G46535	AT2G46060
-9	AT2G46320	AT2G46530	AT2G46070
-8	AT2G46330	AT2G46520	AT2G46080
-7	AT2G46340	AT2G46510	AT2G46090
-6	AT2G46360	AT2G46505	AT1G14620
-5	AT2G46370	AT2G46490	AT2G46140
-4	AT2G46375	AT3G61130	AT2G46160
-3	AT2G46380	AT2G46455	AT1G68960
-2	AT2G46390	AT2G46430	AT2G46230
-1	AT2G46400	AT2G46420	AT2G46370
+1	AT2G46420	AT2G46400	AT2G46495
+2	AT2G46430	AT3G61670	AT2G46500
+3	AT2G46440	AT2G46375	AT2G46510
+4	AT2G46450	AT2G46370	AT2G46535
+5	AT2G46455	AT2G46360	AT2G46540
+6	AT2G46460	AT2G46340	AT2G46550
+7	AT2G46470	AT2G46330	AT2G46570
+8	AT2G46480	AT2G46320	AT2G46580
+9	AT2G46490	AT2G46310	AT2G46590
+10	AT2G46493	AT2G46300	AT2G46600

Given the similarity of the gene structures, the high level of sequence identity and the high level of synteny conservation, it was concluded that *Bra004539* and *Bra039283* are the homologs of *AtCPC* in *B. rapa*. They were given the shortened names *BrCPC1*

and *BrCPC2* for *Bra004539* and *Bra039283*, respectively.

In order to fully characterise the trichome repressors in *B. rapa*, the aim was to generate higher-order mutants of *BrCPC1*, *BrCPC2*, *BrTRY1*, *BrTRY2* and *BrTRY3*. The chromosomal locations of the genes are shown in Figure 4.7.

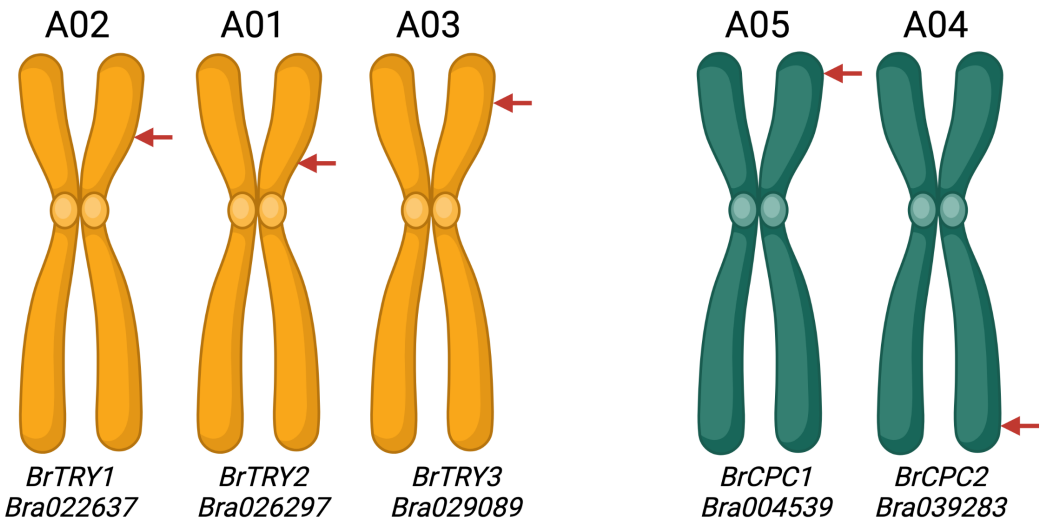

Figure 4.7: Chromosomal location of *BrCPC* and *BrTRY* genes. Arrows indicate the approximate position of each gene - *BrTRY1* on A02, *BrTRY2* on A01, *BrTRY3* on A03, *BrCPC1* on A05 and *BrCPC2* on A04. Figure generated using Biorender.com.

Because the genes are all on different chromosomes, there should be no linkage between them. As discussed in Section 3.2.1, the aim of the overall project was to replicate in *B. rapa* the floral trichome phenotype in *Arabidopsis* by Ó'Maoiléidigh et al. (2018) by introgressing mutant alleles for *BrAG1* and *BrAG2* into mutant lines for the *B. rapa* homologs of the trichome repressors *CPC* and *TRY*. *BrAG1* and *BrAG2* are located on chromosomes 3 and 1 (see Figure 3.5), respectively, which meant that the only mutant line that may prove difficult to obtain is one for *BrTRY3* and *BrAG1* as they are on the same chromosome. However, given that *BrTRY3* lacks part of the MYB-related DNA-binding domain (see above), I concluded that this double mutant was probably not going to be of great importance for this work. All other mutant alleles are on different chromosomes, or in the case of is the case for *BrAG2* and *BrTRY*, on different chromosome

arms, suggesting that a double mutant combination should be easy to generate.

4.2.2 Characterisation of the Expression of *B. rapa* Homologs of *CPC* and *TRY*

As detailed in Section 1.5, following a genome duplication event, there are several possible outcomes for the paralogous genes, namely loss of one, retention of both due to gene dosage effects, neofunctionalization or subfunctionalization, where the function of the ancestral gene becomes split between the two duplicated genes. As there appear to be three homologs of *AtTRY* and two homologs of *AtCPC* in *B. rapa*, the aim was to examine if the homologs were fully functional and if any subfunctionalization may have occurred between them. Similar methods to those used in Section 3.2.2 were employed and reverse transcription-quantitative PCR (RT-qPCR) was used to characterize the level of relative expression of the five homologs across a range of *B. rapa* tissues. As discussed in Sections 1.2 and 1.3, in *Arabidopsis*, *TRY* and *CPC* function in repressing the formation of trichomes in cells adjacent to those that develop trichomes and are also repressing the formation of trichomes on floral organs. The expression profiles of *AtTRY* and *AtCPC* as reported in Winter et al. (2007) are shown in Figure 4.8.

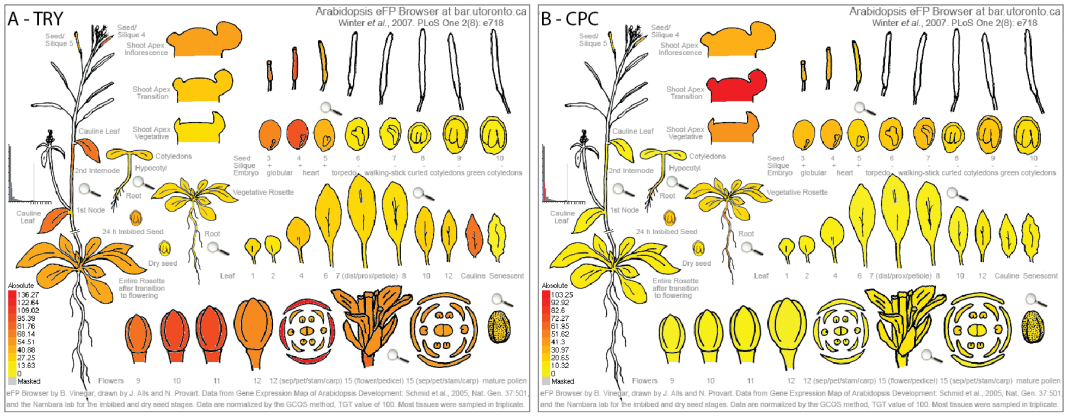

Figure 4.8: Expression profile of *AtTRY* and *AtCPC* in *Arabidopsis*. Expression domains of (A) *AtTRY* and (B) *AtCPC* in *Arabidopsis* from the Arabidopsis eFP Browser (Winter et al., 2007) Yellow indicates no or low expression whilst red indicates higher levels of expression. Expression of *AtTRY* is detected at the highest levels in developing leaves, floral tissue and developing siliques. Expression of *AtCPC* is weaker than *AtTRY* and is detected in root and leaf tissue.

The tissue samples used for the expression analysis of the *B. rapa* homologs were as follows - whole 10-day seedlings; leaf tissue; flowers sized 1 mm, 3 mm, 5 mm, 7 mm, 9 mm; individual floral organs from mature flowers equivalent to stage 13 flowers in *Arabidopsis* (Smyth et al., 1990). All samples were from the R-o-18 accession of *B. rapa*. Three biologically independent sets of samples were collected for analysis. The reference gene (*PROTEIN PHOSPHATASE 2A*) was used in the RT-qPCR experiments for *CPC* and *TRY* homologs that follow differ from that used in Section 3.2.2 for the characterisation of the *AG* homologs (*TUBULINB*). This is because it became clear that PP2A was the better reference gene for use in *B. rapa* as the expression was more consistent across tissue samples compared to *TUBULINB* (Wang et al., 2016).

4.2.2.1 Expression Patterns of *BrTRY1*, *BrTRY2* and *BrTRY3*

The results of the RT-qPCR experiments for *BrTRY1*, *BrTRY2* and *BrTRY3* genes are shown in Figure 4.9. The expression profiles for these genes broadly matched that for *AtTRY* (Figure 4.8A) with expression detected in leaf tissue, developing seedlings and developing flowers. However, the level of expression of the different genes appeared

to vary. In leaves, *BrTRY1* is expressed at a much higher level than *BrTRY2* and at the same approximate level of*BrTRY3*. In contrast, *BrTRY2* is the *TRY* gene with the highest expression in flowers including gynoecia from mature flowers. From the results of this experiment, several conclusions can be drawn. Firstly, *BrTRY3* is expressed at very low levels in all tissues examined. As discussed in Section 4.2.1.1, the protein sequence of *BrTRY3* lacks approximately half of the MYB domain that is known to be essential for the proper function of the class or regulators. This partial truncation paired with the very low levels of expression strongly suggests that *BrTRY3* has lost its function in the control of trichome repression.

A second conclusion from the RT-qPCR experiments is that of the three *AtTRY* homologs, *BrTRY1* is the only one with strong expression in leaves, while *BrTRY2* appears to be the sole *AtTRY* homolog being expressed in developing and mature gynoecia. Thus, there is evidence of *BrTRY1* and *BrTRY2* having undergone subfunctionalization through acquiring marked differences in their expression patterns.

For the aims of this project, the expression of *BrTRY2* in the gynoecium was most relevant because it suggested that targeting this gene could result in the formation of trichomes on seed pods (as detailed in Section 1.3). In contrast, mutations in *BrTRY1* may lead to supernumerary trichomes on leaves.

As discussed in Section 1.3, in order to achieve trichomes on gynoecia a strong knockout of *TRY* is necessary in addition to a knockdown of *AG*. Given the RT-qPCR data, I generated for the *B. rapa* homologs of *TRY*, it appears the *BrTRY2* is the most important homolog to target in order to attempt to generate gynoecia with trichomes as it is the most highly expressed of the three homologs.

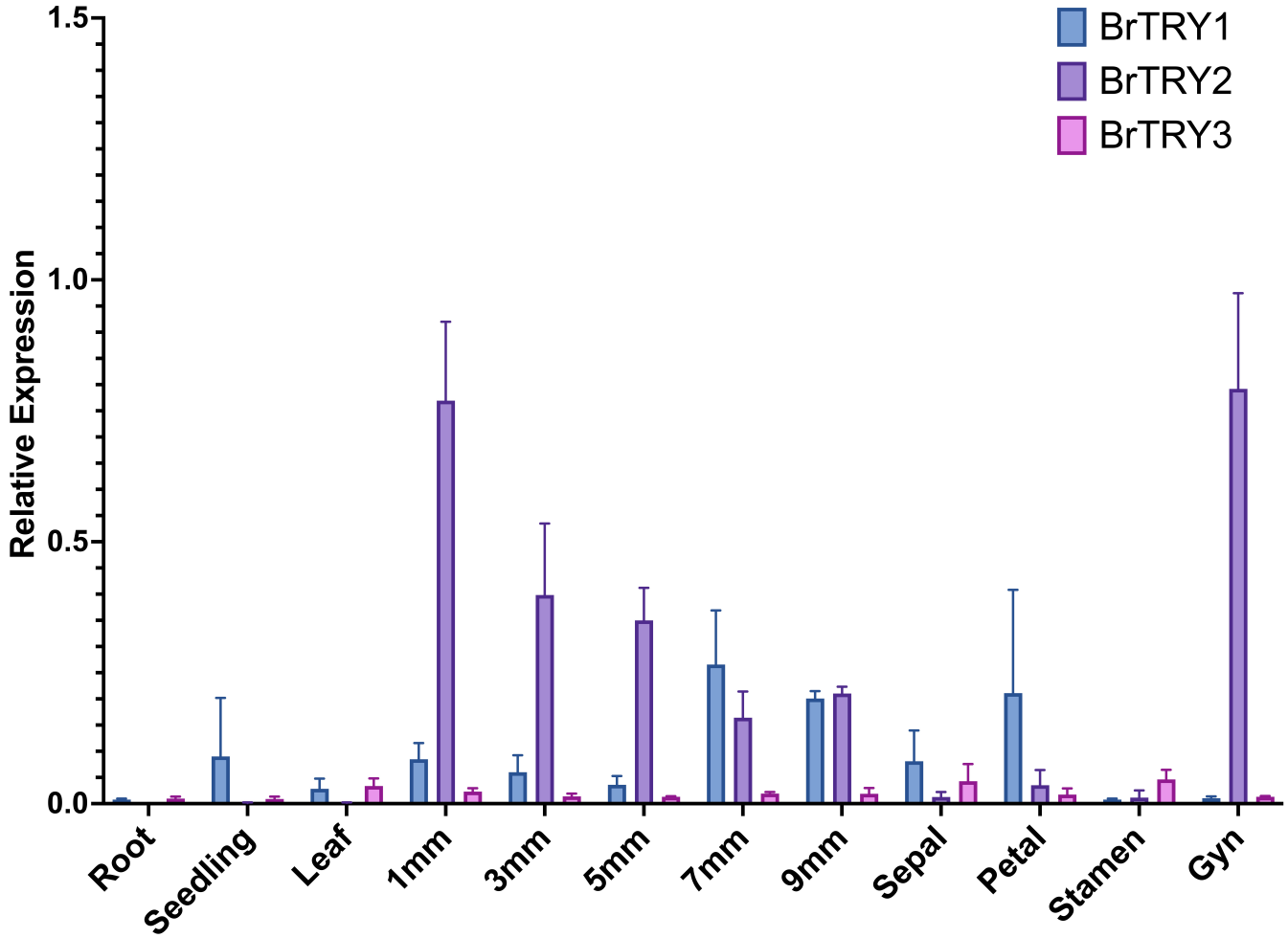

Figure 4.9: Expression of *BrTRY1*, *BrTRY2* and *BrTRY3* as determined by RT-qPCR. Results for *BrTRY1*, *BrTRY2* and *BrTRY3* are shown by blue, purple and pink bars respectively. Values are the mean of three biological replicates. Error bars represent standard deviations. The different tissue types used for the analysis are indicated. Samples from flowers of different stages are sorted by size. Expression values were normalized using a reference gene *PROTEIN PHOSPHATASE 2A*. No further statistical analyses were carried out on this data.

4.2.2.2 Expression Patterns of *BrCPC1* and *BrCPC2*

The results of RT-qPCR experiments to determine the expression strengths and domains of *BrCPC1* and *BrCPC2* expression are shown in Figure 4.10. As can be seen from the data, the expression pattern of the *B. rapa* homologs generally reflect the expression of *AtCPC* as shown in Figure 4.8B. The expression levels detected for *BrCPC1* and *BrCPC2* were lower than that for *BrTRY1-3* which is in line with what is seen in *Arabidopsis* (Wada et al., 1997; Schellmann et al., 2002). In 10-day-old seedlings, *BrCPC2* is expressed at a higher level than *BrCPC1* with the same trend observed in leaf tissue. Low levels of expression of *BrCPC1* and *BrCPC2* were detected in developing flowers of all sizes with expression peaking in 5 mm and 7 mm sized flowers. There was no clear difference in the expression trend of the two genes in the whole flower tissue samples. When examining the expression of *BrCPC1* and *BrCPC2* in individual floral organs, the highest expression was seen in sepals, which are arguably the most leaf-like floral organs. *BrCPC1* and *BrCPC2* expression is similar in sepals, petals and stamens, while *BrCPC2* expression appears much higher in mature gynoecium tissue when compared to *BrCPC1*.

In addition to controlling trichome initiation, *AtCPC* is known to play a key role in the development of root hairs in *Arabidopsis* (Wada et al., 1997). To test whether the *BrCPC* genes are also expressed in roots, I determine their transcript levels in roots of 10-day-old seedlings that were already developing root hairs. I found that *BrCPC1* and *BrCPC2* are both expressed and at relatively similar levels in root tissue. From these results, it became clear that both *BrCPC1* and *BrCPC2* are likely functional as their expression is detected across multiple tissue samples at varying levels. Both homologs are detected at varying levels in developing flowers so it appeared that both may be involved in repressing the formation of trichomes on floral organs. In the context of the overall aims of this project, it appeared that *BrCPC2* would be the gene to target for recreating the floral trichome phenotype seen described in *Arabidopsis* by Ó'Maoiléidigh et al. (2018) as it expressed at a much higher level than *BrCPC1*. Both homologs have high

expression in the sepal tissue of flowers, suggesting that perhaps trichome repression in this tissue is controlled by both homologs.

As shown in Figures 1.7 and 1.8, the knockout of *CPC* alongside the knockout of *TRY* and knockdown of *AG*, results in a dramatic increase in trichomes on gynoecia compared to *TRY* and *AG* alone. From these RT-qPCR experiments, it appears that *Br-CPC2* is the homolog that is important to knock out in efforts to obtain gynoecia with trichomes.

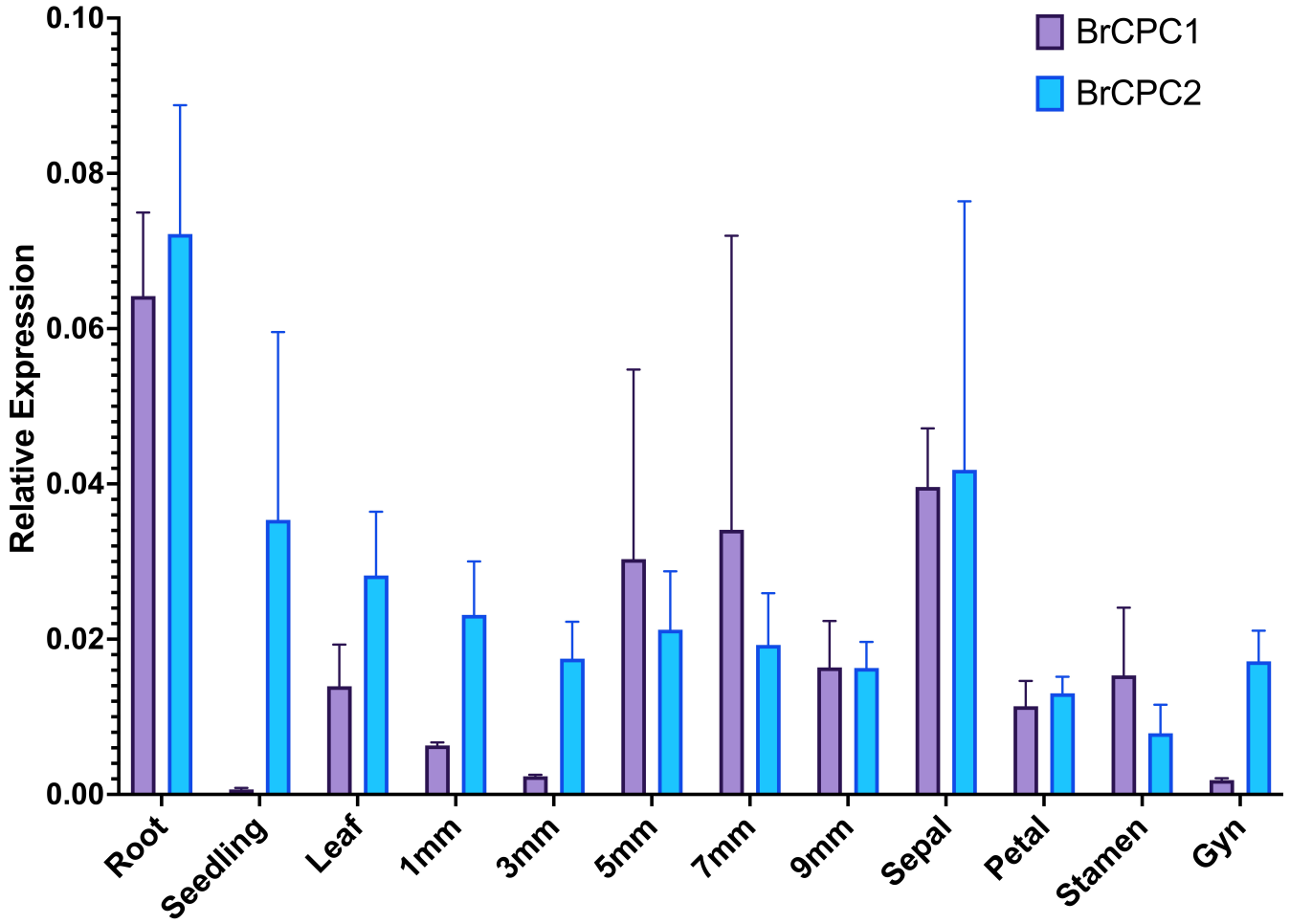

Figure 4.10: Expression of *BrCPC1* and *BrCPC2* as determined by RT-qPCR. Results for *BrCPC1* and *BrCPC2* are shown by purple and blue bars respectively. Values are the mean of three biological replicates. Error bars represent standard deviations. The different tissue types used for the analysis are indicated. Samples from flowers of different stages are sorted by size. Expression values were normalized using a reference gene *PROTEIN PHOSPHATASE 2A*. No further statistical analyses were carried out on this data.

4.2.3 Comparing Expression of *BrCPC* and *BrTRY* Genes in Two Lines of *B. rapa*

As will be discussed in detail in Chapter 5, in addition to working with the inbred line R-o-18, seeds of another *B. rapa* accession were obtained from the Tohoku University Brassica Seed Bank. This accession, named C634, develops trichomes on the leaf blade as well as the leaf margin. This is in contrast to R-o-18, which develops trichomes sparsely on the leaf margin only (Figure 5.1). As the aim of my work was to generate *B. rapa* with trichome on floral organs, the mutant TILLING lines (as discussed in Sections 3.2.6 and 4.2.7), which are in an R-o-18 background, needed to be introgressed into the C634 background to allow for potential trichome formation. I wished to examine the expression levels of the homologs of *TRY* and *CPC* in the C634 accession and see how they compared to the expression levels seen in R-o-18. Using RT-qPCR, the expression of the *BrTRY* and *BrCPC* genes was determined in C634 and compared to R-o-18. To this end, leaf and mature gynoecium tissue from C634 and R-o-18 was used.

As shown in Figure 4.11A, in the case of *BrTRY1-3*, the expression levels of *BrTRY1* and *BrTRY2* in both C634 and R-o-18 leaves are broadly the same. In the case of *BrTRY3*, higher expression is detected in the C634 leaf tissue which could indicate the *BrTRY3* is indeed functional despite its truncated length and is expressed at higher levels in this accession due to the increased numbers of trichomes developing. Figure 4.11A also shows the expression levels of *BrTRY1-2* in C634 and R-o-18 mature gynoecium tissue. A similar trend is seen in C634 to R-o-18 with very low expression of *BrTRY1* and *BrTRY3* compared to very high expression of *BrTRY2*. However, *BrTRY2* is expressed more highly in the C634 accession tissue than in R-o-18.

Figure 4.11B shows that in the case of *BrCPC1* and *BrCPC2*, there is a similar general trend in expression in both accessions with *BrCPC2* expressed at a higher level than *BrCPC1*. However, *BrCPC2* is expressed at an overall higher level in R-o-18 compared

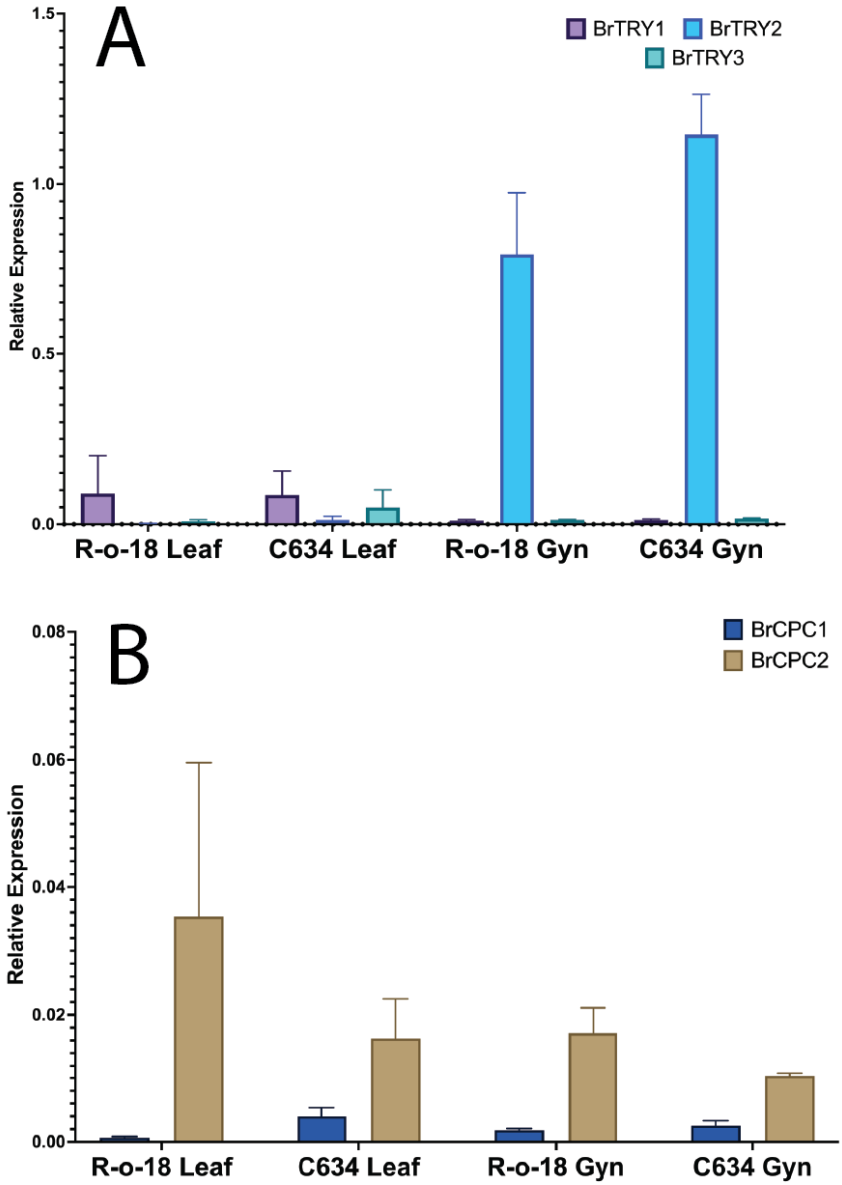

Figure 4.11: Expression of *BrTRY1*, *BrTRY2*, *BrTRY3*, *BrCPC1* and *BrCPC2* in leaf and gynoecium tissue of two *B. rapa* lines. Results for *BrTRY1*, *BrTRY2*, *BrTRY3*, *BrCPC1* and *BrCPC2* are shown in purple, blue, green, dark blue, and tan, respectively. "Gyn" indicates mature gynoecium tissue. Values are the mean of three biological replicates. Relative expression of genes standardized to *PROTEIN PHOSPHATASE 2A*. Error bars represent standard deviations.

---

to C634. In the case of gynoecium tissue, again *BrCPC2* is expressed at a higher level than *BrCPC1* in both accessions but expression of *BrCPC2* is higher in R-o-18 (Figure 4.11B).

These results further support the results from Section 4.2.2 that *BrTRY2* and *BrCPC2*

are the key homologs to target in order to affect trichome repressor activity specifically in the gynoecium. These results show that these expression trends hold true for both the R-o-18 line and C634 accession, so backcrossing the mutant alleles into C634 should provide no complications for the overall aim of this project.

4.2.4 Characterisation of the 3' UTRs of *BrCPCs* and *BrTRYs*

To further characterise the homologs of trichome repressor genes in *B. rapa*, I determined their 3' untranslated regions (UTRs) which had not been previously described. As mentioned in Section 3.2.3, it is known that UTRs can contain *cis*-acting elements which function in post-transcriptional control and also have an effect on the stability, transport and translation efficiency of mRNAs (Mignone et al., 2002; Srivastava et al., 2018). To identify the 3' UTRs, 3' Rapid Amplification of cDNA Ends (RACE)-PCR, as per the protocol described by Scotto-Lavino et al., 2007, was employed. The sequences obtained via 3'RACE-PCR were aligned with the known 3'UTR of *AtTRY* (Figure 4.12) and the 3'UTR of *AtCPC* (Figure 4.13).

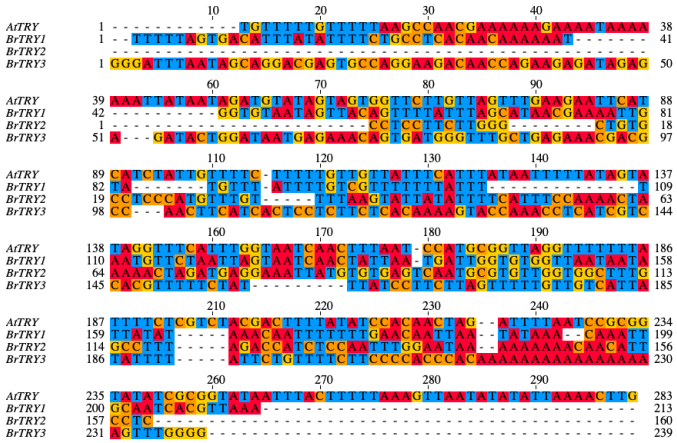

Figure 4.12: Alignment of 3'UTR sequences of *TRY* homologs. Alignment of the 3'UTR sequence of *AtTRY*, *BrTRY1*, *BrTRY2* and *BrTRY3*. The sequence of the 3'UTR of *AtTRY* was obtained from *EnsemblPlants* while the sequence of the 3'UTR of *BrTRY1*, *BrTRY2* and *BrTRY3* were obtained via 3'RACE-PCR. The sequences were aligned using the *Clustal Omega Multiple Sequence Alignment* tool from EMBL-EBI (Madeira et al., 2022). The results were visualised using *Jalview* software.

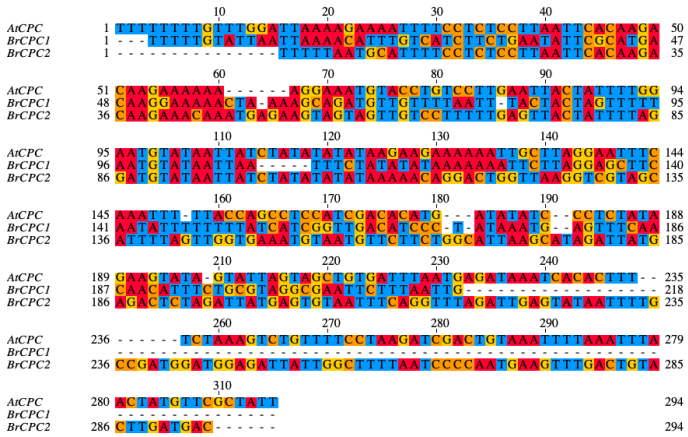

Figure 4.13: Alignment of 3'UTR sequences of *CPC* homologs. Alignment of the 3'UTR sequences of *AtCPC*, *BrCPC1* and *BrCPC2*. The sequence of *AtCPC* was obtained from *EnsemblPlants* while the sequence of the 3'UTR of *BrCPC1* and *BrCPC2* was obtained via 3'RACE-PCR. The sequences were aligned using the *Clustal Omega Multiple Sequence Alignment* tool from EMBL-EBI (Madeira et al., 2022). The results were visualised using *Jalview* software.

The 3'UTR of *AtTRY* is 283 bp in length. The sequences obtained using 3'RACE-PCR showed that the 3'UTRs for *BrTRY1*, *BrTRY2* and *BrTRY3* were 213 bp, 160 bp and 239 bp, in length, respectively. The alignment of the 3' UTR *AtTRY* and *BrTRY1-3* did not reveal regions that could explain the differences in expression that are observed for *BrTRY1*, *BrTRY2* and *BrTRY3*. The percentage identity values are shown in Table 4.5.

Table 4.5: Percentage identity matrix for *TRY* 3'UTRs. Table showing the level of percentage identity between *AtTRY* and the *BrTRY* homologs.

	1	2	3	4
1. AtTRY	100.00	54.68	45.91	42.15
2. BrTRY1	54.68	100.00	45.59	43.22
3. BrTRY2	45.91	45.59	100.00	37.84
4. BrTRY3	42.15	43.22	37.84	100.00

In Section 3.2.3, the level of sequence identity of the 3' UTRs of *AtAG*, *BrAG1* and *BrAG2* was $\approx 70\%$. Here we see that the level of sequence identity is lower for *AtTRY* and its homologs - the highest for *BrTRY1* at $\approx 55\%$, followed by *BrTRY2* with $\approx 46\%$ identity and the lowest is *BrTRY3* at $\approx 42\%$. This could explain the differences in the

expression levels and expression domains seen in Sections 4.2.2 and 4.2.3.

In the case of *AtCPC*, the annotated 3'UTR is 294 bp in length. The sequences obtained via 3'RACE-PCR were 218 bp in the case of *BrCPC1* and 294 bp in the case of *BrCPC2*. Given that 3'UTRs are under less selective pressure to maintain the sequence than coding DNA sequences, it is not surprising that the 3'UTRs for *AtCPC* and *BrCPC2* are the same length but do not share a high percentage of sequence identity. The aligned sequences are shown in Figure 4.13 and again there are no clear differences in the sequences that could explain the difference in expression strength and patterns - one might expect to see regions sequence conservation between the *AtCPC* 3'UTR with the *BrCPCs* 3'UTRs if they were functionally critical to the processing of the mRNA. The percentage identity of the sequences is shown in Table 4.6.

Table 4.6: Percentage identity matrix for *CPC* 3'UTRs. Table comparing the percentages of sequence identity of the 3'UTR of *AtCPC* and the *BrCPC* homologs.

	1	**2**	**3**
1. AtCPC	100.00	60.00	53.85
2. BrCPC1	60.00	100.00	47.57
3. BrCPC2	53.85	47.57	100.00

Similar to the *BrTRY* genes, the level of conservation in the sequences for the *BrCPC* genes are lower than the values seen for *BrAG* genes. For *BrCPC1*, there is $\approx 60\%$ sequence identity and $\approx 54\%$ identity for *BrCPC2*. The differences in the expression levels across different tissues seen for these two genes (Sections 4.2.2 and 4.2.3) could be explained in part by the differences in the 3'UTRs.

A similar analysis to the one conducted in Section 3.2.3 for the *AG* homologs was employed for the 3'UTR sequences of the *CPC* and *TRY* homologs. The 3'UTR sequences were scanned using the Meme-Suite Tool Simple Enrichment Analysis (SEA) (Section 2.8) which identifies motifs that are relatively enriched in the sequences under examination compared with shuffled sequences or control sequences. SEA was used to

scan the 3' UTRs for the following motifs that are described as AU-rich elements.

As shown in Table 4.7, a number of these sites were detected in the sequences of the *TRY* homologs. While most of them appeared to be shared between all of them, several sites were present only in *AtTRY* but absent from *BrTRY1-3*. All three *B. rapa TRY* homologs lack the "AUUUA", "WUUUW" and "$(W)_5UUU(W)_5$". *BrTRY2* alone lacks a "$(W)_4UUU(W)_4$" recognition site.

Table 4.7: ARE binding sites detected in 3'UTR of *TRY* homologs. The table shows the number of sites for each ARE motif detected in the 3'UTR sequences using SEA analysis. W in the motif indicates that the nucleotide could be an A or U base.

Motif Sequence	*AtTRY*	*BrTRY1*	*BrTRY2*	*BrTRY3*
AUUUA	1	0	0	0
WUUUW	1	0	0	0
WWUUUWW	1	1	1	1
$(W)_3UUU(W)_3$	1	1	1	1
$(W)_4UUU(W)_4$	1	1	0	1
$(W)_5UUU(W)_5$	1	0	0	0
UUUGUUU	1	1	1	1
GUUUG	1	1	1	1
AWUAAA	1	1	1	1
Total	**9**	**6**	**5**	**6**

It is striking that *BrTRY1* and *BrTRY3* have the exact same ARE binding site profiles detected in their 3'UTR while simultaneously showing broadly similar expression profiles - expressed in leaf tissue and not expressed in the gynoecium of flowers. Perhaps the absence of the "$(W)_4UUU(W)_4$" recognition site in the 3'UTR of *BrTRY2* is a possible reason for its extremely high expression in gynoecium tissue compared to the other two *TRY* homologs.

The results of SEA analysis of *CPC* homologs are shown in Table 4.8. Both *BrCPC1* and *BrCPC2* have four detected ARE binding sites whilst *AtCPC* has eight. *AtCPC* and *BrCPC1* share a "$(W)_4UUU(W)_4$" and "AWUAAA" recognition site. *BrCPC2* alone possesses a "UUUGUUU" and "GUUUG" recognition site. Although both homologs

share the same number of ARE binding sites, it is possible that the different binding sites contained within their 3'UTRs could be contributing to the difference in expression patterns seen for the two in Sections 4.2.2 and 4.2.3.

Table 4.8: ARE binding sites detected in 3'UTR of *CPC* homologs. The table shows the number of sites for each ARE motif detected in the 3'UTR sequences using SEA analysis. W in the motif indicates that the nucleotide could be an A or U base.

Motif Sequence	*AtCPC*	*BrCPC1*	*BrCPC2*
AUUUA	2	0	0
WUUUW	1	1	1
WWUUUWW	1	1	1
$(W)_3UUU(W)_3$	1	0	0
$(W)_4UUU(W)_4$	1	1	0
$(W)_5UUU(W)_5$	1	0	0
UUUGUUU	0	0	1
GUUUG	0	0	1
AWUAAA	1	1	0
Total	**8**	**4**	**4**

The differences detected in the numbers and types of ARE binding sites in the 3'UTRs of the *B. rapa TRY* and *CPC* homologs could explain the different expression patterns as discussed in Section 4.2.2. However as mentioned previously in Section 3.2.3, it is not possible to conclusively state this without further experimental work such as the deletion and swapping of ARE motifs.

4.2.5 Comparative Analysis of Upstream Promoter Regions of *TRY* Homologs

It has been shown that SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL) transcription factors (of which there are 16 members) are essential in ensuring the correct expression of *AtTRY* in *Arabidopsis* by binding to both elements in the upstream promoter sequence and its second intron (Yu et al., 2010; Xue et al., 2014; Cui et al.,

2022).

I therefore reasoned that employing *JASPAR* software to predict potential SPL binding sites (similar to experimental work as described in Sections 3.2.4 and 3.2.5) may give an insight as to the cause for the differential expression patterns detected for *BrTRY1*, *BrTRY2* and *BrTRY3* as seen in Sections 4.2.2 and 4.2.3. It has also been previously shown that AG plays an essential role in the upregulation of *AtTRY* in floral tissue (Wellmer et al., 2006; Ó'Maoiléidigh et al., 2018). As discussed in Section 1.1.3, MADS-domain transcription factors work in higher-order multimer protein complexes so it was decided to also carry out an analysis of the predicted MADS-domain transcription factor binding sites for the *TRY* homologs.

The aim was to predict both SPL transcription factors and MADS-domain transcription factor binding sites for both the promoter and intronic sequences of *BrTRY1*, *BrTRY2*, and *BrTRY3* and compare the results to the predicted binding sites of the promoter and second intron of *AtTRY*. The results of the promoter sequence analysis are presented here and the intron sequence analysis is discussed in Section 4.2.6.

When examining the upstream promoter sequences, there is a large difference in the number of predicted SPL transcription factor binding sites for the four genes being examined (Table 4.9). There are 18 sites predicted for *AtTRY*, none for *BrTRY1*, two for *BrTRY2*, and 10 for *BrTRY3*. For the *TRY* homologs, the transcription factors that differ most in the number of predicted sites are SPL3, SPL4, SPL9, and SPL10.

Table 4.9: *JASPAR* binding site analysis of promoter regions of *TRY* homologs for SPL transcription factor binding sites. The 2 kb upstream of the start codon for the *TRY* homologs were analyzed using *JASPAR* to scan for predicted SPL transcription factor binding sites. The transcription factor binding sites were determined using a relative profile score threshold of 90%. The gene name and matrix ID for the binding site used are provided.

Gene	Matrix ID	AtTRY	BrTRY1	BrTRY2	BrTRY3
SPL3	MA0577.1	1	0	0	0
	MA0577.2	1	0	0	1
	MA0577.3	0	0	0	1
SPL4	MA1058.1	5	0	1	1
SPL5	MA1059.1	2	0	0	1
SPL8	MA0578.1	2	0	0	0
SPL9	MA1322.1	1	0	1	1
	MA1322.2	2	0	0	1
SPL10	MA1799.1	1	0	0	0
SPL11	MA1056.1	1	0	0	0
SPL12	MA1057.1	2	0	0	1
SPL13A	MA1321.1	0	0	0	1
SPL14	MA0586.2	0	0	0	1
Total		**18**	**0**	**2**	**10**

The predicted MADS-domain transcription factor binding sites for the upstream promoter region are summarised in Table 4.10.

Table 4.10: *JASPAR* binding site analysis of promoter regions of *TRY* homologs for MADS-domain transcription factor binding sites. The 2 kb upstream of the start codon for the *TRY* homologs were analyzed using *JASPAR* to scan for MADS-domain transcription factor binding sites. The transcription factor binding sites were determined using a relative profile score threshold of 90%. The gene name and matrix ID for the binding site used are provided.

Gene	Matrix ID	AtTRY	BrTRY1	BrTRY2	BrTRY3
AG	MA0005.1	0	1	0	1
AGL6	MA1205.1	0	0	0	0
AGL15	MA0548.2	1	0	0	0
AGL27	MA1012.1	1	0	0	0
AGL42	MA1201.1	3	0	2	1
AGL55	MA1202.1	0	2	0	2
AGL63	MA1203.1	1	0	0	0
AP1	MA0940.1	2	2	2	3
AP3	MA0556.1	1	0	0	0
FLC	MA0558.1	1	0	0	0
PI	MA0559.1	1	0	1	1
SEP1	MA0584.1	0	0	0	2
SEP3	MA0563.1	2	0	1	2
SEP4	MA0001.1	1	0	1	1
SHP1	MA0585.1	0	0	0	0
SOC1	MA0554.1	1	0	0	0
SVP	MA0555.1	1	0	0	0
Total		**16**	**5**	**7**	**11**

AtTRY has the most predicted binding sites at 16 with *BrTRY3* again having the most sites of the homologs (11 sites), followed by *BrTRY2* (7 sites) and then *BrTRY1* (5 sites). There is a trend observed in the predicted binding sites for SPL and MADS-domain transcription factors in the upstream promoter regions of *TRY* with *AtTRY* and *BrTRY3* being the most similar despite *BrTRY3* expression being detected only at extremely low levels in Sections 4.2.2 and 4.2.3.

4.2.6 Comparative Analysis of Introns of *TRY* Homologs

In order to further investigate the difference in expression pattern and levels of the *BrTRY1-3*, *JASPAR* software was used to analyze the second intron of *AtTRY* and the

only intron of the *BrTRY1-3* for predicted binding sites of SPL and MADS-domain transcription factor binding sites. The results are displayed in Table 4.11 and Table 4.12.

Table 4.11: *JASPAR* binding site analysis of introns of *TRY* homologs for SPL transcription factor binding sites. The DNA sequence of the second intron of *AtTRY* and the only introns of *BrTRY1-3* were analyzed using the *JASPAR* software to predict SPL transcription factor binding sites. The transcription factor binding sites were determined using a relative profile score threshold of 90%. The gene name and matrix ID for the binding site used are provided.

Gene	Matrix ID	AtTRY	BrTRY1	BrTRY2	BrTRY3
SPL1	MA1055.2	0	1	0	0
SPL3	MA0577.1	0	1	0	0
SPL4	MA1058.1	5	2	2	0
SPL5	MA1059.1	1	2	1	0
SPL8	MA0578.1	0	0	1	0
SPL9	MA1322.1	2	1	2	0
SPL10	MA1799.1	1	0	0	0
SPL12	MA1057.1	2	1	1	0
SPL13A	MA1321.1	0	1	1	0
SPL14	MA0586.2	0	0	0	0
SPL15	MA1320.1	0	0	0	0
Total		**11**	**9**	**8**	**0**

AtTRY, *BrTRY1*, and *BrTRY2* had a similar number of predicted sites for SPL transcription factors - 11, 9 and 8 respectively whilst there were none predicted for *BrTRY3*. The lack of SPL transcription factor binding sites in the *BrTRY3* intronic region could perhaps explain the absence of expression detected in floral and gynoecium tissue.

Perhaps over the scale of time, the SPL transcription factor binding sites in the intronic regions of *BrTRY3* were lost via mutations and with *BrTRY1* and *BrTRY2* still able to fulfil the function of *TRY* there was no negative consequence for the loss of *BrTRY3* resulting in it becoming a pseudogene potentially.

When examining the intronic sequences for MADS-domain transcription factor binding sites, there are a similar number of sites predicted for *AtTRY* and *BrTRY2* and fewer sites for *BrTRY1* and *BrTRY3*. The similarity of sites predicted for *AtTRY* and *BrTRY2*

may explain *BrTRY2* expression at a much higher level in gynoecium tissue as seen in Section 4.2.2.

Table 4.12: *JASPAR* binding site analysis of introns of *TRY* homologs for MADS-domain transcription factor binding sites. The DNA sequence of the second intron of *AtTRY* and the only intron of *BrTRY1-3* were analyzed using the *JASPAR* software to predict SPL transcription factor binding sites. The transcription factor binding sites were determined using a relative profile score threshold of 90%. The gene name and matrix ID for the binding site used are provided.

Gene	Matrix ID	AtTRY	BrTRY1	BrTRY2	BrTRY3
AG	MA0005.1	0	1	1	0
AGL27	MA1012.1	1	0	0	0
AGL3	MA0001.1	0	0	1	0
AGL42	MA1201.1	0	1	1	0
AP1	MA0940.1	2	0	0	1
SEP3	MA0563.1	1	0	1	0
SOC1	MA0554.1	1	0	0	0
Total		**5**	**2**	**4**	**1**

4.2.7 Results of *TRY* Complementation Experiments

In order to functionally characterise the *B. rapa* homologs of *CPC* and *TRY*, I conducted a second line of experiments to determine whether they could functionally replace their *Arabidopsis* homologs.

For the complementation experiments, the mutant line that was chosen for *TRY* was *try-29760*. This allele contains a T-DNA insertion which renders the protein nonfunctional (Tominaga et al., 2008). *Try-29760* plants exhibit an increase in the overall number of trichomes as well as an increase in trichome clustering on the leaf surface. This phenotype can be seen in Figure 4.14.

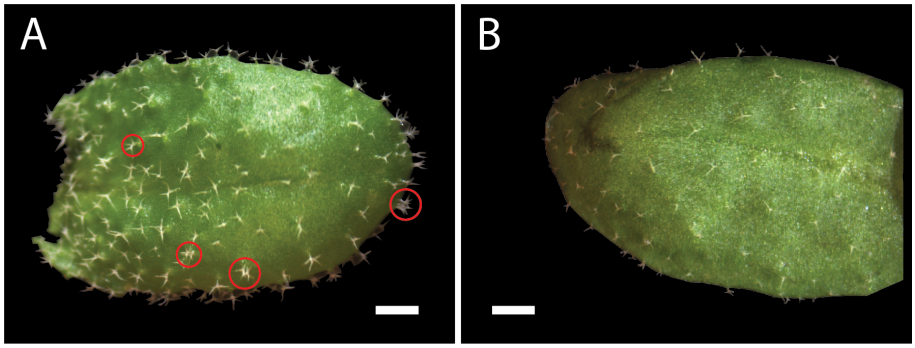

Figure 4.14: Phenotype of *try-29760* mutant plants. (A) A leaf from *try-29760* plant. An increase in trichome numbers and clusters (circled in red) can be seen compared to the leaf from a wild-type plant (accession: Col-0) shown in (B). Scale bars: 1 mm.

The coding DNA sequence (including the 3'UTRs) of *BrTRY1*, *BrTRY2* and *BrTRY3* were placed under the control of the CaMV 35S promoter and the *AtTRY* promoter (kindly provided by Mr. Andrea Finocchio from our laboratory), respectively. *AtTRY* was included in these experiments as a positive control. The constructs were transformed into try-29760 and Col-0 plants for comparison and the results of these experiments are described below.

4.2.7.1 Complementation Experiments Using a p35S:*AtTRY* Construct

AtTRY overexpression lines have been previously generated by Schellmann et al., 2002 and were described as being completely glabrous. In agreement with the previous data, all p35S:*AtTRY* lines I obtained (n=3 for try-29760; n=2 for Col-0) were glabrous on their leaves and stems (Figure 4.15), showing that the general approach I used for the complementation experiments worked.

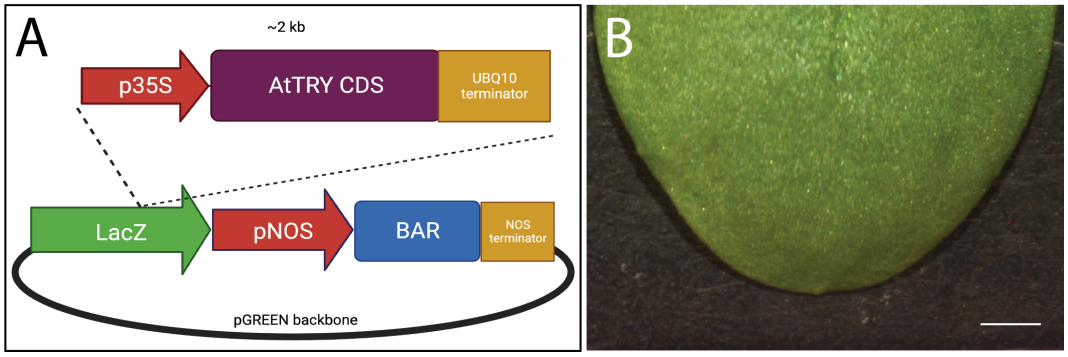

Figure 4.15: Results of overexpression of *AtTRY*. (A) Schematic of the genetic construct used to generate the *AtTRY* overexpression lines. p35S, CaMV 35S promoter sequence; AtTRY CDS, the coding DNA sequence of *AtTRY*; Terminator, the *UBQ10* terminator; LacZ, lacZ a modified version of the β- galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA^{TM} resistance gene; NOS terminator, nopaline synthase terminator sequence. Figure generated in Biorender. (B) A leaf from a p35S:*AtTRY* plant. Scale bar: 1 mm.

4.2.7.2 Complementation Experiments Using a p35S:*BrTRY1* Construct

I next expressed *BrTRY1* from the 35S promoter using the same strategy outlined in the preceding section. All independent lines (n=4 for try-29760; n=1 for Col-0) obtained showed phenotypes (Figure 4.16) similar to those of p35S:*AtTRY* lines in that their leaves and stems were completely glabrous. Thus, BrTRY1 can functionally replace *AtTRY*.

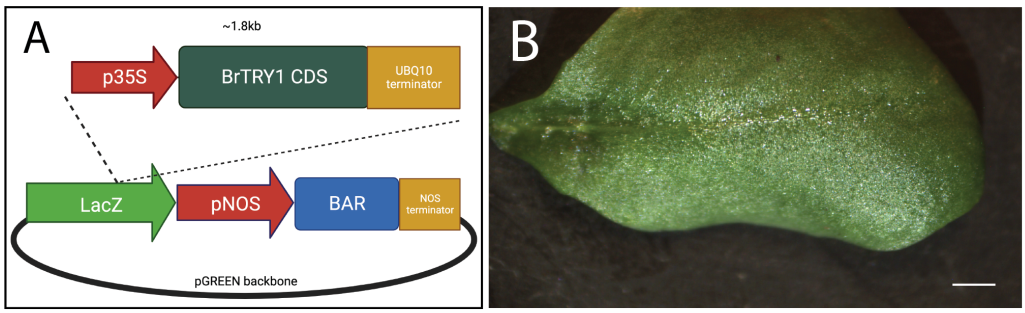

Figure 4.16: Results of overexpression of *BrTRY1*. (A) Schematic of the genetic construct used to generate the *BrTRY1* overexpression lines. p35S, CaMV 35S promoter sequence; BrTRY1 CDS, the coding DNA sequence of *BrTRY1*; Terminator, the *UBQ10* terminator; LacZ, lacZ a modified version of the β- galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA^{TM} resistance gene; NOS terminator, nopaline synthase terminator sequence. Figure generated using Biorender.com. (B) A glabrous leaf from a p35S:*BrTRY1* plant. Scale bar: 1 mm.

4.2.7.3 Complementation Experiments Using a p35S:*BrTRY2* Construct

To test whether *BrTRY2* can also functionally replace *AtTRY*, I overexpressed it from the 35S promoter in *try-29760* and in wild-type plants. All independent lines (n=3 for try-29760; n=3 for Col-0) I obtained produced glabrous leaves and stems (Figure 4.17), suggesting that *BrTRY2* can indeed functionally replace *AtTRY*.

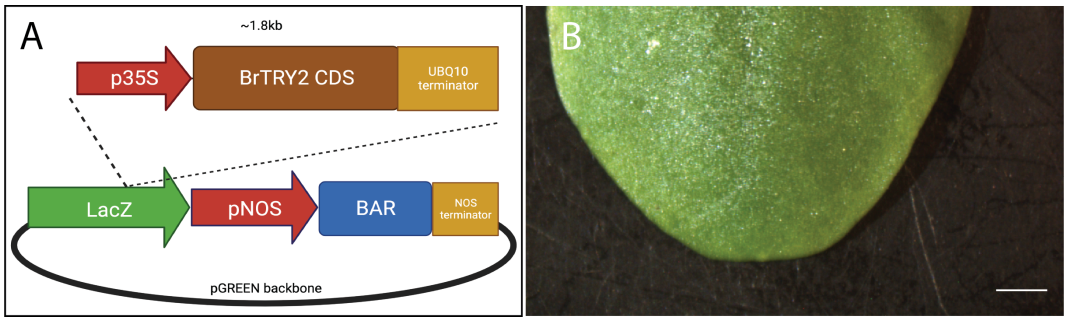

Figure 4.17: Results of overexpression of *BrTRY2*. (A) Schematic of the genetic construct used to generate the *BrTRY2* overexpression lines. p35S, CaMV 35S promoter sequence; BrTRY2 CDS, the coding DNA sequence of *BrTRY2*; Terminator, the *UBQ10* terminator; LacZ, lacZ a modified version of the β- galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA^{TM} resistance gene; NOS terminator, nopaline synthase terminator sequence. Figure generated using Biorender.com (B) A glabrous leaf from a p35S:*BrTRY2* plant. Scale bar: 1 mm.

4.2.7.4 Complementation Experiments Using a p35S:*BrTRY3* Construct

The construct containing the CDS of *BrTRY3* driven by the 35S promoter was cloned and transformed into *try-29760* and Col-0 plants using the floral dip method described in Section 2.7.1. Unfortunately, no transformants were recovered. The floral dip was repeated but transformants were not recovered in time for phenotypic analysis and inclusion as part of this work.

4.2.7.5 Complementation Experiments Using a pTRY:*AtTRY* Construct

In addition to the constitutive over-expression lines described above, I also drove expression of the *TRY* homologs from the regulatory region of *AtTRY* to test whether they could rescue *try-29760* mutants when expressed in the normal domain of *AtTRY*. As a control, I first generated lines carrying a pTRY:*AtTRY* construct. A total of 4 independent lines (n=4 for *try-29760*) were isolated. All lines generated (n=4 for *try-29760*) containing the pTRY:*AtTRY* transgene in a *try-29760* displayed a leaf trichome phenotype that resembled the wild-type (Figure 4.18B). Thus, the construct could rescue the *try-29760* mutant phenotype.

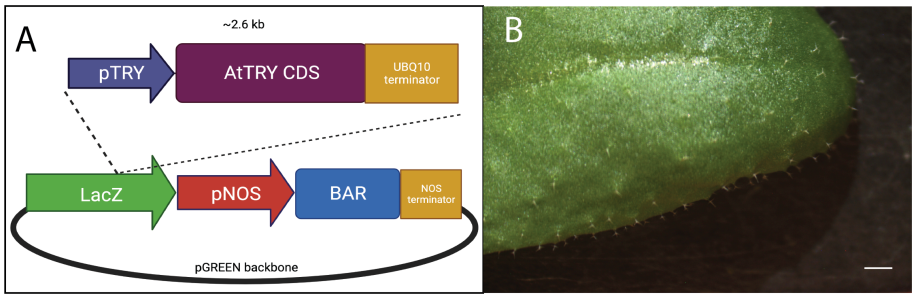

Figure 4.18: Effects of a pTRY:*AtTRY* construct on trichome formation. (A) Schematic of the genetic construct used to generate the pTRY:*AtTRY* lines. pTRY, *AtTRY* promoter sequence, 1.5kb upstream of the start codon; AtTRY CDS, the coding DNA sequence of *AtTRY*; Terminator, the *UBQ10* terminator; LacZ, lacZ a modified version of the β- galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA^{TM} resistance gene; NOS terminator, nopaline synthase terminator sequence. Figure generated using Biorender.com (B) A leaf of a pTRY:*AtTRY* *try-29760* plant showing a rescue of the leaf trichome phenotype of *try-29760* (compared to Figure 4.16A). Scale bar: 1 mm.

4.2.7.6 Complementation Experiments Using a pTRY:*BrTRY1* Construct

After confirming the experimental approach for these complementation experiments, I tested whether a pTRY:*BrTRY1* could also rescue the *try-29760* mutant phenotype. All transgenic lines I obtained (n=4 for *try-29760*; n=2 for Col-0) displayed a leaf trichome phenotype that resembled the wild type (Figure 4.19B), suggesting that even under the native *AtTRY* promoter, *BrTRY1* can functionally replace *AtTRY*.

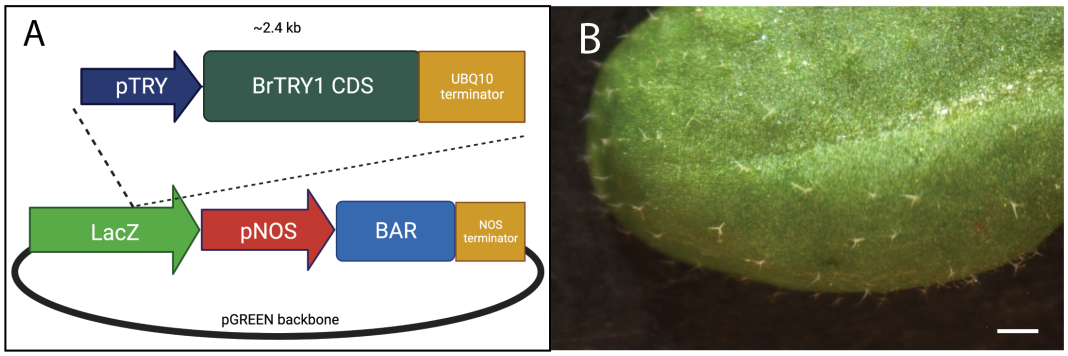

Figure 4.19: Effects of a pTRY:*BrTRY1* construct on trichome formation. (A) Schematic of the genetic construct used to generate the pTRY:*BrTRY1* lines. pTRY, *AtTRY* promoter sequence, 1.5kb upstream of the start codon; BrTRY1 CDS, the coding DNA sequence of *BrTRY1*; Terminator, the *UBQ10* terminator; LacZ, lacZ a modified version of the β- galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA^{TM} resistance gene; NOS terminator, nopaline synthase terminator sequence. Figure generated using Biorender.com (B) A leaf phenotype of pTRY:*BrTRY1* plant *try-29760* plant showing a rescue of the leaf trichome phenotype of *try-29760* (Figure 4.16A). Scale bar: 1 mm.

4.2.7.7 Complementation Experiments Using a pTRY:*BrTRY2* Construct

All lines generated (n=4 for *try-29760*; n=2 for L-*er*) containing the pTRY:*BrTRY2* transgene displayed a leaf trichome phenotype that resembled the wild-type (Figure 4.20B), suggesting that under the native *AtTRY* promoter, *BrTRY2* can also functionally replace *AtTRY*.

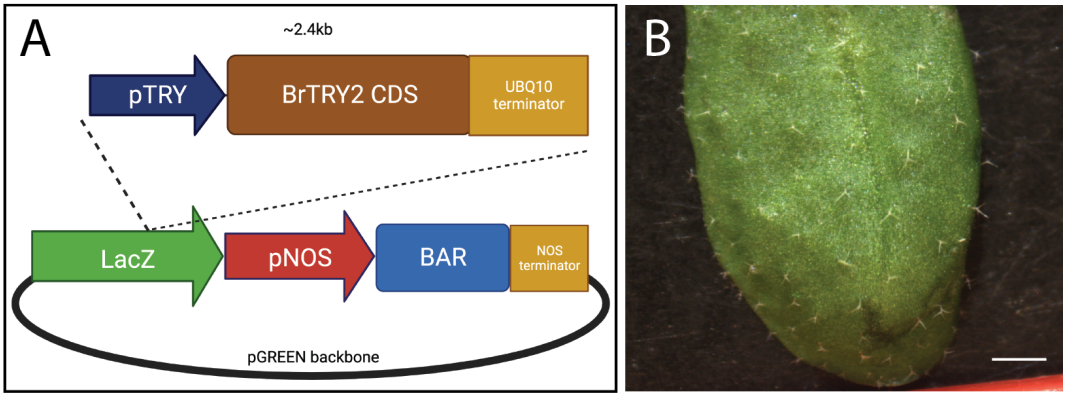

Figure 4.20: Effects of a pTRY:*BrTRY2* construct on trichome formation. (A) Schematic of the genetic construct used to generate the pTRY:*BrTRY2* lines. pTRY, *AtTRY* promoter sequence, 1.5kb upstream of the start codon; BrTRY2 CDS, the coding DNA sequence of *BrTRY2*; Terminator, the *UBQ10* terminator; LacZ, lacZ a modified version of the β- galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA^{TM} resistance gene; NOS terminator, nopaline synthase terminator sequence. Figure generated using Biorender.com. (B) A leaf phenotype of pTRY:*BrTRY2* plant *try-29760* plant showing a rescue of the leaf trichome phenotype of *try-29760* (Figure 4.16A). Scale bar: 1 mm.

4.2.7.8 Complementation Experiments Using a pTRY:*BrTRY3* Construct

All lines generated containing the pTRY:*BrTRY3* transgene (n=3 for *try-29760*) displayed a weak rescue of the *try-29760* mutant phenotype on the basis that it eliminates the trichome clustering phenotype of *try-29760* (Figure 4.21). While the number of trichomes on leaves was still increased as in *try-29760*, clustering of trichomes appeared to be absent in these lines. This is a potentially interesting finding given that *BrTRY3* lacks part of the normal *TRY* sequence (see Section 4.2.1.1). However, the number of independent lines obtained for this analysis was small and the sample size probably needs to be increased to allow any conclusions to be reached about how well *BrTRY3* can replace *AtTRY*.

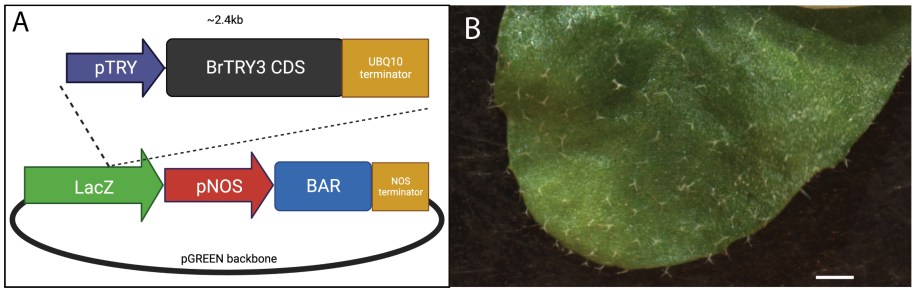

Figure 4.21: Effects of a pTRY:*BrTRY3* construct on trichome formation. (A) Schematic of the genetic construct used to generate the pTRY:*BrTRY3* lines. pTRY, *At-TRY* promoter sequence, 1.5kb upstream of the start codon; BrTRY3 CDS, the coding DNA sequence of *BrTRY3*; Terminator, the *UBQ10* terminator; LacZ, lacZ a modified version of the β- galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA^{TM} resistance gene; NOS terminator, nopaline synthase terminator sequence. Figure generated using Biorender.com (B) A leaf of a pTRY:*BrTRY3 try-29760* plant that does not show a full rescue of the leaf trichome phenotype of *try-29760* (compare to Figure 4.16A). Scale bar: 1 mm.

4.2.8 Results of *CPC* Complementation Experiments

In the case of *CPC*, several mutant lines were considered for complementation experiments. The line that was ultimately chosen was *cpc-1* which has an increased number of trichomes on the leaf and stem surfaces (Figure 4.22). This line contains a T-DNA insertion in *AtCPC* which renders the gene nonfunctional (Wada et al., 1997). As shown in Figure 4.22, leaves and stems of *cpc-1* mutant plants exhibit a marked increase in the number of trichomes when compared to wild-type plants.

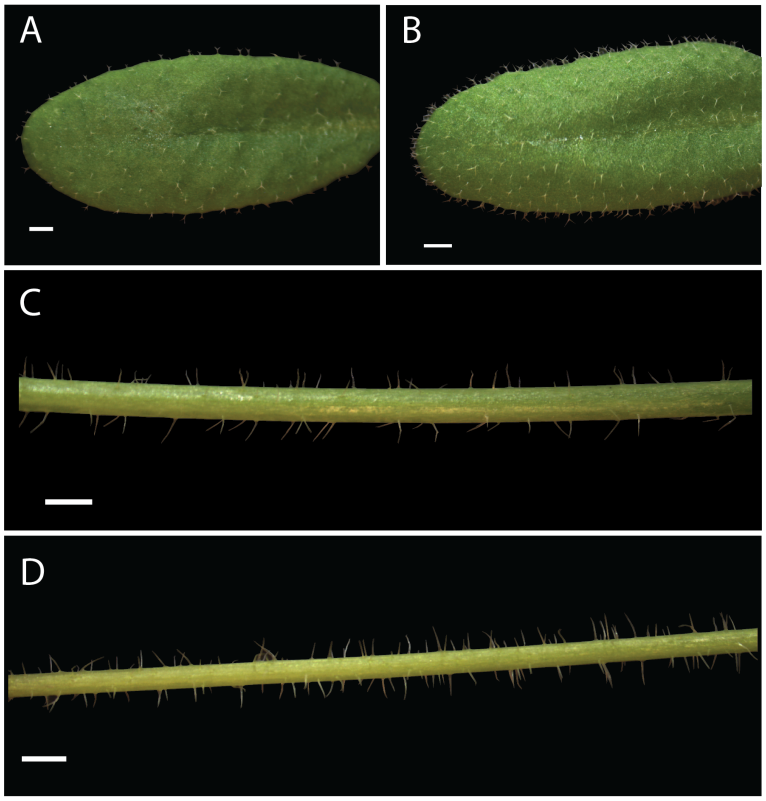

Figure 4.22: Phenotype of *cpc-1* mutant plants. (A) A leaf from a wild-type plant (accession: *Ws*). (B) A leaf from a *cpc-1* plant with an increased number of trichomes. (C) A stem of a *Ws* plant, and (D) a stem of a *cpc-1* plant. Scale bars: 1 mm.

To test whether *BrCPC1* and *BrCPC2* can functionally replace *AtCPC*, their coding sequences (including the 3'UTRs) were placed under the control of the CaMV 35S promoter. *AtCPC* was also included as a positive control. In addition, the CDS of the genes were placed under the control of the *AtCPC* promoter which was kindly provided by Mr. Andrea Finocchio from our laboratory. The constructs were transformed into *cpc-1* and for comparison, into wild-type plants (Accession: *Ws*), and transgenic plants were phenotypically characterized.

4.2.8.1 Complementation Experiments Using a p35S:*AtCPC* Construct

A total of 12 independent lines (n=8 for *cpc-1*; n=4 for *Ws*) were obtained containing the p35S:*AtCPC* construct. All lines had entirely glabrous leaves and stems (Figure 4.23B-C). These phenotypes concur with previously published results for p35S:*AtCPC*

over-expression lines (Schellmann et al., 2002) and showed the validity of the experimental approach.

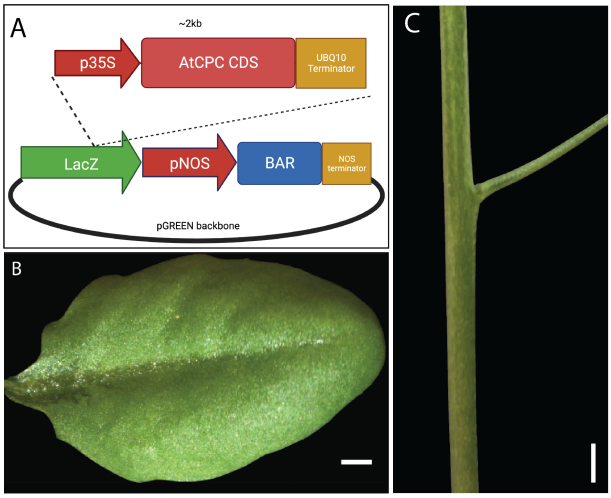

Figure 4.23: Results of overexpression of *AtCPC*. (A) Schematic of the genetic construct used to generate the *AtCPC* overexpression lines. p35S, CaMV 35S promoter sequence; AtCPC CDS, the coding DNA sequence of *AtCPC*; Terminator, the *UBQ10* terminator; LacZ, lacZ a modified version of the β- galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA^{TM} resistance gene; NOS terminator, nopaline synthase terminator sequence. Figure made using Biorender.com (B) A leaf of a typical p35S:*AtCPC* plant lacking any trichomes. (C) The glabrous stem of a p35S:*AtCPC* plant. Scale bars: 1 mm.

4.2.8.2 Complementation Experiments Using a p35S:*BrCPC1* Construct

I next tested whether the overexpression of *BrCPC1* could functionally replace *AtCPC*. Similar to the results of the *AtCPC* overexpression experiment described above, p35S:*BrCPC1* plants (n=4 for *cpc-1*; n=11 for *Ws*) displayed an entirely glabrous phenotype (Figure 4.24C). Thus, it appears that the function of *BrCPC1* is similar to that of *AtCPC*.

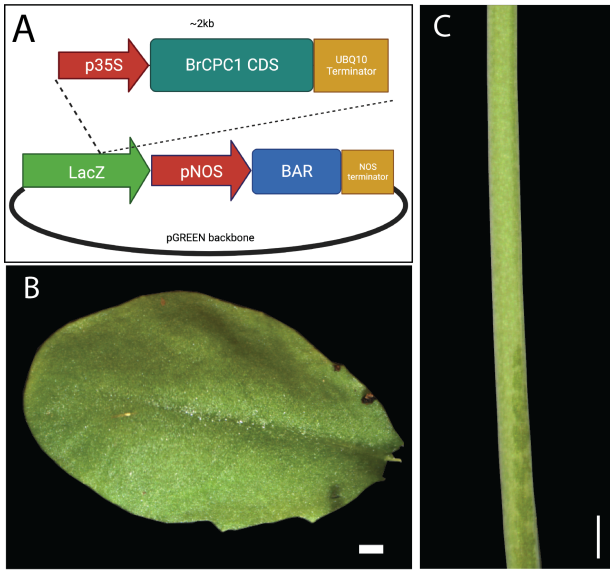

Figure 4.24: Results of overexpression of *BrCPC1*. (A) Schematic of the genetic construct used to generate the *BrCPC1* overexpression lines. p35S, CaMV 35S promoter sequence; BrCPC1 CDS, the coding DNA sequence of *BrCPC1*; Terminator, the *UBQ10* terminator; LacZ, lacZ a modified version of the β- galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA^{TM} resistance gene; NOS terminator, nopaline synthase terminator sequence. Figure made using Biorender.com (B) A leaf of a typical p35S:*BrCPC1* plant lacking any trichomes. (C) The glabrous stem of a p35S:*BrCPC1* plant. Scale bars: 1 mm.

4.2.8.3 Complementation Experiments Using a p35S:*BrCPC2* Construct

To test the effects of *BrCPC2* over-expression, 4 independent lines (n=2 for *cpc-1*; n=2 for *Ws*) carrying a p35S:*BrCPC2* construct were isolated and characterized. As observed for BrCPC1 over-expression, all lines were fully glabrous (Figure 4.25). Thus, *BrCPC2* is also able to functionally replace *AtCPC*.

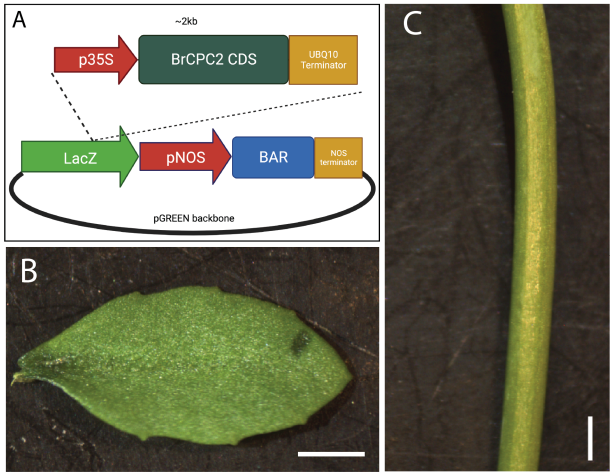

Figure 4.25: Results of overexpression of *BrCPC2*. (A) Schematic of the genetic construct used to generate the *BrCPC2* overexpression lines. p35S, CaMV 35S promoter sequence; BrCPC2 CDS, the coding DNA sequence of *BrCPC2*; Terminator, the *UBQ10* terminator; LacZ, lacZ a modified version of the β- galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA^{TM} resistance gene; NOS terminator, nopaline synthase terminator sequence. Figure made using Biorender.com (B) A leaf of a typical p35S:*BrCPC2* plant lacking any trichomes. (C) The glabrous stem of a p35S:*BrCPC2* plant. Scale bars: 1 mm.

4.2.8.4 Complementation Experiments Using a pCPC:*AtCPC* Construct

In addition to the constitutive over-expression lines described above, I also drove the expression of the *CPC* homologs from the promoter region of *AtCPC* to test whether they could rescue *cpc-1* mutants when expressed in the normal domain of *AtCPC*. As a control, I first generated lines carrying a pCPC:*AtCPC* construct as shown in Figure 4.26. A total of 17 independent lines (n=10 for *cpc-1*; n= 7 for *Ws*) were isolated. The characterization of these lines showed that *cpc-1* plants containing the construct were rescued for their trichome phenotype, suggesting that the *AtCPC* promoter was active in the right domain.

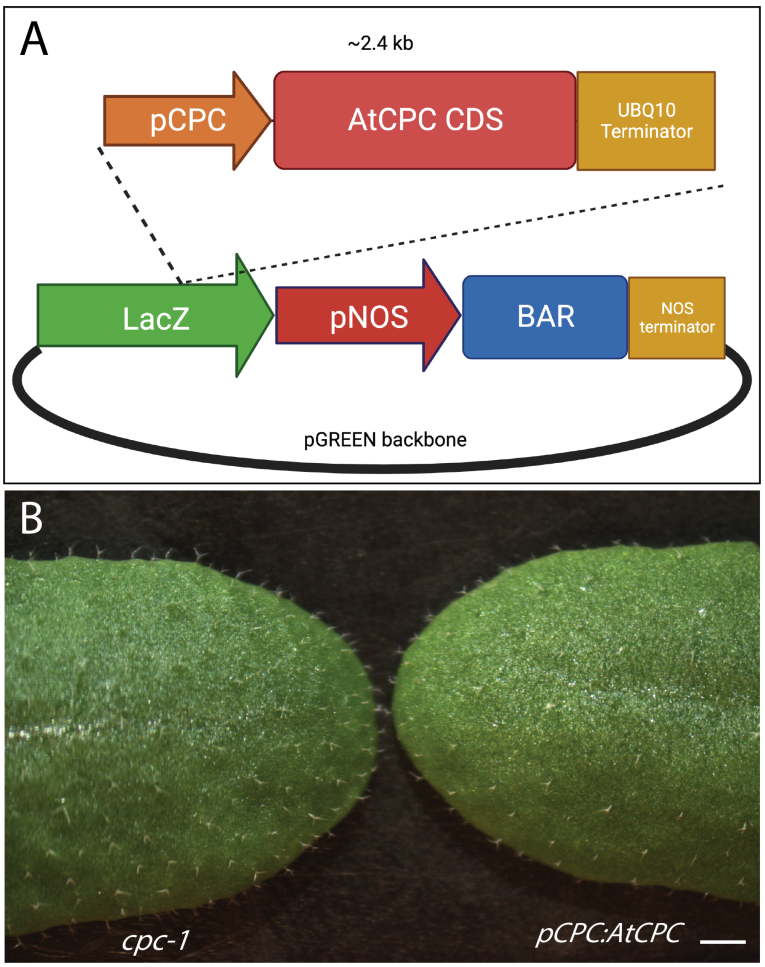

Figure 4.26: Effects of a pCPC:*AtCPC* construct on trichome formation. (A) Schematic of the genetic construct used to generate the pCPC:*AtCPC* lines. pCPC, *AtCPC* promoter sequence, 1.5kb upstream of the start codon; AtCPC CDS, the coding DNA sequence of *AtCPC*; Terminator, the *UBQ10* terminator; LacZ, lacZ a modified version of the β- galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA^{TM} resistance gene; NOS terminator, nopaline synthase terminator sequence. Figure made using Biorender.com (B) A leaf of a typical pCPC:*AtCPC* *cpc-1* plant compared to a *cpc-1* leaf. Scale bar: 1 mm.

4.2.8.5 Complementation Experiments Using a pCPC:*BrCPC1* Construct

All lines generated (n=2 for *cpc-1*; n=2 for *Ws*) containing the pCPC:*BrCPC1* transgene displayed a leaf trichome phenotype that resembled the wild-type (Figure 4.27), suggesting that when driven under the native *AtCPC* promoter, *BrCPC1* can functionally replace AtCPC.

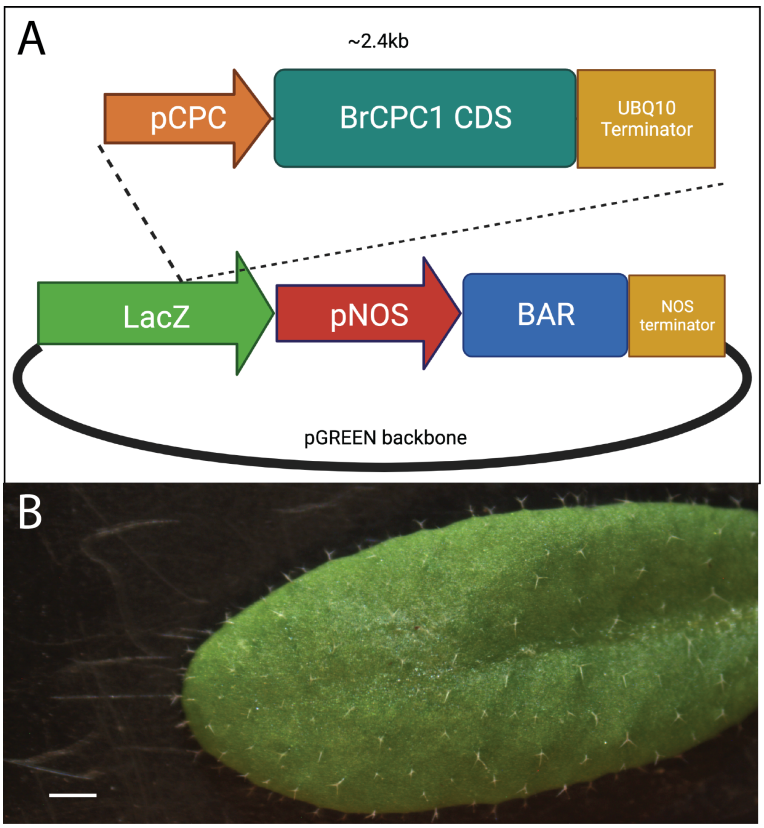

Figure 4.27: Effects of a pCPC:*BrCPC1* construct on trichome formation. (A) Schematic of the genetic construct used to generate the pCPC:*BrCPC1* lines. pCPC, *AtCPC* promoter sequence, 1.5kb upstream of the start codon; BrCPC1 CDS, the coding DNA sequence of *BrCPC1*; Terminator, the *UBQ10* terminator; LacZ, lacZ a modified version of the β- galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA^{TM} resistance gene; NOS terminator, nopaline synthase terminator sequence. Figure made using Biorender.com (B) A leaf of a typical pCPC:*BrCPC1 cpc-1* plant. Scale bar: 1 mm.

4.2.8.6 Complementation Experiments Using a pCPC:*BrCPC2* Construct

All lines generated containing the pCPC:*BrCPC2* transgene (n=5 for *cpc-1*; n=1 for *Ws*) displayed a leaf trichome phenotype that resembled the wild type (Figure 4.28), suggesting that when driven under the native *AtCPC* promoter, *BrCPC2* can functionally replace *AtCPC*.

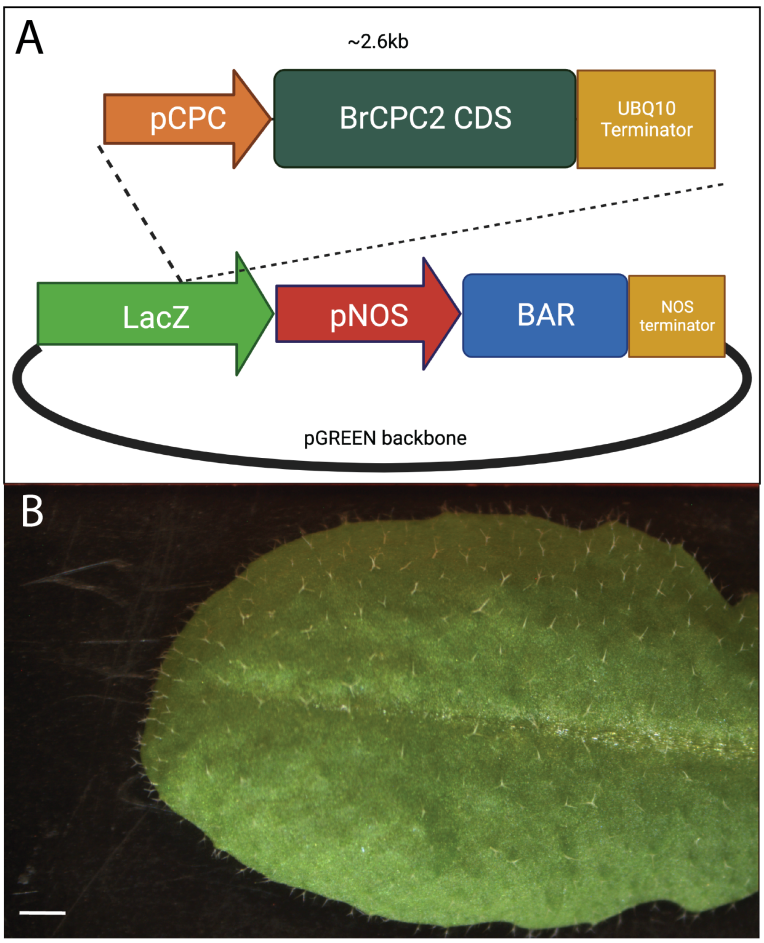

Figure 4.28: Effects of a pCPC:*BrCPC2* construct on trichome formation. (A) Schematic of the genetic construct used to generate the pCPC:*BrCPC2* lines. pCPC, *AtCPC* promoter sequence, 1.5kb upstream of the start codon; BrCPC2 CDS, the coding DNA sequence of *BrCPC2*; Terminator, the *UBQ10* terminator; LacZ, lacZ a modified version of the β- galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA^{TM} resistance gene; NOS terminator, nopaline synthase terminator sequence. Figure made using Biorender.com (B) A leaf of a typical pCPC:*BrCPC2* *cpc-1* plant compared to a *cpc-1* leaf. Scale bar: 1 mm.

4.2.9 Identification of Mutant Alleles for *BrCPC1*, *BrCPC2*, *BrTRY1*, *BrTRY2* and *BrTRY3*

As shown in Section 4.2.1, the *CPC* and *TRY* genes from *B. rapa* encode proteins that have a high percentage of sequence identity to their *Arabidopsis* homologs. The results of RT-qPCR analyses (see Section 4.2.2) showed that generally, their expression patterns also mimicked those of the *Arabidopsis* genes. To further test if the *BrTRY* and *BrCPC* genes indeed function as trichome repressors, I aimed to analyse the effect that

their inactivation has on the trichome formation in *B. rapa*. To this end, *B. rapa* mutant lines were selected from the catalogue of commercially available TILLING lines from RevGenUK at the John Innes Centre in Norwich.

For *BrTRY1*, four mutant lines were available. Two of them contained mutations in intronic sequences and one contained a synonymous mutation. The only other line available was a missense mutation - JI31767-a. The mutation in this line caused the glutamic acid at position 34 to be converted into lysine. The substitution of a negatively charged amino acid with a positively charged amino was deemed to have a possible impact on the overall structure or function of the protein. This line was denoted *brtry1-1*. For *BrTRY2*, seven lines were available. Four contained mutations in intronic regions, one a synonymous mutation and one a missense mutation. The remaining line (JI30284-a) was chosen for this work as it contained a G to A conversion at position 60 in the CDS, leading to an early stop codon at position 20 of the protein sequence. This line was denoted *brtry2-1*. *BrTRY3* had only one mutant line available, which contained a mutation in the intronic region. It was concluded that this line would likely have no effect on the function of the gene and hence was not used in this study. As shown in Section 4.2.2, *BrTRY3* is expressed at extremely low levels when compared to *BrTRY1* and *BrTRY2*, and misses part of the normal TRY protein sequence. Thus, the unavailability of a mutant line for it was deemed unlikely to affect the investigation into the redundancy between the *TRY* homologs or the attempts to generate *B. rapa* with trichomes on floral organs. The mutant alleles discussed above are summarised in Figure 4.29.

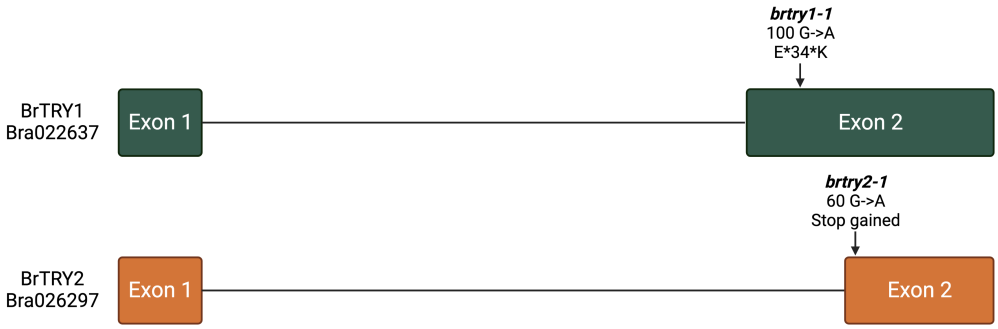

Figure 4.29: Mutant lines for *BrTRY* genes. Schematic showing the mutant lines for *BrTRY1* (top) and *BrTRY2* (bottom). Exons are numbered and shown as coloured boxes; black lines represent introns. The *brtry1-1* line contains a missense mutation at position 34 of the amino acid - a glutamic acid is replaced with a lysine. The *brtry2-1* line contains an early stop codon at position 20 of the protein sequence. Figure generated with BioRender.com.

For *BrCPC1*, 15 mutant lines were available from RevGenUK. Of these 15, five were lines that contained mutations in intronic regions, 4 contained synonymous substitutions and 5 contained missense mutations that resulted in amino acid substitutions not likely to have negative effects on protein structure or function. The one line that appeared suitable for my project was JI32146-b. This line has a C-to-T conversion at position 235 in the CDS which results in an early stop codon at position 79 of the protein sequence. This line was named *brcpc1-1*. Similarly, for *BrCPC2* there were 12 lines available from RevGenUK, however, many of them were determined to unlikely affect protein function. Five were lines that contained mutations in intronic regions, one contained a synonymous mutation and five contained missense mutations that were unlikely to have an impact on protein function. The line chosen for my work was JI32170-a, which contains a G to A conversion at position 72 in the CDS, resulting in an early stop codon at position 24 in the protein. This line was denoted *brcpc2-1*. The mutant alleles described above are shown in Figure 4.30.

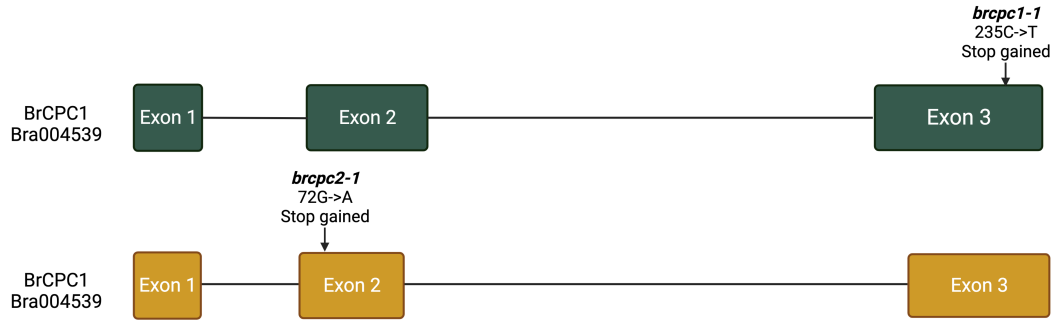

Figure 4.30: Mutant lines for *BrCPC* genes. Schematic showing the mutant lines for *BrCPC1* (top) and *BrCPC2* (bottom). Exons are numbered and shown as coloured boxes; black lines represent introns. The *brcpc1-1* line contains an early stop codon at position 79 of the protein. The *brcpc2-1* contains an early stop codon at position 20 in the protein sequence. Figure generated with BioRender.com.

4.2.10 Generation and Analysis of *BrCPC* and *BrTRY* Single Mutants

As advised by RevGenUK, all TILLING lines obtained were back-crossed with wild-type plants in order to reduce the mutational load of the lines, improve plant health and ensure that any phenotypic differences observed could be attributed to the SNP that defines each line. All *BrCPC* and *BrTRY* mutant lines were back-crossed with R-o-18 or C634 (see Section 5.3.1 for C634) *B. rapa* accession at least 4 times as recommended by RevGenUK.

4.2.10.1 Trichome Phenotype of *brtry1-1* Mutant Plants

Plants homozygous for the *brtry1-1* mutation were initially isolated by sequencing as described in Section 2.3.2 and then by DMAS genotyping as described in Section 2.3.3. The phenotypes described in this section are those observed in homozygous *brtry1-1* mutants that have been introgressed with the *B. rapa* accession C634 to allow for the potential of trichome formation. This is described in greater detail in Chapter 5. Leaves of these plants exhibited an increase in the number of trichomes seen along the leaf

veins on the abaxial leaf side compared to the leaves of wild-type plants (Figure 4.31). Trichome numbers in other areas of the leaf appeared to be unaffected.

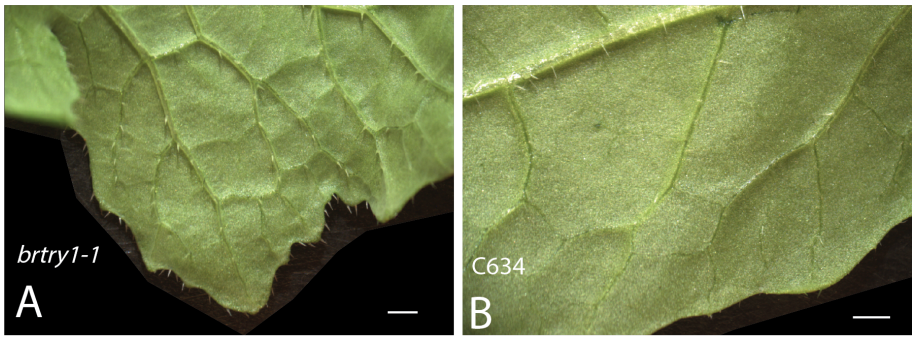

Figure 4.31: Leaf phenotype of *brtry1-1* plants. (A) Abaxial side of a *brtry1-1* leaf. Note the increase in trichome numbers along the leaf veins when compared to a leaf of the C634 accession (B). Scale bars: 1 mm.

4.2.10.2 Trichome Phenotype of *brtry2-1* Mutant plants

Plants homozygous for the *brtry2-1* mutation were isolated by genotyping as described in Section 2.3.2 and then by DMAS genotyping as described in Section 2.3.3. As described above for *brtry1-1*, the phenotypes described in this section are those observed in homozygous *brtry2-1* mutants that had been crossed into the accession C634 to allow for the potential of trichome formation.

The trichome phenotype observed for *brtry2-1* is shown in Figure 4.32. In plants, an overall increase in the number of trichomes was observed on the stem with the spacing between trichomes reduced when compared to wild-type plants. An additional trichome phenotype was observed for *brtry2-1* which is shown in Figure 4.33. Panel (A) shows a typical petiole from a wild-type C634 plant with a small cluster of trichomes present on the abaxial surface of the base of the petiole. Typically, 1-3 trichomes were observed at this location. Panel (B) shows a petiole from a *brtry2-1* plant which formed an increased number of trichomes at the abaxial surface of the base of the petiole with 8-10 trichomes observed on average. Trichomes were also observed developing further up the petiole

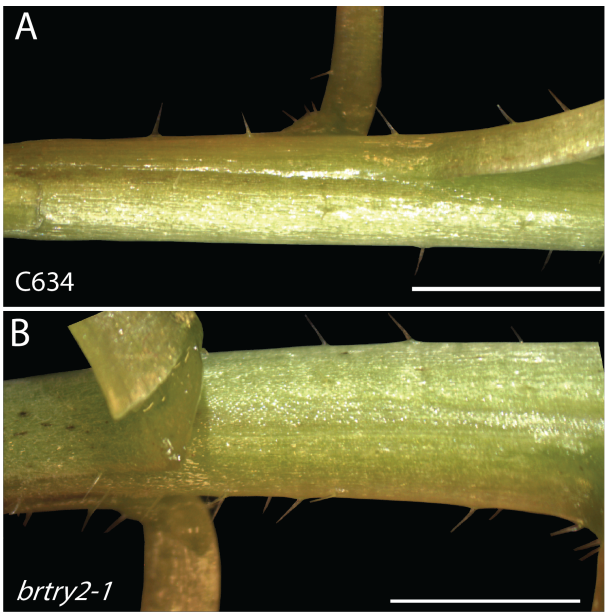

Figure 4.32: Stem trichome phenotype of *brtry2-1* plants. (A) Stem from a wild-type C634 plant. (B) Stem from *brtry2-1* mutant introgressed into a C634 background. Note the increased number of trichomes in the mutant. Scale bars: 5 mm.

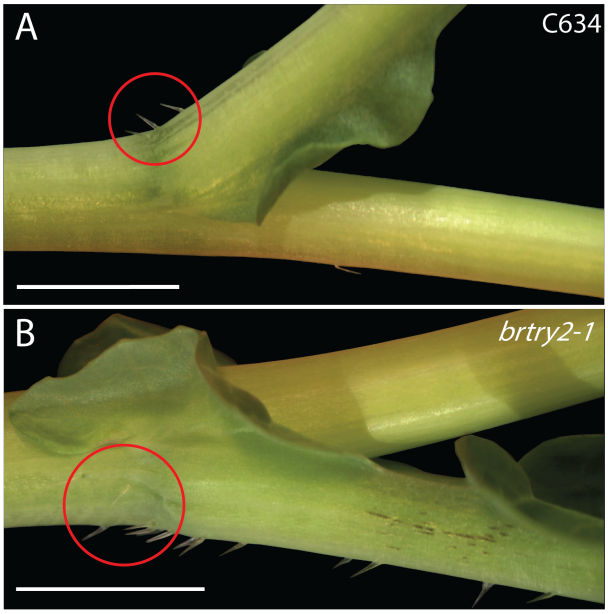

Figure 4.33: Petiole trichome phenotype of *brtry2-1* plants. (A) and (B) show the petiole from a wild-type C634 plant and from *brtry2-1* in a C634 background respectively. (A) The petiole trichome phenotype observed in C634 plants where typically 1-3 trichomes form on the base of the petiole. The red circle highlights the trichomes. (B) The petiole trichome phenotype observed in *brtry2-1* mutants introgressed into the C634 accession. There is an increased number of trichomes on the base of the petiole - typically 8 to 10 trichomes that are located closer together than seen in wild-type C634. The red circle highlights the trichomes. Scale bars: 5 mm.

---

towards the leaf in *brtry2-1* compared to C634. In contrast, no increase in the number of leaf trichomes was detected. Given the results of the expression analysis (Section 4.2.2), which showed that *BrTRY2* is expressed at a low level in leaves, this was not unexpected.

4.2.10.3 Trichome Phenotype of *brcpc1-1* Mutant plants

Plants homozygous for the *brcpc1-1* mutation were isolated by sequencing and by DMAS genotyping as described above. The allele was crossed into the *B. rapa* accession C634 to allow for the potential of trichome formation and then analysed. When compared to wild-type C634 plants, leaves of *brcpc1-1* have a slightly increased ($\approx$ 10%) number of trichomes on leaf margins (Figure 4.34). No precise counts of trichomes were carried out in the scope of this work. In contrast, the number of trichomes on leaf blades appeared to be unaffected.

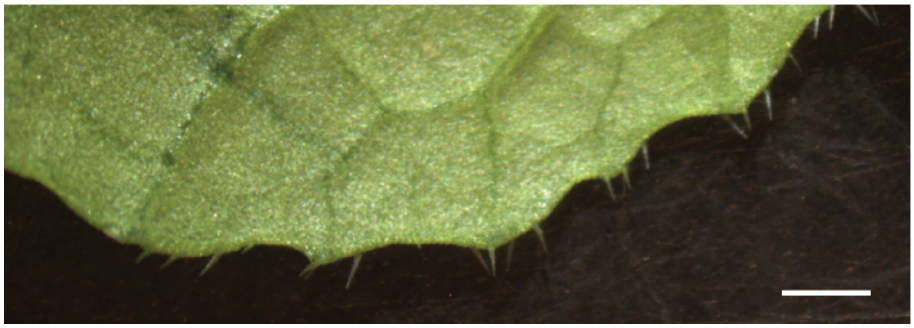

Figure 4.34: Leaf trichome phenotype of *brcpc1-1* plants. The leaves of *brcpc1-1* plants show a slight increase in the number of trichomes on the leaf margin. Scale bar: 1 mm.

4.2.10.4 Trichome Phenotype of *brcpc2-1* Mutant plants

Plants homozygous for the *brcpc2-1* mutation were isolated by sequencing and genotyping and then crossed into the *B. rapa* accession C634 to allow for the potential of trichome formation. Leaves of *brcpc2-1* mutant plants appeared to have a slightly higher number of trichomes with leaf margins being particularly affected (Figure 4.35A). Occasionally, a small trichome cluster formed on leaf margins (Figure 4.35B), a phenotype

that I never observed in any of the other trichome repressor mutant lines or the wild-type.

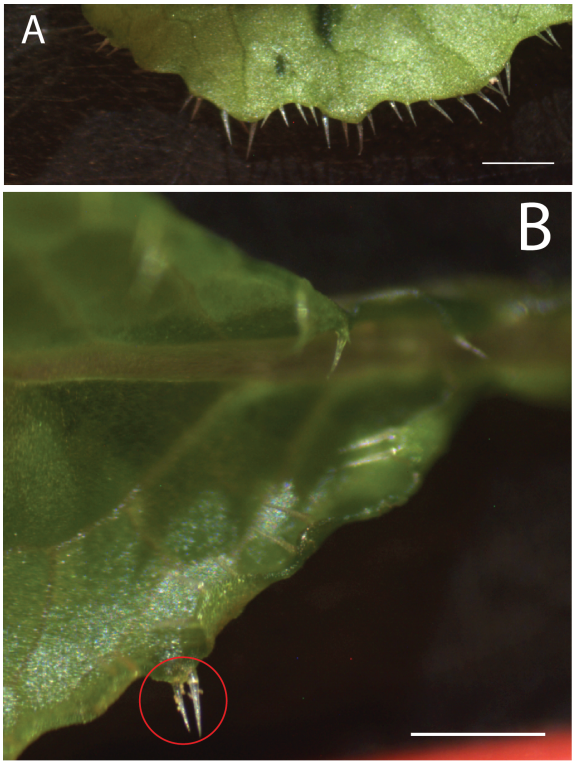

Figure 4.35: Leaf trichome phenotype of *brcpc2-1* plants. (A) The leaves of *brcpc2-1* plants showed an increased number of trichomes on the leaf margin. (B) Occasionally a trichome cluster (circled in red) formed on the leaf margin. Scale bars: 1 mm.

4.2.11 Analysis of *brtry* and *brcpc* Double Mutants

To test for functional redundancy of the trichome repressor genes, I generated double-mutant combinations.

4.2.11.1 Trichome Phenotype of *brtry1-1 brtry2-1* Mutant Plants

When compared to leaves of wild-type C634 plants, or to the leaves of *brtry1-1* and *brtry2-1* single mutants, *brtry1-1 brtry2-1* double-mutant plants exhibited a marked increase in trichome numbers on the leaf margins as well as the abaxial and adaxial surfaces of the leaf blade (Figure 4.36), suggesting that the two TRY homologs act

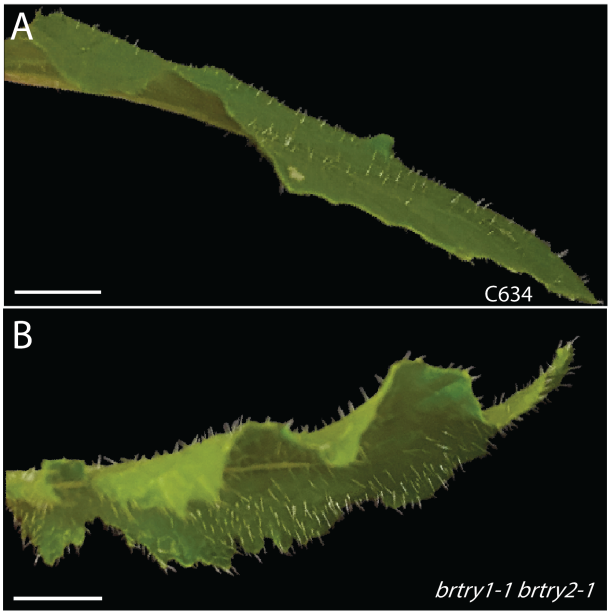

Figure 4.36: Leaf trichome phenotype of *brtry1-1 brtry2-1*. (A) Leaf of a wild-type C634 plant. (B) Leaf of a *brtry1-1 brtry2-1* double-mutant plants. Note the increase in the number of trichomes on both the abaxial and adaxial surfaces of the leaf. Scale bars: 5 mm.

---

redundantly in the control of trichome initiation on leaves. In addition, trichome formation on stems and petioles was not affected in the double mutant with the phenotype mimicking that of the *brtry2-1* single mutant (Section 4.2.10.2).

4.2.11.2 Trichome Phenotype of *brcpc1-1 brcpc2-1* Mutant Plants

Leaves of *brcpc1-1 brcpc2-1* double-mutant plants exhibited a marked increase in the number of trichomes on leaf margins (Figure 4.37A) when compared to the *brcpc1-1* and *brcpc2-1* single mutants or the wild type.

Figure 4.37: Leaf phenotype of *cpc1-1 cpc2-1* double-mutant plant. (A) A leaf of a *brcpc1-1 brcpc2-1* plant. (B) A trichome cluster (marked by a red circle) on the leaf margin of *brcpc1-1 brcpc2-1* plant. Scale bars: 1 mm.

As described above for the *brcpc2-1* single mutant, small trichome clusters occasionally formed on the leaf margins of the double mutant but their frequency appeared to be as low as in the *brcpc2-1* single mutant. There was no increase in trichome numbers observed on the stems of these plants or on the leaf surface.

4.3 Discussion

4.3.1 Identification and Characterisation of *TRY* and *CPC* homologs in *B. rapa*

The aim of this chapter was to identify the homologs of *CPC* and *TRY* in *B. rapa* and identify mutant alleles that could be used in order to test our main working hypothesis - that the mechanism of trichome suppression in flowers is conserved between *Arabidopsis* and *B. rapa*. During the course of this work, I identified two *B. rapa* homologs of *CPC* - *BrCPC1* (*Bra004539*) and *BrCPC2* (*Bra039283*), and three *B. rapa* homologs of *TRY* - *BrTRY1* (*Bra022637*), *BrTRY2* (*Bra026297*) and *BrTRY3* (*Bra029089*). All five of these trichome repressor genes have been characterised in terms of their expression patterns and their ability to fulfil the function of their *Arabidopsis* homologs via a series of complementation experiments.

The RT-qPCR experiments conducted to examine the expression of *BrTRY1*, *BrTRY2* and *BrTRY3* revealed that subfunctionalization had occurred with *BrTRY1* expressed at a higher level than the other two genes in leaf tissue and *BrTRY2* being the only *TRY* homolog whose expression was detected in mature gynoecium. The expression levels of *BrTRY3* were consistently low across all tissue sampled and it was concluded that it was possible that *BrTRY3* could have become a pseudogene. The results of the RT-qPCR experiments suggested that *BrCPC1* and *BrCPC2* are functional given that their expression is detected across all tissue samples used in the experiments with the expression of *BrCPC2* detected at a higher level in leaf tissue compared to *BrCPC1*. *BrCPC2* is also expressed more strongly in mature gynoecium tissue in comparison to *BrCPC1*. The results from Section 4.2.2 indicated that for the purposes of generating trichomes on floral organs, the important genes to obtain strong mutant lines for were BrTRY2 and *BrCPC2*.

The 3'UTRs of the five trichome repressor genes were characterised using 3'RACE-PCR and compared to their *Arabidopsis* genes in an effort to determine the cause of the differential expression patterns found in Section 4.2.2. After examining the percentage identity and alignment of the sequences there was no clear indication as to what may be causing the differences in expression. Subsequently, the sequences were analysed using the MEME suite SEA tool to check for common binding motifs in 3'UTRs that are associated with post-translational modification and processing. There were some differences in the ARE sites detected between the 3'UTR of *AtTRY* and *AtCPC* and their respective homologs but without further experimental evidence, it is difficult to definitively conclude that the differences in expression are solely due to these differences.

Given the results from the expression analysis presented in Section 4.2.2, strongly suggesting that *BrTRY2* is the sole TRY homolog being expressed in the gynoecium, JASPAR software was employed to examine potential differences in the predicted binding sites for SPL and MADS-domain transcription factor binding sites which are known to be important for the proper expression of *AtTRY* in *Arabidopsis*. The promoter sequence and intronic sequences of the *TRY* homologs were analysed. For the predicted MADS-domain transcription factor binding sites for the promoter regions, *BrTRY3* and *AtTRY* had a similar number of sites predicted. The predicted SPL transcription factor binding sites for the promoter sequences revealed that there were far fewer sites predicted for *BrTRY1*, *BrTRY2* and *BrTRY3* than *AtTRY*. *BrTRY3* was the most similar to *AtTRY* in this regard.

For intronic sequences, *AtTRY* had the highest number (11) of SPL transcription factor binding sites at the default relative profile threshold. *BrTRY1* and *BrTRY2* - 9 and 8 sites respectively and no sites were detected *BrTRY3*. It was concluded that the absence of SPL transcription factor binding sites could explain the low levels of expression detected for *BrTRY3* in Section 4.2.2. The MADS-domain transcription factor binding site analysis for the intronic sequences predicted similar numbers for AtTRY and *BrTRY2*

- 5 and 4 sites respectively. It could be the retention of MADS-domain transcription factor binding sites in the *BrTRY2* intron that is contributing to its high expression level in floral and gynoecium tissue.

The results of the complementation experiments showed that *BrTRY1* and *BrTRY2* were able to functionally complement *AtTRY* in a *try-29760* background. When expressed under the control of a CaMV 35S promoter, both *BrTRY1* and *BrTRY2* overexpression resulted in a glabrous leaf phenotype, exactly as observed for p35S:*AtTRY* plants. When placed under the control of the *AtTRY* promoter, both *BrTRY1* and *BrTRY2* could rescue the trichome phenotype of *try-29760* plants. In the case of *BrTRY3*, only data for the expression under the control of the *AtTRY* promoter is available. It seems that *BrTRY3* can only weakly rescue the *try-29760* mutant likely because it is missing a portion of the normal *TRY* sequence as discussed in Section 4.2.1. More independent lines for this construct will need to be generated, in addition to lines for p35S:*BrTRY3*, to better understand the function of *BrTRY3* in the control of trichome initiation.

The *CPC* complementation experiments revealed that when expressed under the control of either the CaMV 35S promoter or the native *AtCPC* promoter, *BrCPC1* and *BrCPC2* are able to rescue the *cpc-1* phenotype. These results concur with the high level of sequence preservation between the proteins (Section 4.2.1) and the expression data showing broad expression in the plant (Section 4.2.2).

4.3.2 Characterisation of BrTRYs and BrCPCs by Examining Mutant Alleles

In order to further test the role of *BrTRY1-3* and *BrCPC1-2* in trichome initiation, mutant lines for these genes were obtained from RevGenUK. The isolation of mutant alleles would also allow testing of whether the floral trichome phenotypes described in *Arabidopsis* (Section 1.3) can be re-created in *B. rapa*. The phenotypic analysis of sin-

gle mutants is detailed in Section 4.2.10 and confirms the basis of selection for alleles as detailed in Section 4.2.9. All mutant alleles chosen for the trichome repressor genes were selected for their suspected potential to eliminate the function of the gene and each of the mutant lines examined during this work showed leaf trichome phenotypes suggesting a loss of function of trichome repressor genes.

The trichome phenotypes of *brtry1-1* are shown in Figure 4.31. The leaves produced by *brtry1-1* plants develop leaves with an increased number of trichomes present on the leaf veins. The expression data in Section 4.2.2 show that *BrTRY1* is expressed in leaf tissue and the results of the single mutant phenotype suggest that *BrTRY1* does indeed have a role in repressing trichome formation given the increased numbers in the mutant. Interestingly, the increase is greater on the abaxial side of the leaf. It has been shown that SPL transcription factors have a role to play in regulating trichome formation in both a time-related and polarity-related manner (Yu et al., 2010; Xue et al., 2014; Cui et al., 2022). There are differences in the SPL transcription factor predicted binding sites for *BrTRY1* compared to *ATRY*, *BrTRY2* and *BrTRY3* as shown in Sections 4.2.5 and 4.2.6. *BrTRY1* lacks any predicted SPL transcription factor binding sites in the promoter regions compared to the other *BrTRY* genes. There is no way to conclude that the absence of these binding sites is the cause of the trichome phenotype without experiments modulating the existing sites in order to examine the effects of removing predicted SPL transcription factor binding sites.

The leaves of *brtry2-1* plants show no increase in trichomes on the leaves compared to C634 which concurs with the expression data from Section 4.2.2 where *BrTRY2* expression is very low in leaf tissue. The *brtry2-1* plants show an increased number of trichomes on both the stems and petioles as shown in Figure 4.32 and Figure 4.33. Stem and petiole tissue were not tissues that were included in the experiments attempting to characterise the expression domains of the trichome repressors but from the mutant phenotype, it is clear that *BrTRY2* functions in repressing trichome formation on

the stems and petioles. Given that this phenotype is the same in the *brtry1-1 brtry2-1* double-mutants, it appears that subfunctionalization may have occurred and *BrTRY2* alone functions in repressing the formation of trichomes on the stem and petioles of *B. rapa*. *BrTRY3* may be involved in repressing trichomes in these tissues but given the unavailability of a suitable mutant line (discussed in Section 4.2.9), it is not possible at this time to investigate this possibility.

The *brtry1-1 brtry2-1* show an increase in the number of trichomes developing on both the adaxial and abaxial surface of the leaves compared to the wild-type C634 as shown in Figure 4.36. The phenotype suggests that *BrTRY1* and *BrTRY2* function partially redundantly in repressing trichome initiation on leaves. The expression of *BrTRY2* was very low in the leaf tissue in expression data in Section 4.2.2. However, the expression data for C634 leaf tissue in Section 4.2.3 shows that expression levels of *BrTRY2* are higher in this accession than the R-o-18 line of *B. rapa*. As detailed in Section 1.2, the expression of trichome repressors such as *TRY* are activated by the formation of a trichome initiation complex. As C634 *B. rapa* plants form many more trichomes than R-o-18, perhaps there is a stronger transcriptional activation signal in this background resulting in the higher expression of *BrTRY2*. It appears that *BrTRY1* acts as the main trichome repressor in leaf tissue and is able to fulfil this function even in the absence of *BrTRY2* as indicated by the lack of leaf trichome phenotype seen in *brtry2-1* plants. In *brtry1-1* mutant plants, *BrTRY2* appears to be partially functionally redundant as the trichome number increases even further in the *brtry1-1 brtry2-1* double mutants.

The trichome phenotype of *brcpc1-1* mutants (Figure 4.34) showed only a slight increase ($\approx 10\%$) in the trichomes that formed on the leaf margin. A dramatic phenotype was not expected as *BrCPC1* was expressed at a lower level to *BrCPC2* in leaf tissue (Section 4.2.2). This phenotype does indicate that *BrCPC1* is functional as a trichome repressor in leaves despite being expressed at a lower level. The mutant phenotype for *brcpc2-1* plants (Figure 4.35) showed an increased number of trichomes developing on

the leaf margin but at a higher frequency than *brcpc1-1* mutant plants. Occasionally, trichome clusters would occur on the leaf margin of *brcpc2-1* plants (Figure 4.35B). This mutant phenotype was never observed in any of the *brtry* single or double mutants plants or in *brcpc1-1* mutant plants. The infrequent trichome cluster was observed in *brcpc1-1 brcpc2-1* double mutants but at the same frequency as *brcpc2-1* plants alone. There was an increased amount of trichomes observed on the leaf margin in the *brcpc1-1 brcpc2-1* double-mutant compared to wild-type and either of the *brcpc* single mutants. There was no increase in the number of trichomes forming on either the stems or petioles of *brcpc1-1*, *brcpc2-1* or *brcpc1-1 brcpc2-1* plants suggesting that *BrCPC1* and *BrCPC2* do not have a role in repressing trichome formation in these tissues.

Generally, the mutant trichome phenotypes broadly agree with the expression data generated for *BrTRY1-3* and *BrCPC1-2* in Sections 4.2.2 and 4.2.3 with a broad increase in the number of trichomes developing in the single and double mutants of the trichome repressor genes. The most interesting result was the trichome cluster phenotype seen in *brcpc2-1* and *brcpc1-1 brcpc2-1* plants. In *Arabidopsis*, *cpc* mutant plants show an increase in trichome numbers on leaves and stem with no clusters observed (Schellmann et al., 2002). *Try-29670* mutants in *Arabidopsis* show trichome clusters with a $\approx$ 5-10% frequency (Schellmann et al., 2002). In *B. rapa*, trichome clusters occur only in *brcpc2* mutants suggesting that perhaps the mechanism inhibiting clustering is controlled by *CPC* in *B. rapa* and not *TRY* like in *Arabidopsis*.

4.4 Future Work

Further work is necessary to examine the level of redundancy that exists between the *B. rapa* homologs of *BrCPC* and *BrTRY* and their roles in repressing trichome formation on leaves and stems. This will be achieved by generating *brcpc brtry* mutants in various combinations in order to fully assess the function that each has in various tissues and the level of functional redundancy that exists between these trichome repressors in *B. rapa*. In *Arabidopsis*, *cpc try* double mutants show a dramatic increase in trichome clusters with $\approx 75\%$ of all trichome initiation sites containing between 20 and 30 clusters (Schellmann et al., 2002).

The *B. rapa* higher-order mutants will be generated and examined to see if a combination of *brtry* and *brcpc* mutant alleles can replicate the dramatic clustering seen in *Arabidopsis*. Modified complementation experiments where the mutant CDS of *brcpc1-1*, *brcpc2-1*, *brtry1-1* and *brtry2-1* are used to attempt a rescue of a suitable *Arabidopsis* mutant could be carried out in order to further characterise the impact of these mutations on gene function. The coding DNA sequence from each of these mutant lines and clone each of them under the control of either pTRY or pCPC (whichever one is applicable). The constructs would then be used to transform either *try* or *cpc Arabidopsis* and transformants examined for any rescue of the mutant phenotype. The *brtry1-1* mutant line results in an amino acid substitution (Figure 4.29) and experiments such as these would allow for a clearer idea of the effect that this mutation is having on the function of the *BrTRY1* protein. The *brcpc1-1* mutant line contains an early stop codon but it occurs in the third exon (Figure 4.30) which may result in an almost fully functional protein still being active in the plant and may explain the mild mutant phenotype seen in Figure 4.34.

In *Arabidopsis*, trichome repressors genes and specifically *AtCPC* play a role in promoting the development of root hairs with *cpc* mutants showing only a few normally shaped root hairs with much fewer hairs forming and those that did form often appearing

stunted (Wada et al., 1997). As the overall aim of this project is to potentially generate *B. rapa* with trichomes on floral organs for the purpose of protection against pests, it is essential that root formation and the fitness of the plants are not affected negatively. The root-hair trichome phenotype of *B. rapa BrTRYs BrCPCs* mutants as well as *brcpc brtry* higher-order mutants will be monitored by growing mutant seedlings on MS agar plates and comparing the root hair phenotype to both C634 and R-o-18 wild-type plants.

For the purpose of generating plants with trichomes on the gynoecia, the mutant alleles that seem key to work with are *brtry2-1* in the case of *TRY* homologs and *brcpc1-1* in the case of *CPC* homologs in *B. rapa*.

Chapter 5 - Investigating Trichome Phenotype Variation in Two Lines of *Brassica rapa* & Generating Plants With Trichomes on Floral Organs

5.1 Introduction

The trichome phenotype of *Arabidopsis* is discussed in Section 1.2 and overall, there is little natural variation in the trichome phenotype observed across accessions of *Arabidopsis* with the exception being those found across the Iberian peninsula such as Don-0 which develops flowers with trichomes present on the gynoecium and subsequent silique (Arteaga et al., 2021). These floral trichomes are the result of mutations in regulatory regions of trichomes repressions such as *TRY* or *TCL1* which result in lower levels of expression as well as a structural mutation in GL1 leading to increased protein activity. The many accessions of *B. rapa* exhibit phenotypic variation in their trichome phenotypes with three main phenotypic categories observed across hundreds of *B. rapa* accessions - "hairy-leaf", "leaf-margin" and "glabrous". There have been no reported *B. rapa* accessions with naturally occurring trichomes on the gynoecia of flowers. Accessions belonging to the "hairy-leaf" group develop trichomes all over the leaf surface, those belonging to the "leaf-margin" group produce trichomes on the leaf margin only with the leaf blade lacking trichomes and those belonging to the "glabrous" group com-

pletely lack trichomes on their leaves.

Work described in Zhang et al. (2018) aimed to identify the genetic cause of the diverse leaf trichome morphology seen across *B. rapa* accessions. They were able to identify *Bra025311* - one of two *B. rapa* homologs of the *Arabidopsis* gene *GLABRA1* - via quantitative trait locus (QTL) analysis as a gene associated with the different leaf trichome phenotypes. They showed by examining over 200 accessions of *B. rapa* that there was a high level of naturally occurring variation at this locus but only one functional SNP was identified with the presence of thymine at position 274 of the coding DNA sequence associated with the "hairy-leaf" phenotype while cytosine at that position associated with the "leaf-margin" phenotype. However, there were no other functional SNPs identified and no explanation of the cause of the "glabrous" mechanism in *B. rapa*. As detailed in Section 1.5.2, the attempt to genetically analyse the cause of trichome variation in *B. rapa* is complicated by the fact there are several homologs for many of the genes involved in trichomes (listed in Table 5.1) with no previously described experiments characterising if all of these homologs are functional in the process of trichome formation.

Table 5.1: *B. rapa* homologs of *Arabidopsis* genes involved in trichome formation. The sequence of genes that are involved in the initiation of trichomes in *Arabidopsis* *GL1*, *GL2*, *GL3*, and *TTG1* were used to perform BLAST searches on the *B. rapa* genome in order to identify likely homologs. The genes, the corresponding TAIR IDs, and their homologs in *B. rapa* with the corresponding BRAD IDs listed.

Gene	TAIR ID		BRAD ID	Abbreviation
GL1	AT3G29760		Bra025311	BrGL1a
			Bra039065	BrGL1b
GL2	AT1G79840		Bra003535	BrGL2
GL3	AT5G41315		Bra025508	BrGL3
TTG1	AT5G24520		Bra09770	BrTTG1a
			Bra029410	BrTTG1b
			Bra029411	BrTTG1c
TTG2	AT2G37260		Bra023112	BrTTG2a
			Bra005210	BrTTG2b

5.2 Aim of the Chapter

As described in Section 1.2, the mechanism of trichome initiation is well understood in *Arabidopsis*. In this chapter, the aim was to investigate the cause of the almost completely glabrous phenotype of the *B. rapa* R-o-18 line and to test whether *B. rapa* plants could be generated that form trichomes on seed pods. The cause behind the different leaf trichome phenotypes in *B. rapa* was examined and the genetic cause of this naturally occurring variation was investigated. Candidate genes were identified in the literature and sequenced in the R-o-18 line which develops trichomes on the leaf margin only and the C634 accession which develops trichomes all over the leaf surface. The two lines were crossed and the progeny were examined for their trichome phenotype. The cause of the residual leaf margin phenotype observed in R-o-18 was investigated and *B. rapa* mutant lines that were completely glabrous obtained and used to determine how the function of trichome formation is split between the homologs of *GL1* in *B. rapa*. Finally, the higher-order mutants of the *B. rapa* homologs of *AG*, *CPC* and *TRY* that were generated in the course of this work are discussed.

5.3 Results

5.3.1 Examining the Genetic Mechanism Behind the Phenotypic Difference between R-o-18 and C634

All of the TILLING lines used for this work detailed in Sections 3.2.7 and 4.2.9, were generated in an inbred line of *B. rapa* subsp. trilocularis yellow sarson termed R-o-18 (Stephenson et al., 2010). An examination of R-o-18 plants showed that leaves of this inbred line form only a few trichomes on their margins while leaf blades are completely glabrous. This suggested that the mechanism for trichome initiation is partially blocked or reduced in its activity in this line. Given the overall aim of my project, i.e. the induction of supernumerary and/or floral trichomes, this result implied that the TILLING mutations would have to be introgressed into a different background where trichome initiation occurs normally. We, therefore, obtained seeds of a *B. rapa* yellow sarson line, termed C634, that had been described in the literature as having a "hairy-leaf" phenotype from the Tohoku University Brassica Seed Bank in Japan (Li et al., 2009).

The phenotypes of R-o-18 and C634 are compared in Figure 5.1. The mature flowers of R-o-18 and C634 are phenotypically very similar with the only difference observed in the pistil length and the ovary of the gynoecium - those of C634 flowers are slightly slimmer and the styles are slightly elongated however, the fertility of the plant is not affected (Figure 5.1A-B). The leaf trichome phenotype of R-o-18 and C634 show a clear difference - a typical leaf of R-o-18 develops trichomes sparsely on the leaf margin only and the blade is completely glabrous while a typical leaf of C634 develops trichomes on the leaf margin and leaf blade (Figure 5.1 C-D). Figure 5.1E-F shows siliques from R-o-18 and C634. The difference in silique phenotypes was attributed to the differences in the gynoecium of the two accessions. The style and ovary of the C634 siliques are longer compared to R-o-18.

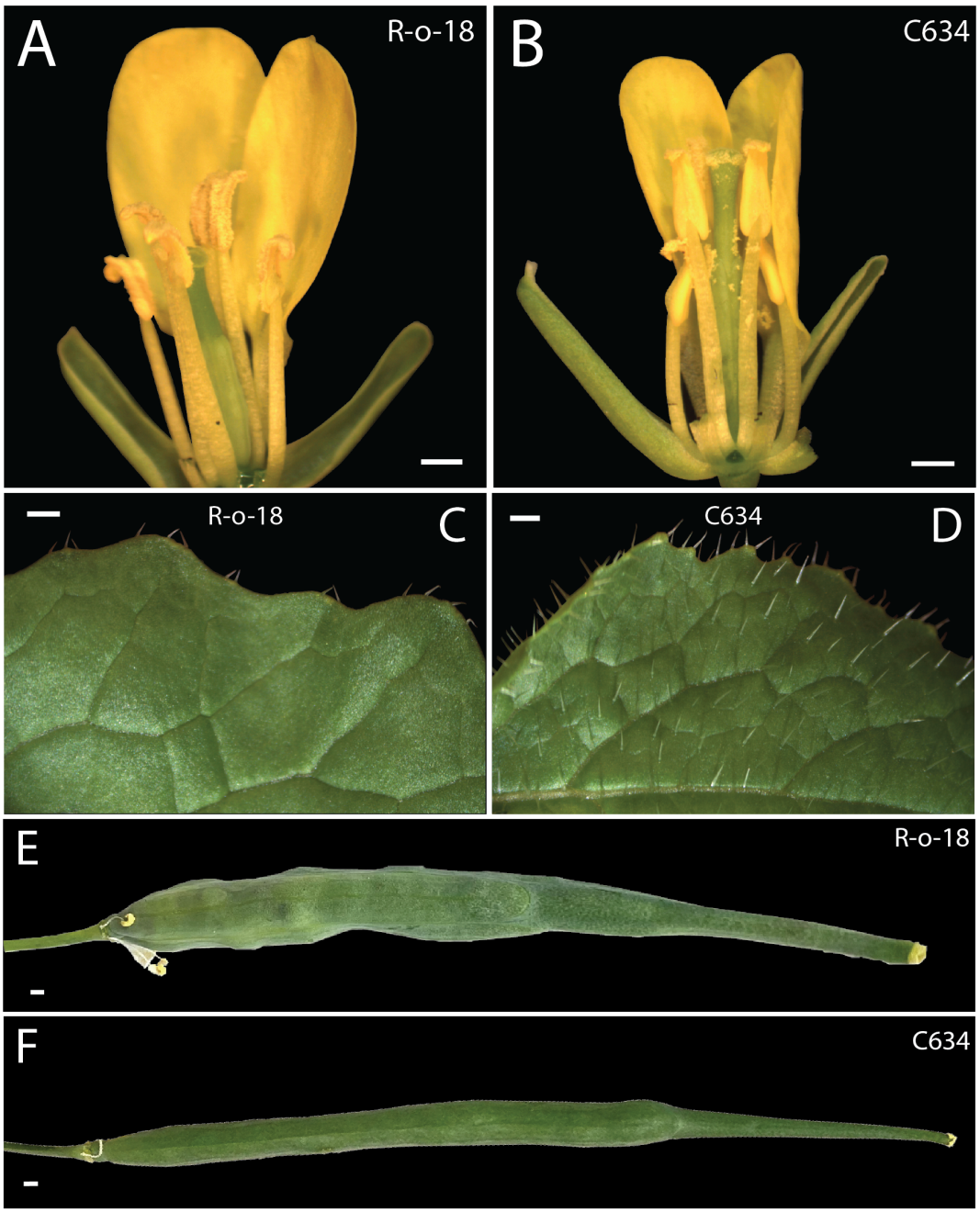

Figure 5.1: Phenotypic differences between R-o-18 line and C634 accession of *B. rapa*. (A) Flower of an R-o-18 plant. (B) Flower of a C634 plant. The flowers of the two lines are nearly identical with the only phenotypic difference being that the gynoecium of C634 is slightly narrower and more elongated compared to R-o-18. (C) A typical leaf of an R-o-18 plant. Trichomes form sparsely on the leaf margins and never on the leaf blade. (D) A typical leaf of a C634 plant with trichomes forming on both the leaf margin and leaf blade. (E) Mature silique of an R-o-18 plant. (F) Silique from a C634 plant. The ovary of the R-o-18 silique is wider and the surface is bumpier than that of the C634 silique. The style of the C634 siliques is longer and thinner than that of R-o-18. Scale bars: 1 mm.

Zhang et al. (2018) had previously mapped the naturally occurring trichome variation in *B. rapa* to the *GL1* homolog *Bra025311* which was named *BrGL1a*. To test whether *BrGL1a* was also the cause of the difference in trichome formation in R-o-18 and C634, I PCR-amplified its genomic region in both lines and analysed the products by agarose gel electrophoresis. Based on the sequence of *BrGL1a* from the *B. rapa* Chiifu-401-42 reference genome, the expected size of the amplicon was approximately 1.8 kb. The results of the PCR reactions I carried out are displayed in Figure 5.2.

Figure 5.2: Results of *BrGL1a* amplification in R-o-18 and C634. Results of PCR which aimed to amplify *BrGL1a* using genomic R-o-18 and C634 DNA as the template. The expected amplicon size was approximately 1.8 kb. Lane 1 contained the R-o-18 amplicon which appears to be approximately 2 kb in length. Lane 2 contained the C634 amplicon which appears to be approximately 1.8 kb in length. Lane L contained a 1 kb size marker (ThermoFisher). The bands corresponding to 1, 1.5 and 2 kb are labelled.

There was a clear difference in the size of the *BrGL1a* amplicon of R-o-18 and C634. While the amplicon for C634 was approximately 1.8 kb and thus of the expected size, the PCR fragment from R-o-18 genomic DNA was markedly larger, around 2 kb. The PCR products were sequenced and sequences were aligned to that of *BrGL1a* from the Chiifu-401-42 reference genome (Figure 5.3).

The amplicon from C634 was found to be 1,734 bp in length and aligned almost

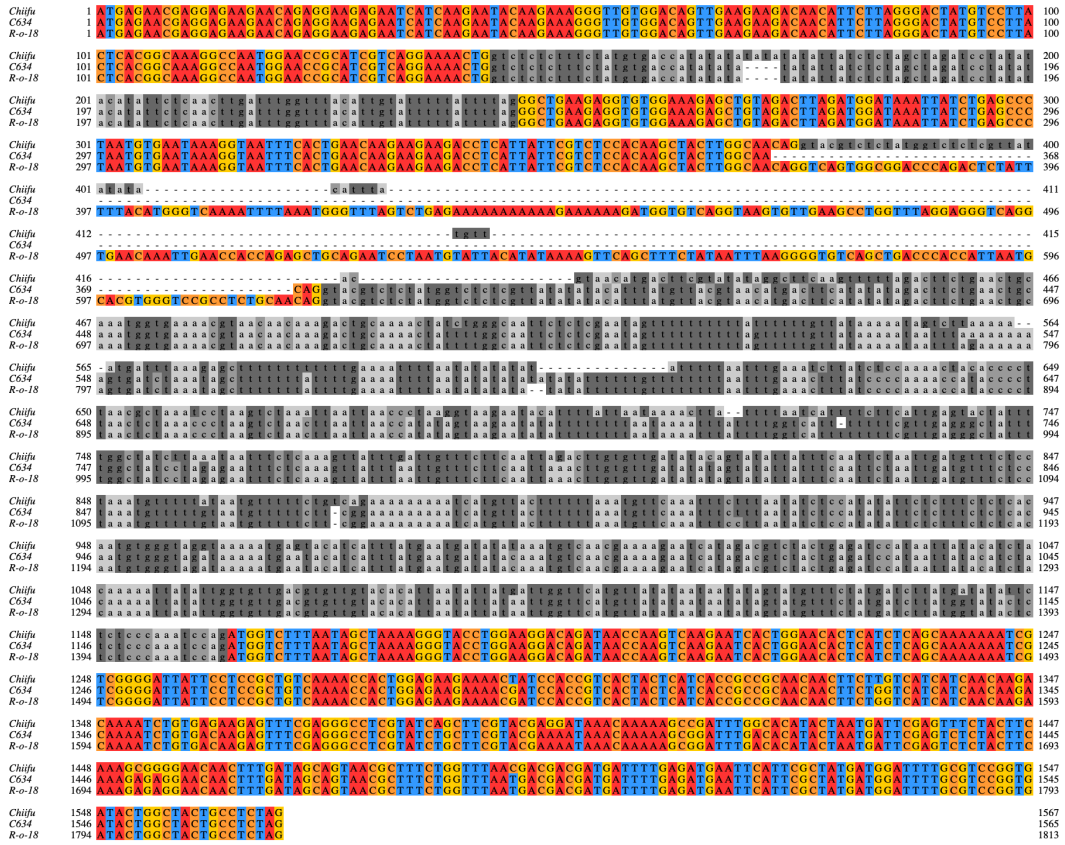

Figure 5.3: Alignment of sequences obtained for *BrGL1a* by PCR amplification from R-o-18 and C634, respectively, to *BrGL1a* from the Chiifu reference genome. Alignment of the sequence of *BrGL1a* from the reference genome (Chiifu) to the sequences obtained via PCR from R-o-18 and C634. Note the insertion of a 249 bp fragment in the second exon of *BrGL1a* in the case of R-o-18. Results were visualised using *Jalview* software. The coding DNA sequences are in upper case and in colour whilst intronic sequences are in lower case and in grey.

---

perfectly with the *BrGL1a* gene sequence Wang et al., 2011. In the case of R-o-18, the amplicon was 1,989 bp in length and from the alignment, there is an insertion of 249 bp in the second exon. To find out where the sequence of the 249 bp insertion found in *BrGL1a* in R-o-18 may have originated, I performed a BLAST search using the program MOLE-BLAST. MOLE-BLAST is a software that produces a list of the closest database neighbours for the sequences submitted by users and generates a phylogenetic tree. There were four databases identified by the search: three from the *B. napus* genome assembly (chromosomes A03, A04 and A08) and two from the *B. rapa* genome assembly (chromosome A04). The sequence of the insert was then used to conduct a BLAST search of the *B. napus* and *B. rapa* genomes. As expected, the top hit

was *BrGL1a* itself. Additional hits are summarised in Table 5.2. From the results of the analysis, it appears that the insert originated from the *B. rapa* genome itself and is likely the consequence of a recombination event that possibly occurred during the inbreeding that led to the generation of the R-o-18 line.

Table 5.2: BLAST search results for the insertion in *BrGL1a* of R-o-18. The insertion found in R-o-18 *BrGL1a* was used to conduct a BLAST search across the *B. rapa* genome to determine if the sequence originated from the *B. rapa* genome itself. The genome location of the search hit, any overlapping genes at that location, the length of the sequence that overlaps in bp, the E-value and the percentage of identity between the insertion and the result sequence are shown.

Genomic Location	Gene(s)	Length	E-val	% ID
A06:26061228-26061476	A06p049300.1	249	1.50E-138	100.0
A05:812250-812490		241	8.70E-134	100.0
A04:22210766-22211007	A04p039340.1	242	1.30E-129	99.2
A08:10316770-10317010	A08g505990.1	241	5.20E-129	99.2
A09:9207037-9207215		179	7.40E-85	97.2
A03:15286914-15287090		177	1.20E-83	97.2
A02:25721377-25721503		127	1.70E-39	91.3
A03:15286852-15286924		73	3.80E-31	98.6
A09:9206973-9207024		52	5.30E-21	100.0
A02:25721263-25721312		50	4.90E-15	96.0
A09:2282112-2282163		52	7.70E-14	94.2
A09:8006397-8006448	A09p015390.1	52	7.70E-14	94.2
A09:9777361-9777412		52	7.70E-14	94.2
A02:12217805-12217856	A02p025510.1	52	7.70E-14	94.2
A09:3632360-3632410	A09p006930.1	51	3.00E-13	94.1

The coding DNA sequence of BrGL1a was translated in silico and the protein sequences were aligned and visualised (Figure 5.4). *BrGL1a* from C634 is 225 amino acids in length and is $\approx$ 94% identical to the protein encoded by the gene in the reference genome. In contrast, *BrGL1a* from R-o-18 is predicted to be 124 amino acids in length as the insertion in the second intron introduces an early stop codon. In addition to being shorter in length, the C-terminal end of *BrGL1a* from R-o-18 would contain missense mutations that among others eliminate a tryptophan at position 92, which has been shown to be critical for the interaction of *BrGL1a* with *Arabidopsis GL3* in yeast two-hybrid and bimolecular fluorescence complementation assays (W. Zhang et al., 2018).

From these results, it was concluded that *BrGL1a* in C634 is fully functional while in R-o-18 the absence of almost half of the protein sequence and of an essential amino acid residue in all probability leads to a complete loss of function.

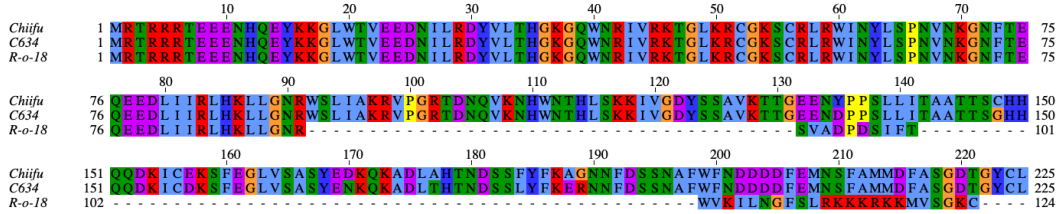

Figure 5.4: Alignment of the Amino Acid sequence of *BrGL1a* from R-o-18 and C634 to Chiifu sequence. Alignment of the three protein sequences shows that *BrGL1a* from C634 is almost identical to *BrGL1a* from Chiifu and both possess the functionally critical tryptophan as identified in Zhang et al. (2018) at position 92. The protein of *BrGL1a* from R-o-18 contains an early stop codon produced as a result of the 249bp insertion in the second exon. The protein is 124 amino acids in length and does not contain a tryptophan at position 92. Results were visualised using *Jalview* software.

As shown in Table 5.1, in addition to *BrGL1a*, there is another homolog of *AtGL1* in *B. rapa*, termed *BrGL1b*. To find out whether there was any sequence variation for this gene between R-o-18 and C634, I PCR-amplified *BrGL1b* from genomic DNA. An alignment of the genomic sequence (Figure 5.5) showed that the exon sequences of *BrGL1b* are identical in the two lines, and the only difference detected is a 2 bp deletion in the first intron of the gene in C634. Thus, *BrGL1b* appears to be unaffected in both lines and is thus unlikely to contribute to the observed differences in trichome initiation.

A question that remained unanswered was whether the trichome margin phenotype

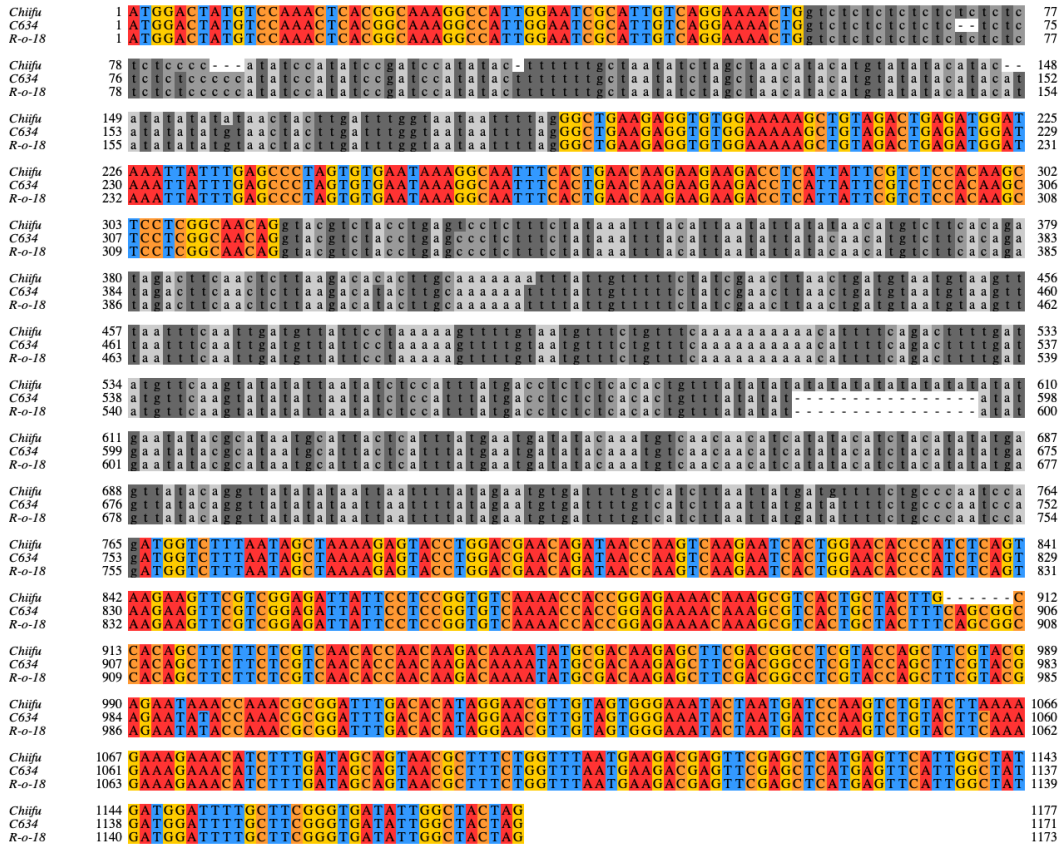

Figure 5.5: Alignment of sequences obtained for *BrGL1b* by PCR amplification from R-o-18 and C634 to *BrGL1b* from the Chiifu reference genome. Alignment of the sequence of *BrGL1b* from the reference genome (Chiifu) to the sequences obtained via PCR from R-o-18 and C634. Results were visualised using *Jalview* software. The coding DNA sequences are in upper case and in colour whilst intronic sequences are in lower case and in grey.

---

of R-o-18 was due to some residual function of the truncated *BrGL1a* protein or whether it was a result of *BrGL1b* activity. This question was addressed through experiments outlined in Section 5.3.3.

5.3.2 Generation of Hybrids between R-o-18 and C634

To determine whether the variation at the *BrGL1a* locus in R-o-18 and C634 was the cause of the difference in the trichome phenotype of the two lines, I carried out reciprocal crosses between C634 and R-o-18 plants. Seeds of the F1 generation of both crosses were planted and examined for the trichome leaf phenotype. All F1 plants from both

crosses displayed the "hairy-leaf" phenotype as seen in the C634 accession (Table 5.3 and Figure 5.6), indicating that it was a dominant trait.

Table 5.3: Phenotypic analysis of the F1 progeny of crosses between C634 and R-o-18. Reciprocal crosses were performed. The line used as a maternal plant in a cross is indicated and the number of plants examined (n) is indicated. All F1 plants displayed a "hairy-leaf" phenotype similar to that seen in the C634 accession.

	n	Hairy-Leaf	Leaf-Margin
R-o-18 maternal	25	25	0
C634 maternal	15	15	0

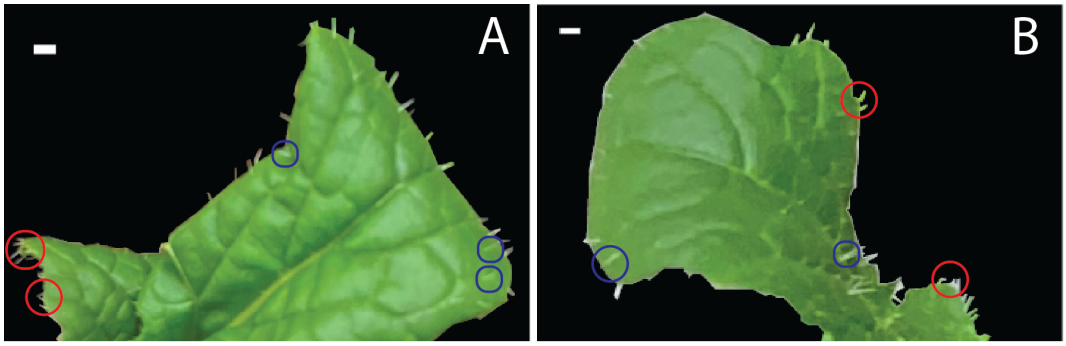

Figure 5.6: Leaf trichome phenotype of the F1 progeny of C634 X R-o-18 crosses. (A) F1 plant from a cross using C634 as a female. (B) F1 plant from a cross using R-o-18 as a female. Red circles indicate trichomes forming on the leaf margin and blue circles indicate trichomes forming on the leaf blade. Scale bars: 1 mm.

PCR-based genotyping confirmed that both *BrGL1a* alleles from C634 and R-o-18, respectively, were present in those plants. The F1 plants were then selfed and F2 progeny was planted and genotyped to determine which *BrGL1a* locus they contained, and examined for their leaf trichome phenotype (Table 5.4).

I found that plants homozygous for *BrGL1a* from R-o-18 displayed the "leaf-margin"

Table 5.4: Phenotypic analysis of F2 progeny of C634 R-o-18 crosses. Phenotypic analysis of the F2 plants of R-o-18 C634 reciprocal crosses. Of the 32 F2 plants examined, 25 displayed the "hairy leaf" phenotype and 7 displayed the "leaf-margin" phenotype.

	n	Hairy-Leaf	Leaf-Margin
F2 C634 x R-o-18	32	25	7

phenotype while plants heterozygous or homozygous for *BrGL1a* from C634 had the "hairy leaf" phenotype. Figure 5.7A shows the leaf trichome phenotype that plants homozygous for R-o-18 *BrGL1a* displayed - "leaf-margin". Figure 5.7B-C show the leaf trichome phenotype of plants heterozygous and homozygous for C634 *BrGL1a* – "hairy leaf".

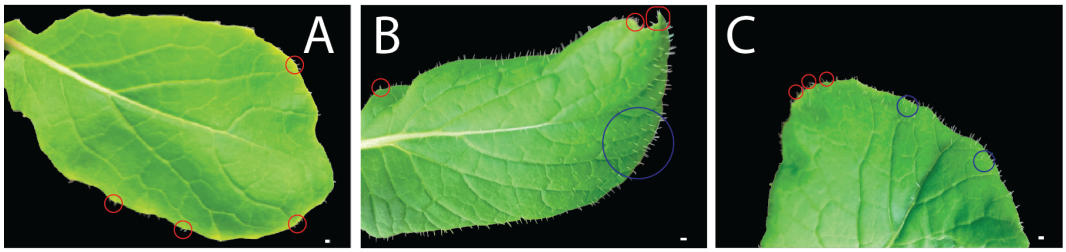

Figure 5.7: Leaf trichome phenotype of F2 progeny of C634 R-o-18 crosses. The leaf trichome phenotypes observed in the F2 progeny of C634 R-o-18 cross. (A) F2 plant genotyped and homozygous for R-o-18 *BrGL1a*. The leaves of these plants formed trichomes on the leaf margin only. (B) F2 plant heterozygous for the two *BrGL1a* alleles. These plants formed leaves with trichomes on the leaf margins and the leaf blade. (C) F2 plant homozygous for the C634 *BrGL1a* allele. The leaves develop trichomes on the leaf margin and the leaf blade. The red circles indicate trichomes on the leaf margins and the blue circles indicate trichomes on the leaf blade. Scale bars: 1 mm.

For a trait acting recessively, a classic Mendelian segregation ratio of 3:1 would be expected in the F2 generation and this is what was observed for F2 plants of the C634 R-o-18 cross. All plants that contained *BrGL1a* from C634 developed leaves with the "hairy-leaf" phenotype. The ratio between the different phenotypic groups was approximately 3:1 in agreement with the idea that a single locus was responsible for the phenotypic differences.

5.3.3 Investigating the Role of *BrGL1b* in the Formation of Trichomes on the Leaves

As discussed in Section 5.3.1, *BrGL1a* in R-o-18 contains an insertion resulting in an early stop codon and likely to result in a complete loss of protein function. However, in contrast to *gl1* mutants in *Arabidopsis*, R-o-18 plants are not fully glabrous and form some trichomes on leaf margins (Oppenheimer et al., 1991). This phenotypic difference between R-o-18 and *gl1* in *Arabidopsis* could be due to residual activity of the *BrGL1a* gene. Alternatively, the formation of trichomes on leaf margins in the absence of *BrGL1a* function could be mediated by *BrGL1b*.

To better understand the function of *BrGL1b* in trichome formation and to assess the degree of functional redundancy between BrGL1a and *BrGL1b*, I obtained a TILLING line (JI33214-b) containing a mutation in *BrGL1b* from RevGenUK. In this line denoted *brgl1b-1*, the guanine at position 249 of the coding DNA has been mutated to an adenine causing the tryptophan at position 83 to become an early stop codon. Thus, it was expected that this mutation would lead to a loss of gene function. Because the TILLING mutant is in the R-o-18 background and hence contains a mutation in *BrGL1a*, I expected that *brgl1b-1* mutant plants would exhibit a reduction in the number of trichomes relative to R-o-18. Indeed, I observed that *brgl1-b* plants were completely glabrous (Figure 5.8). This result suggested that the combined activities of *BrGL1a* and *BrGL1b* are essential for trichome formation.

In order to further examine the role of *BrGL1b* in the formation of trichomes on the

Figure 5.8: Leaf phenotype of *brgl1b-1* mutant plants. A leaf from a brgl1b-1 plant. All leaves, petioles and stems that developed from brgl1-b plants were completely glabrous. Scale bar: 1 mm.

leaf margin, I crossed *brgl1b-1* with C634 plants. Plants of the F1 generation developed trichomes on the leaf blade and the leaf margin similar to the phenotype of the C634 accession (Table 5.5 and Figure 5.9). These plants develop trichomes on the leaf blade and the leaf margin similar to the phenotype of the C634 accession. From these results, it was concluded that the *brgl1-b* mutation was not acting dominantly in the C634 background.

Table 5.5: Phenotypic analysis of F1 progeny of C634 *brgl1b-1* crosses. Counts of the leaf trichome phenotypes of F1 plants from cross between C634 and *brgl1b-1* mutants. Hairy leaf, Leaf margin and glabrous describe the trichome phenotype that could potentially be observed. All F1 plants show the "hairy-leaf" phenotype.

	n	Hairy-Leaf	Leaf-Margin	Glabrous
***brgl1b-1* maternal**	4	4	0	0
C634 maternal	5	5	0	0

Figure 5.9: Leaf trichome phenotype of the F1 progeny of C634 *brgl1b-1* crosses. F1 plants developed leaves with trichomes on the leaf blade and the leaf margin similar to C634. Scale bar: 1 mm.

After selfing of the F1 plants, the seed of the F2 progeny was planted and genotyped to identify plants that were homozygous for *brgl1-b* and either heterozygous or homozygous for the C634 BrGL1a locus. These plants had trichomes on both the leaf margins and leaf blades (Figure 5.10). Thus, in the presence of at least one functional copy of *BrGL1a*, mutations in *BrGL1b* have little effect on trichome formation. At the same time, in the absence of functional *BrGL1a*, *BrGL1b* is able to drive minimal trichome formation on leaf margins only.

Figure 5.10: Leaf trichome phenotype of the F2 progeny of an C634 *brgl1b-1* crosses. F2 plants developed leaves with trichomes on the leaf blade and the leaf margin similar to C634. Scale bar: 1 mm.

5.3.4 Examining the Expression Patterns of Genes Involved in Trichome Formation in R-o-18, C634 and the Hybrid

Following the conclusion that *BrGL1a* variance was the probable cause of the different trichome phenotypes observed in R-o-18 and C634, I conducted a series of RT-qPCR experiments in order to further examine the impact that the insertion present in the R-o-18 allele has on the expression levels of genes related to trichome formation. The expression of genes listed in Table 5.1 was measured in whole leaf, leaf margin and leaf blade tissue in R-o-18, C634 and plants containing one *BrGL1a* allele from R-o-18 and C634 (referred to as 'hybrid" from here on). Four biologically independent sets of samples were collected for analysis. The leaf tissue used was collected from *B. rapa* R-o-18, C634 and Hybrid plants 10 days after germination. For the purpose of these experiments, the leaf margin was defined as 1-2 mm from the edge of the leaf, and the leaf blade was defined as leaf tissue excluding the margin and the petiole. In *Arabidopis*, expression levels of *TTG1* would be expected to be the highest, then *TTG2*, *GL2*, *GL3* at an intermediate level with *GL1* at the lowest level of expression (Winter et al., 2007).

The results of the RT-qPCR experiments for whole leaf tissue are shown in Figure 5.11. In the case of *BrGL2*, *BrGL3*, *BrTTG1c*, *BrTTG2a* and *BrTTG2b*, expression levels are higher in C634 than R-o-18. This makes sense considering how C634 develops trichomes on the whole leaf surface whilst R-o-18 forms trichomes only on the leaf margin and even then quite sporadically. Unexpectedly, the expression levels of many genes tested were found to be highest in the tissue from the hybrid. The exact reason for this is unclear but possible reasons are discussed in Section 5.4.

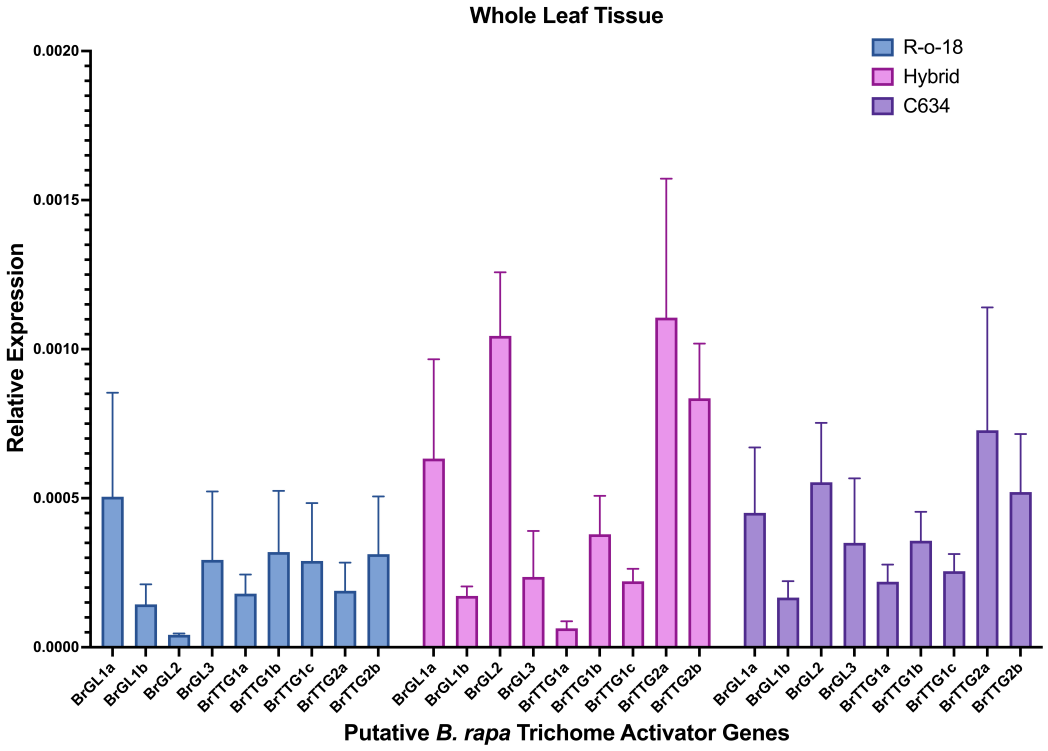

Figure 5.11: Expression of *B. rapa* homologs of genes involved in the control of trichome initiation in *Arabidopsis* in whole leaf tissue. The expression of genes in whole leaf tissue from R-o-18 (BrCF1), C634 (BrCF2), and a hybrid of the two was determined by RT-qPCR. Values shown are the mean of four biological replicates and error bars represent standard deviations. Values were normalized using a reference gene elongation factor-1α (EF1-α) as described in Wang et al., 2016. No further statistical analyses were carried out on this data.

The expression data for isolated leaf blade tissue only is shown in Figure 5.12. A similar pattern is seen compared to whole leaf tissue with expression levels being higher in C634 compared to R-o-18 with *BrGL1a*, *BrGL1b*, *BrTTG1b* and *BrTTG2b* expressed

at higher levels. It is the case again for many of the homologs under examination that expression levels are higher in the hybrid tissue than with R-o-18 and C634.

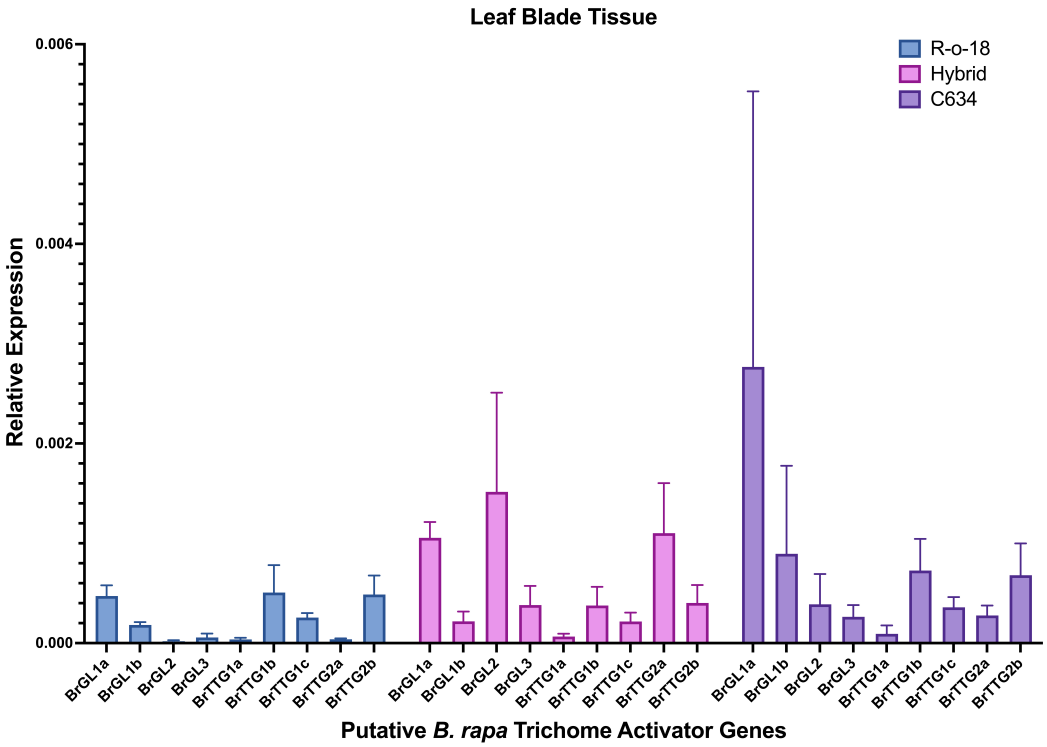

Figure 5.12: Expression in leaf blade tissue of *B. rapa* homologs of genes involved in the control of trichome initiation in *Arabidopsis*. Results for the expression of genes involved in trichome formation leaf blade from R-o-18 (BrCF1), C634 (BrCF2) and a hybrid of the two was determined by RT-qPCR. Values shown are the means of the four biological replicates and error bars represent standard deviations. Values were normalized using a reference gene elongation factor-1α (EF1-α) as described in Wang et al., 2016. No further statistical analyses were carried out on this data.

Figure 5.13 shows the expression data from isolated leaf margin tissue from the three accessions under examination. Again the trend of expression being higher in C634 than in R-o-18 is seen here but the highest expression levels are seen in the tissue from hybrid plants.

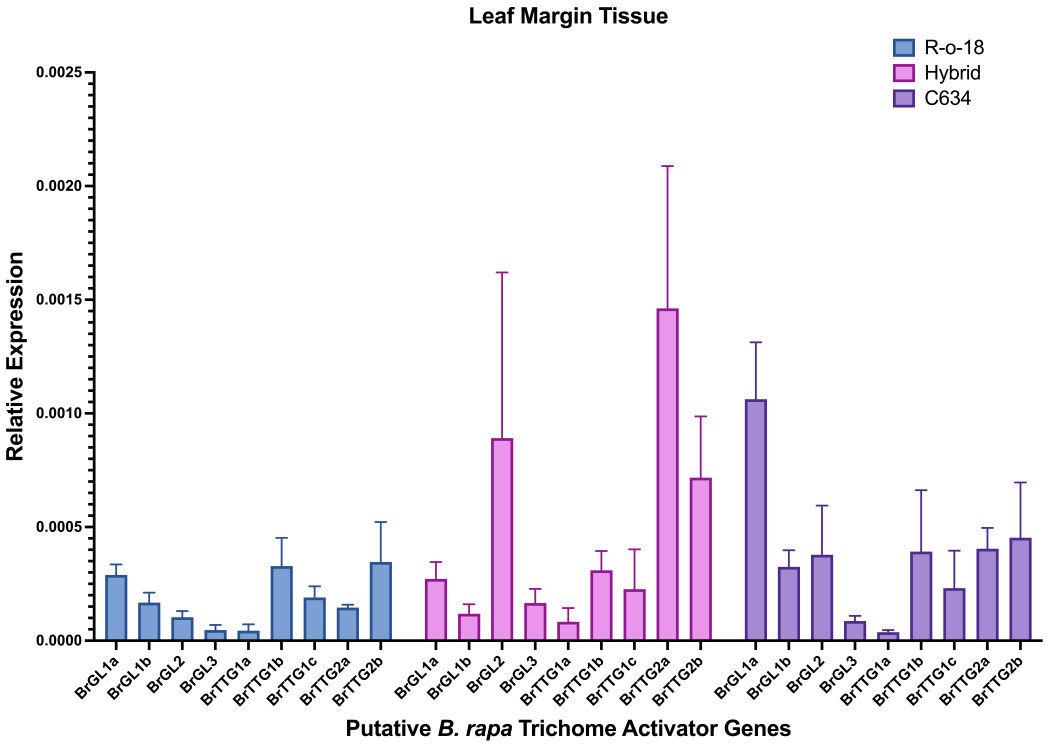

Figure 5.13: Expression in leaf margin tissue of *B. rapa* homologs of genes involved in the control of trichome initiation in *Arabidopsis*. Results for the expression of genes involved in trichome formation in leaf margin tissue from R-o-18 (BrCF1), C634 (BrCF2) and a hybrid of the two was determined by RT-qPCR. Values were normalized using a reference gene elongation factor-1α (EF1-α) as described in Wang et al., 2016. No further statistical analyses were carried out on this data.

When examining the collective results of these experiments, the most interesting of these was the expression data for the homolog of *GL2* (Figure 5.14). As detailed in Section 1.2, in *Arabidopsis* the formation of a trichome activator complex leads to the activation of the expression of *GL2* which in turn activates the expression of genes that lead to trichome development. Given that C634 has two fully functional copies of *BrGL1a*, the hybrid has one functional copy and R-o-18 lacks any, it was expected that expression of *BrGL2* would be highest in C634, with an intermediate level in the hybrid and the lowest level in R-o-18. However, the highest expression of *BrGL2* was detected in the hybrid plant tissue with the second highest in C634 and the lowest levels in R-o-18.

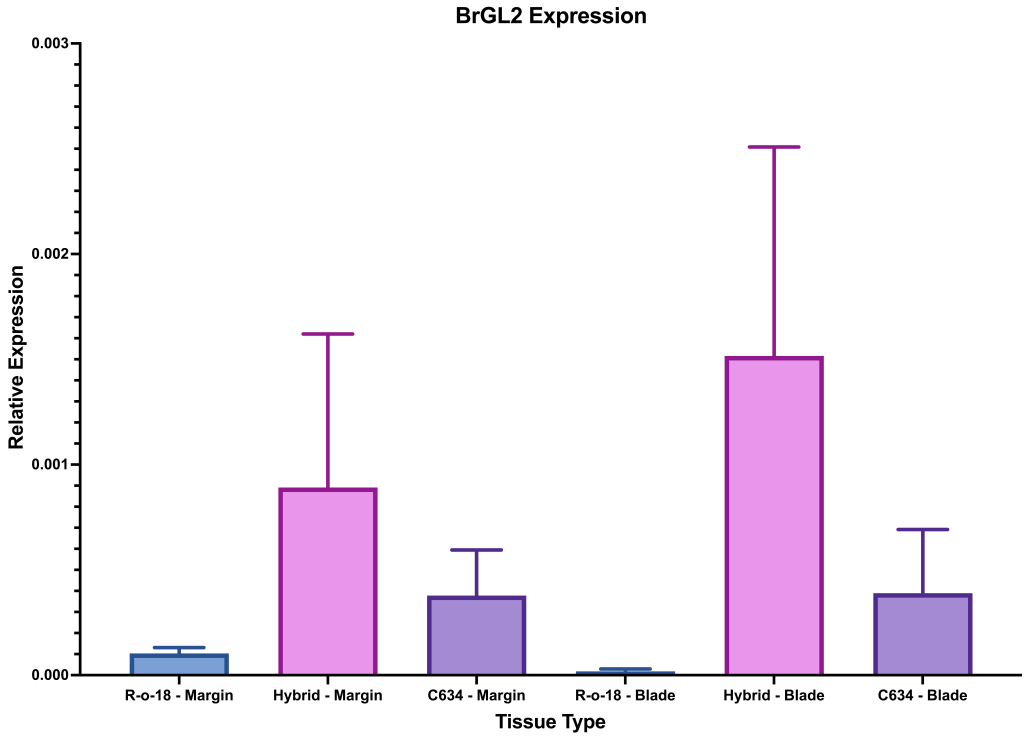

Figure 5.14: *BrGL2* expression. Expression of *BrGL2* across leaf blade and leaf margin tissue across R-o-18 (BrCF1), C634 (BrCF2) and the hybrid of the two. Values are the mean of four biological replicates. Error bars are standard deviation. Relative expression to REF gene. In both the leaf margin and the leaf blade tissue samples, the highest expression of *BrGL2* is seen in the hybrid plant tissue. Values were normalized using a reference gene elongation factor-1α (EF1-α) as described in Wang et al., 2016. No further statistical analyses were carried out on this data.

From these experiments, it was concluded that the genes involved in trichome formation display approximately the same expression profiles in both the leaf margin and leaf blade tissue with expression detected at higher levels in the "hairy leaf" C634 accession than in R-o-18. The lower levels of gene expression however do not explain the sparse trichome on the leaf margin and glabrous leaf blade observed in R-o-18 *B. rapa* as all the components of the MBW activator complex appear to be expressed. Surprisingly, the expression of these genes and particularly *BrGL2* was higher in the hybrid plants which display the same leaf phenotype as C634 for the most part. The potential genetic causes of this are discussed further in Section 5.4.

5.3.5 Double Mutants of *B. rapa* homologs of *AG*, *CPC* and *TRY*

As discussed in Section 1.4, one of the main aims of my SFI/EPA-funded PhD work was to try to induce trichomes on seed pods of *B. rapa* by perturbing C function and the activities of trichome repressors. The work described in Chapter 3 led to the isolation of *brag* mutant alleles that could be used as a sensitised background into which mutant alleles for *TRY* and *CPC* (Chapter 4) could be introgressed. Because of the above-mentioned abnormal trichome phenotype of the R-o-18 background, in which all mutants were generated, they first had to be crossed into the 'hairy" C634 line to ensure they were all present in a genetic background with a functional *BrGL1a*.

Expression data as shown in Section 3.2.2 indicated that both *BrAG1* and *BrAG2* are expressed in the gynoecium but *BrAG1* expression appeared almost five times higher than that of *BrAG2*. It was therefore concluded that *brag1-1* or *brag1-2* would be the most likely single mutant lines able to provide a sensitised background similar to *ag-10* used in the published *Arabidopsis* work. As shown by the expression data obtained in Section 4.2.2, *BrTRY2* appeared to be the most important *TRY* homolog to eliminate the trichome repressor function in order to potentially develop trichomes on the gynoecium. Of the *CPC* homologs, *BrCPC2* is more highly expressed in the gynoecium and it was therefore concluded that given the potential redundancy between the *BrTRY* homologs, it may be necessary to generate a *brag brtry brcpc* triple mutant to potentially see trichomes on the floral organs of *B. rapa*. The plan for the lines that were deemed essential to generate via crossing in order to fully explore the potential of floral trichomes in *B. rapa* are shown in Figure 5.15. The phenotypes of the double mutants generated are described in the sections that follow.

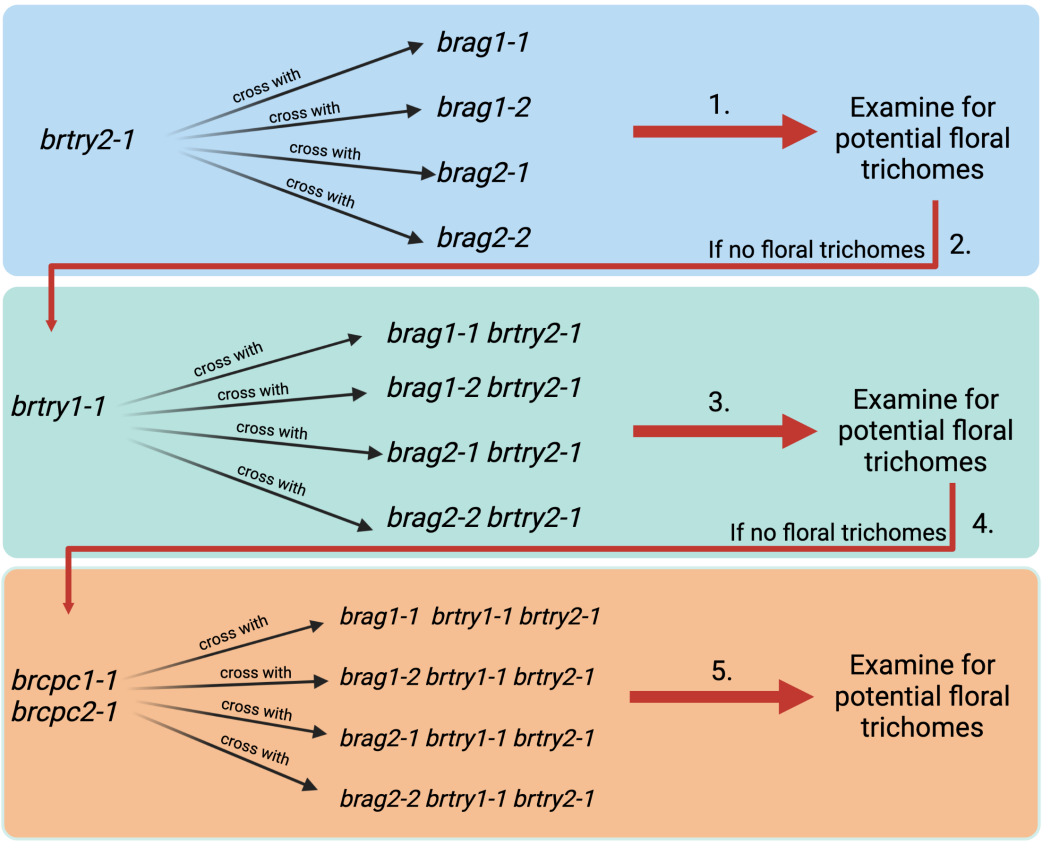

Figure 5.15: *B. rapa* lines to be generated for testing floral trichome formation. Schematic of the *B. rapa ag*, *try* and *cpc* lines to be generated to examine the genetic redundancy between the genetic network controlling the development of floral trichomes in *Arabidopsis* and the genetic network in *B. rapa* 1. Generate *brtry2-1 brag* double mutants in a C634 accession background and examine for potential floral trichome phenotypes. If there is no floral trichome phenotype, then 2. create *brtry1-1 btry2-1 brag* triple mutants and examine for potential floral trichome phenotypes. If there is no floral trichomes phenotype observed in the triple mutant lines, the 4. create *brcpc1-1 brcpc2-1 brtry1-1 brtry2-1 brag* quadruple mutants and examine for floral trichomes. Figure generated using Biorender.com.

Given the potential for functional redundancy among these homologs, it was decided to potentially include other mutant alleles if *brag1 brtry2* double mutants alone did not develop gynoecia with trichomes.

5.3.5.1 Generation of *brag brtry2-1* Double Mutants

Plants carrying the *brtry2-1* mutant allele in combination with one of the mutant alleles for one of the two *BrAG* homologs were isolated by genotyping and flowers were

examined for the occurrence of trichomes on seed pods. Flowers produced by *brtry2-1 brag1-2*, *brtry2-1 brag2-1* and *brtry2-1 brag2-2* double-mutant plants did not develop any trichomes on gynoecia or on any other floral organs (not shown). In contrast, *B. rapa* plants homozygous for *brtry2-1 brag1-1* produced flowers that developed trichomes on the gynoecium (Figure 5.16).

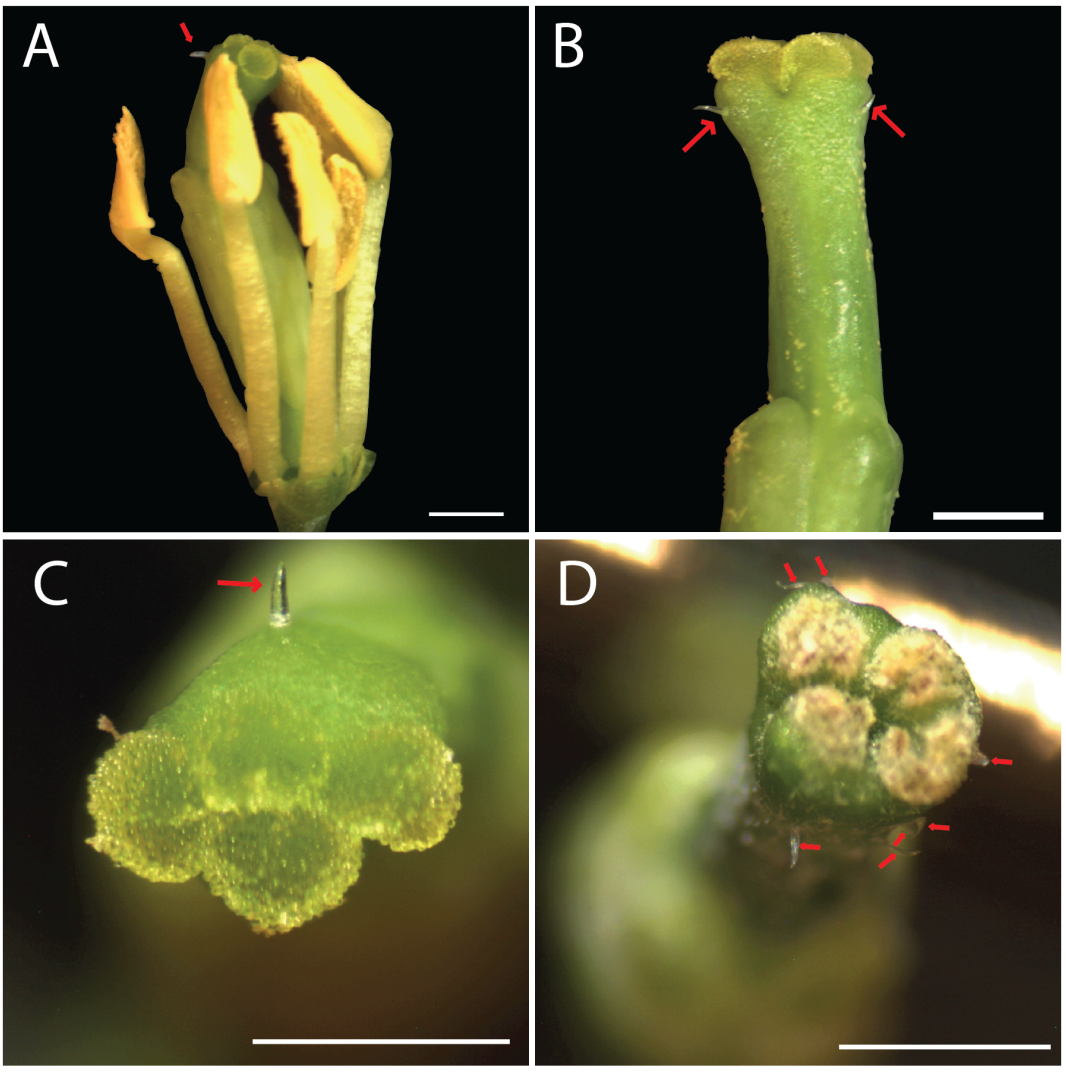

Figure 5.16: Floral phenotype of *brtry2-1 brag1-1* double-mutant plants. (A) A flower of a *brtry2-1 brag1-1* plant with a single trichome on the gynoecium. (B) Other flowers of this genotype developed multiple trichomes on the gynoecium. (C) A top view of the stigma of the flower in (A). (D) A top view of the stigma of the first flower produced by a *brtry2-1 brag1-1* plant with seven trichomes on the style. Red arrows indicate the location of trichomes. Scale bars: 1 mm.

The trichomes developed primarily on the styles of gynoecia which differs from the

trichome phenotype seen in *Arabidopsis ag try* plants where trichomes were formed primarily on carpel valves. A closer examination of *brtry2-1 brag1-1* plants showed that trichomes formed especially on early-arising flowers with the greatest numbers seen in the flowers produced first (5.16D). As the flowering process continued, flowers developed gradually fewer trichomes until flowers without trichomes on the gynoecium were observed. However, flowers with singular trichomes present on the gynoecium were still seen even at late stages of development.

5.4 Discussion

The aim of the work described in this chapter was to determine the cause of the "leaf-margin" phenotype observed in the R-o-18 line and to test whether trichomes could be induced on seed pods through the combination of mutant alleles for the BrAG and the trichome repressor genes described in Chapters 3 and 4, respectively.

A *B. rapa* accession, C634, with a "hairy-leaf" phenotype, was obtained and characterized. The sequence of *Bra025311* (*BrGL1a*) which has previously been shown to be associated with the variation in trichome phenotype (as discussed in 5.1) from both lines was examined. The C634 accession was shown to possess the functionally critical tryptophan amino acid residue at position 92 of the *BrGL1a* protein described by Zhang et al. (2018). Strikingly, a 249bp insertion was identified in the second exon of *BrGL1a* from R-o-18, which introduced an early stop codon and resulted in a truncated 121 amino acid protein that also lacked the essential tryptophan at position 92. The sequence of this insert was used to conduct a BLAST search and it was concluded that the insert originated from within the *B. rapa* genome as shown in Table 5.2.

The C634 accession and R-o-18 line were crossed and the progeny was examined to confirm that the variation at the *BrGL1a* locus was the causative locus for the difference in their trichome phenotypes. All F1 plants showed the C634 "hairy-leaf" phenotype (Figure 5.6) and the segregation ratio shown in Table 5.4 of the F2 (3:1, "hairy-leaf":"leaf-margin") proved that the "hairy-leaf" trait was acting in a dominant fashion. All F2 plants that showed the "hairy-leaf" phenotype were heterozygous or homozygous for the C634 *BrGL1a* locus (Figure 5.7). It was concluded that the variation at the *BrGL1a* locus was the cause of the different trichome phenotypes observed and that it was possible to cross the TILLING SNPs present in an R-o-18 background into C634 in order to assure the lines have the ability to develop trichomes and potentially develop floral trichomes.

The genetic cause of the "leaf-margin" phenotype of R-o-18 was investigated by obtaining a TILLING line containing a strong mutant allele for the other *AtGL1* homolog, *Bra039065* (*BrGL1b*). Plants that were homozygous for *brgl1b-1* developed leaves that were completely glabrous as shown in Figure 5.8. In order to investigate if *BrGL1b* function was essential for the formation of trichomes on the leaf margin, *brgl1b-1* was crossed with C634. The F1 progeny developed trichomes on the leaf margin as shown in Figure 5.9. Figure 5.10 shows a typical leaf of a plant that is homozygous for *brgl1b-1* in a C634 background which develops trichomes on the leaf margin and leaf blade. For these crosses, it was concluded that *BrGL1a* is the *AtGL1* homolog that is driving the formation of trichomes with *BrGL1b* having little to no role in the trichome development process as it alone cannot drive proper trichome formation on the leaf margin or blade. Only in the absence of a functional allele of *BrGL1a*, *BrGL1b* seems to initiate trichome development in an extremely inefficient manner with trichomes developing on the leaf margin only. It was proposed that *BrGL1b* is in the process of becoming a pseudogene as it cannot fulfil the role that *AtGL1* in *Arabidopsis* does. It was not possible to conclude from the work carried out in this study if this was due to a mutation in the coding DNA sequence or a promoter element of *BrGL1b*.

The expression of genes involved in trichome formation in whole leaf, leaf blade and leaf margin tissue were analysed using RT-qPCR in Section 5.3.4. The expression data revealed that the *B. rapa* homologs involved in the formation of the trichome formation are still expressed in R-o-18 despite the small number of trichomes that it produced and expression of these genes is detected at a higher level in C634. Expression of all the genes examined was detected at a higher level in leaf blade tissue than in leaf margin tissue which is logical given that there are more trichomes produced on the surface of the leaf than on the margins. Interestingly, the highest expression levels for some trichome regulatory genes, including *BrGL2*, were detected in the tissue of hybrid plants possibly as a result of heterosis. Heterosis is a phenomenon observed when the progeny

of a cross between species or a cross between diverse varieties of a species exhibit improved or increased function of a biological trait of an offspring (Birchler et al., 2010).

In the final part of this chapter, I presented the results of the analysis of *brag brtry2-1* double-mutant plants in a C634 background. The crossing plan with the combination of mutant alleles (with the expression data from 3.2.2 and 4.2.2 in mind) that have the best chance of reproducing the floral trichome phenotype seen in *Arabidopsis* (as detailed in Section 1.3) is shown in Figure 5.16. Due to issues affecting plant health over the course of this project such as infestations of pests including long periods of thrip and aphid infestations as well as difficulties ensuring optimal plant growth conditions in the facilities available during the course of this work, I was only able to reach the stage of phenotypically analysing *brag brtry* mutant lines. Given the expression data from Sections 3.2.2 and 4.2.2, I prioritised the generation of *brag brtry* lines as I concluded that was the double mutant with the best potential for developing trichomes on the gynoecium of flowers.

Strikingly trichomes were observed forming on the gynoecium of flowers of plants that were homozygous for both the *brag1-1* and *brtry2-1* mutant alleles in the C634 accession background. This was the double mutant that was the most likely to develop floral trichomes given that *BrAG1* and *BrTRY2* are the most highly expressed *AG* and *TRY* homologs in gynoecium tissue. The results here confirm that the genetic network controlling trichome formation as detailed in Ó'Maoiléidigh et al. (2018) is conserved in *B. rapa*. However, there are notable differences. The trichome phenotype observed here appeared weaker than that described for *Arabidopsis* when functionally similar mutant alleles were used. This could be due to a generally less active trichome initiation mechanism in *B. rapa*. Such a reduction appears likely because even the hairy C634 plants form relatively fewer trichomes when compared to *Arabidopsis* wild-type plants. The location of trichome formation also showed clear differences. Whereas in *Arabidopsis*, trichomes were primarily found on carpel valves, in *brag1-1 brtry2-1*

double mutants they were initiated on the style. Interestingly, the trichomes forming on the style in *B. rapa* is similar to the pattern of development seen in the closely related *Sinapis alba* which develops gynoecia and subsequent seed pods with trichomes on the surface. Scanning electron microscope images taken by Frank Wellmer show that trichomes begin to develop on the style of the gynoecia first before developing on the stigma and carpels (Frank Wellmer, unpublished data). Floral trichome formation in *B. rapa* also decreased as plants grew older, which has not been described for *Arabidopsis*. The genetic basis of this acropetal phenotype is currently unclear.

This work, alongside the work contained in Chapters 3 and 4, provides insight into the level of conservation in the genetic network suppressing floral trichome formation between *Arabidopsis* and *B. rapa*. As discussed in Section 1.3, in *Arabidopsis AG* is central to suppressing trichome development on gynoecia, repressing genes essential to trichome formation such as *GL1* and promoting genes that repress trichome formation such as *CPC* and *TRY*. The combined knockdown of *AG* function with the knockout of trichome repressors leads to trichome formation on the gynoecium of the flower. As shown in Figure 5.16, the knockout of *BrAG1* and *BrTRY2-1* function results in the development of trichomes on the gynoecia of flowers in *B. rapa*. However, unlike *Arabidopsis*, the phenotype is much more mild suggesting that both *AG* homologs in *B. rapa* are involved in repressing genes essential to trichome formation and promoting genes involved in trichome repression. The less dramatic phenotype seen in *B. rapa* also suggests a level of redundancy between the homologs of *TRY* or perhaps a greater role of the other R3 MYBs such as *CPC* in suppressing trichome formation on the gynoecia.

5.5 Future Work

The analysis of the *brag1-1 brtry2-1* double mutant showed that trichomes can be initiated on the seed pods of *B. rapa* through genetic manipulations similar to those used in *Arabidopsis*. While this is a promising result, the number of trichomes observed in this background is in all probability too small to provide any meaningful protection against insect pests. A main aim of any future work must therefore be to try and increase the number of trichomes on the seed pods. To this end, the *brtry1-1* mutant allele will be introgressed into the *brag brtry2-1* double-mutant background to test whether the disruption of BrTRY1 leads to a stronger trichome phenotype. In addition, plants that contain mutant alleles for *BrAG1*, *BrTRY* and *BrCPC* will also be generated in different combinations, with the hope that this will lead to an increase in floral trichome numbers. Treatments of *brag brtry2-1* plants with BAP will be carried out in order to test if the *B. rapa* homologs of *AG* also act to limit the cytokinin response in a similar manner to *AtAG* in *Arabidopsis*.

In addition to attempting to increase trichome numbers on gynoecia, it will be important to assess the mutant lines for agronomically important traits, in particular fertility and seed set, which may be affected in the lines I generated due to the partial disruption of C function. Furthermore, because regulators of trichome formation also play a role in the formation of root hairs, the roots of the mutants need to be investigated and overall plant health needs to be monitored. Complementation experiments examining the rescue of a *gl1 Arabidopsis* mutant by *BrGL1a* from R-o-18 and C634 will be investigated. If the 249bp insertion is affecting the ability of R-o-18 *BrGL1a* (detailed in Section 5.3.1) to interact with BrGL3, in complementation experiments one would expect to see no restoration of trichomes in *Arabidopsis gl1* mutant lines when transformed with the R-o-18 *BrGL1a* expressed under its native promoter and restoration of trichomes when transformed with C634 *BrGL1a* under its native promoter.

Chapter 6 - Investigating the Role of LEAFY in Auxin Modulation in Floral Meristems

6.1 The Role of Auxin in the Floral Meristems

Auxin is a plant hormone that affects nearly every aspect of plant growth and development, including cell division and elongation, tropic responses and flowering (Lee et al., 2019). In *Arabidopsis*, auxin synthesis is carried out via two key pathways - the tryptophan-independent pathway and the tryptophan-dependent pathway which is the mechanism which has been studied in much greater detail (Cao et al., 2019). This pathway involves the processing of tryptophan by aminotransferases including TRYPTOPHAN AMINOTRANSFERASE OF ARABIDOPSIS (TAA) and members of the TRYPTOPHAN AMINOTRANSFERASE RELATED (TAR) family. The products from these reactions are then processed by flavin monooxygenase of the *YUCCA* (*YUC*) family (of which there are 11 members) into indole-3-acetic acid (IAA), which is the most common form of auxin (Zheng et al., 2013). *TAA* and *TAR* gene members are expressed broadly across the plant while *YUC* genes are expressed in specific tissue types with *YUCCA1* (*YUC1*) and *YUCCA4* (*YUC4*) being expressed in the inflorescence meristem (Cao et al., 2019; Blakeslee et al., 2019; Matthes et al., 2019).

As detailed in Section 1.1.1, the expression of floral meristem identity genes as well as the event known as auxin maxima are essential for the initiation and the specification of floral meristems. This accumulation of auxin is due to its polarized transport in the inflorescence meristem via the PIN-FORMED1 (PIN1) efflux carrier following its phosphorylation by PINOID (PIN) (Sassi and Vernoux, 2013). The auxin maxima event causes the release of MONOPTEROUS/AUXIN RESPONSE FACTOR5 (MP/ARF5) from Aux/IAA repression (Denay et al., 2017). MP/ARF5 in turn directly induces the expression of *LEAFY* (*LFY*) which in turn activates upregulates expression of *APETALA1* (*AP1*) (Saddic et al., 2006; Kaufmann et al., 2010; Yamaguchi et al., 2013).

Following the auxin maxima event, PIN1 relocalizes to the basal pole of the cell and acts in conjunction with other auxin-influx carriers likely members of the AUX1/LAX family, resulting in an auxin sink. It has been shown that this auxin sink is necessary for the correct phyllotaxis of the emerging flowers of the developing plant (Kierzkowski et al., 2013). Previous studies have investigated the role of LFY in the control of auxin biosynthesis and response (Li et al., 2013; C. Winter et al., 2011). In particular, Li et al. (2013) proposed a model in which LFY plays a key role in decreasing auxin levels in the floral meristem after auxin maxima and this is achieved *via* the modulation of expression of genes involved in auxin biosynthesis and transport as listed in Table 6.1. Li et al. (2013) places emphasis on the downregulation of the auxin biosynthesis gene *YUC4* following LFY activation.

Table 6.1: Genes involved in auxin biosynthesis and response studied by Li et al. (2013). Genes under examination in W. Li et al., 2013. Gene names, aliases and functions are listed.

Gene	Alias	Function
YUCCA1	YUC1	Involved in auxin biosynthesis
YUCCA4	YUC4	Involved in auxin biosynthesis
PIN-FORMED 1	PIN1	Auxin efflux carrier
INDOLE-3-ACETIC ACID INDUCIBLE 1	IAA1	Repressor of auxin-inducible gene expression
INDOLE-3-ACETIC ACID INDUCIBLE 17	IAA17	Repressor of auxin-inducible gene expression
INDOLE-3-ACETIC ACID INDUCIBLE 29	IAA29	Repressor of auxin-inducible gene expression
TRYPTOPHAN AMINOTRANSFERASE RELATED 1	TAR1	Involved in auxin biosynthesis
TRYPTOPHAN AMINOTRANSFERASE RELATED 2	TAR2	Involved in auxin biosynthesis

There were several potential issues identified in the work carried out in Li et al. 2013. The work was carried out using a newly constructed line - p35S::GVG-6xUAS::LFY - to examine the effect of LFY activation on the expression levels of genes involved in auxin synthesis and transport. However, this construct was not in a *lfy* mutant background therefore induction using DEX would not be the first time LFY was becoming active in the plant as this would have occurred during the process of flower development. Another potential limitation of the work was that after its publication it was shown that some LFY target genes require the activity of AP1 and together the two in tandem act to antagonistically regulate the expression of LFY targets (Goslin et al., 2017). It is therefore not possible to emphatically state that LFY alone is responsible for the modulation of gene expression of auxin-related genes. Additionally, in experiments previously conducted in our laboratory with a line that allows a specific activation of LFY, results were obtained that were in disagreement with the data published by Li et al. (2013). Thus, there is a need to verify or reconcile the available data for the role of LFY in controlling auxin responses in floral meristems.

6.1.1 Approaches to Investigating Genetic Processes in the Floral Meristem

The Floral Induction System (FIS) is a transgenic plant line first described by Wellmer et al. (2006), which allows the specific activation of flower formation and the collection of hundreds of developmentally synchronised floral buds from a single plant. This line is based on plants that contain loss of function mutations in *AP1* and *CAL*. *Ap1 cal* double mutants undergo an over-proliferation of the inflorescence meristem which leads to a cauliflower-like appearance shown in Figure 6.1A (Ferrándiz et al., 2000). To induce flowering in this background, a construct was introduced into *ap1 cal* plants consisting of *AP1* fused to the hormone binding domain of the rat glucocorticoid receptor (GR) under the control of the constitutive CaMV 35S promoter.

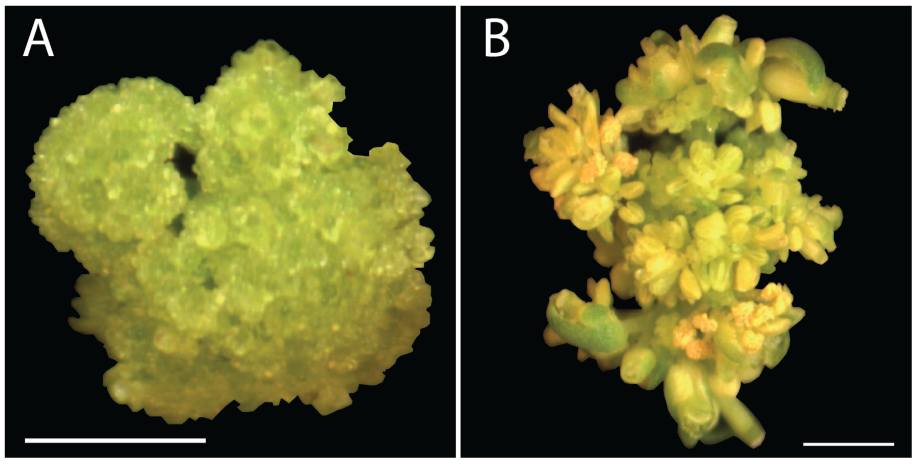

Figure 6.1: A floral induction system. (A) Inflorescence of a floral induction system plant in the absence of DEX. Meristems accumulate over time, leading to a cauliflower-like phenotype. (B) Inflorescence of a floral induction system plant 7 d after DEX treatment. Meristems are converted into floral meristems and flowers begin to form. Scale bars: 1 mm.

The molecular mechanism underlying the FIS system is detailed in Figure 6.2. Treatment of the FIS plants with the steroid hormone dexamethasone (DEX) results in the AP1-GR fusion protein being translocated to the nucleus and activating flower development in a synchronous manner as seen in Figure 6.1B. The same GR inducible

system has been applied to LFY in an *ap1 cal* mutant background in order to allow the examination of the effect of LFY activation on the floral meristem. However, in contrast to the p35S:AP1-GR *ap1 cal* line, activation of the LFY-GR fusion protein in p35S:LFY-GR *ap1 cal* plants does not lead to flower formation due to the lack of *AP1* and *CAL* activity (Goslin et al., 2017).

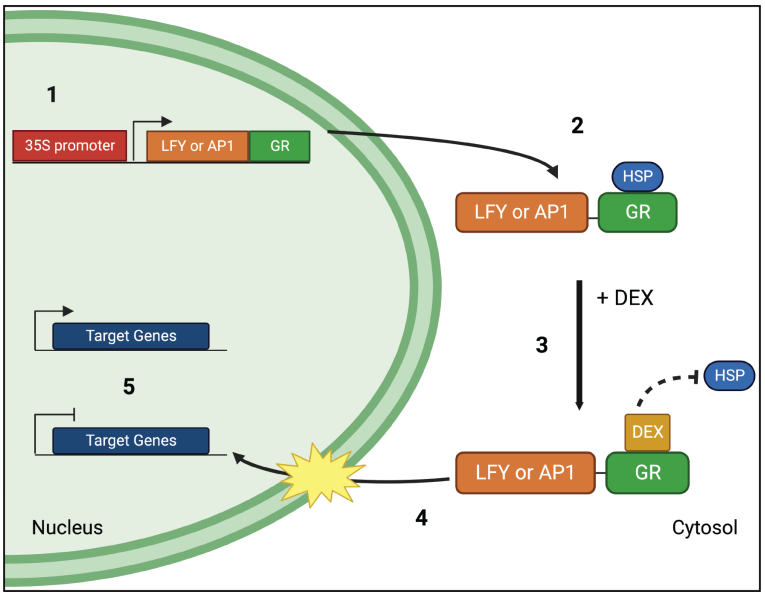

Figure 6.2: Molecular mechanism underlying the floral induction system and the line used for a specific activation of the LFY transcription factor. 1. AP1 or LFY are fused to the hormone-binding domain of the rat glucocorticoid receptor (GR). The fusion proteins are expressed in planta under the control of the constitutive 35S promoter. 2. After synthesis, the fusion proteins are retained in the cytoplasm by interactions with heat shock proteins (HSP). 3. Once DEX is applied, the HSPs dissociate from the fusion proteins, allowing their nuclear import. 4. The fusion proteins translocate to the nucleus. 5. The transcription factors that are part of the fusion proteins control the expression of their target genes.

6.2 Aims of the Chapter

As discussed in Section 6.1, the phytohormone auxin plays a key role in the control of the emergence of floral meristems but how exactly auxin impacts the gene network that initiates the floral meristem is an active area of research. A model was proposed by Li et al. (2013) which posits that LFY has a key role in downregulating auxin levels in the floral meristem and this was achieved by causing the downregulation of the auxin biosynthesis genes *YUC4*. A plant line with inducible LFY activation was used to examine changes in the relative expression of genes involved in auxin response in the floral meristem which are summarised in Table 6.1. These genes have crucial roles in the production and transportation of auxin in *Arabidopsis* and specifically the meristem of the plant. Upon activation of LFY, *YUC4* was downregulated and this is the experimental evidence on which Li et al. 2019 base their conclusion that LFY is reducing auxin levels *via YUCCA4*.

From the experiments presented by Li et al. (2013), it may not be possible to conclude that LFY alone is acting to perturb the expression of genes involved in auxin biosynthesis. The aim of the work carried out in this chapter was therefore to investigate the role that *LFY* has on genes related to auxin production and transportation. To this end, the well-characterised line p35S:LFY-GR *ap1-1 cal-1* was employed. The use of this line would allow for LFY activity to only begin upon treatment of the plants with dexamethasone and allow for confidence that the effects of are solely due to LFY as the line contains a null allele for *AP1*.

6.3 Results

6.3.1 LFY-Dependent Regulation of Genes Involved in Auxin Responses

The *Arabidopsis* line p35S:LFY-GR *ap1-1 cal-1* was used to carry out similar relative expression experiments as described by Li et al. (2013). Plants were treated with a solution containing DEX in order to activate LFY or were mock-treated. After the treatment, floral meristem tissue was collected at 2-, 4- or 8-hour time-points. RT-qPCR was employed to examine the effect that activation of LFY had on the expression of the genes listed in Table 6.1. The results of this analysis are summarised in Figure 6.3. Table 6.2 summarises the expression trend for each of the genes in the data generated using p35S:LFY-GR *ap1-1 cal-1* and directly compares it to the expression trends reported by Li et al. (2013).

Table 6.2: Comparison of RT-qPCR results for auxin-related genes. The results generated in this study using the p35S::LFY-GR *ap1-1 cal-1* line are compared to results reported by Li et al. (2013). General trends for the observed expression changes across the different time points are listed. *: only after 12 hours.

Gene	This study	Li et al. (2013)
YUC1	Upregulated	Downregulated
YUC4	Upregulated	Downregulated
PIN1	Downregulated	Upregulated
IAA1	Upregulated	No change
IAA17	Upregulated	No change
IAA29	Upregulated	Upregulated
TAR1	Upregulated	Downregulated
TAR2	Downregulated	Downregulated *

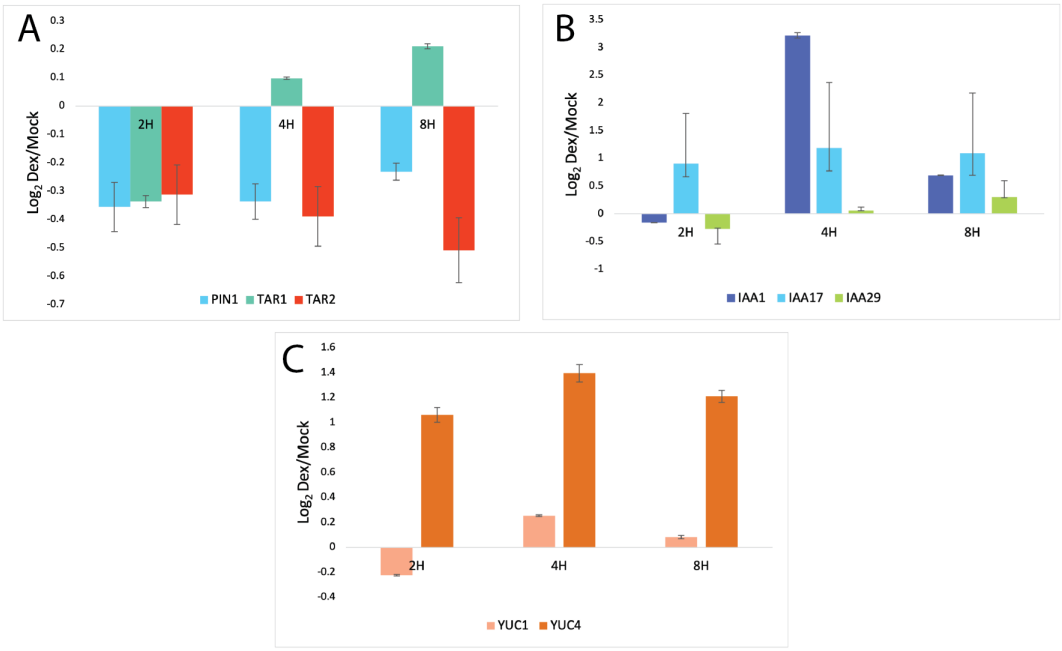

Figure 6.3: Effect of LFY activation on the expression of auxin-related genes. p35S:LFY-GR *ap1-1 cal-1* plants were treated either with a DEX-containing or with a mock solution. Subsequently, meristematic tissue was collected 2-, 4- and 8- hours post-treatment for RT-qPCR analysis. (A) Results for *PIN1*, *TAR1*, and *TAR2*. (B) Results for *IAA1*, *IAA17* and *IAA29*. (C) Results for *YUC1* and *YUC4*. Values shown are mean of four biological replicates. Log2-transformed fold-change values (DEX/mock) are shown. Error bars indicate SEs from three independent measurements.

As can be seen in Figure 6.3 and Table 6.2, the results obtained from the RT-qPCR data using the p35S::LFY-GR *ap1-1 cal-1* line shows in many cases different trends when compared to data generated in Li et al. (2013). Notably, *YUC4* appeared to be upregulated in data generated using the p35S::LFY-GR *ap1-1 cal-1* line which is in direct contradiction to the data from the 2013 paper. However, it was unclear whether this could be due to the absence of *AP1* activity in the data generated during the course of this study compared to the other data set.

6.3.2 Assessing LFY Function in the Regulation of Auxin-Related Genes through RNA Sequencing

An experiment previously conducted in our laboratory by Dr. Beibei Zheng involved carrying out RNA-sequencing in *ap1 cal* plants with a series of inducible constructs. One dataset was generated with LFY activation alone with DEX treatment and another dataset where AP1 was activated alone with dihydrotestosterone (DHT) treatment was generated. A third dataset where both LFY and AP1 were simultaneously activated through a combined DEX/DHT treatment. The data set was analysed to examine the expression data for the genes in Table 6.1. The results are shown in Figure 6.4 and compared to those of Li et al. (2013) in Table 6.3.

Table 6.3: Comparison of expression data for auxin-related genes. The results from RNA sequencing experiments of tissue after activation of both LFY and AP1 are compared to the results reported by Li et al. (2013). General trends for the observed expression changes across the different time points are listed. *: only after 12 hours.

Gene	DEX vs MOCK	DEX/DHT vs MOCK	Li et al. 2013
YUC1	Upregulated	Upregulated	Downregulated
YUC4	Upregulated	Upregulated	Downregulated
PIN1	Downregulated	Downregulated	Upregulated
IAA1	Upregulated	Slightly upregulated	No change
IAA17	Upregulated	Upregulated	No change
IAA29	Upregulated	Downregulated	Upregulated
TAR1	Upregulated	Downregulated	Downregulated
TAR2	Downregulated	Downregulated	Downregulated *

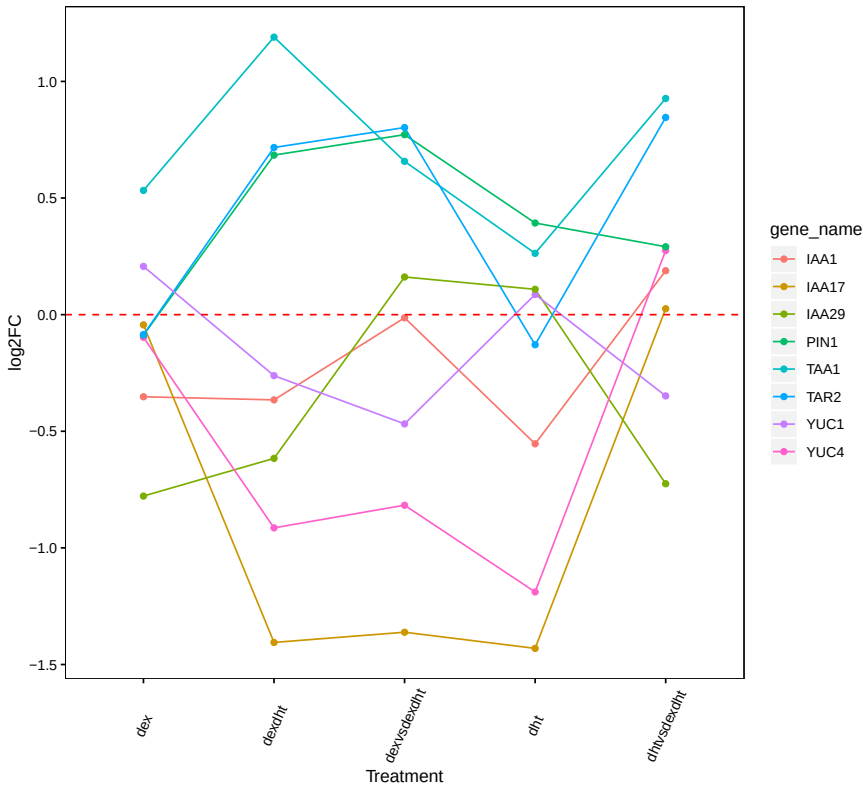

Figure 6.4: RNA sequencing results for auxin-related genes. The results from RNA sequencing experiments of tissue after activation of both LFY and AP1. LFY activity alone was induced with dexamethasone (DEX), AP1 activity alone was induced with dihydrotestosterone (DHT) and then both were induced using DEX and DHT. Activation of LFY alone (dex), activation of LFY and AP1 (dexdht), AP1 alone (dht) are shown. LFY activation alone is compared to LFY and AP1 activation (dexvsdexdht). AP1 activation alone is compared to LFY and AP1 activation (dhtvsdexdht). Log2-transformed fold-change values (DEX/mock) are shown.

When examining RNA-seq data where LFY and AP1 are both activated, it can be seen that again the trends do not agree with the results obtained using the p35S::GV G6xUAS::LFY line. The trends for the genes listed in Table 6.1 within the RNA-seq data sets are summarised in Figure 6.4. Additionally, the RNA-seq counts for *YUC4* are shown in Figure 6.5 where the number of counts detected in LFY alone, AP1 alone, and when both are activated compared to the mock treatment set. When LFY alone is activated *YUC4* is upregulated but critically *YUC4* is also upregulated also when both LFY and AP1 are activated. Notably, *YUC4* is only downregulated when AP1 alone is activated. Based on these results it appeared unlikely that LFY would downregulate auxin biosynthesis in floral meristems.

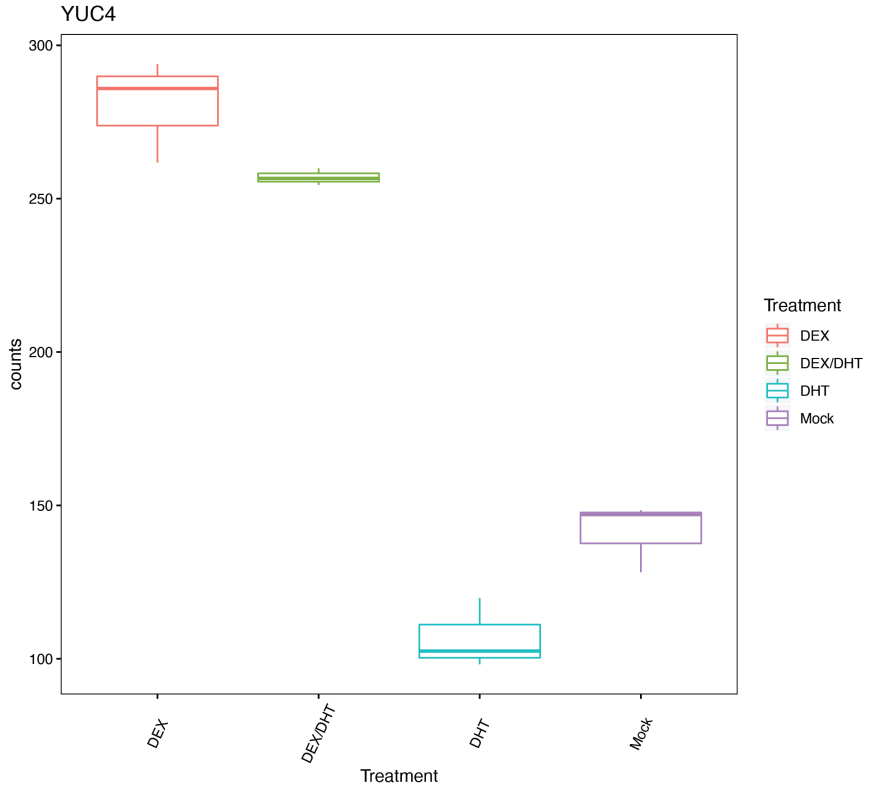

Figure 6.5: Read counts from RNA sequencing for *YUC4*. Read counts obtained in RNA sequencing dataset previously generated in the Wellmer laboratory. 'DEX' represents samples where a LFY-GR fusion protein was activated by DEX treatment; 'DHT' represents the samples where an AP1-AR fusion protein was activated by DHT treatment; and 'DEX/DHT' refers to samples where both LFY-GR and AP1-AR were activated. The results for a mock-treated sample ('Mock') are shown as well.

6.3.3 Quantification of IAA in Floral Meristems

To directly test the role of LFY in regulating auxin in the floral meristem, I decided to quantify the amount of IAA after a specific activation of LFY. To this end, p35S::LFY-GR *ap1-1 cal-1* and p35S::AP1-GR *ap1-1 cal*-1 plants were treated with a DEX-containing or with a mock solution and meristematic tissue was collected 8 and 16 hours after the treatment. These time points were chosen in order to measure both the direct and indirect effects of LFY activation on auxin levels. In total, four biologically independent sets of samples were generated. The tissue was prepared for analysis as described in Section 2.11, and IAA measurements were carried out by Prof. Stefan Pollmann at the

Universidad Politécnica de Madrid in Spain.

From the results of the IAA measurements (Figure 6.6), it appears that there is no significant difference in IAA levels in the DEX-treated samples when compared to the mock-treated samples in either of the plant lines or either of the two-time points taken.

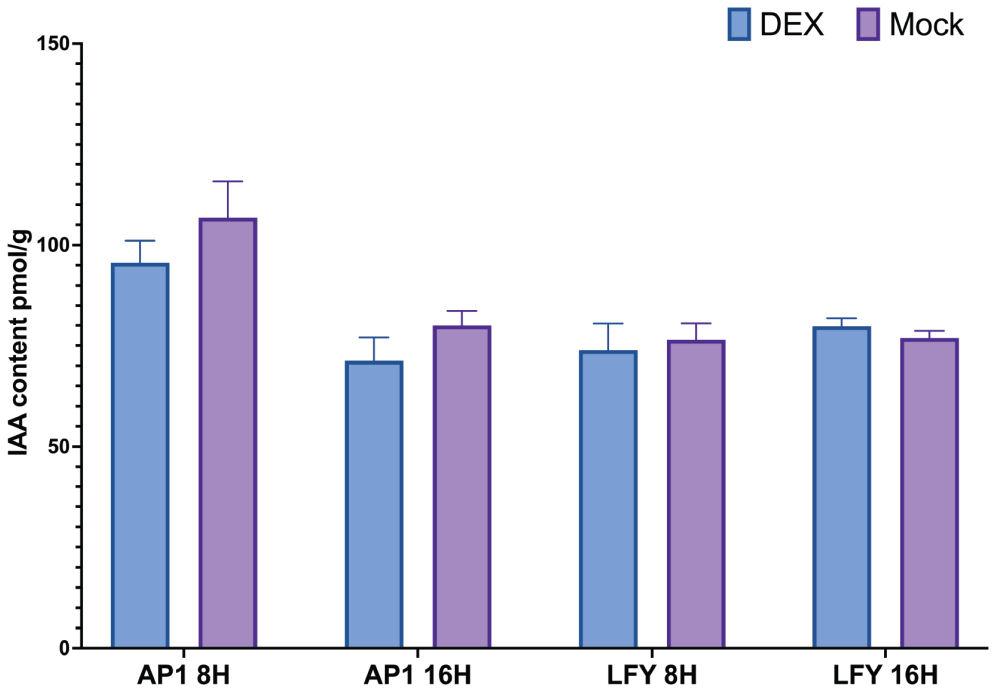

Figure 6.6: Effects of LFY and AP1 activation on IAA levels in meristematic tissue. Inflorescences of p35S:LFY-GR *ap1-1 cal-1* or of p35S:AP1-GR *ap1-1 cal-1* plants were treated either with a DEX-containing solution or with a mock solution and tissue was collected 8 and 16 hours after the treatment. The results of IAA measurement for the DEX-treated samples are shown in blue and for the mock-treated samples in purple. Error bars represent standard errors. Data were provided by Prof. Stefan Pollmann, Universidad Politécnica de Madrid, Spain.

In the case of p35S::AP1-GR *ap1-1 cal-1*, the levels of IAA detected were higher at the 8-hour than at the 16-hour time-point, but the values for the DEX- and mock-treated tissue samples were similar. The levels of IAA detected in meristems of p35S::LFY-GR *ap1-1 cal-1* plants were nearly identical across DEX- and mock-treated samples at both the 8 and 16-hour time-points. From the results of these experiments, it seems that neither the activation of LFY nor the activation of AP1 results in discernible changes in the

concentration of IAA in floral meristems. This provides further experimental support for the idea that LFY is not controlling auxin levels in floral meristems by downregulating the expression of the auxin biosynthesis gene *YUC4*, as previously suggested (W. Li et al., 2013). However, it is important to note that this data provides no insight into the effect of LFY or AP1 activation of the localization of IAA within the meristem. As discussed in Sections 1.1.1 and 6.1, the movement and location of auxin within the meristem are critical to the formation of the floral meristem and to ensure correct phyllotaxis of emerging flowers so perhaps LFY and AP1 are influencing auxin localization within the floral meristem as opposed to controlling the levels of auxin within the floral meristem.

6.4 Discussion

The aim of this chapter was to examine the role of LFY in the modulation of the levels of auxin in the floral meristem and compare the results obtained to previously published work by Li et al. (2013). This was accomplished by carrying out RT-qPCR experiments to examine how the activation of LFY influences the expression of genes involved in the biosynthesis and transport of auxin. The results of the RT-qPCR experiments show that activation of LFY does indeed have an impact on the expression levels of the genes under examination. There is no universal trend observed with some genes involved in auxin biosynthesis such as *YUC1*, *YUC4* and *TAR1* being upregulated following the activation of LFY but *TAR2* being downregulated. However, the upregulation of *YUC4* observed in the experiments carried out here directly contradicts the model proposed by Li et al. (2013) which is that LFY downregulates auxin in the floral meristem by downregulating *YUC4*.

Given that it has been previously shown that LFY and AP1 can control the expression of some genes in an antagonistic manner, a previously generated RNA-seq dataset where both LFY and AP1 were activated was examined to assess the expression trends for the auxin biosynthesis genes. The results confirmed that LFY activation alone resulted in *YUC4* upregulated and revealed that co-induction of LFY and AP1 also resulted in the upregulation of *YUC4*. These results were again in contradiction to the model that LFY is downregulating auxin levels in the floral meristem by downregulating *YUC4*.

Following the results of the RT-qPCR experiments and analysis of the RNA-seq data, it was decided to examine the effects of LFY to AP1 activation on the amount of auxin in the floral meristem. It was hoped that data on how the activation of LFY and AP1 impacted auxin levels would provide a clearer picture of how these two genes were affecting auxin in the floral meristem. To this end, p35S:LFY-GR *ap1-1 cal-1* and

p35S:AP1-GR *ap1-1 cal-1* tissue that had been DEX treated and then collected 8- or 16-hours following DEX induction had the amount of IAA present quantified. Surprisingly, the results of the quantification experiments revealed that the activation of LFY and AP1 seems to have no significant impact on the levels of IAA in the floral meristem. However, these results lend further support that the model of LFY downregulating auxin biosynthesis via the downregulation of *YUC4* is incorrect.

6.5 Future Work

During the course of this work, in an attempt to confirm the results of the auxin quantification experiments and further investigate what occurs in terms of auxin accumulation and localization in the floral meristem upon activation of LFY and AP1, DR5 imaging lines were generated. DR5 is an auxin response element first described in Ulmasov et al., 1997 that when fused to a reporter gene such as GUS or GFP can be used to visualise auxin. Both DR5::GUS and DR5::GFP were introgressed into the p35S::LFY-GR *ap1-1 cal-1* and p35S::AP1-GR *ap1-1 cal-1*. Plants will be treated with DEX or a mock solution and their respective reporters used to assay the effect that LFY or AP1 activation has. DR5::GFP confocal imaging under the same conditions will be carried out by Nathanaël Prunet at the University of North Carolina and I will complete the β-Glucuronidase (GUS) reporter assays. One would expect to see similar results to those in Figure 6.6 and that no difference would be detected in auxin levels following LFY and AP1 activation. However, these experiments may provide insight into the impact of LFY and AP1 activation on auxin localization within the floral meristem.

Appendix

7.1 List of Oligonucleotides Used in this Study:

Table 7.1: Oligonucleotides used in the course of this study. Table lists the details of the oligonucleotides used in the course of this study. The unique identifier, the purpose and the sequence of each.

Name	Category	Sequence
oCF15	Genotyping	ACCTATCTCTTCATCAGCACTACC
oCF16	Genotyping	TTGACGATCTTAACCAAGCACC
oCF18	Genotyping	TAGTTTGATTCTCGCTGGGAGG
oCF27	Genotyping	TGGCTTAGGAGCAGCAAACATG
oCF28	Genotyping	ATCTGCCAAATTGCGTCAAC
oCF30	Genotyping	CCAACAAGAATCGGCCAAATTGC
oCF31	Genotyping	TTCCACCCATAAGAGGGTTCC
oCF32	Genotyping	AGAAGAGAAGGGAAAGAGAGACAC
oCF33	Genotyping	GAGACACAGCAAGGCCAAAG
oCF34	Genotyping	GACAGAGTAATGGGCTTACGAAC
oCF35	Genotyping	TTTCTTCTCTTTGTGCTCGTGTCG
oCF36	Genotyping	AATGCTACGACCTTAACCAGTCC
oCF37	Genotyping	AAATGAGCCAACAGGAGGAAG
oCF38	Genotyping	CCGAGCATTTAATCCGGTTTG
oCF39	Genotyping	AGTGAGCAGTATCGAATGGAAG

oCF40	Genotyping	CGTCTGATTTCAAGAGTTTATGGTG
oCF54	DMAS	CGACCACTAACATAACCCTAACACG
oCF64	DMAS	AGGGACGTGGACGATGAAGTTTG
oCF72	DMAS	ATGGATCGAGGGTGAGAGTGTG
oCF78	qPCR	CCGTGGCCGTCTCTATGAATACTC
oCF79	qPCR	GTCCGTTAATTTCTGCCACGGTTC
oCF100	qPCR	CTGGTAATAACTGGGCGAAAGGTC
oCF101	qPCR	TGTGCCTCCTCCAAGAGAGTG
oCF133	Genotyping	TGAGAATCAAAGGATGAGAACGAGG
oCF134	Genotyping	GCAAGTGCATCTTGGCAGG
oCF139	qPCR	ACCAAATGGAGCTAGGAGGAGAATC
oCF140	qPCR	CCGATATTGCTTTCTTGTACCTCTCAATTG
oCF146	DMAS	GAGGCAAGAAATGTGTGCTGTTTCTC
oCF173	qPCR	GCCTACTTGGCTTCAGGGAG
oCF174	qPCR	TGGTGGACCCCTTGATTTCG
oCF175	qPCR	GTCACCGGTAAAACTCCCGT
oCF176	qPCR	TCACGGCTTGCGTCACTTTA
oCF181	qPCR	GCATCACCTGGTCCCTCATT
oCF182	qPCR	ACGCCATGAACAACCCAAGA
oCF183	qPCR	TCCTCCTCCTGCAAAAACACA
oCF184	qPCR	TGGAGCTCCGTCCATACTCA
oCF187	qPCR	ACCATGGGCAAACATGGAGG
oCF188	qPCR	CGAGCATCCAATCACCGTCT
oCF191	qPCR	ACATGGAAGATCTTTGCGGA
oCF192	qPCR	ATGGTCCGATTTGAACGCCT
oCF195	qPCR	CGGAACCATCAGAGAGACGG
oCF196	qPCR	TTCACCAATGCCCACCCAAT
oCF199	qPCR	TGTGGGAATGTAACGGGCAA

oCF200	qPCR	ATGCAAATGCTGGTTGTGGC
oCF203	qPCR	AAGCGGTTGTGGAGAACATGATACG
oCF204	qPCR	TGGAGAGCTTGATTTGCGAAATACCG
oCF232	qPCR	TTTGACTTGCACGACAAGGAACAC
oCF233	qPCR	TCCGAGCCAGCAGAGTAAACC
oCF234	qPCR	TCACCATGGGTCGTACTGACAC
oCF235	qPCR	GTCGCGCAAGTCAAAGATCCTG
oCF240	qPCR	AGCTTCCTCGAGGACTACAACAAC
oCF241	qPCR	TCGGCAACTCCCAGATAAGAGC
oCF284	qPCR	CGAAGAGTAAGTAGCATCGAATGG
oCF285	qPCR	CGGCTATCAATTCCACCTGTCTC
oCF286	qPCR	TCCGAAGAAGTGAGTAGCATCGAG
oCF287	qPCR	TCCCACCTGTCTCCAACGAG
oCF291	qPCR	TATAGACTTGTCGGTGATAGGTGG
oCF292	qPCR	GAAGGGTAAATAGAGGGACGTGG
oCF295	qPCR	ATGTACAGACTCGTAGGCGATAGG
oCF296	qPCR	TGGAAGATGGAGGCAAGAAATGTG
oCF299	qPCR	ACTCTCCATGACTCTGAAGAAGTG
oCF300	qPCR	TCGACAGATGAGATCTTCTTCTTG
oCF313	qPCR	GTTTCATGGACGTGGAGCAATGG
oCF314	qPCR	CCCAAACATCAACTGAATGGCAC
oCF317	qPCR	GGTGAAGGAATGCCTGGACG
oCF318	qPCR	TCCTCTGCAATGTGTTCTGTGG
oCF319	qPCR	CATGTCCACAGGTGGAAGGAAATG
oCF320	qPCR	TTACCCATATTAGCCAATGCGGAG
oCF322	qPCR	TTCTGTCTCATCTCCGCGGGTG
oCF323	qPCR	CTGGAAGCATATTAGCCAACGCAG
oCF407	Cloning	ATGGCTTACCAAATGGAGCTAGG

oCF409	Cloning	AACAGGTCTCAGGCTTTATGGCCTACCAAATGGAGCTAG
oCF411	Cloning	AACAGGTCTCAGGCTTTATGGCTTACCAAATGGAGCTAGG
oCF415	Cloning	AACAGGTCTCAGGCTTTATGGCGTACCAATCGGAGCTAG
oCF440	Cloning	TACACTCACACACATACACTTG
oCF441	Cloning	AACAGGTCTCACTGAGGTAGAATTTCACATCGCCATAAG
oCF442	Cloning	AACAGGTCTCACTGATACACTCACACACATACACTTG
oCF444	Cloning	AACAGGTCTCACTGAAATCATTAATAGAGCAGGAATTGG
oCF451	Cloning	AACAGGTCTCATGTTAATTGATAGGGTCAAATCGACCAC
oCF460	Genotyping	AGCAAGGCATATGAAGAAGAGAAG
oCF461	Genotyping	TGAACACCCAACTACAGCTC
oCF463	Cloning	AACAGGTCTCAACCTGTTTCTTCTTCTTCTCGTGCTCTG
oCF475	Cloning	AACAGGTCTCAGGCTTTATGTTTCGTTCAGACAAGGC
oCF477	Cloning	AACAGGTCTCAGGCTTTATGGATAGACGACGTCGG
oCF479	Cloning	AACAGGTCTCAGGCTTTATGGATAGACGTCGCAGG
oCF480	Cloning	AACAGGTCTCACTGAAAGCCAATAATCTCCATCCATCG
oCF481	Cloning	AACAGGTCTCAGGCTTTATGGATAACACTGACCGTCG
oCF482	Cloning	AACAGGTCTCACTGACCTAACCGCATGGATTAAAGTTG
oCF483	Cloning	AACAGGTCTCAGGCTTTATGACAGAACAAGAAGAAGATATT
oCF484	Cloning	AACAGGTCTCACTGAAAACTGCAACTATTACACC
oCF485	Cloning	AACAGGTCTCAGGCTTTATGAGCCAACAGGAGGAAG
oCF486	Cloning	AACAGGTCTCACTGAGATGGTCTAAAGGCCAAAGC
oCF487	Cloning	AACAGGTCTCAGGCTTTATGGATAACACCGACCGTC
oCF488	Cloning	AACAGGTCTCACTGATGGACGATGAGGTTTGGTAC
oCF499	Cloning	AACAGGTCTCACTGACTTGTCTTGTGAATTAAGGAGAGG
oCF500	Cloning	AACAGGTCTCACTGCCTTGTCATGCGAATATTCAG
oCF509	dCAPS	ATAGAATTACCTTCTTGGATCGG
oCF510	dCAPS	GATATATTAATATATGTTGATAAATCACTTA
oCF511	Genotyping	TCACTGAACAAGAAGAAGACCTC

oCF512	Genotyping	GTAGCCTAGTAGCCAATATCACC
oCF524	DMAS	AGCATGTCAAACATGGCCgACC
oCF525	DMAS	AGCATGTCAAACATGGCCgACt
oCF526	DMAS	GATTTCGATCTTTCCTCTCCCTGC
oCF530	DMAS	CAATAGGGTCAATGTCTCCCgAAG
oCF531	DMAS	CAATAGGGTCAATGTCTCCCgAAa
oCF532	DMAS	AAATTACCTTCTTGGATCGGATTCG
oCF542	DMAS	CGATTGGGTCAATGTCTCtCgAAG
oCF543	DMAS	CGATTGGGTCAATGTCTCtCgAAa
oCF544	DMAS	TCCTGCACAGACTTTACAAATGAGC
oCF547	DMAS	CCGTCTCTATGAGTACTCAAACAgCAGG
oCF548	DMAS	CCGTCTCTATGAGTACTCAAACAgCAGa
oCF551	DMAS	TGTTTTCTTGCAAAATACCAGtTGG
oCF552	DMAS	TGTTTTCTTGCAAAATACCAGtTGa
oCF553	DMAS	ACGGTGTTGTTTTTGCCAACgAAC
oCF554	DMAS	ACGGTGTTGTTTTTGCCAACgAAt
oCF555	DMAS	CCTTGTCATGCGAATATTCAGAAGATG
oCF556	DMAS	GCAGAAGTGAGTAGCATCGAtTGG
oCF557	DMAS	GCAGAAGTGAGTAGCATCGAtTGa
oCF558	DMAS	GAACAGATAACCAAGTCAAGAATCgCTGG
oCF559	DMAS	GAACAGATAACCAAGTCAAGAATCgCTGt
oCF560	DMAS	CGGAGGAATAATCTCCGACGAACTTC
oCF577	DMAS	TGCCAGGAAGACAGCCgGAAG
oCF578	DMAS	TGCCAGGAAGACAGCCgGAAa

7.2 Buffers and Solutions

TAE Buffer

50X stock solution:

242 g Tris base

57.1 mL glacial acetic acid

37.2g $Na_2EDTA.2H_2O$

Made up to 1 litre with ddH_2O

***Agrobacterium* colony PCR lysis buffer**

10 mM Tris-HCl (pH8.0)

0.1% (v/v) Triton X-100

CTAB Buffer

100 mM Tris-HCl (pH8.0)

20 mM EDTA

1.4 M NaCl

2% (w/v) CTAB

Facultative 1% PVP 40,000

T.E Buffer

10 mM Tris-HCl (pH 8.0)

1mM EDTA

Edwards Preparation Extraction Buffer

200 mM Tris-HCl (pH8.0)

25 mM EDTA

250 mM NaCl

0.5% (w/v) SDS

7.3 BLAST Results

The following table shows the top 25 BLAST search results for *AtAG* that overlapped with a *B. rapa* gene ID:

Table 7.2: BLAST results for the search of *AtAG* genomic sequence against the *B. rapa* chifuu genome. The BLAST results were sorted by the E-value and the results that did not overlap with a gene were disregarded. The top 25 hits are shown in the table. *Bra012564* occurs 10 times and *Bra013364* shows up 12 times in the top 25 results.

BRAD ID	Length (bp)	E-value	% ID
Bra013364	205	3.20E-82	94.1
Bra013364	198	1.90E-80	94.4
Bra012564	182	1.20E-75	95.1
Bra012564	124	5.50E-53	96.8
Bra013364	141	4.90E-44	91.5
Bra013364	121	4.90E-44	94.2
Bra012564	107	1.90E-40	95.3
Bra012564	118	7.40E-40	93.2
Bra013364	92	6.80E-34	95.7
Bra013364	120	1.70E-31	90.0
Bra012564	119	6.60E-31	89.9
Bra013364	85	1.00E-29	95.3
Bra013364	73	2.50E-27	97.3
Bra012564	92	2.40E-24	91.3
Bra012564	84	2.40E-24	92.9
Bra004927	83	9.50E-24	92.8
Bra013364	70	3.80E-23	95.7
Bra013364	81	1.50E-22	92.6
Bra013364	73	1.50E-22	94.5
Bra012564	70	9.20E-21	94.3
Bra012564	70	9.20E-21	94.3
Bra012564	81	8.90E-18	90.1
Bra003356	156	3.50E-17	82.7
Bra007419	156	3.50E-17	82.7
Bra013364	53	2.20E-15	96.2

The following table shows the top 25 BLAST search results for *AtTRY* that overlapped with a *B. rapa* gene ID:

Table 7.3: BLAST results for *AtTRY* genomic sequences against the *B.rapa* chifuu genome. The BLAST results were sorted by the E-value and the results that did not overlap with a BRAD gene ID were disregarded. The top 25 results are shown in the table. *Bra022637* occurs four times while *Bra029089* and *Bra026297* both occur twice.

BRAD ID	Length (bp)	E-value	% ID
Bra022637	106	9.20E-36	93.4
Bra029089	98	9.20E-36	94.9
Bra029089	87	4.90E-22	90.8
Bra022637	113	7.60E-21	86.7
Bra022637	58	1.10E-16	94.8
Bra022637	51	4.20E-10	92.2
Bra026297	65	6.60E-09	87.7
Bra026297	44	0.0015	88.6
Bra011616	25	0.095	96.0
Bra031385	24	0.38	95.8
Bra029696	20	0.38	100.0
Bra037887	24	0.38	95.8
Bra038661	28	0.38	92.9
Bra010815	20	0.38	100.0
Bra025589	24	0.38	95.8
Bra037332	19	1.5	100.0
Bra039065	23	1.5	95.7
Bra029467	23	1.5	95.7
Bra019716	19	1.5	100.0
Bra025311	23	1.5	95.7
Bra010551	23	1.5	95.7
Bra040939	22	5.9	95.5
Bra004588	26	5.9	92.3
Bra039501	22	5.9	95.5
Bra032231	18	5.9	100.0

The following table shows the top 25 BLAST search results for *AtCPC* that overlapped with a *B. rapa* gene ID:

Table 7.4: BLAST results for *AtCPC* genomic sequences against the *B. rapa* chifuu genome: The BLAST results were sorted by E-value and those that did not overlap with a BRAD gene ID were disregarded. The top 25 results are shown in the table. *Bra004539* and *Bra039283* both occur three times.

BRAD ID	Length (bp)	E-value	% ID
Bra004539	127	1.10E-31	89.0
Bra039283	126	4.50E-31	88.9
Bra004539	94	1.10E-28	92.6
Bra039283	94	2.70E-26	91.5
Bra004539	57	5.40E-09	89.5
Bra039283	32	5.20E-06	96.9
Bra017237	23	0.005	100.0
Bra029089	54	0.02	85.2
Bra007264	21	0.079	100.0
Bra032652	21	0.079	100.0
Bra039283	57	0.079	84.2
Bra005267	24	0.31	95.8
Bra015389	20	0.31	100.0
Bra007264	20	0.31	100.0
Bra023069	24	0.31	95.8
Bra025554	20	0.31	100.0
Bra034945	24	0.31	95.8
Bra004825	19	1.2	100.0
Bra039441	23	1.2	95.7
Bra009618	19	1.2	100.0
Bra016517	23	1.2	95.7
Bra026297	19	1.2	100.0
Bra023172	19	1.2	100.0
Bra005703	19	1.2	100.0
Bra001212	19	1.2	100.0

7.4 Genomic Sequence Alignments

Figure 7.1 shows the sequence alignment for *AtAG*, *Bra012564* and *Bra013364*.

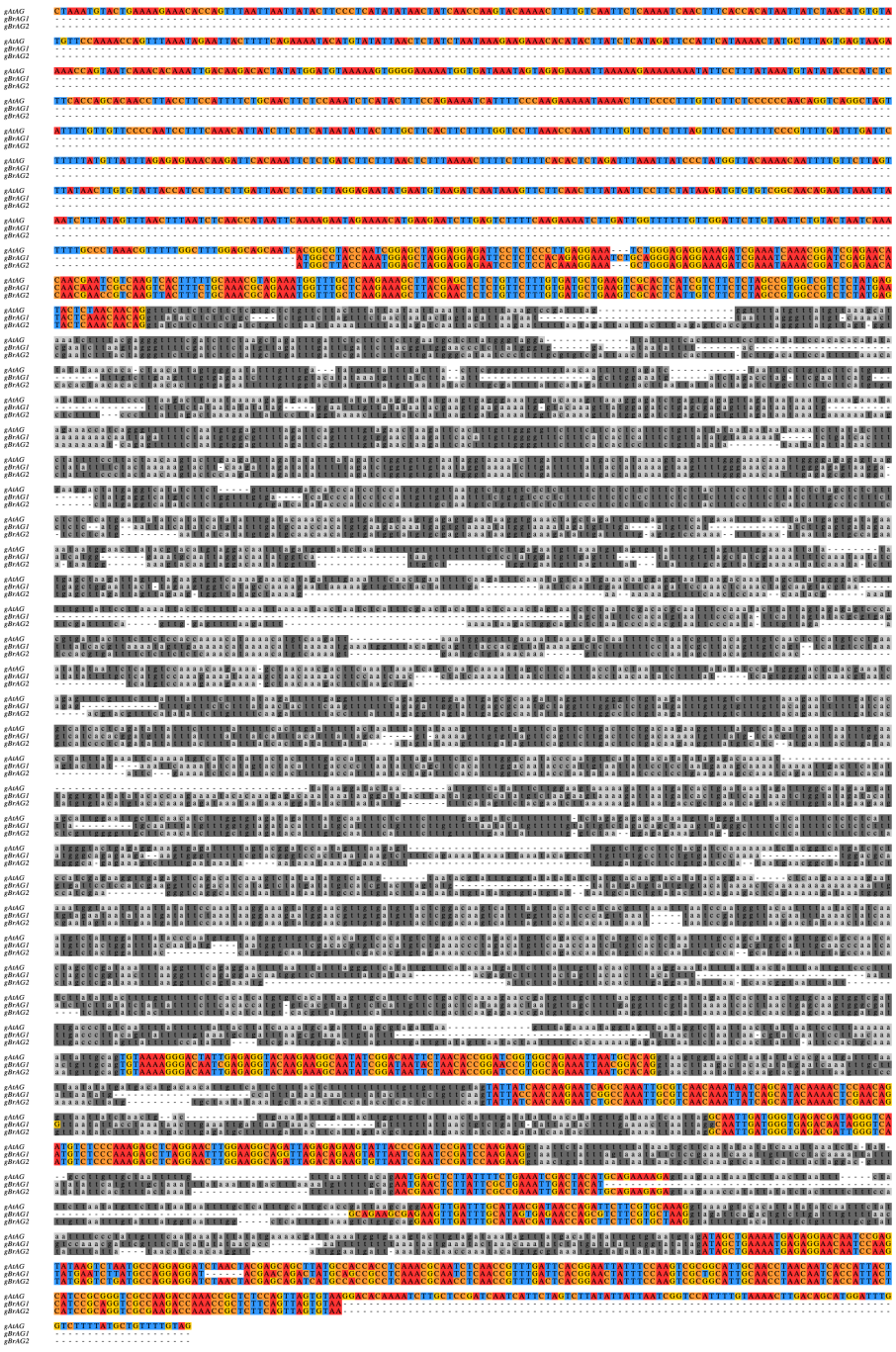

Figure 7.1: Alignment of the genomic sequence of *AtAG*, *Bra012564* and *Bra013364*. The sequences from the start codon to the stop codon of AtAG, Bra012564 and Bra013364 were aligned using the EMBL-EBI Clustal Omega Multiple Sequence Alignment tool. The results were visualised using *Jalview* software. The coding DNA sequences are in upper case and in colour whilst intronic sequences are in lower case and in grey.

Figure 7.2 shows the sequence alignment for *AtTRY*, *Bra022637*, *Bra026297* and *Bra029089*.

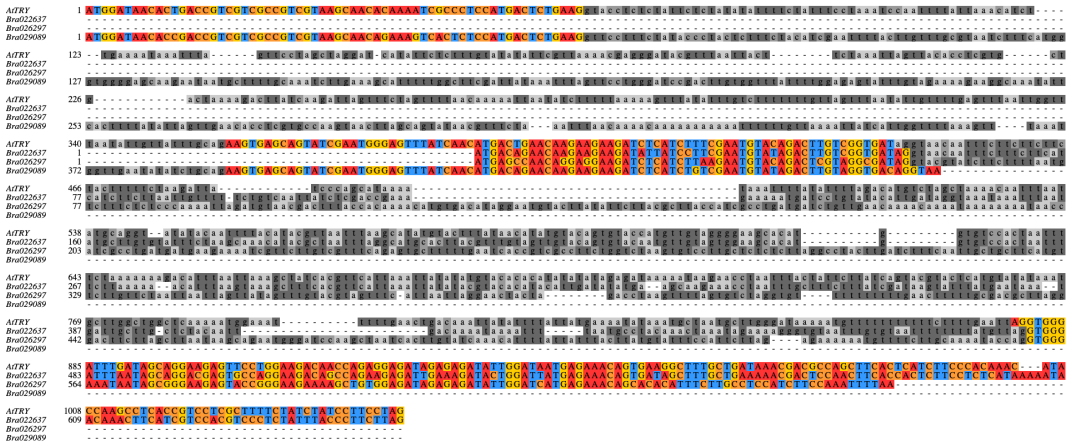

Figure 7.2: Alignment of the genomic sequence of *AtTRY*, *Bra022637*, *Bra026297* and *Bra029089*. The sequences from the start codon to the stop codon of *AtTRY*, *Bra022637*, *Bra026297* and *Bra029089* were aligned using the EMBL-EBI Clustal Omega Multiple Sequence Alignment tool. The results were visualised using *Jalview* software. The coding DNA sequences are in upper case and in colour whilst intronic sequences are in lower case and in grey.

Figure 7.3 shows the sequence alignment for *AtCPC*, *Bra004539* and *Bra039283*.

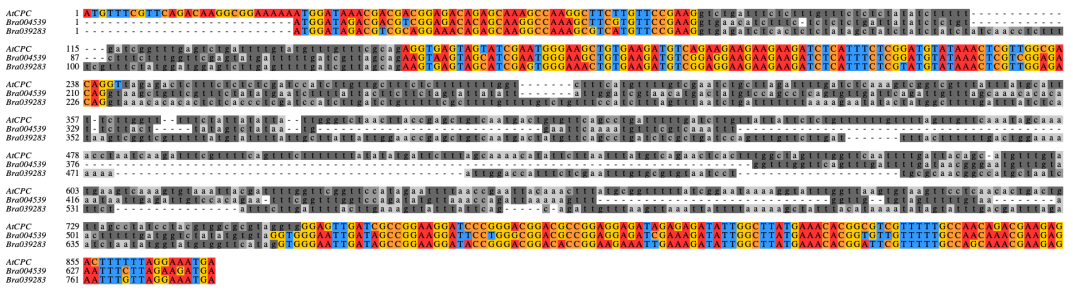

Figure 7.3: Alignment of the genomic sequence of *AtCPC*, *Bra004539* and *Bra039283*. The sequences from the start codon to the stop codon of *AtCPC*, *Bra004539* and *Bra039283* were aligned using the EMBL-EBI Clustal Omega Multiple Sequence Alignment tool. The results were visualised using *Jalview* software. The coding DNA sequences are in upper case and in colour whilst intronic sequences are in lower case and in grey.

7.5 JASPAR Binding Matrices

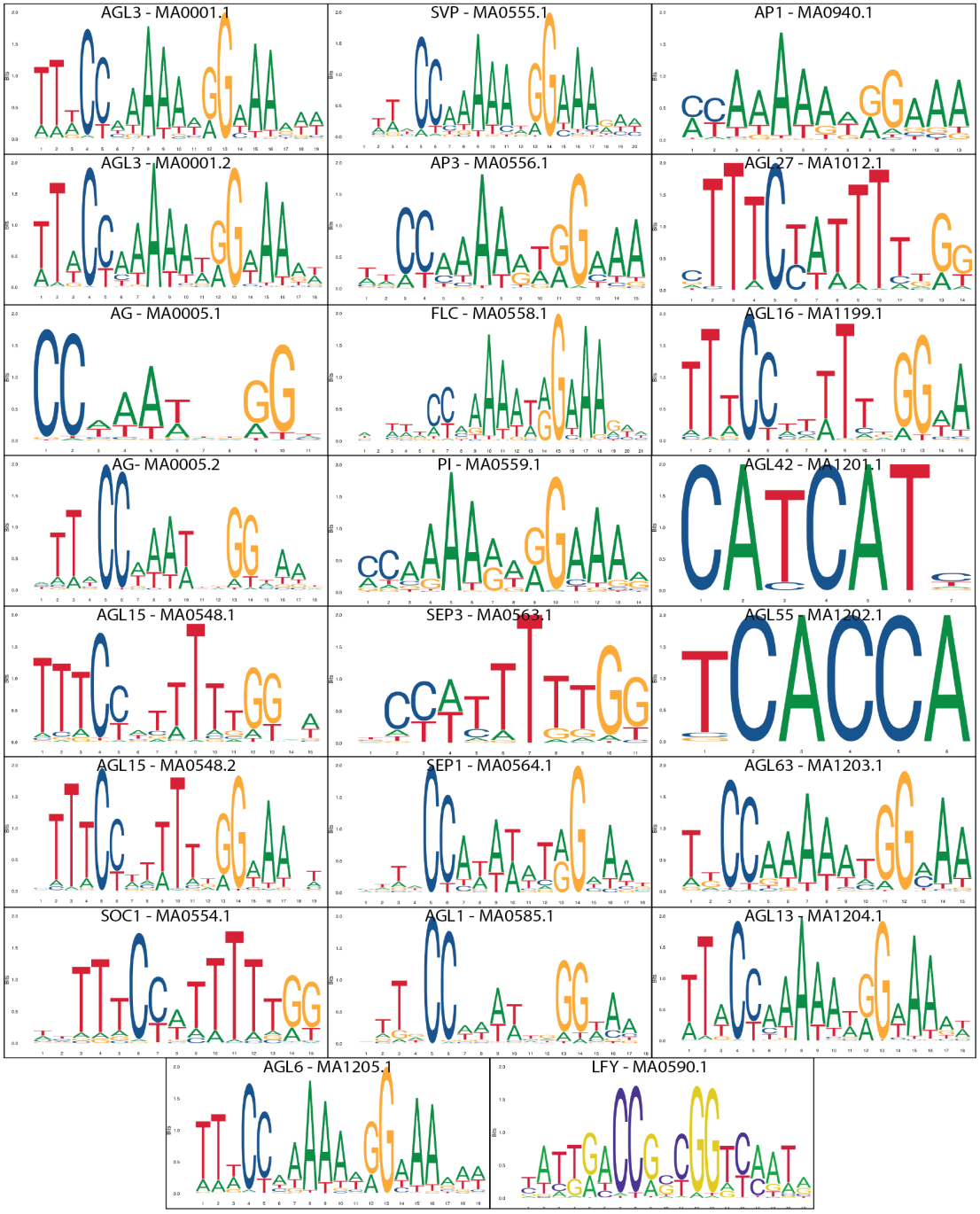

Figure 7.4: JASPAR binding site matrices used for analysis of sequences of *AtAG*, *BrAG1* and *BrAG2*. Sequences of the binding matrices from JASPAR used to predict transcription factor binding sites in the promoter and intronic sequences of *AtAG*, *BrAG1* and *BrAG2*.

Bibliography

Abe, M., Kobayashi, Y., Yamamoto, S., Daimon, Y., Yamaguchi, A., Ikeda, Y., Ichinoki, H., Notaguchi, M., Goto, K., & Araki, T. (2005). FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. *Science*, *309*(5737), 1052–1056. https://doi.org/10.1126/science.1115983

Alahakoon, U. I., Taheri, A., Nayidu, N. K., Epp, D., Yu, M., Parkin, I., Hegedus, D., Bonham-Smith, P., & Gruber, M. Y. (2016). Hairy Canola (Brasssica napus) revisited: Down-regulating TTG1 in an AtGL3-enhanced hairy leaf background improves growth, leaf trichome coverage, and metabolite gene expression diversity. *BMC Plant Biology*, *16*(1). https://doi.org/10.1186/s12870-015-0680-5

Alavanja, M. C. (2009). Pesticides use and exposure extensive worldwide. *Reviews on Environmental Health*, *24*(4), 303–309. https://doi.org/10.1515/REVEH.2009.24.4.303

Alvarez and Smyth. (2002). CRABS CLAW and SPATULA genes regulate growth and pattern formation during gynoecium development in Arabidopsis thaliana. *Internation Journal of Plant Sciences*, *163*(1), 17–41. https://doi.org/https://www.jstor.org/stable/10.1086/324178

Arteaga, N., Méndez-Vigo, B., Fuster-Pons, A., Savic, M., Murillo-Sánchez, A., Picó, F. X., & Alonso-Blanco, C. (2022). Differential environmental and genomic architectures shape the natural diversity for trichome patterning and morphology in different Arabidopsis organs. *Plant, Cell & Environment*, (February), 1–18. https://doi.org/10.1111/pce.14308

Arteaga, N., Savic, M., Méndez-Vigo, B., Fuster-Pons, A., Torres-Pérez, R., Oliveros, J. C., Picó, F. X., & Alonso-Blanco, C. (2021). MYB transcription factors drive

evolutionary innovations in Arabidopsis fruit trichome patterning. *The Plant Cell*, *33*(3), 548–565. https://doi.org/10.1093/plcell/koaa041

Azhakanandam, S., Nole-Wilson, S., Bao, F., & Franks, R. G. (2008). SEUSS and AINTEGUMENTA mediate patterning and ovule initiation during gynoecium medial domain development. *Plant Physiology*, *146*(3), 1165–1181. https://doi.org/10.1104/pp.107.114751

Balkunde, R., Deneer, A., Bechtel, H., Zhang, B., Herberth, S., Pesch, M., Jaegle, B., Fleck, C., & Hülskamp, M. (2020). Identification of the trichome patterning core network using data from weak ttg1 alleles to constrain the model space. *Cell Reports*, *33*(11). https://doi.org/10.1016/j.celrep.2020.108497

Becker, A., & Ehlers, K. (2016). Arabidopsis flower development - of protein complexes, targets, and transport. *Protoplasma*, *253*(2), 219–230. https://doi.org/10.1007/s00709-015-0812-7

Benková, E., Michniewicz, M., Sauer, M., Teichmann, T., Seifertová, D., Jürgens, G., & Friml, J. (2003). Local, efflux-dependent auxin gradients as a common module for plant organ formation. *Cell*, *115*(5), 591–602. https://doi.org/10.1016/S0092-8674(03)00924-3

Birchler, J. A., & Yang, H. (2022). The multiple fates of gene duplications: Deletion, hypofunctionalization, subfunctionalization, neofunctionalization, dosage balance constraints, and neutral variation. *The Plant Cell*, *34*(7), 2466–2474. https://doi.org/10.1093/plcell/koac076

Birchler, J. A., Yao, H., Chudalayandi, S., Vaiman, D., & Veitia, R. A. (2010). Heterosis. *The Plant Cell*, *22*(7), 2105–2112. https://doi.org/10.1105/tpc.110.076133

Blakeslee, J. J., Spatola Rossi, T., & Kriechbaumer, V. (2019). Auxin biosynthesis: Spatial regulation and adaptation to stress. *Journal of Experimental Botany*, *70*(19), 5041–5049. https://doi.org/10.1093/jxb/erz283

Blanc, G., & Wolfe, K. H. (2004a). Functional divergence of duplicated genes formed by polyploidy during Arabidopsis evolution. *The Plant Cell*, *16*(7), 1679–1691. https://doi.org/10.1105/tpc.021410

Blanc, G., & Wolfe, K. H. (2004b). Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. *The Plant Cell*, *16*(7), 1667–1678. https://doi.org/10.1105/tpc.021345

Bodnaryk, R. P. (1996). Physical and chemical defences of pods and seeds of white mustard (Sinapis alba L.) against tarnished plant bugs, Lygus lineolaris (Palisot de Beauvois) (Heteroptera: Miridae). *Canadian Journal of Plant Science*, *76*(1), 33–36. https://doi.org/10.4141/cjps96-006

Bowers, J. E., Chapman, B. A., Rong, J., & Paterson, A. H. (2003). Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature*, *422*(6930), 433–438. https://doi.org/10.1038/nature01521

Bowman, J. L., Smyth, D. R., & Meyerowitz, E. M. (1989). Genes directing flower development in Arabidopsis. *The Plant Cell*, *1*(1), 37–52. https://doi.org/10.1105/tpc.1.1.37

Bowman, J. L., Smyth, D. R., & Meyerowitz, E. M. (1991). Genetic interactions among floral homeotic genes of Arabidopsis. *Development*, *112*(1), 1–20. https://doi.org/10.1242/dev.112.1.1

Bowman, J. L., & Smyth, D. R. (1999). CRABS CLAW, a gene that regulates carpel and nectary development in Arabidopsis, encodes a novel protein with zinc finger and helix-loop-helix domains. *Development*, *126*(11), 2387–96. https://doi.org/10.1016/j.palaeo.2004.10.017

Bramsiepe, J., Wester, K., Weinl, C., Roodbarkelari, F., Kasili, R., Larkin, J. C., Hülskamp, M., & Schnittger, A. (2010). Endoreplication controls cell fate maintenance. *PLoS Genetics*, *6*(6), 1–14. https://doi.org/10.1371/journal.pgen.1000996

Briar, S. S., Shrestha, G., Sharma, A., Miller, J. H., & Reddy, G. V. (2019). Effect of nitrogen fertilization on flea beetle (Phyllotreta cruciferae) and cabbage seedpod weevil (Ceutorhynchus obstrictus) injury and crop yield in dry land canola. *Phytoparasitica*, *47*(5), 637–645. https://doi.org/10.1007/s12600-019-00762-x

Brockington, S. F., Yang, Y., Gandia-Herrero, F., Covshoff, S., Hibberd, J. M., Sage, R. F., Wong, G. K., Moore, M. J., & Smith, S. A. (2015). Lineage-specific gene

radiations underlie the evolution of novel betalain pigmentation in Caryophyllales. *New Phytologist*, *207*(4), 1170–1180. https://doi.org/10.1111/nph.13441

Busch, M. A., Bomblies, K., & Weigel, D. (1999). Activation of a floral homeotic gene in Arabidopsis. *Science*, *285*(5427), 585–587. https://doi.org/10.1126/science.285.5427.585

Cai, C., Wang, X., Liu, B., Wu, J., Liang, J., Cui, Y., Cheng, F., & Wang, X. (2017). Brassica rapa genome 2.0: A reference upgrade through sequence re-assembly and gene re-annotation. *Molecular Plant*, *10*(4), 649–651. https://doi.org/10.1016/j.molp.2016.11.008

Cao, X., Yang, H., Shang, C., Ma, S., Liu, L., & Cheng, J. (2019). The roles of auxin biosynthesis YUCCA gene family in plants. *International Journal of Molecular Sciences*, *20*, 1–19. https://doi.org/10.3390/ijms20246343

Castro-Mondragon, J. A., Riudavets-Puig, R., Rauluseviciute, I., Berhanu Lemma, R., Turchi, L., Blanc-Mathieu, R., Lucas, J., Boddie, P., Khan, A., Perez, N. M., Fornes, O., Leung, T. Y., Aguirre, A., Hammal, F., Schmelter, D., Baranasic, D., Ballester, B., Sandelin, A., Lenhard, B., … Mathelier, A. (2022). JASPAR 2022: The 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Research*, *50*, 165–173. https://doi.org/10.1093/nar/gkab1113

Causier, B., Schwarz-Sommer, Z., & Davies, B. (2010). Floral organ identity: 20 years of ABCs. *Seminars in Cell and Developmental Biology*, *21*(1), 73–79. https://doi.org/10.1016/j.semcdb.2009.10.005

Chang, W., Guo, Y., Zhang, H., Liu, X., & Guo, L. (2020). Same actor in different stages: Genes in shoot apical meristem maintenance and floral meristem determinacy in Arabidopsis. *Frontiers in Ecology and Evolution*, *8*, 1–12. https://doi.org/10.3389/fevo.2020.00089

Chen, Q., Payyavula, R. S., Chen, L., Zhang, J., Zhang, C., & Turgeon, R. (2018). FLOWERING LOCUS T mRNA is synthesized in specialized companion cells in Arabidopsis and Maryland Mammoth tobacco leaf veins. *Proceedings of the*

National Academy of Sciences of the United States of America, *115*(11), 2830–2835. https://doi.org/10.1073/pnas.1719455115

Churchman, M. L., Brown, M. L., Kato, N., Kirik, V., Hülskamp, M., Inzé, D., De Veylder, L., Walker, J. D., Zheng, U., Oppenheimer, D. G., Gwin, T., Churchman, J., & Larkin, J. C. (2006). SIAMESE, a plant-specific cell cycle regulator, controls endoreplication onset in Arabidopsis thaliana. *The Plant Cell*, *18*(11), 3145–3157. https://doi.org/10.1105/tpc.106.044834

Clark, J. W., & Donoghue, P. C. (2018). Whole-genome duplication and plant macroevolution. *Trends in Plant Science*, *23*(10), 933–945. https://doi.org/10.1016/j.tplants.2018.07.006

Clough, S. J., & Bent, A. F. (1998). Floral dip: a simplified method for Agrobacterium mediated transformation of Arabidopsis thaliana. *The Plant Journal*, *16*(6), 735–743. https://doi.org/10.1046/J.1365-313X.1998.00343.X

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

Colombo, M., Brambilla, V., Marcheselli, R., Caporali, E., Kater, M. M., & Colombo, L. (2010). A new role for the SHATTERPROOF genes during Arabidopsis gynoecium development. *Developmental Biology*, *337*(2), 294–302. https://doi.org/10.1016/j.ydbio.2009.10.043

Conner, J., & Liu, Z. (2000). LEUNIG, a putative transcriptional corepressor that regulates AGAMOUS expression during flower development. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(23), 12902–12907. https://doi.org/10.1073/pnas.230352397

Cui, W., Chen, Z., Shangguan, X., Li, T., Wang, L., Xue, X., & Cao, J. (2022). TRY intron2 determined its expression in inflorescence activated by SPL9 and MADS-box genes in Arabidopsis. *Plant Science*, *321*(May), 111311. https://doi.org/10.1016/j.plantsci.2022.111311

Daum, G., Medzihradszky, A., Suzaki, T., & Lohmann, J. U. (2014). A mechanistic framework for noncell autonomous stem cell induction in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(40), 14619–14624. https://doi.org/www.pnas.org/cgi/doi/10.1073/pnas.1406446111

Denay, G., Chahtane, H., Tichtinsky, G., & Parcy, F. (2017). A flower is born: an update on Arabidopsis floral meristem formation. *Current Opinion in Plant Biology*, *35*, 15–22. https://doi.org/10.1016/j.pbi.2016.09.003

Desvoyes, B., Ramirez-Parra, E., Xie, Q., Chua, N. H., & Gutierrez, C. (2006). Cell type-specific role of the retinoblastoma/E2F pathway during Arabidopsis leaf development. *Plant Physiology*, *140*(1), 67–80. https://doi.org/10.1104/pp.105.071027

Ditta, G., Pinyopich, A., Robles, P., Pelaz, S., & Yanofsky, M. F. (2004). The SEP4 gene of Arabidopsis thaliana functions in floral organ and meristem identity. *Curr Opin Plant Biol*, *14*, 1935–1940. https://doi.org/10.1016/j

Drews, G. N., Bowman, J. L., & Meyerowitz, E. M. (1991). Negative regulation of the Arabidopsis homeotic gene AGAMOUS by the APETALA2 product. *Cell*, *65*(6), 991–1002. https://doi.org/10.1016/0092-8674(91)90551-9

Edwards, K., Johnstone, C., & Thompson, C. (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Research*, *19*(6), 1349. https://doi.org/10.13031/aim.20141900217

Esch, J. J., Chen, M., Sanders, M., Hillestad, M., Ndkium, S., Idelkope, B., Neizer, J., & Marks, M. D. (2003). A contradictory GLABRA3 allele helps define gene interactions controlling trichome development in Arabidopsis. *Development*, *130*, 5885–5894. https://doi.org/10.1242/dev.00812

European Commission. (2013). Press Release - Bees & Pesticides: Commission to proceed with plan to better protect bees. *European Comission Press Release*, (April 29), 1–2. https://doi.org/PressRelease-Bees{\&}Pesticides:Commissiontoproceedwithplantobetterprotectbees.(April29),1âĂŞ2.

Favaro, R., Pinyopich, A., Battaglia, R., Kooiker, M., Borghi, L., Ditta, G., Yanofsky, M. F., Kater, M. M., & Colombo, L. (2003). MADS-box protein complexes control carpel and ovule development in Arabidopsis. *The Plant Cell*, *15*, 2603–11. https://doi.org/10.1105/tpc.015123

Ferrándiz, C., Gu, Q., Martienssen, R., & Yanofsky, M. F. (2000). Redundant regulation of meristem identity and plant architecture by FRUITFULL, APETALA1 and CAULIFLOWER. *Development*, (127), 725–734. https://doi.org/https://doi.org/10.1242/dev.127.4.725

Fletcher, J. C., Brand, U., Running, M. P., Simon, R., & Meyerowitz, E. M. (1999). Signaling of cell fate decisions by CLAVATA3 in Arabidopsis shoot. *Science*, *283*(March), 1911–1914. https://doi.org/10.1126/science.283.5409.1911

Fomeju, B. F., Falentin, C., Lassalle, G., Manzanares-Dauleux, M. J., & Delourme, R. (2015). Comparative genomic analysis of duplicated homoeologous regions involved in the resistance of Brassica napus to stem canker. *Frontiers in Plant Science*, *6*(September), 1–14. https://doi.org/10.3389/fpls.2015.00772

Franks, R. G., Wang, C., Levin, J. Z., & Liu, Z. (2002). SEUSS, a member of a novel family of plant regulatory proteins, represses floral homeotic gene expression with LEUNIG. *Development*, *129*(1), 253–263. https://doi.org/10.1242/dev.129.1.253

Friml, J., Vieten, A., Sauer, M., Weijers, D., Schwarz, H., Hamann, T., Offringa, R., & Jürgens, G. (2003). Efflux-dependent auxin gradients establish the apical-basal axis of Arabidopsis. *Nature*, *426*(6963), 147–153. https://doi.org/10.1038/nature02085

Friml, J., Yang, X., Michniewicz, M., Weijers, D., Quint, A., Tietz, O., Benjamins, R., Ouwerkerk, P. B., Ljung, K., Sandberg, G., Hooykaas, P. J., Palme, K., & Offringa, R. (2004). A PINOID-dependent binary switch in apical-basal PIN polar targeting directs auxin efflux. *Science*, *306*(5697), 862–865. https://doi.org/10.1126/science.1100618

Gan, L., Xia, K., Chen, J. G., & Wang, S. (2011). Functional characterization of TRICHOMELESS2, a new single-repeat R3 MYB transcription factor in the regulation of trichome patterning in Arabidopsis. *BMC Plant Biology*, *11*, 1–12. https://doi.org/10.1186/1471-2229-11-176

Goslin, K., Zheng, B., Serrano-Mislata, A., Rae, L., Ryan, P. T., Kwaśniewska, K., Thomson, B., Ó'Maoiléidigh, D. S., Madueño, F., Wellmer, F., & Graciet, E. (2017). Transcription factor interplay between LEAFY and APETALA1/CAULIFLOWER during floral initiation. *Plant Physiology*, *174*(2), 1097–1109. https://doi.org/10.1104/pp.17.00098

Goto, K., Kyozuka, J., & Bowman, J. L. (2001). Turning floral organs into leaves, leaves into floral organs. *Current Opinion in Genetics and Development*, *11*(4), 449–456. https://doi.org/10.1016/S0959-437X(00)00216-1

Green, J. M. (2018). The rise and future of glyphosate and glyphosate-resistant crops. *Pest Management Science*, *74*(5), 1035–1039. https://doi.org/10.1002/ps.4462

Gregis, V., Sessa, A., Colombo, L., & Kater, M. M. (2006). AGL24, SHORT VEGETATIVE PHASE, and APETALA1 redundantly control AGAMOUS during early stages of flower development in Arabidopsis. *The Plant Cell*, *18*(6), 1373–1382. https://doi.org/10.1105/tpc.106.041798

Gruber, M. Y., Wang, S., Ethier, S., Holowachuk, J., Bonham-Smith, P. C., Soroka, J., & Lloyd, A. (2006). "HAIRY CANOLA" - Arabidopsis GL3 induces a dense covering of trichomes on Brassica napus seedlings. *Plant Molecular Biology*, *60*(5), 679–698. https://doi.org/10.1007/s11103-005-5472-0

Gustafson-Brown, C., Savidge, B., & Yanofsky, M. F. (1994). Regulation of the Arabidopsis floral homeotic gene APETALA1. *Cell*, *76*(1), 131–143. https://doi.org/10.1016/0092-8674(94)90178-3

Hanano, S., & Goto, K. (2011). Arabidopsis TERMINAL FLOWER1 is involved in the regulation of flowering time and inflorescence development through transcriptional repression. *Plant Cell*, *23*(9), 3172–3184. https://doi.org/10.1105/tpc.111.088641

Hong, R. L., Hamaguchi, L., Busch, M. A., & Weigel, D. (2003). Regulatory elements of the floral homeotic gene AGAMOUS identified by phylogenetic footprinting and shadowing. *The Plant Cell*, *15*(6), 1296–1309. https://doi.org/10.1105/tpc.009548

Honma, T., & Goto, K. (2000). The Arabidopsis floral homeotic gene PISTILLATA is regulated by discrete cis-elements responsive to induction and maintenance signals. *Development*, *127*(10), 2021–2030. https://doi.org/10.1242/dev.127.10.2021

Honma, T., & Goto, K. (2001). Complexes of MADS-box proteins are suffcient to convert leaves into floral organs. *Nature*, *409*, 525–529. https://doi.org/10.1038/35054083

Hülskamp, M., Miséra, S., & Jürgens, G. (1994). Genetic dissection of trichome cell development in Arabidopsis. *Cell*, *76*(3), 555–566. https://doi.org/10.1016/0092-8674(94)90118-X

Inoue, H., Nojima, H., & Okayama, H. (1990). High efficiency transformation of Escherichia coli with plasmids. *Gene*, *96*(1), 23–28. https://doi.org/10.1016/0378-1119(90)90336-P

Irish, V. F., & Sussex, I. M. (1990). Function of the apetala-1 gene during Arabidopsis floral development. *The Plant Cell*, *2*(8), 741–753. https://doi.org/10.1105/tpc.2.8.741

Ishida, T., Kurata, T., Okada, K., & Wada, T. (2008). A genetic regulatory network in the development of trichomes and root hairs. *Annual Review of Plant Biology*, *59*(1), 365–386. https://doi.org/10.1146/annurev.arplant.59.032607.092949

Ito, T., Wellmer, F., Yu, H., Das, P., Ito, N., Alves-Ferreira, M., Riechmann, J. L., & Meyerowitz, E. M. (2004). The homeotic protein AGAMOUS controls microsporogenesis by regulation of SPOROCYTELES. *Nature*, *430*, 356–360. https://doi.org/10.1038/nature02733

Ito, T., Ng, K. H., Lim, T. S., Yu, H., & Meyerowitz, E. M. (2007). The homeotic protein AGAMOUS controls late stamen development by regulating a jasmonate

biosynthetic gene in Arabidopsis. *The Plant Cell*, *19*(11), 3516–3529. https://doi.org/10.1105/tpc.107.055467

Jang, S., Marchal, V., Panigrahi, K. C., Wenkel, S., Soppe, W., Deng, X. W., Valverde, F., & Coupland, G. (2008). Arabidopsis COP1 shapes the temporal pattern of CO accumulation conferring a photoperiodic flowering response. *EMBO Journal*, *27*(8), 1277–1288. https://doi.org/10.1038/emboj.2008.68

Ji, L., Liu, X., Yan, J., Wang, W., Yumul, R. E., Kim, Y. J., Dinh, T. T., Liu, J., Cui, X., Zheng, B., Agarwal, M., Liu, C., Cao, X., Tang, G., & Chen, X. (2011). ARGONAUTE10 and ARGONAUTE1 regulate the termination of floral stem cells through two microRNAs in Arabidopsis. *PLoS Genetics*, *7*(3), 1–14. https://doi.org/10.1371/journal.pgen.1001358

Jiang, W. K., Liu, Y. L., Xia, E. H., & Gao, L. Z. (2013). Prevalent role of gene features in determining evolutionary fates of whole-genome duplication duplicated genes in flowering plants. *Plant Physiology*, *161*(4), 1844–1861. https://doi.org/10.1104/pp.112.200147

Johansson, M., & Staiger, D. (2015). Time to flower: Interplay between photoperiod and the circadian clock. *Journal of Experimental Botany*, *66*(3), 719–730. https://doi.org/10.1093/jxb/eru441

Johnston, J. S., Pepper, A. E., Hall, A. E., Chen, Z. J., Hodnett, G., Drabek, J., Lopez, R., & Price, H. J. (2005). Evolution of genome size in Brassicaceae. *Annals of Botany*, *95*(1), 229–235. https://doi.org/10.1093/aob/mci016

Karabourniotis, G., Kyparissis, A., & Yiannis, M. A. N. E. T. A. S. (1993). Leaf hairs of Olea europaea protect underlying tissues against ultraviolet-B radiation damage. *Environmental and Experimental Botany*, *33*(3), 341–345. https://doi.org/10.1016/0098-8472(93)90035-E

Kaufmann, K., Wellmer, F., Muiñ, J. M., Ferner, T., Wuest, S. E., Kumar, V., Serrano-Mislata, A., Madueño, F., Kraiewski, P., Meyerowitz, E. M., Angenent, G. C., & Riechmann, J. L. (2010). Orchestration of floral initiation by APETALA1. *Science*, *328*(5974), 85–89. https://doi.org/10.1126/science.1185244

Kessler, A., & Baldwin, I. T. (2002). Plant responses to insect herbivory: The emerging molecular analysis. *Annual Review of Plant Biology*, *53*, 299–328. https://doi.org/10.1146/annurev.arplant.53.100301.135207

Kierzkowski, D., Lenhard, M., Smith, R., & Kuhlemeier, C. (2013). Interaction between meristem tissue layers controls phyllotaxis. *Developmental Cell*, *26*(6), 616–628. https://doi.org/10.1016/j.devcel.2013.08.017

Kim, C. K., Seol, Y. J., Perumal, S., Lee, J., Waminal, N. E., Jayakodi, M., Lee, S. C., Jin, S., Choi, B. S., Yu, Y., Ko, H. C., Choi, J. W., Ryu, K. Y., Sohn, S. H., Parkin, I., & Yang, T. J. (2018). Re-exploration of U's triangle Brassica species based on chloroplast genomes and 45S nrDNA sequences. *Scientific Reports*, *8*(1), 1–11. https://doi.org/10.1038/s41598-018-25585-4

Kinoshita, A., & Richter, R. (2020). Genetic and molecular basis of floral induction in Arabidopsis thaliana. *Journal of Experimental Botany*, *71*(9), 2490–2504. https://doi.org/10.1093/JXB/ERAA057

Kirik, V., Bouyer, D., Schöbinger, U., Bechtold, N., Herzog, M., Bonneville, J. M., & Hülskamp, M. (2001). CPR5 is involved in cell proliferation and cell death control and encodes a novel transmembrane protein. *Current Biology*, *11*(23), 1891–1895. https://doi.org/10.1016/S0960-9822(01)00590-5

Kirik, V., Simon, M., Huelskamp, M., & Schiefelbein, J. (2004). The ENHANCER of TRY and CPC1 gene acts redundantly with TRIPTYCHON and CAPRICE in trichome and root hair cell patterning in Arabidopsis. *Developmental Biology*, *268*(2), 506–513. https://doi.org/10.1016/j.ydbio.2003.12.037

Kirik, V., Simon, M., Wester, K., Schiefelbein, J., & Hulskamp, M. (2004). Enhancer of TRY and CPC 2 (ETC2) reveals redundancy in the region-specific control of trichome development of Arabidopsis. *Plant Molecular Biology*, *55*(3), 389–398. https://doi.org/10.1007/s11103-004-0893-8

Klucher, K. M., Chow, H., Reiser, L., & Fischer, R. L. (1996). The AINTEGUMENTA gene of Arabidopsis required for ovule and female gametophyte development is

related to the floral homeotic gene APETALA2. *The Plant Cell*, *8*(2), 137–153. https://doi.org/10.1105/tpc.8.2.137

Krizek & Fletcher, J. (2005). Molecular mechanisms of flower development: An armchair guide. *Nature Reviews Genetics*, *6*(9), 688–698. https://doi.org/10.1038/nrg1675

Kwaśniewska, K., Breathnach, C., Fitzsimons, C., Goslin, K., Thomson, B., Beegan, J., Finocchio, A., Prunet, N., Ó'Maoiléidigh, D. S., & Wellmer, F. (2021). Expression of KNUCKLES in the stem cell domain is required for its function in the control of floral meristem activity in Arabidopsis. *Frontiers in Plant Science*, *12*(July), 1–14. https://doi.org/10.3389/fpls.2021.704351

Kwon, C. S., Chen, C., & Wagner, D. (2005). WUSCHEL is a primary target for transcriptional regulation by SPLAYED in dynamic control of stem cell fate in Arabidopsis. *Genes and Development*, *19*(8), 992–1003. https://doi.org/10.1101/gad.1276305

Lai, X., Vega-Léon, R., Hugouvieux, V., Blanc-Mathieu, R., van der Wal, F., Lucas, J., Silva, C. S., Jourdain, A., Muino, J. M., Nanao, M. H., Immink, R., Kaufmann, K., Parcy, F., Smaczniak, C., & Zubieta, C. (2021). The intervening domain is required for DNA-binding and functional identity of plant MADS transcription factors. *Nature Communications*, *12*(1), 1–13. https://doi.org/10.1038/s41467-021-24978-w

Lam, W. K. F., & Pedigo, L. P. (2001). Effect of trichome density on soybean pod feeding by adult bean leaf beetles (Coleoptera: Chrysomelidae). *Journal of Economic Entomology*, *94*(6), 1459–1463. https://doi.org/10.1603/0022-0493-94.6.1459

Lampropoulos, A., Sutikovic, Z., Wenzl, C., Maegele, I., Lohmann, J. U., & Forner, J. (2013). GreenGate - A novel, versatile, and efficient cloning system for plant transgenesis. *PLoS ONE*, *8*(12). https://doi.org/10.1371/journal.pone.0083043

Larkin, J. C., Young, N., Prigge, M., & Marks, M. D. (1996). The control of trichome spacing and number in Arabidopsis. *Development*, *122*, 997–1005. https://doi.org/10.1242/dev.122.3.997

Laubinger, S., Marchal, V., Le Gourrierec, J., Wenkel, S., Andrian, J., Jang, S., Kulajta, C., Braun, H., Coupland, G., & Hoecker, U. (2006). Arabidopsis SPA proteins regulate photoperiodic flowering and interact with the floral inducer CONSTANS to regulate its stability. *Development*, *133*(22), 4608. https://doi.org/10.1242/dev.02691

Laux, T., Mayer, K. F. X., Berger, J., Jürgens, G., Genetik, L., & München, L.-m.-u. (1996). The WUSCHEL gene is required for shoot and floral meristem integrity in Arabidopsis. *Development*, 87–96. https://doi.org/10.1242/dev.122.1.87

Lee, Baum, S. F., Alvarez, J., Patel, A., Chitwood, D. H., & Bowman, J. L. (2005). Activation of CRABS CLAW in the nectaries and carpels of Arabidopsis. *The Plant Cell*, *17*(1), 25–36. https://doi.org/10.1105/tpc.104.026666

Lee, Hirakawa, T., Yamaguchi, N., & Ito, T. (2019). The roles of plant hormones and their interactions with regulatory genes in determining meristem activity. *International Journal of Molecular Sciences*, *20*(16), 4065. https://doi.org/10.3390/ijms20164065

Lenhard, M., Bohnert, A., Ju, G., Laux, T., Morgenstelle, A. D., & Tu, D. (2001). Termination of stem cell maintenance in Arabidopsis floral meristems by interactions between WUSCHEL and AGAMOUS. *Cell*, *105*, 805–814. https://doi.org/10.1105/tpc.111.091538

Levin, D. A. (1973). The role of trichomes in plant defense. *The Quarterly Review of Biology*, *48*(1), 3–15. https://doi.org/10.1086/407484

Li, Kitashiba, H., Inaba, K., & Nishio, T. (2009). A Brassica rapa Linkage Map of EST-based SNP Markers for Identification of Candidate Genes Controlling Flowering Time and Leaf Morphological Traits. *DNA Research*, *16*(6), 311–323. https://doi.org/10.1093/dnares/dsp020

Li, W., Zhou, Y., Liu, X., Yu, P., Cohen, J. D., & Meyerowitz, E. M. (2013). LEAFY controls auxin response pathways in floral primordium formation. *Science Signaling*, *6*(270). https://doi.org/10.1126/scisignal.2003937

Liu, Chen, Er, Soo, & Kumar. (2008). Direct interaction of AGL24 and SOC1 integrates flowering signals in Arabidopsis. *Development*, *135*(8), 1481–1491. https://doi.org/10.1242/dev.020255

Liu, Franks, R., & Klink, V. (2000). Regulation of gynoecium marginal tissue formation by LEUNIG and AINTEGUMENTA. *The Plant Cell*, *12*(10), 1879–1891. https://doi.org/10.1105/tpc.12.10.1879

Liu, Kim, Y. J., Müller, R., Yumul, R. E., Liu, C., Pan, Y., Cao, X., Goodrich, J., & Chen, X. (2011). AGAMOUS terminates floral stem cell maintenance in Arabidopsis by directly repressing WUSCHEL through recruitment of Polycomb Group proteins. *The Plant Cell*, *23*(10), 3654–3670. https://doi.org/10.1105/tpc.111.091538

Lorković, Z. J., Wieczorek Kirk, D. A., Lambermon, M. H., & Filipowicz, W. (2000). Pre-mRNA splicing in higher plants. *Trends in Plant Science*, *5*(4), 160–167. https://doi.org/10.1016/S1360-1385(00)01595-8

Lysak, M. A., Koch, M. A., Pecinka, A., & Schubert, I. (2005). Chromosome triplication found across the tribe Brassiceae. *Genome Research*, *15*(4), 516–525. https://doi.org/10.1101/gr.3531105

Ma, C. S., Zhang, W., Peng, Y., Zhao, F., Chang, X. Q., Xing, K., Zhu, L., Ma, G., Yang, H. P., & Rudolf, V. H. (2021). Climate warming promotes pesticide resistance through expanding overwintering range of a global pest. *Nature Communications*, *12*(1), 1–10. https://doi.org/10.1038/s41467-021-25505-7

Madeira, F., Pearce, M., Tivey, A. R., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A., & Lopez, R. (2022). Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Research*, *50*(W1), W276–W279. https://doi.org/10.1093/nar/gkac240

Maes, L., & Goossens, A. (2010). Hormone-mediated promotion of trichome initiation in plants is conserved but utilizes species- and trichome-specific regulatory mechanisms. *Plant Signaling and Behavior*, *5*(2), 205–207. https://doi.org/10.4161/psb.5.2.11214

Matthes, M. S., Best, N. B., Robil, J. M., Malcomber, S., Gallavotti, A., & McSteen, P. (2019). Auxin EvoDevo: Conservation and diversification of genes regulating auxin biosynthesis, transport, and signaling. *Molecular Plant*, *12*(3), 298–320. https://doi.org/10.1016/j.molp.2018.12.012

Mayer, K. F., Schoof, H., Haecker, A., Lenhard, M., Jürgens, G., & Laux, T. (1998). Role of WUSCHEL in regulating stem cell fate in the Arabidopsis shoot meristem. *Cell*, *95*(6), 805–815. https://doi.org/10.1016/S0092-8674(00)81703-1

Mayr, C. (2019). What are 3UTRs doing? *Cold Spring Harbor Perspectives in Biology*, *11*(10). https://doi.org/10.1101/cshperspect.a034728

McBride, K. E., & Summerfelt, K. R. (1990). Improved binary vectors for Agrobacterium-mediated plant transformation. *Plant Molecular Biology*, *14*(2), 269–276. https://doi.org/10.1007/BF00018567

McCallum, C. M., Comai, L., Greene, E. A., & Henikoff, S. (2000). Targeted screening for induced mutations. *Nature Biotechnology*, *18*, 455–457. https://doi.org/10.1038/74542

Mendes, M. a., Guerra, R. F., Berns, M. C., Manzo, C., Masiero, S., Finzi, L., Kater, M. M., & Colombo, L. (2013). MADS domain transcription factors mediate short-range DNA looping that is essential for target gene expression in Arabidopsis. *The Plant Cell*, *25*(7), 2560–2572. https://doi.org/10.1105/tpc.112.108688

Meyerowitz, E. M., & Pruitt, R. E. (1985). Arabidopsis thaliana and plant molecular genetics. *Science*, *229*(4719), 1214–1218. https://doi.org/10.1126/science.229.4719.1214

Mignone, F., Gissi, C., Liuni, S., & Pesole, G. (2002). Untranslated regions of mRNAs. *Genome Biology*, *3*(3), 1–10. https://doi.org/10.1186/gb-2002-3-3-reviews0004

Mizukami, Y., Huang, H., Tudor, M., Hu, Y., & Ma, H. (1996). Functional domains of the floral regulator AGAMOUS: Characterization of the DNA binding domain and analysis of dominant negative mutations. *The Plant Cell*, *8*(5), 831–845. https://doi.org/10.1105/tpc.8.5.831

Mizukami, Y., & Ma, H. (1992). Ectopic expression of the floral homeotic gene AGAMOUS in transgenic Arabidopsis plants alters floral organ identity. *Cell*, *71*, 119–131. https://doi.org/10.1016/0092-8674(92)90271-d

Mockaitis, K., & Estelle, M. (2008). Auxin receptors and plant development: A new signaling paradigm. *Annual Review of Cell and Developmental Biology*, *24*, 55–80. https://doi.org/10.1146/annurev.cellbio.23.090506.123214

Morohashi, K., & Grotewold, E. (2009). A systems approach reveals regulatory circuitry for Arabidopsis trichome initiation by the GL3 and GL1 selectors. *PLoS Genetics*, *5*(2). https://doi.org/10.1371/journal.pgen.1000396

Mun, J. H., Kwon, S. J., Yang, T. J., Seol, Y. J., Jin, M., Kim, J. A., Lim, M. H., Kim, J. S., Baek, S., Choi, B. S., Yu, H. J., Kim, D. S., Kim, N., Lim, K. B., Lee, S. I., Hahn, J. H., Lim, Y. P., Bancroft, I., & Park, B. S. (2009). Genome-wide comparative analysis of the Brassica rapa gene space reveals genome shrinkage and differential loss of duplicated genes after whole genome triplication. *Genome Biology*, *10*(10). https://doi.org/10.1186/gb-2009-10-10-r111

Nagaharu, U. (1935). Genome analysis in Brassica with special reference to the experimental formation of B. napus and peculiar mode of fertilization. *Jpn J Bot*, (7), 389–452.

Neff, M. M., Neff, J. D., Chory, J., & Pepper, A. E. (1998). dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: Experimental applications in Arabidopsis thaliana genetics. *Plant Journal*, *14*(3), 387–392. https://doi.org/10.1046/j.1365-313X.1998.00124.x

Nole-Wilson, S., & Krizek, B. A. (2006). AINTEGUMENTA contributes to organ polarity and regulates growth of lateral organs in combination with YABBY genes. *Plant Physiology*, *141*(3), 977–987. https://doi.org/10.1104/pp.106.076604

Obermeier, C., Mason, A. S., Meiners, T., Petschenka, G., Rostás, M., Will, T., Wittkop, B., & Austel, N. (2022). Perspectives for integrated insect pest protection in oilseed rape breeding. *Theoretical and Applied Genetics*, *135*(11), 3917–3946. https://doi.org/10.1007/s00122-022-04074-3

Ó'Maoiléidigh, D. S., Graciet, E., & Wellmer, F. (2014). Gene networks controlling Arabidopsis thaliana flower development. *New Phytologist*, 1–15. https://doi.org/10.1111/nph.12444

Ó'Maoiléidigh, D. S., Stewart, D., Zheng, B., Coupland, G., & Wellmer, F. (2018). Floral homeotic proteins modulate the genetic program for leaf development to suppress trichome formation in flowers. *Development*, *145*, 1–12. https://doi.org/10.1242/dev.157784

Ó'Maoiléidigh, D. S., Wuest, S. E., Rae, L., Raganelli, A., Ryan, P. T., Kwaśniewska, K., Das, P., Lohan, A. J., Loftus, B., Graciet, E., & Wellmer, F. (2013). Control of reproductive floral organ identity specification in Arabidopsis by the C function regulator AGAMOUS. *The Plant Cell*, *25*(7), 2482–2503. https://doi.org/10.1105/tpc.113.113209

Oppenheimer, D. G., Herman, P. L., Sivakumaran, S., Each, J., & Marks, M. D. (1991). A myb gene required for leaf trichome differentiation in Arabidopsis is expressed in stipules. *Cell*, *67*, 483–493. https://doi.org/10.1016/0092-8674(91)90523-2

Orzech, K. M., & Nichter, M. (2008). From resilience to resistance: Political ecological lessons from antibiotic and pesticide resistance. *Annual Review of Anthropology*, *37*, 267–282. https://doi.org/10.1146/annurev.anthro.37.081407.085205

Pajoro, A., Madrigal, P., Muiño, J. M., Matus, J. T., Jin, J., Mecchia, M. A., Debernardi, J. M., Palatnik, J. F., Balazadeh, S., Arif, M., Ó'Maoiléidigh, D. S., Wellmer, F., Krajewski, P., Riechmann, J. L., Angenent, G. C., & Kaufmann, K. (2014). Dynamics of chromatin accessibility and gene regulation by MADS-domain transcription factors in flower development. *Genome Biology*, *15*(3). https://doi.org/10.1186/gb-2014-15-3-r41

Panchy, N., Lehti-Shiu, M., & Shiu, S. H. (2016). Evolution of gene duplication in plants. *Plant Physiology*, *171*(4), 2294–2316. https://doi.org/10.1104/pp.16.00523

Payne, Johnson, S. D., & Koltunow, A. M. (2004). KNUCKLES (KNU) encodes a C2H2 zinc-finger protein that regulates development of basal pattern elements of the Arabidopsis gynoecium. *Development*, *131*(15), 3737–3749. https://doi.org/10.1242/dev.01216

Payne, Zhang, F., & Lloyd, A. M. (2000). GL3 encodes a bHLH protein that regulates trichome development in Arabidopsis through interaction with GL1 and TTG1. *Genetics*, *156*(3), 1349–1362. https://doi.org/10.1093/genetics/156.3.1349

Pelaz, S., Ditta, G. S., Baumann, E., Wisman, E., & Yanofsky, M. F. (2000). B and C floral organ identity functions require SEPALLATA MADS-box genes. *Nature*, *405*(5), 200–203. https://doi.org/10.1038/35012103

Pelaz, S., Tapia-López, R., Alvarez-Buylla, E. R., & Yanofsky, M. F. (2001). Conversion of leaves into petals in Arabidopsis. *Current Biology*, *11*(3), 182–184. https://doi.org/10.1016/S0960-9822(01)00024-0

Perazza, D., Herzog, M., Hülskamp, M., Brown, S., Dorne, A. M., & Bonneville, J. M. (1999). Trichome cell growth in Arabidopsis thaliana can be derepressed by mutations in at least five genes. *Genetics*, *152*(1), 461–476. https://doi.org/10.1093/genetics/152.1.461

Périlleux, C., Bouché, F., Randoux, M., & Orman-Ligeza, B. (2019). Turning meristems into fortresses. *Trends in Plant Science*, *24*(5), 431–442. https://doi.org/10.1016/j.tplants.2019.02.004

Popp, J., Pető, K., & Nagy, J. (2013). Pesticide productivity and food security.: A review. *Agronomy for Sustainable Development*, *33*(1), 243–255. https://doi.org/10.1007/s13593-012-0105-x

Quiros, C. F., Grellet, F., Sadowski, J., Suzuki, T., Li, G., & Wroblewski, T. (2001). Arabidopsis and Brassica comparative genomics: Sequence, structure and gene content in the ABI1-Rps2-Ck1 chromosomal segment and related regions. *Genetics*, *157*(3), 1321–1330. https://doi.org/10.1093/genetics/157.3.1321

Reinhardt, D., Pesce, E. R., Stieger, P., Mandel, T., Baltensperger, K., Bennett, M., Traas, J., Friml, J., & Kuhlemeier, C. (2003). Regulation of phyllotaxis by po-

lar auxin transport. *Nature*, *426*(6964), 255–260. https://doi.org/10.1038/nature02081

Riechmann, J. L., Ito, T., & Meyerowitz, E. M. (1999). Non-AUG initiation of AGAMOUS mRNA translation in Arabidopsis thaliana. *Molecular and Cellular Biology*, *19*(12), 8505–8512. https://doi.org/10.1128/mcb.19.12.8505

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

Rieu, P., Turchi, L., Thévenon, E., Zarkadas, E., Nanao, M., Chahtane, H., Tichtinsky, G., Lucas, J., Blanc-Mathieu, R., Zubieta, C., Schoehn, G., & Parcy, F. (2023). The F-box protein UFO controls flower development by redirecting the master transcription factor LEAFY to new cis-elements. *Nature Plants*, *9*(2), 315–329. https://doi.org/10.1038/s41477-022-01336-2

Sablowski, R. (2007). Flowering and determinacy in Arabidopsis. *Journal of Experimental Botany*, *58*(5), 889–907. https://doi.org/10.1093/jxb/erm015

Saddic, L. A., Huvermann, B., Bezhani, S., Su, Y., Winter, C. M., Kwon, C. S., Collum, R. P., & Wagner, D. (2006). The LEAFY target LMI1 is a meristem identity regulator and acts together with LEAFY to regulate expression of CAULIFLOWER. *Development*, *133*(9), 1673–1682. https://doi.org/10.1242/dev.02331

Sassi, M., & Vernoux, T. (2013). Auxin and self-organization at the shoot apical meristem. *Journal of Experimental Botany*, *64*(9), 2579–2592. https://doi.org/10.1093/jxb/ert101

Savidge, B., Rounsley, S. D., & Yanofsky, M. F. (1995). Temporal relationship between the transcription of two Arabidopsis MADS box genes and the floral organ identity genes. *The Plant Cell*, *7*(6), 721–733. https://doi.org/10.2307/3870174

Sawa, S., Watanabe, K., Goto, K., Kanaya, E., Morita, E. H., & Okada, K. (1999). FILAMENTOUS FLOWER, a meristem and organ identity gene of Arabidopsis,

encodes a protein with a zinc finger and HMG-related domains. *Genes and Development*, *13*(17), 2337. https://doi.org/10.1101/gad.13.9.1079

Schellmann, S., Schnittger, A., Kirik, V., Wada, T., Okada, K., Beerman, A., Thumfahrt, J., Jürgens, G., & Hülskamp, M. (2002). TRIPTYCHON and CAPRICE mediate lateral inhibition during trichome and root hair patterning in Arabidopsis. *EMBO Journal*, *21*(19), 5036–5046. https://doi.org/10.1093/emboj/cdf524

Schiessl, S. (2020). Regulation and subfunctionalization of flowering time genes in the allotetraploid oil crop Brassica napus. *Frontiers in Plant Science*, *11*, 1–11. https://doi.org/10.3389/fpls.2020.605155

Schnittger, A., & Hülskamp, M. (2002). Trichome morphogenesis: A cell-cycle perspective. *Philosophical Transactions of the Royal Society B*, *357*, 823–826. https://doi.org/10.1098/rstb.2002.1087

Schoof, H., Lenhard, M., Haecker, A., Mayer, K. F., Jürgens, G., & Laux, T. (2000). The stem cell population of Arabidopsis shoot meristems is maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. *Cell*, *100*(6), 635–644. https://doi.org/10.1016/S0092-8674(00)80700-X

Scotto-Lavino, E., Du, G., & Frohman, M. A. (2007). 3 End cDNA amplification using classic RACE. *Nature Protocols*, *1*(6), 2742–2745. https://doi.org/10.1038/nprot.2006.481

Searle, I., He, Y., Turck, F., Vincent, C., Fornara, F., Kröber, S., Amasino, R. A., & Coupland, G. (2006). The transcription factor FLC confers a flowering response to vernalization by repressing meristem competence and systemic signaling in Arabidopsis. *Genes and Development*, *20*(7), 898–912. https://doi.org/10.1101/gad.373506

Shang, E., Ito, T., & Sun, B. (2019). Control of floral stem cell activity in Arabidopsis. *Plant Signaling and Behavior*, *14*(11). https://doi.org/10.1080/15592324.2019.1659706

Shannon, S., & Meeks-Wagner, D. R. (1991). A mutation in the Arabidopsis TFL1 gene affects inflorescence meristem development. *The Plant Cell*, *3*(9), 877–892. https://doi.org/https://doi.org/10.1105/tpc.3.9.877

Shin, Y. H., Lee, H. M., & Park, Y. D. (2022). CRISPR/Cas9-mediated editing of AGAMOUS-like genes results in a late-bolting phenotype in chinese cabbage (Brassica rapa ssp. pekinensis). *International Journal of Molecular Sciences*, *23*(23). https://doi.org/10.3390/ijms232315009

Shpanev, A. M., & Moseyko, A. G. (2021). Cruciferous flea beetles (Phyllotreta spp.; Coleoptera, Chrysomelidae) in spring rape crops in leningrad province. *Entomological Review*, *101*(2), 174–180. https://doi.org/10.1134/S0013873821020032

Sitaraman, J., Bui, M., & Liu, Z. (2008). LEUNIG_Homolog and LEUNIG perform partially redundant functions during Arabidopsis embryo and floral development. *Plant Physiology*, *147*(2), 672–681. https://doi.org/10.1104/pp.108.115923

Smaczniak, C., Immink, R. G. H., Muiño, J. M., Blanvillain, R., Busscher, M., Busscher-Lange, J., Dinh, Q. D. P., Liu, S., Westphal, A. H., Boeren, S., et al. (2012). Characterization of MADS domain transcription factor complexes in Arabidopsis flower development. *Proceedings of the National Academy of Sciences*, *109*(5), 1560–1565. https://doi.org/10.1073/pnas.1112871109

Smyth, D. R., Bowman, J. L., & Meyerowitz, E. M. (1990). Early flower development in Arabidopsis. *The Plant Cell*, *2*(8), 755–767. https://doi.org/10.1105/tpc.2.8.755

Somssich, M., Je, B. I., Simon, R., & Jackson, D. (2016). CLAVATA-WUSCHEL signaling in the shoot meristem. *Development*, *143*(18), 3238–3248. https://doi.org/10.1242/dev.133645

Song, X.-M., Wang, J.-P., Sun, P.-C., Ma, X., Yang, Q.-H., Hu, J.-J., Sun, S.-R., Li, Y.-X., Yu, J.-G., Feng, S.-Y., Pei, Q.-Y., Yu, T., Yang, N.-S., Liu, Y.-Z., Li, X.-Q., Paterson, A. H., & Wang, X.-Y. (2020). Preferential gene retention increases the robustness of cold regulation in Brassicaceae and other plants after polyploidization. *Horticulture Research*, *7*(1). https://doi.org/10.1038/s41438-020-0253-0

Soroka, J. J., Holowachuk, J. M., Gruber, M. Y., & Grenkow, L. F. (2011). Feeding by flea beetles (Coleoptera: Chrysomelidae; Phyllotreta spp.) is decreased on canola (Brassica napus) seedlings with increased trichome density. *Journal of Economic Entomology*, *104*(1), 125–136. https://doi.org/10.1603/EC10151

Srivastava, A. K., Lu, T., Zinta, G., Lang, Z., & Zhu, J.-K. (2018). UTR dependent control of gene expression in plants. *Trends in Plant Science*, *23*(3), 248–259. https://doi.org/10.1016/j.tplants.2017.11.003.UTR

Stephenson, P., Baker, D., Girin, T., Perez, A., Amoah, S., King, G. J., & Østergaard, L. (2010). A rich TILLING resource for studying gene function in Brassica rapa. *BMC Plant Biology*, *10*, 1–10. https://doi.org/10.1186/1471-2229-10-62

Sun, B., Xu, Y., Ng, K. H., & Ito, T. (2009). A timing mechanism for stem cell maintenance and differentiation in the Arabidopsis floral meristem. *Genes and Development*, *23*(15), 1791–1804. https://doi.org/10.1101/gad.1800409

Sun, B., Zhou, Y., Cai, J., Shang, E., Yamaguchi, N., Xiao, J., Looi, L. S., Wee, W. Y., Gao, X., Wagner, D., & Ito, T. (2019). Integration of transcriptional repression and polycomb-mediated silencing of WUSCHEL in floral meristems. *The Plant Cell*, *31*(7), 1488–1505. https://doi.org/10.1105/tpc.18.00450

Szymanski, D. B., Lloyd, A. M., & Marks, M. D. (2000). Progress in the molecular genetic analysis of trichome initiation and morphogenesis in Arabidopsis. *Trends in Plant Science*, *5*(5), 214–219. https://doi.org/10.1016/S1360-1385(00)01597-1

Telfer, A., Bollman, K. M., & Poethig, R. S. (1997). Phase change and the regulation of trichome distribution in Arabidopsis thaliana. *Development*, *124*(3), 645–654. https://doi.org/10.1242/dev.124.3.645

Teng, J., Wang, J., Zhang, L., Wei, C., Shen, S., Xiao, Q., Yue, Y., Hao, Y., Ge, W., & Wang, J. (2022). Paleopolyploidies and genomic fractionation in major eudicot clades. *Frontiers in Plant Science*, *13*, 1–15. https://doi.org/10.3389/fpls.2022.883140

The Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature*, *408*(6814), 796–815. https://doi.org/10.1038/35048692

Theißen, G. (2001). Development of floral organ identity : stories from the MADS house. *Curr Opin Plant Biol*, *4*, 75–85. https://doi.org/10.1016/s1369-5266(00)00139-4

Tominaga, R., Iwata, M., Sano, R., Inoue, K., Okada, K., & Wada, T. (2008). Arabidopsis CAPRICE-LIKE MYB 3 (CPL3) controls endoreduplication and flowering development in addition to trichome and root hair formation. *Development*, *135*(7), 1335–1345. https://doi.org/10.1242/dev.017947

Troshin, P. V., Procter, J. B., & Barton, G. J. (2011). Java bioinformatics analysis web services for multiple sequence alignment-JABAWS:MSA. *Bioinformatics*, *27*(14), 2001–2002. https://doi.org/10.1093/bioinformatics/btr304

Turck, F., Fornara, F., & Coupland, G. (2008). Regulation and identity of florigen: Flowering Locus T moves center stage. *Annual Review of Plant Biology*, *59*, 573–594. https://doi.org/10.1146/annurev.arplant.59.032607.092755

Ulmasov, T., Murfett, J., Hagen, G., & Guilfoyle, T. J. (1997). Aux/lAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. *The Plant Cell*, *9*, 1963–1971. https://doi.org/10.1105/tpc.9.11.1963

Van Den Bosch, R. (1976). Public health advantages of biological insect controls. *Environmental Health Perspectives*, *14*, 161–163. https://doi.org/10.2307/3428374

Vekemans, D., Proost, S., Vanneste, K., Coenen, H., Viaene, T., Ruelens, P., Maere, S., Van De Peer, Y., & Geuten, K. (2012). Gamma paleohexaploidy in the stem lineage of core eudicots: Significance for MADS-BOX gene and species diversification. *Molecular Biology and Evolution*, *29*(12), 3793–3806. https://doi.org/10.1093/molbev/mss183

von Goethe, J. W. (1790). Versuch die Metamorphose der Pflanzen zu erklären.

Wada, T., Kurata, T., Tominaga, R., Koshino-Kimura, Y., Tachibana, T., Goto, K., Marks, M. D., Shimura, Y., & Okada, K. (2002). Role of a positive regulator of root hair development, CAPRICE, in Arabidopsis root epidermal cell differentiation. *Development*, *129*(23), 5409–5419. https://doi.org/10.1242/dev.00111

Wada, T., Tachibana, T., Shimura, Y., & Okada, K. (1997). Epidermal cell differentiation in Arabidopsis determined by a myb homolog, CPC. *Science*, *277*(5329), 1113–1116. https://doi.org/10.1126/science.277.5329.1113

Waese, J., Fan, J., Pasha, A., Yu, H., Fucile, G., Shi, R., Cumming, M., Kelley, L. A., Sternberg, M. J., Krishnakumar, V., Ferlanti, E., Miller, J., Town, C., Stuerzlinger, W., & Provart, N. J. (2017). ePlant: Visualizing and exploring multiple levels of data for hypothesis generation in plant biology. *The Plant Cell*, *29*(8), 1806–1821. https://doi.org/10.1105/tpc.17.00073

Wagner. (1991). Secreting glandular trichomes: More than just hairs. *Plant Physiology*, *96*(3), 675–679. https://doi.org/10.1104/pp.96.3.675

Wagner, Sablowski, R. W. M., Meyerowitz, E. M., Wagner, D., Sablowski, R. W. M., & Meyerowitzt, E. M. (1999). Transcriptional activation of APETALA 1 by LEAFY. *Science*, *285*(5427), 582–584. https://doi.org/10.1126/science.285.5427.582

Wagner, Wang, E., & Shepard, R. (2004). New approaches for studying and exploiting an old protuberance , the plant trichome. *Annals of Botany*, *93*, 3–11. https://doi.org/10.1093/aob/mch011

Walker, J. D., Oppenheimer, D. G., Concienne, J., & Larkin, J. C. (2000). SIAMESE, a gene controlling the endoreduplication cell cycle in Arabidopsis thaliana trichomes. *Development*, *127*(18), 3931–3940. https://doi.org/10.1242/dev.127.18.3931

Wang & Chen, J. G. (2014). Regulation of cell fate determination by single-repeat R3 MYB transcription factors in Arabidopsis. *Frontiers in Plant Science*, *5*, 1–11. https://doi.org/10.3389/fpls.2014.00133

Wang, Cui, H. M., Huang, T. H., Liu, T. K., Hou, X. L., & Li, Y. (2016). Identification and validation of reference genes for RT-qPCR analysis in non-heading Chinese cabbage flowers. *Frontiers in Plant Science*, *7*, 1–12. https://doi.org/10.3389/fpls.2016.00811

Wang, K., Ndathe, R. W., Kumar, N., Zeringue, E. A., Kato, N., & Larkin, J. C. (2020). The CDK inhibitor SIAMESE targets both CDKA;1 and CDKB1 complexes to establish endoreplication in trichomes. *Plant Physiology*, *184*(1), 165–175. https://doi.org/10.1104/PP.20.00271

Wang, Kwak, S. H., Zeng, Q., Ellis, B. E., Chen, X. Y., Schiefelbein, J., & Chen, J. G. (2007). TRICHOMELESS1 regulates trichome patterning by suppressing GLABRA1 in Arabidopsis. *Development*, *134*(21), 3873–3882. https://doi.org/10.1242/dev.009597

Wang, Shen, C., Meng, P., Tan, G., & Lv, L. (2021). Analysis and review of trichomes in plants. *BMC Plant Biology*, (1), 1–11. https://doi.org/10.1186/s12870-021-02840-x

Wang, Sun, R., Wu, J., Liu, S., Bai, Y., Mun, J. H., Bancroft, I., Cheng, F., Huang, S., Li, X., Hua, W., Wang, J., Wang, X., Freeling, M., Pires, J. C., Paterson, A. H., Chalhoub, B., Wang, B., Hayward, A., … Zhang, Z. (2011). The genome of the mesopolyploid crop species Brassica rapa. *Nature Genetics*, *43*(10), 1035–1040. https://doi.org/10.1038/ng.919

Wang, Wang, X., & Paterson, A. H. (2012). Genome and gene duplications and gene expression divergence: A view from plants. *Annals of the New York Academy of Sciences*, *1256*(1), 1–14. https://doi.org/10.1111/j.1749-6632.2011.06384.x

Wang, Zhou, Q., Meng, Z., Abid, M. A., Wang, Y., Wei, Y., Guo, S., Zhang, R., & Liang, C. (2022). Multi-dimensional molecular regulation of trichome development in Arabidopsis and Cotton. *Frontiers in Plant Science*, *13*, 1–13. https://doi.org/10.3389/fpls.2022.892381

Weigel, D., Alvarez, J., Smyth, D. R., Yanofsky, M. F., & Meyerowitz, E. M. (1992). LEAFY controls floral meristem identity in Arabidopsis. *Cell*, *69*(5), 843–859. https://doi.org/10.1016/0092-8674(92)90295-N

Wellmer, F., Alves-Ferreira, M., Dubois, A., Riechmann, J. L., & Meyerowitz, E. M. (2006). Genome-wide analysis of gene expression during early Arabidopsis flower development. *PLoS Genetics*, *2*(7), 1012–1024. https://doi.org/10.1371/journal.pgen.0020117

Wellmer, F., & Riechmann, J. L. (2010). Gene networks controlling the initiation of flower development. *Trends in Genetics*, *26*(12), 519–527. https://doi.org/10.1016/j.tig.2010.09.001

Wellmer, F., Riechmann, J. L., Alves-Ferreira, M., & Meyerowitz, E. M. (2004). Genome-wide analysis of spatial gene expression in Arabidopsis flowers. *Plant Cell*, *16*(5), 1314–1326. https://doi.org/10.1105/tpc.021741

Wendel, J. F. (2000). Genome evolution in polyploids. *Plant Molecular Biology*, *42*(1), 225–249. https://doi.org/10.1007/978-94-011-4221-2{_}12

Wester, K., Digiuni, S., Geier, F., Timmer, J., Fleck, C., & Hülskamp, M. (2009). Functional diversity of R3 single-repeat genes in trichome development. *Development*, *136*(9), 1487–1496. https://doi.org/10.1242/dev.021733

Wheeler, A. G., & Krimmel, B. A. (2015). Mirid (Hemiptera: Heteroptera) specialists of sticky plants: Adaptations, interactions, and ecological implications. *Annual Review of Entomology*, *60*, 393–414. https://doi.org/10.1146/annurev-ento-010814-020932

Wigge, P. A., Kim, M. C., Jaeger, K. E., Busch, W., Schmid, M., Lohmann, J. U., & Weigel, D. (2005). Integration of spatial and temporal information during floral induction in Arabidopsis. *Science*, *309*(5737), 1056–1059. https://doi.org/10.1126/science.1114358

Wils, C. R., & Kaufmann, K. (2017). Gene-regulatory networks controlling inflorescence and flower development in Arabidopsis thaliana. https://doi.org/10.1016/j.bbagrm.2016.07.014

Winter, C., Austin, R. S., Blanvillain-Baufumé, S., Reback, M. A., Monniaux, M., Wu, M. F., Sang, Y., Yamaguchi, A., Yamaguchi, N., Parker, J. E., Parcy, F., Jensen, S. T., Li, H., & Wagner, D. (2011). LEAFY target genes reveal floral regulatory logic, cis motifs, and a link to biotic stimulus response. *Developmental Cell*, *20*(4), 430–443. https://doi.org/10.1016/j.devcel.2011.03.019

Winter, Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V., & Provart, N. J. (2007). An "electronic fluorescent pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS ONE*, *2*(8), 1–12. https://doi.org/10.1371/journal.pone.0000718

Wolfe, K. H. (2001). Yesterday's polyploids and the mystery of diploidization. *Nature reviews*, *2*, 333–341. https://doi.org/10.1038/35072009

Wu, H. W., Deng, S., Xu, H., Mao, H. Z., Liu, J., Niu, Q. W., Wang, H., & Chua, N. H. (2018). A noncoding RNA transcribed from the AGAMOUS (AG) second intron binds to CURLY LEAF and represses AG expression in leaves. *New Phytologist*, *219*(4), 1480–1491. https://doi.org/10.1111/nph.15231

Wuest, S. E., O'Maoileidigh, D. S., Rae, L., Kwasniewska, K., Raganelli, A., Hanczaryk, K., Lohan, A. J., Loftus, B., Graciet, E., & Wellmer, F. (2012). Molecular basis for the specification of floral organs by APETALA3 and PISTILLATA. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(33), 13452–13457. https://doi.org/10.1073/pnas.1207075109

Xu, Y., Qian, Z., Zhou, B., & Wu, G. (2019). Age-dependent heteroblastic development of leaf hairs in Arabidopsis. *New Phytologist*, *224*(2), 741–748. https://doi.org/10.1111/nph.16054

Xue, X. Y., Zhao, B., Chao, L. M., Chen, D. Y., Cui, W. R., Mao, Y. B., Wang, L. J., & Chen, X. Y. (2014). Interaction between two timing microRNAs controls trichome distribution in Arabidopsis. *PLoS Genetics*, *10*(4). https://doi.org/10.1371/journal.pgen.1004266

Yadav, R. K., Perales, M., Gruel, J., Girke, T., Jönsson, H., & Venugopala Reddy, G. (2011). WUSCHEL protein movement mediates stem cell homeostasis in the

Arabidopsis shoot apex. *Genes and Development*, *25*(19), 2025–2030. https://doi.org/10.1101/gad.17258511

Yamaguchi, N., Wu, M. F., Winter, C. M., Berns, M. C., Nole-Wilson, S., Yamaguchi, A., Coupland, G., Krizek, B. A., & Wagner, D. (2013). A molecular framework for auxin-mediated initiation of flower primordia. *Developmental Cell*, *24*(3), 271–282. https://doi.org/10.1016/j.devcel.2012.12.017

Yamaguchi, N., Wu, M. F., Winter, C. M., & Wagner, D. (2014). LEAFY and polar auxin transport coordinately regulate Arabidopsis flower development. *Plants*, *3*(2), 251–265. https://doi.org/10.3390/plants3020251

Yan, W., Chen, D., Schumacher, J., Durantini, D., Engelhorn, J., Chen, M., Carles, C. C., & Kaufmann, K. (2019). Dynamic control of enhancer activity drives stage-specific gene expression during flower morphogenesis. *Nature Communications*, *10*(1), 1–16. https://doi.org/10.1038/s41467-019-09513-2

Yang & Ye, Z. (2013). Trichomes as models for studying plant cell differentiation. *Cellular and Molecular Life Sciences*, *70*(11), 1937–1948. https://doi.org/10.1007/s00018-012-1147-6

Yanofsky, M. F., Ma, H., Bowman, J. L., Drews, G. N., Feldmann, K. A., & Meyerowitz, E. M. (1990). The protein encoded by the Arabidopsis homeotic gene agamous resembles transcription factors. *Nature*, *346*(6279), 35–39. https://doi.org/10.1038/346035a0

Yu, N., Cai, W. J., Wang, S., Shan, C. M., Wang, L. J., & Chena, X. Y. (2010). Temporal control of trichome distribution by microRNA156-targeted SPL genes in Arabidopsis thaliana. *The Plant Cell*, *22*(7), 2322–2335. https://doi.org/10.1105/tpc.109.072579

Zhang, Gennaro, D. D., Lin, G., Chai, J., & Shpak, E. D. (2021). ERECTA family signaling constrains CLAVATA3 and WUSCHEL to the center of the shoot apical meristem. *Development*, *148*(5). https://doi.org/10.1242/dev.189753

Zhang, W., Mirlohi, S., Li, X., & He, Y. (2018). Identification of functional single-nucleotide polymorphisms affecting leaf hair number in Brassica rapa. *Plant Physiology*, *177*(2), 490–503. https://doi.org/10.1104/pp.18.00025

Zhao, M., Morohashi, K., Hatlestad, G., Grotewold, E., & Lloyd, A. (2008). The TTG1-bHLH-MYB complex controls trichome cell fate and patterning through direct targeting of regulatory loci. *Development*, *135*(11), 1991–1999. https://doi.org/10.1242/dev.016873

Zheng, Z., Guo, Y., Novák, O., Dai, X., Zhao, Y., Ljung, K., Noel, J. P., & Chory, J. (2013). Coordination of auxin and ethylene biosynthesis by the aminotransferase VAS1. *Nature Chemical Biology*, *9*(4), 244–246. https://doi.org/10.1038/nchembio.1178

Zhou, Y., Yan, A., Han, H., Li, T., Geng, Y., Liu, X., & Meyerowitz, E. M. (2018). HAIRY MERISTEM with WUSCHEL confines CLAVATA3 expression to the outer apical meristem layers. *Science*, *361*(6401), 502–506. https://doi.org/10.1126/science.aar8638