# A case for developmental genetics to increase yield of oilseed rape

By Marie J. Brüser

Thesis submitted for the Degree of Doctor of Philosophy

University of East Anglia
John Innes Centre

September 2017

## Abstract

*B. napus* is a widely-grown crop and its uses are widespread. The seeds used mostly for oil production are contained within cylindrical fruits called pods. These strongly resemble the well-studied fruiting structures of the model plant *Arabidopsis*, in which patterning genes, hormones and cellular growth dynamics are well characterized. The ovary of the fruit is partitioned into two valves separated by the replum and the valve margin, which is required for fruit opening and seed dispersal. On the inside are the ovules, which post-fertilization develop into seeds. Of benefit for *B. napus* crop improvement would be the production of larger seeds as well as more seeds per pod.

The aim of this work was to improve yield of oilseed rape by targeting fruit growth. To this end, the growth of the *B. napus* pod was described in detail. This formed the basis of the study as it is imperative to understand the entire fruiting structure before attempting any improvements. Fruit length correlated consistently with seed number, making it a valuable trait to study. A correlation between stomata density and fruit length was investigated using several members of the *Brassicaceae.* Next, to identify new genes contributing to pod growth, a genome-wide association study (GWAS) was carried out on a diverse *B. napus* population. This yielded *DPb,* a cell-cycle gene. Further analysis lead to the identification of *TSN1,* a gene with similar expression levels as the former. Functions of the DP/E2F family with TSN1 and TSN2 were then studied in detail in the *Arabidopsis* fruit. Results uncovered an overlapping function in seed development and fruit growth of these previously not associated genes. Overall, this work thus demonstrates how findings in a crop species can be translated to the model to increase the fundamental knowledge of biological processes.

## Acknowledgements

This thesis would not have come together the way it has without the help and support of a key number of people.

Up and foremost I would like to thank my family for their support, interest and encouragement. Your curiosity about my project kept me fascinated with it and also helped me phrase my work in easy terminology. You always supported me through my tough times and listened to me if I felt negative about my project. And at the same time you would attempt to grasp my enthusiasm and would let me tell you all about my results, even if we both knew that I couldn't convey it all without explaining some basics first. Our holidays together also always helped to relax me, to put things into perspective and to come back to Norwich full of enthusiasm again. And thanks for rising to the challenge of finding and photographing a bok choi for me (Thanks to Leo for the picture!). Vielen Dank an jeden Einzelnen von euch!

Next, I would like to thank my fellow lab and office members. We support each other really well and I always felt comfortable discussing my worries and my successes with all of you! Thanks for picking me up when I needed it, for going on countless tea breaks and for having beers at the rec together. I want to especially mention Laila Moubayidin, who has been one of the few people who has been here for the entire duration of my project. You are an inspiration to me and I have thoroughly enjoyed watching you grow into a future group leader and also seeing you appreciate my growth and changes. Nicola Stacey, things would have been very different for me with another bench neighbour! Having your optimism, your open ear and your calmness next to me has been amazing! Thank you for helping me so much in the lab and being so incredibly patient with all of us. And then there is Mikhaela Neequaye, whom I sit next to, often write to in the evenings and hope to spend every weekend with. I am so happy you joined our lab and am incredibly grateful for our friendship! Thanks so much for your help and happy memories together at work and outside in the real world, you'll do great and I know I will hear all about it as we are definitely going to stay in touch!

Thanks to all my other cherished friends at the JIC come and gone! I have so many great memories of fun times in the lab, at tea breaks, in the rec, celebrations, parties, holidays... the list goes on. I will keep those memories close to my heart and look forward to making more!

Thanks also to my Master student Marlène Carrière who tackled *B. napus* with me and worked really hard to deliver some great results which come up in several chapters of this thesis. It was a joy to work with you and to share results with you as they happened!

I really need to mention Rachel Wells in this list too. Thank you so much for teaching me pretty much everything about *B. napus* and the GWAS, for always putting your stuff aside to help me and for taking me to Cereals and meeting all the industry experts. It is safe to say my Ph.D. project would be very different without you!

I would also like to thank horticultural services who helped me understand *B. napus* and how to grow it. Thanks for growing countless *B. napus* and *Arabidopsis* plants for me, often with short notice!

Thank you to the BBSRC for funding this project. I am very grateful to have been eligible for the funding and hope you will continue giving this chance to other non-British students too. I would like to also mention the excellent Crop Improvement Research Club (CIRC) whose meetings I am grateful to have attended. I really enjoyed being able to talk to industry representatives and seeing my research through their eyes.

Which brings me to my industrial partner Limagrain UK and my supervisor Vasilis Gegas. Thanks for supervising me, for teaching me about the oilseed rape industry and giving me the opportunity to learn about plant breeding. I have really appreciated having this link between research and academia and hope to continue working at the junction of both.

And finally, Judith Irwin and Lars Østergaard, my John Innes supervisors. Thanks so much for your constructive criticism, your belief in me and your constant support. I always felt confident coming to either of you with issues (and successes) and really appreciate that a lot! Thanks for editing my thesis, Lars, it has been a joy seeing you improving my writing by simply exchanging a word here and there or restructuring a sentence.

# Table of Contents


**1 GENERAL INTRODUCTION 15**

**1.1 FRUIT DEVELOPMENT 16**
1.1.1 THE IMPORTANCE OF FRUITS 16
1.1.2 OILSEED RAPE YIELD IMPROVEMENTS OVER TIME 16
1.1.3 BRASSICACEAE FRUITS 19
1.1.4 DISSECTING GROWTH AND SHAPE MECHANISMS 22
**1.2 GENES REQUIRED FOR THE PATTERNING OF THE *ARABIDOPSIS* SILIQUE 23**
1.2.1 FRUITFUL (FUL) 24
1.2.2 REPLUMLESS (RPL) 25
1.2.3 SHATTERPROOF 1/2 (SHP1/2) 25
1.2.4 INDEHISCENT (IND) 26
1.2.5 HECATE1, 2 AND 3 (HEC1,2,3) 26
1.2.6 SPATULA (SPT) 27
1.2.7 ALCATRAZ (ALC) 27
**1.3 HORMONES REQUIRED FOR CORRECT PATTERNING AND GROWTH OF THE *BRASSICACEAE* FRUIT 28**
1.3.1 AUXIN 28
1.3.2 CYTOKININ 30
1.3.3 GIBBERELLIN 32
**1.4 NATURAL VARIATION, DISCONTINUOUS AND CONTINUOUS TRAITS 34**
**1.5 AIM OF THE THESIS 34**

**2 GENERAL MATERIALS AND METHODS 36**

**2.1 GENERAL METHODS USED THROUGHOUT THE THESIS 37**
2.1.1 PLANT MATERIAL 37
2.1.2 DNA EXTRACTIONS 37
2.1.2.1 Edwards' quick DNA extraction for Arabidopsis 37
2.1.2.2 Arabidopsis DNA extraction carried out by the DNA extraction service 37
2.1.2.3 DNA extraction for B. napus and B. rapa 38
2.1.3 GENERAL PCR PROTOCOLS 38
2.1.4 SEQUENCING REACTIONS 39
2.1.5 E. COLI HEAT SHOCK AND ELECTROPORATION TRANSFORMATION 39
2.1.6 AGROBACTERIUM ELECTROPORATION TRANSFORMATION 39
2.1.7 ARABIDOPSIS FLORAL DIP TRANSFORMATION 40
2.1.8 MINIPREP, GEL EXTRACTION AND PCR CLEAN UP 40
2.1.9 GUS STAINING 40
2.1.10 CONFOCAL MICROSCOPY 40
2.1.11 SCANNING ELECTRON MICROSCOPY (SEM) 40
2.1.12 SEED MEASUREMENTS 41

**3 GENERAL GROWTH MECHANISMS WITHIN THE *B. NAPUS* POD 42**

**3.1 INTRODUCTION 43**
**3.2 MATERIALS AND METHODS 44**
3.2.1 PLANT MATERIAL AND SAMPLING 44
3.2.2 RANDOM IMPACT TESTING 45
3.2.3 OPTICAL PROJECTION TOMOGRAPHY (OPT) 46
3.2.4 TISSUE CULTURE AND TRANSFORMATION OF RV31 46
**3.3 RESULTS 47**
3.3.1 GROWTH DYNAMICS IN THE EARLY STAGES 48

3.3.2 The direction of cell growth can be inferred using sector analysis 50
3.3.3 Assessing differences in various B. napus accessions 52
3.3.4 Flow Cytometry 55
3.3.5 Longer pods contain more seeds 56
3.3.6 Plants with larger pods produce fewer fruits 56
3.3.7 Different ways of finding novel variation 57
3.3.8 Pod strength differed widely between accessions 57
**3.4 Discussion 59**
3.4.1 Correlations between yield parameters are consistent 59
3.4.2 B. napus fruits grow steadily in the early stages 60
3.4.3 Vernalisation time affects pod length 61
3.4.4 Pre-fertilization differences are larger than the final pod length differences 61
3.4.5 Pods have different lengths due to varied cell numbers 62
3.4.6 Different plant material can be used for different purposes 63
3.4.7 Pod-shatter resistance as a trait with much potential for improvement 63
**3.5 Concluding Remarks 64**

**4 THE CORRELATION BETWEEN STOMATA DENSITY AND POD LENGTH IN THE *BRASSICACEAE* 66**

**4.1 Introduction 67**
4.1.1 Aims and hypothesis 69
**4.2 Material and Methods 69**
4.2.1 Plant material and sampling 69
4.2.2 Rosette size measurements 70
4.2.3 Leaf and pod prints using nail varnish 70
**4.3 Results 70**
4.3.1 B. napus accessions have different pod lengths and stomata indices 70
4.3.2 B. oleracea ICE1 overexpressing plants have more stomata and shorter pods 71
4.3.3 Arabidopsis stomata density mutants display shorter siliques 73
**4.4 Discussion 77**
4.4.1 Testing the hypothesis in B. oleracea 77
4.4.2 Differences between Arabidopsis and B. oleracea scrmD lines 79
4.4.3 Total stomata numbers are more relevant than clusters 79
4.4.4 Mutants of TMM show different phenotypes 79
4.4.5 Early plant growth is affected in stomata mutants 80
4.4.6 Gynoecium development at early stages is different 80
4.4.7 Final fruit length in Arabidopsis 81
4.4.8 Water use efficiency may be affected 81
4.4.9 Carbon assimilation rates may be different 81
4.4.10 Stomata density and crop improvement 82
**4.5 Concluding Remarks 82**

**5 A GENOME-WIDE ASSOCIATION STUDY TO IDENTIFY CANDIDATE GENES UNDERLYING KEY YIELD TRAITS 83**

**5.1 Introduction 84**
**5.2 Materials and Methods 85**
5.2.1 Plant material and phenotyping 85
5.2.2 mRNA-seq and marker scoring 86
5.2.3 Mixed Linear Modelling-SNP analysis 86
5.2.4 Gene expression marker (GEM) analysis 87
**5.3 Results 87**

5.3.1 THREE COMPONENTS UNDERLIE SEED AND POD TRAITS IN B. NAPUS 87
5.3.2 A GWAS IDENTIFIES THE SAME CANDIDATE GENE IN SEVERAL DATASETS 88
5.3.3 OTHER MEMBERS OF THE DP/E2F NETWORK DID NOT SCORE HIGHLY IN THE GWAS 92
5.3.4 TWO OTHER GENES WERE PRESENT IN THE 2014 AND 2015 DATASETS OF VALVE LENGTH 92
5.3.5 GWAS ON OTHER YIELD TRAITS WERE ALSO CARRIED OUT 93
5.3.6 USING AT5G03415 TO IDENTIFY CO-EXPRESSED GENES 93
5.3.7 PARALOGUES OF AT5G03415 IN B. NAPUS 94
5.3.8 A SUBSET OF ACCESSIONS WAS USED FOR FURTHER ANALYSIS OF BNA.DPB 99
5.3.9 BNAA3.DPB PROTEIN SEQUENCES DO NOT SHOW MANY DIFFERENCES 99
5.3.10 DPB PROMOTER SEQUENCES HIGHLIGHT AN ACCESSION FROM A DIFFERENT BACKGROUND AND A HOMEOLOGOUS EXCHANGE 100
5.3.11 SEVERAL PUTATIVE TRANSCRIPTION FACTOR BINDING SITES ARE CONSERVED IN DPB PROMOTERS 103
5.3.12 THE ROLE OF DPB IN B. RAPA 104
**5.4 DISCUSSION 104**
5.4.1 THE GENE OF INTEREST WAS CHOSEN FOR ITS PUBLISHED FUNCTION 104
5.4.2 OTHER GENES THAT WERE HIGHLIGHTED IN THE GWAS USING SNPS 105
5.4.3 REASONS FOR THE IDENTIFICATION OF DIFFERENT HOMOEOLOGUES DEPENDING ON THE DATA SET USED 106
5.4.4 CANDIDATES THAT WERE PRESENT IN MORE THAN ONE YEARS' DATA 107
5.4.5 THE GWAS USING AT5G03415 EXPRESSION AS A TRAIT HIGHLIGHTED A HOMOEOLOGOUS EXCHANGE 107
5.4.6 CANDIDATE GENES WITH EXPRESSION PATTERNS SIMILAR TO THAT OF AT5G03415 107
5.4.7 PROTEIN SEQUENCES OF AT5G03415 SHOWED LITTLE VARIATION BETWEEN ACCESSIONS 108
5.4.8 SEQUENCING OF AT5G03415 PROMOTER SEQUENCES HIGHLIGHTED NIN AS DIFFERENT 108
5.4.9 CONSERVED PUTATIVE BINDING SITES MAY PROVIDE INFORMATION ABOUT DPB FUNCTION 108
**5.5 CONCLUDING REMARKS 109**

**6 OVERLAPPING FUNCTIONS OF E2F AND TSN IN FRUIT DEVELOPMENT 111**

**6.1 INTRODUCTION 112**
6.1.1 TUDOR-STAPHYLOCCOCAL-NUCLEASE (TSN) 114
6.1.2 ADENOVIRUS E2 PROMOTER BINDING FACTORS (E2FS) 115
6.1.2.1 E2Fa 116
6.1.2.2 E2Fb 117
6.1.2.3 E2Fc 118
6.1.2.4 E2Fs and programmed cell death 119
**6.2 MATERIALS AND METHODS 120**
6.2.1 PLANT MATERIALS AND GROWTH CONDITIONS 120
6.2.2 SAMPLING, MEASURING AND STATISTICS 120
6.2.3 POLLEN TUBE STAINING 120
6.2.4 FLOW CYTOMETRY 121
6.2.5 CLONING OF A PE2FB:GUS REPORTER 121
6.2.6 YEAST TRANSFORMATION 121
6.2.7 YEAST SELECTION PLATES 122
6.2.8 RNA EXTRACTION AND CDNA SYNTHESIS 122
6.2.9 QUANTITATIVE REAL-TIME PCR ANALYSIS (RT-QPCR) 122
**6.3 RESULTS 122**
6.3.1 TSN DOUBLE MUTANTS HAVE A FRUIT PHENOTYPE 122
6.3.2 E2F MUTANTS MOSTLY DO NOT SHOW A FRUIT PHENOTYPE 123
6.3.3 PLANTS HARBOURING MUTATIONS IN E2F AND TSN GENES ARE AFFECTED IN FRUIT LENGTH AND SEED SET 125
6.3.4 SEED SIZE AND SEED NUMBER CORRELATE NEGATIVELY 128
6.3.5 DEFECTS HAVE A MATERNAL ORIGIN 129
6.3.6 DEFECTS OCCUR POST FERTILIZATION 131

6.3.7 Seed loss occurs after fertilization 132
6.3.8 No differences in ploidy are present 134
6.3.9 E2F and TSN are expressed at the same place in embryo development 134
6.3.10 E2F and TSN proteins do not interact directly 135
6.3.11 Metacaspase 4 is downregulated in e2f and tsn mutants 137
6.3.12 A feedback loop between E2F and TSN is present in Arabidopsis 138
**6.4 Discussion 138**
6.4.1 tsn mutant phenotypes 138
6.4.2 e2f mutant phenotypes 139
6.4.3 Ovule development and fertilization require further study 139
6.4.4 The nature of abortions 140
6.4.5 TSN and E2F proteins do not form a dimer 141
6.4.6 Gene expression was affected in mutants 142
6.4.7 E2F cis-elements in the TSN promoters 142
6.4.8 Hormonal changes are unlikely to be the cause of the phenotypes 143
6.4.9 E2F and TSN mode of action 144
6.4.10 Differences between heterozygous and homozygous mutants 146
**6.5 Concluding remarks 146**

**7 GENERAL DISCUSSION 148**

7.1.1 Summary of findings 149
7.1.2 Cell-cycle control links all chapters 150
7.1.3 Applying the results back to the crop 151
7.1.4 Applicability of the GWAS 151
7.1.5 Using the Brassicaceae family in research 152
**7.2 Further work 153**
**7.3 Concluding Remarks 154**

**8 REFERENCES 155**

**9 APPENDIX 175**

## List of Table and Figures


FIG. 1.1. OILSEED RAPE YIELD....17
FIG. 1.2. THE TRIANGLE OF U....18
FIG. 1.3. BRASSICACEAE EVOLUTION....19
FIG. 1.4. FRUIT DEVELOPMENT IN *B. NAPUS*....19
FIG. 1.5. PATTERNING OF THE FRUIT....21
FIG. 1.6. MATURE OVULE....22
FIG. 1.7. GENES REQUIRED FOR FRUIT DEVELOPMENT....29
TABLE 2.1 GENERAL PCR PROGRAM....39
FIG. 3.1. AVENUES OF YIELD IMPROVEMENT....43
FIG. 3.2. CORRELATIONS BETWEEN YIELD RELATED TRAITS....48
FIG. 3.3. *B. NAPUS* RV31 POD GROWTH....49
FIG. 3.4. *B. NAPUS* RV31 GYNOECIUM WIDTH....49
FIG. 3.5. RV31 GYNOECIUM LENGTH VS WIDTH....50
FIG. 3.6. SECTOR ANALYSIS CONSTRUCTS....51
FIG. 3.7. SECTORS IN THE STYLE....52
TABLE 3.1. VALVE LENGTHS OF DIFFERENT *B. NAPUS* ACCESSIONS....53
FIG. 3.8. FINAL POD LENGTHS....53
FIG. 3.9. STAGE 10 GYNOECIA LENGTHS....54
FIG. 3.10. PRE-FERTILIZATION CELL ARRANGEMENTS....54
FIG. 3.11. STAGE 17 CELLULAR DETAIL....55
FIG. 3.12. PLOIDY OF PODS....55
TABLE 3.2. SUMMARY OF SEED AND POD TRAITS....56
TABLE 3.3. PLANT ARCHITECTURE....57
TABLE 3.4. SUMMARY STATISTICS OF THREE B. NAPUS POPULATIONS....58
FIG. 3.13. EXAMPLE CURVES FROM THE RANDOM IMPACT TEST....59
FIG. 4.1. STOMATA DEVELOPMENT IS TIGHTLY REGULATED....68
TABLE 4.1. STOMATA ON *B. NAPUS* PODS OF DIFFERENT LENGTHS....71
FIG. 4.2. STOMATA ON *B. OLERACEA* LEAVES AND PODS....72
FIG. 4.3. POD LENGTHS IN *B. OLERACEA*....73
TABLE 4.2. TRANSFORMED PLANTS HAD ON AVERAGE ONE LESS SEED PER POD....73
FIG. 4.4. STOMATA MUTANTS HAD SMALLER ROSETTE SIZES....74
FIG. 4.5. SILIQUE LENGTHS OF STOMATA MUTANTS....75
FIG. 4.6. STOMATA ARRANGEMENTS ON LEAVES AND SILIQUES....76
FIG. 4.7. STOMATA ARRANGEMENTS ON SILIQUES....77
FIG. 4.8. STAGE 12 GYNOECIA....78
FIG. 5.1. THE DFFS ACCESSIONS WERE SPLIT INTO FOUR POPULATIONS....87
FIG. 5.2. A PRINCIPAL COMPONENT ANALYSIS DID NOT IDENTIFY DISTINCT GROUPS IN THE DATASET....88
FIG. 5.3. AN OUTLIER AFFECTED RESULTS....89
FIG. 5.4. GWAS CANDIDATE GENES ON A3 AND C3....90
FIG. 5.5. THE EXPRESSION OF AT5G03415 CORRELATED WITH VALVE AND POD LENGTH....91
TABLE 5.1. 2014 VALVE LENGTH GWAS GEM RESULTS....92
TABLE 5.2. SUMMARY OF THE GWAS USING GEMS....95
FIG. 5.6. MANHATTAN PLOT OF THE CALCULATED TRAIT OF SEED AREA X SEEDS PER POD....94
TABLE 5.3. THE TOP 20 HITS FOR EVERY OTHER TRAIT....96
TABLE 5.4. GWAS USING GEMS WITH THE EXPRESSION OF A_JCVI_18795 AND C_ES940427....97
FIG. 5.7. GWAS CARRIED OUT USING THE EXPRESSION OF THE UNIGENE A_JCVI_18795....98
TABLE 5.5. PROTEIN SEQUENCE IDENTITIES....98
FIG. 5.8. POSITIVE CORRELATION OF *BNAA3.DPB* AND *BNAA3.TSN1*....99
TABLE 5.6. *B.NAA3.DPB* EXPRESSION AND VALVE LENGTHS....100
FIG. 5.9. PROTEIN SEQUENCES SHOWED VERY LITTLE VARIATION....101

Table 5.7. Putative transcription factor sites ..... 102
Fig. 5.10. PCR amplification of the BnaA3.DPb promoter ..... 103
Table 5.8. Five *B. rapa* TILLING mutants ..... 104
Fig. 6.1. The Arabidopsis fruit is made of several distinct parts ..... 112
Fig. 6.2. The embryo ..... 113
Fig. 6.3. The E2Fs promote the G1/S transition of the cell cycle ..... 116
Fig. 6.4. *Arabidopsis TSN* mutants had fruit phenotypes ..... 123
Fig. 6.5. Single mutants ..... 124
Fig. 6.6. *e2f* double mutants ..... 124
Fig. 6.7. *e2f tsn* double heterozygous plants had shorter pods ..... 125
Fig. 6.8. The *E2F* zygocity was causal ..... 126
Fig. 6.9. *DPB* single mutants did not have a phenotype ..... 127
Table 6.1. Higher order *e2f/tsn* mutants showed a similar trend to double mutants ..... 128
Table 6.2. Seed size ..... 129
Fig. 6.10. Ovary sizes did not differ ..... 130
Fig. 6.11. The segregation ratio ..... 130
Fig. 6.12. Pollen viability was not affected ..... 131
Fig. 6.13. Seeds were aborted shortly after fertilization ..... 132
Fig. 6.14. Mutants did not show a different cellular arrangement at stage 17 (a). ..... 133
Fig. 6.15. *E2Fa*, *E2Fb* and *TSN2* were expressed in the fruit at the same place ..... 135
Fig. 6.16. E2F and TSN proteins did not interact ..... 136
Fig. 6.17. Gene expression levels in mutants ..... 137
Fig. 6.18. Model of the TSN/MC4/TSN network in *Arabidopsis* ..... 145
Appendix Table 1. STRUCTURE output ..... 176
Appendix Table 2. Trait files used for the GWAS ..... 178
Appendix Table 3. Primers used in Chapter 5 ..... 182
Appendix Table 4. Primers used in Chapter 6 ..... 183
Appendix Fig.1. Alignment of the *B.naA3.DPb* promoters ..... 185
Appendix Fig.2. Alignment of the *BnaC3.Dpb* promoters ..... 188
Appendix. ImageJ Macros ..... 190
Appendix Table 5. Silique lengths *e2f tsn* double mutants ..... 192
Appendix Table 6. Seeds per silique of *e2f tsn* double mutants ..... 192
Appendix Table 7. Silique lengths of overexpressing plants ..... 192
Appendix Table 8. Silique lengths of crosses ..... 193
Appendix Table 9. Plant height ..... 194

**List of Abbreviations**

| Abbreviation | Full word |
|---|---|
| **A.** | Agrobacterium |
| **ABD** | Aberrant Pollen Development |
| **ALC** | ALCATRAZ |
| **AmbxCom** | AmberxCommanche |
| **ANOVA** | Analysis of Variance |
| **ARF** | Auxin Response Factor |
| **ARR** | Arabidopsis Response Regulator |
| **AXR** | Auxin response factor |
| **B.** | Brassica |
| **bHLH** | basic Helix-Loop-Helix |
| **BY-2** | Bright Yellow-2 |
| **bZIP** | Basic domain leucine zipper |
| **Cab** | Cabriolet |
| **CAP** | Capulet |
| **CK** | Cytokinin |
| **CKX** | Cytokinin oxidase/dehydrogenase |
| **CMM** | Carpel Margin Meristem |
| **CMS** | Cytoplasmic Male Sterile |
| **CNX** | Co-factor for Nitrate reductase and Xanthine dehydrogenase |
| **Col** | Columbia |
| **Dar** | Darmor |
| **DAS** | Days After Sowing |
| **DFFS** | Diversity Fixed Foundation Set |
| **DH** | Double Haploid |
| **DP** | Dimer Protein |
| **Dup** | Duplo |
| **DZ** | Dehiscence Zone |
| **E. coli** | Escherichia coli |
| **E2F** | Adenovirus E2 promoter binding Factor |
| **EDGAR** | Experimental Design Generator and Randomiser |
| **EMS** | Ethyle Methanesulfonate |
| **en*a*** | endocarp layer a |
| **en*b*** | endocarp layer b |
| **EPF** | EPIDERMAL PATTERNING FACTOR |
| **ETI** | Effector-Triggered Immunity |
| **ETO** | Ethylene overproducer |
| **ETOH** | Ethanol |
| **ETT** | ETTIN |
| **Exp** | Expert |
| **FAD** | FATTY ACID DESATURASE |
| **FAE** | FATTY ACID ELONGASE |
| **FLC** | Flowering Locus C |
| **FLP** | FOUR LIPS |

| | |
|---|---|
| **FUL** | FRUITFUL |
| **GA** | Gibberellin |
| **GA20ox** | Gibberellin 20-oxidase |
| **GA3ox** | Gibberellin 3-oxidase |
| **GAI** | GA-Insensitive |
| **GEM** | Gene Expression Marker |
| **GFP** | Green Fluorescent Protein |
| **GID** | GA-Insensitive Dwarf |
| **GMC** | Guard Mother Cell |
| **GME** | Gametophytic Maternal-Effect |
| **GUS** | ß-glucuronidase |
| **GWAS** | Genome Wide Association Study |
| **Han** | Hanna |
| **HAP** | Hours After Pollination |
| **Hs** | Homo sapiens |
| **IAA** | Indole-3-Acetic Acid |
| **ICE** | INDUCER OF CBF EXPRESSION1 |
| **IND** | INDEHISCENT |
| **KLU** | KLUH |
| **KNOX** | Knotted 1-like homeobox |
| **L.** | Lepidium |
| **Landsberg** | Ler |
| **LL** | Lignified Layer |
| **MC** | Metacaspase |
| **MLM** | Mixed Linear Model |
| **MMC** | Meristemoid Mother Cell |
| **mybp** | million years before present |
| **NASC** | European Arabidopsis Stock Centre |
| **Nin** | NingYou |
| **NLS** | Nuclear Localisation Signal |
| **OPT** | Optical Projection Tomography |
| **OSR** | Oilseed rape |
| **P.** | Picea |
| **PB** | Processing Body |
| **PCD** | Programmed Cell Death |
| **PDCD** | Programmed Cell Death Protein |
| **PUFA** | Polyunsaturated Fatty Acid |
| **QTL** | Quantitative Trait Loci |
| **Qui** | Quinta |
| **Ram** | Ramses |
| **RBR** | RETINOBLASTOMA RELATED |
| **RGA** | Repressor of gi1-3 |
| **RGL** | RGA-Like |
| **RIT** | Random Impact Test |

| | |
|---|---|
| **RPKM** | Reads per kb per million aligned reads |
| **RPL** | REPLUMLESS |
| **RT-qPCR** | Quantitative real-time PCR |
| **SAM** | Shoot Apical Meristem |
| **SAR** | Systemic Acquired Resistance |
| **ScrmD** | ScreamD |
| **SDD1** | STOMATAL DENSITY AND DISTRIBUTION 1 |
| **SEM** | Scanning Electron Microscope |
| **SG** | Stress Granule |
| **SHP1** | SHATTERPROOF 1 |
| **SHP2** | SHATTERPROOF 2 |
| **si** | Stomatal Index |
| **SIR** | Sirtinol |
| **SL** | Separation Layer |
| **SNP** | Single Nucleotide Polymorphism |
| **SOSR** | Spring Oilseed Rape |
| **SPCH** | SPEACHLESS |
| **SPT** | SPATULA |
| **SPY** | SPINDLY |
| **TAA1** | Tryptophan Aminotransferase of Arabidopsis 1 |
| **TCC** | Tissue Culture Control |
| **TCS** | Two-Component System |
| **TEC3** | TARGETS UNDER ETTIN CONTROL 3 |
| **TF** | Transcription Factor |
| **TGA** | TGACG Sequence-specific Binding Protein |
| **TLL** | Tetratricopeptide-repeat thioredoxin-like |
| **TMM** | TOO MANY MOUTHS |
| **TPR** | Tetratricopeptide Repeat |
| **TSN** | TUDOR STAPHYLOCCOCAL NUCLEASE |
| **TSW** | Thousand Seed Weight |
| **UBP** | UBIQUITIN-SPECIFIC PROTEASE |
| **UPL** | Ubiquitin Protein Ligase |
| **WOSR** | Winter Oilseed Rape |
| **Ws** | Wassilwskija |
| **WUE** | Water Use Efficiency |
| **Zho** | Zhongshuang |

# Chapter 1

## 1 General Introduction

## 1.1 Fruit Development

### *1.1.1 The importance of fruits*

Flowering plants (angiosperms) have evolved in wonderful ways and dominate many of our landscapes. This has been enabled by their successful reproduction strategies which involve seed dispersal with the help of wind, water and other species. Furthermore, different seed sizes have evolved ranging from very small seeds, such as those of *Arabidopsis* to very large ones such as those of the palm tree.

Humans rely on fruits and seeds for their survival. Most of the calories we intake come from either eating fruits themselves or seeds carried within them. Therefore, breeding efforts have been put towards the production of plants with bigger fruits or grains, a higher proportion of harvestable to non-harvestable parts and making the edible organs more nutritious and tasty.

However, humans rely on relatively few species for their calorie intake. In the UK, most agricultural land is devoted to the production of cereals, followed by temperate grasses and oilseeds (DEFRA, 2017). Over 600, 000 hectares of land were used to grow oilseeds in 2016. This group is mainly made up of oilseed rape (OSR) where the seeds contained in pods are of economic importance and are crushed to produce oil. This can be used for several purposes such as human consumption as vegetable oil and margarines, for biodiesel production, as a lubricant as well as to make plastics. The left-overs after the crushing are commonly used as high-protein animal feed.

### *1.1.2 Oilseed rape yield improvements over time*

Oilseed rape (OSR) has been grown in the UK mainly since the second world war. As a young crop, OSR is still not fully domesticated and still contains many ‘wild’ characters such as indeterminate growth and seed dispersal. As a result, the yield is not as high as for instance for cereals such as wheat and barley and the harvest index has not changed significantly over the last 25-30 years (Fig. 1.1) (AHDB, 2016).

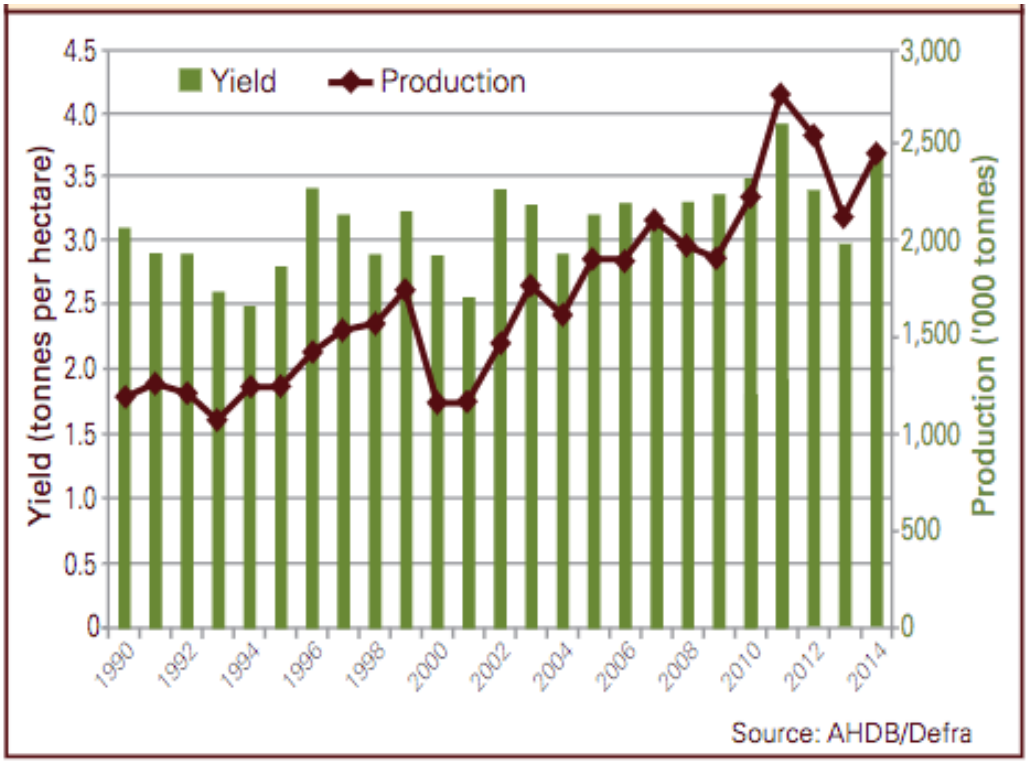



Fig. 1.1. Oilseed rape yield. The yield and production of *B. napus* has increased over time with current yields at around 3.5t/ha. Graph and data from AHDB (2016)

Oilseed rape cultivars are either spring or winter lines. The difference is that the spring varieties do not require vernalisation to flower and their lifecycle therefore is much shorter. The drawback of the spring varieties however is that the yields are much lower than the ones achieved by growing winter oilseed rape (WOSR). WOSR has higher yields that are roughly twice as high as those from spring oilseed rape (SOSR). The WOSR plants are also larger and are more robust and can therefore withstand biotic and abiotic stresses better. Furthermore, OSR cultivars are mostly hybrid varieties with few cultivars still being open pollinated. This is because hybrids yield better and it is easier to introduce new valuable traits using the hybrid way of producing plants.

OSR has been bred for several characteristics. The gross output is the most important factor and is calculated by the yield and the oil content (AHDB, 2016). Traits such as resistance to lodging, disease resistance and maturity dates are also important. The Agriculture and Horticulture Development Board (AHDB) prepares a yearly booklet highlighting the most important traits of cultivars. They make notes of traits such as the stem stiffness, the shortness of the stem, flowering time, timing of maturity, glucosinolate content, resistance to light leaf spot and phoma stem canker as well as the above-mentioned lodging and yield traits. Other traits some breeding companies have bred for are pod shatter resistance as well as different oil profiles such as high

erucic acid rape (HEAR) and high oleic, low linolenic (HOLL) rape. Clearly the yield of oilseed rape depends on many factors and breeding needs to incorporate all the known traits to produce cultivars that are able to yield highly.

In the UK and Western Europe, *B. napus* is the primary species grown for rapeseed oil production. *B. napus* belongs to the Brassica genus of the *Brassicaceae* family. The genetic relationship among domesticated members of the Brassica genus has been described in the triangle of U, which include six closely related Brassica species; the three diploids *Brassica rapa* (AA), *B. nigra* (BB) and *B. oleracea* (CC) and the three allotetraploids *B. juncea* (AABB), *B. carinata* (BBCC) and *B. napus* (AACC) (Fig. 1.2) (U, 1935). Whereas *B. oleracea* is the only species among those which is not grown for seed production, the others are used to produce both mustard and oil except for *B. napus* which is only grown for making oil.

The model plant *Arabidopsis thaliana* is also part of the *Brassicaceae*. *B. napus* and *Arabidopsis* are estimated to have diverged approximately 43 million years ago (Fig. 1.3) (Beilstein et al., 2010). During this evolution, a genome triplication occurred, giving rise to *B. rapa* and *B. oleracea*. These two species then hybridized roughly 10,000 years ago and gave rise to *B. napus* (Niu et al., 2009, Ostergaard and King, 2008). Due to the triplication and subsequent hybridisation, genes that are present once in *Arabidopsis* can have up to six gene copies in *B. napus* (Fig. 3). This evolutionary closeness is useful for scientific purposes as genes in *B. napus* often have the same function as those in *Arabidopsis*.

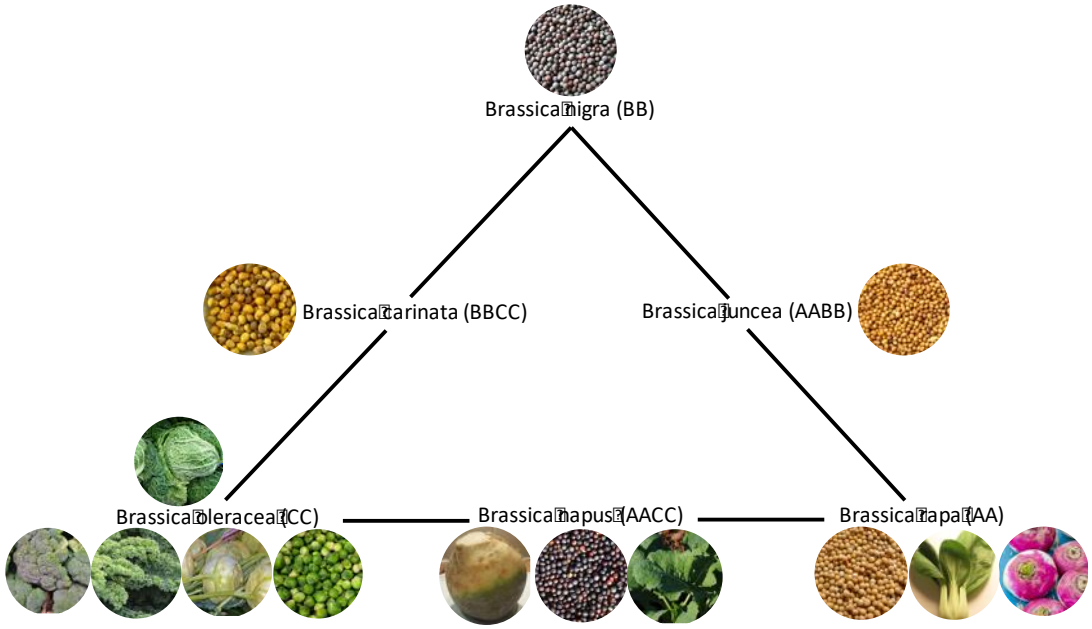



Fig. 1.2. The triangle of U describes how different Brassica species are linked. Most are grown for their seeds but many other uses exist.

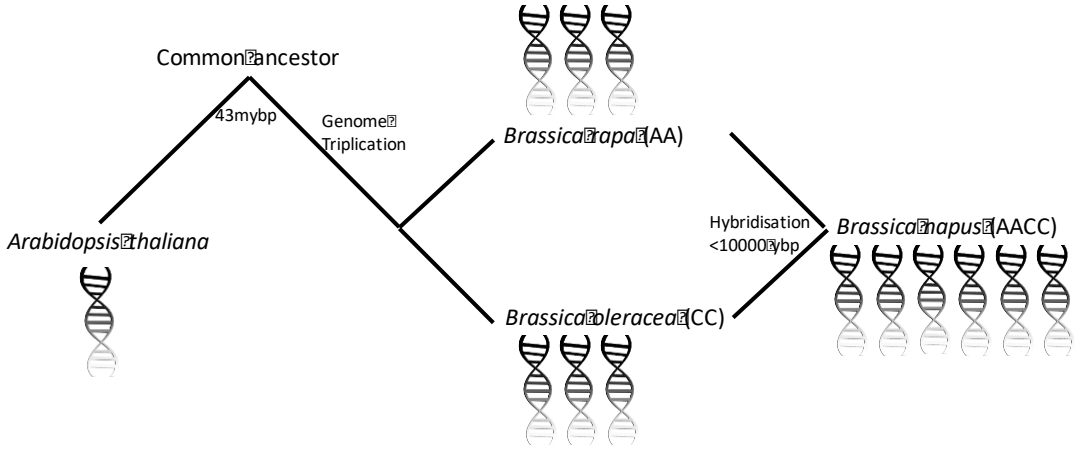



Fig. 1.3. Brassicaceae evolution. *B. napus* and the model plant *Arabidopsis* have diverged roughly 43mybp. More recently, *B. rapa* and *B. oleracea* hybridised to give rise to the allotetraploid *B. napus* with its A and C genomes. ybp, years before present; mybp, million years before present

### *1.1.3 Brassicaceae fruits*

*Brassicaceae* seeds are contained in cylindrical structures called pods. This fruit is typical of more than 3,000 species of *Brassicaceae* (Ferrandiz et al., 2000). The *Arabidopsis* fruit development has been described in detail from the moment the initial floral buttress is formed. The growth has been divided into stages. Stages 1 – 12 describe the growth pre-fertilization beginning with the floral buttress and ending with the mature gynoecium ready for fertilization (Smyth et al., 1990). At that point, the petals are the same length as the stamens. It takes almost two weeks from initiation to stage 12 (Smyth et al., 1990). After fertilization, the fruit continues growing and reaches its full length and thickness roughly 10 days after fertilization (Robles and Pelaz, 2005). The development of a *B. napus* fruit is partitioned into the same stages as the *Arabidopsis* fruit except for the addition of a stage 14b, and is shown in Figure 1.4.


Fig. 1.4. Fruit development in *B. napus*. Similar stages are found in *Arabidopsis* development.

The early fruit consists of a gynoecium composed of two valves. At the top of the gynoecium is the stigmatic tissue which consists of specialised epidermal cells (Fig. 1.5). Their function is to capture pollen and begin the process of pollen germination. The style connects the stigma to the ovary with its two valves. Separating the valves is the replum. At the boundary of the replum are a few rows of narrow cells, called the valve margin, which is where the fruit will eventually open (Robles and Pelaz, 2005). At the base, the gynophore connects the gynoecium to the receptacle and the rest of the plant.

After fertilization the gynoecium goes through a number of changes and specialized tissues are formed. The valve margins, a region of roughly seven cells, develop the dehiscence zone (DZ) where the valves eventually separate from the replum (Ferrandiz et al., 1999) (Fig. 1.5). This region expands more slowly than the valve cells, leading to a constriction at the valve margin (Rajani and Sundaresan, 2001). The dehiscence zone is made of two types of tissues; the non-lignified cells close to the replum (Separation layer (SL) or endocarp layer a (en*a*)) and the lignified layer (LL), made up of thick-walled cells, next to the valve (Child et al., 2003). These arise through unequal divisions which form the small, thin-walled SL cells, and larger LL cells, which begin forming thickened walls at stage 17a and then lignify at stage 17b (Ferrandiz, 2002, Wu et al., 2006). Once the lignification process has occurred, cells in the en*a* layer start degrading and secreting cell wall-degrading enzymes, such as polygalacturonase, which degrade the pectin-rich separation layer (Child et al., 1998, Petersen et al., 1996, Rajani and Sundaresan, 2001). These processes together create a spring-like tension that leads to the opening of the fruit upon maturity (Rajani and Sundaresan, 2001).

The inside of the fruit is also made of several distinct tissues. These arise from the carpel margin meristem (CMM) which gives rise to the placenta, ovules, septum, transmitting tract as well as the stigma and style seen from the outside. Defects in any of these tissues leads to problems with fertilization and can subsequently mean the abortion of ovules or seeds.

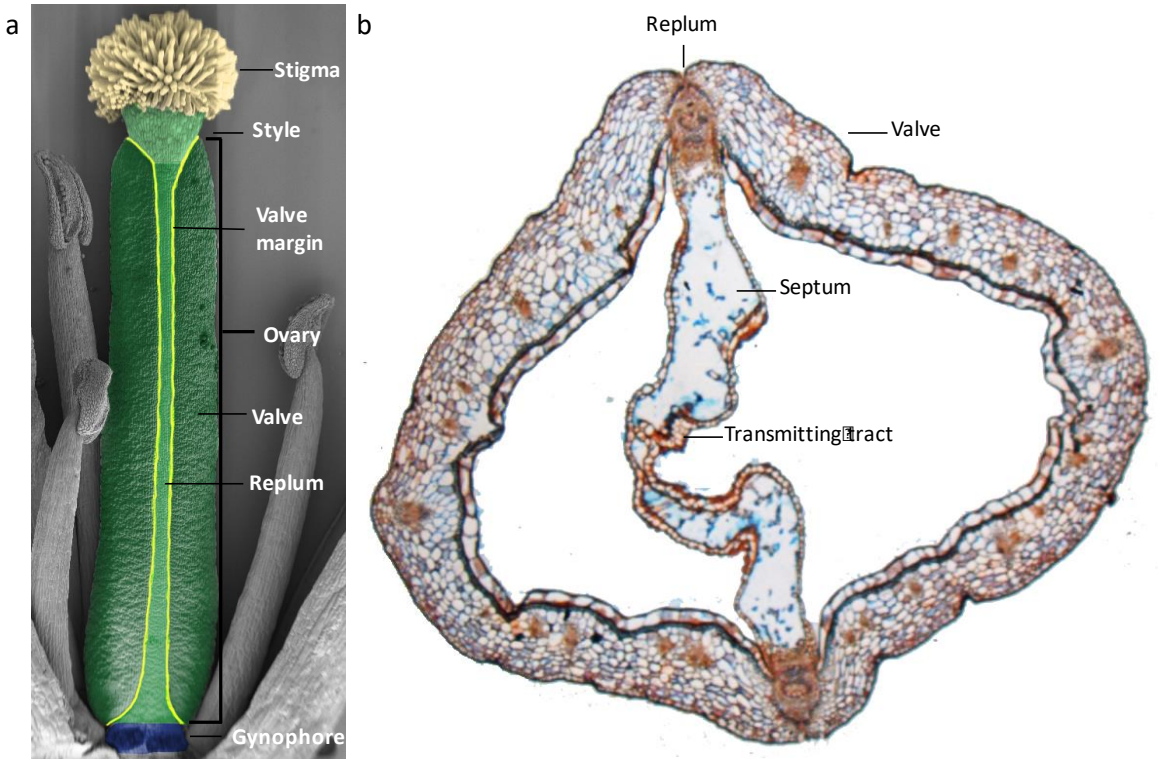



Fig. 1.5. Patterning of the fruit. The gynoecium is split into the stigma, the style and the ovary (a). A transverse section of the fruit displays the tissues found on the inside such as the septum and transmitting tract (b). The *Arabidopsis* fruit (a) has a much larger replum than *B. napus* fruits (b). False colouring in (a) added in Photoshop.

The ovules are derived from the placental tissue within the ovary wall. The ovule contains the nucellus, which will hold the embryo sac, the integuments, which surround the final embryo sac, and the funiculus, which connects the ovule to the ovary (Reiser and Fischer, 1993). The development of the ovule occurs in two stages; the megasporogenesis and the magagametogenesis. During megasporogenesis four megaspore nuclei are produced. They then, through megagametogenesis develop into the mature embryo sac: the megagametophyte (Reiser and Fischer, 1993). The embryo sac contains six haploid cells: one egg cell, two synergid cells and three antipodal cells as well as the diploid central cell (Fig. 1.6). Upon fertilization, the pollen tube discharges its contents into either of the synergid cells (Reiser and Fischer, 1993). From there, one sperm nucleus fertilizes the egg cell and the other the central cell. The egg cell then develops into a zygote and will turn into the embryo. The central cell generates the endosperm which supplies the embryo with nutrients (Reiser and Fischer, 1993).

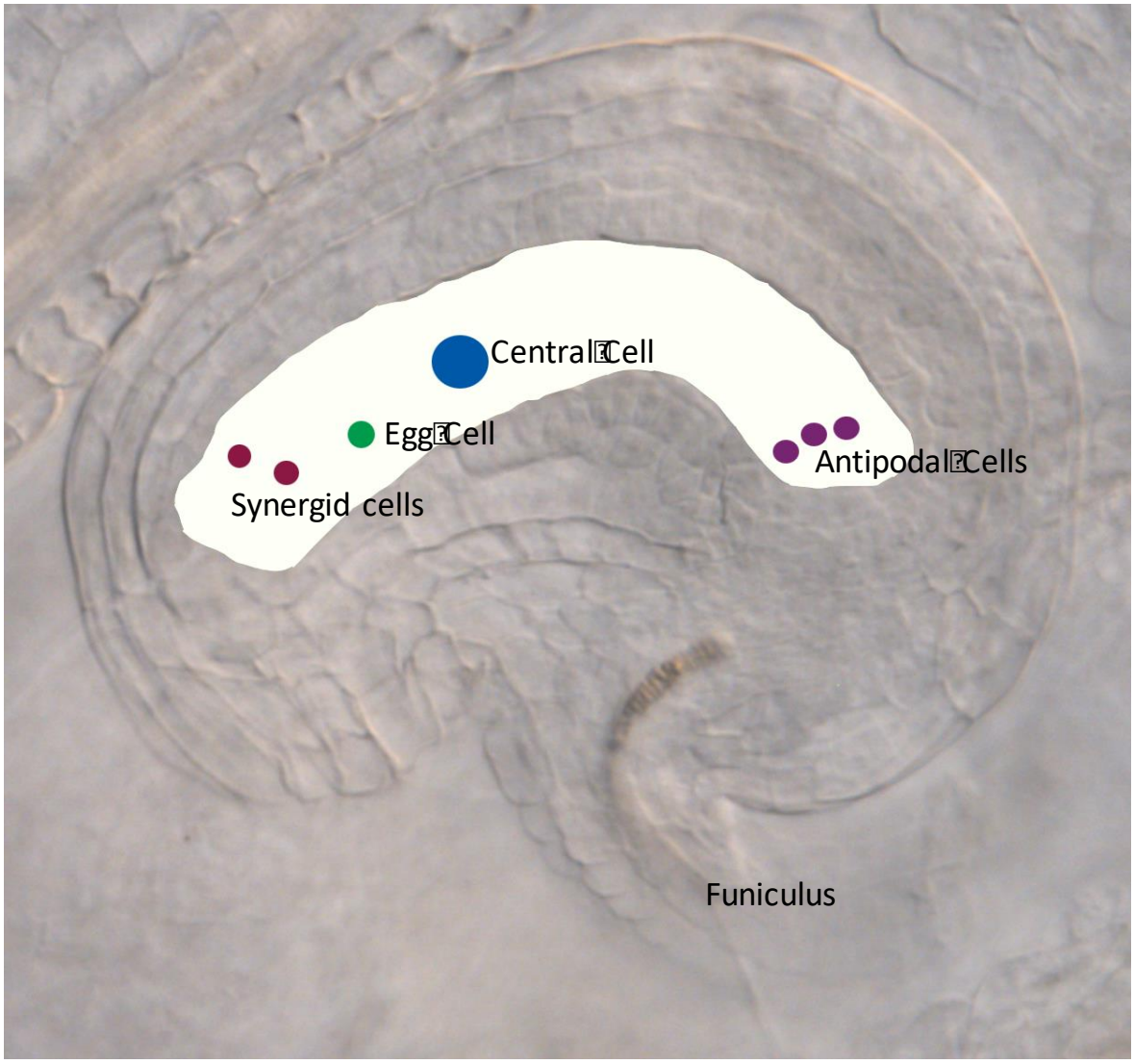



Fig. 1.6. Mature ovule. The ovule contains the embryo sac (light green) with the seven cells required for correct fertilization. The ovule is connected to the central tissue by the funiculus.

### *1.1.4* Dissecting growth and shape mechanisms

The growth of the *Arabidopsis* fruit has been described and modelled in detail by Eldridge et al. (2016). The cylindrical fruit shape is determined from the first stages of development and does not change. The growth of the fruit however, can be split into two main phases; a short phase of rapid growth in the longitudinal axis lasting two days at around stage 6 followed by a slower, steady growth to maturity (Eldridge et al., 2016). The growth in width is constant throughout development. Using a heat-inducible system, sectors of RFP and CFP were induced and cell growth tracked. Cells in the valve divide two to three times in the proximodistal axis with zero or one division in the circumferential axis. The growth rate in the replum however is higher and occurs only in the longitudinal axis. In contrast, growth at the top and the bottom of the fruit is

slower. All divisions occur along the proximodistal axis of the fruit, leading to the long, slender final shape.

Other *Brassicaceae* fruits have more complex shapes. This is true for *Capsella rubella* for instance which develops heart-shaped fruits. These fruits first grow cylindrically like *Arabidopsis*, then develop into a flattened oval shape and finally, post fertilization, acquire a heart-shape (Eldridge et al., 2016). The early dynamics of growth are therefore the same as those present in *Arabidopsis* with growth being mainly longitudinal. The dynamics in the style and the replum are the same as in *Arabidopsis* since their shapes are the same. Conserved are also the regions of reduced growth at the base and apex of the fruit. However, to produce the obloid spheroid shape, *Capsella* fruit growth changes to be isotropic. That is, cells divide in the longitudinal and latitudinal axis at the same rate (Eldridge et al., 2016). Furthermore, to create the rounded shape, cells do not grow along the proximodistal axis but instead grow outwards at the base of the fruit and angle back towards the replum at the apex. The final stage of growth in *Capsella* creates the characteristic heart-shape of the fruit. To grow from the flattened oval into a heart-stage, the dynamics change again. At the base of the fruit, cells grow along the longitudinal axis but this changes with distance from the base. Cells higher up in the fruit grow diagonally. This change in growth dynamics is required to create the more complex shape of the *Capsella* fruits, which was also shown using computational modelling (Eldridge et al., 2016).

This study highlights some of the mechanisms required or growing different shapes. It also emphasizes the importance of the interplay between growth speed and direction. Using the knowledge gained from the two different shapes, the authors could acquire shapes of other *Brassicaceae* fruits such as the spherical fruit of *Neslia paniculata,* the mature oblate sphere of *Lepidium campestre* as well as the fruit of *Alyssum* which is also a flattened spheroid but its flat side is opposite to the one found in *L. campestre* (Eldridge et al., 2016). Such studies highlight the importance of understanding general growth mechanisms and the basics before attempting to make size, shape or other changes to an organ.

### 1.2 Genes required for the patterning of the *Arabidopsis* silique

Observations on growth rate and direction such as those outlined above can be very helpful in gaining an understanding of the growth dynamics. However, understanding the genetic basis underlying such growth is also required to grasp the entire system. Some such genes have already

been studied but some remain elusive. Many of the genes known are involved in the development of both internal and external tissues.

### *1.2.1 FRUITFUL (FUL)*

One important gene for fruit growth and shape is *FRUITFUL* (*FUL*). *Arabidopsis ful* mutants are indistinguishable from wild type before fertilization but fail to elongate after fertilization (Ferrandiz et al., 2000). *Capsella ful* mutant fruit are similar to wild type before fertilization with the exception of an extended style (Eldridge et al., 2016), post fertilization however, many more differences arise. *ful* fruits do not acquire the heart shape but instead remain rounded and growth too is stunted (Eldridge et al., 2016). This gene is therefore required for post-fertilization growth in both species and likely also in other *Brassicaceae.*

*FRUITFUL* is also required for the correct patterning of the fruit *and is required for correct* valve margin development where it represses the valve margin identity genes *SHATTERPROOF1/2* (*SHP1/2*), *INDEHISCENT* (*IND*) and *ALCATRAZ* (*ALC*) from being expressed in the valves (Ferrandiz et al., 2000, Liljegren et al., 2004). Mutants therefore develop lignified valve cells instead of the normal arrangement (Ferrandiz et al., 2000, Gu et al., 1998). This conversion of tissue means the fruits cannot open the way they would normally. On the other hand, overexpressing *FUL* leads to the loss of the valve margin as the entire fruit takes on valve identity (Ferrandiz et al., 2000). This also renders the fruit resistant to shattering. These results are true in *Arabidopsis* but have also been translated to crop plants such as *Brassica juncea* where the overexpression of the *Arabidopsis FUL* gene under the 35S promoter lead to shatter resistant fruits (Ostergaard et al., 2006). It has therefore been demonstrated several times that the same mechanisms in fruit patterning are present in the *Brassicaceae* and that knowledge from the model can be translated to the crop.

*FUL* is required for post fertilization elongation of the fruit (Ferrandiz et al., 2000, Gu et al., 1998). Therefore, *ful* mutants have smaller siliques than wild-type plants (80% smaller) and a reduced seed set (by 25%) (Gu et al., 1998). The size difference is caused by a defect in cell differentiation in the replum and valves. The epidermal cells fail to expand and no stomata form (Gu et al., 1998). The size restriction causes the pods to split open a few days after fertilization. *ful* mutants fail to develop a normal DZ and are indehiscent due to a conversion of valve cells to lignified cells (Ferrandiz et al., 2000, Gu et al., 1998). Over-expression of *FUL* in *Arabidopsis* gives

rise to fruits that completely lack valve margin tissues, which is in agreement with the role of FUL in repressing the expression of valve margin identity genes (Ferrandiz et al., 2000).

### *1.2.2 REPLUMLESS (RPL)*

The replum formation requires the gene *REPLUMLESS* (*RPL*) (Girin et al., 2009, Roeder et al., 2003). It moreover restricts valve expansion to control the meeting point of the valves.

*RPL* is widely expressed and is present in stems, inflorescences, in the style, the pedicel and in the sepal vasculature (Roeder et al., 2003). *RPL* prevents the genes that are active in the valve margin, such as *INDEHISCENT* (*IND*) and *SHATTERPROOF1/2* (*SHP1/2*), from being expressed in the replum, thereby aiding the correct patterning of the fruit (Groszmann et al., 2011, Liljegren et al., 2004, Roeder et al., 2003). This prevents the replum from adopting valve margin identity. However, *RPL* is not necessary for the replum formation as such as the *shatterproof1/2 rpl* triple mutant has a replum (Roeder et al., 2003). *rpl* mutant fruits display lignified valve margin cells in the replum region and are half the size of wild-type fruit (Roeder et al., 2003). They also display reduced shatter susceptibility (Dinneny et al., 2005).

*Arabidopsis* fruits have a much broader replum than pods from members of the Brassica genus. This is due to a single SNP difference between the species in the *RPL* gene changing one nucleotide from a cytosine (C) to a thymine (T) ~3kb upstream in the promoter region. Interestingly, the replum of *B. oleracea* is enlarged when expressing a version of the *B. oleracea RPL* gene containing a C at this position (as in *Arabidopsis*) suggesting that a cis-element in this part of the promoter is required for replum-specific *RPL* expression and activity (Arnaud et al., 2011).

### *1.2.3 SHATTERPROOF 1/2 (SHP1/2)*

*FUL* represses the activity of the redundant MADS-box genes *SHATTERPROOF1* and *SHATTERPROOF 2* (*SHP1/2*) (Ferrandiz et al., 2000, Liljegren et al., 2000, Rajani and Sundaresan, 2001). *SHP1/2* are expressed in the valves, in developing ovules, the septum and the style (Flanagan et al., 1996, Liljegren et al., 2000, Savidge et al., 1995). *SHP1/2* are repressed by *RPL* and *FUL* and are positively regulated by the carpel identity gene *AGAMOUS* (Ferrandiz et al., 2000, Savidge et al., 1995).

*SHP1/2* control the development of the dehiscence zone and the lignification of the valve margin (Liljegren et al., 2000). Loss of function mutants (*shp1/2* double mutant) do not develop a dehiscence zone, show reduced lignification and fail to shatter in *Arabidopsis* and in *L. campestre* (Dinneny et al., 2005, Lenser and Theissen, 2013, Liljegren et al., 2000).

### *1.2.4 INDEHISCENT (IND)*

Downstream of the *SHP1/2* genes are the *INDEHISCENT* (*IND*) and the *ALCATRAZ* basic helix-loop-helix-type (bHLH) transcription factor genes (Dinneny et al., 2005). *IND* was first unknowingly described by Sundaresan et al. (1995) who reported the expression pattern of a valve margin marker called *GT140*. Like *SHP1/2*, *IND* is repressed by *RPL* and *FUL* which leads to expression only in the valve margin tissues (Ferrandiz et al., 2000, Liljegren et al., 2000, Robles and Pelaz, 2005, Sundberg and Ostergaard, 2009).

*IND* is required for DZ differentiation, fruit opening and lignification (Liljegren et al., 2004) and is expressed in anthers, pollen, transmitting tract and the dehiscence zone (Kay et al., 2013). After fertilization, the expression levels reduce and are targeted to the valve margin only (Kay et al., 2013). Furthermore, IND represses replum formation and *ind* fruits show enlarged repla (Girin et al., 2010). IND is also required for valve margin development and in *ind* mutants the valve margins do not form correctly and fail to shatter, a phenotype also described in *L. campestre* (Lenser and Theissen, 2013, Liljegren et al., 2004). This is in part due to the lack of the auxin minimum which is necessary for pod shatter (Sundberg and Ostergaard, 2009). The mechanism of this is described in detail in the auxin section of this introduction.

### *1.2.5 HECATE1, 2 and 3 (HEC1,2,3)*

The *HECATE1,2* and *3* (*HEC1, 2, 3*) genes encode bHLH transcription factors (Gremski et al., 2007). Pre-fertilization, all three genes are expressed in the developing stigma, septum and transmitting tract (Gremski et al., 2007). *HEC1* and *2* are also present at low levels in the ovules and *HEC3* in the funiculus. The expression of *HEC1* and *2* disappears post fertilization but *HEC3* continues to be expressed in the funiculus and the transmitting tract. *hec1* and *hec2* single mutants as well as their double mutant do not have a fruit phenotype (Gremski et al., 2007). Loss of *hec3* however leads to a 40% reduction in fertility which is due to a maternal defect. This is in part because *HEC3*

is required for the attraction of pollen to the ovules (Kay et al., 2013). The defects are exacerbated in the *hec1 hec3* double mutant where only 17% of wild-type seed set is present. Triple mutants are strongly affected in fertility due to a complete lack of stigmatic tissue, a cleft in the style as well as septum and transmitting tract defect (Gremski et al., 2007). The three genes therefore function redundantly and are required for correct patterning of the fruit.

### *1.2.6 SPATULA (SPT)*

Another gene that is closely associated to *IND* is *SPATULA* (*SPT*), encoding another bHLH factor. *SPT* is needed for septum, style and stigma formation as well as valve margin and dehiscence zone development (Groszmann et al., 2011). *SPT* is expressed in carpel margins, leaves, petals, stamens, roots, stigmatic papillae, young ovule primordia, the style, the transmitting tract and in the peripheral zone of shoot apical meristems (Heisler et al., 2001). At later stages *SPT* expression becomes restricted to the valve margins and the dehiscence zone (Heisler et al., 2001).

*spt* mutant fruits show increased shatter resistance due to defects in valve margin formation, have reduced stigma and styles, lack the transmitting tract cells and display a split phenotype at the apical end of the gynoecium (Girin et al., 2011, Groszmann et al., 2011, Reyes-Olalde et al., 2013). Furthermore, the fruits have fewer ovules and in the strong *spt-2* mutant, only a quarter of seeds develop because, due to the lack of transmitting tract, the pollen tubes cannot make their way to the ovules successfully (Alvarez and Smyth, 1999).

*SPT* is directly and positively regulated by *IND* and they are closely linked to several hormones (Girin et al., 2011). IND and SPT proteins interact and control the direction of auxin transport (Girin et al., 2011). They are therefore both involved in the formation of the auxin minimum needed for dehiscence and in correct development of the carpels (Girin et al., 2011). *spt ind* double mutants fail to create a radialised style due to defects in auxin patterning (Moubayidin and Ostergaard, 2014). Cytokinin treatment can abolish the split phenotype but the septum at the inside of the fruit does not fuse (Reyes-Olalde et al., 2017). *SPT* also induces cytokinin signalling at the CMM and septum primordia by regulating the expression of *ARR1* and *ARR12* (Reyes-Olalde et al., 2017).

### *1.2.7 ALCATRAZ (ALC)*

*ALCATRAZ* (*ALC*) encodes a protein with a *bHLH* domain that has recently diverged from a *SPATULA*-like ancestor and is required for separation layer development (Groszmann et al., 2011, Robles and Pelaz, 2005). *ALC* expression can be found in the stigmatic papillae prior to fertilization but post-fertilization it is found only in the valve margin due to repression by *FUL* (Rajani and Sundaresan, 2001, Robles and Pelaz, 2005). In the internal organs, *ALC* is expressed in the septum, the ovules as well as at the boundary of the stigma and the transmitting tract (Groszmann et al., 2011). *alc* mutants fail to develop a fully defined separation layer and therefore the fruits rupture instead of having the valves detaching (Rajani and Sundaresan, 2001, Robles and Pelaz, 2005, Lenser and Theissen, 2013).

The genes described here are all necessary for correct gynoecium patterning and development. The interaction between them can be seen in Fig. 1.7. In summary, most genes play a role in the formation of the septum and transmitting tract tissues. *HEC1/2/3, SPT* and *ALC* are needed for stigma development pre-fertilization and *SPT, FUL, RPL* and *SHP1/2* for style formation. *SHP1/2, IND, ALC* and *SPT* are restricted to the valve margin by *RPL* from the replum and *FUL* from the valve side.

Many other genes have been shown to affect fruit development but the ones presented here are some of the key factors involved in the correct patterning. While many of the genes interact with each other, many also have a cross-talk with hormones and their role will be described next.

## 1.3 Hormones required for correct patterning and growth of the *Brassicaceae* fruit

### *1.3.1 Auxin*

The processes involved in pod formation and dehiscence are extensive and require proper positioning. Several of the hormones involved have been identified but even so some of the details remain elusive. The most-, and first-studied hormone in this area is auxin (Nemhauser et al., 2000).

Auxin is involved in many plant processes and as such is necessary for correct gynoecium growth and patterning (Sundberg and Ostergaard, 2009). Auxin maxima at the apical end of ovules and primordia are necessary for this (Benkova et al., 2003, Kuusk et al., 2002, Moubayidin and

Ostergaard, 2014). The transmitting tract formation, pollen tube germination and growth also require correct auxin distribution (Wu et al., 2006).

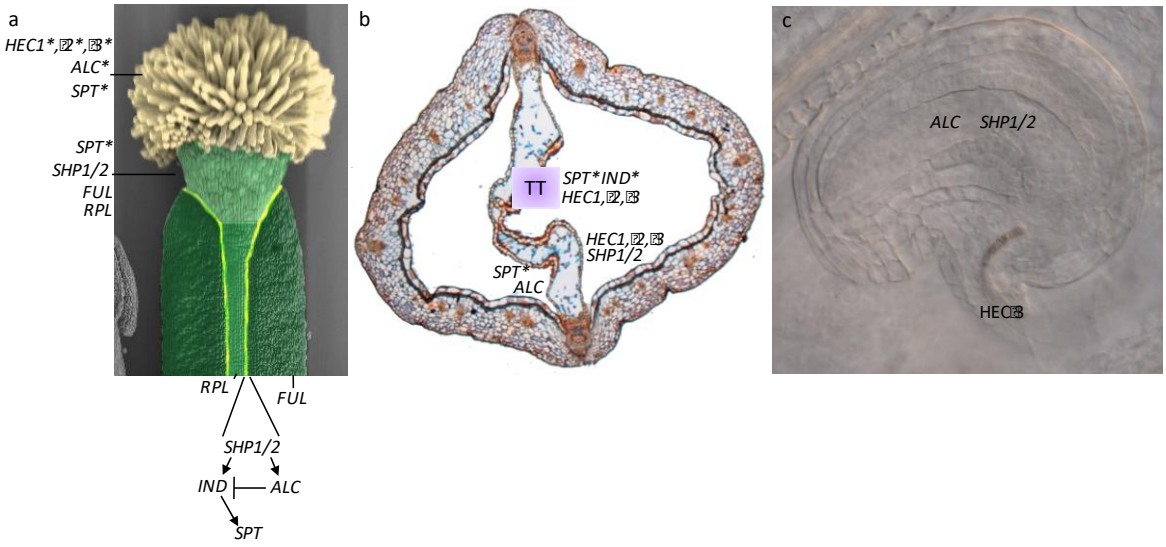



Fig. 1.7. Genes required for fruit development. Genes expressed in the fruit are often involved in the formation of organs inside the gynoecium and correct fruit patterning on the outside. Many of the genes interact with each other and therefore a mutation in one can have widespread effects in other parts of the fruit. (a) The outside of the gynoecium with its stigma (light yellow), style (light green), replum (green), valve margin (bright yellow) and valves (dark green). (b) The inside of the gynoecium with the septum and the transmitting tract. (c) The ovule and funiculus. * genes are only expressed pre-fertilization

Use of the *DR5rev::GFP* reporter for auxin signalling has made visualisation of auxin-signalling patterns possible (Benkova et al., 2003). At early stages of gynoecium development it can be seen as two foci at the edges of the gynoecium (Larsson et al., 2013). Four foci then emerge which later form a ring and in mutants with defective style development these foci fail to join and the style does not acquire its normal radialised pattern (Moubayidin and Ostergaard, 2014). The DR5::GFP signal is strongest just before pollen maturation at which point it can be seen in the developing stigma. After pollen maturation the signal decreases and fades (Larsson et al., 2013).

The *Arabidopsis* genome encodes 23 Auxin response factors (ARFs) (Rademacher et al., 2011) some of which have been described as having functions during fruit development (Goetz et al., 2006, Sessions et al., 1997). Under low auxin conditions these are repressed by Aux/IAA proteins (Tiwari et al., 2004). The repression is then lost in high auxin conditions where auxin is bound by the TIR1/AFB class proteins which in turn interact with the Aux/IAAs, leading to ubiquitylation of

the Aux/IAAs and their subsequent degradation in the 26S Proteasome (Dharmasiri et al., 2005, Kepinski and Leyser, 2005). ETTIN/ARF3 (ETT) is slightly different than the other ARFs in that it lacks the PB1 domain required for the interaction with Aux/IAAs (Guilfoyle, 2015). *ett* mutants have severe fruit phenotypes in which the style does not fuse at the top and tissue identities are lost, leading to aberrant stigma formation (Sessions et al., 1997, Simonini et al., 2016). ETT can heterodimerize with IND and the phenotype of the double *ett ind* mutant exacerbates either of the singles mutants at the apex of the gynoecium (Simonini et al., 2016). The protein interaction between ETT and IND is auxin sensitive so that, upon increase in auxin, the two proteins no longer bind to each other (Simonini et al., 2016). This sensing of auxin by ETT is required for proper development of the gynoecium (Simonini et al., 2016). In addition to IND, ETT has been found to interact with a suite of other proteins belonging to different transcription factor families in a similarly IAA-sensitive manner (Simonini et al., 2016). Some of these interactions have roles in other aspects of plant development and in a recent genome-wide study downstream target genes, such as *RPL,* were identified and verified (Simonini et al., 2017).

Dynamic distribution of auxin is required for formation of the separation layer. Auxin is present in the fruits after fertilization and a maximum has been observed in the valve margin until roughly stage 16 that may be required for an asymmetric cell division that defines the separation and lignified layers (van Gelderen et al., 2016). Depletion of auxin to form a local auxin minimum in the valve margin is then required for correct development. This was shown in *Arabidopsis* (Sorefan et al., 2009) and correlated with tissue-specific auxin measurements in *B. napus* (Chauvaux et al., 1997). Artificially increasing auxin levels in the valve margins in *Arabidopsis*, leads to a lack of proper valve margin tissue formation and fruits become indehiscent (Sorefan et al., 2009). IAA is produced in the seed and therefore, in parthenocarpic fruit auxin levels stay low throughout (Chauvaux et al., 1997). Since then it has been suggested that as seeds become dormant auxin levels drop due to a decrease both in biosynthesis and in transport, allowing the ripening of the fruit (McAtee et al., 2013). This decrease in auxin in the dehiscence zone correlates with an increase in cell-wall-degrading enzyme activity and may allow pods to become responsive to ethylene produced by the seeds (Child et al., 1998, Meakin and Roberts, 1990). Auxin therefore functions as the control mechanism to ensure the timely maturation of seed and pod (Sundberg and Ostergaard, 2009).

### *1.3.2 Cytokinin*

Auxin is not the only hormone involved in fruit patterning. Cytokinins (CKs), N6-substituted adenine derivatives, also play an important role (Argueso et al., 2010). These signalling molecules are required throughout plant development such as root and shoot meristem control, shoot initiation, leaf senescence, nutrient mobilization and seed germination (Lindsay et al., 2006, Muller and Sheen, 2008, Werner et al., 2003).

Using a two-component system tagged with GFP (TCS::GFP) CK signalling in *Arabidopsis* can be visualised and used to study the distribution of CK signalling during developmental processes. During gynoecium development, the TCS::GFP reporter can be observed before fertilization. The signal is detected in the medial regions and later, at fertilization, at the boundary between the valves and the replum, which later develops into the valve margins (Marsch-Martinez et al., 2012). Exogenously adding CKs to fruits from mutants with defects in dehiscence (*ind* and *shp1/2* mutants) can restore valve margin formation and dehiscence (Marsch-Martinez et al., 2012).

Cytokinin oxidase/dehydrogenases (CKXs) are responsible for CK breakdown (Bartrina et al., 2011). The *Arabidopsis* genome encodes seven CKXs (*AtCKX1 – AtCKX7*) with different expression patterns for each (Werner et al., 2003). Mutation of two *CKX* genes (*CKX3* and *CKX5*) leads to an increase in cytokinin content, a delay in reproductive meristem cell differentiation and subsequently to the formation of larger inflorescence meristems and more flowers (Bartrina et al., 2011). Carpels also display higher ovule and seed density which translates to an increase of 55% in overall yield (Bartrina et al., 2011). Similar results have been obtained in *B. napus* where an increase in CK levels leads to an increase in thousand-seed weight and produces plants with 52% higher yields (Roeckel et al., 1998). A decrease in cytokinin leads to the opposite and creates flowers displaying asynchronous fruit ripening and fewer seeds (Werner et al., 2003).

CK signalling requires type-B Arabidopsis response regulator proteins (ARR). These are transcription factors that control the expression of the type-A *ARRs* which are CK target genes and part of a negative-feedback loop regulating cytokinin levels. There are 12 type-B *ARRs* and 10 type-A *ARRs* in *Arabidopsis* (Hwang et al., 2012). Disruptions in cytokinin signalling can lead to a wide range of defects, ranging from root to shoot. For example, the triple type-B *arr1 arr10 arr12* mutant has fewer seeds due to fewer ovules, a reduced replum, septum fusion defects, less transmitting tract tissue, coupled with fewer fruits per plants and an overall shorter growth (Reyes-Olalde et al., 2017).

Cytokinins and auxins are closely linked and affect each other's activity domains. The first hypothesis about this CK/auxin interaction was formulated by Skoog and Miller (1957) who found that both hormones are needed for root and shoot growth and formation. In the root auxin and cytokinin have antagonistic roles (Argueso et al., 2010). This is however reversed in the shoot apical meristem (SAM) where the two hormones function cooperatively (Argueso et al., 2010, Jones et al., 2010). Altering CK levels has an effect not only on CK-related genes, such as the CKXs, but also on auxin biosynthesis (*IAA1*) genes, auxin transporters (PIN7, TAA1) and auxin response genes (*ARF6* and *AINTEGUMENTA*), showing a strong interaction between the two hormonal pathways (Reyes-Olalde et al., 2013, Reyes-Olalde et al., 2017).

### *1.3.3 Gibberellin*

The third hormone of importance in gynoecium patterning is gibberellin (GA). Gibberellins in plants are derived from geranyl geranyl diphosphate (Hu et al., 2008). There are several bioactive forms and in *Arabidopsis* GA4 is the major one (Hu et al., 2008). Gibberellins are involved in the growth and function of various reproductive tissues (Swain and Singh, 2005). They are necessary for pollen tube elongation, seed development, silique elongation as well as DZ specification and valve margin development (Hu et al., 2008, Kay et al., 2013). Mutants with very low GA concentrations fail to develop normal flowers, are semi-dwarfed, are delayed in flowering, have darker leaves, are infertile and have smaller siliques due to a reduced seed set (Fleet and Sun, 2005, Hu et al., 2008, Swain and Singh, 2005).

There are several genes and enzymes necessary for bioactive GA production. The enzymes GA 20-oxidase (GA20ox) and GA 3-oxidase (GA3ox), which are both involved in the late stages of biosynthesis, are especially important for the control of GA levels (Fleet and Sun, 2005). *Arabidopsis* has four *GA3ox* genes (*GA3ox1 – GA3ox4*) which all have unique expression patterns in the fruit and are required for patterning and growth (Hu et al., 2008). After fertilization, GA biosynthesis is triggered in the placenta (by *Ga3ox1*), and the endosperm (by *Ga3ox4*) leading to the elongation of the silique.

Gibberellin expression is perceived by soluble receptor proteins named GA-INSENSITIVE DWARF1-3 (GID1-3) (Hauvermale et al., 2012, Nakajima et al., 2006). These regulate growth by degrading DELLA proteins which are growth-repressors (Santner and Estelle, 2009). Five DELLA protein genes are present in *Arabidopsis* (*GA-Insensitive* [*GAI*], *Repressor of gi1-3* [*RGA*], *RGA-like1* [*RGL1*], *RGL2*

and *RGL3*) (Fleet and Sun, 2005). Loss of function of DELLA genes leads to an increase in GA and a taller, slender phenotype (Hauvermale et al., 2012). DELLA degradation also involves auxin. Auxin produced by the ovules leads to GA biosynthesis in the pericarp which in turn causes the release of growth repression by degradation of DELLAs (Fuentes et al., 2012, Sundberg and Ostergaard, 2009, Zhao, 2010).

Gibberellin is required for the valve margin development and, together with *IND* and *ALC,* enables correct fruit opening. The GA biosynthesis gene *GA3ox1* is, at stage 15, expressed at the same location as *IND,* namely in the valve margins and the septum (Arnaud et al., 2010, Liljegren et al., 2004). In *ind-1* mutants this expression is reduced, strengthening the link between the two (Arnaud et al., 2010). Moreover, reducing GA expression, for example in the *ga4-1* mutant or when expressing the gibberellin inactivating gene *Ga2ox* under the IND promoter, leads to incorrect valve margin development and subsequent less dehiscent fruits (Arnaud et al., 2010). The defects are not as severe as in *ind* single mutants however. This is due to the fact that the LL is unaffected in GA mutants and only the SL does not form correctly (Arnaud et al., 2010). The development of the SL specifically requires *ALC* and indeed GA and *ALC* act together in that region. ALC can bind with the DELLAs GAI, RGA and RGL2, preventing activation of its downstream targets (Arnaud et al., 2010). Upon GA increase, the DELLAs are degraded, ALC is released and the SL is specified. This pathway demonstrates the tight link between a hormone and two transcription factors. They are all required to function correctly for the patterning of the valve margin and hence for correct fruit opening.

A connection also exists between gibberellins and cytokinins. CKs have a cell division activation role and GAs promote cell elongation. Knotted 1-like homeobox (KNOX) proteins control the balance between these two hormones in the meristem. High CK levels couples with low GA in the SAM for instance is important to prevent cell differentiation and maintain pluripotent cell fate (Hay and Tsiantis, 2010).

Interplay between several hormones is necessary for gynoecium patterning in *Arabidopsis*. Auxin, cytokinin and gibberellins are only some of the many hormones involved and display complex interactions.

Due to the strict control of each factor involved in fruit patterning and the many interactions between hormones, between genes and between the two, alterations in the levels of one can

have widespread effects on development. Mistakes in early fruit growth can significantly affect seed development and therefore yield. Consequently, it is important to understand the individual parts to grasp what other factors are involved. This knowledge can then be used to produce crops with better yields.

### 1.4 Natural variation, discontinuous and continuous traits

As generations go on, genetic mutations arise. Some can be fatal and some may cause a disadvantage and some may give the individual an advantage over its peers. This is true for all living beings in all kingdoms. The variation thus arising is termed natural variation. In plant breeding, natural variation has been exploited for a long time to cross different plants to make progeny with improved features and traits. In science too, the variation is useful. Here, the origin of variation is further studied and can lead to advances in knowledge which can sometimes be exploited in breeding too.

Variation can be either continuous or discontinuous. Discontinuous variation results in just two phenotypes – such as plants producing seed or sterile plants. Continuous variation however gives rise to a number of phenotypes which are typically present in a distribution that can be plotted as a bell-shape curve – such as plant height (Griffiths, 2000). In the breeding of higher-yielding plants, most traits are continuous. Some however, such as disease resistance tend to be discontinuous where plants are either completely resistant or completely susceptible.

In breeding, only a fraction of the tools are allowed to be used compared to in science. Therefore, to uncover the function of a gene, knock-out mutants are often worked on where the effect of the gene is much more noticeable than when only reducing its levels. Furthermore, specific mutations can be acquired by various ways to investigate a gene function or a specific mutation further. This can be useful in identifying a region of a gene to select for in breeding, for example. While this describes an approach from the lab to the field, likely from a model species to a crop, the other way is also an option. This is due to the fact that natural variation in the wild occurs all the time and using techniques such as QTL studies and GWAS this variation can be studied and the cause of small differences identified. This can then be the basis of research.

### 1.5 Aim of the thesis

The aim of this Ph.D. project was to identify new avenues to increase yield of oilseed rape by studying the pod and seed growth. This was tackled first by gaining an understanding of how the pod grows and which traits correlate with each other. This knowledge was important to gain an overview of pod growth and understand the relationship between different traits. One such correlation, between stomata and pod length, was further investigated. From there, a genome wide association study (GWAS) was carried out on key traits to identify genes previously not associated to specific traits. This yielded candidate genes that were followed up in *Arabidopsis* to infer their role in fruit development.

# Chapter 2

## 2 General Materials and Methods

## 2.1 General methods used throughout the thesis

### *2.1.1 Plant material*

*Arabidopsis thaliana* ecotype Col-0 was used in all experiments unless otherwise specified. All mutants used were also in that background. Plants were grown in controlled environment rooms at 22°C and a 16h photoperiod. Plants were grown in 70% relative humidity and light levels around 150µmols at tray height.

### *2.1.2 DNA extractions*

#### *2.1.2.1 Edwards' quick DNA extraction for* Arabidopsis

One small young leaf was collected in an Eppendorf tube. The tissue was roughly ground using blue plastic grinders. 200µL Genomic extraction buffer (200mM Tris HCL pH7.5, 250mM NaCl, 25mM EDTA, 0.5% SDS) was added and the material ground some more. This was then centrifuged at 13000rpm for 5 minutes. The supernatant was added to 150µL propan-2-ol and, after being shaken for 2min, spun for 7min. The supernatant of that spin was poured off and the remaining DNA pellet washed in 50µL 70% ethanol. This was then air-dried and the DNA resuspended in 30-50µL $H_2O$ (modified from Edwards et al. (1991). DNA was stored at -20°C.

#### *2.1.2.2* Arabidopsis *DNA extraction carried out by the DNA extraction service*

Young *Arabidopsis* leaf tissue was collected and given to the DNA extraction service run by Richard Goram. He used a modified version of Palletta et al. (2003). 240µL isopropanol was added to each collection tube and frozen. Extraction buffer (0.1M Tris-HCL, pH 7.5, 0.05 EDTA pH8, 1.25% SDS) was pre-heated to 65°C and 10mg/ml RNaseA added prior to use. 333µL warm extraction buffer was then added to each tube and the plates were shaken for 3min at 30Hz. Then, plates were pulse spun at 3000rpm before being incubated at 65°C for 45min. Plates were then cooled for 15min in the fridge before 167µL 6M ammonium acetate was added. Plates were shaken for 15sec and left in the freezer for 10-15min. Plates were then centrifuged for 15min at 5000rpm. 400µL of the supernatant was pipetted into cold tubes with 240µL isopropanol per tube. This was then shaken for 15sec before being pulse spun. Samples were stored in the freezer for 10-15min for the DNA to precipitate and were then centrifuged for 15min at 5000rpm to

pellet the DNA. Next, the supernatant was tipped off and all the fluid removed by tipping tubes upside down on a paper towel. The remaining DNA pellet was then washed in 350µL of 70% ethanol and centrifuged for 15min at 5000 rpm before discarding the supernatant. Plates were left to air-dry overnight or at 65°C for 30min before the pellet was resuspended in 200µL of water. This was then left to dissolve at room temperature for an hour or for 15min at 65°C before a final vortex and a spin for 20min at 5000rpm was carried out.

#### *2.1.2.3 DNA extraction for* B. napus *and* B. rapa

Young leaf tissue was collected in an Eppendorf tube and 500µL 65°C pre-heated extraction buffer added (0.1M Tris-HCL, pH 7.5, 0.05 EDTA pH8, 1.25% SDS). The plant material was then ground using blue plastic grinders. Next, tubes were incubated at 65°C for 30-60min. Then, tubes were cooled in the fridge for 15min before adding 250µL 6M ammonium acetate (stored at 4°C). The tubes were shaken for 15sec and stored for 30min in the fridge. Then, the tubes were centrifuged for 15min. Subsequently, 600µL of the supernatant was added 360µL cold propan-2-ol. The tubes were shaken for 15min and then left for the DNA to precipitate for 5min. The samples were then centrifuged for 15min and the supernatant tipped off. The remaining fluid was discarded by tipping the tubes upside down on paper towels. Pellets were washed in 500µL of 70% ethanol and the pellets left to dry overnight. The next day the pellet was resuspended in 50µl $H_2O$.

### *2.1.3 General PCR protocols*

PCR reactions were prepared to a final volume of 20µl with 1x PCR buffer, 0.2mM dNTPS each, 1uM each of the forward and reverse primers, 1-2µL of Taq Polymerase and 50-100ng of gDNA. PCR reactions were carried out in a GSTORM Thermocycler and products were analysed using gel electrophoresis (1-2% agrose gels). An example PCR program used with Go Taq G2 Flexi DNA Polymerase (Promega UK Ltd) is presented below in Table 2.1.

Table 2.1 General PCR program

| | |
|---|---|
| 1 | 110°C Heated Lid |
| 2 | 94°C for 2min |
| 3 | 94°C for 20 sec |
| 4 | 60°C for 1min |
| 5 | 72°C for 30sec |
| 6 | Repeat steps 3-5 35 times |
| 7 | 72°C 5min |

### *2.1.4 Sequencing Reactions*

Samples were sent for sequencing with Eurofins (https://www.eurofinsgenomics.eu/) using the Mix2Seq service. A total volume of 17µL per tube was sent off. This included 2µL of the primer (at 10µM). Plasmid DNA was sent at a concentration of 50-100 ng/µL, PCR products (300-1000bp) at a concentration of 5ng/µL and longer PCR products at a concentration of 10ng/µL. Samples were viewed on the Eurofins website the day after sending them.

### *2.1.5* E. coli *heat shock and electroporation transformation*

*E. coli* DH5α competent cells were used. For heat-shocking 5µL of the plasmid or DNA ligation was added to 50µL competent cells and kept on ice for 20-30min. Cells were then heat-shocked at 42°C for 45sec before being returned to ice for 2-3min. 400µL SOC or LB were then added and the mixture shaken at 37°C for over an hour. Cells were then streaked on a plate with LB and the selection used. Plates were kept at 37°C overnight. The same was done for electro-competent cells except that they were subjected to a pulse of 4-6ms in a Biorad GenePulser (settings 2.5kV, 25µFD and 200 Ohms) instead of the heat-shocking.

### *2.1.6* Agrobacterium *electroporation transformation*

50-200ng plasmid was added to 50µL *Agrobacterium tumefaciens* strain GV3101 (25µL/ml gentamycin, 25µL rifampicin) competent cells and kept on ice for 20-30min. The cells were transferred to a pre-chilled cuvette and subjected to a pulse of 8-12ms using a Biorad GenePulser (settings 2.5kV, 25µFD and 400 Ohms). 400µL LB or SOC were added immediately and cells were

incubated at 28°C shaking for over 2hours. Cells were then streaked on selection plates (with cell-specific and plasmid specific antibiotics). The plates were then incubated at 28°C for 2 days.

### *2.1.7* Arabidopsis *floral dip transformation*

A single colony of *A. tumefaciens* containing the plasmid to be dipped was grown in 10mL LB with the appropriate antibiotics and incubated at 28°C overnight. 24h later, 2mL of that culture were transferred to 200mL LB with antibiotics and this was again grown at 28°C for two days. The culture was then centrifuged at 4000rpm for 25min and the supernatant poured off. The pellet was resuspended in 5% sucrose and 90µL Silwet in a total volume of 300mL per construct. Three pots containing nine plants each were dipped for roughly 1min. Before dipping, already formed siliques were removed from the plants. After dipping, the plants were tied inside bags and kept lying down away from direct sunlight for a day. Plants were then returned upright and the bags removed.

### *2.1.8 Miniprep, gel extraction and PCR clean up*

Miniprep, gel extraction and PCR clean up were carried out using the Qiagen kits and using their instructions.

### *2.1.9 GUS staining*

Organs were sampled in X-Gluc Buffer Solution (1mg/mL X-Gluc, 0.1M Sodium Phosphate Buffer, 1mM $K_3[Fe(CN)_6]$ + $K_4[Fe(CN)_6]$ Solution, in 1mM EDTA and 0.001% Triton) and vacuum infiltrated for 3 minutes. Samples were then incubated at 37°C overnight. The reaction was stopped by changing the solution with 70% ETOH. Samples were imaged using the Leica DM6000 microscope.

### *2.1.10 Confocal Microscopy*

Confocal microscopy of various GFP reporters was carried out using either using a Leica SP5 or a Leica SP8X microscope. Z-stacks were combined using Fiji (Schindelin et al., 2012).

### *2.1.11 Scanning Electron Microscopy (SEM)*

Stage 17 siliques were fixed in FAA solution (50% Ethanol, 3.7% Formaldehyde, 5% Acetic Acid). Samples were placed in FAA, vacuum infiltrated for 15 minutes and incubated for 4 hours at room temperature. Samples were then dehydrated using 50%, 60%, 70%, 80%, 90%, 95% and 100% Ethanol for 30 minutes each. Two more dehydrations with dry ethanol were carried out before using the Critical Point Dryer. Siliques were then gold coated and examined using Zeiss Supra 55 SEM using an acceleration voltage of 3 kV.

### *2.1.12 Seed measurements*

Seeds were measured using a Marvin Seed Analyser. All the seeds from twenty hand-opened pods were analysed. Mean seed surface area, width and length as well as a seed count were measured by the machine. This was then used to infer the mean seeds per pod.

# Chapter 3

## 3 General growth mechanisms within the *B. napus* pod

## 3.1 Introduction

Oilseed rape is a widely grown crop. Chapter 1 introduced its uses as well as what is known about some of the underlying mechanisms enabling a pod to grow. Oilseed rape has been bred for its seeds which are crushed for the oil. Therefore, some of the ways to increase yield of oilseed rape is to increase the seed size, the seed number per pod or the pods per plant. Another way to add value to the crop is to increase the quality of the oil (Fig. 1).

Fig. 3.1. Avenues of yield improvement. The yield of oilseed rape plants can be improved in several ways such as increasing the size of the seeds, the seeds per pod and the pods per plant. Scale bars are 1cm

Many of the factors involved in oil quality are known. Several are identified in *Arabidopsis* but only some have also been studied in *B. napus* to the point where crop improvement is possible. An early example is the discovery of *FATTY ACID ELONGASE1.1* and *1.2 (FAE1.1* and *FAE1.2)* which control the erucic acid contents (Fourmann et al., 1998). Erucic acids are desirable for the production of lubricants, paints and inks but undesirable for nutritional purposes. Therefore, cultivars have been bred for high as well as low levels of it. Similarly, *FATTY ACID DESATURASE 2* (*FAD2*) affects the levels of polyunsaturated fatty acids (PUFAs) which are undesirable. Reducing PUFA contents was indeed achieved by specific mutations in the *BnaC.FAD2.a* copy of the gene in

*B. napus* generated by EMS mutagenesis (Wells et al., 2014). Several accessions with PUFA levels of around 6% were generated which was a marked reduction from the 16% and 29% found in varieties Cabriolet and Tapidor, respectively (Wells et al., 2014).

This chapter will explore how a pod grows and how some of the yield-related traits correlate. This is done to understand the underlying mechanisms of pod growth which then allow a targeted approach to increase yield while being aware of the other factors which may be affected when focusing on one specific trait.

## 3.2 Materials and Methods

### *3.2.1 Plant material and sampling*

The 12U0245 double haploid (DH) population developed by Limagrain UK was grown in a field in Lincolnshire by Limagrain UK. The parents of the population are 210FR00309 and 211EU00114. In 2013/2014 there was one replicate per line in small plots. I selected this population to work on from several DH populations grown in the same field. I selected it based on pod length differences that were apparent from walking through the population. The rationale for working on this population was to eventually carry out a QTL analysis (the project then developed differently and no QTL study was done). Hence, in the 2014/2015 growing season the plants were grown in larger plots that were 6m x 1m with a guard of the cultivar Anastasia surrounding the plot. The plot design was done by Limagrain. To make up for only one replicate being present, I sampled each plot at three positions; from the left edge, the middle and the right edge. Well developed, representative pods were chosen from the main raceme or the secondary raceme. Fifty pods were sampled to ensure that at least 20 returned intact to the John Innes Centre for measuring. Pods were photographed and measured using ImageJ (Schneider et al., 2012). A macro was developed for pod measuring for the lines harvested in 2015 (the Macro can be found in Appendix).

The Diversity Fixed Foundation Set (DFFS) made of 87 accessions was grown in a triplicated design in a polytunnel at the John Innes Centre in 2013/2014 and again in 2014/2015. The plants were grown with six weeks of vernalisation time, regardless of their type. This ensured constant conditions for all and allowed flowering for most of them. Fewer accessions were grown in the second year due to elimination of those plants that did not flower in the first year. Plants were

sampled and measurements were made by Rachel Wells and Marlène Carrière, an MSc student under my supervision. I measured pod width. A full list of the accessions grown can be found in the Appendix Table 1.

The winter oilseed rape varieties were grown by Dekalb UK in a field near Cambridge in the 2013/2014 growing season. Varieties included Avatar, DK Cabernet, DK Extrovert, Excalibur and PR46W21, Compass, DK Excellium, DK Expower, DK Sensei and V295OL. The first five in this list were also used for pod strength measurements. 40 pods per plant were sampled in the field from three different plants for each variety. Pod and seed measurements were carried out as for the DH population.

The spring oilseed RV31 was used for the growth curve and plants were grown in a greenhouse at constant 19°C. The first inflorescence of three plants were sampled every three days from the moment the first buds were formed. Gynoecia were measured using either optical projection tomography (OPT) or using a light microscope. The timing of the first flower of the main inflorescence was given the timepoint 0 and subsequent flowers on the same inflorescence were timed using the plastochron (time between two flowers opening). The plastochron was calculated by counting the number of open flowers every day for 15 primary inflorescences. These values were averaged and gave a plastochron of 8.76 hours for RV31 under these conditions. The length of the gynoecia imaged using OPT were measured using the measuring tool in VolViewer (https://www.kitware.com/volview/).

### *3.2.2 Random Impact Testing*

The Random Impact Test was carried out as described by Morgan et al. (1998). Twenty pods per accession were grouped. Pods were calibrated in a calibration chamber at 23°C and 50% relative humidity for 5 days. Post-calibration 20 pods were placed into a cylindrical container of 20cm diameter and a height of 15cm. Five ball bearings of 1cm diameter were added. The container was then shaken horizontally at 4.95 Hz over a stroke of 51mm for 8secs. After each 8-second interval pods were analysed for being either damaged or broken. A pod is termed damaged if it is not intact but still contains the seed; the most common example of a damaged pod is a lost style. A pod is determined as broken when the seeds are lost. The test was over once all pods are broken. A decay curve was fitted using the data points from the test. Two separate curves were made for the damaged pods and the broken pods. Using the curve, a half-life where 50% of pods

are broken or damaged was determined. This value can then be used to rank accessions and for statistical analysis. First, correlations with pod weight and pod length were analysed. This could have an effect on results as the ball bearings may hit larger and heavier pods more often. If a factor did correlate covariates were calculated to adjust the half-lives to remove this effect. An Analysis of Variance (ANOVA) was then used to determine significant differences between varieties.

### *3.2.3 Optical Projection Tomography (OPT)*

Samples were collected in 100% ethanol and stored at 4°C until imaged. Samples were then rehydrated (70%, 50%, 30%, 2 x $H_2O$ for 30 min at room temperature) before being embedded in 1% low-melting point agarose (as in Sharpe et al. (2002)). Mounted specimens were then dehydrated overnight in 100% methanol before being cleared for 24 hours in 1:2 benzyl alcohol and benzyl benzoate (Sigma-Aldrich). Specimens smaller than 1cm were scanned with a prototype OPT device described by Lee et al. (2006). Larger specimens were scanned using a Commercial Scanner Biooptimics 3001. Images were visualised in 3D in a freely available software called VolViewer (https://www.kitware.com/volview/).

### *3.2.4 Tissue Culture and transformation of RV31*

Sterilized seeds (100% ethanol 2 min, 15% sodium hypochlorite 15 min, 3 water washes) of RV31 were sown on germination medium (20 seeds per 90 mm petri dish) and grown at 23°C with 16h day lengths for 4 days. Overnight cultures of bacterial suspension were spun down at 3000 rpm for 5min before being resuspended in 10ml liquid MS media. An O.D. of 0.1 to 0.3 was required. Cotyledons from 4-day old plants were excised by slicing through the petiole just above the meristem. The cotyledons were then placed onto co-cultivation medium petri dishes. Once all were excised cotyledons were inoculated, one by one, by briefly dipping the cut end of the petiole into the *Agrobacterium* suspension. Cotyledons were then returned to the co-cultivation plates (4.71g/L M&S medium, 30g sucrose, 4mg/L BAP, 2.5g/L gelrite, pH5.8) and sealed with micropore tape before being transferred to a 23°C culture room with 16h our days for 72 hours. The cotyledons were then moved to selection medium in deep petri dishes (4.71g/L M&S medium, 30g/L sucrose, 4mg/L BAP, 500mg/L MES, 8g/L bacto agar, pH5.8, 600mg/L augmentin, 15mg/L kanamycin) and moved back to the growth room. Two control plates without kanamycin including plants dipped and plants not dipped were also kept. After 2 weeks, explants were transferred to

fresh medium (4.71g/L M&S medium, 30g/L sucrose, 4mg/L BAP, 500mg/L MES, 8g/L bacto agar, pH5.8, 300mg/L augmentin, 15mg/L kanamycin). During this subculture, any white escape shoots were removed. Green transgenic shoots were excised and transferred to 100mL jars (75x50mm) containing 25mL of medium (4.71g/L M&S medium, 30g/L sucrose, 2.5g/L gelrite, 300mg/L augmentin, 15mg/L kanamycin, 0.5mg/L IBA, pH5.8). Shoots were maintained at 23°C with 16-hour day lengths until roots developed. After root elongation plants were transferred to sterile peat pots for further root growth before being transferred to the greenhouse. Plants were transferred to soil and maintained under shade within a propagator for the first week. This enabled the plants to adjust to the reduced humidity and increased light intensity. Once budding, plants were heat-shocked for 4 min and analysed for GFP expression using a Leica M205FA stereo microscope. When budding, plants were covered in plastic bags to prevent outcrossing. Seeds were collected and the next generation sown. GUS staining was done on leaf samples to test expression and plants were heat-shocked at budding and analysed for GFP expression at later stages. Leaf samples were sent to iDNA to get a measure of copy number. Plants with two copies of each construct were kept and grown until they were homozygous for each. Seed was bulked by growing plants in 2L pots in the greenhouse.

### 3.3 Results

First, to gain an understanding of the processes underlying pod development, pod length, pod width, seed surface area, seeds per pod and thousand seed weight (TSW) were measured on a field-grown double haploid population over two years, on a diversity fixed foundation set (DFFS) (measurements on the DFFS were done by Rachel Wells and colleagues) and on 11 winter oilseed rape varieties grown in the field in 2013/2014. The DFFS included spring and winter oilseed rape lines (SOSR/WOSR), kales, swedes, fodder rape and some exotic lines. Correlations were determined using GenStat and assembled in Fig. 3.2. In brief, longer pods correlated positively with seed number while smaller pods had larger and heavier seeds (negative correlation between pod length and TSW). Negative correlations were found between seed area and seeds per pod as well as seeds per pod and TSW. The DH population had a positive correlation between seed area and pod width in both growing years. One dataset showed a negative correlation between seed area and pod length and between pod length and TSW and a positive correlation between TSW and pod width. A positive correlation was present in the 2014 DH population between pod length and pod width.

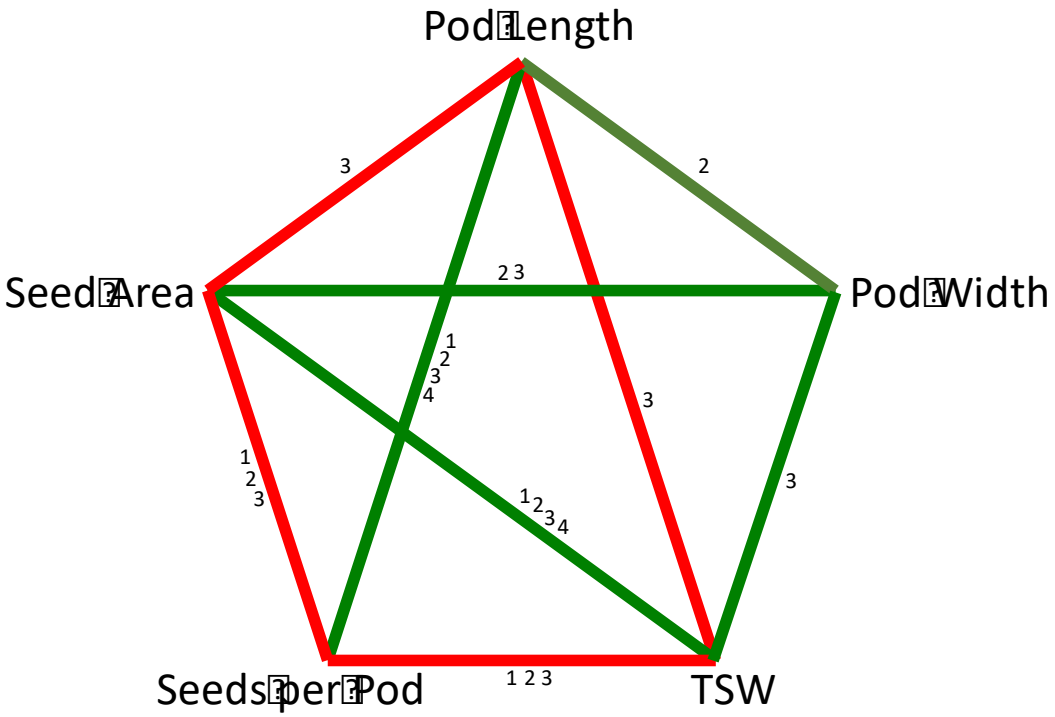



Fig. 3.2. Correlations between yield related traits were often found in several datasets. Those that were consistently present were the positive correlations of pod length to seeds per pod and the seed area to thousand seed weight (TSW). Green lines are positive correlations, red lines are negative ones. Numbers denote the dataset the correlation was found in: (1) 2014 DFFS, (2) 2014 DH population, (3) 2015 DH population and (4) 11 WOSR varieties grown in the 2013/2014 season (note that pod width was not measured in this dataset). Correlations were determined using GenStat.

### *3.3.1 Growth dynamics in the early stages*

The correlations explained how the seed and pod traits were connected but they did not give information on how the pod itself grew. To elucidate this, RV31 inflorescences were sampled and gynoecia measured every three days from the moment buds were present. Measurements were stopped when fruits started elongating post fertilization as the differences in pod growth were too large. The differences arose due to differences in fertilization. The results were graphed on a log scale and showed constant growth throughout (Fig. 3.3). A straight trend-line through the gynoecia length measurements had a fit of 97%, supporting a very constant growth trend. The width was also constant throughout development and the $R^2$ value here was 85% (Fig. 3.4).

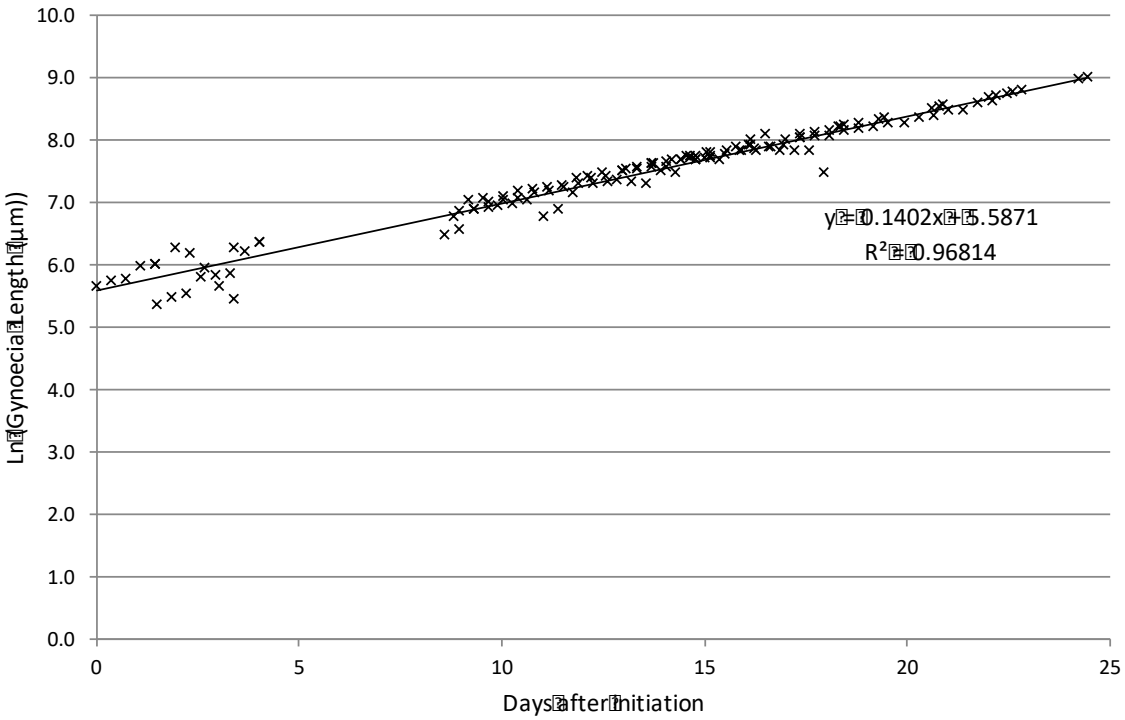



Fig. 3.3. *B. napus* RV31 pod growth. Gynoecium length was measured from the moment of initiation. Pod length grows steadily throughout development. Measurements were stopped when pods began elongating post-fertilization as large differences between plants were then present.

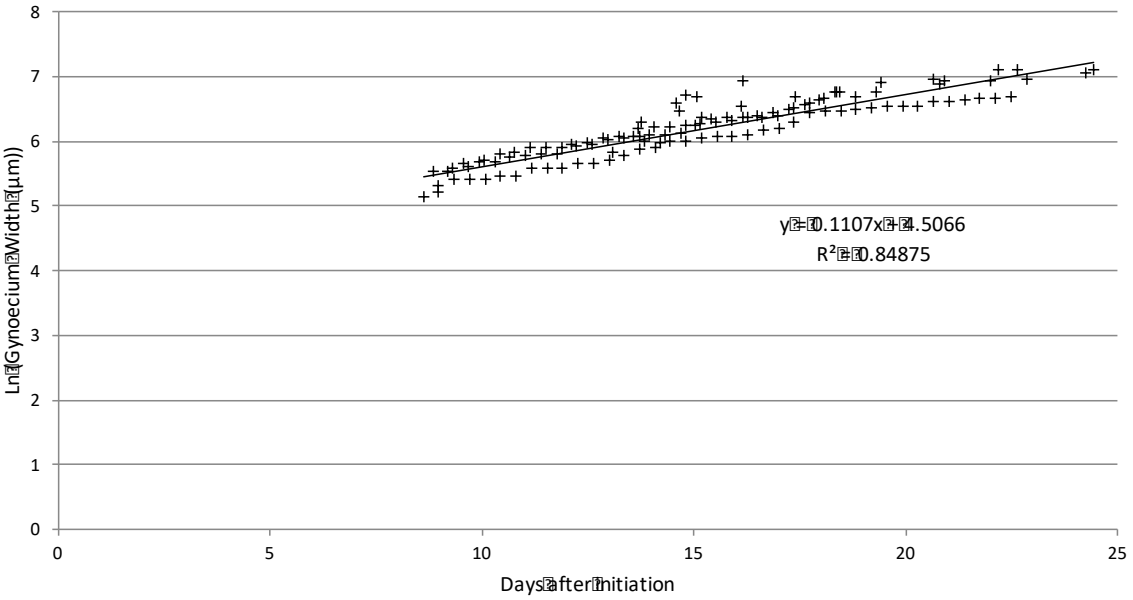



Fig. 3.4. *B. napus* RV31 gynoecium width. Gynoecium width was determined over time using OPT images or photographs and showed a steady growth. Measurements were taken using Volviewer and ImageJ.

Plotting length versus width can give an idea of whether the fruit is growing anisotropically, *i.e.* more in one direction or the other. If the width is on the y-axis and the length on the x-axis, the slope of the trend-line can give an indication as to which parameter is growing more. If the slope is less than 1, the fruit is growing more in the length than in width whereas if the slope is more than 1, the width is growing faster. Plotting the relationship for the RV31 data supported the shape of a cylindrical fruit by showing that the fruit grew more in length than in width, as demonstrated by a slope of 0.79 (Fig. 3.5).

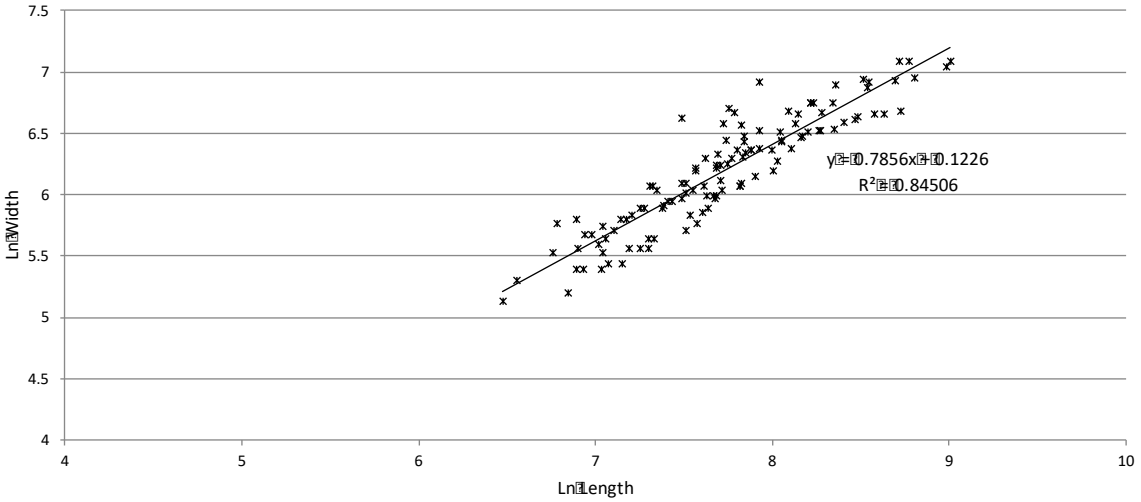



Fig. 3.5. RV31 gynoecium length vs width. The fruit grew more in length than in width. This also confirmed the steady growth displayed in Fig. 3.2 and Fig. 3.3.

### *3.3.2 The direction of cell growth can be inferred using sector analysis*

The direction of growth can be analysed using the sector analysis system. This has been used previously to describe the growth of *Arabidopsis* petals, *Arabidopsis* fruits and *Capsella* fruits (Eldridge et al., 2016, Sauret-Gueto et al., 2013). *B. napus* accession RV31 was transformed using tissue culture with a *Cre* recombinase under the control of a heat-shock promoter (Gallois et al., 2002). Another plant was transformed with a construct under a 35S promoter driving a GUS construct surrounded by *lox* sites and followed by mGFP5. Transformed plants were then crossed and double homozygotes created. Under normal conditions plants will express the GUS construct but upon heat-shock over 37°C the *Cre* recombinase is activated and cuts along the *lox* sites to generate cells expressing GFP (Fig. 3.6). This signal is then passed on to any subsequently dividing

cells creating sectors. Analysing the sectors several days after heat-shocking can give information on the direction of growth.

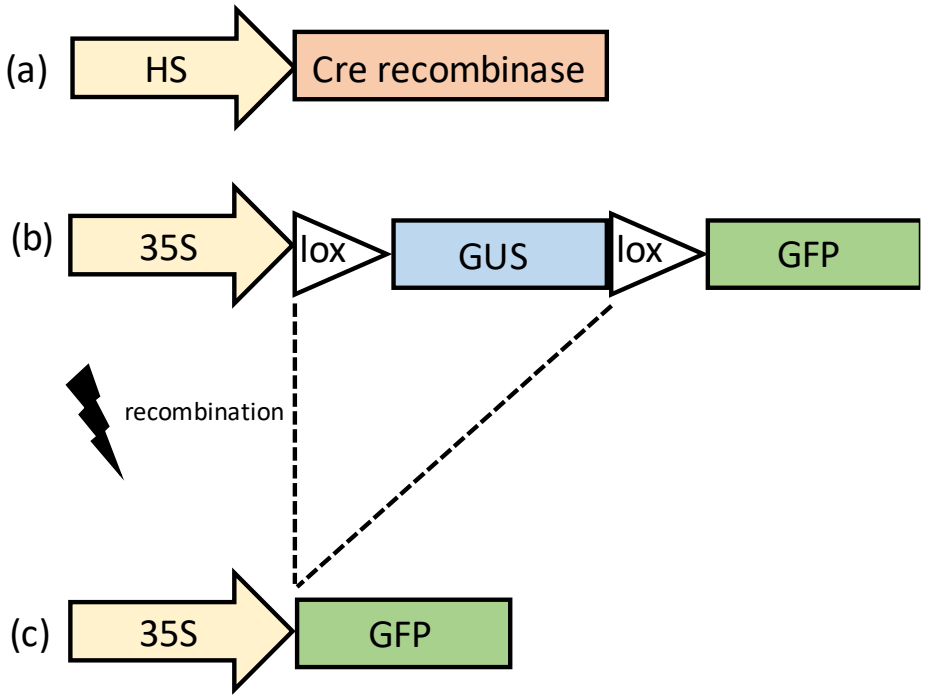



Fig. 3.6. Sector analysis constructs transformed into *B. napus* RV31 for the sector analysis. Constructs (a) and (b) were transformed into RV31 using tissue culture. Plants were crossed until double homozygous plants were present. (c) Upon heatshock, the cre recombinase cuts along the lox sites and some cells will subseqeuntly express the GFP instead of the GUS.

*Arabidopsis* and *B. napus* have the same cylindrical-shaped fruits and in *Arabidopsis*, cell files occur in straight vertical lines along the fruit (Eldridge et al., 2016). The same was found in *B. napus* where sectors were present in long, narrow cell files. Also in the style, which is much larger in *B. napus* than in *Arabidopsis*, sectors showed long cell files that were mostly one cell in width (Fig. 3.7). This shows that cells tend to divide mostly along the same axis leading to strong anisotropic growth primarily in length but not width.

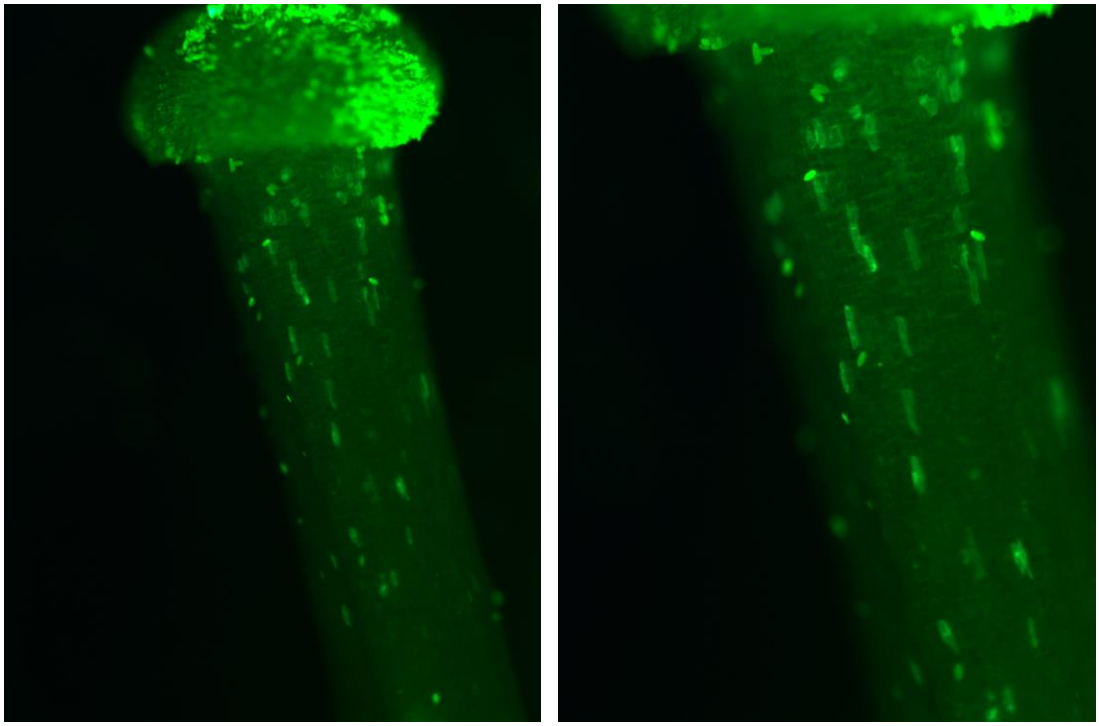


Fig. 3.7. Sectors in the style. Pods of the sector analysis displayed cells expressing GFP post heat-shocking. Sectors were formed as cells divided and passed the GFP signal on to the new cells. The direction of growth can be gleaned from the sectors. RV31 styles and stigma are visible here. The growth in the style was in one direction with narrow sectors of one or two cell files across.

### *3.3.3 Assessing differences in various* B. napus *accessions*

Once I knew the general growth dynamics of *B. napus* pods were established, I wanted to know how pod length differences between accessions arise. To this end, I chose four lines from the DFFS for further study. The lines were chosen for their varied valve lengths when grown in a polytunnel at the JIC with six weeks vernalisation. AmberxCommanche (AmbxCom) and Darmor (Dar) had large valves while Quinta (Qui) and NingYou (Nin) had short ones (Table 3.1). For detailed measurements, three plants per accession were grown in 1L pots with 10 weeks of vernalisation to ensure a fairly even flowering time. These plants hence had four more weeks vernalisation time than those grown previously in the polytunnel. The plants different in several growth characteristics such as for example plant height which was reduced for the plants with more vernalisation time (not measured). Fifty pods per plant were sampled and valve lengths measured using ImageJ (by Marlène Carrière). While the longer vernalisation time achieved more even flowering, it caused a change in final valve length where Nin was larger than Dar, although this was not significant. Qui also was one centimetre larger than it had been previously. This reduced the differences between the final valve and pod lengths (Table 3.1 and Fig. 3.8).

Table 3.1. Valve lengths of different *B. napus* accessions. The lines were chosen for their varied lengths when grown in 2014 in a polytunnel with six weeks vernalisation. When grown again in the greenhouse with ten weeks vernalisation the valve lengths did not differ as much anymore.

| Line | Spring/Winter | Valve Length 2014 | Valve Length GH |
|---|---|---|---|
| **Quinta** | W | 4.19 | 4.9 |
| **NingYou** | Chinese | 5.64 | 5.87 |
| **AmbxCom** | W | 6.81 | 5.62 |
| **Darmor** | W | 7.50 | 5.75 |

Next, Marlène and I wanted to see whether these differences were already present pre-fertilization. To this end, stag 10 gynoecia were imaged using SEM. At this stage, the gynoecia are just starting to develop the stigmatic papillae. The length was measured and again Qui was the shortest. At this stage, however, AmbxCom was significantly larger than Nin and Dar (Fig. 3.9). This was lost later on when Dar and Nin acquired a total length that was comparable to that of AmbxCom (Fig. 3.8).

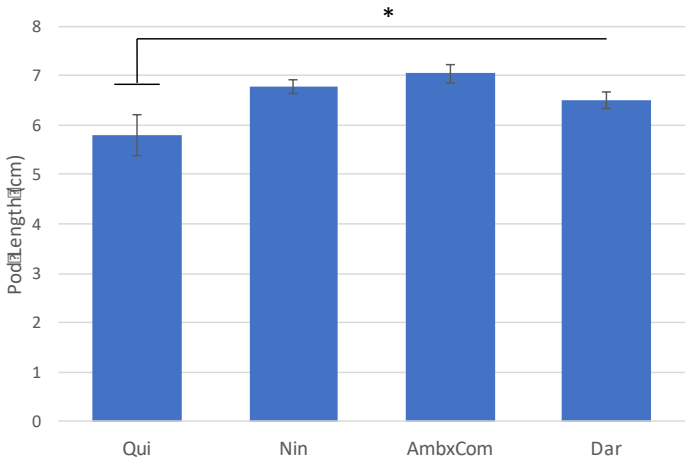



Fig. 3.8. Final pod lengths of greenhouse-grown plants with 10 weeks vernalisation. Nin, chosen previously for its small valves, had large pods and Qui too had larger pods than when previously grown with six weeks of vernalisation. Plant averages were compared and error bars are standard deviations. *p<0.05

To further evaluate what constitutes the differences in gynoecium length, the cellular arrangement was observed. Cells were measured at stage 10 and the number of cells over 180µM were measured at the base, middle and top of the gynoecia. Cell numbers were the same regardless of accession and location on the gynoecium, suggesting a constant cell size (data not shown). The total number of cells per gynoecia was then calculated, making the assumption that cell sizes were constant throughout the gynoecium. Plotting the results told us that the assumption was correct as the graph was the same as when plotting the pod length at stage 10

(Figs. 3.9 and 3.10). We can therefore say that, at least at the early stages, the differences in pod length were due to cell numbers and not cell sizes.

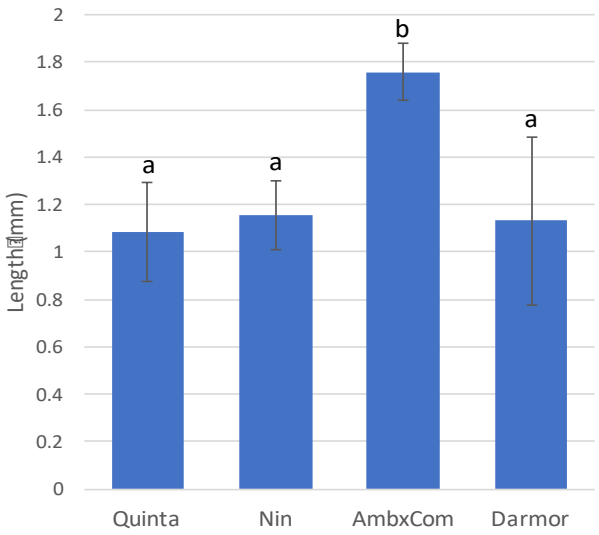



Fig. 3.9. Stage 10 gynoecia lengths. AmbxCom had significantly longer gynoecia than the other three lines examined. Measurements were taken using ImageJ on images taken using the SEM.

Measurements made on pods at stage 17 showed differences in cell lengths. AmbxCom and Dar had significantly smaller cells than the other two lines ($p<0.05$) (Fig. 3.11). As these pods were chosen for their larger pods and they have smaller cells it can be concluded that larger pods arise due to increased cell division rather than cell size. These measurements were also carried out by Marlène Carrière.

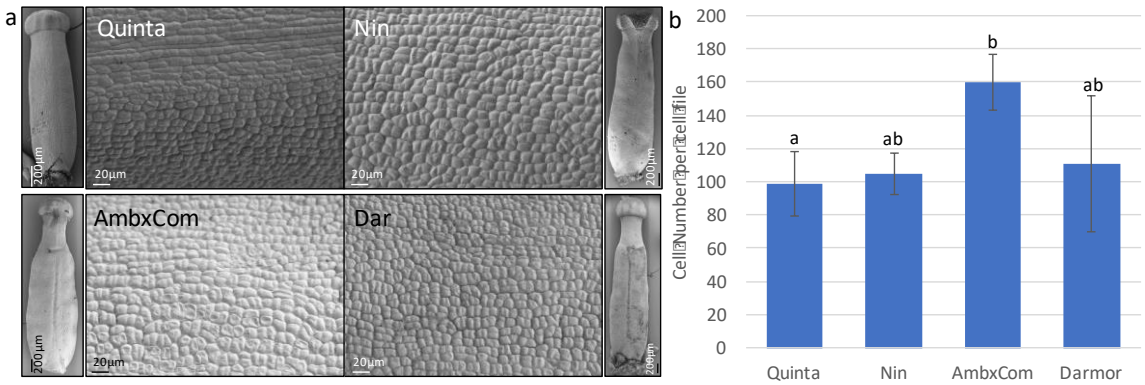



Fig. 3.10. Pre-fertilization cell arrangements were similar (a). SEM images of the entire fruit and a zoomed in section show the individual cells (b). The longer gynoecium of AmbxCom was therefore due to more cells and not due to larger cells. Cell measurements were taken at different places along the gynoecium and were constant. These values were then used to calculate the total cells per gynoecium. Error bars are standard deviations and letters denote significantly different lines (determined by ANOVA, Tukey's 5% multiple comparison).

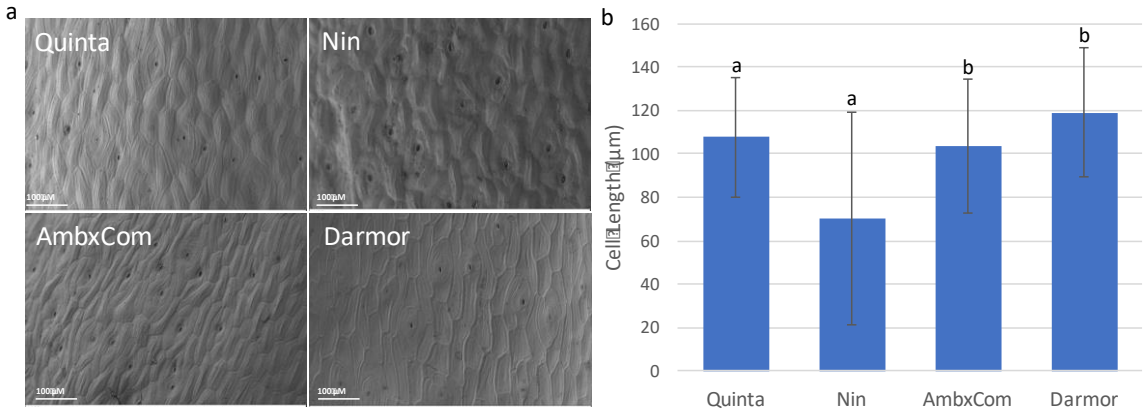



Fig. 3.11. Stage 17 cellular detail. The cellular arrangement did not differ much as seen in the SEM images (a). Cell lengths differed slightly between the lines with Quinta and Nin having larger cells than AmbxCom and Darmor (b). Error bars are standard deviations. Letters denote different significance groups (Tukey's 5% as determined by ANOVA).

### *3.3.4 Flow Cytometry*

Flow cytometry was then carried out to see whether the pods had different levels of DNA ploidy. Pods 15 days after flowering were used. At this stage, they were fully elongated. Roughly 1cm of valve material was used for the analysis. The accessions had 2C, 4C and 8C cells mainly with roughly the same distribution in the different lines (Fig. 3.12). The highest fraction in all except Nin were 4C cells. Only Dar had a small section of 16C cells but not more than 2% of the total. This suggested that endoreduplication did not play a major role in pod elongation in *Brassica napus*.

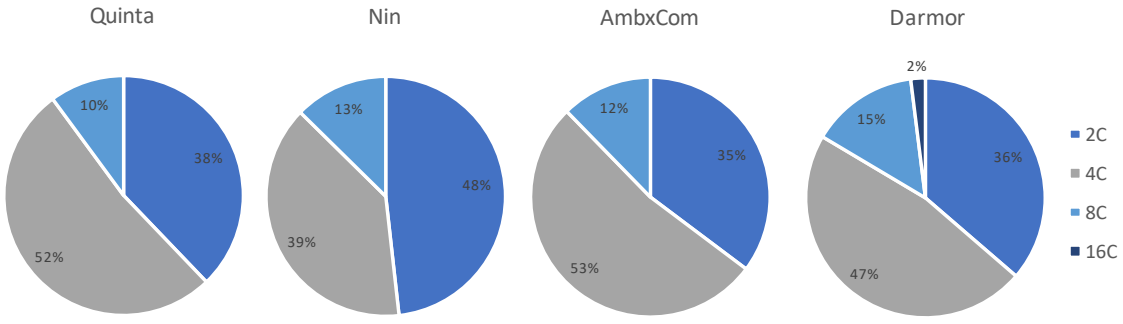



Fig. 3.12. Ploidy of pods. All accessions behaved similarly, which is in accordance with the cell sizes not differing much either.

### *3.3.5 Longer pods contain more seeds*

Next the mean amount of seeds per mature pod was analysed. Quinta had very few seeds and AmbxCom the most with 23 seeds per pod on average (Table 3.2). Seeds per pod correlated positively with the final pod length as well as the length at stage 10 (pre-fertilization), suggesting that length pre-fertilization can be used to infer about the final length of the fruit. The valve lengths did not match the trend entirely however. Here, AmbxCom was not the largest. This, combined with the fact that it had the most seeds, suggests that AmbxCom's pods were better filled than those of Nin and Dar with their bigger valves but fewer seeds.

Table 3.2. Summary of seed and pod traits. Significantly different means are denoted by letters as determined by ANOVA and Tukey's 5% intervals. Stats were carried out in GenStat.

| Accession | Seeds per Pod | Mean seed surface area ($mm^2$) | Mean valve length (cm) | Mean pod length (cm) |
|---|---|---|---|---|
| Qui | 9.59 | 4.68 | 4.895 a | 5.77 a |
| Nin | 19.81 | 3.22 | 5.869 b | 6.79 b |
| Amb x Com | 23.81 | 3.6 | 5.621 b | 7.05 b |
| Dar | 21.26 | 3.26 | 5.75 b | 6.51 b |

### *3.3.6 Plants with larger pods produce fewer fruits*

Next, to determine whether plants compensate for producing smaller pods by setting more pods in total thereby ending up with a comparable number of seeds per plant, I counted pods per plant. The total number of branches was recorded to give a measure of the differences in plant architecture. Fertilized pods were counted six months after being removed from vernalisation. Most plants were entirely dry, some still had very few pods that were slightly green. The mean overall pod number ranged from 279 pods (AmbxCom) to 448 (Qui) per plant (Table 3.3). Only Qui was significantly different from the others with more pods ($p<0.05$). Qui was also the line with the smallest pods so this could be a mechanism to produce a similar amount of overall yield. Total yield however was not measured.

In terms of the branches there was a clear and significant difference with Nin having more than twice as many as AmbxCom and Dar (Table 3.3). Each branch was however filled with fewer pods.

This was also mirrored in the total pods found on the primary raceme where Nin had fewer than the other accessions (Table 3.3).

Table 3.3. Plant architecture differed as did numbers of pods. Total pods, branches and pods on the primary raceme differed significantly between accessions. Letters denote different significant groups as determined by ANOVA and Tukey's 5% interval (using GenStat).

| Line | Total Pods | Total Branches | Pods on primary raceme |
|---|---|---|---|
| **Qui** | 448.33 b | 30.49 b | 51.67 b |
| **Nin** | 319.25 a | 44.5 c | 27.75 a |
| **AmxCom** | 279.4 a | 19.5 a | 60.2 c |
| **Dar** | 284.17 a | 19.17 a | 45.5 b |

### *3.3.7 Different ways of finding novel variation*

Scientists often use different techniques and different plant materials to identify genes underlying agronomic traits than breeders. In this study, a diversity fixed foundation set (DFFS) was described to assess the variation in yield traits related to seed and pod characteristics. The accessions included in the DFFS comprised plants from different geographical origins with different vernalisation requirements that were bred for different aspects. This meant that much variation was present in the population which was reflected in the measured traits which displayed large standard deviations in all traits assessed (Table 4.4). A double haploid population was also phenotyped. This was a population developed by a breeding company (Limagrain UK) and was made purposefully to develop new oilseed rape varieties. This population comprised plants that differed in pod and seed traits. The DH population was made up of 95 accessions while the DFFS had 88 measured lines. Twelve winter OSR varieties were also measured (Avatar, Compass, DK Exalte, DK Exclaim, DK Cabernet, DK Excellium, DK Expower, DK Extrovert, DK Sensei, Excalibur, PR46W21 and V295OL) and here the least variation was found. This could be due to the few varieties tested but also likely due to the fact that OSR on the market are all bred for high yield and therefore do not differ that widely.

### *3.3.8 Pod strength differed widely between accessions*

The trait that had the highest standard deviations for all populations tested was pod strength (Table 4). Pod strength is an important characteristic as the pods of oilseed rape can open pre-

maturely, for example during a hailstorm, and seed is then lost. This is inconvenient for several reasons: firstly, the prematurely dispersed seeds lead to a reduction in yield and secondly, the seed that have fallen to the ground can germinate in the following year giving rise to volunteer plants which contaminate crops grown in following years. Pod strength was measured using the Random Impact Test (RIT) developed at the John Innes Centre. Here, twenty pods were shaken in a container and the time for half of all pods to break or damage was determined. Two different values can be determined; the time it takes for a pod to be damaged, but not lose seed and the time for a pod to break, at which point seeds are lost. These are both important values as the former is important in the field because once pods are no longer closed pathogens and insects can enter the pod and the seeds may get damaged. The latter is important as once a pod loses seed, it can no longer be harvested. Figure 3.13 shows the decay curves for a strong and a weak variety for the break point. Two biological and three technical repeats were used, leading to a total of 6 runs per variety. The RIT supplies a value that is easy to compare between lines and also allows ranking. The strongest marketed varieties were as strong as the strongest line of the DH population (Table 3.4). However, the DFFS had one outlier which was much stronger than the other accessions. This was XiangYou which took almost twice as long to shatter in the RIT assay as the strongest of the DH plants or the varieties tested. This demonstrates the variation present in the DFFS and the value of using accessions from different backgrounds to find novel variation.

Table 3.4. Summary statistics of three B. napus populations. The OSR varieties had the smallest variation within their traits as displayed by their small standard deviations. Large differences in pod strength were present in all populations, showing that much amelioration can still be achieved.

| **Population** | **2014 DH** | | | **2014 DFFS** | | | **OSR Varieties** | | |
|---|---|---|---|---|---|---|---|---|---|
| | Min | Max | StDev | Min | Max | StDev | Min | Max | StDev |
| Pod Length (cm) | 3.23 | 9.04 | 1.03 | 3.49 | 10.67 | 1.04 | 5.99 | 8.69 | 0.70 |
| Pod Width (cm) | 0.43 | 0.67 | 0.05 | 0.42 | 2.27 | 0.54 | Not measured | | |
| Seed Surface Area ($mm^2$) | 6.35 | 9.19 | 0.60 | 4.26 | 8.57 | 0.79 | 5.52 | 6.51 | 0.43 |
| Seeds per Pod | 3.71 | 30.44 | 5.10 | 5.70 | 31.29 | 5.59 | 14.45 | 29.60 | 3.91 |
| TSW (g) | 3.91 | 8.31 | 0.78 | 2.13 | 8.20 | 1.15 | 3.92 | 6.84 | 0.70 |
| RIT50 Damaged Pods | 0.83 | 22.80 | 4.17 | 0.75 | 46.96 | 5.55 | 9.82 | 25.00 | 6.21 |
| RIT50 Broken Pods | 4.70 | 112.5 | 19.91 | Not measured | | | 18.30 | 74.96 | 24.26 |

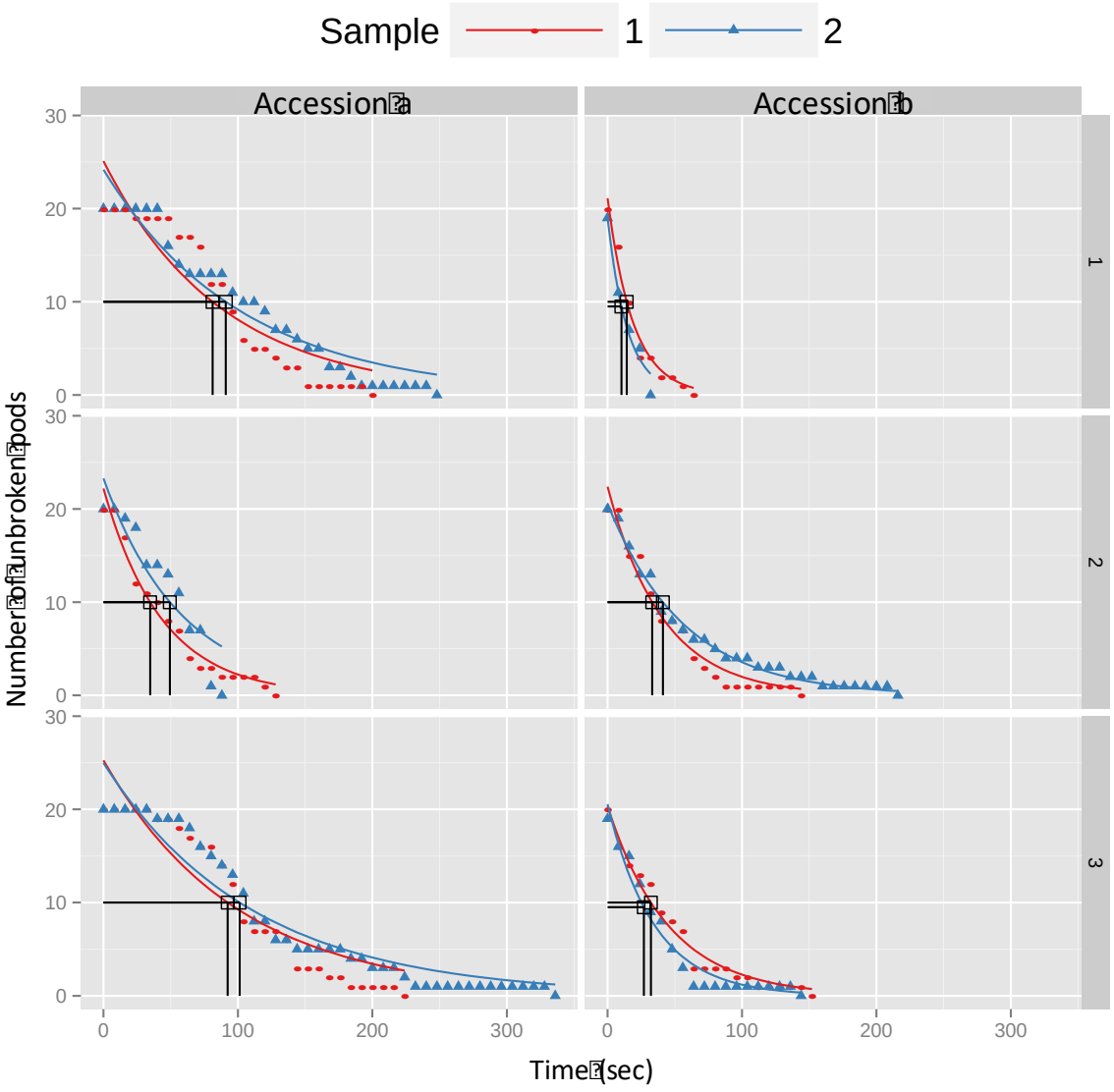



Fig. 3.13. Example curves from the Random Impact Test. A strong and a weak accession are shown. The test begins with 20 intact pods and is carried out until all pods are broken. The black lines denote the half-point at which half the pods are broken. This value can be used to compare accessions.

## 3.4 Discussion

### *3.4.1 Correlations between yield parameters are consistent*

The relationship between pod length and seeds per pod occurred in all populations used here. This was also present in several studies using Chinese OSR populations (Qi et al., 2014, Wang et al., 2016, Zhang et al., 2011). From work in *Arabidopsis*, it is known that seed-derived signals are

needed for fruit growth and they involve several hormones such as gibberellin and auxin, which could explain this phenomenon (Fuentes et al., 2012). Only one dataset, the 2015 DH one, had a negative correlation between pod length and TSW. This was a strong correlation however with $p<0.001$ and a correlation coefficient of -0.5612. Several other studies show a positive correlation between these two traits however (Qi et al., 2014, Wang et al., 2016, Yang et al., 2012). It is noteworthy that the studies found were all conducted in China where the growing conditions are likely to be different to the ones in the UK. It is possible that the environmental effect could give rise to these different correlations. This is supported by the data presented here where the DH population grown in two consecutive years showed different trends in the two different years. The effect of environmental conditions was also seen with the DFFS plants grown with different vernalisation times having different morphologies and pod lengths.

A negative relationship between seed area and seeds per pod was present in the DFFS and in the DH population. Understanding and breaking down that correlation could be an interesting avenue for research and could give rise to plants with higher yield.

Furthermore, a negative relationship between seeds per pod and total pod number per plant was present in the four more closely studied accessions. This could be due to assimilate distribution within the plant. In two Chinese studies, pod number was shown to significantly correlate with seed yield, suggesting that pod number could be an important characteristic to focus breeding efforts on (Shi et al., 2009, Shi et al., 2015). Certainly, the breeding of plants with many, well-filled pods is desirable for crop improvement.

### *3.4.2* B. napus *fruits grow steadily in the early stages*

The RV31 *B. napus* gynoecium grew steadily at a constant pace in the early stages. This differs to results reported in *Arabidopsis* and *Capsella* where a detailed analysis on the growth of the fruit highlighted two stages; a rapid growth stage from stage 6 that lasts two days and then a slowed, steady growth until maturity (Eldridge et al., 2016). The final shape of the *Arabidopsis* and *B. napus* fruits are very similar and therefore it is interesting to find a difference in growth. Perhaps it has to do with the larger size of the *B. napus* fruit.

It would have been interesting to have more information on post-fertilization growth. For example, by hand-pollinating with different amounts of pollen and study the effect on final

elongation. It is known that fully-fertilized fruits grow longer than partially-fertilized fruits. Moreover, it would be interesting to test whether fruits grow constantly and then stop or whether different growth rates are present throughout development. This work only looked at one accession, perhaps studying other accessions could already aid in answering this question.

The sector analysis lines have so far only been preliminarily analysed, mainly to confirm that the system worked. This could be taken further to do a more detailed analysis and modelling of the growth, such as has been done for fruits by Eldridge et al. (2016). Although, both *B. napus* and *Arabidopsis* fruits are cylindrical in structure, there are particular differences in style and replum morphology. The growth dynamic of these tissues would therefore be interesting to study in detail. Such analyses could answer questions regarding style emergence and growth rate and provide information on any growth restrictions involved in forming the highly reduced Brassica replum.

### *3.4.3 Vernalisation time affects pod length*

Several accessions of the DFFS were grown with both six and ten weeks vernalisation. In addition to a difference in flowering time, pod length also differed under these conditions. The effect of vernalisation on flowering time is well studied (Duncan et al., 2015, Irwin et al., 2016, Shindo et al., 2006). In Brassica, it is through the silencing of *FLOWERING LOCUS C* (*FLC*), that a plant changes from vegetative to a reproductive program (Michaels and Amasino, 1999). Cold is required for this silencing to occur and it is thought that differences in cold duration affect the transition of flowering through differences in *FLC* suppression (Shindo et al., 2006). The effect of cold on reproductive success and subsequent yield however is less well understood but has been demonstrated in wheat (Steinfort et al., 2017a, Steinfort et al., 2017b). Wheat cultivars grown with longer vernalisation times produced more grains and had a higher overall yield (Steinfort et al., 2017a, Steinfort et al., 2017b). Spring oilseed rape varieties which do not require vernalisation also yield much lower than winter oilseed rape lines, suggesting a link between cold requirement and yield (2t/ha and 3.1t/ha average in 2016, respectively) (Wynn et al., 2016). It would be interesting to trial different vernalisation times to see how plants are affected and whether an "ideal" time exists where all accessions maximise pod numbers, seeds per pod and overall yield.

### *3.4.4 Pre-fertilization differences are larger than the final pod length differences*

Pre-fertilization, the gynoecia of one accession (AmbxCom) was significantly larger than gynoecia from the rest of the accessions tested. This effect was lost later during pod development with Nin and Dar being almost as long. Nin and Dar were no different from Qui at pre-fertilization stages, but were significantly larger at the fully mature stage. This likely is due to the effect of fertilization on fruit elongation. Hand-pollinating the plants myself could have been good to ensure maximum pollination and therefore also maximum pod growth. This could have eliminated the effect of differential pollination. The seed spacing may also affect the final fruit length though and this could also be assessed in the different accessions. Measuring ovule numbers of the lines could also be interesting to have a measure of total seed potential.

Measuring ovule sizes could also be used to identify whether there is a correlation between their size and the final pod size. The positive effect of ovule size on seed size and subsequent silique size has been shown in several studies in *Arabidopsis. arf2* mutants for instance have larger ovules due to more cells in the integument layers which translates to larger seeds (Schruff et al., 2006). The cytochrome P450 KLUH (KLU) also affects seed size through the ovule size (Adamski et al., 2009). *KLU* overexpressing plants have larger ovules due to more cells in the outer integument. The seeds that develop from these ovules are also larger than wild-type seeds and the siliques are wider than wild-type ones. In agreement with these observations, the opposite phenotype is seen in loss-of-function plants (Adamski et al., 2009). Furthermore, DA1, a ubiquitin binding protein, and its network have been linked to organ size. DA1 has an effect on plant size by controlling the period of cell division (Li et al., 2008). *da1-1* ovules are enlarged and so are the subsequently formed seeds and siliques (Li et al., 2008). A more pronounced phenotype is found in *eod3-1*, which is an enhancer of *da1-1*. *EOD3* encodes the cytochrome P450/CYP78A6 and when dominantly overexpressed, the entire plant is larger than wild type, including cotyledons, seeds and siliques (Fang et al., 2012). UBIQUITIN-SPECIFIC PROTEASE (UBP15) is closely associated with DA1 as the latter stabilises UBP15 (Du et al., 2014). UBP15 regulates ovule and seed size through the integument length which is decreased in *ubp15* mutants (Du et al., 2014). With this in mind, it would be interesting to measure ovule size in the different lines shown here to see whether they were the cause of the enlarged final seed size.

### *3.4.5 Pods have different lengths due to varied cell numbers*

Overall, I show that cell size does not vary between the lines and that final pod length is instead a result of varied cell numbers. Larger cells often arise due to endoreduplication and therefore have

higher level of ploidy. The flow cytometry data which does not show strong differences between the accessions supports the data on cell size. Perhaps using accessions with more varied pod lengths would show different results but from the results presented here, it seems cell division is the primary factor in pod elongation.

### *3.4.6 Different plant material can be used for different purposes*

Here, a diversity fixed foundation set (DFFS) made up of very diverse plants and a doubled haploid (DH) population were used. The latter is often used in breeding for the development of new lines and the introgression of new variation. Crosses between a high-yielding and a non-commercial parent are thus often used to give rise to plants with increased variation, novel traits and desirable characteristics. These can then be used for breeding and a new variety can be made fairly rapidly. The use of a more diverse set of germplasm, such as the DFFS, allows the analysis of more variation. Several methods exist to exploit the variation (more on this in Chapter 4). Due to the fact that the DFFS does not include current breeding material, it is perhaps better suited to research purposes where a gene of interest underlying a trait can be studied, validated and then the knowledge passed on to breeders, possibly even with markers for a specific trait.

### *3.4.7 Pod-shatter resistance as a trait with much potential for improvement*

Varieties on the market have been bred for high yield. As a consequence, they do not differ drastically in most traits, such as those related to yield used in this study. However, one trait where large differences are still present is the pod strength, which is important to avoid premature pod opening and hence seed loss in the field, a process known as pod shatter. Pod shatter is a source of significant yield loss for farmers of OSR with an average of 11-25% of the harvest (Price et al., 1996); however, losses higher than 50% can be experienced under adverse weather conditions (MacLeod, 1981). With the changing climate and unpredictable weather patterns, it is becoming increasingly important for breeding companies to address this problem. Pod strength can be improved using several different avenues. Two companies; Dekalb and Limagrain, have developed markers for the trait and achieved this through the hybrid Ogura CMS breeding system first described by Ogura (1968). In this system, a cytoplasmic male sterile line (CMS) is crossed with a line carrying restorer (*Rf*) genes which will then produce functional pollen. This system is used in the production of hybrid OSR.

Another avenue that can be used to reduce pod shattering is the translation of knowledge from *Arabidopsis* to *B. napus*. Genes required for pod opening and the correct development of the dehiscence zone are well-characterised in *Arabidopsis* and include the transcription factors *INDEHISCENT (IND)* and *SHATTERPROOF1/2 (SHP1/2).* In *shp1/2* the dehiscence zone does not form correctly and fruits are indehiscent (Liljegren et al., 2000). The same effect is present in *ind* mutant fruits where valve margins fail to form (Liljegren et al., 2004). The same mechanisms have been found in *L. campestre* and mutants in these two genes are also indehiscent (Lenser and Theissen, 2013).

That the gene functions are conserved in the *Brassicaceae* was first demonstrated by Girin et al. (2010). Here, it was shown that *IND* is present in *B. oleracea, B. rapa* and in *B. napus*, expressed in the same manner and has the same function (Girin et al., 2010). Furthermore, mutants with significantly lower *IND* expression levels in *B. oleracea* and *B. rapa* displayed a lack of valve margin tissue (Girin et al., 2010). This effect was dose-dependent so that plants with slightly reduced levels still formed all tissues but had a smaller valve margin, leading to stronger fruits which are still able to open. This suggests that this method can be used for crop improvement where fruits should be stronger but still able to shatter.

While most mutations explained above render pods so hard that they do not shatter, for agriculture, pods need to be strong enough to withstand wind and rain but weak enough to open in the combine when harvested (Bruce et al., 2001). This is a tight balance. Using the knowledge from *Arabidopsis* outlined above, Bayer CropScience has recently developed varieties with a mutation in the *IND* gene which are harder to open (CropScience, 2014). This was done using chemical mutagenesis and backcrossing to get the correct amount of shatter resistance. The first variety with this mutation was called InVigor$^{TM}$ L140P and sold in Canada since the 2015 growing season. More varieties have been developed since. These avenues show that improvement in pod strength is possible and can be tackled from different sides; both through breeding using diverse material, but perhaps more efficiently through a fundamental understanding of the genes that control the developmental process involved in dehiscence zone formation.

### 3.5 Concluding Remarks

Pod growth in *B. napus* is controlled by several factors. Some general correlations seem to exist and have been shown here to be present in several distinct populations. The effect of the

environment on fruit growth is substantial however. Understanding how pods and seeds grow has to be the first step to then be able to improve these traits and produce crop plants with higher yield. Pod strength is a good place to improve this as many varieties do not have strong pods and premature pod opening can lead to large yield losses.

# Chapter 4

## 4 The correlation between stomata density and pod length in the *Brassicaceae*

## 4.1 Introduction

All living beings need energy to grow. Plants get theirs through photosynthesis which requires $CO_2$ and releases $O_2$. The gas exchange occurs on the surface of the plant through pores called stomata. The process occurs in the green organs of plants. Photosynthesis therefore is mainly carried out in leaves and the assimilates are then redistributed throughout the plant. In *B. napus*, the developing pods are also green and also are relevant for growth. The leaves of the plants senesce and abscise before the seeds in the fruits are fully mature. Furthermore, the dense canopy of *B. napus* prevents light from reaching the leaves. Hence, photosynthesis carried out in the pods themselves is of relevance to seed filling and therefore yield in *B. napus*.

Stomata grow through a specialised program which has been thoroughly studied in *Arabidopsis* (Fig. 4.1). Meristemoid mother cells (MMCs) are protodermal cells that divide to create a small triangular cell called a meristemoid (Bergmann and Sack, 2007). This then divides again and creates the non-stomatal cells surrounding stomata which eventually turn into guard mother cells (GMCs). A final symmetric division creates terminally differentiated guard cells (Bergmann and Sack, 2007). Stomata patterning is under tight control and stomata are usually separated by at least one epidermal cell from any other stoma (Geisler et al., 2000). Mistakes in this closely controlled pathway can lead to changes in spacing of stomata, often leading to altered densities or stomata clustering.

The transcription factors necessary for the different transitions in stomata development have been characterized. *SPEACHLESS (SPCH),* is required for the change from a MMC to a meristemoid (MacAlister et al., 2007). The next step, whereby the meristemoid develops into a GMC is guided by *MUTE* (Pillitteri et al., 2007). Finally, *FAMA* is required for the division giving rise to the guard cells and thus a fully formed stomata (Ohashi-Ito and Bergmann, 2006). Loss of any of these genes leads to aborted or undeveloped fruits.

STOMATAL DENSITY AND DISTRIBUTION1 (SDD1) is required for the correct spatial distribution of stomata. *SDD1* is mainly expressed in GMCs and only expressed very lowly during the other stages (von Groll et al., 2002). *sdd1* mutants have a higher stomata density with knock-out mutants showing many more stomata. Over-expressing lines display two- to three-fold fewer pores (Berger and Altmann, 2000, von Groll et al., 2002). In the *sdd-1* mutant most stomata are still at least one

cell apart but in some organs, such as in cotyledons, up to 45% are present in clusters (Berger and Altmann, 2000, von Groll et al., 2002).

EPIDERMAL PATTERNING FACTOR1 (EPF1) controls the asymmetric cell divisions required for proper stomata development (Hara et al., 2007). The activity of EPF1 is dependent on *TOO MANY MOUTHS* (*TMM*) and over expressing the gene leads to plants producing no stomata at all (Hara et al., 2007). On the other hand, mutants in the *EPF1* gene have a higher stomata density and also display clustered stomata (Hara et al., 2007). *EPIDERMAL PATTERNING FACTOR2 (EPF2)* is expressed earlier than *EPF1* and controls the entry of cells into the stomatal lineage (Hunt and Gray, 2009). Mutants have more stomata but do not display clustering. The *epf1epf2* double mutant exacerbates the phenotypes of both single mutants and gives rise to plants with more stomata and more clusters (Hunt and Gray, 2009).

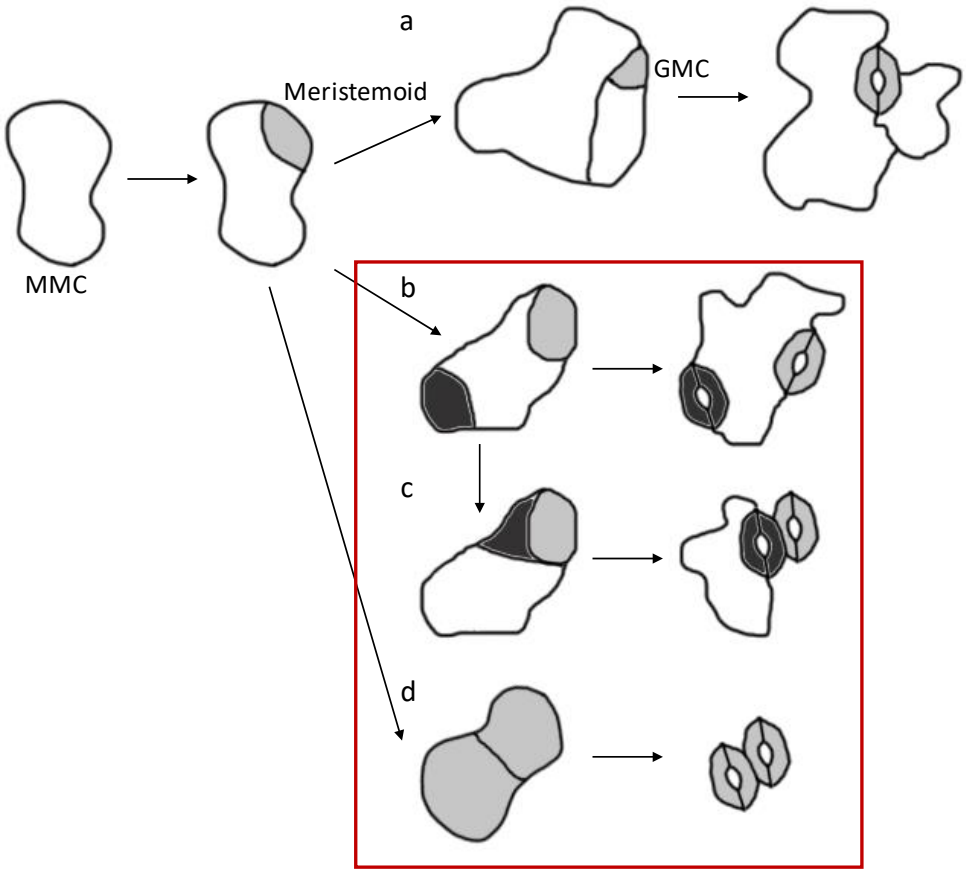



Fig. 4.1. Stomata development is tightly regulated. A MMC divides to create a meristemoid which subsequently creates the GMC. The latter divides to create a fully developed stoma. (a) Normally, stomata are at least one epidermal cell apart from one another. However, mistakes in the divisions can lead to higher density stomata which no longer follow the one-cell rule (b). Clustered stomata can also appear (c). In very severe cases, all cells assume stomata identity (d).

*FOUR LIPS (FLP)* is also required for correct stomata development and densities. In *flp* mutants stomata sometimes appear clustered with 1-3% more clusters in siliques than in wild-type fruits (Geisler et al., 1998, Lai et al., 2005). This is due to extra divisions of GMCs giving rise to two stomata closely next to each other.

*TOO MANY MOUTHS* (*TMM*) encodes a receptor-like protein and is required for the control of division planes. In the *tmm-1* mutant the formation of stomata directly next to already formed ones is possible. This creates plants with many clustered stomata (Nadeau and Sack, 2002). Phenotypes differ depending on the organ and in *tmm-1* leaves, up to 55% of stomata are present in clusters (Vrablova et al., 2017) whereas siliques and stems do not have stomata in *tmm-1*. However, different phenotypes have been reported (Geisler et al., 1998, Yan et al., 2014b).

Perhaps the largest effect on stomata development is caused by the dominant overexpression of *INDUCER OF CBF EXPRESSION1* (*ICE1*), also termed *ScreamD* (*scrmD*), which encodes a bHLH-leucine zipper protein. *scrmD* plants show severe growth defects and reduced fertility due to the fact that they are made of mainly stomata, leading to wrinkly leaves (Kanaoka et al., 2008). The *ICE1* gene acts in a dosage dependent manner and therefore intermediate phenotypes are possible.

Many more genes involved in the correct development and spacing of stomata have been described but the work in this chapter will address the genes introduced above only.

### *4.1.1 Aims and hypothesis*

In Chapter 3 I carried out experiments to establish how pods grow and how different traits interact. From SEM pictures of large and small pods, I made the observation that stomata densities varried between large and small pods. This Chapter explores that relationship and tests the hypothesis that larger pods have lower stomata densities.

## 4.2 Material and Methods

### *4.2.1 Plant material and sampling*

The oilseed rape accessions used (Quinta, NingYou, Darmor and ColumbusxNickel) were grown in a polytunnel and sampled and measured by Rachel Wells and colleagues. SEM images were taken by myself.

*Brassica oleracea* was supplied by Dana McGregor. DH1012 rapid-cycling, self-compatible plants were transformed by Bract with 35S driving wildtype *Arabidopsis* ICE1 cDNA (35S:ICE1), 35S driving *Arabidopsis* cDNA with the scrmD mutation (35S:scrmD) (as in Kanaoka et al. (2008)) and with 35S driving *Arabidopsis* ICE1 cDNA with Serine 403 mutated to Alanine which leads to an enhanced protein stability and transactivational activity (35S::ICE1-S403A) (Bohuon et al., 1996, Miura et al., 2011). A tissue culture control (TCC) line was used for comparison to transformed plants. The plants were grown in a greenhouse. I sampled the plants and measured the pod lengths using the macro in ImageJ that I developed (same as in Chapter 3).

The *Arabidopsis* mutants were all obtained from Dana McGregor. They included Col-0, *epf1epf2, flp-1, tmm-1, scrmD, ice1,* Landsberg (Ler), *sdd-1*, Wassilwskija (Ws) and *atrlp17-1* (another allele of *TMM*). They were grown in a controlled environment chamber at 22°C. I sampled fully developed pods and measured them using ImageJ. I also took the SEM images.

### *4.2.2 Rosette size measurements*

*Arabidopsis* plants were photographed from above and rosette sizes measured. The pictures were loaded into ImageJ and I wrote a macro to measure the green area (in Appendix).

### *4.2.3 Leaf and pod prints using nail varnish*

Leaf and pod prints were made for the *B. oleracea* plants. Samples were coated in nail varnish and left to dry. Then the varnish was peeled off using sticky tape and the tape stuck to a microscope slide. This was then imaged using a light microscope.

## 4.3 Results

### *4.3.1* B. napus *accessions have different pod lengths and stomata indices*

Four accessions from the Diversity Fixed Foundation Set (DFFS) were chosen based on the length of their pods. Quinta (Qui) and NingYou (Nin) had small pods while Darmor (Dar) and ColumbusxNickel (ColxNic) had large pods (Table 4.1). Mature fruits were sampled and cells observed using SEM images. The total number of cells and stomata was determined on an area of 0.33mm$^2$ on four images per accession. Qui and Nin had more stomata per surface area (Table 4.1). The stomata index (si) was also calculated as described by Salisbury (1927). The si = number of stomata/(number of epidermal cells + number of stomata) x 100. Qui and Nin had a stomatal index of 20.18% and 20.47%, respectively whereas Dar and ColxNic had a lower si roughly round 18%. Although these numbers were not significantly different from each other, the trend is intriguing and I therefore hypothesised that larger pods have fewer stomata. As both the stomata numbers in a specific area and the stomata index showed a difference and they were found to correlate in *Arabidopsis* (here and in Delgado et al. (2011)), in further experiments I only counted the stomata in a specific area.

Table 4.1. Stomata on *B. napus* pods of different lengths. The accessions differed in stomata per area and in the stomata index with the larger pods having fewer stomata.

| Accession | Pod Length (cm) | Stomata per 0.33mm$^2$ | Cells per 0.33m$^2$ | Stomata Index (%) |
|---|---|---|---|---|
| Qui | 4.19 | 15.25 | 61.00 | 20.18 |
| Nin | 5.64 | 17.50 | 69.00 | 20.47 |
| Dar | 7.50 | 11.25 | 49.75 | 17.91 |
| ColxNic | 7.56 | 12.00 | 55.25 | 17.89 |

### *4.3.2 B. oleracea ICE1 overexpressing plants have more stomata and shorter pods*

The results obtained in *B. napus* suggest that there may be a connection between stomata density and final pod size. In order to further test this hypothesis, I decided to take advantage of resources developed in the close relative *B. oleracea*. Dana McGregor had previously developed 35S::ICE1 *B. oleracea* lines (overexpressing the Arabidopsis *ICE1* cDNA), 35S::ICE1-S403A (overexpressing a stabilised version of the ICE1 protein) and 35S::scrmD (overexpressing the *Arabidopsis* cDNA harbouring the *scrmD* dominant mutation known to increase stomata number). In summary, all lines overexpressed forms of *ICE1*. Stomata densities were determined on the leaves and the pods to determine whether the plants did indeed produce more stomata. Leaf and pod stomata were only significantly increased in 35S:scrmD compared to the tissue culture

control (TCC) (Fig. 4.2). The other two lines also had more stomata but not significantly so. In a given area, many more stomata were present on the leaves than on the pods.

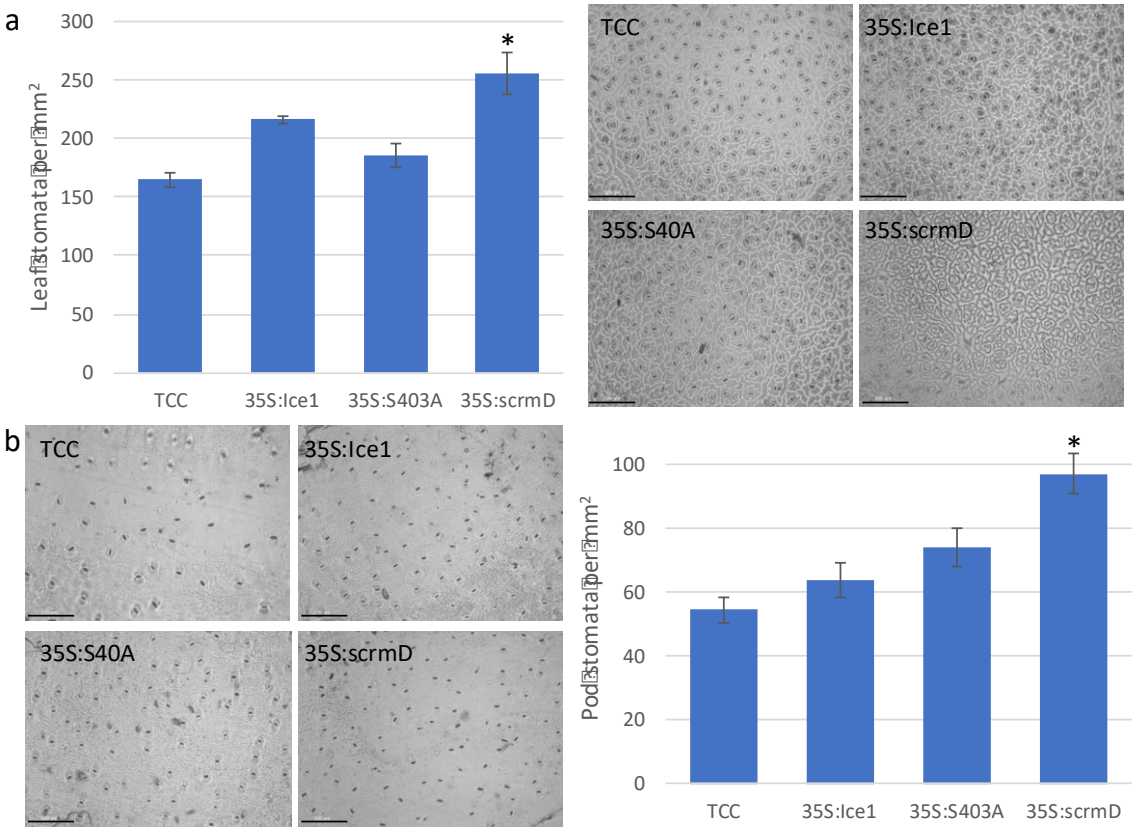



Fig. 4.2. Stomata on *B. oleracea* leaves and pods. 35S:scrmD was the only transgenic to display significantly more stomata on the leaves (a) and on the pods (b) ($p<0.05$). The tissue culture control (TCC) always had the fewest stomata. Stomata were counted on prints made using nail varnish. Bars are standard errors.

Once confirmed that more stomata were present on tissue from the transformed plants, pod lengths were measured on the progeny of the tissue culture transformed plants (T2). Four plants of the TCC control were used to measure 171 pods. All other lines had at least nine plants and 286 pods measured. If the hypothesis that more stomata gives rise to shorter pods is true, pods should be shorter in the transgenics. This was indeed the case with all three lines having significantly shorter pods (Fig. 4.3). 35S:scrmD, the line with the most stomata on the leaves and on the pods showed the largest growth reduction with a reduction of 12%. These results were mirrored in the seeds per pod where the transformed plants had one seed less on average with no significant changes in seed size or thousand seed weight (Table 4.2).

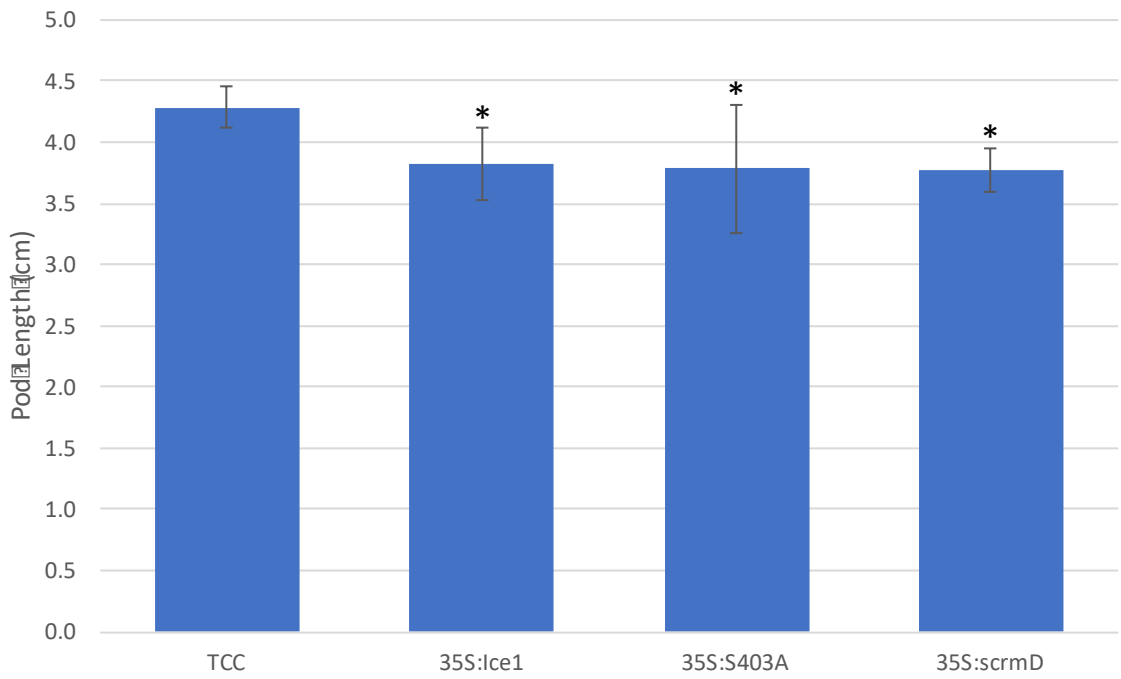



Fig. 4.3. Pod lengths in *B. oleracea.* Transgenic plants overexpressing *ICE1* all had significantly shorter pods ($p<0.05$). Error bars are standard errors.

Table 4.2. Transformed plants had on average one less seed per pod. Thousand seed weight (TSW) and seed surface area were not affected much. None of the parameters differed significantly to the TCC. Analysis was carried out using a Marvin seed analyser and statistics in GenStat using a general linear regression.

| Line | Seeds per Pod | TSW (g) | Seed surface area (mm) |
|---|---|---|---|
| TCC | 6.633 | 2.860 | 3.339 |
| 35S:Ice1 | 5.082 | 2.961 | 3.328 |
| 35S:S403A | 5.930 | 2.561 | 3.117 |
| 35S:scrmD | 5.944 | 2.554 | 3.047 |

### *4.3.3 Arabidopsis stomata density mutants display shorter siliques*

Next, to test the hypothesis further and to look at more mutants, *Arabidopsis* was used. The mutants were supplied by Dana McGregor. I analysed Col-0 as a control and *epf1epf2, flp-1, scrmD, tmm-1, ice1* in the Col-0 background. I also analysed *sdd-1* in the Ler background and compared it to Ler. Lastly, Ws and *atrlp17-1* were grown as the mutation was in the Ws ecotype.

Clear differences in growth were visible and to quantify this, the rosette area was measured. To this end, plants of each line were photographed from above and the green area was determined using ImageJ. All mutants were smaller than the associated wild type but *epf1epf2, scrmD, ice1* and *sdd-1* were significantly smaller ($p<0.001$) (Fig. 4.4).

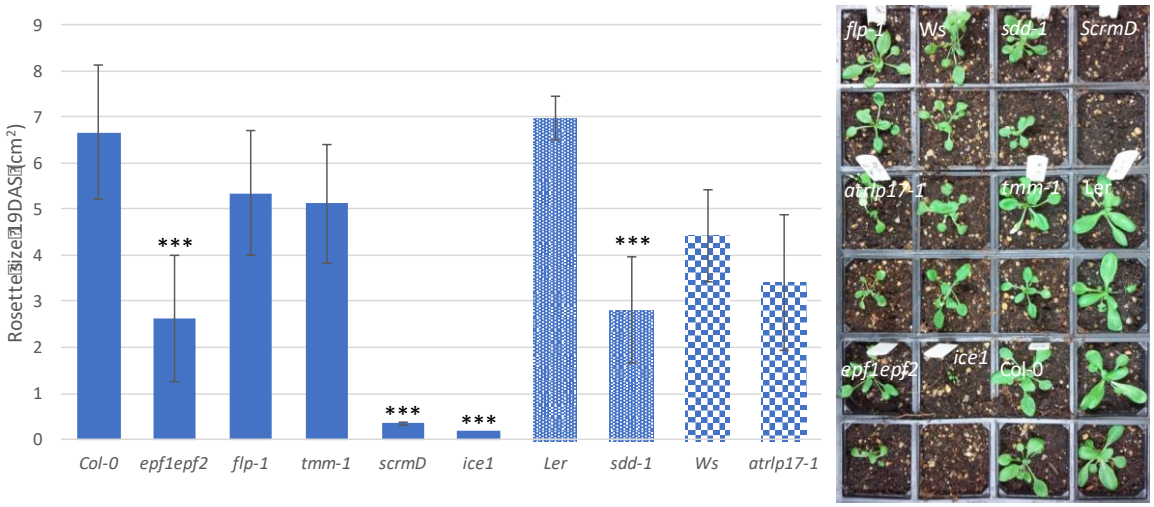



Fig. 4.4. Stomata mutants had smaller rosette sizes (19 DAS). *epf1epf2, scrmD, ice1* and *sdd-1* were significantly smaller ($p<0.001$). Bar fill patterns show which ecotype the mutation was in. Error bars are standard deviations.

Next, siliques from six different plants were measured (at least 62 siliques per line). The silique lengths were more similar than the rosette sizes. The only mutants with silique lengths significantly different from the associated wild-type control were *epf1epf2, sdd-1* and *atrlp17-1* (Fig. 4.5). *scrmD* and *ice1* also had very small siliques but they were not included in the graph as they only produced very few pods. It is interesting that while the rosette sizes differed substantially in size, the silique sizes were not that different.

Cellular arrangements of leaves and siliques were observed using SEM. Most mutants had the published phenotype in the leaves and the same pattern in the silique (Fig. 4.6). *tmm-1* has been described as having siliques without stomata whereas *tmm-6* has some (Geisler et al., 1998, Yan et al., 2014b). Here, *tmm-1* stomata formed in the *tmm-1* silique (Fig. 4.6f). However, *atrlp17-1,* a different *tmm* allele, had hardly any stomata in the siliques (Fig. 4.6r). No silique phenotypes have been published except for *flp-1* and *tmm-1* where siliques behave slightly differently than the leaves (Geisler et al., 1998).

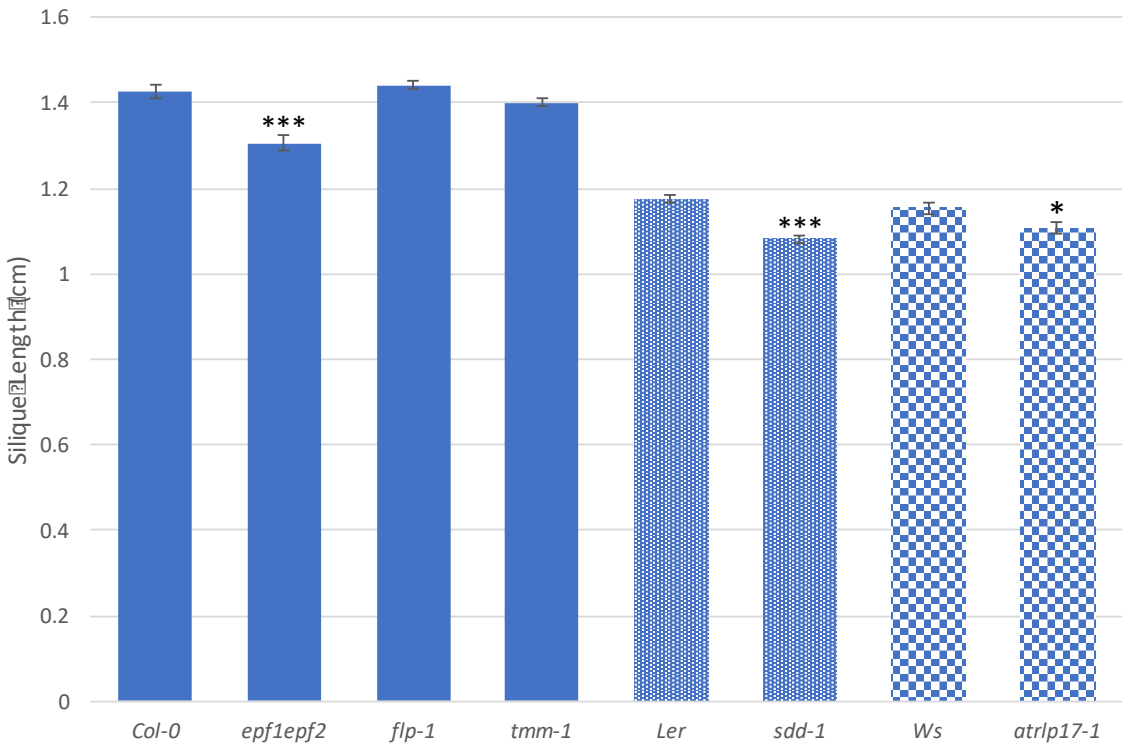



Fig. 4.5. Silique lengths of stomata mutants. *epf1epf2, sdd-1* and *atrlp17-1* had significantly smaller siliques than the wild type. The siliques of *scrmD* and *ice1* were also smaller than the Col-0 wild type. However, too few siliques were produced by the plants to include them. Fill patterns show which ecotype the mutation is in and mutant lengths were compared to their respective wildtype. *p<0.05, ***p<0.001. Standard errors are shown.

To determine which factor could affect the silique length, stomata per $mm^2$ and the percentage of stomata that were found in clusters were counted by Dana McGregor and myself. There were three lines with significantly more stomata per $mm^2$; *scrmD, ice1* and *sdd-1*. *atrlp17-1* was also significantly different with fewer stomata than its wild type Ws (Fig 4.7a). Stomata clustering was less than 10% in the wild types and ranged up to almost 80% in *scrmD* (Fig. 4.7b). The three mutants *tmm-1, scrmD* and *ice1* had significantly more stomata clusters than Col-0. It is interesting to note that while *tmm-1* had many more clustered stomata it did not have more stomata overall. *scrmD* and *ice1* were the only two mutants that had more stomata and more clusters. Those two mutants were the most different from the wild type and growth was also impaired most. The fact that *sdd-1* had more stomata, shorter siliques but not more clusters suggests that it is the stomata number that affects the silique length and not the clustering. This is supported by the *tmm-1* mutant which displayed many more stomata in clusters, the same number of stomata and no difference in silique length. Therefore, the number of stomata per area had a larger effect on silique length than the stomata clustering.

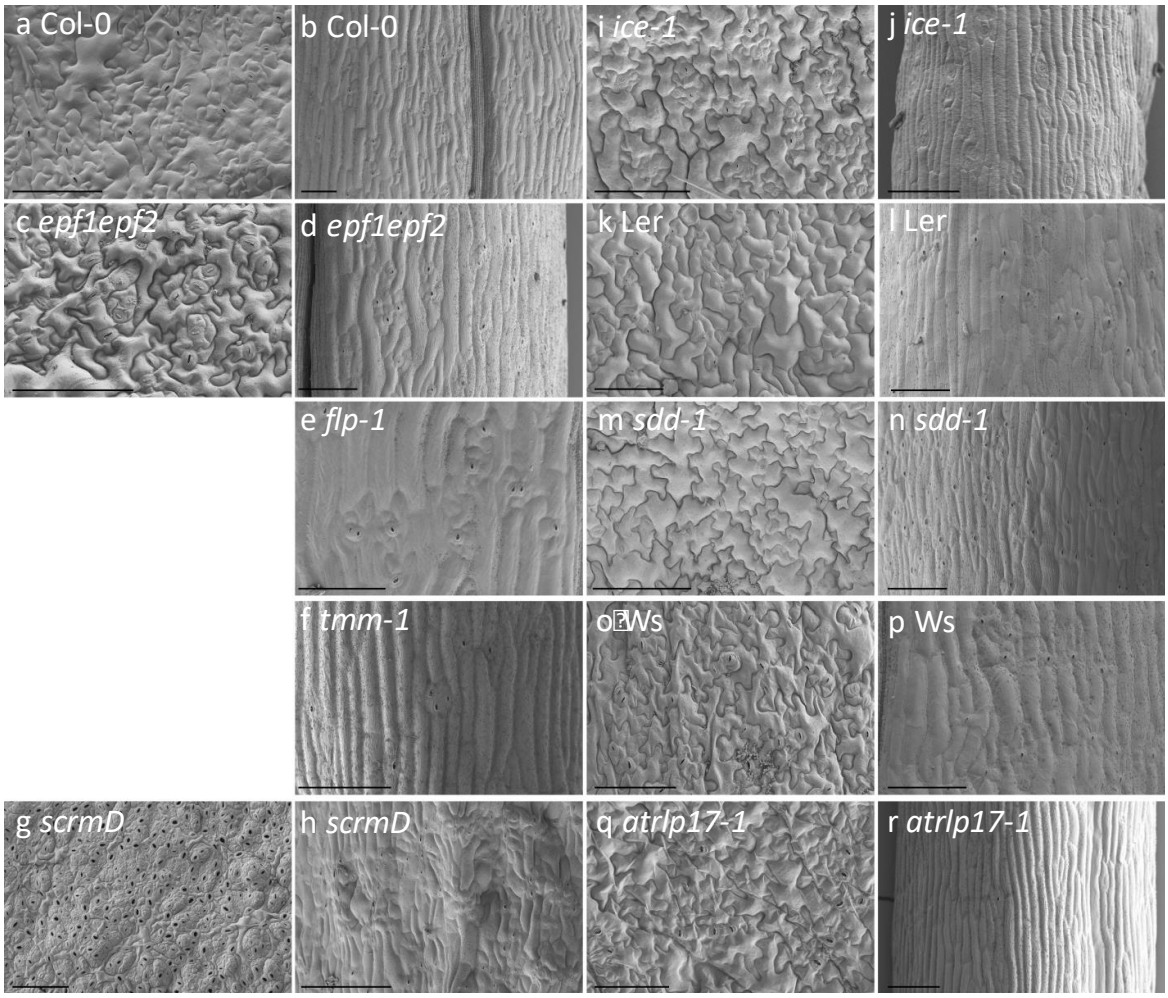



Fig. 4.6. Stomata arrangements on leaves and siliques. Similar patterns in the siliques (b,d,e,f,h,j,l,n,p,r) as in the leaves (a,c,g,l,k,m,o,q) were found. a, b Col-0; c, d *epf1efp2;* e *flp-1*; f *tmm-1;* g,h *scrmD*; i,j *ice-1*; k,l Ler; m,n *sdd-1*; o,p Ws; q,r *atrlp17-1*. No leaf images were taken of *flp-1* and *tmm-1.* Scale bars are 100µm

Next, to test the basis for this effect, stage 12 gynoecia were analysed using SEM images. At that stage, meristemoids are visible but have not developed into stomata yet. In wild-type gynoecia cells are aligned in clearly definable columns growing roughly in straight lines (Fig. 4.8). In the *epf1epf2* and *sdd-1* mutants, this arrangement was altered and cell files were not clearly distinguishable (Fig. 4.8d, n). However, the overall appearance of the gynoecium was not altered. The effect also seemed to largely disappear later on with stage-17 fruits looking more like the respective wild types (Col-0 and Ler). The most striking mutant was s*crmD* which produced leaves made entirely of stomata (Fig. 4.6g). In the siliques however some non-stomata cells were present (Fig. 4.6h). The gynoecium looked different from the one of Col-0 in that it was wider (Fig 4.8i). The cellular arrangement was vastly different with much smaller, rounder cells growing with no obvious direction of growth (Fig 4.8j). This was not lost later on as the fully mature s*crmD* fruit still

had a different cellular patterning with many small cells growing in many different directions (compare Fig. 4.8b and 4.8j). The overall fruits were also dramatically shorter. Most mutants however did not show a noticeable difference in the gross cellular patterning.

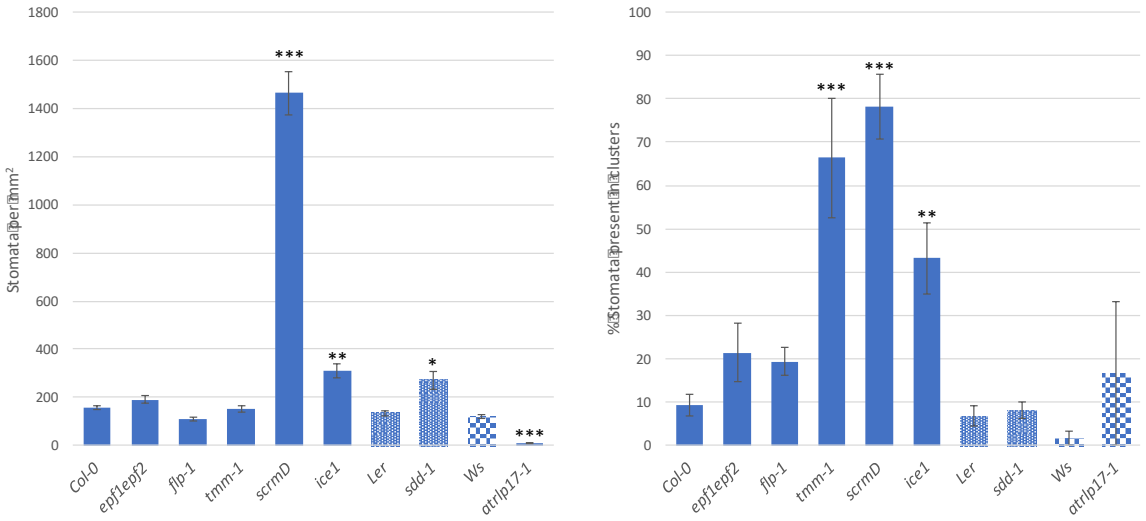



Fig. 4.7. Stomata arrangements on siliques. (a) *scrmD, ice1* and *sdd-1* had significanlty more stomata per mm$^2$ than their respective wild types. The mutants a*trlp17-1* had significantly fewer stomata per mm$^2$ with only 12 stomata found in total over the entire area looked at (34.4mm$^2$). (b) Significantly more stomata clusters were present in *tmm-1, scrmD* and *ice1* than in their Col-0 wild type. Columns are filled according to their background – all filled columns are plants in the Col-0 background, *sdd-1* is in the Ler ecotype and *atrlp17-1* was in the Ws ecotype. Stars denote significant differences to the respective wild type as determined using GenStat. *p<0.05, **p<0.01, ***p<0.001

## 4.4 Discussion

This chapter's aim was to investigate the relationship between stomata numbers and pod length. The hypothesis to test was that plants with higher stomata density had shorter fruits. The hypothesis was based on observations in *B. napus* where this relationship was present after analysing four different accessions. Other members of the *Brassicaceae* were then also used.

### *4.4.1 Testing the hypothesis in B. oleracea*

Whether plants with more stomata on the leaves also had more stomata on the fruits was so far untested in *B. oleracea* and rarely discussed in *Arabidopsis* (apart from in *flp-1* and *tmm*). *B. oleracea ICE1* overexpressing plants displayed more stomata on the leaves and on the pods. Many

more stomata were present on the leaves than on the same area on the pods, which could be due to the fact that there were more cells in total present on the leaves in a given area than on the pods. A potential change in stomata density could be assessed in the future but would require higher resolution images so that the non-stomata cells can be counted accurately. The pods of the transformed plants were also shorter than in wild type. The data based on lines in *B. oleracea* therefore support the hypothesis of a negative correlation between stomata density and pod length.

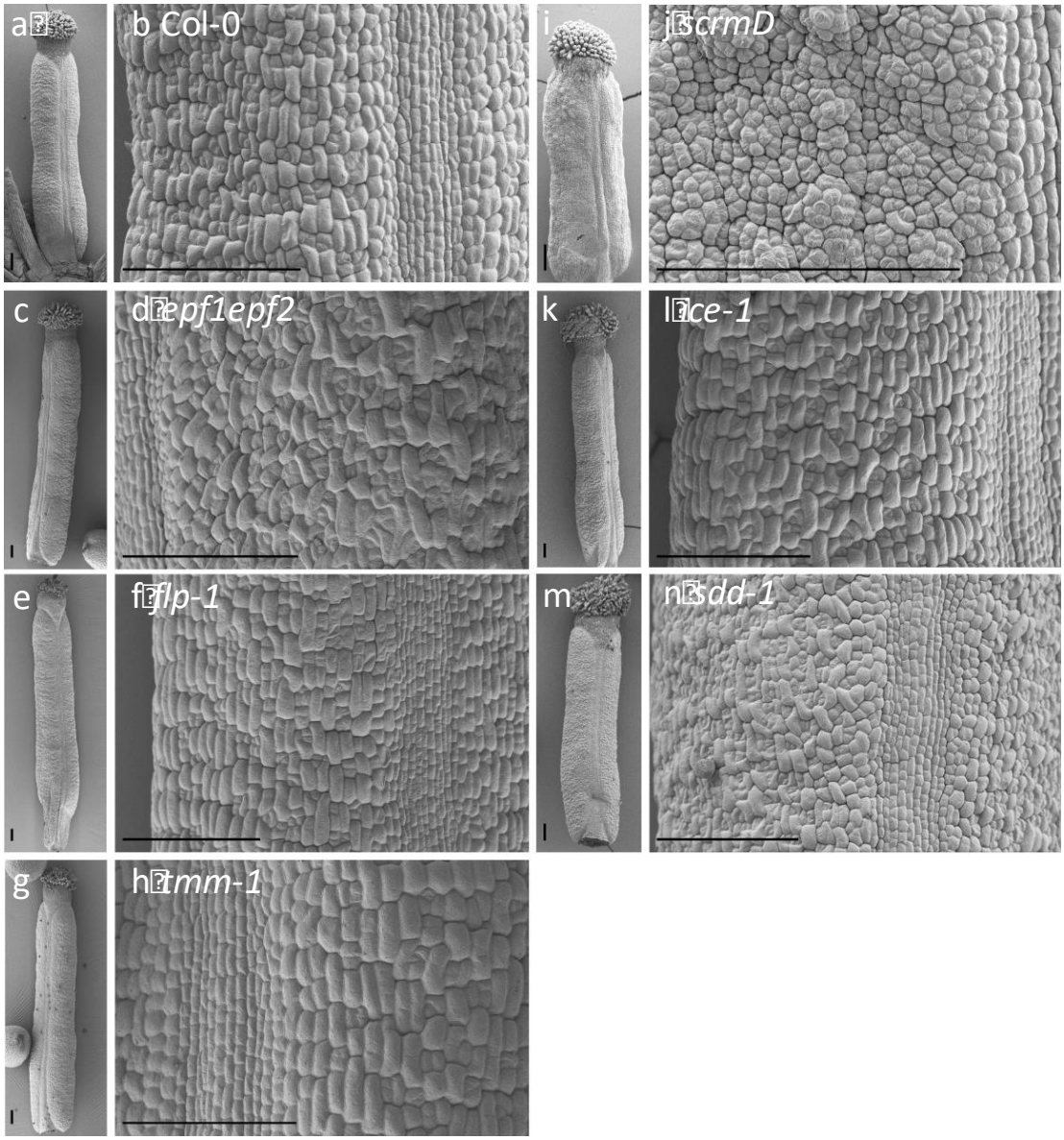



Fig. 4.8. Stage 12 gynoecia mostly do not differ much from the wild-type ones. However, when looking at the cellular detail *epf1epf2* (c,d) and *sdd-1* (m,n) had a less structured cellular pattern than Col-0 (a,b). The *scrmD* fruit was wider and the cellular arrangement was different with much rounder cells (i,j). a, b Col-0; c, d *epf1efp2;* e,f *flp-1*; g,h *tmm-1;* i,j *scrmD*; k,l *ice-1*; m,n *sdd-1.* No images at this stage were taken for Ler, Ws or *atrlp17-1.* Scale bars are 100µm

### *4.4.2 Differences between Arabidopsis and B. oleracea scrmD lines*

The *scrmD* mutant in *Arabidopsis* is known to have leaves made up entirely of stomata (Kanaoka et al., 2008). Interestingly, the equivalent *B. oleracea* mutants did not show this trend. The leaves of all three transformed lines still had epidermal cells present. This could be due to the fact that *ICE1* works in a dose-dependent manner and therefore phenotypes can be different depending on the expression level (Kanaoka et al., 2008). The difference could also have arisen because the *Arabidopsis* gene was used in the constructs and not the native *B. oleracea* copy. Perhaps transformations with the native version would lead to a different phenotype.

### *4.4.3 Total stomata numbers are more relevant than clusters*

Next, *Arabidopsis* was used to study a wider array of stomata mutants. Several different phenotypes were found with some mutants displaying a higher number of total stomata, some had a similar number of stomata but organised in clusters and some having a combination of both.

The *sdd-1, ice1* and *scrmD* mutants were the only ones with more stomata than wild type. They all displayed significantly shorter fruits, supporting the hypothesis that more stomata lead to shorter fruits. The *tmm-1* mutant had more clustered stomata coupled with same total number of stomata as wild type. No difference in silique length was present, suggesting that total stomata number was a more important factor than stomata clustering. This effect is likely to be due to plants with clustered stomata behaving differently than plants with normally-spaced ones; stomata conductance is negatively affected when clusters are present. Overall, one cluster of stomata has been proposed to function like one stoma by itself (Dow et al., 2014). Clusters also cannot open and close as efficiently as non-clustered stomata and therefore they are slower to adapt to changes in the environment and have a lower rate of photosynthesis (Dow et al., 2014, Vrablova et al., 2017).

### *4.4.4 Mutants of TMM show different phenotypes*

The *tmm-1* results were not the same as those published which suggest that siliques are 50% smaller than wild type (Geisler et al., 1998). The mutant is also described as not having many stomata on the siliques and a gradient being present with more stomata in the middle than at the

top and bottom ends of the fruit (Geisler et al., 1998). Another feature described was more clustered stomata in the silique which the results here support (Geisler et al., 1998). The *atrlp17-1* mutant, which is another *TMM* allele, displayed fewer stomata on the siliques and a 4% growth reduction. This allele therefore behaved more closely to the *tmm-1* phenotype described by Geisler et al. (1998). Another study, which worked on different *tmm* mutants found that *tmm-6* had similar stomata numbers in the silique as wild type, showing that different phenotypes for *tmm* mutants are possible (Yan et al., 2014b).

### *4.4.5 Early plant growth is affected in stomata mutants*

Not just fruit length was affected in the stomata mutants analysed here. The mutants that displayed shorter fruits also had smaller rosette sizes 19 days after sowing. As with the silique lengths, *flp1* and *tmm-1* were least affected. Final plant sizes were not recorded but these early results show that overall plant growth was negatively affected. It could be that the mutants grow slower but reach the same final size. This is supported by the fact that the *sdd-1* mutant does not display differences in leaf area, general shape or overall morphology when mature (Berger and Altmann, 2000). Furthermore, *tmm-1* primordia have similar cell numbers and cell division rates as wild-type plants (de Marcos et al., 2016). The mutant also takes longer to develop stomata; while the wild-type primordia begin making stomata by day 1 after initiation, *tmm-1* does not make fully developed stomata until three days after initiation, which shows that differences in growth are present (de Marcos et al., 2016). *epf1* and *flp* mutants have the same leaf fresh weight as Col-0 (Lawson et al., 2014). This study however also noted smaller fresh weights in *sdd-1* and *tmm*, so it cannot be ruled out that plants are affected in growth parameters.

### *4.4.6 Gynoecium development at early stages is different*

Furthermore, looking at the early stage gynoecia showed some developmental effects. While most gynoecia were not different to wild type, the ones with the most stomata were. The fruits of *epf1epf2, sdd-1* and s*crmD* seemed to have different division planes and a more uneven growth. This seemed to be largely rescued at maturity (stage 17) in the *epf1epf2* and *sdd-1* whereas fruits from *scrmD* still looked very different. The reason for the differences could be that these genes affect stomata density, an effect that would arise early whereas clusters would only appear later on. This could also explain the lack of differences in the mutants with more clustered stomata.

Measuring the sizes of stage-12 gynoecia could be useful to determine if the changes affected growth early on. This could also aid to answer the question of whether the mutants were smaller in all areas of development or whether the phenotypes are mostly present in the fully developed fruits.

### *4.4.7 Final fruit length in Arabidopsis*

While the results here suggest that more stomata lead to shorter fruits, they do not explain why such an effect should be present. It could be that the fruits contain the same number of cells but stomata are smaller than epidermal cells and do not elongate like the others and would therefore restrict the growth of the fruit. If all fruits are the same size pre-fertilization but are later different, this could be the case. This hypothesis is supported by Delgado et al. (2011), who found that in cotyledons and leaves of different *Arabidopsis* accessions, due to larger pavement cells, fewer stomata are present and organs are larger overall (Delgado et al., 2011). That study also found the same trend as reported here for the fruit with a negative relationship between stomata index and size of cotyledons in *Arabidopsis* (Delgado et al., 2011).

### *4.4.8 Water use efficiency may be affected*

An alternative hypothesis is that the difference in stomata patterning affects carbon assimilation and water use efficiency (WUE). These are difficult theories to test as plants react differently in different environments and therefore results are sometimes contradictory.

Using *epf1epf2* and *EPF2* overexpressing plants, which have many more and fewer stomata, respectively, it has been demonstrated that WUE correlates negatively with stomata density (Franks et al., 2015, Hepworth et al., 2015) suggesting that having fewer stomata is an advantage in terms of WUE.

### *4.4.9 Carbon assimilation rates may be different*

Whether stomata density affects $CO_2$ assimilation rates is debatable. Lawson et al. (2014) showed that plants with low stomata densities have higher $CO_2$ assimilation rates compared to those with a higher density. On the other hand, a study focussing on mutants with different levels of clustering found that clustering increases carbon assimilation rates, but is also coupled to higher

transpiration rates (Dow et al., 2014). However other studies found no differences in $CO_2$ assimilation rates between plants with different stomata densities and this effect was coupled to specific light intensities (Franks et al., 2015, Schluter et al., 2003, Vrablova et al., 2017). This was because the ratio of stomatal conductance to mesophyll conductance is kept stable, so, while when more stomata are present more $CO_2$ can access the leaf, the photosynthesis rate remains constant (Vrablova et al., 2017). These results cannot explain why stomata densities affect fruit size but they do not rule out that $CO_2$ assimilation rate and/or water efficiency could play a role.

### *4.4.10 Stomata density and crop improvement*

My results and those of others suggest that the stomata density could be a trait targeted for crop improvement. This has indeed been done in barley where overexpressing the orthologue of *EPF1* gives rise to plants with fewer stomata (Hughes et al., 2017). The plants exhibit increased WUE, but no changes in photosynthetic efficiency, nor plant growth (Hughes et al., 2017). Instead, plants have a slightly higher seed number and overall yield under normal conditions. Under drought conditions, these small increases are exacerbated, highlighting the advantage of the fewer stomata with respect to WUE. In the previous chapter, I showed that seed number correlates positively with pod length. Although seed number was not assessed in this chapter, it is possible that a reduced stomata density in the pod will also correlated with an increase in seed number.

## 4.5 Concluding Remarks

The hypothesis that more stomata lead to shorter fruits was tested in three different species. I showed that the phenotypes related to stomata density that was previously described in leaves for several of the mutants is also evident in fruits. Furthermore, stomata density indeed correlates negatively with fruit length. This was consistent for *Arabidopsis*, *B. oleracea* and *B. napus*. However, the reason for this effect remains unknown and effects on the rest of the plant were not studied. It is possible that slight changes in stomata density could be favourable for crop improvement whereas large changes are likely to have too many pleiotropic effects on the overall plant development.

# Chapter 5

## 5 A genome-wide association study to identify candidate genes underlying key yield traits

## 5.1 Introduction

*Brassica napus* is a species that has been bred for many different agricultural purposes. Some of these are fodder rape, swede, kale and winter and spring oilseed rape (OSR). These are vastly different plants bred for their leaves, roots or seeds. The large amount of variation thus found can be used to identify genes underlying specific traits. Several methods exist to identify and establish which genes are important for these traits, each with their own advantages and disadvantages.

Quantitative Trait Loci (QTL) mapping methods are often used to define a region underlying a trait of interest. The method requires a population that is closely related and stems from two parents. The sequenced lines are phenotyped and the output of the analysis is a genomic region underlying a trait of interest. The region can be large and contain many genes. Narrowing down the region to one gene that affects the desired trait is a lengthy process which requires time and money, hence QTLs are often not fully followed up (Doerge, 2002). Problems also include the fact that due to different genetic maps of different quality, studies can often not be fully compared. However, advantages of QTL analyses are that the resolution of gene sequencing does not need to be very high, making it affordable and a good choice for complex, unsequenced genomes.

Another way of going from phenotype to genotype is using Genome Wide Association Studies (GWAS) which allow comparisons of plants from a more diverse background (Brachi et al., 2011). This method is becoming increasingly popular with the advent of next generation sequencing technology. In a GWAS, plants of the same species can be compared almost regardless of their relatedness. This sets it apart from QTL studies which need a population derived from two known parents. Furthermore, the GWAS compares single-nucleotide polymorphisms (SNPs) to each other and can therefore detect genes that correlate to a trait. This is much more focused than the region acquired from a QTL study. On the other hand, rare alleles do not get picked up, as SNPs with a frequency of less than 5% are discarded from the analysis (Brachi et al., 2011). This is also the reason why it is important to consider the plant material used in a GWAS. It allows comparison of a large subset of plants but the population studied needs to be carefully selected to include the right amount of variation (Bazakos et al., 2017). The predictive power of GWAS increases with more samples. Ideally, QTL and GWAS are used in conjunction to identify a gene underlying a trait. A gene picked up in the GWAS that is located in a QTL is a good indication of a causal gene affecting the studied phenotype (Bazakos et al., 2017). Using both together can therefore eliminate false positives quickly as well as identify false negatives (Brachi et al., 2010).

So far, the methods presented have correlated a trait to DNA sequences. However, it is also possible to correlate a trait to gene expression. To this end, RNA-Seq can be used to generate Gene Expression Markers (GEMs) which can then be correlated to any phenotypic data available (Harper et al., 2012). This constitutes of a regression analysis finding correlations between a trait value and expression and this method can also pinpoint an exact gene. Combining the GWAS using GEMs with that using SNPs creates a powerful tool to identify variation underlying a trait.

Several examples exist where GWAS has been used to successfully identify genes underlying a specific trait. In *B.napus* a GWAS using just 53 accessions was successfully used to confirm known genes contributing to seed glucosinolate content (Harper et al., 2012). A GWAS using the GEMs further identified genes in the glucosinolate biosynthesis pathway that were known orthologues in *Arabidopsis* (Harper et al., 2012). The crops where GWAS has been most successful are rice and maize (Bazakos et al.,2017), where genes not previously associated with a phenotype such as grain width, grain quality, tassel length, ear length and flowering time have been identified (Huang et al., 2010, Wallace et al., 2014, Yang et al., 2014, Huang et al., 2012). Other crops where GWAS has been used successfully include wheat (Miller et al., 2016), sugarcane (Racedo et al., 2016) as well as the model plant *Arabidopsis* (Atwell et al., 2010, Brachi et al., 2010).

In this chapter a GWAS was carried out using SNPs and GEMs to find genes underlying pod and seed traits in *B. napus*. A candidate gene correlating with pod length was chosen for further analysis because it came up when using multiple different trait files in multiple analyses and because of its known function in *Arabidopsis*. The coding region and promoters were sequenced in several accessions to determine why the expression of the gene differed between lines. Furthermore, putative transcription factor binding sites were determined and those conserved between *Arabidopsis* and *B. napus* are described.

s

## 5.2 Materials and Methods

### *5.2.1 Plant material and phenotyping*

A subset of a Diversity Fixed Foundation Set (DFFS) called the ASSYST panel (described by Bus et al. (2011)), was grown in a polytunnel at JIC. This subset of 88 lines included five kales, five Chinese “Semi-winter” OSR lines, one synthetic, one Chinese winter vegetable, five modern winter oilseed rape UK cultivars (WOSR), fourteen spring oilseed rapes (SOSR), six swedes, four

winter fodder, 45 other, older UK WOSR cultivars and two unspecified cultivars. They were grown in a replicated design with six weeks vernalisation. Several traits on these plants were measured by Rachel Wells and colleagues. These traits included valve length, beak length, seeds per pod, weight of seeds per pod, thousand seed weight (TSW), seed surface area, seed length and seed width. I myself measured pod width and Marlène Carrière, a Masters student under my supervision, measured total pod lengths. All traits measured showed a normal distribution.

### *5.2.2 mRNA-seq and marker scoring*

Leaf RNA for all lines was extracted and Illumina mRNA-Seq of 80-bp reads obtained (Harper et al., 2012). This was mapped onto a reference genome of Cabriolet. This was carried out by Andrea Harper, Rachel Wells and Ian Bancroft.

### *5.2.3 Mixed Linear Modelling-SNP analysis*

Associative transcriptomics were carried out as described by Harper et al. (2012). SNPs present at minor allele frequency (<5%) were removed, leaving 144,131 SNP markers for the analysis. To correct for population structure, the SNP data set was analysed with the Bayesian clustering algorithms implemented in STRUCTURE The SNP dataset was used to construct a Q-matrix. This queried the relatedness of the accessions using the sequencing data as input. It was set to attempt to fit 1-5 populations within the dataset. The best fit was found when four distinct populations were fitted (Fig. 5.1). The kinship matrix thus made was used for the further analysis. This is necessary to reduce the number of false-positives picked up later on (Pritchard et al., 2000). This allowed the identification of groups clustering accessions from different ancestral backgrounds. The output from the STRUCTURE analysis was used as a Q-matrix for the associative transcriptomics analysis afterwards. The trait data, SNP data and Q-matrix were incorporated into a Mixed Linear Model (MLM) algorithm performed using TASSEL V3.0 (Bradbury et al., 2007, Harper et al., 2012, Miller et al., 2016). This calculated how related each pair of accessions were to each other from their genotypes. This method was already set up and protocols written. I ran the analysis.

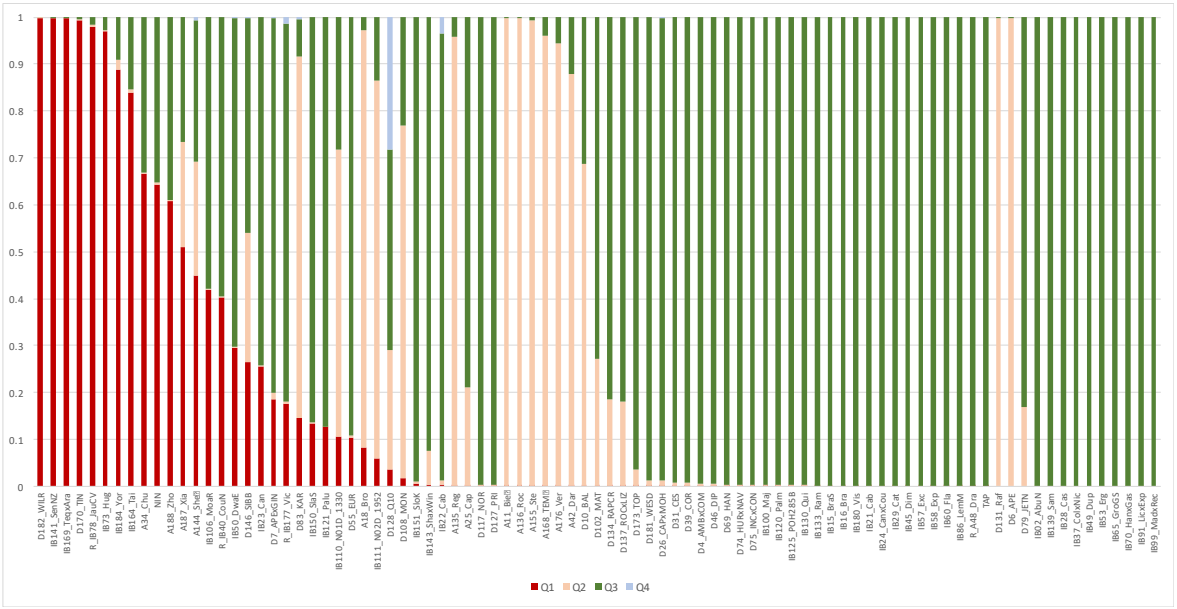



Fig. 5.1. The DFFS accessions were split into four populations for the GWAS. Red (Q1), pink (Q2), green (Q3), light blue (Q4).

### *5.2.4 Gene expression marker (GEM) analysis*

The reads per kb per million aligned reads (RPKM) were quantified and normalized and the transcripts with a RPKM of less than 0.4 were filtered out. That left 189,116 GEM markers used for the analysis. The dataset was then used for a linear regression correlating the RPKM values and the measured trait values. Results were plotted in a Manhattan Plot using the obtained P values (as $-\log_{10}P$) in pseudomolecule order. This was done using R (cran.r-project.org). The methods were developed by Harper, Bancroft and Wells. Again the protocols were there and I ran the analysis myself.

## 5.3 Results

### *5.3.1 Three components underlie seed and pod traits in* B. napus

To understand more of the underlying processes making up yield in *B. napus* I carried out a principal component (PC) analysis using the measured pod and seed traits. This yielded three main components. PC1 was made up of thousand seed weight (TSW) (Eigen value of 0.444), seed area (0.436), seed length (0.428) and seed width (0.445). This can be summarized as the seed

component. PC2 included pod length (0.493), seeds per pod (0.540) and the weight of seeds per pod (0.445). This can be described as pod filing components. The last PC included the negative correlation of beak length (-0.446) and pod width (0.854).

To see whether the DFFS was made up of distinct groups of plants, PCs were plotted against each other (Fig. 5.2). This can be done to isolate separate groups showing distinct characteristics. As the DFFS included some accessions bred for seed yield such as the oilseeds and some bred for the leaves such as the fodder rapes, it is possible that the difference could be seen in a PC. However, no distinct groups were found and all types of *B. napus* included in the DFFS clustered together.

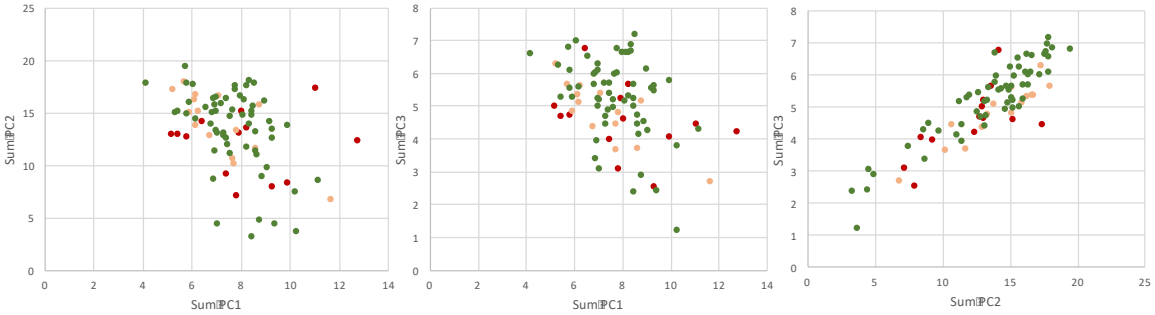



Fig. 5.2. A principal component analysis did not identify distinct groups in the dataset. PC1 included TSW, seed area, seed length and seed width. PC2 was made up of pod length, seeds per pod and the weight of the seeds per pod. In PC3 were beak length and pod width. Traits were measured on 87 accessions of the DFFS. Accessions were coloured by their predominant group according to STRUCTURE (Q1 (Red), Q2 (Pink) and Q3 (Green)).

### *5.3.2 A GWAS identifies the same candidate gene in several datasets*

To find genes underlying yield in *B. napus* a GWAS was carried out using several of the measured traits. A candidate gene was chosen from these for several reasons. The expression marker for the gene came highly correlated in the GWAS using the GEMs with several different datasets (2014 valve length data, the pod length data and the 2015 dataset), there was a small peak in the GWAS using the SNPs surrounding this gene and the known gene function in *Arabidopsis* in the cell cycle suggested it could be involved in fruit growth. Final valve length gave the most promising candidates and was used for further study. Valve length ranged from 3.39cm to 10.67cm with an overall mean of 6.19cm.

As the analyses gave several potential candidates, I tested whether the genes truly did correlate with the trait of interest. To this end, the expression of a gene of interest was plotted against

valve length. After doing this for several genes, it quickly became obvious that due to an outlier, Zhong Shuang (Zho), with a valve length of 10.67cm, the correlations that were coming up were not true correlations. Upon removal of Zho in the graphs, the correlations often broke down completely (Fig. 5.3). Therefore, the GWAS was done again on the same dataset without the Zho data point. In doing so, the outputs changed.

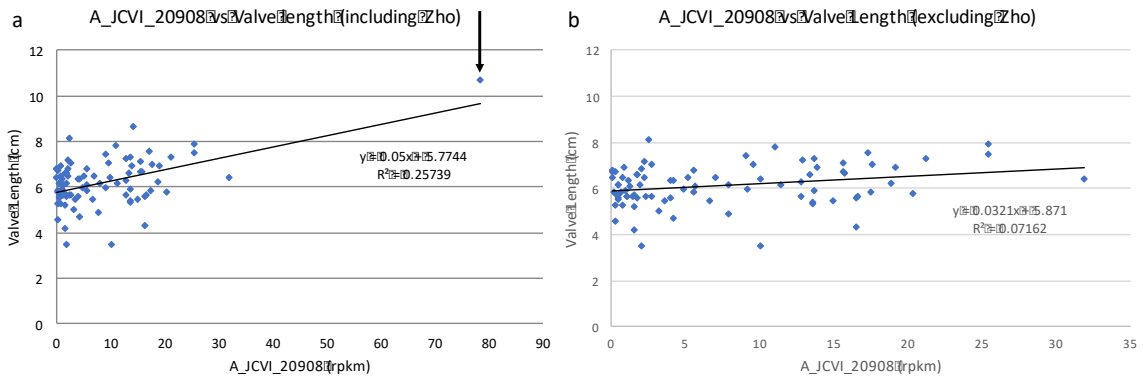



Fig. 5.3. An outlier affected results (a). Upon removal of the outlier (Zho), the correlations often broke down (b). Arrow in (a) points at the Zho data point.

One candidate gene was picked up in SNP and GEM associations, which has been annotated as being most similar to the *Arabidopsis* gene with the systematic name, *At5G03415* (Fig. 5.4). This gene encodes a potential DIMER PROTEIN b (DPb) homologue (see further description below) and ranked 7$^{th}$ among the strongest genes in the GEM-based GWAS (p=3.93E-06). Moreover, it had also been picked up when Zho was included in the analysis – in 10$^{th}$ highest place (p=1.32E-6) (Table 5.1). Interestingly however, not the same gene homoeologues were found. The A3 homoeologue was picked up when excluding Zho and C3 homoeologue when including it (*BnaA3.DPb* and *BnaC3.DPb* according to the agreed gene nomenclature for the Brassica genus (Ostergaard and King, 2008)). The expression of *BnaA3.DPb* was negatively correlated to valve length and *BnaC3.DPb* was positively correlated (with slopes of -1.1 and 1.022 when using all Keder data, respectively) (Fig. 5.5). The same candidate was found using the full pod length. Here, it was the 10$^{th}$ highest correlated gene in the GEM and located on the A3 chromosome (p=4.51E-05). The RPKM values for the GEMs picked up ranged from 0.368 – 2.233rpkm for *BnaA3.DPb* and from 0 – 1.949rpkm for *BnaC3.Dpb*. *At5G03415* is well characterized in *Arabidopsis* and plays a role in the cell cycle. It is therefore widely expressed and a role in pod growth is plausible but has not yet been reported.

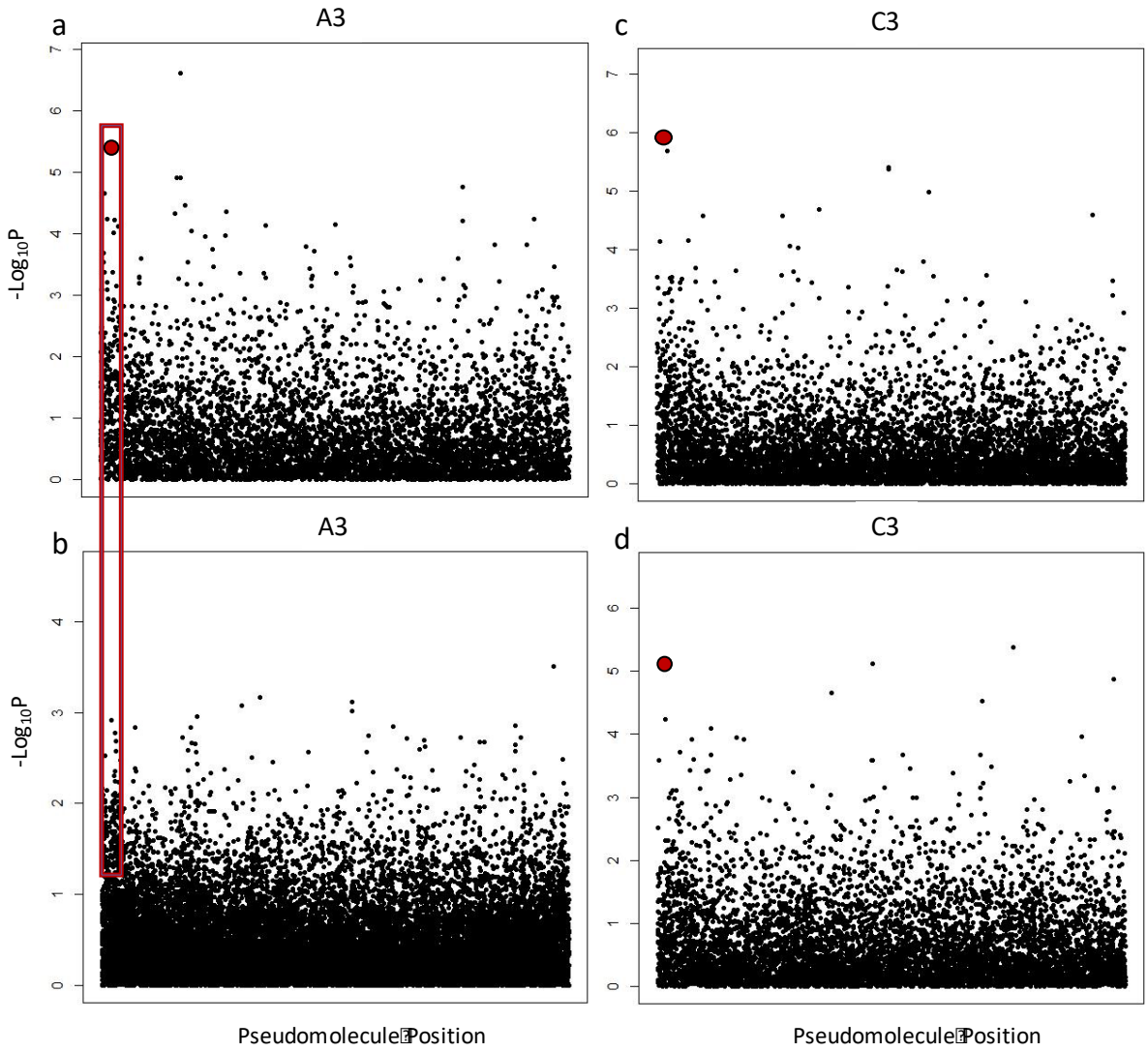



Fig. 5.4. GWAS candidate genes on A3 and C3. (a) GWAS using GEM markers and the valve lengths (excluding Zho) from 2014 gave *DPbA3.DPb* at the top. A small peak was also present in the GWAS using SNPs using the same dataset (b). The gene was also picked up on the C3 chromosome when including the outlier Zho in the dataset (c). Also in the valve data from 2015 which excluded Zho, the same marker came up on C3 when doing the GWAS using GEMs (d). Red dot denotes the marker of interest.

When carrying out the GWAS using the SNP markers, a peak in the same region, at the beginning of chromosome A3 and C3, was present. This peak was not very large but included *At5G03415* and strengthened it as a gene of interest that may affect valve length in *B. napus* (Fig. 5.4). The SNP to the left of the peak was *At5G01450* and the top of the peak was *At5G03560*. These two flanked *At5G03415*. The two genes located in the SNP peak did not score highly in the GEM

analysis with *At5G01450* being the 2150th best correlated gene and *At5G03560* the 8579th. Due to the low score in the latter, the genes were not chosen for further analysis.

As my research is focused on valve length and pod and seed traits it was pertinent to also do the GWAS using only the oilseed rape accessions, the only type of *B. napus* where the economically important trait is in the fully developed fruit. *At5G03415* did not correlated highly in that dataset with neither of the gene copies being in the top 20 genes (Table 5.1). In fact, they were the 162nd (*BnaC3.DPb*) and 9347th (*BnaA3.DPb*) best associated genes.

The analyses were then re-done on valve length measurements from plants grown 2015. This dataset only included 80 lines as those with very late flowering time were not included in the trial, this excluded Zho. Here, *BnaC3.DPb* was again found and correlated 7th strongest (Fig. 5.4 and Table 5.2). This was the only gene consistently picked up which made it a strong candidate gene to follow up in further experiments.

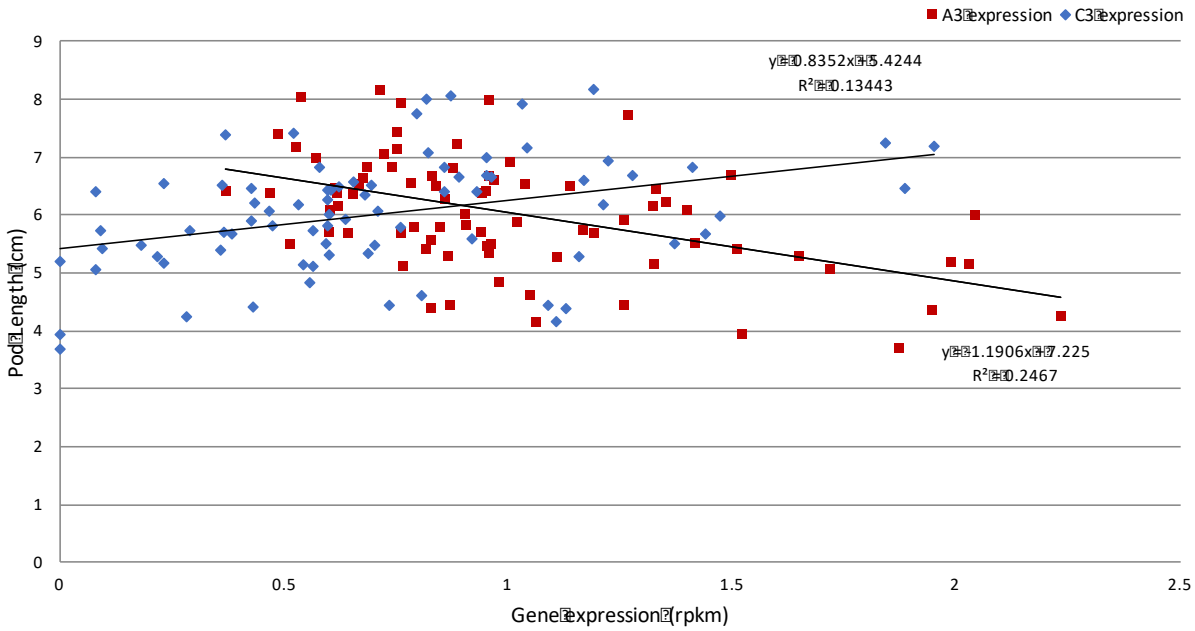



Fig. 5.5. The expression of At5G03415 correlated with valve and pod length in several datasets. Depending on the dataset used, the A3 or the C3 copy was better correlated. When plotting their expression patterns versus the pod lengths, *BnaA3.DPb* (unigene A_JCVI_18795) (red squares) was negatively correlated ($R^2$=0.247) whereas *BnaC3.DPb* (C_ES940427) (blue diamonds) showed a positive correlation with a lower $R^2$ value (0.134).

Table 5.1. 2014 valve length GWAS GEM results (excluding Zho). The top 20 best associated genes were considered to find target to work on further.

| | Valve Length 2014 (Excluding Zho) | | |
|---|---|---|---|
| | Unigene | AGI | Gene Name |
| 1 | A_EE463732 | AT2G43280.1 | FAR1-RELATED SEQUENCES-RELATED FACTOR4 |
| 2 | A_JCVI_32594 | AT5G20970.1 | HSP20-like chaperone superfamily protein |
| 3 | A_AM389200 | AT3G50810.1 | CASP-LIKE PROTEIN 5C2 |
| 4 | C_JCVI_27752 | AT5G41940.1 | Ypt/Rap-GAP domain of gyp1p superfamily protein |
| 5 | A_JCVI_28885 | AT4G26620.1 | Sucrase/ferredoxin-like family protein |
| 6 | A_JCVI_36750 | AT2G19110.1 | ARABIDOPSIS HEAVY METAL ATPASE 4 |
| 7 | A_JCVI_18795 | AT5G03415.1 | DIMER PROTEIN B |
| 8 | C_JCVI_5066 | AT1G29850.1 | PROGRAMMED CELL DEATH PROTEIN 5 |
| 9 | A_JCVI_20444 | AT5G59613.1 | ATP synthase |
| 10 | C_EX121700 | AT5G43880.1 | TON1 RECRUITING MOTIF 21 |
| 11 | A_JCVI_40510 | AT1G03080.1 | NETWORKED 1D |
| 12 | A_JCVI_9255 | AT5G20990.1 | CHLORATE RESISTANT 6, SIRTINOL 4<br>CO-FACTOR FOR NITRATE REDUCTASE AND XANTHINE DEHYDROGENASE |
| 13 | A_JCVI_9444 | AT5G20165.1 | Kish-A-like protein |
| 14 | A_JCVI_40734 | AT1G54390.4 | INHIBITOR OF GROWTH 2 |
| 15 | C_EV131911 | AT5G65380.1 | MATE efflux family protein |
| 16 | A_JCVI_25578 | AT2G24640.1 | UBIQUITIN-SPECIFIC PROTEASE 19 |
| 17 | A_JCVI_6607 | AT5G60710.1 | Zinc finger (C3HC4-type RING finger) family protein |
| 18 | A_JCVI_7140 | AT3G21350.1 | MED6 |
| 19 | A_JCVI_11094 | AT1G76560.1 | CP12 DOMAIN-CONTAINING PROTEIN 3 |
| 20 | C_JCVI_39738 | AT5G65380.1 | MATE efflux family protein |

### *5.3.3 Other members of the DP/E2F network did not score highly in the GWAS*

In *Arabidopsis,* it is known that DPb forms heterodimers with the E2Fs to mediate the $G_1$/S transition of the cell cycle and to activate downstream target genes (Magyar et al., 2000). Therefore, I asked whether genes of the E2F family scored highly in the GWAS using the GEMs. There are three E2Fs in the Brassicas that DPb can form a dimer with; E2Fa-c (Mariconti et al., 2002). None of them scored highly in the dataset of the 2014 valve length. This could suggest a novel role for *DPb*, which is independent of the function with E2F.

### *5.3.4 Two other genes were present in the 2014 and 2015 datasets of valve length*

Two other genes were identified in the 2015 datasets and in some of the sets from 2014. These was the orthologue of *At1G29850*, which was the 14$^{th}$ most correlated gene in 2015 and was also picked up in the 2014 datasets of the valve length (without Zho) and full pod length (Table 5.1). In both of those analyses, the gene was the 8$^{th}$ most highly correlated. The candidate gene in

question is a potential orthologue of the *PROGRAMMED CELL DEATH PROTEIN 5* (*PDCD5*) gene. The other gene that came up highly correlated twice (15$^{th}$ in 2015 and 12$^{th}$ in the 2014 valve length dataset without Zho), is the orthologue of *At5G20990*, which has several names; *CO-FACTOR FOR NITRATE REDUCTASE XANTHINE DEHYDROGENASE* (*CNX*) and *SIRTINOL 4* (*SIR4*). Neither of these genes have annotated functions that support an obvious role in pod growth and hence they were not further studied.

### *5.3.5 GWAS on other yield traits were also carried out*

Further GWAS were carried out using some of the other measured traits. The 20 best correlated genes from the GEM using the data for seed area, seeds per pod, pod strength, beak length and pod width are displayed in Table 5.3. Furthermore, as seed size and seed number correlated, a new trait was calculated, namely seeds per pod x seed area. This trait was also used for the GWAS using SNPs and GEMs. The GWAS with SNPs gave a very strong peak on chromosome A8 and C3 and a smaller peak on the A3 and C7 chromosomes (Fig. 5.6a). The gene located at the peak of chromosomes A8 and C3 was the orthologue of *At3G46100; HISTIDYL-tRNA SYNTHETASE 1* for which several SNP markers were present at the top of the peak. Not much is known about the role of *At3G46100* but other tRNA synthases have been shown to be involved in ovule and seed development and mutants display aborted ovules and seeds in *Arabidopsis* (Berg et al., 2005). The GWAS with the GEMs did not pick up the same candidate. Its best correlated gene was the *At4G38600* orthologue*; Ubiquitin Protein Ligase 3* (*UPL3*) on Chromosome C3 (Fig. 6b). *UPL3* has mainly been described for its role in trichome development (Downes et al., 2003). This gene also scored highly in the GWAS on seed size and Charlotte Miller (Bevan lab, JIC) characterised it further. She identified a connection to seed size without affecting seeds per pod (personal communication). As none of these candidates were found in the GEM and the SNP correlations, the pod length candidate was the only one followed up.

### *5.3.6 Using* At5G03415 *to identify co-expressed genes*

Once I had identified *DPb* as a strong candidate gene for regulating pod length, I attempted to find other genes with a similar expression pattern in the DFFS. The rationale was that two genes that are regulated in a similar way could be working together. To this end, the expression of *At5G03415* was used as the trait for the GWAS instead of the previously used phenotypic traits such as valve length. As *BnaA3.DPb* and *BnaC3.DPb* had come up, the analysis was done with the

expression of both markers (Table 5.4). The region at the beginning of the A3 and C3 chromosomes correlated strongly to the expression of *DPb* (Fig. 5.7). In that region one gene in particular caught my interest as its published function suggests that it could be involved in similar processes as *At5G03415*. This gene is *At5G07350* and is annotated to encode a Tudor-SN Protein (Table 5.4), which is also involved in controlling the cell cycle (see further description below). It correlated 15$^{th}$ highest in *BnaC3.DPb* dataset and 42$^{nd}$ and 47$^{th}$ highest when using *BnaA3.DPb* as the trait. When plotting *At5G03415* expression against that of *At5G07350* a correlation of 22-23% was found consistently, regardless of which copies were correlated (Fig. 5.8).

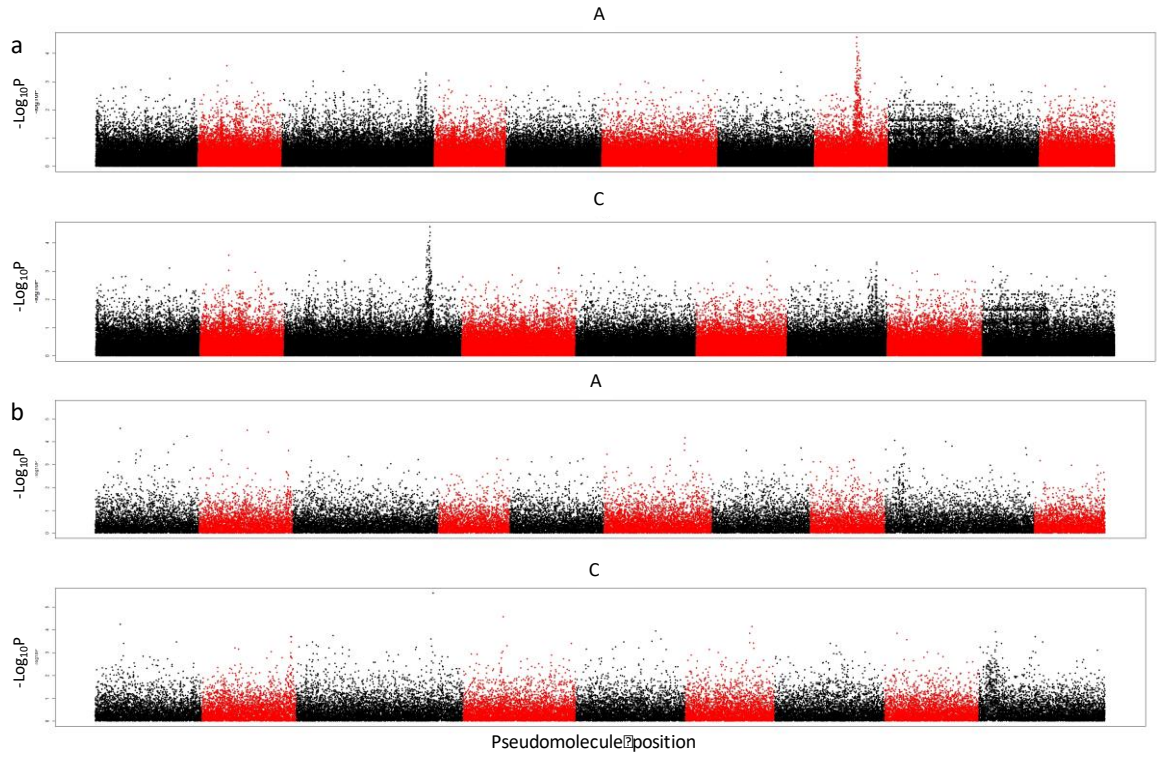



Fig. 5.6. Manhattan plot of the calculated trait of seed area x seeds per pod. Strong SNP peaks were present (a). At the top of the SNP peaks on A8 and C3 is *At4G46100*. The GWAS using GEMs had its strongest-correlating gene also on C3 but slightly to the right of the SNP peak (b). Manhattan plots rank the GEMs or SNPs according to their correlation (using $-Log_{10}P$) and order them along the pseudomolecule. The colour changes denote the chromosomes with the ten A chromosomes at the top and the nine C chromosomes at the bottom.

### *5.3.7 Paralogues of At5G03415 in B. napus*

There are four copies of *DPb* in *Brassica napus* located on Chromosomes A3, C3, C2 and A10. The GWAS picked up the A3 and C3 copies making it likely that either is dominant or that they function together.

Table 5.2. Summary of the GWAS using GEMs and the different datasets. Highlighted is *At5G03415* which was found in several different datasets.

| | **Valve Length, All Keder** | | **Valve Length, No Zho** | | **Valve Length, OSR 2014** | | **Full Length 2014** | | **Valve Length, all 2015** | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Unigene | AGI | Unigene | AGI | Unigene | AGI | Unigene | AGI | Unigene | AGI |
| **1** | C_EX074554 | AT1G13060.1 | A_EE463732 | AT2G43280.1 | A_ES921547 | AT5G03870.1 | A_EE463732 | AT2G43280.1 | C_DY010973 | AT3G03330.1 |
| **2** | A_JCVI_20908 | AT1G61800.1 | A_JCVI_32594 | AT5G20970.1 | A_JCVI_25721 | AT3G10870.1 | A_JCVI_28885 | AT4G26620.1 | A_JCVI_6607 | AT5G60710.1 |
| **3** | C_JCVI_20908 | AT1G61800.1 | A_AM389200 | AT3G50810.1 | A_JCVI_39441 | AT2G17620.1 | A_JCVI_32594 | AT5G20970.1 | C_JCVI_25752 | AT1G23560.1 |
| **4** | C_EV147906 | AT1G61800.1 | C_JCVI_27752 | AT5G41940.1 | A_JCVI_1398 | AT2G42570.1 | A_EV145332 | NA | A_JCVI_26222 | AT1G25370.1 |
| **5** | A_EV117664 | AT5G24580.1 | A_JCVI_28885 | AT4G26620.1 | A_JCVI_20908 | AT1G61800.1 | C_JCVI_27752 | AT5G41940.1 | A_JCVI_18238 | AT5G46190.1 |
| **6** | C_JCVI_21739 | AT3G21560.1 | A_JCVI_36750 | AT2G19110.1 | C_JCVI_2085 | AT3G22840.1 | A_AM389200 | AT3G50810.1 | A_JCVI_38959 | AT3G17950.1 |
| **7** | C_DV643282 | AT3G16150.1 | A_JCVI_18795 | AT5G03415.1 | C_JCVI_20908 | AT1G61800.1 | C_JCVI_12508 | AT5G18180.1 | C_ES940427 | AT5G03415.1 |
| **8** | A_EE517757 | AT5G09300.1 | C_JCVI_5066 | AT1G29850.1 | C_DV643282 | AT3G16150.1 | C_JCVI_5066 | AT1G29850.1 | C_JCVI_28299 | AT3G04810.1 |
| **9** | A_JCVI_15524 | AT3G45140.1 | A_JCVI_20444 | AT5G59613.1 | C_JCVI_19085 | AT5G26280.1 | C_DY025151 | AT3G04680.1 | A_EX121358 | AT5G38260.1 |
| **10** | C_ES940427 | AT5G03415.1 | C_EX121700 | AT5G43880.1 | C_EV147906 | AT1G61800.1 | A_JCVI_18795 | AT5G03415.1 | A_JCVI_26382 | AT1G12830.1 |
| **11** | C_JCVI_19823 | AT3G16150.1 | A_JCVI_40510 | AT1G03080.1 | A_JCVI_15524 | AT3G45140.1 | A_JCVI_9444 | AT5G20165.1 | C_AM395826 | AT3G15980.2 |
| **12** | A_JCVI_39441 | AT2G17620.1 | A_JCVI_9255 | AT5G20990.1 | C_JCVI_21739 | AT3G21560.1 | A_JCVI_25578 | AT2G24640.1 | A_CX280968 | AT2G28930.1 |
| **13** | C_JCVI_24791 | AT5G04470.1 | A_JCVI_9444 | AT5G20165.1 | A_JCVI_9170 | AT3G21560.1 | C_JCVI_42542 | AT2G32070.1 | A_JCVI_19822 | AT2G41680.1 |
| **14** | A_JCVI_35992 | AT3G09210.1 | A_JCVI_40734 | AT1G54390.4 | C_EX074554 | AT1G13060.1 | C_JCVI_36991 | AT3G43590.1 | C_JCVI_5066 | AT1G29850.1 |
| **15** | A_CX193364 | NA | C_EV131911 | AT5G65380.1 | C_JCVI_7629 | AT2G47880.1 | C_EV073899 | AT3G02555.1 | A_JCVI_9255 | AT5G20990.1 |
| **16** | A_JCVI_25721 | AT3G10870.1 | A_JCVI_25578 | AT2G24640.1 | A_JCVI_15758 | AT5G17220.1 | C_AM389238 | AT3G01610.1 | C_JCVI_18657 | AT4G00355.3 |
| **17** | C_JCVI_25721 | AT3G10870.1 | A_JCVI_6607 | AT5G60710.1 | A_EX138466 | AT5G62390.1 | A_JCVI_8987 | AT5G11910.1 | A_JCVI_36993 | AT4G35510.1 |
| **18** | C_JCVI_19085 | AT5G26280.1 | A_JCVI_7140 | AT3G21350.1 | A_EE539916 | AT5G60400.3 | A_JCVI_7579 | NA | C_JCVI_1863 | AT5G03290.1 |
| **19** | A_JCVI_4644 | AT1G72290.1 | A_JCVI_11094 | AT1G76560.1 | C_EV170929 | AT3G28740.1 | A_JCVI_40510 | AT1G03080.1 | C_JCVI_15128 | AT5G06150.1 |
| **20** | C_JCVI_4644 | AT1G72290.1 | C_JCVI_39738 | AT5G65380.1 | A_EE517757 | AT5G09300.1 | A_JCVI_17178 | NA | A_JCVI_38749 | AT5G04710.1 |

Table 5.3. The top 20 hits for every other trait used for the GWAS using the GEMs.

| | Seed Area | | Seeds per Pod | | Pod Strength | | Beak Length (no Zho) | | Pod Width | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Unigene | AGI | Unigene | AGI | Unigene | AGI | Unigene | AGI | Unigene | AGI |
| 1 | A_JCVI_3634 | AT1G33170.1 | A_JCVI_1611 | AT3G11930.2 | A_DY017774 | NA | C_JCVI_34134 | AT1G29740.1 | A_EE510149 | AT5G53045.1 |
| 2 | A_CD827525 | AT2G40460.1 | C_AM389238 | AT3G01610.1 | A_JCVI_3301 | AT5G07470.1 | A_EX056678 | AT3G14650.1 | A_JCVI_20569 | AT3G52250.1 |
| 3 | A_JCVI_7926 | NA | C_JCVI_9950 | AT3G27890.1 | C_JCVI_3301 | AT5G07470.1 | A_EE527512 | AT5G03720.1 | C_JCVI_34747 | AT5G10910.1 |
| 4 | A_JCVI_2068 | AT5G26280.1 | C_JCVI_23676 | AT5G62890.1 | C_JCVI_20470 | AT3G19000.1 | A_JCVI_34134 | AT1G29740.1 | A_JCVI_20549 | AT5G53045.1 |
| 5 | C_JCVI_3634 | AT1G33170.1 | A_JCVI_5227 | AT1G62540.1 | A_EE442039 | AT5G02610.1 | A_EX044469 | AT3G14690.1 | C_EE510149 | AT5G53045.1 |
| 6 | A_JCVI_16339 | AT2G22240.1 | A_ES998075 | AT2G41250.1 | A_EX090552 | AT3G26744.1 | C_JCVI_39746 | AT5G48360.1 | C_JCVI_20549 | AT5G53045.1 |
| 7 | A_JCVI_9358 | AT5G66460.1 | C_EV068994 | AT5G47820.1 | A_JCVI_20470 | AT3G19000.1 | A_EX022680 | AT5G46520.1 | A_JCVI_578 | AT5G61500.1 |
| 8 | C_EV028699 | AT1G70830.2 | A_JCVI_18606 | AT1G17210.1 | A_JCVI_32263 | AT3G19000.1 | A_EX098518 | AT5G46260.1 | C_EX087365 | AT5G66730.1 |
| 9 | A_JCVI_34762 | AT1G04770.1 | C_EX043867 | AT5G62430.1 | C_JCVI_10044 | AT3G19000.2 | C_JCVI_38689 | AT1G12220.1 | A_JCVI_9473 | AT3G52090.1 |
| 10 | A_EV208510 | AT4G03200.1 | A_JCVI_20744 | AT1G08510.1 | A_JCVI_10044 | AT3G19000.2 | A_EV153614 | AT3G14860.2 | A_JCVI_6768 | AT4G24680.1 |
| 11 | C_ES929581 | AT1G32660.1 | A_JCVI_4008 | AT1G29850.2 | C_JCVI_13315 | AT5G26220.1 | C_AM389512 | AT3G22770.1 | C_EE482690 | AT5G06560.1 |
| 12 | A_JCVI_494 | AT5G54160.1 | C_JCVI_1126 | AT1G54020.2 | A_JCVI_12931 | AT2G28790.1 | C_ES941913 | AT2G30590.1 | A_EX087365 | AT5G66730.1 |
| 13 | C_JCVI_11679 | AT1G73480.1 | A_JCVI_5069 | AT3G06180.1 | C_JCVI_15878 | AT5G08280.1 | C_JCVI_4610 | AT3G63120.1 | A_EV196661 | AT5G05580.1 |
| 14 | A_JCVI_11679 | AT1G73480.1 | C_EX104820 | AT5G62090.1 | A_JCVI_5622 | AT1G12080.2 | A_AM388847 | AT1G52360.1 | C_EV120597 | AT2G36570.1 |
| 15 | A_EE475977 | AT2G22240.1 | C_JCVI_17418 | NA | C_JCVI_28098 | AT1G72290.1 | A_JCVI_10791 | AT1G52340.1 | A_JCVI_9940 | AT4G22140.2 |
| 16 | C_EV189813 | AT3G21670.1 | C_JCVI_33692 | AT1G69370.1 | C_JCVI_12931 | AT2G28790.1 | C_JCVI_21227 | NA | C_EE534088 | AT4G12540.1 |
| 17 | A_JCVI_33578 | AT5G09760.1 | A_JCVI_32594 | AT5G20970.1 | C_JCVI_25015 | AT4G29020.1 | A_JCVI_16979 | AT3G20870.1 | A_DY009766 | AT5G08420.1 |
| 18 | C_JCVI_35192 | AT5G01510.1 | A_EV183049 | AT1G64780.1 | A_JCVI_25015 | AT4G29020.1 | A_EE524516 | AT5G13740.1 | C_JCVI_23461 | AT1G15660.1 |
| 19 | C_JCVI_39168 | AT2G31750.1 | A_JCVI_1126 | AT1G54020.2 | A_JCVI_13315 | AT5G26220.1 | A_EX122419 | AT1G29720.1 | A_JCVI_27011 | AT1G63640.1 |
| 20 | C_JCVI_27883 | AT3G18190.1 | A_CD816665 | AT4G27020.1 | C_JCVI_11340 | AT5G57180.2 | C_EX072990 | AT2G15730.1 | A_JCVI_5417 | AT2G44970.1 |

Table 5.4. GWAS using GEMs with the expression of A_JCVI_18795 and C_ES940427 as traits gave candidate genes with expression patterns similar to that of *At5G03415.* Highlighted are the Tudor-SN Proteins of interest for further study.

| | (A_JCVI_18795) *At5G03415* A3 expression as trait | | | (C_ES940427) *At5G03415* C3 expression as trait | | |
|---|---|---|---|---|---|---|
| | Unigene | AGI code | Gene Name | Unigene | AGI code | Gene Name |
| **1** | A_JCVI_24288 | AT5G04410.1 | NAC DOMAIN CONTAINING PROTEIN 78 AND 2 | A_JCVI_26818 | AT5G12410.1 | THUMP domain-containing protein |
| **2** | A_JCVI_34962 | AT5G05210.1 | Surfeit locus protein 6 | C_JCVI_33620 | AT5G01910.1 | Myelin transcription factor |
| **3** | A_JCVI_27666 | AT3G09035.1 | Concanavalin A-like lectin family protein | C_JCVI_9579 | AT5G01350.1 | UvrABC system C |
| **4** | A_EX041450 | AT5G01030.1 | Enolase | A_JCVI_3552 | AT5G04280.1 | RNA-binding family protein |
| **5** | A_JCVI_16103 | AT5G01290.1 | mRNA capping enzyme family protein | C_JCVI_2175 | AT5G09650.1 | PYROPHOSPHORYLASE 6 |
| **6** | A_JCVI_9579 | AT5G01350.1 | UvrABC system C | C_JCVI_39644 | AT5G02590.1 | Tetratricopeptide repeat (TPR)-like superfamily protein |
| **7** | C_JCVI_37610 | AT5G01260.2 | Carbohydrate-binding-like foci | C_JCVI_14123 | AT5G14910.1 | Heavy metal transport/detoxification superfamily protein |
| **8** | C_JCVI_34802 | AT5G05740.1 | ETHYLENE-DEPENDENT GRAVITROPISM-DEFICIENT AND YELLOW-GREEN-LIKE 2 | C_JCVI_6727 | AT5G05670.2 | Signal recognition particle binding protein |
| **9** | A_JCVI_4080 | AT5G04430.1 | BINDING TO TOMV RNA1 | C_JCVI_40776 | AT5G11720.1 | Glycosyl hydrolase family 31 |
| **10** | C_JCVI_24800 | AT5G03340.1 | CELL DIVISION CYCLE 48C | A_JCVI_32451 | AT5G01920.1 | Protein kinase superfamily protein |
| **11** | C_JCVI_37182 | AT5G10110.1 | DNA-directed RNA polymerase subunit beta | C_JCVI_15794 | AT5G05180.1 | Myosin heavy chain, striated protein |
| **12** | A_JCVI_21365 | AT5G04080.1 | Cysteine-rich TM module stress tolerance protein | C_JCVI_24065 | AT5G03330.1 | Cysteine proteinases superfamily protein |
| **13** | A_JCVI_6523 | AT5G01010.1 | Retinal-binding protein | A_JCVI_24065 | AT5G03330.1 | Cysteine proteinases superfamily protein |
| **14** | A_JCVI_5226 | AT5G07240.1 | IQ-domain 24 | C_EE468587 | AT5G11480.1 | ENGB-2 |
| **15** | A_JCVI_22067 | AT5G13020.1 | EMSY-LIKE 3 | C_JCVI_16166 | AT5G07350.1 | Tudor-SN Protein 1 |
| **16** | C_EE521134 | NA | | C_JCVI_16721 | AT5G04420.1 | Galactose oxidase/kelch repeat superfamily protein |
| **17** | A_EE442388 | AT5G06110.1 | GONIDIALESS A/ZUOTIN RELATED FACTOR A1,A2,1B | C_JCVI_20268 | AT5G08580.1 | Calcium -binding EF hand family protein |
| **18** | C_JCVI_20697 | AT5G03120.1 | Transmembrane protein | C_JCVI_32451 | AT5G01920.1 | Protein kinase superfamily protein |
| **19** | A_JCVI_10595 | AT5G04750.1 | F1F0-ATPase inhibitor protein | C_ES944210 | AT5G08280.1 | HYDROXYMETHYLBILANE SYNTHASE, RUGOSA 1 |
| **20** | A_JCVI_5893 | AT3G14230.2 | RELATED TO AP2 2 | C_JCVI_39098 | AT5G04110.1 | DNA GYRASE B3 |
| | | | | | | |
| **42** | A_JCVI_16166 | AT5G07350.1 | Tudor-SN Protein 1 | | | |
| **47** | A_ES938255 | AT5G07350.1 | Tudor-SN Protein 1 | | | |

The oilseed rape Cabriolet has been sequenced and aligning its *DPb* protein sequences to each other and to the *Arabidopsis DPb* sequence showed *BnaA3.DPb* being the most conserved copy with amino-acid identity to *Arabidopsis DPb* of 70.6% (Table 5.5). *BnaA3.DPb* is very similar to *BnaC3.DPb* (82.8% identify) but *BnaC2.DPb* is more similar to the *Arabidopsis DPb* (69.9% identity), as opposed to 65.5% identity for *BnaC3.DPb*. *BnaA10.DPb* is vastly different to the others with only 30.9% identity to *Arabidopsis DPb*.

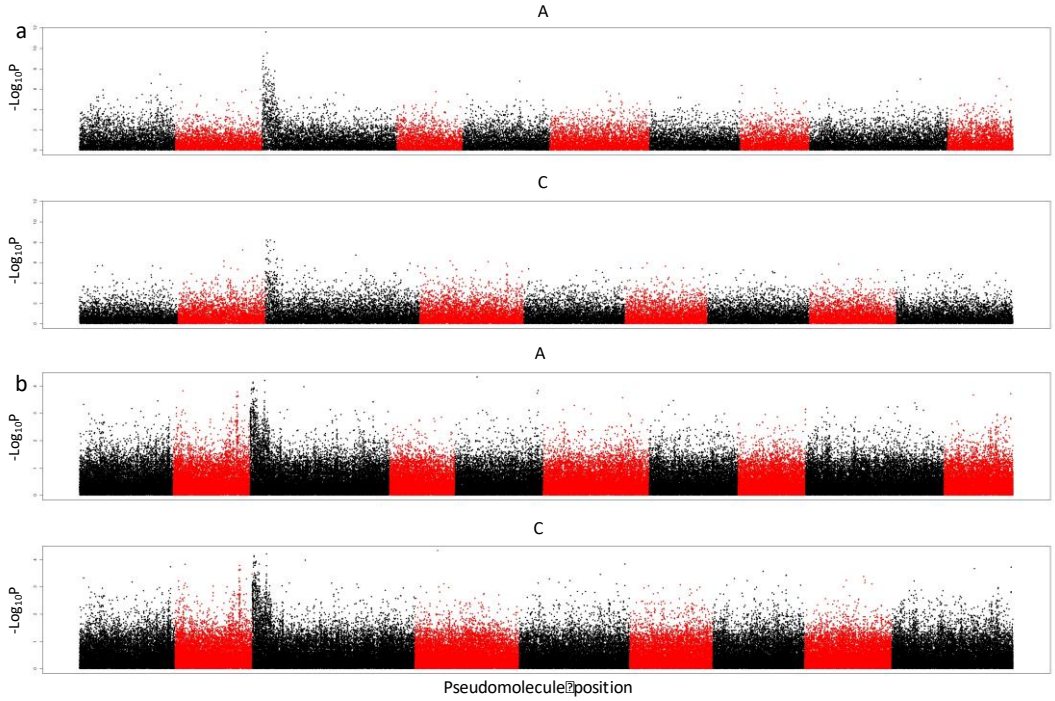



Fig. 5.7. GWAS carried out using the expression of the unigene A_JCVI_18795 (*BnaA3.DPb)* as the trait. The GWAS using the GEMs (a) as well as the SNPs (b) showed a strong peak/region correlated at the beginning of chromosomes A3 and C3.

Table 5.5. Protein sequence identities. The *B. napus* protein sequences were compared to each other as well as to *Arabidopsis* DPb. BnaA3.DPb was the most conserved protein while BnaA10.DPb differed widely from the others. Sequences are from Ensembl Plants and alignments were made using ClustalW.

| Protein | BnaC3.DPb | BnaA3.DPb | BnaA10.DPb | *Arabidopsis* DPb |
|---|---|---|---|---|
| **BnaC2.DPb** | 69.1 | 66.7 | 29.1 | 69.9 |
| **BnaC3.DPb** | | 82.8 | 26.7 | 65.5 |
| **BnaA3.DPb** | | | 28.3 | 70.6 |
| **BnaA10.DPb** | | | | 30.9 |

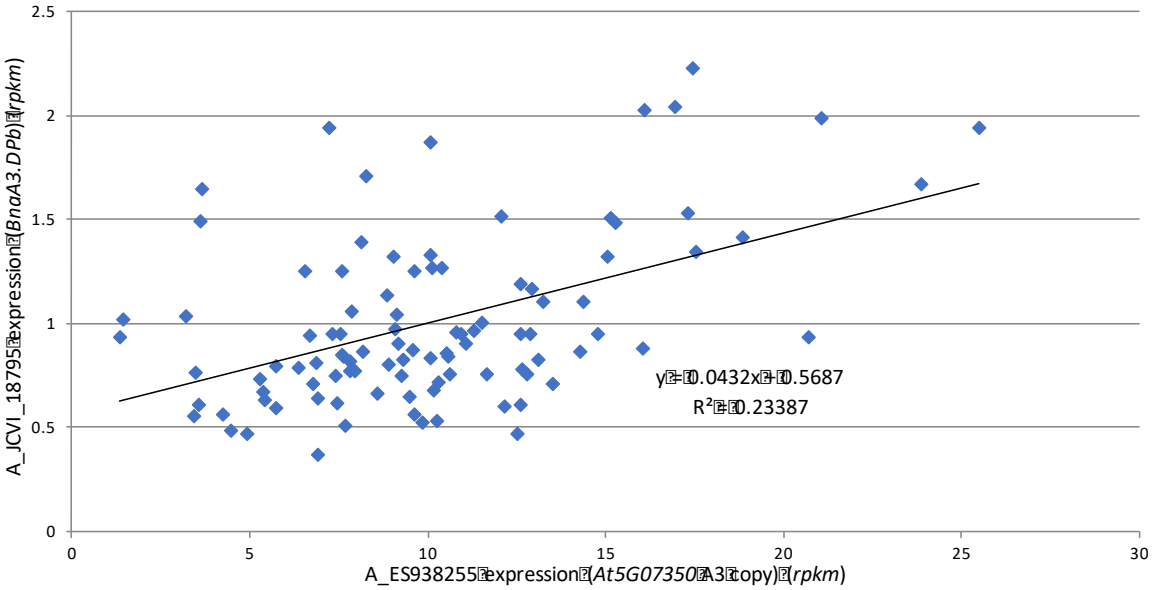



Fig. 5.8. Positive correlation of *BnaA3.DPb* and *BnaA3.TSN1.* The expression of the A_JCVI_18795 unigene which corresponds to *BnaA3.DPb* was positively correlated to the GEM A_ES938255 corresponding to the A3 copy of the *At5G07350* orthologue.

### *5.3.8 A subset of accessions was used for further analysis of Bna.DPb*

To gain an understanding as to why the *B. napus* orthologue of *At5G03415* was expressed differentially and therefore found in the GWAS using the GEMs, eight lines from the DFFS were chosen for analysis. Two were spring OSR (SOSR) accessions, five lines were winter oilseed rape (WOSR) accessions and one Chinese accession termed semi-winter. The latter does not have an obligate requirement for flowering but will flower faster and be more productive with a period of vernalisation. The WOSR lines were chosen for having different pod lengths (three large, three short) and for their differences in *BnaA3.DPb* expression. Therefore, the lines either highly expressed *BnaA3.DPb* and had short pods or large pods with low *BnaA3.DPb* expression. The lines chosen were AmberxCommanche (AmbxCom), Darmor (Dar), Quinta (Qui), Expert (Exp), Ramses (Ram), NingYou (Nin), Hanna (Han) and Duplo (Dup) (Table 5.6). In addition, the sequenced Cabriolet (Cab) was also included. It had high expression of *BnaA3.DPb* with rather long pods. The SOSR accession had to be dismissed due to a mix-up in the seed set which gave rise to very different plants supposedly of the same genotype. Expert was also excluded as it did not germinate well.

### *5.3.9* BnaA3.DPb *protein sequences do not show many differences*

The region surrounding *DPb* was found in the SNP-based GWAS and therefore I questioned whether there were differences in the coding sequences between lines. To this end, *BnaA3.DPb* sequences were ascertained in the chosen lines. The coding sequences showed some differences. However, once translated to protein sequences, the differences became minimal (Fig. 5.9). Quinta had three different amino acids at the beginning of the protein but these were outside of functionally defined regions (DNA binding domain, dimerization domain, E2F-binding domain) (Magyar et al., 2000). Nin had three unique mutations – one being a deletion. Overall however, the protein sequences did not differ much and did not suggest that sequence variation in the coding sequence was causal for the differences in valve length.

Table 5.6. *B.naA3.DPb* expression and valve lengths of chosen accessions.

| Accession | Spring/Winter | *BnaA3.DPb* expression | Valve Length |
|---|---|---|---|
| **Hanna** | S | 1.7155 | 5.2823 |
| **Duplo** | S | 0.9767 | 5.5997 |
| **AmbxCom** | W | 0.4873 | 6.8073 |
| **Darmor** | W | 0.5379 | 7.4953 |
| **Ramses** | W | 0.6142 | 6.8163 |
| **Expert** | W | 1.3212 | 6.2930 |
| **Quinta** | W | 1.5208 | 4.1910 |
| **NingYou** | Chinese | 1.9445 | 5.6363 |
| **Cabriolet** | W | 1.4949 | 6.6967 |

### *5.3.10* DPb *promoter sequences highlight an accession from a different background and a homeologous exchange*

To gain an insight into the differences in expression found between the lines, the promoters of *BnaA3.DPb* were sequenced using primers in the upstream gene and in *BnaA3.DPb* itself. This spanned a region of roughly 570 base pairs. There were several differences between the lines, the most pronounced being a ~700 base pair insertion in Nin (full alignment in Appendix). The other five promoters were over 90% similar to each other. The differences were too small to be able to draw a conclusion as to why some lines had lower *DPb* expression than others. Again the accession from China differed most to the others. PCR assays and sequencing were done by Marlène Carrière under my supervision.

The region between *BnaC3.DPb* and its upstream gene was also sequenced and comprised ~650 base pairs. The accessions were very similar except for the Qui promoter which had various insertions and deletions along the sequence (Fig. 5.10 and full alignment in Appendix). In fact, the

*BnaC3.DPb* promoter of Qui is the exact same as its *BnaA3.DPb* promoter. *BnaC3.DPb* in Quinta was not picked up at all in the mRNA-Seq for the transcriptome data. This difference in promoter sequence could explain that.

```
Qui             MTTTASNSTHNPHEDSPVKNPPSTTVSAQSVSSSGSMGSPSEQTTTVATTPASDDPASQL
Nin             MTTPPSNSIHNPHEDSPVKNPPSTTVSAQSVSSSGSMGSPSEQTTTVATTPASDDPASQL
Ram             MTTPPSNSIHNHHEDSPVKNPPSTTVSAQSVSSSGSMGSPSEQTTTVATTPASDDPASQL
AmbxCom         MTTPPSNSIHNHHEDSPVKNPPSTTVSAQSVSSSGSMGSPSEQTTPVATTPASDDPASQL
Cab             MTTPPSNSIHNHHEDSPVKNPPSTTVSAQSVSSSGSMGSPSEQTTTVATTPASDDPASQL
Dar             MTTPPSNSIHNHHEDSPVKNPPSTTVSAQSVSSSGSMGSPSEQTTTVATTPASDDPASQL
                ***..*** ** ********************************.**************

Qui             ASGSGKKKRRGQRASGPDKNGRGLRQFSLKVCEKVESKGRTTYNEVADELVAEFALPNND
Nin             ASGSGKKKRRGQRASGPDKNGRGLRQFSLKVCEKVESKGRTTYNEVADELVAEFALPNND
Ram             ASGSGKKKRRGQRASGPDKNGRGLRQFSLKVCEKVESKGRTTYNEVADELVAEFALPNND
AmbxCom         ASGSGKKKRRGQRASGPDKNGRGLRQFSLKVCEKVESKGRTTYNEVADELVAEFALPNND
Cab             ASGSGKKKRRGQRASGPDKNGRGLRQFSLKVCEKVESKGRTTYNEVADELVAEFALPNND
Dar             ASGSGKKKRRGQRASGPDKNGRGLRQFSLKVCEKVESKGRTTYNEVADELVAEFALPNND
                ************************************************************

Qui             DTSPDQQQYDEKNIRRRVYDALNVLMAMDIVSKYQKEILWRGLPRSSLSDIEELKAERLS
Nin             DTSPDQQQYDEKNIRRRVYDALNVLMAMDIVSKYQKEILWRGLPRSSLSDIEELKAERLS
Ram             DTSPDQQQYDEKNIRRRVYDALNVLMAMDIVSKYQKEILWRGLPRSSLSDIEELKAERLS
AmbxCom         DTSPDQQQYDEKNIRRRVYDALNVLMAMDIVSKYQKEILWRGLPRSSLSDIEELKAERLS
Cab             DTSPDQQQYDEKNIRRRVYDALNVLMAMDIVSKYQKEILWRGLPRSSLSDIEELKAERLS
Dar             DTSPDQQQYDEKNIRRRVYDALNVLMAMDIVSKYQKEILWRGLPRSSLSDIEELKAERLS
                ************************************************************

Qui             LRSRIEKKTEYSQELEEQYVGLQNLIRRNENLYSSGNAPNGGVTLPFILVQTRPNATVEV
Nin             LRSRIEKKTEYSQELEEQYVGLQNLIRRNEHLYSSGNAPNGGVTLPFILVQTRPNATVEV
Ram             LRARL-KKTEYSQELEEQYVGLQNLIRRNENLYSSGNAPNGGVTLPFILVQTRPNATVEV
AmbxCom         LRSRIEKKTEYSQELEEQYVGLQNLIRRNENLYSSGNAPNGGVTLPFILVQTRPHATVEV
Cab             LRSRIEKKTEYSQELEEQYVGLQNLIRRNENLYSSGNAPNGGVTLPFILVQTRPNATVEV
Dar             LRSRIEKKTEYSQELEEQYVGLQNLIRRNENLYSSGNAPNGGVTLPFILVQTRPNATVEV
                **:*: ************************:************************:*****

Qui             EISEDMQLVHFDFNTTPFELHDDNFVLKTMKFCDQPPLNNGHNNNSHETSHSFVPEENKE
Nin             EISEDMQLVHFDFNTTPFELHDDNFVLKTMKFCDQPPLNNGHNNN-HEASHSFVPEENKE
Ram             EISEDMQLVHFDFNTTPFELHDDNFVLKTMKFCDQPPLNNGHNNNSHETSHSFVPEENKE
AmbxCom         EISEDMQLVHFDFNTTPFELHDDNFVLKTMKFCDQPPLNNGHNNNSHETSHSFVPEENKE
Cab             EISEDMQLVHFDFNTTPFELHDDNFVLKTMKFCDQPPLNNGHNNNSHETSHSFVPEENKE
Dar             EISEDMQLVHFDFNTTPFELHDDNFVLKTMKFCDQPPLNNGHNNNSHETSHSFVPEENKE
                ****************************************** **:************

Qui             GLSTDPKLPQQVDTDQSHHQLHAQPQIIPTPVTISASAASNAPVTSPPLSGITKSSVKTE
Nin             GLSTDPKLPQQVDTDQSHHQLHAQPQIIPTPVTISASAASNAPVTSPPLSGITKSSVKTE
Ram             GLSTDPKLPQQVDTDQSHHQLHAQPQIIPTPVTISASAASNAPVTSPPLSGITKSSVKTE
AmbxCom         GLSTDPKLPQQVDTDQSHHQLHAQPQIIPTPVTISASAASNAPVTSPPLSGITKSSVKTE
Cab             GLSTDPKLPQQVDTDQSHHQLHAQPQIIPTPVTISASAASNAPVTSPPLSGITKSSVKTE
Dar             GLSTDPKLPQQVDTDQSHHQLHAQPQIIPTPVTISASAASNAPVTSPPLSGITKSSVKTE
                ************************************************************

Qui             N
Nin             N
Ram             N
AmbxCom         N
Cab             N
Dar             N
                *
```

Fig. 5.9. Protein sequences showed very little variation between the studied accessions.

Table 5.7. Putative transcription factor sites that were present in the *BnaA3.DPb* promoter of the six sequenced accessions. Highlighted are the two that were also present in all different copies of *At5G03415* in *Arabidopsis* and *B. napus.*

| | TF-Putative Site | Sequence | Binding Gene | Published Function | Publication |
|---|---|---|---|---|---|
| **1** | CARGCW8AT | CTTATTAAAG | AGL15 | Floral transition, embryo development | (Adamczyk et al., 2007, Perry et al., 1996) |
| **2** | ARR1AT | CGATT, AATCT, AGATT | ARR1 | Cytokinin response | (Taniguchi et al., 2007) |
| **3** | At1G23380 | AGTCA | KNAT6 | Inflorescence architecture, lateral root formation | (Belles-Boix et al., 2006, Dean et al., 2004) |
| **4** | At1G51600 | AATCT | GATA28/ ZML2 | Response to light | (Reyes et al., 2004, Shaikhali et al., 2012, Shikata et al., 2004) |
| **5** | SORLIP1AT | GCCAC | | In promoters of light-induced genes | (Alvarez-Canterbury et al., 2014) |
| **6** | At3G48160 | ACCGCGCGAA | DEL-1 | Cell division, endoreduplication, defence | (Chandran et al., 2014, Vlieghe et al., 2005) |
| **7** | MYBCOREATCYCB1 | CCGAT, AACCG, CCGTC | CYCB1;1 | Cell cycle | (Mironov et al., 1999, Weingartner et al., 2003) |
| **8** | At5G12870 | AACCA | MYB 46 | Secondary wall biosynthesis, defence against *Botrytis* | (Ramirez et al., 2011, Zhong et al., 2007) |
| **9** | At1G43700 | ACCGCT | VIP1 | Regulator of osmosensory signalling | (Tsugama et al., 2012) |
| **10** | At1G77920 | TGCCG, CGTTA, | TGA7 | Defence response to bacterium | (Hepworth et al., 2005, Jakoby et al., 2002, Kesarwani et al., 2007) |
| **11** | At2G22300 | CCGCGC | CAMTA3 | Stress response, defence | (Doherty et al., 2009) |
| **12** | U01377 | CCGTC, CCGAT | COR15A | Cold response, salt/osmotic stress response | (Liu et al., 2014) |
| **13** | At1G09030 | CCGAT, ATTGC | NF-YB4 | Embryo development, stress response | (Swain et al., 2017, Zhao et al., 2016) |

### *5.3.11 Several putative transcription factor binding sites are conserved in DPb promoters*

The differences in *BnaA3.DPb* promoters between the lines could suggest differences in gene regulation but could also have a role in its function. Using Plantpan2.0 putative transcription factor binding sites were identified that were conserved in all six *B. napus* accessions. Having the same ones in all lines suggests they are conserved and that they may be important. Thirteen putative binding sites were found conserved in the six genotyped accessions and *Arabidopsis*. The identified sites can be split into two main categories; plant defence and cell differentiation and growth (Table 5.7).

Next, to infer more about DPb's function and the genes it may interact with, the conserved binding sites between the *DPb* promoters of *Arabidopsis* and the *B. napus* C2, C3, A3, A10 copies were determined.

Only two sites were conserved between all five *DPb* promoters; the site for *At1G77920* and one for *MYBCOREATCYCB1*. Interestingly, these genes are involved in very different processes; *CYCB1;1* is involved in cell cycling, which is related to the role published for DPb. *At1G77920* encodes a bZIP transcription factor known to be involved in defence response to bacteria.

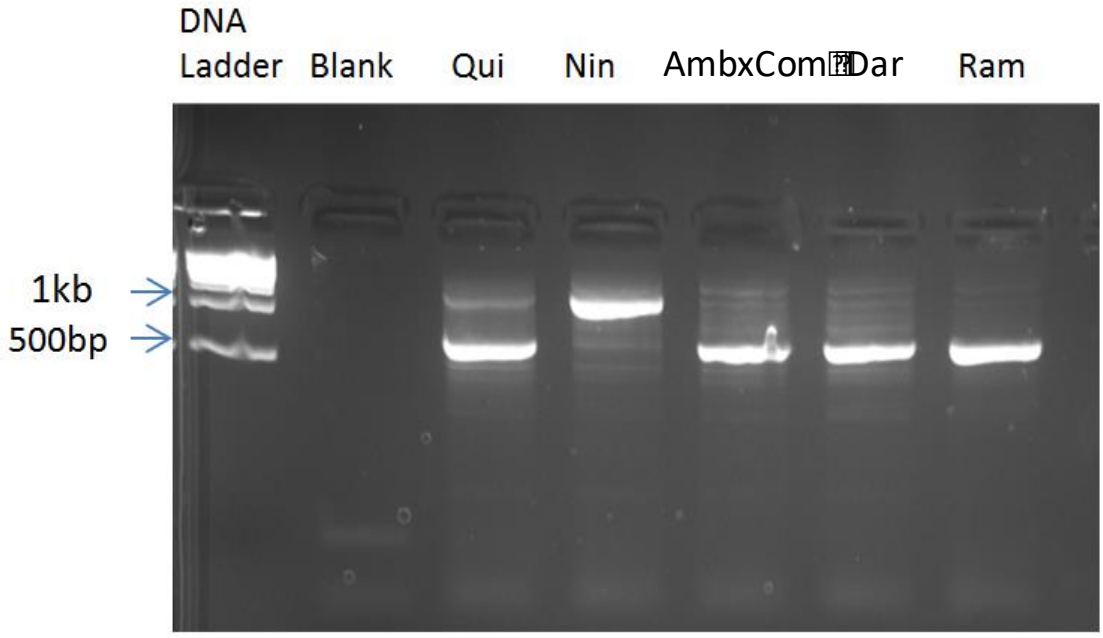



Fig. 5.10. PCR amplification of the *BnaA3.DPb* promoter. Nin had a longer intergenic region than the other lines. PCR carried out by Marlène Carrière. 1kb ladder was used.

### *5.3.12 The role of DPb in* B. rapa

*B. rapa* contains the A genome which hybridised with the C genome from *B. oleracea* to give rise to the allopolyploid *B. napus* containing the A and C genomes. A *B. rapa* EMS mutant population was produced at the JIC and developed into a TILLING platform for reverse genetics (Stephenson et al., 2010). A subset comprising 1154 lines were sequenced using exome-capture methods. In the database found on the RevGen website (https://www.jic.ac.uk/technologies/genomic-services/revgenuk-tilling-reverse-genetics/), several plants contained mutations in the *At5G03415* orthologue (*BraA3.DPb*) located on the A3 chromosome, *Bra005763*. Five accessions were purchased and grown at the JIC (Table 5.8). These lines contained a large amount of background mutations and therefore had to be backcrossed to the parent of the population, R-o-18, several times to be able to observe a potential phenotype caused by a mutation in the gene of interest. Due to time constraints, I only managed to backcross mutated lines once to wild type and most still exhibited strong pleiotropic phenotypes due to the mutated background. Many plants were infertile and no phenotype could be attributed to *BraA3.DPb*. Further backcrossing would be required to allow proper analysis of these mutants.

Table 5.8. Five *B. rapa* TILLING mutants were purchased. They were chosen for the severity of the mutations.

| | Codon Change | Protein Change | Type of Mutation |
|---|---|---|---|
| **1** | G - A | | Splice acceptor site |
| **2** | Cag -Tag | Gln - Stop | Premature Stop |
| **3** | Gag - Aag | Glu - Lys | Missense |
| **4** | Gtg - Atg | Val - Met | Missense |
| **5** | gGa - gAa | Gly - Glu | Missense |

## 5.4 Discussion

### *5.4.1 The gene of interest was chosen for its published function*

The GWAS allowed identification of genes correlating with traits related to pod morphology and growth. The method is subjective as it is the person carrying out the study who decides which candidate to follow up. The rationale for choosing *DPb* here came from the fact that it was picked up in multiple analyses. For a gene to come up when running the GWAS using both SNPs and GEMs is rare – in fact the candidate gene was the only one found to rank highly in both analyses.

Furthermore, many of the genes that correlated to the yield traits tested were uncharacterized proteins of unknown function. *DPb* therefore stood out in that respect too as some of its functions in *Arabidopsis* are known. Its published role in *Arabidopsis* is in the $G_1$/S transition of the cell cycle where it forms a heterodimer with the E2F family of transcription factors (de Jager et al., 2001, Vandepoele et al., 2002). As I am interested in pod growth, where cell divisions are occurring, I reasoned that it could be an interesting candidate to study as nothing was known on its function in fruit development. However, the RPKM of *At5G03415* had a range from 0 – 2.233 which is not very large compared to other genes such as for example the gene *At5G07350*, which had values ranging from 3.102 to 32.114 for its unigene A_JCVI_16166.

### *5.4.2 Other genes that were highlighted in the GWAS using SNPs*

Many factors are known to affect pod growth and fertility which include many biotic and abiotic stresses (Hua et al., 2016, Liu et al., 2015, Walton et al., 2012). This was mirrored in the GWAS where the lack of clear peaks showed that valve length is controlled by many factors. While there was a small SNP peak around *At5G03415*, the gene itself was not at the top of the peak. It is not uncommon for a GWAS peak not to be located exactly on the gene of interest but close to it (Huang et al., 2010). The gene at the top of the peak was *At5G03560*, encoding an uncharacterized tetratricopeptide repeat (TPR)-like superfamily protein. A different TPR-like superfamily protein encoded by *At5G02590*, was also found to correlate with the expression of *At5G03415* when doing the GWAS using GEMS and was also at the top of the SNP peak when using the same dataset (Table 5.4 and Fig. 5.5). Moreover, another TRP-like superfamily protein (*At2G02150*) orthologue was found in a QTL study on pod number and seed number, located on chromosome A6 (Shi et al., 2015). These proteins are found in all kingdoms and are mediators of protein-protein interactions (D'Andrea and Regan, 2003). In plants, links to several hormones have been made; *ETHYLENE OVERPRODUCER 1 (ETO1*) which contains TPR motifs at its C-terminus, is involved in ethylene biosynthesis (Yoshida et al., 2005) and *SPINDLY (SPY)* with TPR motifs at its N-terminal, is a gene linking gibberellin and cytokinin responses in plants (Greenboim-Wainberg et al., 2005). A link to ABA response has also been made through the TPR-containing gene *TETRATRICOPEPTIDE-REPEAT THIOREDOXIN-LIKE 1 (TTL1)* (Schapire et al., 2006). Hormones are known to play a role in fruit development and it is not unlikely that TPR-like superfamily proteins found here also play a role. However, *At5G03560* did not score highly in the GEM correlation where it ranked only at the 8579$^{th}$ place. It would be interesting to study *Arabidopsis* mutants in these genes to elucidate their potential role in fruit development.

Another gene that scored highly in the SNP correlation was *At5G01450*, *Aberrant Pollen Development 2 (ABD2)*. In *Arabidopsis* it, together with its three paralogues, plays a role in pollen mitosis II and plants defective in *ABD* have defective pollen which would affect the final size of the pod (Luo et al., 2012). However, since this gene also did not score highly in the GEM analysis and since the peak on the Manhattan plot of the SNP analysis was not that high or obvious, this gene was also not worked on further.

### *5.4.3 Reasons for the identification of different homoeologues depending on the data set used*

Interestingly, depending on the data set used in the GWAS, *BnaA3.DPb* or *BnaC3.DPb* appeared. Not much is known on how the duplication of genes impacts their expression. When genes are duplicated, they can retain their function, gain a new function or be deleterious. Some studies have shown that gene copies can work in an additive manner, this is the case for instance for *FLOWERING LOCUS C* in *B. rapa* (Schranz et al., 2002). Here, *BnaC3.DPb* was picked up more often than *BnaA3.DPb*, suggesting it to be correlated more strongly. However, when plotting valve length versus gene expression the expression of *BnaA3.DPb* always correlated stronger than the expression of *BnaC3.DPb*, with $R^2$ values of 24.67 and 16.24, respectively (in the dataset of all lines without the outlier Zho (the same pattern, but a lower correlation is present when plotting the data including Zho)). However, the p-value from the GEM was smaller for the association to *BnaC3.DPb* than the one to *BnaA3.DPb* (1.32E-06 and 3.93E-06, respectively).

Genes can also work together in a different manner; by compensating for the loss of each other. This could be the case here whereby *BnaC3.DPb* is the dominant copy but if the expression is low, *BnaA3.DPb* expression is upregulated. This could also explain their inverted correlation to pod length; *BnaC3.DPb* having a positive correlation and *BnaA3.DPb* a negative one.

The candidate gene was found in almost all datasets tested and over two years (2014 valve length, with and without the outlier Zho, 2014 pod length, 2015 pod length). It was not however picked up when running the analysis with just the 70 oilseed rape accessions. This could be because when including fewer lines the analysis loses statistical power. It could also be because the gene had already been selected for when producing oilseed rape lines with large pods with many seeds.

### *5.4.4 Candidates that were present in more than one years' data*

The *B. napus* orthologue of *At5G03415* was found in both years of growing the DFFS. Only two other genes were present in both 2014 and 2015. These were the orthologues of *At1G29850, PROGRAMMED CELL DEATH PROTEIN 5* (*PDCD5*) and of *At5G20990, CO-FACTOR FOR NITRATE REDUCTASE AND XANTHINE DEHYDROGENASE1* (*CNX1*) also called SIRTINOL 4*. PDCD5* plays a role in DNA damage responses induced by UV-B radiation and also in age-induced cell death (Ferreyra et al., 2016). *pdcd5* mutants also senesce slower (Ferreyra et al., 2016). The *SIRTINOL* genes are regulators of auxin signalling but SIR4/CNX1 is also in the biosynthesis pathway of the molybdenum cofactor (Moco) which is needed for many vital processes in the plant (Dai et al., 2005, Kaufholdt et al., 2016, Zhao et al., 2003). Neither of the genes were worked on further in this study.

### *5.4.5 The GWAS using* At5G03415 *expression as a trait highlighted a homoeologous exchange*

The GWAS using A3 and C3 *At5G03415* expression as the trait gave rise to Manhattan plots that strongly highlighted the beginning of chromosomes A3 and C3 where the gene itself is located, for both the correlation using the SNPs and the GEMs. This suggests either that genes in this region function as a cluster and are regulated together or that a homoeologous recombination event occurred there. The latter seems to be the explanation as a recombination in this region has been described by He et al. (2016). Furthermore, the fact that Qui's *BnaC3.DPb* expression was zero and that its promoter is identical to the one of the A3 copy suggest a recombination event. This could be tested further by looking into the region more closely and analysing the differences between lines with known recombination events and those without.

### *5.4.6 Candidate genes with expression patterns similar to that of* At5G03415

The GEM-based GWAS was done using A3 and C3 *At5G03415* expression as the trait with the aim to identify possible interactors. Many of the genes that were in the top hits of the analysis were uncharacterized genes with not much information about them. Two genes stood out because of their function. *At5G03340*; CELL DIVISION CYCLE 48C, came up 10th best correlated against *BnaA3.DPb*. It could have been interesting to follow up as it is involved in cell-cycle control, similar to DPb. TUDOR-SN PROTEIN1 (TSN1), *At5G07350*, stood out as it was found in the GWAS using the unigenes on both the A3 and the C3 chromosomes. The known functions of it were also

promising; in *Arabidopsis*, it has been shown to be involved in gibberellin biosynthesis (Liu et al., 2010, Yan et al., 2014a), a hormone necessary for proper growth including pod growth (Fuentes et al., 2012). TSN1 also plays a role in programmed cell death (PCD), particularly in the embryo suspensor (Sundstrom et al., 2009). This latter role is relevant as cell cycle genes, including DPb, often regulate the cell cycle as well as PCD (Wang et al., 2014, Wu et al., 2009). In humans TSN has been linked directly to E2F expression as well as the $G_1$/S transition of the cell cycle (Elbarbary et al., 2017, Su et al., 2015). These known gene functions suggested that perhaps the DPb and TSN pathways overlap in development. This question will be addressed in Chapter 6.

### *5.4.7 Protein sequences of* At5G03415 *showed little variation between accessions*

The protein sequences of *BnaA3.DPb* exhibited a high level of conservation across the chosen lines of the DFFS. This is in accordance with the genes scoring highly in the GEMs. Gene expression is different between the lines and that is most likely to be affected due to variation in the promoter or other non-coding regions and not in the protein sequence itself.

### *5.4.8 Sequencing of* At5G03415 *promoter sequences highlighted Nin as different*

Several differences were found when comparing the *BnaA3.DPb* promoters. This result suggests that the subset is large enough to identify differences. Nin, the Chinese semi-winter line was quite distinct compared to the others as it had a large insertion. It is perhaps not surprising that it is the most different as it originates from a different geographical area and will have come from a different gene pool. Furthermore, the breeding of Chinese cultivars often includes crosses with *B. rapa* which could have changed the genome further (Bancroft et al., 2011). The differences in the conserved parts are miniscule however and this is reflected by the GEMs which did not show gene expression changes on a large scale.

### *5.4.9 Conserved putative binding sites may provide information about DPb function*

Conserved putative binding sites were found in the *DPb* promoters. Many were binding sites for transcription factors involved in controlling the cell cycle. They therefore seem likely to be functional cis-elements as they fit with a role of DPb in the cell cycle. The link to the B-type cyclins for instance seems intuitive as both are involved in the same process. The cyclins are necessary for the progression of the cell cycle (Mironov et al., 1999). However, while DPb and the E2Fs are

mainly known as mediators of the $G_1$/S transition, CyclinB1;1 is a marker of $G_2$/M (Mironov et al., 1999). However, a direct link between CyclinB1;1 and E2Fc exists in which expression of the former is reduced in plants overexpressing the latter (Del Pozo et al., 2007).

Several genes were found that do not have a role in the cell cycle but instead function in plant defence. Many pathogens affect the cell cycle but it would be interesting to follow this up by for instance looking at changes in *DPb* expression after inoculation with a pathogen. A study on the effect of *Cabbage leaf curl virus* in *Arabidopsis* for instance found over 300 cell cycle genes affected by infection, including *E2Fc* (Ascencio-Ibanez et al., 2008). Another study showed similar results in which plants mutated in the three *E2Fs* lacked effector-triggered immunity (ETI) and were more susceptible to virulent pathogens (Wang et al., 2014). Further investigating the link between pathogen response and the cell cycle could be very interesting.

*TGACG SEQUENCE-SPECIFIC BINDING PROTEIN 7* (*TGA7*) was one of the two genes identified through the analysis of putative binding sites in all *Arabidopsis* and *B. napus DPbs*. TGA7 is a basic domain leucine zipper (bZIP) transcription factor and *Arabidopsis* has ten of them in five clades (Jakoby et al., 2002). The main published function of these proteins is in plant defence, in particular in the systemic acquired resistance (SAR) (Kesarwani et al., 2007, Zhang et al., 2003). However, a link to floral organ formation has also been made as the TGA family is able to interact with NPR1-like genes, of which *Arabidopsis* has five (Hepworth et al., 2005). BOP1 and BOP2 are part of this latter group and TGA7 has been shown to form a protein bond using a yeast two-hybrid system (Hepworth et al., 2005). Like DPb, TGA7 therefore also seems to be involved in plant defence and plant development. It would be interesting to study the link between TGA7 and DPb in these two contexts.

## 5.5 Concluding Remarks

A *B. napus* orthologue of the *Arabidopsis* gene *At5G03415* was found repeatedly to be associated with pod morphology and growth in the GWAS study presented here. It would be interesting to elucidate the role it plays in fruit development. To understand more about the role of this gene in regulating pod growth, Chapter 6 explores the DPb/E2F network in *Arabidopsis* and the role the genes play in fruit growth. *B. rapa* and *B. oleracea* are the diploid progenitors of the allotetraploid *B. napus* and it could be interesting to study the gene function of *DPb* in those species. It would also be simpler than in *B. napus* as they are not polyploids and potential redundancy between

*DPb* paralogues would be less of a problem. The *B. rapa* exom-capture mutants discussed here still had too many background mutations for proper phenotypic analysis, however, with more backcrosses it may be possible to assess the involvement of *BraA3.DPb* in *B. rapa* fruit development. Other genes found through the GWAS could also be interesting to study as they may also contribute to improving yield of *B. napus.*

# Chapter 6

## 6 Overlapping functions of E2F and TSN in fruit development

## 6.1 Introduction

Reproduction is arguably the most important goal of any living being. The making of a seed in plants or a foetus in animals is a tightly regulated process and mistakes are often fatal. The model plant *Arabidopsis* reproduces by making many small seeds contained in a fruit called silique. The fruit develops from a central, cylindrical gynoecium. It can be split into three main parts; the ovary containing the ovules and later seeds, the style and the stigmatic papillae, on which the pollen land (Fig. 6.1). Upon germination, pollen tubes guide the sperm cells towards the ovules to facilitate fertilization and subsequent seed development. When mature, the female gametophyte contained in the ovules contains a central cell, an egg cell and two synergid cells which are all required for correct growth of a seed (Fig. 6.2) (Drews and Yadegari, 2002). The synergid cells are required for the attraction and recognition of pollen tubes and for the correct release of the sperm cells (Yang et al., 2010). Once the pollen has reached an ovule, the egg cell and central cell are fertilized by two different sperm cells. From there, the central cell develops into the endosperm, the egg cell into the seed embryo and the seed coat is derived from the ovule integuments, surrounding the other parts (Drews and Yadegari, 2002). Interestingly this process of making new life also involves cell death and problems in the cell death pathway can be just as fatal as those involved in cell proliferation.

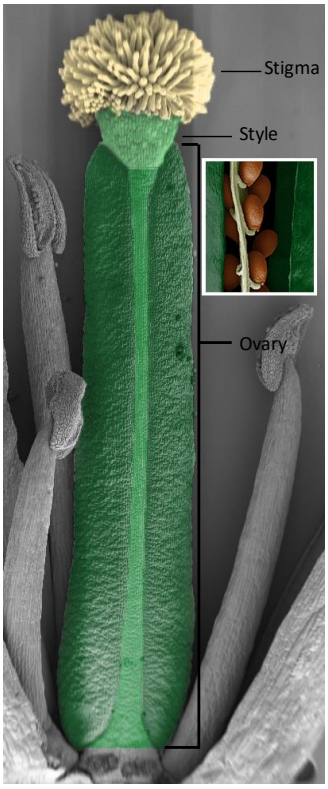



Fig. 6.1. The Arabidopsis fruit is made of several distinct parts. At the top is the stigmatic tissue (yellow) upon which the pollen lands. Beneath it is the style (light green) and the ovary (dark green) in which the seeds (inlet) later develop. False coloured SEM picture of a Col-0 fruit.

In the proper development of a seed several different cell death pathways are activated. After pollen lands on the tip of the gynoecium, the transmitting tract undergoes programmed cell death (PCD), which is the intentional removal of cells. The transmitting tract is located in the stylar region of the gynoecium and as it degrades, it makes space for the pollen tubes, guiding them towards the ovules (Crawford et al., 2007). The pollen tubes themselves are next to degrade after releasing the sperm cells in the ovules (Kessler and Grossniklaus, 2011). After fertilization, the zygote develops into two parts; the apical and basal parts, also called the embryo proper and the suspensor (Fig. 6.2). The embryo proper will develop into the mature seed. The suspensor is necessary to anchor the embryo, to supply it with nutrients and for the production of hormones (Kawashima and Goldberg, 2010, Smertenko and Bozhkov, 2014). It is only needed for the initial steps of development and degenerates when the embryo proper is at the late globular stage and is completely abolished at the end of the torpedo stage (Bozhkov et al., 2005). One of the final parts to be removed is the endosperm. It develops alongside the zygote but undergoes PCD at the last stages of development where, in *Arabidopsis*, only one layer of the endosperm remains (Van Hautegem et al., 2015). The final cell type that needs to die for the production of a functional seed is the hard seed coat made up of several layers (Daneva et al., 2016).

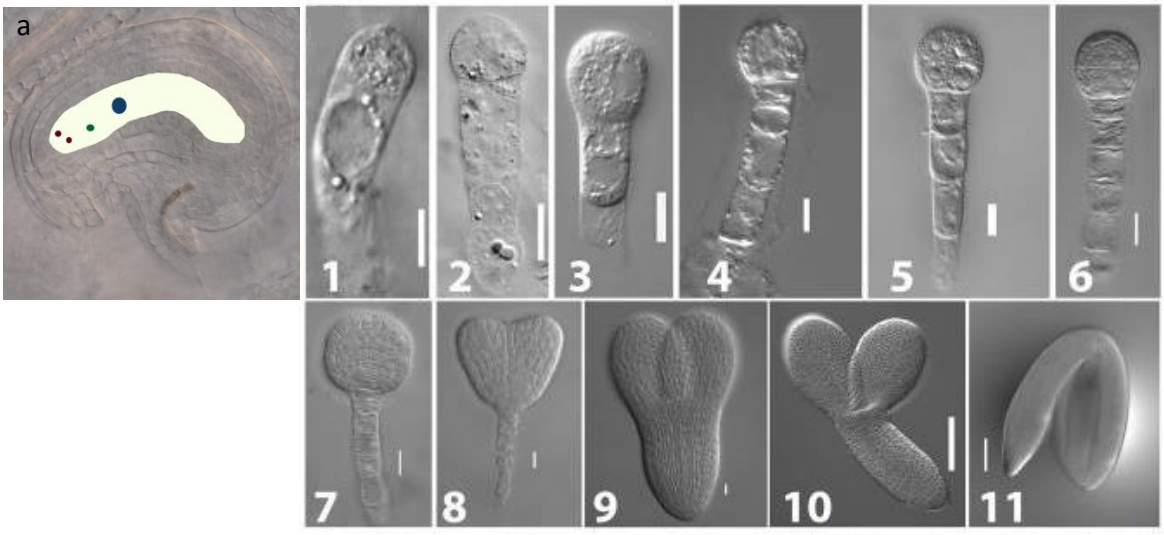



Fig. 6.2. The embryo sac (light green) is located in the ovule. Inside it are the central cell (blue), the two synergid cells (red) and the egg cell (green). As the embryo develops, it goes through several distinct stages. Stage 1; elongated stage, (2) one-cell, (3) two cell, (4) quadrant, (5) octant, (6) dermatogen, (7) globular, (8) heart, (9) torpedo, (10) bend, (11) mature stage. Embryos adapted from Xiang et al. (2011)

### *6.1.1 Tudor-Staphyloccocal-Nuclease (TSN)*

Many genes involved in cell death in *Arabidopsis* are known. Two of these are the functionally redundant paralogues *Tudor-Staphyloccocal-Nuclease 1* and *2* (*TSN1* and *TSN2*). They are RNA-binding proteins located in the cytoplasm (Sundstrom et al., 2009). They have the same expression pattern and are found in roots, leaves, flowers, siliques and mature seeds (Frey et al., 2010, Gutierrez-Beltran et al., 2015, Liu et al., 2010). TSN1 and TSN2 (called TSN from here on) contain four consecutive Staphyloccocoal-Nuclease (SN) domains at the 5' end followed by a Tudor domain and one more SN domain. TSN promotes cell proliferation and is cleaved upon cell death. The cleavage has been shown to be carried out by the metacaspase mcII-Pa in *Picea abies*. In mammals too TSN is cleaved by caspases (Sundstrom et al., 2009). In *P. abies* and in *Arabidopsis* TSN plays a role in fertility; *Arabidopsis tsn* heterozygous mutants ($tsn1^{+/-}$ $tsn2^{+/-}$) have reduced pollen numbers and few developed seeds due to early embryo suspensor abortions (Sundstrom et al., 2009). A model for the function of TSN in plants has been proposed whereby TSN helps to suppress cell death by stabilising the expression of protease inhibitors (Tsiatsiani et al., 2011). These in turn suppress cell-death proteases (Tsiatsiani et al., 2011). Once TSN is cleaved, the repressive action is lost and proteases are activated, leading to PCD.

TSN also function in processes that do not involve PCD. TSN have been found to play a role in heat and salt stress responses (Frey et al., 2010, Gutierrez-Beltran et al., 2015, Yan et al., 2014a). They are part of stress granules (SGs) and processing bodies (PBs), structures produced upon stress. These are needed to protect cells from death and repress the translation of mRNAs (Gutierrez-Beltran et al., 2015). In *tsn* mutant plants SGs and PBs can still form but do so slower and fewer are present in total. Furthermore, ectopic cell death occurs in the roots (Gutierrez-Beltran et al., 2015).

TSN are required for the biosynthesis of gibberellin (GA) (Liu et al., 2010, Yan et al., 2014a). Plants deficient in *TSN* show reduced levels of *Ga20ox3*, an enzyme involved in gibberellin biosynthesis. This translates to 20% less GA in *tsn2* mutants and a reduced germination rate which can be rescued upon gibberellin addition (Liu et al., 2010). Similar effects are seen in *tsn1* plants where, under salt stress, plants display the same phenotypes as plants defective in *ga20ox3-1* and *ga20ox3-2*, namely slower growth rates (Yan et al., 2014a). Furthermore, overexpression of *TSN1* generates plants with bigger rosettes and earlier flowering, phenotypes also associated with higher GA levels (Yan et al., 2014a).

*TSN* is conserved throughout kingdoms and is also present in animals. Plant and animal TSN share the same structure of four SN domains followed by a Tudor domain and a final SN domain. Human TSN (HsTSN) is present in the nucleus and the cytoplasm, while plant TSN is absent from the nucleus. HsTSN has been demonstrated to play a role in the cell cycle, namely in the $G_1$/S transition (Elbarbary et al., 2017, Su et al., 2015). Overexpression of *HsTSN* promotes the $G_1$/S transition and knocking it down causes $G_1$ arrest (Su et al., 2015). A direct interaction with E2F-1 promotes the activation of *E2F*-downstream targets such as *CYCLIN A*, promoting the cell cycle progression (Su et al., 2015). Furthermore, TSN degrades miRNAs which repress $G_1$/S-necessary genes such as *CDK2*, *Cyclin D1, E2F1* and *E2F2* (Elbarbary et al., 2017). More is known about *HsTSN* than the plant orthologue and no cell cycle role has been found for *TSN* in plants thus far.

### *6.1.2 Adenovirus E2 promoter binding factors (E2Fs)*

The balance between cell proliferation and cell death is tightly regulated in all living beings. Therefore, genes active in cell division often also have a role in PCD, as does *TSN*. Another set of genes characterized as having a role in both are the *Adenovirus E2 promoter binding factors* (*E2Fs)*. These have been found in many species such as humans, *Drosophila*, wheat, carrot, tomato, tobacco and *Arabidopsis* (Abraham and del Pozo, 2012, Albani et al., 2000, Ramirez-Parra et al., 1999, Sekine et al., 1999). There are three typical *E2Fs* in *Arabidopsis*; *E2Fa – c* (de Jager et al., 2001, Vandepoele et al., 2002). These are transcription factors (TFs) containing a DNA binding domain, a DP heterodimerization domain, a RBR binding domain as well as a transactivation domain (de Jager et al., 2001, del Pozo et al., 2006, Mariconti et al., 2002). The three paralogues are all involved in the $G_1$/S transition of the cell cycle. In quiescent cells and during the $G_1$ phase of the cell cycle, the E2Fs repress cell cycle-regulated promoters and then later activate the genes required for the $G_1$/S transition (Fig. 3) (Mariconti et al., 2002). It is widely accepted that E2Fa and E2Fb promote the cell cycle, whereas the function of E2Fc is ambiguous. Some results suggest that E2Fc, due to its shortened C-terminal transactivation domain, forms a repressor complex promoting endoreduplication (Cheng et al., 2013, de Jager et al., 2009, del Pozo et al., 2002, Gutierrez et al., 2002, Magyar et al., 2005) while others have found a role in gene activation (Mariconti et al., 2002). All three are bound by RETINOBLASTOMA-RELATED (RBR) which prevents the activation of downstream targets because RBR shares the E2F transactivation domain (Ramirez-Parra et al., 1999, Uemukai et al., 2005). To release the E2Fs, RBR gets phosphorylated by CDKA and CYCD (Fig. 3) (Boniotti and Gutierrez, 2001, Nakagami et al., 1999, Nakagami et al., 2002). A heterodimer between an E2F and a DIMER PROTEIN (DP) then enables downstream

target activation (Mariconti et al., 2002). There are two *Arabidopsis DPs*; *DPa* and *DPb* (de Jager et al., 2001, Vandepoele et al., 2002). All E2Fs can bind to both DPs but E2Fa and E2Fb preferentially bind to DPa whereas E2Fc preferentially binds to DPb (Kosugi and Ohashi, 2002b). The E2Fs function through their consensus site of TTT(C/G)(C/G)CG(C/G)(C/G) and the site is mostly situated within 400 bp of the start codon (Ramirez-Parra et al., 2003, Vlieghe et al., 2005). E2Fa and E2Fb have nuclear localization signals (NLS) but E2Fc lacks this site (de Jager et al., 2001). While they contain an NLS it is thought that there is a nuclear-export signal which dominates the localization. The E2F proteins are therefore located in the cytoplasm as well as in the nucleus (Kosugi and Ohashi, 2002b). Only when DPa, and not DPb, binds to E2Fa or E2Fb, they move from the cytoplasm to the nucleus, probably because the binding of DPa covers the export signal (Kosugi and Ohashi, 2002b). E2Fc does not move to the nucleus upon coexpression with DPa or DPb, possibly due to its truncated N-terminus (Kosugi and Ohashi, 2002b).

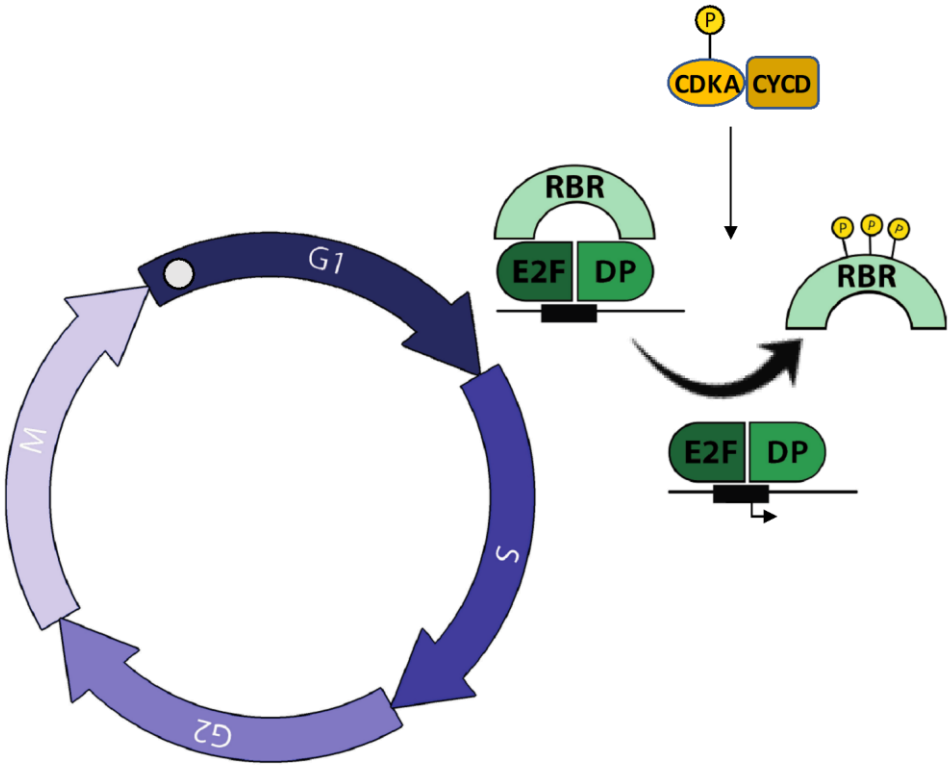



Fig. 6.3. The E2Fs promote the G1/S transition of the cell cycle. The E2F/DP heterodimer is bound by RBR suppressing their action. Upon phosphorylation of RBR by CDKA and CYCD, E2F downstream targets are activated.

#### 6.1.2.1 E2Fa

In plants, the E2Fs have mostly been characterized for their role in the cell cycle progression. *E2Fa* is strongly expressed in the shoot apical meristem, young leaf primordia, the vascular tissues of

leaves and buds as well as in roots (De Veylder et al., 2002, He et al., 2004). While it is strongly present in dividing cells, it is also found in differentiated tissues undergoing endoreduplication such as the root cortex and the hypocotyl of 5-day-old plants (De Veylder et al., 2002). *E2Fa* is transcriptionally regulated during the cell cycle to mediate the $G_1$/S transition and its expression peaks shortly before the S-phase and then rapidly decreases (de Jager et al., 2001, Mariconti et al., 2002). Overexpressing *E2Fa* produces plants with larger cotyledons due to smaller but more cells than wildtype plants (De Veylder et al., 2002, He et al., 2004, Magyar et al., 2005). In leaves however, 33% fewer cells are found while in older leaves larger cells are present (He et al., 2004).

*E2Fa*-overexpressing plants are dwarfed with narrower leaves (He et al., 2004) and phenotypes are enhanced when overexpressing *E2Fa* in conjunction with *DPa*. Some tissues show more cells and others more endoreduplication, hinting at distinct roles for E2Fa and DPa (De Veylder et al., 2002). Overexpressing *E2Fa* and *DPa* prolongs mitotic activity (Magyar et al., 2005) and *Arabidopsis* protoplasts transformed with *35S::E2Fa* or with *35S::DPa* are able to reenter S-phase (Rossignol et al., 2002). However, the production of a constitutively expressed *E2Fa*, by the deletion of the RBR1-binding domain, produces cells that differentiate prematurely (Magyar et al., 2012). Overall, E2Fa is believed to be an activator of transcription but some genes are also downregulated by it (de Jager et al., 2009). This downregulation is likely through a different mechanism than the classical E2F function as these genes do not have enriched E2F-sites in the promoters (de Jager et al., 2009). E2Fa is therefore seen as having a bimodal role where it can induce cell division as well as cell growth.

#### 6.1.2.2 E2Fb

*E2Fb* is also maximally expressed at the $G_1$/S transition but, unlike *E2Fa*, its levels stay constant throughout the other cell cycle phases (Mariconti et al., 2002). *E2Fb* is widely expressed in roots, the SAM, young leaves and their vascular tissues, trichomes as well as inflorescence meristems, sepals, petals, styles, anthers, pollen grains and ovaries (Sozzani et al., 2006).

Plants overexpressing *E2Fb* have shorter primary roots, a higher density of thicker root hairs, shorter hypocotyls, a lack of trichomes and slightly larger cotyledons with smaller epidermal cells but more total cells (Sozzani et al., 2006). These phenotypes are very similar to those seen in plants overexpressing *E2Fa*. Plants overexpressing *E2Fb* in conjunction with *DPa* have longer mitotic activities, shorter cell doubling times and inhibited cell growth (Magyar et al., 2005). The lack of cell growth suggests a role for E2Fb not only in the $G_1$/S transition but also in the $G_2$/M

phase of the cell cycle (Magyar et al., 2005). This is further supported by $G_2$/M specific gene upregulation in *E2Fb*-overexpressing plants (Sozzani et al., 2006).

E2Fb, unlike the other two E2Fs, is stabilized by auxin (Magyar et al., 2005). Removal of auxin from growth medium destabilizes the protein whereas addition of it increases E2Fb protein stability (Magyar et al., 2005). Plants with higher *E2Fb* levels also display higher levels of the *AUXIN RESISTANT 1* (*AXR1*) gene providing a link with auxin. Overexpressing *E2Fb* removes the auxin sensitivity so that cells in Bright Yellow-2 (BY-2) cultures continue dividing at normal rates also in auxin starved conditions (Magyar et al., 2005).

#### 6.1.2.3 E2Fc

*E2Fc* is also expressed in a cell cycle dependent way and transcripts peak at the $G_1$/S transition (de Jager et al., 2001, Mariconti et al., 2002). It is strongly expressed in cotyledons and in the shoot and root meristems (del Pozo et al., 2002). In flowers, *E2Fc* is present in the young, developing buds and then in pollen grains and vascular tissues (del Pozo et al., 2002). It is found in both dividing and in differentiated cells (del Pozo et al., 2002). Reducing *E2Fc* levels only changes gene expression levels in mature leaves and not in young ones, suggesting E2Fc functions by repressing E2F-regulated genes in mature differentiated cells (Magyar, 2008).

The activity of E2Fc differs from the other E2Fs in several ways and it has mainly been characterized as a promoter of endoreduplication as opposed to cell division. This is due to several reports showing that plants overexpressing *E2Fc* and its preferential binding partner *DPb* have increased endoreduplication (del Pozo et al., 2006), that knockout mutants have less endoreduplication coupled with more cell division (del Pozo et al., 2006) and that when overexpressing *E2Fc* and *DPa* 60% of the genes affected are downregulated (de Jager et al., 2009). Overexpressors and RNAi lines of *E2Fc* have a dwarf phenotype but for different reasons. Overexpressing the transcription factor leads to a delay in cell division while RNAi lines display many more yet smaller cells than wildtype plants (del Pozo et al., 2006).

None of the genes downregulated in *E2Fc/DPa*-overexpressing lines contain E2F-sites, suggesting an added layer of complexity in *E2F* gene regulation. While E2Fs are mostly considered as activators of $G_1$/S transition, the fact that overexpressing lines activate and repress different sets of genes suggests a complicated yet tightly controlled network of E2F function.

### *6.1.2.4* E2Fs *and programmed cell death*

Few papers have addressed the role of E2Fs in other processes than the cell cycle such as, for example in plant PCD. However, a link to PCD was made through pathogen responses (Wang et al., 2014). Here, upon pathogen recognition, RBR is hyperphosphorylated leading to over-activation of E2Fs and subsequent cell death (Wang et al., 2014). In accordance with this, the triple *e2f* mutant is defective in effector-triggered immunity (ETI) and more susceptible to virulent pathogens, providing a link between the *E2Fs* and plant immunity through the cell death pathway (Wang et al., 2014). Furthermore, a link to DNA damage was briefly mentioned in regard to *E2Fc* and *CYCB1;1*, the latter of which is downregulated in plants overexpressing *E2Fc* (Del Pozo et al., 2007). This reduction is overcome when DNA damage is present (Del Pozo et al., 2007), showing that DNA damage can induce a different set of genes, distinct from the response of healthy cells.

More is known about the role of the E2Fs in PCD in mammalian systems. There are eight *E2Fs* in humans and *E2Fa* and *E2Fb* equivalents are *E2F1- E2F5* while *E2Fc*'s orthologue is *E2F6* (de Jager et al., 2009). The consensus binding site is also conserved and is TTTCGCGCG (Zheng et al., 1999). Moreover, the domain structure is conserved (Magyar et al., 2005) as are the *DPs* (Wu et al., 1995).

*E2F*-mediated upregulation of pro-apoptotic genes has been shown in animal systems with *E2F1* having the highest apoptotic activity (Nahle et al., 2002, Pediconi et al., 2003, Wu et al., 2009). This is linked to DNA damage as this leads to E2F1 stabilization due to phosphorylation and acetylation (Lin et al., 2001, Pediconi et al., 2003). The E2F acetylated form then activates pro-apoptotic genes as opposed to cell-cycle ones (Pediconi et al., 2003). Caspases-3 and -7 are upregulated in the $G_1$/S transition by *E2F1-3* (Nahle et al., 2002). Caspases are cysteine proteases which cleave target proteins required for PCD. Their accumulation itself does not have any effects as these caspases need to be activated by post-translational modifications (Boatright and Salvesen, 2003). Therefore, cells have a strong protection mechanism in which caspases and other proteins required for apoptosis accumulate and can be rapidly activated in case of DNA damage recognition.

This chapter investigates the role of TSN in embryo development in *Arabidopsis*. Furthermore, a link is made to the E2F family of cell-cycle regulators. Data are presented demonstrating that the *TSN* and *E2F* pathways overlap and mutants in both gene families are defective in fertility due to embryo abortion.

## 6.2 Materials and Methods

### *6.2.1 Plant materials and growth conditions*

T-DNA insertion lines for *E2Fa-3* (SALK_104099), *E2Fb-1* (SALK_103138C), *E2Fc-1* (SALK_073459), *TSN1* (SALK_045179), *TSN2* (SALK_062222C), and *DPb* (SALK_066506c) were acquired from the European Arabidopsis Stock Centre (NASC) (Arabidopsis.info). Additionally, *e2fb-2* (SALK_120959), *e2fa-1* (MPIZ_244), *e2fa-2* (GABI-248E09), *e2f-2$^{-/-}$e2fb-1$^{-/-}$ HA-E2Fb/DPa* overexpressing lines and *HA-E2Fb$^{\Delta RBR}$/DPa* overexpressing lines were sent to me by Zoltan Magyar. Lines overexpressing *E2Fc* and one overexpressing *DPb* were received from Crisanto Gutierrez. Lines were selected by PCR and homozygotes identified. Higher order mutants were made by crossing and selection by PCR. Plants were grown in controlled environment rooms with 16 hours light and constant 22°C. Plants were randomized after genotyping using the Experimental Design Generator and Randomiser (EDGAR) (www.edgarweb.org.uk).

### *6.2.2 Sampling, measuring and statistics*

Stage 17 siliques were sampled from plants when they were mature. The longest siliques were chosen to get a measure of maximum silique length. Ten siliques were removed from the plant and photographed. Lengths were then measured using ImageJ 1.48v (Schneider et al., 2012). At the same time, plant height was photographed and measured using ImageJ. Seeds per silique were determined after clearing siliques using 100% Ethanol. Mature seeds were harvested and imaged using a light microscope. Images were analysed using an ImageJ macro I wrote (See Appendix). Values were normalised using the Col-0 wildtype values of that experiment. Plant height means were used for normalisation and statistics. To determine when seeds got aborted, pre-emasculated gynoecia were fertilized and analysed once a day for aborted seeds. Statistics were carried out using GenStat (VSN International, 2011) General Linear Models (GLM).

### *6.2.3 Pollen tube staining*

Aniline blue staining of pollen tubes was carried out as in Jiang et al. (2005). Flowers of Col-0 and *tsn1$^{-/-}$ tsn2$^{-/-}$* were emasculated and pollinated 24 hours later. Pistils were collected 2, 6 and 24 hours after pollination (hap) (as in (Crawford and Yanofsky, 2011)) and fixed in ethanol:acetic acid (3:1) for two hours at room temperature. Three washes with distilled water followed. Subsequently, pistils were softened in a solution of 8M NaOH overnight after which they were

again washed three times in distilled water. Pistils were then stained in aniline blue solution (0.1 % aniline blue in 0.1M $K_2HPO_4$-KOH buffer, pH 11) for 3 – 5 hours in the dark. Stained pistils were imaged using a Leica DM6000 microscope.

### *6.2.4 Flow Cytometry*

Approximately 0.5cm$^2$ of leaf tissue or one stage 17 silique was sampled and finely chopped in 1mL of CyStain® UV Ploidy extraction buffer. The solution was then filtered into the sample tube to which 1mL of staining buffer was added. Samples were analysed with BD FACSMelody™ cell sorter.

### *6.2.5 Cloning of a* pE2Fb:GUS *reporter*

The *pE2Fb:GUS* reporter was constructed by amplifying the *E2Fb* promoter (960 bp) by PCR. It was then cloned into L0 pICH41295 using Golden Gate. This was then cloned into the L1 pICH47742 and L2 pICSL4723 with a GUS (L0 pICH75111) and NOS (L0 pICH41421) cassette. All constructs were received from Synbio TSL (http://synbio.tsl.ac.uk/).

### *6.2.6 Yeast Transformation*

Yeast strain AH109 was used for all yeast transformations. Construct pGAD424 (which includes the GAL4 AD domain) and pGBT9 (with the GAL4 BD domain) were transformed into AH109. To do this, yeast colonies were grown up for two days in 20mL YPDA (400mL water, 8g peptide, 8g glucose (2%), 4g yeast extract, 6mL of 0.2% adenine hemisulphate) at 28°C shaking at 180rpm. The cultures were then spun down at 4000rpm for 10min and the supernatant was poured off. The remaining pellet was washed in 10ml sterile water and spun again for 10min at 4000rpm. Next, the pellet was resuspended in 1.5ml LiTE (0.1M Lithium Acetate and 1xTE Buffer (0.01M Tris HCl and 1mM EDTA), pH 7.5) before being spun for 2minutes at 5000rpm. After removing the supernatant, the pelleted cells were resuspended in 1.2mL LiTE. For the transformation 1µL plasmid DNA was pipetted into an Eppendorf tube. 50µL yeast was added followed by 8µL Salmon or Herring sperm DNA, 10µL DMSO and 300µL PEG/LiTE (50% Polythylene Glycol and 1x LiTE solution). The solution was gently mixed without inverting the tube and incubated at 28°C for 25min. Next, samples were heat-shocked at 42°C for 15min before being centrifuged for 2min at 5000rpm. The supernatant was carefully removed before the pellet was resuspended in 100µL

sterile water. All of the 100µL were spread out on a selection plate of YSD-W-L (0.616g of YSD-W-L, 10.75g SD broth (2% glucose), 8g agar in 400ml water) and incubated at 28°C for 3 – 4 days.

### *6.2.7 Yeast Selection Plates*

Plates of YSD-W-L-H-A were prepared (0.556g YSD-W-L-H-A powder, 10.75g SD broth (2% glucose), 8g agar in 400ml water, pH 5.8) and four plates for each interaction set up with no 3AT, 2.5mM 3AT, 5mM 3AT and 10mM 3AT. Three yeast colonies were added to 100µL water and three drops of 10µL were placed into each plate. This was done with three biological replicates. Plates were placed at 28°C for two weeks and checked daily.

### *6.2.8 RNA extraction and cDNA synthesis*

Inflorescences with gynoecia up to stage 12 were collected in liquid nitrogen and ground using a pistil and mortar. RNA was extracted using the RNAeasy® Plant Mini Kit from Qiagen using the manufacturer's protocol. cDNA was synthesized using 5µL RNA, 1µl oligo dT, 1µL dNTPs and 5µL $H_2O$. This was kept at 65°C for 5min and tubes were then cooled on ice. 4µL 5x Buffer and 2µL 0.1M DTT were added. This was then kept at 37°C for 2min. 1µL M-MLVRT was added and mixed using the pipette. The mixture of 19µL was then incubated at 37°C for 50min and subsequently kept at 70°C for 15min. RNA was stored at -20°C until needed.

### *6.2.9 Quantitative Real-Time PCR analysis (RT-qPCR)*

All RNA samples were diluted to be the same strength (100-150ng/µL). Gene expression was analysed using the BioRad CFX96 machine using SYBR Green JumpStart Taq Ready Mix PCR Master Mix (Sigma-Aldrich). All primers were tested using a dilution series and by calculating the primer efficiency. Primers used can be found in Appendix Table 2.

## 6.3 Results

### *6.3.1* tsn *double mutants have a fruit phenotype*

To investigate the role of *TSN* in fruit development, *tsn* loss-of-function mutants were obtained. Homozygous single mutants did not show a fruit phenotype, which is in accordance with the

literature (Sundstrom et al., 2009). In contrast, double heterozygous and double homozygous lines both exhibited significant growth defects. Although double mutant plants had no obvious defects during vegetative development, both heterozygous and homozygous double mutants produced shorter fruits with fewer seeds (Fig. 6.4). *tsn1*$^{-/-}$ *tsn2*$^{-/-}$ siliques were roughly 60% as long as Col-0 fruits with a quarter of the seed set. As silique length correlated with seeds per silique, the former was used to infer about the latter in the rest of the experiments. These results point to a role of *TSN* in fertility.

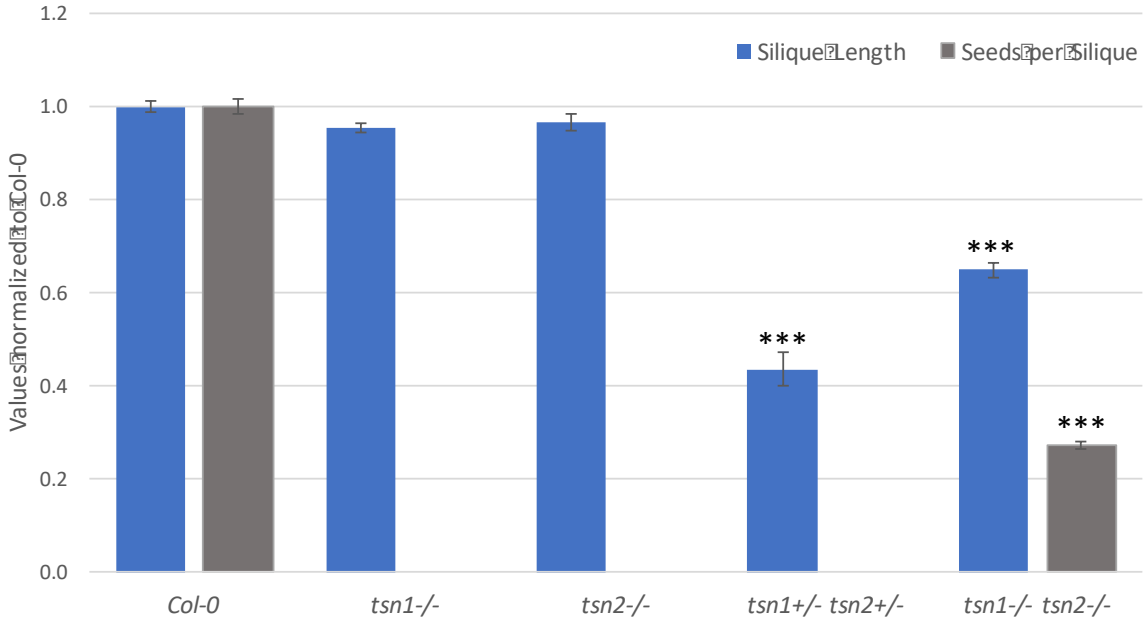



Fig. 6.4. *Arabidopsis tsn* mutants had fruit phenotypes. Single mutants did not show a reduction in silique length or seeds per silique. The double heterozygous and double homozygous lines were shorter (blue bars) and had fewer seeds (grey bars). The missing seeds per silique is due to the seeds not being counted. The plants did produce seed. All data is normalized to Col-0. Error bars are StErrors. Statistics were carried out using General Linear Regression Models in GenStat. *** p<0.001 compared to Col-0.

### *6.3.2* e2f *mutants mostly do not show a fruit phenotype*

Next, I wanted to investigate the role of the E2F family of transcription factors in reproduction. To this end, I analysed *e2f* TDNA-mutants. The two *e2fb* and one *e2fc* mutants did not have a fruit phenotype (Fig. 6.5). While *e2fa-1* and *e2fa-3* did not have a difference in silique length, *e2fa-2* was significantly shorter with a silique length of 93% that of wildtype (p=0.019) (Fig. 6.5). Two *e2fa*$^{-/-}$*e2fb*$^{-/-}$ combinations were tested. *e2fa-2*$^{-/-}$*e2fb-1*$^{-/-}$ displayed a reduction of 40% (p<0.001) while *e2fa-3*$^{+/-}$*e2fb-1*$^{+/-}$ showed a reduction of 10% (p=0.01). All subsequent crosses were made

using *e2fa-3* and *e2fb-1* as they were the first lines acquired. Single, the other double or triple mutants in *E2Fs* did not have a fruit phenotype (Fig. 6.6) or vegetative defects.

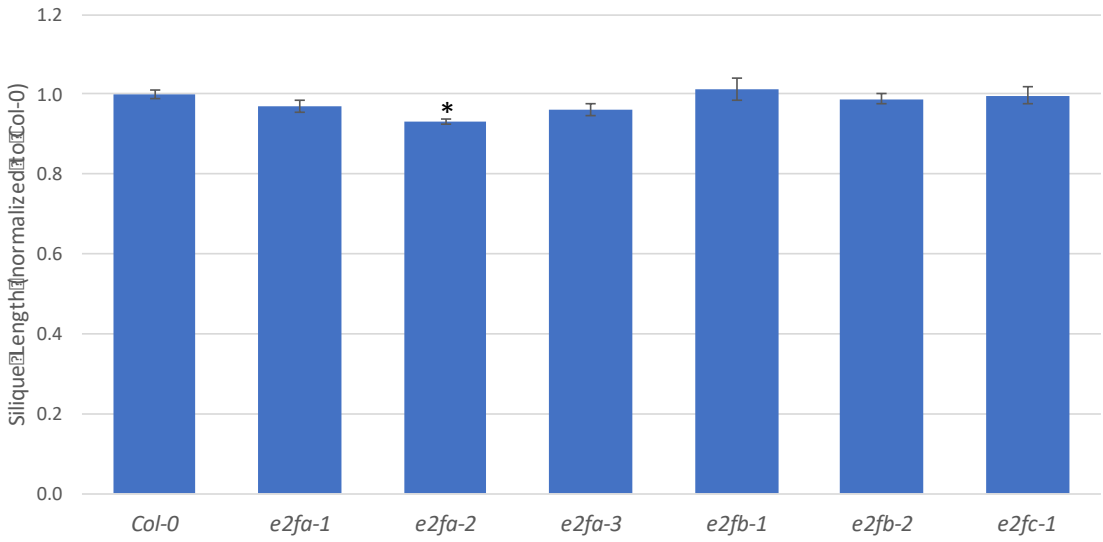



Fig. 6.5. Single mutants did not have a fruit phenotype apart from *e2fa-2* which had significantly reduced silique lengths (93% of wild type lengths). Error bars are standard errors. * p<0.05

*E2Fs* have been overexpressed and published. Plants overexpressing E2Fb/DPa, E2Fc and DPb did not show a difference in silique length or in plant development. HA-E2Fb$^{\Delta RBR}$/DPa however displayed significantly reduced siliques that were 8% shorter of wild type (p=0.019) (data not shown, in the appendix). This is not a very strong phenotype, however.

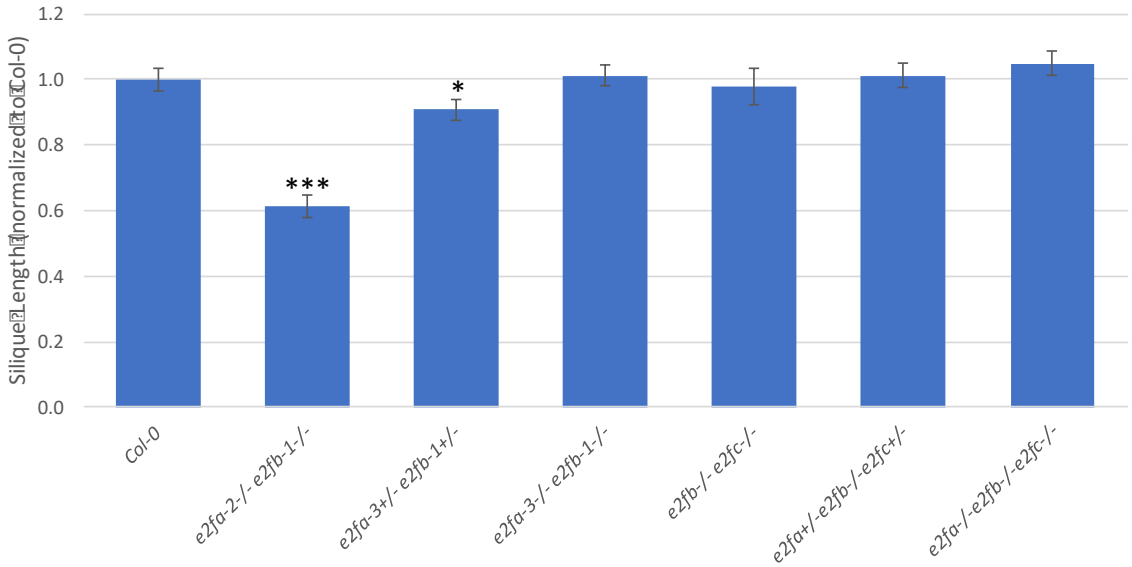



Fig. 6.6. *e2f* double mutants did not show strong fruit phenotypes apart from e2fa-2-/- e2fb-1-/- where siliques were reduced by 40%. Error bars are standard deviations. Silique lengths are normalized to Col-0. * p<0.05, ***p<0.001

### *6.3.3 Plants harbouring mutations in* E2F *and* TSN *genes are affected in fruit length and seed set*

To elucidate whether *TSN* and *E2F* function together in fruit development, crosses were carried out the test for genetic interactions. Double heterozygous mutants involving one *tsn* and one *e2f* mutant all showed significantly reduced fruit lengths (Fig. 6.7). Plants heterozygous for any *e2f* and *tsn1* had siliques 60% as long as wildtype. This was the same reduction in length as *tsn1*$^{-/-}$ *tsn2*$^{-/-}$ plants. Crosses with *tsn2* were not as strongly affected with silique lengths of 80% compared to wild type. Fruit length reductions were consistent for all *e2f* mutants; none of them appeared to be dominant over the others. This result suggests that *TSN* and *E2F* function together in fruit development.

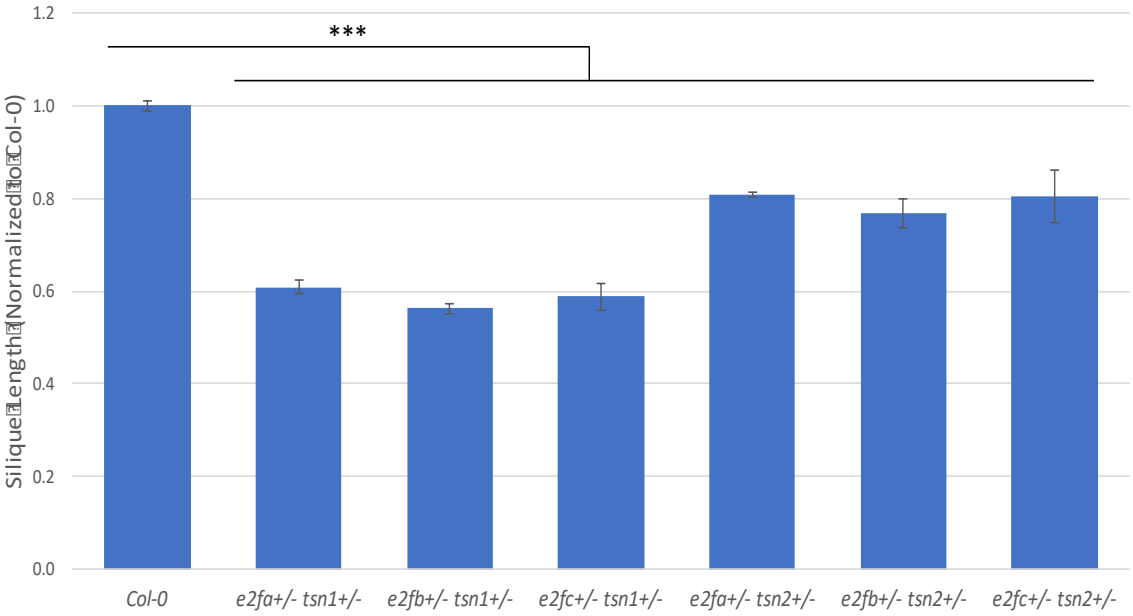



Fig. 6.7. *e2f tsn* double heterozygous plants had shorter pods. Pods in *tsn1*$^{+/-}$ were roughly 60% as long as Col-0 pods. *tsn2*$^{+/-}$ plants were less severely affected with pod lengths reduced by only 20%. All lines were significantly shorter than Col-0. Error bars are standard errors. Regression analysis was carried out using GenStat. ***p<0.001

Intriguingly, plants homozygous for *e2f* mutations had a fully rescued phenotype (Fig. 6.8). To determine whether one gene was more influential in this than the other, I grew the four different possible zygocities of double mutants: het het, het hom, hom het and hom hom e2f tsn mutants. This showed that the phenotype was not dependent on the TSN as plants with heterozygous and homozygous tsn mutations had the silique length phenotype. However, only plants with a

heterozygous E2F showed the phenotype. The e2f homozygous fruits had a full seed set and were indistinguishable from Col-0.

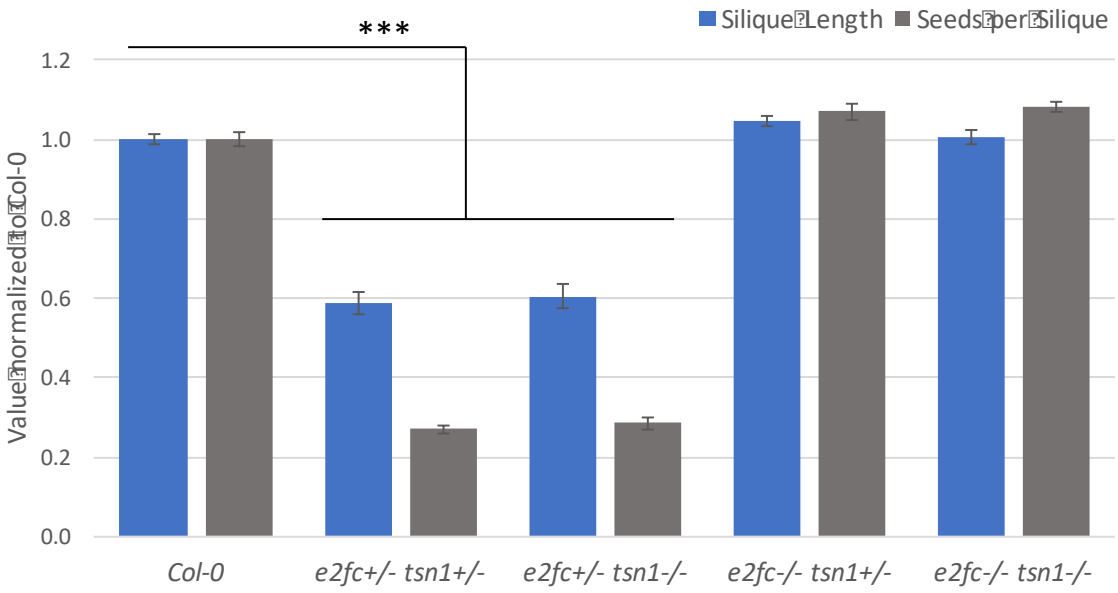



Fig. 6.8. The *E2F* zygocity was causal. If it was heterozygous, plants had shorter pods and fewer seeds but if the *E2F* was homozygous for the mutation the phenotype was rescued. This effect was consistent for all *e2f/tsn* crosses (data in Appendix Table 3). Error bars are standard errors. Statistics were performed using GenStat. *** $p<0.001$

Plants mutated in more than two genes did not have an exaggerated phenotype. Even quintuple heterozygous plants produced fruits that were 80% in length compared to wild type (Table 6.1). In fact, the genotype with the shortest siliques (*e2fb*$^{+/-}$ *e2fc*$^{+/-}$ *tsn1*$^{+/-}$) was 55.9 % compared to wild type fruits and had less than a quarter of the seeds per silique (Table 6.1). The fact that the phenotypes did not get worse in higher order mutants suggests redundancy between the genes. The effect seen in double mutants was still present in higher order ones with lines where all *E2Fs* in the mutant were heterozygous displaying short fruits and those with all *E2Fs* homozygous for the mutation had fully filled siliques. One line did not follow this trend: *e2fa*$^{+/-}$*e2fc*$^{+/-}$*tsn1*$^{-/-}$tsn2$^{-/-}$ had long fruits and with a full seed set. However, it should be noted that this is based on the analysis of just one plant. It was also the only line where both *TSN* were homozygous for the mutation but the *E2Fs* were not. It was less simple to find the trend in mutants that had multiple *E2Fs* mutated and at different zygocyties. Seven lines with at least one *E2F* homozygous for the mutation had a full seed set and five lines had fewer seeds (Table 6.1).

*DPa* and *DPb* have been less well characterized than the *E2Fs*. As DPb is the preferential binding partner of E2Fc, double mutants were generated. Neither the double heterozygote nor the double homozygote showed a reduction in silique length or seed number (Fig. 6.9). Similarly, in combination with the *tsn1* mutants as in *dpb*$^{+/-}$*tsn1*$^{+/-}$ no phenotype was observed. However, *dpb*$^{+/-}$*tsn2*$^{+/-}$ produced siliques that were 20% shorter than wild type based on data from two plants. These plants also had half as many seeds as wild-type plants. This suggests a difference in the importance of *TSN1* and *TSN2*. Interestingly the triple *dpb*$^{+/-}$*e2fc*$^{+/-}$*tsn1*$^{+/-}$ mutant did not show a fruit phenotype (Fig. 6.9). This suggests that reduction in *DPb* rescued the phenotype of *e2fc*$^{+/-}$ *tsn1*$^{+/-}$. It would be interesting to see whether this also happens in the combination with *TSN2*. The *dpb e2fa-c* quadruple heterozygous mutant was different to wild type and to the *e2fa-c* mutants with a 15% reduction in seeds and a 7% reduction in silique length (Fig. 69). These phenotypes were not as severe as any of the *e2f/tsn* mutants tested, however.

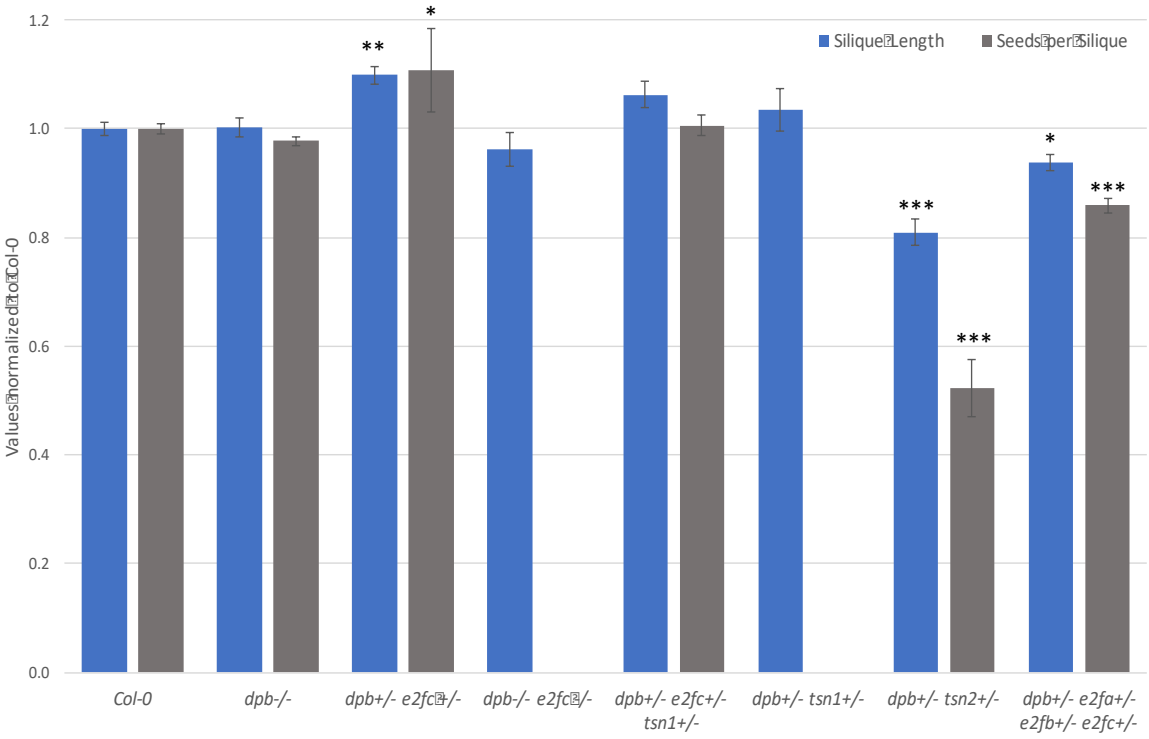



Fig. 6.9. *dpb* single mutants did not have a phenotype. Plants heterozygous for *dpb* and *e2f* had significantly longer siliques with more seeds. A strong reduction in silique length and seed filling was found in conjunction with *tsn2*, but not *tsn1*. A slight but significant phenotype was present in *dpb e2fa-c* quadruple heterozygous lines. *p<0.05; **p<0.01; ***p<0.001.

Table 6.1. Higher order *e2f/tsn* mutants showed a similar trend to double mutants. Lines with all *e2fs* heterozygous showed a significant reduction in pod length and seeds per silique, except for *e2fb*$^{+/-}$ *e2fc*$^{+/-}$ *tsn1*$^{+/-}$ *tsn2*$^{+/-}$ which did not display a fruit phenotype. No significant phenotypes were found in lines where all *e2fs* were homozygous for the mutation, apart from *e2fb*$^{-/-}$*e2fc*$^{-/-}$ *tsn2*$^{-/-}$ where there was a significant reduction in silique length but not in seeds per silique. All values are normalized to Col-0. Ten pods per plant were measured for the pod length. Stars denote the significance level difference to Col-0 (*<0.05; **<0.005; ***<0.001).

| Genotype | Mean Silique Length Norm. | Seeds per Silique Norm. | Plants analysed |
|---|---|---|---|
| Col-0 | 1.000 | 1.000 | 10 |
| *e2fa*$^{+/-}$ *e2fb*$^{+/-}$ *tsn1*$^{+/-}$ | 0.769*** | 0.529*** | 7 |
| *e2fa*$^{+/-}$ *e2fb*$^{-/-}$ *tsn1*$^{+/-}$ | 0.899* | 0.964 | 7 |
| *e2fa*$^{-/-}$ *e2fb*$^{-/-}$ *tsn1*$^{-/-}$ | 1.029 | 0.978 | 10 |
| *e2fb*$^{+/-}$ *e2fc*$^{+/-}$ *tsn1*$^{+/-}$ | 0.559*** | 0.239*** | 3 |
| *e2fb*$^{-/-}$ *e2fc*$^{+/-}$ *tsn1*$^{+/-}$ | 0.984 | 0.918 | 3 |
| *e2fb*$^{-/-}$ *e2fc*$^{-/-}$ *tsn1*$^{-/-}$ | 1.028 | 0.963 | 10 |
| *e2fa*$^{+/-}$ *e2fb*$^{+/-}$ *tsn2*$^{+/-}$ | 0.698*** | 0.347*** | 1 |
| *e2fb*$^{+/-}$ *e2fc*$^{+/-}$ *tsn2*$^{+/-}$ | 0.761*** | 0.336*** | 9 |
| *e2fb*$^{+/-}$ *e2fc*$^{-/-}$ *tsn2*$^{+/-}$ | 0.734*** | | 2 |
| *e2fb*$^{-/-}$ *e2fc*$^{-/-}$ *tsn2*$^{-/-}$ | 0.889* | 0.971 | 2 |
| *e2fa*$^{+/-}$ *tsn1*$^{-/-}$ *tsn2*$^{+/-}$ | 0.763** | 0.536*** | 2 |
| *e2fa*$^{-/-}$ *tsn1*$^{-/-}$ *tsn2*$^{-/-}$ | 0.979 | 0.969 | 9 |
| *e2fb*$^{-/-}$ *tsn1*$^{-/-}$ *tsn2*$^{+/-}$ | 1.078 | | 2 |
| *e2fc*$^{-/-}$ *tsn1*$^{-/-}$ *tsn2*$^{-/-}$ | 0.969 | 0.999 | 10 |
| *e2fa*$^{+/-}$ *e2fb*$^{+/-}$ *tsn1*$^{+/-}$ *tsn2*$^{+/-}$ | 0.718** | 0.355*** | 1 |
| *e2fa*$^{-/-}$ *e2fb*$^{-/-}$ *tsn1*$^{+/-}$ *tsn2*$^{-/-}$ | 0.916 | | 1 |
| *e2fb*$^{+/-}$ *e2fc*$^{+/-}$ *tsn1*$^{+/-}$ *tsn2*$^{+/-}$ | 0.682*** | 0.351*** | 4 |
| *e2fa*$^{+/-}$ *e2fc*$^{+/-}$ *tsn1*$^{-/-}$ *tsn2*$^{+/-}$ | 0.806*** | | 1 |
| *e2fa*$^{+/-}$ *e2fc*$^{+/-}$ *tsn1*$^{-/-}$ *tsn2*$^{-/-}$ | 0.949 | 0.949 | 1 |
| *e2fa*$^{+/-}$ *e2fb*$^{+/-}$ *e2fc*$^{+/-}$ *tsn1*$^{+/-}$ *tsn2*$^{+/-}$ | 0.790*** | 0.373*** | 9 |
| *e2fa*$^{+/-}$ *e2fb*$^{-/-}$ *e2fc*$^{+/-}$ *tsn1*$^{-/-}$ *tsn2*$^{+/-}$ | 0.689*** | | 1 |
| *e2fa*$^{+/-}$ *e2fb*$^{-/-}$ *e2fc*$^{+/-}$ *tsn1*$^{+/-}$ *tsn2*$^{+/-}$ | 1.024 | 0.979 | 5 |
| *e2fa*$^{+/-}$ *e2fb*$^{+/-}$ *e2fc*$^{+/-}$ *tsn1*$^{-/-}$ *tsn2*$^{+/-}$ | 0.726** | 0.409*** | 3 |

### *6.3.4 Seed size and seed number correlate negatively*

A negative correlation between seed size and seed number is often found in plants. This was shown in this work in *B. napus* accessions but has also been reported in *Arabidopsis* mutants

(Sornay et al., 2015). Seed sizes in short-silique *e2f tsn* mutants were significantly larger than those of wild type or plants with homozygous *e2f* mutations (Table 6.2).

The consistent results found for all *E2Fs* and in all other mutants suggests a functional overlap between the *E2F* and *TSN* pathways in fruit development.

Table 6.2. Seed size followed the opposite trend than seed number; plants with fewer seeds per silique had significantly larger seeds (*e2fc$^{+/-}$ tsn1$^{+/-}$ and e2fc$^{+/-}$ tsn1$^{-/-}$*). Seed sizes were determined using a Macro in ImageJ used on pictures taken with a light microscope. *** p<0.001

| Line | Mean Seed Surface Area (mm$^2$) | Number of Seeds |
|---|---|---|
| Col-0 | 0.171 | 416 |
| *e2fc+/- tsn1+/-* | 0.210*** | 406 |
| *e2fc+/- tsn1-/-* | 0.231*** | 373 |
| *e2fc-/- tsn1+/-* | 0.167 | 257 |
| *e2fc-/- tsn1-/-* | 0.171 | 315 |

### *6.3.5 Defects have a maternal origin*

As silique length is the final visible step in a long developmental pathway, I wanted to narrow down when development was affected. To this end, I asked whether development was affected before fertilization and measured ovary sizes. A difference in ovary length could suggest that ovules do not develop normally and are aborted early. No differences in ovary length were found (Fig. 6.10a). Next, I tested whether the defects had a maternal origin by making reciprocal crosses between Col-0 and the mutants *e2fc$^{+/-}$tsn1$^{-/-}$*, *e2fb$^{+/-}$tsn1$^{+/-}$* and *tsn1$^{-/-}$tsn2$^{-/-}$.* All three sets showed reduced silique lengths when the mutation was of maternal origin whereas using mutant pollen to fertilize Col-0 gave rise to full length fruits (Fig. 6.10b for *e2fc/tsn1*, others in Appendix). This result suggested that the defects were of maternal origin and that the pollen was not negatively affected.

Maternal defects mainly lead to non-viable homozygous plants since all ovules containing the mutation abort. Therefore, the nature of defects can often be determined by the segregation ratio which for a single maternal-effect mutant should be 2:1 where two plants are heterozygous and one is wildtype. This diverges from the Mendelian segregation ratio of 1:2:1 since the homozygous mutation is lethal. To see whether the segregation ratio was disturbed in the

mutants, 357 progeny plants of an *e2fb*$^{+/-}$*tsn1*$^{+/-}$ parent were genotyped by Pauline Stephenson. All possible mutant combinations were present and the segregation ratio differed to that of the Mendelian pattern (Fig. 6.11). However, the segregation did not suggest why seeds were aborted. Throughout the study and for all mutant combinations tested, all genotypes were present. As the phenotype is maternal, a different experiment could have perhaps been more informative. Crossing an *e2f tsn* double heterozygous plant with Col-0 pollen and genotyping the subsequent generation would give a better insight into the maternal genotypes passed on.

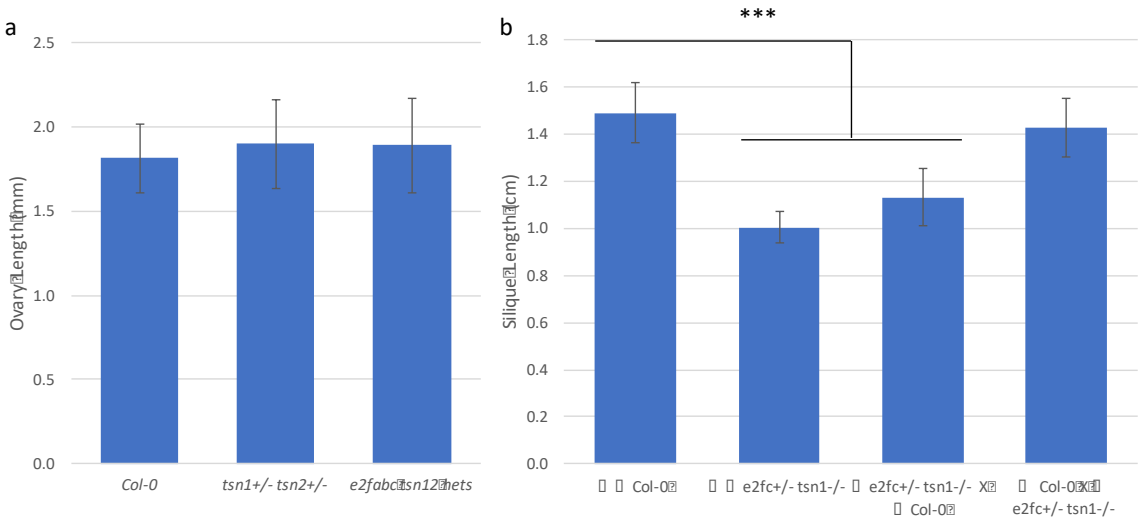



Fig. 6.10. Ovary sizes did not differ between mutant plants (a). The defects were of maternal origin (b). Crosses using Col-0 pollen on e2fc$^{+/-}$ tsn$^{-/-}$ gynoecia gave rise to shorter fruits whereas using the mutant pollen on the wild type did not show a fruit length reduction. (b) Error bars are standard deviations. *** p<0.001

| | | |
|---|---|---|
| 6% Wt Wt 0% | 6% Wt Hom 5% | 13% Hom Het 7% |
| 12% Wt Het 3% | 6% Hom Wt 8% | 13% Het Hom 23% |
| 13% Het Wt 8% | 25% Het Het 43% | 6% Hom Hom 3% |

Fig. 6.11. The segregation ratio of seeds from an *e2fb*$^{+/-}$ *tsn*$^{+/-}$ parent plant did not follow the Mendelian segregation ratio. All genotypes were present demonstrating divergence from the usual ratio found in maternal mutations where no homozygous plants are retrieved. 357 plants were genotyped. The large number is the one found by genotyping, the smaller green one is the expected percentage if a Mendelian segregation ratio had been present.

### *6.3.6 Defects occur post fertilization*

To rule out that the defects were paternal, fully developed siliques were cleared to show where seeds were located. Plants mutated in pollen development often have shorter fruit with seeds only found in the apical part of the silique due to pollen germination defects (Crawford and Yanofsky, 2011, Jiang et al., 2005). Seeds were found located throughout the fruit with gaps over the entire length from where seeds had been aborted (Fig. 6.12a). To eliminate pollen defects as a cause, pollen tubes were stained 2, 6 and 24 hours after pollination (hap) (Fig. 6.12b). The results showed a similar density in pollen tube growth between *tsn1*$^{-/-}$ *tsn2*$^{-/-}$ and wild type and clearly demonstrated that the *tsn1*$^{-/-}$ *tsn2*$^{-/-}$ pollen tubes could reach the end of the gynoecium to fertilize ovules all along the fruit. However, fewer pollen were found on the mutant plants (not measured). Combined, these results allowed us to conclude that the *tsn* and *e2f* mutant fruit were not defective in fertilization.


Fig. 6.12. Pollen viability was not affected in *tsn e2f* mutants. (a) Mutant siliques had seeds distributed randomly along the fruit (b) Pollen germination was not negatively affected. Pollen tube density was not reduced and pollen tubes reached the entire length of the silique. Scale bar in (a) is 2mm

### *6.3.7 Seed loss occurs after fertilization*

Since the ovary sizes were the same in wild type and in the mutant and a full set of ovules was present before fertilization, I next tested when seeds were being aborted. At fertilization all ovules appeared the same size (Fig. 6.13a, d). Shortly thereafter however, some seeds appeared smaller than others (Fig. 6.13b, e). Later, the aborted seeds appeared as small white rounded shapes and were clearly distinguishable from developing seeds (Fig. 6.13c). This suggests abortions occurring right after pollination, possibly due to a problem in fertilization.

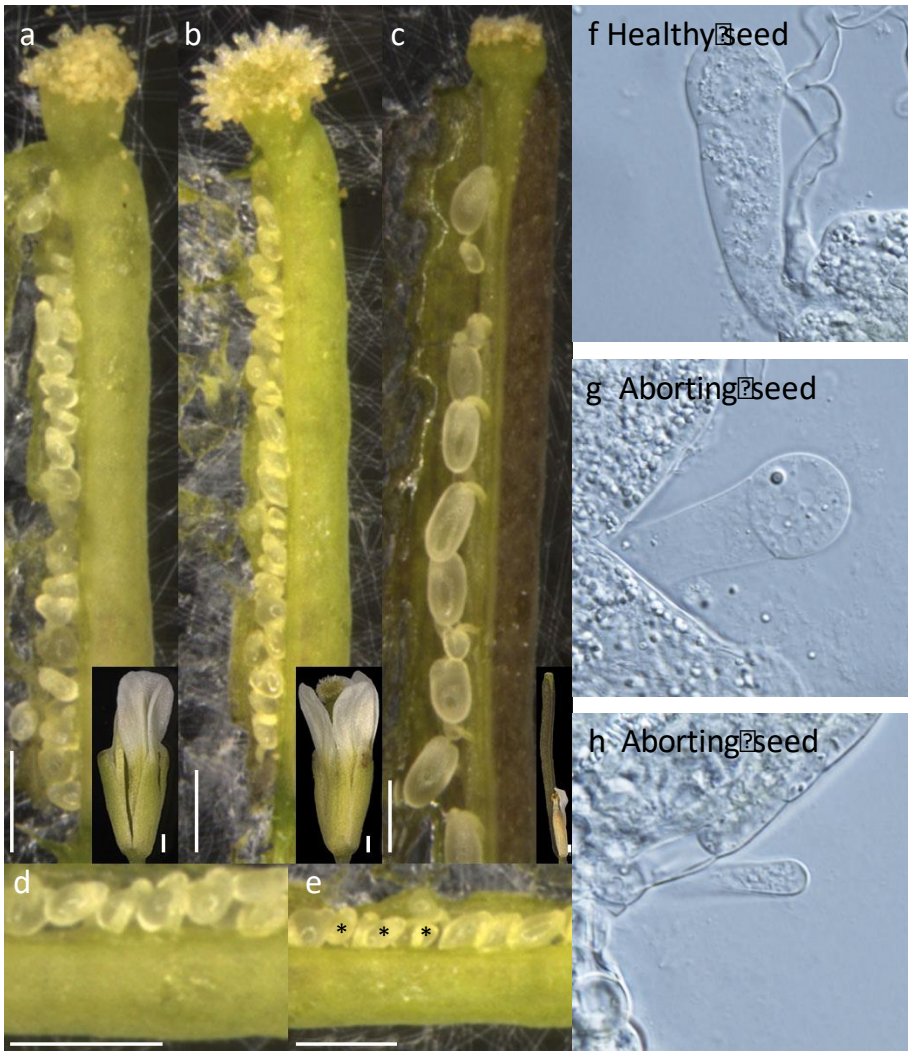



Fig. 6.13. Seeds were aborted shortly after fertilization. At fertilization (a and d), all ovules were the same size. Soon thereafter slightly smaller seeds were visible (b and e). The aborted seeds were clearly distinguishable at later time points (c). Embryos developed and were not different at the two-cell stage (f-h). a-e are *e2fb*$^{+/-}$*tsn2*$^{+/-}$, f-g *tsn1*$^{-/-}$ *tsn2*$^{-/-}$. Scale bars are 250µm. Stars highlight aborted seeds.

A second abortion of seeds occurred later on. At 96 hap, clearly identifiable small, aborting mutant seeds were present. At that time, the embryos of healthy seeds were at the early heart stage, suggesting that the abortions occurred before the heart stage of the embryo was reached. To ensure that fertilization of the egg cell was occurring, embryos of aborted and healthy seeds were imaged at early stages. Embryos were present in the aborting seeds and did not look different from healthy ones (Fig. 6.13 f-h). The fact that these seeds had embryos suggests that some seeds are aborted after successful fertilization of at least the egg cell. More images at different embryo developmental stages and of the endosperm will be required to comment on why seeds abort.

Taken together, these data suggest that *TSN* and the *E2Fs* play a role in seed development. Mutants showed maternal effects that were however not lethal. Pollen defects were not the cause of the reduction in seed set. The defects were visible post-fertilization, leading to fewer developed seeds and the shorter final silique length.

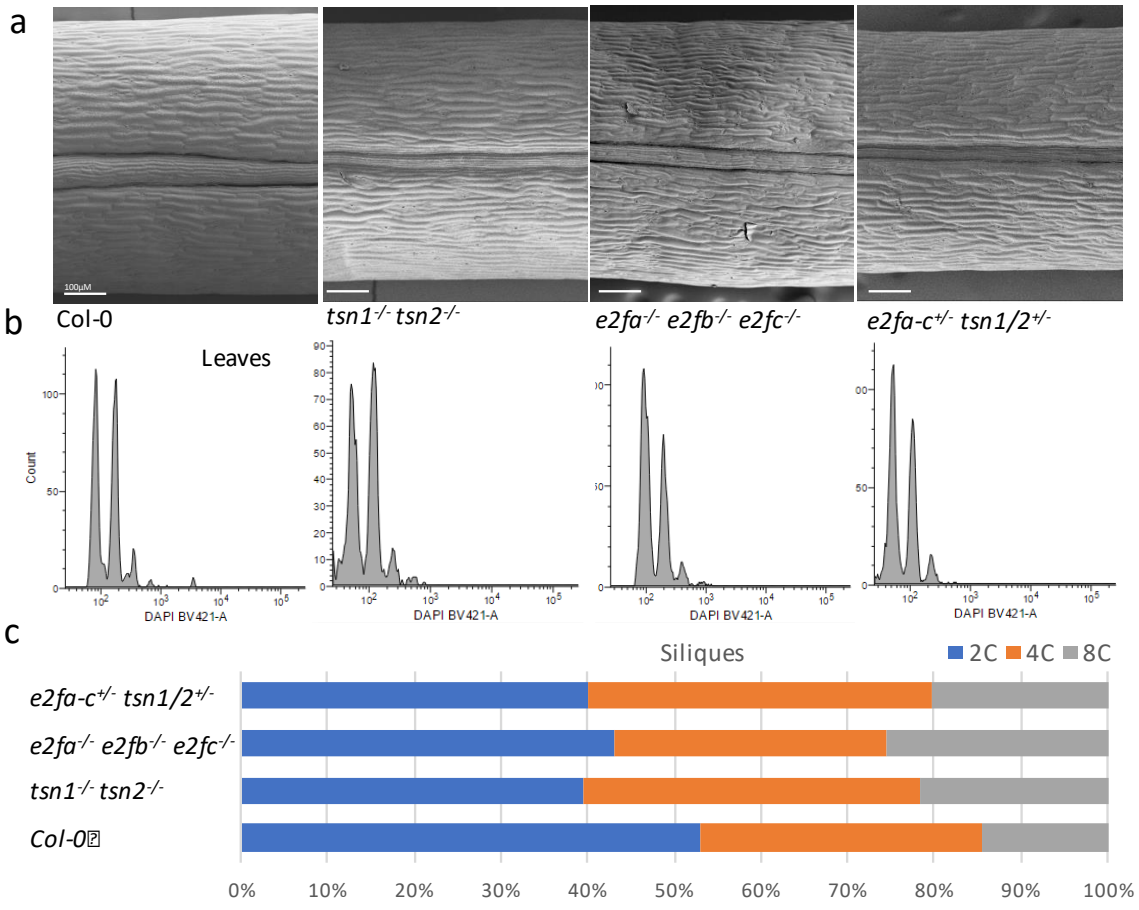



Fig. 6.14. Mutants did not show a different cellular arrangement at stage 17 (a). The ploidy of mutants was also not significantly affected in leaves (b) or in siliques (c). Three plants per genotype were used for flow cytometry and a total of 2000 counts per sample were taken.

### *6.3.8 No differences in ploidy are present*

*TSN* and *E2F* are involved in the cell cycle (de Jager et al., 2001, Elbarbary et al., 2017, Mariconti et al., 2002, Su et al., 2015). Cell cycle mutants often have a different ploidy level than wildtype and a different cellular patterning. This is often because cells are able to complete the $G_1$/S transition but are unable to enter mitosis afterwards, leading to larger cells with a higher DNA content (Castellano et al., 2001, Imai et al., 2006, Roeder et al., 2012). To determine whether this was the case in these mutants, SEM pictures of siliques were taken. Valve cells in mature wild-type siliques (stage 17) are narrow and elongated (Col-0 in Fig. 6.14), however, no difference to this morphology was observed in any of the mutants (Fig. 6.14). This suggested that the smaller silique length observed in the mutants was due to fewer cells rather than reduced cell length. Flow cytometry was used to determine ploidy. No differences were found in the leaves or the siliques with ploidies of 2C, 4C and 8C being present (Fig. 6.14).

### *6.3.9* E2F *and* TSN *are expressed at the same place in embryo development*

To further investigate how *E2F* and *TSN* genes interact during reproductive tissue development, I tested whether they were expressed at the same developmental time and place. GFP reporter lines for *TSN1* and *TSN2* were received from P. N. Moschou and *E2Fa* from Laszlo Bogre. A *pE2Fb:GUS* line was generated by myself. These lines were analysed and all genes were found to be ubiquitously expressed in the fruit (Fig. 6.15). Expression was seen in the silique wall, the seed coat and funiculus. *E2Fa*, *E2Fb* and *TSN2* were also expressed in the embryo and the suspensor until the heart stage, where defects had already occurred in mutated plants (Fig. 6.15). These results suggested that the interaction of the genes could be in the suspensor or the embryo, leading to the abortion of seeds. However, expression pre-fertilization was not looked at nor was the endosperm examined. *pTSN1:TSN1:GFP* plants did not show any GFP expression in the shoot (data not shown). The literature reports that *TSN1* and *TSN2* are co-expressed (Frey et al., 2010), however, since this line did not produce any signal it was not possible to assess here whether this is the case. At least the results presented here show that *E2Fa, E2Fb* and *TSN2* expression overlap, which is compatible with the hypothesis that the function of these gene families overlap.

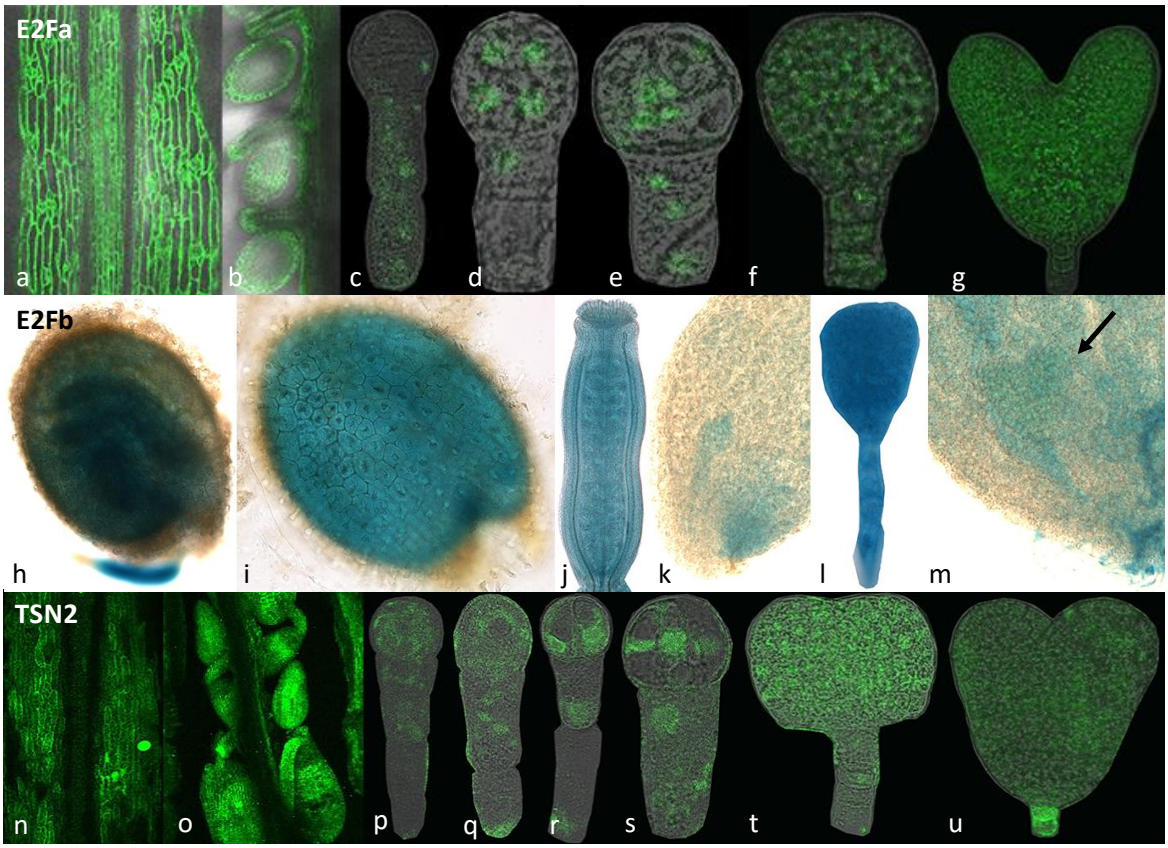



Fig. 6.15. *E2Fa*, *E2Fb* and *TSN2* were expressed in the fruit at the same place. *E2Fa*, *E2Fb* and *TSN2* were all present in the silique wall as well as the seed coat and the funiculus (a, b, h, i, j, n, o). *E2Fa* was expressed in the early developing embryo at the two-cell stage, the octant and the dermatogen stages in the nuclei mainly (c, d, e). In the globular and heart stages, it was expressed in the embryo proper (apical part) and in the suspensor (basal part) (f, g). *E2Fb* was also expressed in the entire embryo at the globular stage and heart stage (other stages were not seen) (k, l, m). *TSN2* was expressed in the early stages of embryo development from the one-cell to the globular stage in the embryo proper and in the suspensor (p, q, r, s, t). Stronger expression was seen in the suspensor than in the embryo proper at heart stage (u). The arrow in (m) shows the heart-stage embryo in the seed.

### *6.3.10 E2F and TSN proteins do not interact directly*

Next, to find out more about the nature of the interaction, a Yeast 2-Hybrid system was used to test whether the proteins interacted directly. As a positive control, yeast was grown on YSD-W-L media where all transformed yeast should be able to grow. However, on media lacking histidine and alanine, the yeast could only grow if a protein-protein interaction was present between the two tested genes. This was due to the fact that the upon interaction, the AD and BD domain come

in close proximity and genes are activated which can overcoming the lack of histidine and alanine. In human cells Tudor-SN and E2F-1 form a dimer and facilitate the $G_1$/S transition (Su et al., 2015). TSN proteins are made up of SN domains and a Tudor domain. The Tudor domain has been found to be responsible for protein-protein interactions (Selenko et al., 2001). To narrow down a possible location of interaction, truncated versions of TSN1 and TSN2 were made. The full length TSN1 was tested against E2Fa and E2Fb and no interaction was found (data not shown). This was also true for the truncated version consisting of the first three SN domains (data not shown). E2Fa and E2Fb were used as they had been shown to interact with DPa and DPb (for E2Fb) and with RBR (for E2Fa) and had therefore been confirmed to be functional (Fig.6.16). A truncated version of TSN2 containing the Tudor domain and the final SN domain was also tested against DPa, DPb and E2Fc (Fig. 6.16). Here too, no interaction was found. Although these negative results do not exclude that TSN, DP and E2Fs interact *in planta*, I conclude that the proteins do not interact directly in this yeast assay.

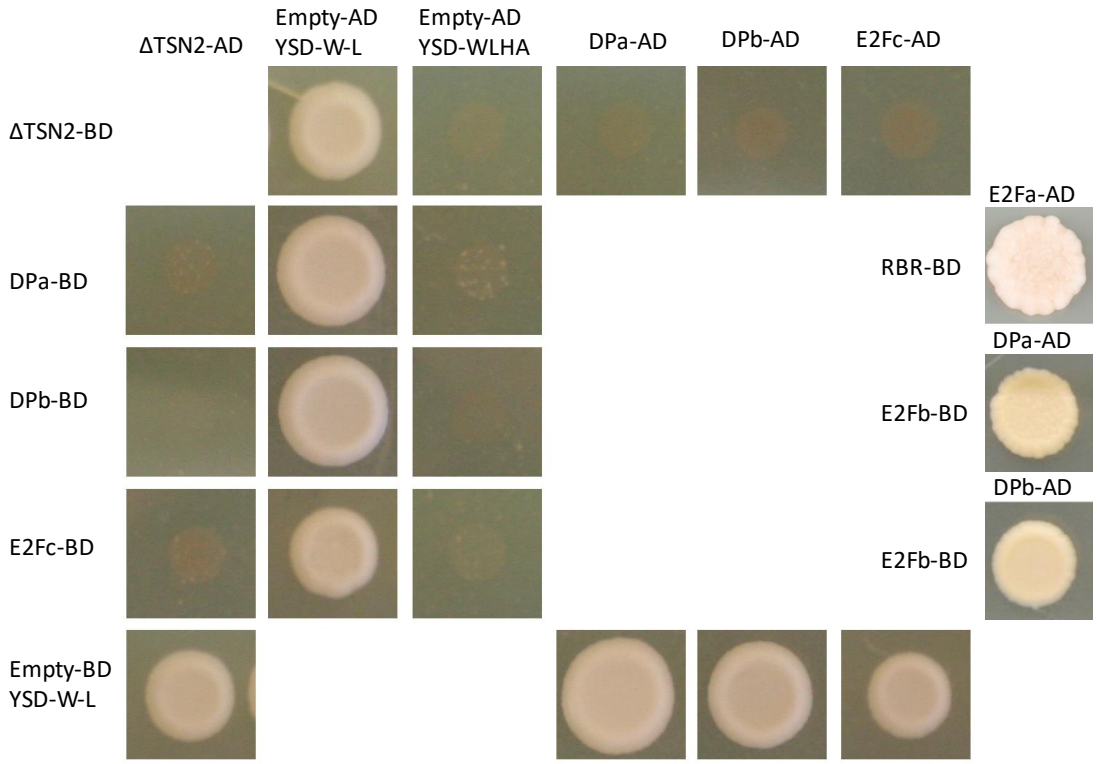



Fig. 6.16. E2F and TSN proteins did not interact. A truncated version of TSN2 made of the Tudor and the fifth SN domain did not interact with DPa, DPb or E2Fc. These were also the results found when testing a full length TSN1 against E2Fa and E2Fb as well as a truncated version of TSN1 containing the first three SN domains (data not shown). Yeast colonies were plated on YSD-WLHA with 2.5mM, 5mM and 10mM 3AT. Results shown here are from 10mM plates. Plates were kept at 37°C for two weeks to grow. Interactions left blank were not tested. E2Fa-RBR, DPa-E2Fb and DPb-E2Fb were positive controls.

### *6.3.11 Metacaspase 4 is downregulated in e2f and tsn mutants*

In mammalian systems, programmed cell death (PCD) is well studied. Whereas PCD roles for *E2Fs* are only starting to emerge in plants, they are well documented in humans where *E2Fs* upregulate caspases needed for PCD (Nahle et al., 2002). Caspase-3 has been shown to cleave TSN (Sundstrom et al., 2009). This cleavage then leads to cell death (Sundstrom et al., 2009). This last part has also been demonstrated in *P. abies* where metacaspaseII-Pa (mcII-Pa) cleaves TSN (Sundstrom et al., 2009).

The *Arabidopsis* genome encodes nine metacaspase genes and *metacaspase-4* (*MC4*) is the closest homologue to *mcII-Pa* and is most widely expressed (Tsiatsiani et al., 2011). As the relationship between the genes mentioned above have not been demonstrated in *Arabidopsis*, I used RT-qPCR to demonstrate that *MC4* was downregulated in *e2f* triple homozygous lines. The levels of *MC4* were significantly downregulated with a 40% downregulation compared to wild-type plants (Fig. 6.17a). Intriguingly *MC4* levels were also significantly downregulated in *tsn1*$^{-/-}$ *tsn2*$^{-/-}$. These data support *E2F* regulation of *MC4* and also suggest a feedback loop between *TSN* and *MC4*.

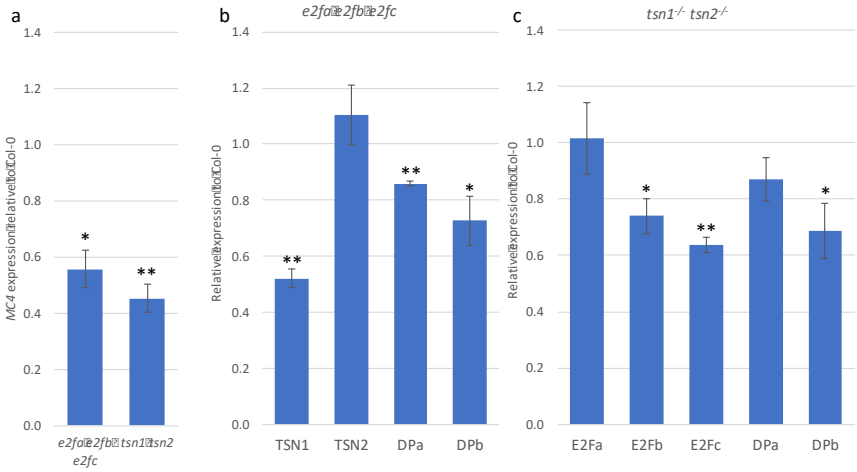



Fig. 6.17. Gene expression levels in mutants were sometimes affected as tested by RT-qPCR. In the *e2f* triple homozygous mutant and in *tsn1*$^{-/-}$ *tsn2*$^{-/-}$, *MC4* levels were strongly reduced (a). In *e2fa*$^{-/-}$ *e2fb*$^{-/-}$ *e2fc*$^{-/-}$ the expression levels of *TSN1* were roughly half the amount they were in wild type but *TSN2* was not affected (b). The levels of the dimerization partners *DPa* and *DPb* were also significantly downregulated (b). Expression levels of *E2Fb*, *E2Fc* and *DPb* were significantly downregulated in *tsn1*$^{-/-}$ *tsn2*$^{-/-}$. *E2Fa* and *DPa* expression levels were not significantly changed (c). Values are means of three biological replicates using inflorescences with fruits up to stage 12. The values are relative to *E1Fα* and *PP2A* in the mutant and Col-0. *p<0.05; **p<0.01

### *6.3.12 A feedback loop between* E2F *and* TSN *is present in* Arabidopsis

To understand the entire system, *TSN* levels in *e2f* triple mutants were investigated. *TSN1* was significantly downregulated but *TSN2* was not (Fig. 6.17b). This is interesting as there were no differences in the phenotypes found apart from a slightly more severe defect when *TSN1* was one of the mutated genes. The dimerization partners *DPa* and *DPb* were also significantly downregulated. Furthermore, *E2Fb, E2Fc and DPb* were significantly downregulated in *tsn1*$^{-/-}$ *tsn2*$^{-/-}$ mutant plants (Fig. 6.17c). *E2Fa* and *DPa* levels were not different. This suggests a loop where *E2Fs* regulate the expression of *TSN* who in turn affect *E2F* levels.

## 6.4 Discussion

This chapter aimed to test whether *TSN* and *E2F* were involved in the regulation of a similar pathway involved in reproductive tissue development. Both groups of genes had previously been linked to the cell cycle (Elbarbary et al., 2017, Polyn et al., 2015) and cell death (Daneva et al., 2016, Wang et al., 2014) in plants and in animals. Results in this chapter provide evidence for a functional overlap of *TSN* and the *E2Fs* in plants, specifically during reproduction.

### *6.4.1* tsn *mutant phenotypes*

*Arabidopsis tsn* mutants have been characterized and published. My results align with the literature in that single mutants did not show a phenotype (Liu et al., 2010, Sundstrom et al., 2009). Double mutants have also been characterized by Sundstrom et al. (2009) who reported pollen development defects. Results presented here clearly demonstrated that the pollen is viable, fully functional and able to fertilize ovules all along the gynoecium. However, pollen was not counted. Sundstrom et al. (2009) described the double heterozygous *tsn1 tsn2* to have abortions at two time points, similar as the *e2f tsn* mutants described in this study.

Studies in plants agree that *TSN1* and *TSN2* function redundantly. However, some papers suggest a slightly stronger effect of one gene over the other. Liu et al. (2010) suggest a more prominent role of *TSN2* over *TSN1* in seed germination and gibberellin biosynthesis. This is opposed to Sundstrom et al. (2009) and Frey et al. (2010) who show *TSN1* to be the slightly more dominant copy. My results align with the latter as the double mutants of *E2F* and *TSN1* showed a more

severe phenotype than the *e2f tsn2* lines. Another difference between the two genes was seen in the expression analysis in *e2f* triple mutant lines where *TSN1* and not *TSN2* was downregulated.

### *6.4.2* e2f *mutant phenotypes*

Phenotypes for *e2f* mutants were not very strong apart from *e2fa-2*$^{-/-}$*e2fb-1*$^{-/-}$ which had a 40% reduction in fruit length. It is possible that the lines used in this study still have residual activities of the genes and the *e2fa-2* is the line with the lowest levels of *E2Fa*, explaining why it is the only single mutant to show a significant reduction in silique length. No silique phenotype were found in the *e2f* triple homozygotes. This is contradictory to results by Wang et al. (2014) who describe reduced fertility but no other vegetative defects, which seems reminiscent of the phenotype observed in *e2f/tsn* mutants. Mutant combinations with *e2fa-2* could help solve this discrepancy. However, as phenotypes are consistent and of the same strengths for all *E2Fs* it is unlikely the results will be dramatically different.

The *E2F* gene family has been well studied in *Arabidopsis* and the three *E2Fs* have been given slightly different roles in development. Their expression patterns also slightly differ. Results presented here did not show a difference between the three *E2Fs*. Mutations in each gene gave the same phenotype and the same level of severity. They were also all able to rescue the phenotype in higher order mutants. Furthermore, the expression of *E2Fa* and *E2Fb* in embryos was the same. *E2Fb, E2Fc, DPb, TSN2* and *RBR1* are expressed throughout all stages of seed development, while *E2Fa* is absent in the mature embryo stage (Le et al., 2010). This supports the expression patterns described here. In my work, only the RT-qPCR data suggested a difference in the *E2Fs* with *E2Fa* being the only *E2F* not downregulated in *tsn1*$^{-/-}$ *tsn2*$^{-/-}$. It is possible that the defects in *e2f* mutants are only manifested in the fruit because problems in reproductive tissue development are fatal unlike those later in plant development, which may be rectified.

### *6.4.3 Ovule development and fertilization require further study*

A full set of ovules was present in the mutants, however, that does not signify that ovule development occurred in the correct manner. Therefore, ovules will be imaged in future work to determine whether the correct patterning is present. While embryos have been imaged from mutants, I am unable to comment on whether all ovules had been fertilized and produced

embryos. The embryos I have imaged have not been morphologically different to wildtype ones (data not shown) but I do not know if they were embryos from healthy seeds or not.

### *6.4.4 The nature of abortions*

The mutation appeared to have a maternal effect as determined through reciprocal crosses. However, ovary size was not affected, a full set of ovules was present at fertilization and embryos were present. Some mutants with ovule defects have similar phenotypes as the mutants presented here. The *four lips-7 (flp-7)* and *flp-1/myb88* mutants for instance do not develop ovules properly due to abnormal meiosis in the megaspore mother cell which leads to an embryo sac that is not formed properly (Makkena et al., 2012). This results in abortion after fertilization. Before fertilization however, more ovules are present in the mutant than in the wild type, demonstrating that a full silique at fertilization is not a sign that ovules have developed properly. Furthermore, this mutation was not lethal and some seeds still developed normally, as in my lines. However, the lines used did not completely abolish the gene functions (Makkena et al., 2012).

It could also be possible that the effect is only present after fertilization. There are three main types of female gametophyte mutants. The first is where the seed is initiated in the absence of fertilization and mutants in this group include *fertilization-independent endosperm, medea* and *fertilization-independent seed2* (Yadegari and Drews, 2004). The second type of mutant is one that affects the embryo and endosperm development after fertilization. Genes in this group are called gametophytic maternal-effect (GME) genes. In most GEM mutants, the embryo is aborted as early as the zygote stage but some have been found that abort later (Pagnussat et al., 2005). The last category are genes involved in imprinting (Yadegari and Drews, 2004).

My phenotype seems to match the second explanation best where the defects are only visible post fertilization. One characterized mutant fitting in this category is *capulet2 (cap2)*. In *cap2* heterozygous mutants, half the seeds are aborted even though the ovule development is the same as in wild type (Grini et al., 2002). The endosperm development in these mutants is aberrant in that fewer divisions take place and nuclei are of different sizes (Grini et al., 2002). However, the embryo develops properly but is aborted at various stages up until the heart stage (Grini et al., 2002). Another mutant that has an embryo that develops until the globular stage is *glauce* (Leshem et al., 2012, Ngo et al., 2007). In this mutant, the central cell in the endosperm

does not get correctly fertilized. In fact, fertilization in general is slower than in wild type but reaches the same number of ovules eventually (Leshem et al., 2012). The egg cell fertilization is not impaired and the embryo therefore develops correctly until the globular stage at which point defects are visible and the seed is aborted (Ngo et al., 2007). As the central cell fails to get fertilized, the single nucleus is still visible and is enlarge (Ngo et al., 2007). This mutation too is lethal in the homozygous state, however, one in six times the central cell does get correctly fertilized (Ngo et al., 2007). These two mutants are some of the few that I have found that allow the embryo to develop past the zygote stage. What they have in common is the problem of endosperm fertilization. Whether this is the case in my mutants is yet to be determined. No difference in embryo development timing was so far noticed (data not shown). Even though embryos were present in aborted seeds, I will test that egg cells are fertilized correctly by crossing a mutant plant with pollen of a CYCA3;2::GUS line, as described by Ngo et al. (2007). If fertilization occurs correctly, the marker should be present in the developing embryo of all seeds.

Most GME mutants or maternal effect mutants are homozygous lethal. However, I could recover all genotypes, suggesting that the mutation can be overcome sometimes. This could be because the mutants are not full knock-outs but seems unlikely as the phenotypes are always of the same strength, regardless of the mutation.

If ovule and fertilization occur correctly, there could be another explanation for the abortion of embryos. Literature has linked *TSN* to accelerated suspensor death in *P. abies* and in *Arabidopsis* (Sundstrom et al., 2009). The data here is supportive of this hypothesis as expression of *E2Fa, E2Fb* and *TSN2* was found in the suspensor and because embryos abort before the heart stage, as described by Sundstrom et al. (2009)**.** It is also possible that problems are encountered in the embryo proper. Such defects would also lead to embryo and seed abortion. To test this, seeds should be visualised within the timeframe of embryo abortion. This could aid in answering whether the embryo is being aborted due to earlier defects, because of the suspensor being aborted prematurely or because of aberrant endosperm development. PCD markers could be used to pinpoint the timing and location of abortion. If the defects are due to accelerated PCD of the embryo suspensor, the results would align with the TSN published role but that would not explain why the effect is maternal.

### *6.4.5 TSN and E2F proteins do not form a dimer*

Another unanswered question is how the *E2Fs* and *TSN* function together. Yeast 2-hybrid results indicated that neither TSN1 nor TSN2 interact with E2Fs. Not all possible interactions were tested however and this could be done in the future. Furthermore, in humans, E2F and TSN form a triplet with GCN5, which have a histone acetyltransferase activity (Su et al., 2015). The *Arabidopsis* orthologue if GCN5 is *At3G54610* but it has not been shown to play a role in the cell cycle (Vlachonasios et al., 2003). Moreover, the interaction in human cells occurs in the nucleus and plant TSN is not found in the nucleus, suggesting that the system in plants must work differently. However, a yeast 3-hybrid system could be used to test whether the *Arabidopsis* GCN5 is able to form a triplet with E2F and TSN. It is unlikely however as GCN5 is only present in the nucleus where TSN is not. Perhaps TSN and E2F do need a third protein to form a triple bond. This could be one of the DPs which could also be tested using the yeast 3-hybrid system.

### *6.4.6 Gene expression was affected in mutants*

The RT-qPCR results showed a reduction in *MC4* levels in *e2f* and *tsn* mutants. Furthermore, several of the genes were downregulated in *e2f* and in *tsn* mutants. Only *TSN1* but not *TSN2* was downregulated in *e2f* mutants. This could be an indication that *TSN1* and *TSN2* have slightly different functions in fruit development. However, the RT-qPCR was carried out with entire inflorescences including many different tissues and cell types. It is therefore possible that some effects were not picked up due to a high background of unrelated tissue. Repeating the experiment using embryos or developing seeds would likely produce more precise datasets. The results could also be verified using the reporter lines. Comparing the signal of *pTSN2:TSN2:GFP* in an *e2f* mutant background to the signal strength in wild type could show whether *TSN2* expression is affected or not.

### *6.4.7 E2F cis-elements in the TSN promoters*

Given that *TSN1* expression is reduced in the *e2f* triple mutant, it is possible that E2Fs directly regulate *TSN1* expression. Most genes regulated by E2Fs contain the E2F-consensus binding site (5'-TTTCGCGCG-3') in their promoter and mostly within the first 400bp of the start codon (Vandepoele et al., 2005). *TSN1* has a *cis*-acting element 77 base pairs from its ATG. Interestingly, downstream of the *TSN2* gene is a region of 4.5 Kb, which contains 25 potential E2F-binding sites. These are densely clustered (located within a region of 710 bases) and ~2.1 kb downstream of the *TSN2* start codon. The long distance between the *TSN2* start codon and a cluster of E2F cis-

element could be the reason for the downregulation of *TSN1* but not *TSN2* in *e2f* triple homozygous lines. However, some genes regulated by the E2Fs do not contain the binding site so an alternative mechanism may be at play here.

### *6.4.8 Hormonal changes are unlikely to be the cause of the phenotypes*

Hormones play an important role in fruit elongation and proper seed development. *TSN* have been linked to gibberellin biosynthesis where *tsn* mutants have lower gibberellin levels (Liu et al., 2010, Yan et al., 2014a). *E2Fb* and *E2Fc* have been shown to be partly regulated by auxin (del Pozo et al., 2002, Magyar et al., 2005). In plants grown on auxin-starved media, E2Fb is degraded and found at much lower levels. Upon re-addition of auxin, E2Fb is stabilized and more mitosis occurs (Magyar et al., 2005). E2Fc on the other hand is stabilized in the *auxin response mutant 1-12* (*axr1-12)* and a larger average cell size is the consequence (del Pozo et al., 2002). In *Arabidopsis* fruits, upon fertilization, an auxin signal is released from the ovules leading to gibberellin biosynthesis and subsequent fruit elongation (Fuentes et al., 2012). Reduced fertilization therefore leads to lower levels of auxin and gibberellin. A reduction in gibberellin due to the mutation in *TSN* is unlikely to be the cause here as gibberellin mutant fruits typically display a reduction in silique length regardless of the seeds inside. Seeds are therefore tightly packed inside the silique. The fruits here however display a constant correlation between fruit length and seed number, suggesting that it is not a lack of gibberellin that caused the reduced silique length.

While I have demonstrated that the pollen can reach ovules and the phenotype is not due to pollen defects, I have not shown that fertilization can occur successfully. However, if we assume that it can occur, the initial gibberellin and auxin signals should be present. It is however possible that, once seeds are aborted, a reduction in auxin is present. This could affect E2Fb and E2Fc stability. E2Fb would then be degraded, leading to less cell division. At the same time, an increase in E2Fc stability would lead to increased cell sizes. The SEM pictures suggest that the cellular arrangement is not altered much and that therefore siliques are different lengths due to a difference in cell number as opposed to size. This is against the phenotype that would be present if auxin was low enough to affect E2Fb and E2Fc stability. Therefore, it is also unlikely that altered auxin levels are the cause of the phenotype.

Altered levels of hormones have been linked to seed size however. A strong link to auxin has been made linking *AUXIN RESPONSE FACTOR 2* (*ARF2)* to integument size which then translates to an

altered seed size (Schruff et al., 2006). Mutants in *ARF2* display larger seeds due to larger integuments. This is in part due to the fact that fewer gynoecia get fertilized but even when fertilization is controlled, plants produce larger seeds (Hughes et al., 2008). The link between total seed number and seed size has often been described. This was also the case in this study where the mutants with fewer seeds displayed larger ones. It is unknown why a reduction in seed set leads to larger seeds being formed. However, two main theories exist. It could be the assimilate redistribution which means that fewer seeds get more nutrients or it could be due to a reduction of space constraint leading to seeds that enlarge more (Sornay et al., 2015). It would be very interesting to study this link as knowledge about this could be used for crop improvement for the production of plants with larger seeds without reducing seed numbers.

### *6.4.9 E2F and TSN mode of action*

The data presented here strongly suggest that *E2Fs* and *TSN* functions link. The RT-qPCR results demonstrated that the metacaspase MC4 is down-regulated in the *e2f* triple homozygous mutant and in the *tsn* double mutant. This suggests that perhaps the PCD pathway in plants is the same in animals. In that kingdom, it has been shown that E2Fs regulate caspases (Nahle et al., 2002) and that caspase-3 cleaves TSN (Sundstrom et al., 2009). In the plant kingdom, this cleavage has also been demonstrated in *P. abies* where mcII-Pa cleaves TSN in the embryo suspensor (Sundstrom et al., 2009). *mcII-Pa* is only found in the suspensor and not expressed anywhere else. *Arabidopsis* metacaspases are expressed more widely and *MC4*, the closest orthologue to *mcII-Pa* is the most widely expressed metacaspase in *Arabidopsis* (Tsiatsiani et al., 2011). In future experiments, it will be tested whether MC4 can cleave TSN1 or TSN2. This will be done using an in-vitro protein expression system. The TSN protein will be c-myc tagged and therefore a western blot will be able to show whether the TSN protein has been cleaved or not. It would also be worth testing the other five *Arabidopsis* type II metacaspases (Piszczek and Gutman, 2007). However, this method will only demonstrate whether TSN is cleaved, not how often. Some differences in kingdoms have been demonstrated whereby the human HsTSN is only cleaved once and the copy in *P. abies* is cleaved four times (Sundstrom et al., 2009). Knowing how often TSN is cleaved could lead to a study investigating whether the cleaved parts have a role. Not much is known about this so far except that cleaved TSN can no longer stabilize mRNAs.

The data presented was combined with knowledge from the animal kingdom to produce a model. I suggest a network of *E2F*, *MC4* and *TSN* in *Arabidopsis* (Fig. 6.18) where the E2Fs

transcriptionally regulate *MC4*, as they do in human cells. Furthermore, they directly regulate transcription of *TSN1*. TSN transcriptionally regulates *E2F* and *MC4*. This is supported by a recent publication in human cells, where TSN can regulate *E2F* transcription levels by degrading mature miRNAs needed for *E2F* transcription (Elbarbary et al., 2017). While this link is untested here, the fact that the other parts of the pathway seem to be conserved in plants and in animals, supports the theory that this effect could also be relevant for *Arabidopsis*. And finally, MC4 may cleave TSN as has been demonstrated in *P. abies* and in animals, generating a model with several feedback loops.

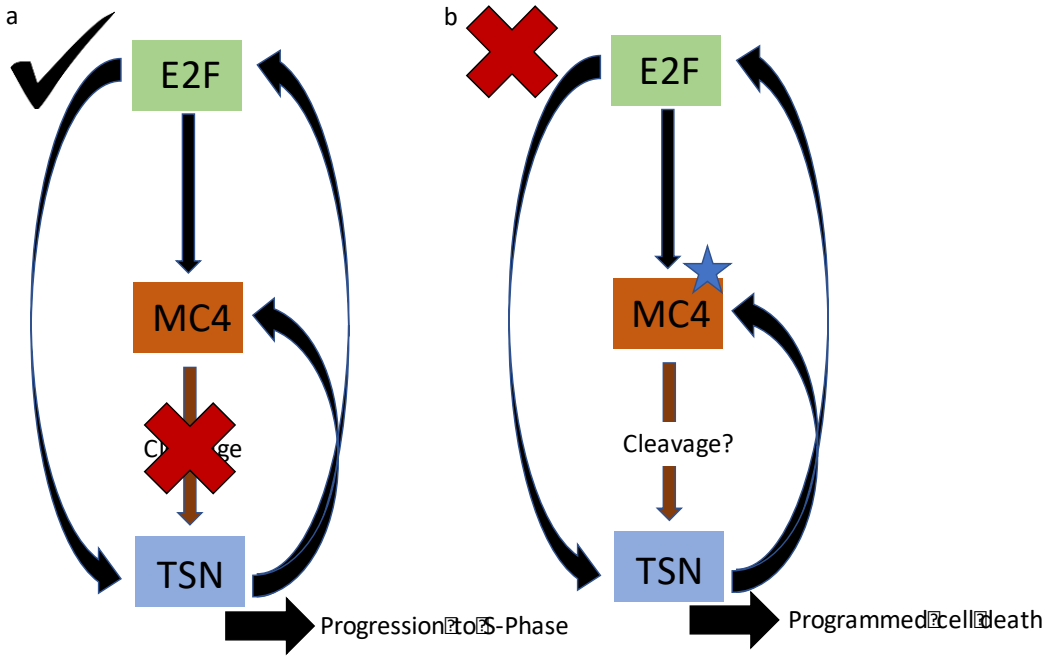



Fig. 6.18. Model of the TSN/MC4/TSN network in *Arabidopsis*. When no damage is present, *MC4* is upregulated by E2F but not activated. Therefore, no cleavage of TSN occurs and the cell progresses to S-phase (a). Upon DNA damage or other perceived abnormality, MC4 undergoes secondary activation and TSN is cleaved, aborting the cell division and leading to programmed cell death (b).

However, while this gives rise to a new network, it is not clear what causes the phenotypes I observed. Once the cleavage is confirmed, the use of *mc4* TDNA mutant lines could help to test the model. This would enable me to identify if it is only MC4 cleaving TSN or if other caspases are involved. Furthermore, blocking the cleavage using metacaspase inhibitors such as leupeptin, antipain or tosyl-lysyl-chloromethylketone should be able to rescue the phenotype. Alternatively, expressing a non-cleavable TSN version could be used to test the effect of TSN gain-of-function.

### *6.4.10 Differences between heterozygous and homozygous mutants*

I consistently found that plants heterozygous for *e2f* mutations had aborted seeds whereas those with homozygous *e2f* mutations were healthy. It is rare that a heterozygous mutation causes more effects than a homozygous one. To my knowledge, the only other example of such a phenomenon was reported by Sotelo-Silveira et al. (2013) who investigated the role of cytochromes P450 *CYP78A9* and *CYP78A8*. Single mutants have slightly shorter siliques and fewer seeds, but not significantly so. Double heterozygous mutants of these genes display siliques that are 60% as long as wildtype with 20% of seeds. Double homozygous lines also show the defect but it is less severe with silique lengths reduced by only 15% and half the amount of seeds (Sotelo-Silveira et al., 2013). The abortions are due to ovule defects and gynoecia pre-fertilization are already smaller than those of the wild types. This makes my phenotypes distinct from their study. However, it is interesting that both studies are in the fruit. The authors do not comment on the phenotype and do not offer an explanation.

When all three *e2f* mutations were homozygous, the phenotype was rescued. I have been unable to find other examples of such an effect. I can hypothesise however that the complete loss of an *E2F* triggers more responses than when it is heterozygous. Such a response could be the upregulation of the other *E2Fs*. This is unlikely however as the $e2fa^{+/-}e2fb^{-/-}e2fc^{+/-}tsn1^{+/-}tsn2^{+/-}$ quintuple mutant produced a full seed set. It is therefore more likely that the rescue of the phenotype occurs through a mechanism distinct to the *E2Fs*. Whether it is a cell cycle gene that ensures division is correctly carried out or whether it is through the cell death pathway is unknown. One aspect of the *E2F* network that this work does not address is the existence of atypical *E2Fs* which are characterized as inhibitors. There are three of these in *Arabidopsis* and they do not have a dimerization domain but instead have a duplicated DNA binding domain, allowing binding of DNA without *DPa* or *DPb* (Kosugi and Ohashi, 2002a, Mariconti et al., 2002, Vandepoele et al., 2002). Whether the atypical *E2Fs* play a role in the phenotypes described here could be tested by carrying out RNA-Seq with heterozygous and homozygous *e2f/tsn* lines. This approach could also show which other genes are affected by the mutations and could hold the answer to the phenotypes presented here.

## 6.5 Concluding remarks

In this Chapter I presented evidence to suggest that the *TSN* and *E2F* pathways overlap. Both of these gene families have been shown to function in cell division and in cell death. *tsn e2f* mutant plants showed normal vegetative development but displayed shorter siliques with fewer seeds. I found that this was due to abortions of seeds at two different time points; just after fertilization and later on, when embryos were present. I propose a model whereby the *E2Fs* and *TSN* are linked through the metacaspase MC4. This is the first time *E2Fs* have been linked to seed development in plants and the first time E2F and TSN have been studied together.

# Chapter 7

## 7 General Discussion

### *7.1.1 Summary of findings*

The aim of this project was to identify ways to increase yield of oilseed rape by focusing on the fruit. This was tackled by identifying how the fruit grows and how different traits relate and influence each other. Through observation of different fruit traits (Chapter 3), the hypothesis that pods with more stomata were shorter was tested (Chapter 4). Next, a GWAS was carried out on several yield-related traits and possible gene candidates underlying traits were identified (Chapter 5). The most promising gene from that study was then used to infer its role on fruit growth in the model plant *Arabidopsis* (Chapter 6).

In summary, most pod and seed traits in *B. napus* were strongly correlated. Longer pods had more seeds while there was usually a trade-off between seed number and seed size and therefore also seed weight. Pods grew steadily in both width and length. On the cellular detail, cells divided more in the apical-basal axis as opposed to the medio-lateral one. Pod growth was largely dependent on cell division as opposed to cell growth.

Using four *B. napus* accessions, *B. oleracea* transformed with three different constructs overexpressing *ICE1* and a series of stomata mutants in *Arabidopsis*, the hypothesis that more stomata lead to shorter fruits was tested. The data suggested the hypothesis was correct with the relationship being present in all three species. It would be interesting in the future to unravel the mechanism underlying this effect.

Next, to find genes not previously associated with fruit development, a Genome-Wide Association Study (GWAS) was carried out. The results confirmed the use of SNP markers and gene expression markers as a possible way to identify target genes affecting a trait of interest.

In the GWAS on pod length, the orthologue of *DPb* was chosen as the best candidate and *TUDOR-SN1* was picked as a gene with a similar expression pattern as the former. E2Fs and TSN were concluded to function together as plants with a mutated *E2F* and one *TSN* had a more pronounced phenotype than mutants in a single gene. Plants grew normally but the fruit length was reduced due to seed abortion. A network was proposed in which the E2Fs and TSN are linked via the metacaspase MC4.

### *7.1.2 Cell-cycle control links all chapters*

While the result chapters in this thesis all tell their own stories, there is something linking them all together. Chapter 3 explained that longer pods arose due to more cell divisions, Chapter 4 saw a relationship between stomata and pod length and Chapter 5 and 6 described cell cycle genes involved in correct fruit and seed development. In fact, the cell cycle is a recurring theme. Particularly RETINOBLASTOMA-RELATED (RBR), which is a pivotal regulator of the $G_1$/S transition, can be linked to all the functions discussed. Chapter 6 briefly discussed that it binds the E2Fs and releases them for downstream target activation. However, RBR has also been linked directly to stomata associated genes.

*RBR* and other cell cycle genes are involved in all stages of stomata development. The initial formative cell division leading to stomata development for instance is regulated by CDKA;1 and RBR (Weimer et al., 2012). Accordingly, *cdka;1* knock-down mutants do not form stomata at all, while loss of function mutations are lethal (Iwakawa et al., 2006, Weimer et al., 2012). *CDKB1;1* and *CDKB1;2* are also required later on for the division giving rise to the fully developed guard cells (Xie et al., 2010).

Moreover, RBR controls the expression of *SPCH*, *SDD-1*, *TMM* and *EPF2*, all genes required for normal stomata patterning (Borghi et al., 2010). The interplay between SPCH and RBR is required for the asymmetric division in stomata development (Weimer et al., 2012). Furthermore, RBR forms heterodimers with FAMA, FLP and its paralogue MYB88, which are all needed for the final steps of correct stomata development (Lee et al., 2014). In the dimer, GMC fate is maintained and the final division leading to the fully functional guard cells is regulated (Borghi et al., 2010, Lee et al., 2014). Additionally, many cell cycle genes such as *CDKB1;1, CDK1;3, CDC6, CYCB1;3* and *CYCD4;1* are targets of FLP and MYB88, displaying another direct link between genes associated with stomata development and the cell cycle (Lee et al., 2014, Xie et al., 2010). *CDKB1;1* is also a target of E2Fa and in fact, the E2Fa and FLP/MYB88 *cis*-regulatory elements in the *CDKB1;1* promoter overlap (Xie et al., 2010).

These results all demonstrate the interplay between genes characterized in stomata development and the cell cycle. Therefore, the link between stomata density and fruit length could be caused by differentially expressed cell cycle genes. To determine whether this is the case, more experiments are needed. Water use efficiency, gas exchange measurements and photosynthetic

efficiency could be measured to inform on the effects of changing the stomata arrangements. Cell division rates, cell sizes and numbers could suggest whether the cell cycle has the more dominant effect. For crop improvement, it would be good to understand the relationship between stomata development and the cell cycle more to enable the production of plants with higher photosynthetic efficiency, lower water-use and larger fruits.

### *7.1.3 Applying the results back to the crop*

This project began with the aim of finding yield-associated genes in *B. napus* with the goal of increasing outputs in the crop. Genes related to fruit development were indeed identified and results in *Arabidopsis* helped uncover their functions in the fruit. Future work will focus on translating this knowledge back to oilseed rape and assessing the effects on yield. It is worth noting that the cell cycle, of course, is crucial for all stages of plant development and modifying the activity of genes involved in this process may have pleiotropic effects. It will therefore be important to use and develop tools that are specific to the development of fruits and seeds.

### *7.1.4 Applicability of the GWAS*

The GWAS was successfully used to identify genes involved in fruit development and *Arabidopsis* mutants had a fruit size phenotype. Many genes are required for the growth and development of the fruit and have been thoroughly characterized (see Introduction). The genes *FUL, RPL, HEC1/2/3, SPT, CKX3* and *CKX5* for instance are all required for normal fruit growth, as determined by their mutants (Alvarez and Smyth, 1999, Bartrina et al., 2011, Ferrandiz et al., 2000, Gremski et al., 2007, Gu et al., 1998, Roeder et al., 2003). Moreover, *IND, SHP1/2* and *ALC* are required for the development of the valve margin and hence fruit opening (Dinneny et al., 2005, Girin et al., 2011, Liljegren et al., 2000, Liljegren et al., 2004, Rajani and Sundaresan, 2001). Therefore, the GWAS could have been expected to yield some of these genes in the analysis of the various factors such as fruit length and pod strength. However, since the transcriptomics were carried out on leaf tissue, some of the genes with an expression pattern restricted to the fruit were not identified in the RNAseq. This includes the *HEC1/2/3, SHP1/2, SPT, IND* and *CKX3* genes and shows a clear drawback of using the leaf transcriptome to infer about fruit growth. However, *RPL* scored at 22$^{nd}$ place in the GWAS using GEMs for the pod length in 2014 and correlated 174$^{th}$ best in the association to valve length only. *Arabidopsis rpl* mutant fruits are half as large as wild type and are less dehiscent (Roeder et al., 2003). Therefore, *RPL* affects pod strength negatively

and pod length positively. Furthermore, RPL binds ETT in an IAA-sensitive manner and therefore contributes to fruit patterning (Simonini et al., 2016). Together, ETT and RPL control the expression of a UDP-glycosyltransferase (*TARGETS UNDER ETTIN CONTROL3* (*TEC3*)) required for correct phyllotactic patterning (Simonini et al., 2017). In *rpl, ett* and *tec3* mutants, primordia appear at a different angle than the 137.5° found in wild type plants (Simonini et al., 2017). Whether this change has an effect on overall productivity of the plant has not been tested. The relationship between pod strength and fruit size suggests that *RPL* may not be a suitable candidate for crop improvement as increases in pod size would likely lead to decreases in pod strength. However, *RPL* has other wide-ranging effects on phylotaxy, meristem maintenance and stem growth which may lead to a change in yield (Bencivenga et al., 2016, Roeder et al., 2003, Simonini et al., 2017). Overall, the fact that *RPL* was identified in the GWAS is yet another example supporting the strength of the method.

### *7.1.5 Using the* Brassicaceae *family in research*

This project began working solely on *B. napus* and then went on to demonstrate the beauty of comparative genetics and the practicality of working on a model species. It showed how research can go from crop to model instead of the more commonly found model to crop.

Working on a crop has many advantages and findings may be used to generate new crops with higher yields. Working on *Arabidopsis* is different for various reasons. The most obvious being the timeframe required for one generation which is around 6-7 weeks for *Arabidopsis* and roughly six months for *B. napus*. This of course means that in the same timeframe, more results can be achieved working on *Arabidopsis* than on *B. napus*. This is reflected in this Ph.D. where more time was spent working on *B. napus* than *Arabidopsis* and yet more mechanistic and genetic results were obtained in *Arabidopsis*. The model plant also allowed me to have more detailed results to enable the study of two previously unconnected gene families. The availability of mutants and the simpler genome are also of advantage. Furthermore, the ease of transformation makes the generation of new mutants and reporter lines possible. Starting from a crop that takes more than six months to grow, is allotetraploid and is best transformed by tissue culture has made me really appreciate the increased simplicity and speed of working with *Arabidopsis*. I am grateful to have worked on both species and to have learned to appreciate the model.

The thesis would not have been possible without the work done in the crop, however. The advantage of using the GWAS to identify genes underlying traits not previously associated with their function was demonstrated. This approach was successfully used for pod length where the DP/E2F network was found. The link to *TSN* would not have been investigated if the GWAS had not identified it as a possible candidate with similar expression levels as *BnaA3.DPb*. Furthermore, Charlotte Miller (Bevan lab, JIC) has used the same approach to identify a gene linked to seed size (personal communication). These results confirm the benefits of GWAS and suggest that the function of many more genes could be unravelled using this method.

Furthermore, *B. napus* is an excellent species to study for yield as the yield-associated parameters related to the fruit are much larger than those in *Arabidopsis.* This is advantageous for several reasons. For one, it makes handling and measuring of some parameters much easier. Seeds per pod and seed size for instance are easier to measure in *B. napus* than in *Arabidopsis*. Furthermore, since the organs are so much bigger, more variation is noticeable which perhaps in *Arabidopsis* would be overlooked. Crop improvement is often on a relatively small scale, one more seed per pod for instance could make a large difference in overall yield. Assuming each plant has around 300 pods, that is 300 more seeds per plant and a lot more per hectare. When using *Arabidopsis* however, such small differences are often overlooked or not seen as a substantial phenotype to study.

### 7.2 Further work

Some questions remain regarding the interplay of *TSN* and the *E2Fs*. The cause of abortion will be determined in future experiments a well as whether MC4 cleaves TSN. These questions will be addressed within the next months.

I am also curious about other factors which would require more work.

A consistently present relationship was the one between seed number and fruit length. This was true for different accessions of *B. napus* as well as the *Arabidopsis* mutants with seed abortions. Not much is known about this relationship apart from the fact that seeds emit hormonal signals for pod growth (Dorcey et al., 2009, Fuentes et al., 2012). Hormones and hormone inhibitors could be used to identify to what extent this relationship is dependent on gibberellin and auxin. It

would be interesting to identify the genes associated with this phenomenon and to elucidate whether their function is seed-related or not.

Not much is known either about why seeds grow larger when there are fewer. The hypotheses about this include an altered nutrient allocation as well as a reduction in space constraint allowing seeds to grow larger (Sornay et al., 2015). The former hypothesis could be tested by growing fruits off the plant in an in-vitro method (Lardon et al., 1993) and controlling the amount of nutrients the fruits receive. The latter could be investigated using gibberellin mutants for instance which are defective in fruit elongation. *fruitful* mutants could also be used to this end as they also fail to elongate post fertilization (Gu et al., 1998). If the space constraint is the main player, seeds would be smaller in those fruits. Ideally, an inducible system would be used to eliminate pre-fertilization effects.

Another relationship that was consistently present in all species studied was the negative correlation between seed number and seed size. This is a drawback for breeding where increases in both seed number and size would be preferable. Studying that relationship could be very interesting and would find instant application. The GWAS results could be used to this end. Using *Arabidopsis* to study the role of *At3G46100* or other candidates that were found in the GWAS on the calculated trait of seeds per pod x seed area could lead to the identification of genes not yet known to play a role in seed development. The knowledge could then be transferred back to the crop. With luck, a universal gene could be identified with the same function in *B. napus* as in other crops such as wheat or rice, where a large grain size is also desirable.

### 7.3 Concluding Remarks

The aim of this Ph.D. thesis was to find ways of increasing *B. napus* yield by focusing on pod growth characteristics. The project has described how pod-growth traits are related and genes were identified that have an impact on fruit development. These were then further studied in *Arabidopsis*. This work has shown how results in a crop can lead to the identification of general mechanisms in fruit growth. Often efforts focus on the translation of knowledge from a model species to a crop species but the work presented here demonstrates that the opposite direction is also a very productive strategy to increase our fundamental knowledge of biological processes.

# 8 References

Abraham, Z. & Del Pozo, J. C. 2012. Ectopic Expression of E2FB, a Cell Cycle Transcription Factor, Accelerates Flowering and Increases Fruit Yield in Tomato. *Journal of Plant Growth Regulation,* 31**,** 11-24.

Adamczyk, B. J., Lehti-Shiu, M. D. & Fernandez, D. E. 2007. The MADS domain factors AGL15 and AGL18 act redundantly as repressors of the floral transition in Arabidopsis. *Plant Journal,* 50**,** 1007-1019.

Adamski, N. M., Anastasiou, E., Eriksson, S., O'neill, C. M. & Lenhard, M. 2009. Local maternal control of seed size by KLUH/CYP78A5-dependent growth signaling. *Proceedings of the National Academy of Sciences of the United States of America,* 106**,** 20115-20120.

AHDB 2016. Oilseed rape guide.

Albani, D., Mariconti, L., Ricagno, S., Pitto, L., Moroni, C., Helin, K. & Cella, R. 2000. DcE2F, a functional plant E2F-like transcriptional activator from Daucus carota. *Journal of Biological Chemistry,* 275**,** 19258-19267.

Alvarez, J. & Smyth, D. R. 1999. CRABS CLAW and SPATULA, two Arabidopsis genes that control carpel development in parallel with AGAMOUS. *Development,* 126**,** 2377-2386.

Alvarez-Canterbury, A. M. R., Flores, D. J., Keymanesh, K., To, K. & Brusslan, J. A. 2014. A double SORLIP1 element is required for high light induction of ELIP genes in Arabidopsis thaliana. *Plant Molecular Biology,* 84**,** 259-267.

Argueso, C. T., Raines, T. & Kieber, J. J. 2010. Cytokinin signaling and transcriptional networks. *Current Opinion in Plant Biology,* 13**,** 533-539.

Arnaud, N., Girin, T., Sorefan, K., Fuentes, S., Wood, T. A., Lawrenson, T., Sablowski, R. & Ostergaard, L. 2010. Gibberellins control fruit patterning in Arabidopsis thaliana. *Genes & Development,* 24**,** 2127-2132.

Arnaud, N., Lawrenson, T., Ostergaard, L. & Sablowski, R. 2011. The Same Regulatory Point Mutation Changed Seed-Dispersal Structures in Evolution and Domestication. *Current Biology,* 21**,** 1215-1219.

Ascencio-Ibanez, J. T., Sozzani, R., Lee, T. J., Chu, T. M., Wolfinger, R. D., Cella, R. & Hanley-Bowdoin, L. 2008. Global analysis of Arabidopsis gene expression uncovers a complex array of changes impacting pathogen response and cell cycle during geminivirus infection. *Plant Physiology,* 148**,** 436-454.

Atwell, S., Huang, Y. S., Vilhjalmsson, B. J., Willems, G., Horton, M., Li, Y., Meng, D. Z., Platt, A., Tarone, A. M., Hu, T. T., Jiang, R., Muliyati, N. W., Zhang, X., Amer, M. A., Baxter, I., Brachi, B., Chory, J., Dean, C., Debieu, M., De Meaux, J., Ecker, J. R., Faure, N., Kniskern, J. M., Jones, J. D. G., Michael, T., Nemri, A., Roux, F., Salt, D. E., Tang, C. L., Todesco, M., Traw, M. B., Weigel, D., Marjoram, P., Borevitz, J. O., Bergelson, J. & Nordborg, M. 2010. Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. *Nature,* 465**,** 627-631.

Bancroft, I., Morgan, C., Fraser, F., Higgins, J., Wells, R., Clissold, L., Baker, D., Long, Y., Meng, J. L., Wang, X. W., Liu, S. Y. & Trick, M. 2011. Dissecting the genome of the polyploid crop oilseed rape by transcriptome sequencing. *Nature Biotechnology,* 29**,** 762-U128.

Bartrina, I., Otto, E., Strnad, M., Werner, T. & Schmulling, T. 2011. Cytokinin Regulates the Activity of Reproductive Meristems, Flower Organ Size, Ovule Formation, and Thus Seed Yield in Arabidopsis thaliana. *Plant Cell,* 23**,** 69-80.

Bazakos, C., Hanemian, M., Trontin, C., Jimenez-Gomez, J. M. & Loudet, O. 2017. New Strategies and Tools in Quantitative Genetics: How to Go from the Phenotype to the Genotype. *In:* Merchant, S. S. (ed.) *Annual Review of Plant Biology, Vol 68.* Palo Alto: Annual Reviews.

Beilstein, M. A., Nagalingum, N. S., Clements, M. D., Manchester, S. R. & Mathews, S. 2010. Dated molecular phylogenies indicate a Miocene origin for Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America,* 107**,** 18724-18728.

Belles-Boix, E., Hamant, O., Witiak, S. M., Morin, H., Traas, J. & Pautot, V. 2006. KNAT6: an Arabidopsis homeobox gene involved in meristem activity and organ separation. *Plant Cell,* 18**,** 1900-1907.

Bencivenga, S., Serrano-Mislata, A., Bush, M., Fox, S. & Sablowski, R. 2016. Control of Oriented Tissue Growth through Repression of Organ Boundary Genes Promotes Stem Morphogenesis. *Developmental Cell,* 39**,** 198-208.

Benkova, E., Michniewicz, M., Sauer, M., Teichmann, T., Seifertova, D., Jurgens, G. & Friml, J. 2003. Local, efflux-dependent auxin gradients as a common module for plant organ formation. *Cell,* 115**,** 591-602.

Berg, M., Rogers, R., Muralla, R. & Meinke, D. 2005. Requirement of aminoacyl-tRNA synthetases for gametogenesis and embryo development in Arabidopsis. *Plant Journal,* 44**,** 866-878.

Berger, D. & Altmann, T. 2000. A subtilisin-like serine protease involved in the regulation of stomatal density and distribution in Arabidopsis thaliana. *Genes & Development,* 14**,** 1119-1131.

Bergmann, D. C. & Sack, F. D. 2007. Stomatal development. *Annual Review of Plant Biology.* Palo Alto: Annual Reviews.

Boatright, K. M. & Salvesen, G. S. 2003. Mechanisms of caspase activation. *Current Opinion in Cell Biology,* 15**,** 725-731.

Bohuon, E. J. R., Keith, D. J., Parkin, I. a. P., Sharpe, A. G. & Lydiate, D. J. 1996. Alignment of the conserved C genomes of Brassica oleracea and Brassica napus. *Theoretical and Applied Genetics,* 93**,** 833-839.

Boniotti, M. B. & Gutierrez, C. 2001. A cell-cycle-regulated kinase activity phosphorylates plant retinoblastoma protein and contains, in Arabidopsis, a CDKA/cyclin D complex. *Plant Journal,* 28**,** 341-350.

Borghi, L., Gutzat, R., Futterer, J., Laizet, Y., Hennig, L. & Gruissem, W. 2010. Arabidopsis RETINOBLASTOMA-RELATED Is Required for Stem Cell Maintenance, Cell Differentiation, and Lateral Organ Production. *Plant Cell,* 22**,** 1792-1811.

Bozhkov, P. V., Suarez, M. F., Filonova, L. H., Daniel, G., Zamyatnin, A. A., Rodriguez-Nieto, S., Zhivotovsky, B. & Smertenko, A. 2005. Cysteine protease mcll-Pa executes programmed cell death during plant embryogenesis. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 14463-14468.

Brachi, B., Faure, N., Horton, M., Flahauw, E., Vazquez, A., Nordborg, M., Bergelson, J., Cuguen, J. & Roux, F. 2010. Linkage and Association Mapping of Arabidopsis thaliana Flowering Time in Nature. *Plos Genetics,* 6**,** 17.

Brachi, B., Morris, G. P. & Borevitz, J. O. 2011. Genome-wide association studies in plants: the missing heritability is in the field. *Genome Biology,* 12**,** 8.

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y. & Buckler, E. S. 2007. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics,* 23**,** 2633-2635.

Bruce, D. M., Hobson, R. N., Morgan, C. L. & Child, R. D. 2001. Threshability of shatter-resistant seed pods in oilseed rape. *Journal of Agricultural Engineering Research,* 80**,** 343-350.

Bus, A., Korber, N., Snowdon, R. J. & Stich, B. 2011. Patterns of molecular variation in a species-wide germplasm set of Brassica napus. *Theoretical and Applied Genetics,* 123**,** 1413-1423.

Castellano, M. M., Del Pozo, J. C., Ramirez-Parra, E., Brown, S. & Gutierrez, C. 2001. Expression and stability of Arabidopsis CDC6 are associated with endoreplication. *Plant Cell,* 13**,** 2671-2686.

Chandran, D., Rickert, J., Huang, Y., Steinwand, M. A., Marr, S. K. & Wildermuth, M. C. 2014. Atypical E2F transcriptional repressor DEL1 acts at the intersection of plant growth and immunity by controlling the hormone salicylic acid. *Cell Host Microbe,* 15**,** 506-13.

Chauvaux, N., Child, R., John, K., Ulvskov, P., Borkhardt, B., Prinsen, E. & Vanonckelen, H. A. 1997. The role of auxin in cell separation in the dehiscence zone of oilseed rape pods. *Journal of Experimental Botany,* 48**,** 1423-1429.

Cheng, Y., Cao, L., Wang, S., Li, Y. P., Shi, X. Z., Liu, H., Li, L. X., Zhang, Z. L., Fowke, L. C., Wang, H. & Zhou, Y. M. 2013. Downregulation of multiple CDK inhibitor ICK/KRP genes upregulates the E2F pathway and increases cell proliferation, and organ and seed sizes in Arabidopsis. *Plant Journal,* 75**,** 642-655.

Child, R. D., Chauvaux, N., John, K., Ulvskov, P. & Van Onckelen, H. A. 1998. Ethylene biosynthesis in oilseed rape pods in relation to pod shatter. *Journal of Experimental Botany,* 49**,** 829-838.

Child, R. D., Summers, J. E., Babij, J., Farrent, J. W. & Bruce, D. M. 2003. Increased resistance to pod shatter is associated with changes in the vascular structure in pods of a resynthesized Brassica napus line. *J Exp Bot,* 54**,** 1919-30.

Crawford, B. C. W., Ditta, G. & Yanofsky, M. F. 2007. The NTT gene is required for transmitting-tract development in carpels of Arabidopsis thaliana. *Current Biology,* 17**,** 1101-1108.

Crawford, B. C. W. & Yanofsky, M. F. 2011. HALF FILLED promotes reproductive tract development and fertilization efficiency in Arabidopsis thaliana. *Development,* 138**,** 2999-3009.

Cropscience, B. 2014. Strong seed pods. *Bayer research,* 26.

D'andrea, L. D. & Regan, L. 2003. TPR proteins: the versatile helix. *Trends in Biochemical Sciences,* 28**,** 655-662.

Dai, X. H., Hayashi, K., Nozaki, H., Cheng, Y. F. & Zhao, Y. D. 2005. Genetic and chemical analyses of the action mechanisms of sirtinol in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 3129-3134.

Daneva, A., Gao, Z., Van Durme, M. & Nowack, M. K. 2016. Functions and Regulation of Programmed Cell Death in Plant Development. *In:* Schekman, R. (ed.) *Annual Review of Cell and Developmental Biology, Vol 32.* Palo Alto: Annual Reviews.

De Jager, S. M., Menges, M., Bauer, U. M. & Murray, J. a. H. 2001. Arabidopsis E2F1 binds a sequence present in the promoter of S-phase-regulated gene AtCDC6 and is a

member of a multigene family with differential activities. *Plant Molecular Biology,* 47**,** 555-568.

De Jager, S. M., Scofield, S., Huntley, R. P., Robinson, A. S., Den Boer, B. G. W. & Murray, J. a. H. 2009. Dissecting regulatory pathways of G1/S control in Arabidopsis: common and distinct targets of CYCD3;1, E2Fa and E2Fc. *Plant Molecular Biology,* 71**,** 345-365.

De Marcos, A., Trivino, M., Fenoll, C. & Mena, M. 2016. Too many faces for TOO MANY MOUTHS? *New Phytologist,* 210**,** 779-785.

De Veylder, L., Beeckman, T., Beemster, G. T. S., Engler, J. D., Ormenese, S., Maes, S., Naudts, M., Van Der Schueren, E., Jacqmard, A., Engler, G. & Inze, D. 2002. Control of proliferation, endoreduplication and differentiation by the Arabidopsis E2Fa-DPa transcription factor. *Embo Journal,* 21**,** 1360-1368.

Dean, G., Casson, S. & Lindsey, K. 2004. KNAT6 gene of Arabidopsis is expressed in roots and is required for correct lateral root formation. *Plant Molecular Biology,* 54**,** 71-84.

Defra 2017. Agriculture in the United Kingdom 2016.

Del Pozo, J. C., Boniotti, M. B. & Gutierrez, C. 2002. Arabidopsis E2Fc functions in cell division and is degraded by the ubiquitin-SCFAtSKP2 pathway in response to light. *Plant Cell,* 14**,** 3057-3071.

Del Pozo, J. C., Diaz-Trivino, S., Cisneros, N. & Gutierrez, C. 2006. The balance between cell division and endoreplication depends on E2FC-DPB, transcription factors regulated by the ubiquitin-SCFSKP2A pathway in Arabidopsis. *Plant Cell,* 18**,** 2224-2235.

Del Pozo, J. C., Diaz-Trivino, S., Cisneros, N. & Gutierrez, C. 2007. The E2FC-DPB Transcription Factor Controls Cell Division, Endoreplication and Lateral Root Formation in a SCF-Dependent Manner. *Plant signaling & behavior,* 2**,** 273-4.

Delgado, D., Alonso-Blanco, C., Fenoll, C. & Mena, M. 2011. Natural variation in stomatal abundance of Arabidopsis thaliana includes cryptic diversity for different developmental processes. *Annals of Botany,* 107**,** 1247-1258.

Dharmasiri, N., Dharmasiri, S., Weijers, D., Lechner, E., Yamada, M., Hobbie, L., Ehrismann, J. S., Jurgens, G. & Estelle, M. 2005. Plant development is regulated by a family of auxin receptor F box proteins. *Developmental Cell,* 9**,** 109-119.

Dinneny, J. R., Weigel, D. & Yanofsky, M. F. 2005. A genetic framework for fruit patterning in Arabidopsis thaliana. *Development,* 132**,** 4687-4696.

Doerge, R. W. 2002. Mapping and analysis of quantitative trait loci in experimental populations. *Nature Reviews Genetics,* 3**,** 43-52.

Doherty, C. J., Van Buskirk, H. A., Myers, S. J. & Thomashow, M. F. 2009. Roles for Arabidopsis CAMTA Transcription Factors in Cold-Regulated Gene Expression and Freezing Tolerance. *Plant Cell,* 21**,** 972-984.

Dorcey, E., Urbez, C., Blazquez, M. A., Carbonell, J. & Perez-Amador, M. A. 2009. Fertilization-dependent auxin response in ovules triggers fruit development through the modulation of gibberellin metabolism in Arabidopsis. *Plant Journal,* 58**,** 318-332.

Dow, G. J., Berry, J. A. & Bergmann, D. C. 2014. The physiological importance of developmental mechanisms that enforce proper stomatal spacing in Arabidopsis thaliana. *New Phytologist,* 201**,** 1205-1217.

Downes, B. P., Stupar, R. M., Gingerich, D. J. & Vierstra, R. D. 2003. The HECT ubiquitin-protein ligase (UPL) family in Arabidopsis: UPL3 has a specific role in trichome development. *Plant Journal,* 35**,** 729-742.

Drews, G. N. & Yadegari, R. 2002. Development and function of the angiosperm female gametophyte. *Annual Review of Genetics,* 36**,** 99-124.

Du, L., Li, N., Chen, L. L., Xu, Y. X., Li, Y., Zhang, Y. Y., Li, C. Y. & Li, Y. H. 2014. The Ubiquitin Receptor DA1 Regulates Seed and Organ Size by Modulating the Stability of the Ubiquitin-Specific Protease UBP15/SOD2 in Arabidopsis. *Plant Cell,* 26**,** 665-677.

Duncan, S., Holm, S., Questa, J., Irwin, J., Grant, A. & Dean, C. 2015. Seasonal shift in timing of vernalization as an adaptation to extreme winter. *Elife,* 4**,** 11.

Edwards, K., Johnstone, C. & Thompson, C. 1991. A SIMPLE AND RAPID METHOD FOR THE PREPARATION OF PLANT GENOMIC DNA FOR PCR ANALYSIS. *Nucleic Acids Research,* 19**,** 1349-1349.

Elbarbary, R. A., Miyoshi, K., Myers, J. R., Du, P., Ashton, J. M., Tian, B. & Maquat, L. E. 2017. Tudor-SN-mediated endonucleolytic decay of human cell microRNAs promotes G1/S phase transition. *Science,* 356**,** 859-862.

Eldridge, T., Langowski, L., Stacey, N., Jantzen, F., Moubayidin, L., Sicard, A., Southam, P., Kennaway, R., Lenhard, M., Coen, E. S. & Ostergaard, L. 2016. Fruit shape diversity in the Brassicaceae is generated by varying patterns of anisotropy. *Development,* 143**,** 3394-3406.

Fang, W. J., Wang, Z. B., Cui, R. F., Li, J. & Li, Y. H. 2012. Maternal control of seed size by EOD3/CYP78A6 in Arabidopsis thaliana. *Plant Journal,* 70**,** 929-939.

Ferrandiz, C. 2002. Regulation of fruit dehiscence in Arabidopsis. *J Exp Bot,* 53**,** 2031-8.

Ferrandiz, C., Liljegren, S. J. & Yanofsky, M. F. 2000. Negative regulation of the SHATTERPROOF genes by FRUITFULL during Arabidopsis fruit development. *Science,* 289**,** 436-438.

Ferrandiz, C., Pelaz, S. & Yanofsky, M. F. 1999. Control of carpel and fruit development in arabidopsis. *Annual Review of Biochemistry,* 68**,** 321-354.

Ferreyra, M. L. F., Casadevall, R., D'andrea, L., Abdelgawad, H., Beemster, G. T. S. & Casati, P. 2016. AtPDCD5 Plays a Role in Programmed Cell Death after UV-B Exposure in Arabidopsis. *Plant Physiology,* 170**,** 2444-2460.

Flanagan, C. A., Hu, Y. & Ma, H. 1996. Specific expression of the AGL1 MADS-box gene suggests regulatory functions in Arabidopsis gynoecium and ovule development. *Plant Journal,* 10**,** 343-353.

Fleet, C. M. & Sun, T. P. 2005. A DELLAcate balance: the role of gibberellin in plant morphogenesis. *Current Opinion in Plant Biology,* 8**,** 77-85.

Fourmann, M., Barret, P., Renard, M., Pelletier, G., Delourme, R. & Brunel, D. 1998. The two genes homologous to Arabidopsis FAE1 co-segregate with the two loci governing erucic acid content in Brassica napus. *Theoretical and Applied Genetics,* 96**,** 852-858.

Franks, P. J., Doheny-Adams, T. W., Britton-Harper, Z. J. & Gray, J. E. 2015. Increasing water-use efficiency directly through genetic manipulation of stomatal density. *New Phytologist,* 207**,** 188-195.

Frey, N. F. D., Muller, P., Jammes, F., Kizis, D., Leung, J., Perrot-Rechenmann, C. & Bianchi, M. W. 2010. RNA Binding Protein Tudor-SN Is Essential for Stress Tolerance and Stabilizes Levels of Stress-Responsive mRNAs Encoding Secreted Proteins in Arabidopsis. *Plant Cell,* 22**,** 1575-1591.

Fuentes, S., Ljung, K., Sorefan, K., Alvey, E., Harberd, N. P. & Ostergaard, L. 2012. Fruit Growth in Arabidopsis Occurs via DELLA-Dependent and DELLA-Independent Gibberellin Responses. *Plant Cell,* 24**,** 3982-3996.

Gallois, J. L., Woodward, C., Reddy, G. V. & Sablowski, R. 2002. Combined SHOOT MERISTEMLESS and WUSCHEL trigger ectopic organogenesis in Arabidopsis. *Development,* 129**,** 3207-3217.

Geisler, M., Nadeau, J. & Sack, F. D. 2000. Oriented asymmetric divisions that generate the stomatal spacing pattern in Arabidopsis are disrupted by the too many mouths mutation. *Plant Cell,* 12**,** 2075-2086.

Geisler, M., Yang, M. & Sack, F. D. 1998. Divergent regulation of stomatal initiation and patterning in organ and suborgan regions of the Arabidopsis mutants too many mouths and four lips. *Planta,* 205**,** 522-530.

Girin, T., Paicu, T., Stephenson, P., Fuentes, S., Korner, E., O'brien, M., Sorefan, K., Wood, T. A., Balanza, V., Ferrandiz, C., Smyth, D. R. & Ostergaard, L. 2011. INDEHISCENT and SPATULA Interact to Specify Carpel and Valve Margin Tissue and Thus Promote Seed Dispersal in Arabidopsis. *Plant Cell,* 23**,** 3641-3653.

Girin, T., Sorefan, K. & Ostergaard, L. 2009. Meristematic sculpting in fruit development. *Journal of Experimental Botany,* 60**,** 1493-1502.

Girin, T., Stephenson, P., Goldsack, C. M. P., Kempin, S. A., Perez, A., Pires, N., Sparrow, P. A., Wood, T. A., Yanofsky, M. F. & Ostergaard, L. 2010. Brassicaceae INDEHISCENT genes specify valve margin cell fate and repress replum formation. *Plant Journal,* 63**,** 329-338.

Goetz, M., Vivian-Smith, A., Johnson, S. D. & Koltunow, A. M. 2006. AUXIN RESPONSE FACTOR8 is a negative regulator of fruit initiation in Arabidopsis. *Plant Cell,* 18**,** 1873-1886.

Greenboim-Wainberg, Y., Maymon, I., Borochov, R., Alvarez, J., Olszewski, N., Ori, N., Eshed, Y. & Weiss, D. 2005. Cross talk between gibberellin and cytokinin: The Arabidopsis GA response inhibitor SPINDLY plays a positive role in cytokinin signaling. *Plant Cell,* 17**,** 92-102.

Gremski, K., Ditta, G. & Yanofsky, M. F. 2007. The HECATE genes regulate female reproductive tract development in Arabidopsis thaliana. *Development,* 134**,** 3593-3601.

Griffiths, A. M., Jh. Suzuki, Dt. Lewontin, Rc. Gelbart, Wm. 2000. An Introduction to Genetic Analysis. 7th edition.

Grini, P. E., Jurgens, G. & Hulskamp, M. 2002. Embryo and endosperm development is disrupted in the female gametophytic capulet mutants of Arabidopsis. *Genetics,* 162**,** 1911-1925.

Groszmann, M., Paicu, T., Alvarez, J. P., Swain, S. M. & Smyth, D. R. 2011. SPATULA and ALCATRAZ, are partially redundant, functionally diverging bHLH genes required for Arabidopsis gynoecium and fruit development. *Plant Journal,* 68**,** 816-829.

Gu, Q., Ferrandiz, C., Yanofsky, M. F. & Martienssen, R. 1998. The FRUITFULL MADS-box gene mediates cell differentiation during Arabidopsis fruit development. *Development,* 125**,** 1509-1517.

Guilfoyle, T. J. 2015. The PB1 Domain in Auxin Response Factor and Aux/IAA Proteins: A Versatile Protein Interaction Module in the Auxin Response. *Plant Cell,* 27**,** 33-43.

Gutierrez, C., Ramirez-Parra, E., Castellano, M. M. & Del Pozo, J. C. 2002. G(1) to S transition: more than a cell cycle engine switch. *Current Opinion in Plant Biology,* 5**,** 480-486.

Gutierrez-Beltran, E., Moschou, P. N., Smertenko, A. P. & Bozhkov, P. V. 2015. Tudor Staphylococcal Nuclease Links Formation of Stress Granules and Processing Bodies with mRNA Catabolism in Arabidopsis. *Plant Cell,* 27**,** 926-943.

Hara, K., Kajita, R., Torii, K. U., Bergmann, D. C. & Kakimoto, T. 2007. The secretory peptide gene EPF1 enforces the stomatal one-cell-spacing rule. *Genes & Development,* 21**,** 1720-1725.

Harper, A. L., Trick, M., Higgins, J., Fraser, F., Clissold, L., Wells, R., Hattori, C., Werner, P. & Bancroft, I. 2012. Associative transcriptomics of traits in the polyploid crop species Brassica napus. *Nature Biotechnology,* 30**,** 798-802.

Hauvermale, A. L., Ariizumi, T. & Steber, C. M. 2012. Gibberellin Signaling: A Theme and Variations on DELLA Repression. *Plant Physiology,* 160**,** 83-92.

Hay, A. & Tsiantis, M. 2010. KNOX genes: versatile regulators of plant development and diversity. *Development,* 137**,** 3153-3165.

He, S. S., Liu, J., Xie, Z., O'neill, D. & Dotson, S. 2004. Arabidopsis E2Fa plays a bimodal role in regulating cell division and cell growth. *Plant Mol Biol,* 56**,** 171-84.

He, Z., Wang, L., Harper, A. L., Havlickova, L., Pradhan, A. K., Parkin, I. A. & Bancroft, I. 2016. Extensive homoeologous genome exchanges in allopolyploid crops revealed by mRNAseq-based visualization. *Plant Biotechnol J*.

Heisler, M. G. B., Atkinson, A., Bylstra, Y. H., Walsh, R. & Smyth, D. R. 2001. SPATULA, a gene that controls development of carpel margin tissues in Arabidopsis, encodes a bHLH protein. *Development,* 128**,** 1089-1098.

Hepworth, C., Doheny-Adams, T., Hunt, L., Cameron, D. D. & Gray, J. E. 2015. Manipulating stomatal density enhances drought tolerance without deleterious effect on nutrient uptake. *New Phytol*.

Hepworth, S. R., Zhang, Y. L., Mckim, S., Li, X. & Haughn, G. 2005. BLADE-ON-PETIOLE-dependent signaling controls leaf and floral patterning in Arabidopsis. *Plant Cell,* 17**,** 1434-1448.

Hu, J. H., Mitchum, M. G., Barnaby, N., Ayele, B. T., Ogawa, M., Nam, E., Lai, W. C., Hanada, A., Alonso, J. M., Ecker, J. R., Swain, S. M., Yamaguchi, S., Kamiya, Y. & Sun, T. P. 2008. Potential sites of bioactive gibberellin production during reproductive growth in Arabidopsis. *Plant Cell,* 20**,** 320-336.

Hua, Y. P., Zhou, T., Ding, G. D., Yang, Q. Y., Shi, L. & Xu, F. S. 2016. Physiological, genomic and transcriptional diversity in responses to boron deficiency in rapeseed genotypes. *Journal of Experimental Botany,* 67**,** 5769-5784.

Huang, X. H., Wei, X. H., Sang, T., Zhao, Q. A., Feng, Q., Zhao, Y., Li, C. Y., Zhu, C. R., Lu, T. T., Zhang, Z. W., Li, M., Fan, D. L., Guo, Y. L., Wang, A., Wang, L., Deng, L. W., Li, W. J., Lu, Y. Q., Weng, Q. J., Liu, K. Y., Huang, T., Zhou, T. Y., Jing, Y. F., Li, W., Lin, Z., Buckler, E. S., Qian, Q. A., Zhang, Q. F., Li, J. Y. & Han, B. 2010. Genome-wide association studies of 14 agronomic traits in rice landraces. *Nature Genetics,* 42**,** 961-U76.

Huang, X. H., Zhao, Y., Wei, X. H., Li, C. Y., Wang, A., Zhao, Q., Li, W. J., Guo, Y. L., Deng, L. W., Zhu, C. R., Fan, D. L., Lu, Y. Q., Weng, Q. J., Liu, K. Y., Zhou, T. Y., Jing, Y. F., Si, L. Z., Dong, G. J., Huang, T., Lu, T. T., Feng, Q., Qian, Q., Li, J. Y. & Han, B. 2012.

Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nature Genetics,* 44**,** 32-U53.

Hughes, J., Hepworth, C., Dutton, C., Dunn, J. A., Hunt, L., Stephens, J., Waugh, R., Cameron, D. D. & Gray, J. E. 2017. Reducing Stomatal Density in Barley Improves Drought Tolerance without Impacting on Yield. *Plant Physiology,* 174**,** 776-787.

Hughes, R., Spielman, M., Schruff, M. C., Larson, T. R., Graham, I. A. & Scott, R. J. 2008. Yield assessment of integument-led seed growth following targeted repair of AUXIN RESPONSE FACTOR 2. *Plant Biotechnology Journal,* 6**,** 758-769.

Hunt, L. & Gray, J. E. 2009. The Signaling Peptide EPF2 Controls Asymmetric Cell Divisions during Stomatal Development. *Current Biology,* 19**,** 864-869.

Hwang, I., Sheen, J. & Muller, B. 2012. Cytokinin Signaling Networks. *In:* Merchant, S. S. (ed.) *Annual Review of Plant Biology, Vol 63.* Palo Alto: Annual Reviews.

Imai, K. K., Ohashi, Y., Tsuge, T., Yoshizumi, T., Matsui, M., Oka, A. & Aoyama, T. 2006. The A-type cyclin CYCA2;3 is a key regulator of ploidy levels in Arabidopsis endoreduplication. *Plant Cell,* 18**,** 382-396.

Irwin, J. A., Soumpourou, E., Lister, C., Ligthart, J. D., Kennedy, S. & Dean, C. 2016. Nucleotide polymorphism affecting FLC expression underpins heading date variation in horticultural brassicas. *Plant Journal,* 87**,** 597-605.

Iwakawa, H., Shinmyo, A. & Sekine, M. 2006. Arabidopsis CDKA;1, a cdc2 homologue, controls proliferation of generative cells in male gametogenesis. *Plant Journal,* 45**,** 819-831.

Jakoby, M., Weisshaar, B., Droge-Laser, W., Vicente-Carbajosa, J., Tiedemann, J., Kroj, T., Parcy, F. & B, Z. I. P. R. G. 2002. bZIP transcription factors in Arabidopsis. *Trends in Plant Science,* 7**,** 106-111.

Jiang, L. X., Yang, S. L., Xie, L. F., Puah, C. S., Zhang, X. Q., Yang, W. C., Sundaresan, V. & Ye, D. 2005. VANGUARD1 encodes a pectin methylesterase that enhances pollen tube growth in the Arabidopsis style and transmitting tract. *Plant Cell,* 17**,** 584-596.

Jones, B., Gunneras, S. A., Petersson, S. V., Tarkowski, P., Graham, N., May, S., Dolezal, K., Sandberg, G. & Ljung, K. 2010. Cytokinin Regulation of Auxin Synthesis in Arabidopsis Involves a Homeostatic Feedback Loop Regulated via Auxin and Cytokinin Signal Transduction. *Plant Cell,* 22**,** 2956-2969.

Kanaoka, M. M., Pillitteri, L. J., Fujii, H., Yoshida, Y., Bogenschutz, N. L., Takabayashi, J., Zhu, J. K. & Torii, K. U. 2008. SCREAM/ICE1 and SCREAM2 specify three cell-state transitional steps leading to Arabidopsis stomatal differentiation. *Plant Cell,* 20**,** 1775-1785.

Kaufholdt, D., Baillie, C. K., Bikker, R., Burkart, V., Dudek, C. A., Von Pein, L., Rothkegel, M., Mendel, R. R. & Hansch, R. 2016. The molybdenum cofactor biosynthesis complex interacts with actin filaments via molybdenum insertase Cnxl as anchor protein in Arabidopsis thaliana. *Plant Science,* 244**,** 8-18.

Kawashima, T. & Goldberg, R. B. 2010. The suspensor: not just suspending the embryo. *Trends in Plant Science,* 15**,** 23-30.

Kay, P., Groszmann, M., Ross, J. J., Parish, R. W. & Swain, S. M. 2013. Modifications of a conserved regulatory network involving INDEHISCENT controls multiple aspects of reproductive tissue development in Arabidopsis. *New Phytologist,* 197**,** 73-87.

Kepinski, S. & Leyser, O. 2005. The Arabidopsis F-box protein TIR1 is an auxin receptor. *Nature,* 435**,** 446-451.

Kesarwani, M., Yoo, J. M. & Dong, X. N. 2007. Genetic interactions of TGA transcription factors in the regulation of pathogenesis-related genes and disease resistance in Arabidopsis. *Plant Physiology,* 144**,** 336-346.

Kessler, S. A. & Grossniklaus, U. 2011. She's the boss: signaling in pollen tube reception. *Current Opinion in Plant Biology,* 14**,** 622-627.

Kosugi, S. & Ohashi, Y. 2002a. E2Ls, E2F-like repressors of Arabidopsis that bind to E2F sites in a monomeric form. *J Biol Chem,* 277**,** 16553-8.

Kosugi, S. & Ohashi, Y. 2002b. Interaction of the Arabidopsis E2F and DP proteins confers their concomitant nuclear translocation and transactivation. *Plant Physiol,* 128**,** 833-43.

Kuusk, S., Sohlberg, J. J., Long, J. A., Fridborg, I. & Sundberg, E. 2002. STY1 and STY2 promote the formation of apical tissues during Arabidopsis gynoecium development. *Development,* 129**,** 4707-4717.

Lai, L. B., Nadeau, J. A., Lucas, J., Lee, E. K., Nakagawa, T., Zhao, L. M., Geisler, M. & Sack, F. D. 2005. The Arabidopsis R2R3 MYB proteins FOUR LIPS and MYB88 restrict divisions late in the stomatal cell lineage. *Plant Cell,* 17**,** 2754-2767.

Lardon, A., Triboiblondel, A. M. & Dumas, C. 1993. A MODEL FOR STUDYING POLLINATION AND POD DEVELOPMENT IN BRASSICA-NAPUS - THE CULTURE OF ISOLATED FLOWERS. *Sexual Plant Reproduction,* 6**,** 52-56.

Larsson, E., Franks, R. G. & Sundberg, E. 2013. Auxin and the Arabidopsis thaliana gynoecium. *Journal of Experimental Botany,* 64**,** 2619-2627.

Lawson, S. S., Pijut, P. M. & Michler, C. H. 2014. Comparison of arabidopsis stomatal density mutants indicates variation in water stress responses and potential epistatic effects. *Journal of Plant Biology,* 57**,** 162-173.

Le, B. H., Cheng, C., Bui, A. Q., Wagmaister, J. A., Henry, K. F., Pelletier, J., Kwong, L., Belmonte, M., Kirkbride, R., Horvath, S., Drews, G. N., Fischer, R. L., Okamuro, J. K., Harada, J. J. & Goldberg, R. B. 2010. Global analysis of gene activity during Arabidopsis seed development and identification of seed-specific transcription factors. *Proceedings of the National Academy of Sciences of the United States of America,* 107**,** 8063-8070.

Lee, E., Lucas, J. R. & Sack, F. D. 2014. Deep functional redundancy between FAMA and FOUR LIPS in stomatal development. *Plant Journal,* 78**,** 555-565.

Lee, K., Avondo, J., Morrison, H., Blot, L., Stark, M., Sharpe, J., Bangham, A. & Coen, E. 2006. Visualizing plant development and gene expression in three dimensions using optical projection tomography. *Plant Cell,* 18**,** 2145-2156.

Lenser, T. & Theissen, G. 2013. Conservation of fruit dehiscence pathways between Lepidium campestre and Arabidopsis thaliana sheds light on the regulation of INDEHISCENT. *Plant Journal,* 76**,** 545-556.

Leshem, Y., Johnson, C., Wuest, S. E., Song, X. Y., Ngo, Q. A., Grossniklaus, U. & Sundaresan, V. 2012. Molecular Characterization of the glauce Mutant: A Central Cell-Specific Function Is Required for Double Fertilization in Arabidopsis. *Plant Cell,* 24**,** 3264-3277.

Li, Y. H., Zheng, L. Y., Corke, F., Smith, C. & Bevan, M. W. 2008. Control of final seed and organ size by the DA1 gene family in Arabidopsis thaliana. *Genes & Development,* 22**,** 1331-1336.

Liljegren, S. J., Ditta, G. S., Eshed, H. Y., Savidge, B., Bowman, J. L. & Yanofsky, M. F. 2000. SHATTERPROOF MADS-box genes control seed dispersal in Arabidopsis. *Nature,* 404**,** 766-770.

Liljegren, S. J., Roeder, A. H. K., Kempin, S. A., Gremski, K., Ostergaard, L., Guimil, S., Reyes, D. K. & Yanofsky, M. F. 2004. Control of fruit patterning in Arabidopsis by INDEHISCENT. *Cell,* 116**,** 843-853.

Lin, W. C., Lin, F. T. & Nevins, J. R. 2001. Selective induction of E2F1 in response to DNA damage, mediated by ATM-dependent phosphorylation. *Genes & Development,* 15**,** 1833-1844.

Lindsay, D. L., Sawhney, V. K. & Bonham-Smith, P. C. 2006. Cytokinin-induced changes in CLAVATA1 and WUSCHEL expression temporally coincide with altered floral development in Arabidopsis. *Plant Science,* 170**,** 1111-1117.

Liu, D., Li, W. C., Cheng, J. F. & Hou, L. 2014. Expression analysis and functional characterization of a cold-responsive gene COR15A from Arabidopsis thaliana. *Acta Physiologiae Plantarum,* 36**,** 2421-2432.

Liu, J., Hua, W., Hu, Z. Y., Yang, H. L., Zhang, L., Li, R. J., Deng, L. B., Sun, X. C., Wang, X. F. & Wang, H. Z. 2015. Natural variation in ARF18 gene simultaneously affects seed weight and silique length in polyploid rapeseed. *Proceedings of the National Academy of Sciences of the United States of America,* 112**,** E5123-E5132.

Liu, S. J., Jia, J. H., Gao, Y., Zhang, B. Y. & Han, Y. Z. 2010. The AtTudor2, a protein with SN-Tudor domains, is involved in control of seed germination in Arabidopsis. *Planta,* 232**,** 197-207.

Luo, G., Gu, H. Y., Liu, J. J. & Qu, L. J. 2012. Four Closely-related RING-type E3 Ligases, APD1-4, are Involved in Pollen Mitosis II Regulation in Arabidopsis. *Journal of Integrative Plant Biology,* 54**,** 814-827.

Macalister, C. A., Ohashi-Ito, K. & Bergmann, D. C. 2007. Transcription factor control of asymmetric cell divisions that establish the stomatal lineage. *Nature,* 445**,** 537-540.

Macleod, J. 1981. Havesting in Oilseed Rape. *Cambridge: Cambridge Agricultural Publishing***,** 107-120.

Magyar, Z. 2008. Keeping the Balance Between Proliferation and Differentiation by the E2F Transcriptional Regulatory Network is Central to Plant Growth and Development. *In:* Bogre, L. & Beemster, G. (eds.) *Plant Cell Monographs.* Springer-Verlag Berlin, Heidelberger Platz 3, D-14197 Berlin, Germany.

Magyar, Z., Atanassova, A., De Veylder, L., Rombauts, S. & Inze, D. 2000. Characterization of two distinct DP-related genes from Arabidopsis thaliana. *Febs Letters,* 486**,** 79-87.

Magyar, Z., De Veylder, L., Atanassova, A., Bako, L., Inze, D. & Bogre, L. 2005. The role of the Arabidopsis E2FB transcription factor in regulating auxin-dependent cell division. *Plant Cell,* 17**,** 2527-2541.

Magyar, Z., Horvath, B., Khan, S., Mohammed, B., Henriques, R., De Veylder, L., Bako, L., Scheres, B. & Bogre, L. 2012. Arabidopsis E2FA stimulates proliferation and endocycle separately through RBR-bound and RBR-free complexes. *EMBO J,* 31**,** 1480-93.

Makkena, S., Lee, E., Sack, F. D. & Lamb, R. S. 2012. The R2R3 MYB Transcription Factors FOUR LIPS and MYB88 Regulate Female Reproductive Development. *Journal of Experimental Botany,* 63**,** 5545-5558.

Mariconti, L., Pellegrini, B., Cantoni, R., Stevens, R., Bergounioux, C., Cella, R. & Albani, D. 2002. The E2F family of transcription factors from Arabidopsis thaliana. Novel and conserved components of the retinoblastoma/E2F pathway in plants. *J Biol Chem,* 277**,** 9911-9.

Marsch-Martinez, N., Ramos-Cruz, D., Reyes-Olalde, J. I., Lozano-Sotomayor, P., Zuniga-Mayo, V. M. & De Folter, S. 2012. The role of cytokinin during Arabidopsis gynoecia and fruit morphogenesis and patterning. *Plant Journal,* 72**,** 222-234.

Mcatee, P., Karim, S., Schaffer, R. & David, K. 2013. A dynamic interplay between phytohormones is required for fruit development, maturation, and ripening. *Frontiers in Plant Science,* 4.

Meakin, P. J. & Roberts, J. A. 1990. DEHISCENCE OF FRUIT IN OILSEED RAPE (BRASSICA-NAPUS L) .2. THE ROLE OF CELL-WALL DEGRADING ENZYMES AND ETHYLENE. *Journal of Experimental Botany,* 41**,** 1003-1011.

Michaels, S. D. & Amasino, R. M. 1999. FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell,* 11**,** 949-956.

Miller, C. N., Harper, A. L., Trick, M., Werner, P., Waldron, K. & Bancroft, I. 2016. Elucidation of the genetic basis of variation for stem strength characteristics in bread wheat by Associative Transcriptomics. *Bmc Genomics,* 17**,** 11.

Mironov, V., De Veylder, L., Van Montagu, M. & Inze, D. 1999. Cyclin-dependent kinases and cell division in plants - The nexus. *Plant Cell,* 11**,** 509-521.

Miura, K., Ohta, M., Nakazawa, M., Ono, M. & Hasegawa, P. M. 2011. ICE1 Ser403 is necessary for protein stabilization and regulation of cold signaling and tolerance. *Plant Journal,* 67**,** 269-279.

Morgan, C. L., Bruce, D. M., Child, R., Ladbrooke, Z. L. & Arthur, A. E. 1998. Genetic variation for pod shatter resistance among lines of oilseed rape developed from synthetic B-Napus. *Field Crops Research,* 58**,** 153-165.

Moubayidin, L. & Ostergaard, L. 2014. Dynamic Control of Auxin Distribution Imposes a Bilateral-to-Radial Symmetry Switch during Gynoecium Development. *Current Biology,* 24**,** 2743-2748.

Muller, B. & Sheen, J. 2008. Cytokinin and auxin interaction in root stem-cell specification during early embryogenesis. *Nature,* 453**,** 1094-U7.

Nadeau, J. A. & Sack, F. D. 2002. Control of stomatal distribution on the Arabidopsis leaf surface. *Science,* 296**,** 1697-1700.

Nahle, Z., Polakoff, J., Davuluri, R. V., Mccurrach, M. E., Jacobson, M. D., Narita, M., Zhang, M. Q., Lazebnik, Y., Bar-Sagi, D. & Lowe, S. W. 2002. Direct coupling of the cell cycle and cell death machinery by E2F. *Nature Cell Biology,* 4**,** 859-864.

Nakagami, H., Kawamura, K., Sugisaka, K., Sekine, M. & Shinmyo, A. 2002. Phosphorylation of retinoblastoma-related protein by the cyclin D/cyclin-dependent kinase complex is activated at the G1/S-phase transition in tobacco. *Plant Cell,* 14**,** 1847-1857.

Nakagami, H., Sekine, M., Murakami, H. & Shinmyo, A. 1999. Tobacco retinoblastoma-related protein phosphorylated by a distinct cyclin-dependent kinase complex with Cdc2/cyclin D in vitro. *Plant Journal,* 18**,** 243-252.

Nakajima, M., Shimada, A., Takashi, Y., Kim, Y. C., Park, S. H., Ueguchi-Tanaka, M., Suzuki, H., Katoh, E., Iuchi, S., Kobayashi, M., Maeda, T., Matsuoka, M. & Yamaguchi, I. 2006. Identification and characterization of Arabidopsis gibberellin receptors. *Plant Journal,* 46**,** 880-889.

Nemhauser, J. L., Feldman, L. J. & Zambryski, P. C. 2000. Auxin and ETTIN in Arabidopsis gynoecium morphogenesis. *Development,* 127**,** 3877-88.

Ngo, Q. A., Moore, J. M., Baskar, R., Grossniklaus, U. & Sundaresan, V. 2007. Arabidopsis GLAUCE promotes fertilization-independent endosperm development and expression of paternally inherited alleles. *Development,* 134**,** 4107-4117.

Niu, Y., Wu, G. Z., Ye, R., Lin, W. H., Shi, Q. M., Xue, L. J., Xu, X. D., Li, Y., Du, Y. G. & Xue, H. W. 2009. Global Analysis of Gene Expression Profiles in Brassica napus Developing Seeds Reveals a Conserved Lipid Metabolism Regulation with Arabidopsis thaliana. *Molecular Plant,* 2**,** 1107-1122.

Ogura, H. 1968. Studies on the new male-sterility in Japanese radish, with special reference to the utilization of this sterility towards the practical raising of hybrid seeds. *Mem. Fac. Agric. Kagoshima Univ,* 6**,** 39-78.

Ohashi-Ito, K. & Bergmann, D. C. 2006. Arabidopsis FAMA controls the final proliferation/differentiation switch during stomatal development. *Plant Cell,* 18**,** 2493-2505.

Ostergaard, L., Kempin, S. A., Bies, D., Klee, H. J. & Yanofsky, M. F. 2006. Pod shatter-resistant Brassica fruit produced by ectopic expression of the FRUITFULL gene. *Plant Biotechnology Journal,* 4**,** 45-51.

Ostergaard, L. & King, G. J. 2008. Standardized gene nomenclature for the Brassica genus. *Plant Methods,* 4.

Pagnussat, G. C., Yu, H. J., Ngo, Q. A., Rajani, S., Mayalagu, S., Johnson, C. S., Capron, A., Xie, L. F., Ye, D. & Sundaresan, V. 2005. Genetic and molecular identification of genes required for female gametophyte development and function in Arabidopsis. *Development,* 132**,** 603-614.

Pediconi, N., Ianari, A., Costanzo, A., Belloni, L., Gallo, R., Cimino, L., Porcellini, A., Screpanti, I., Balsano, C., Alesse, E., Gulino, A. & Levrero, M. 2003. Differential regulation of E2F1 apoptotic target genes in response to DNA damage. *Nature Cell Biology,* 5**,** 552-558.

Perry, S. E., Nichols, K. W. & Fernandez, D. E. 1996. The MADS domain protein AGL15 localizes to the nucleus during early stages of seed development. *Plant Cell,* 8**,** 1977-1989.

Petersen, M., Sander, L., Child, R., Vanonckelen, H., Ulvskov, P. & Borkhardt, B. 1996. Isolation and characterisation of a pod dehiscence zone-specific polygalacturonase from Brassica napus. *Plant Molecular Biology,* 31**,** 517-527.

Pillitteri, L. J., Sloan, D. B., Bogenschutz, N. L. & Torii, K. U. 2007. Termination of asymmetric cell division and differentiation of stomata. *Nature,* 445**,** 501-505.

Piszczek, E. & Gutman, W. 2007. Caspase-like proteases and their role in programmed cell death in plants. *Acta Physiologiae Plantarum,* 29**,** 391-398.

Polyn, S., Willems, A. & De Veylder, L. 2015. Cell cycle entry, maintenance, and exit during plant development. *Current Opinion in Plant Biology,* 23**,** 1-7.

Price, J. S., Hobson, R. N., Neale, M. A. & Bruce, D. M. 1996. Seed losses in commercial harvesting of oilseed rape. *Journal of Agricultural Engineering Research,* 65**,** 183-191.

Pritchard, J. K., Stephens, M., Rosenberg, N. A. & Donnelly, P. 2000. Association mapping in structured populations. *American Journal of Human Genetics,* 67**,** 170-181.

Qi, L. P., Mao, L., Sun, C. M., Pu, Y. Y., Fu, T. D., Ma, C. Z., Shen, J. X., Tu, J. X., Yi, B. & Wen, J. 2014. Interpreting the genetic basis of silique traits in Brassica napus using a joint QTL network. *Plant Breeding,* 133**,** 52-60.

Racedo, J., Gutierrez, L., Perera, M. F., Ostengo, S., Pardo, E. M., Cuenya, M. I., Welin, B. & Castagnaro, A. P. 2016. Genome-wide association mapping of quantitative traits in a breeding population of sugarcane. *Bmc Plant Biology,* 16**,** 16.

Rademacher, E. H., Moller, B., Lokerse, A. S., Llavata-Peris, C. I., Van Den Berg, W. & Weijers, D. 2011. A cellular expression map of the Arabidopsis AUXIN RESPONSE FACTOR gene family. *Plant Journal,* 68**,** 597-606.

Rajani, S. & Sundaresan, V. 2001. The Arabidopsis myc/bHLH gene ALCATRAZ enables cell separation in fruit dehiscence. *Current Biology,* 11**,** 1914-1922.

Ramirez, V., Agorio, A., Coego, A., Garcia-Andrade, J., Hernandez, M. J., Balaguer, B., Ouwerkerk, P. B. F., Zarra, I. & Vera, P. 2011. MYB46 Modulates Disease Susceptibility to Botrytis cinerea in Arabidopsis. *Plant Physiology,* 155**,** 1920-1935.

Ramirez-Parra, E., Frundt, C. & Gutierrez, C. 2003. A genome-wide identification of E2F-regulated genes in Arabidopsis. *Plant J,* 33**,** 801-11.

Ramirez-Parra, E., Xie, Q., Boniotti, M. B. & Gutierrez, C. 1999. The cloning of plant E2F, a retinoblastoma-binding protein, reveals unique and conserved features with animal G(1)/S regulators. *Nucleic Acids Research,* 27**,** 3527-3533.

Reiser, L. & Fischer, R. L. 1993. THE OVULE AND THE EMBRYO SAC. *Plant Cell,* 5**,** 1291-1301.

Reyes, J. C., Muro-Pastor, M. I. & Florencio, F. J. 2004. The GATA family of transcription factors in Arabidopsis and rice. *Plant Physiology,* 134**,** 1718-1732.

Reyes-Olalde, J. I., Zuniga-Mayo, V. M., Montes, R. a. C., Marsch-Martinez, N. & De Folter, S. 2013. Inside the gynoecium: at the carpel margin. *Trends in Plant Science,* 18**,** 644-655.

Reyes-Olalde, J. I., Zuniga-Mayo, V. M., Serwatowska, J., Montes, R. a. C., Lozano-Sotomayor, P., Herrera-Ubaldo, H., Gonzalez-Aguilera, K. L., Ballester, P., Ripoll, J. J., Ezquer, I., Paolo, D., Heyl, A., Colombo, L., Yanofsky, M. F., Ferrandiz, C., Marsch-Martinez, N. & De Folter, S. 2017. The bHLH transcription factor SPATULA enables cytokinin signaling, and both activate auxin biosynthesis and transport genes at the medial domain of the gynoecium. *Plos Genetics,* 13**,** 31.

Robles, P. & Pelaz, S. 2005. Flower and fruit development in Arabidopsis thaliana. *International Journal of Developmental Biology,* 49**,** 633-643.

Roeckel, P., Oancia, T. & Drevet, J. R. 1998. Phenotypic alterations and component analysis of seed yield in transgenic Brassica napus plants expressing the tzs gene. *Physiologia Plantarum,* 102**,** 243-249.

Roeder, A. H. K., Cunha, A., Ohno, C. K. & Meyerowitz, E. M. 2012. Cell cycle regulates cell type in the Arabidopsis sepal. *Development,* 139**,** 4416-4427.

Roeder, A. H. K., Ferrandiz, C. & Yanofsky, M. F. 2003. The role of the REPLUMLESS homeodomain protein in Patterning the Arabidopsis fruit. *Current Biology,* 13**,** 1630-1635.

Rossignol, P., Stevens, R., Perennes, C., Jasinski, S., Cella, R., Tremousaygue, D. & Bergounioux, C. 2002. AtE2F-a and AtDP-a, members of the E2F family of transcription factors, induce Arabidopsis leaf cells to re-enter S phase. *Mol Genet Genomics,* 266**,** 995-1003.

Salisbury, E. J. 1927. On the causes and ecological significance of stomatal frequency, with special reference to the woodland flora. *Philos Trans R Soc Lond B Biol Sci,* 216**,** 1-65.

Santner, A. & Estelle, M. 2009. Recent advances and emerging trends in plant hormone signalling. *Nature,* 459**,** 1071-1078.

Sauret-Gueto, S., Schiessl, K., Bangham, A., Sablowski, R. & Coen, E. 2013. JAGGED Controls Arabidopsis Petal Growth and Shape by Interacting with a Divergent Polarity Field. *Plos Biology,* 11.

Savidge, B., Rounsley, S. D. & Yanofsky, M. F. 1995. TEMPORAL RELATIONSHIP BETWEEN THE TRANSCRIPTION OF 2 ARABIDOPSIS MADS BOX GENES AND THE FLORAL ORGAN IDENTITY GENES. *Plant Cell,* 7**,** 721-733.

Schapire, A. L., Valpuesta, V. & Botella, M. A. 2006. TPR Proteins in Plant Hormone Signaling. *Plant Signaling & Behavior,* 1**,** 229-230.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J. Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P. & Cardona, A. 2012. Fiji: an open-source platform for biological-image analysis. *Nature Methods,* 9**,** 676-682.

Schluter, U., Muschak, M., Berger, D. & Altmann, T. 2003. Photosynthetic performance of an Arabidopsis mutant with elevated stomatal density (sdd1-1) under different light regimes. *Journal of Experimental Botany,* 54**,** 867-874.

Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods,* 9**,** 671-675.

Schranz, M. E., Quijada, P., Sung, S. B., Lukens, L., Amasino, R. & Osborn, T. C. 2002. Characterization and effects of the replicated flowering time gene FLC in Brassica rapa. *Genetics,* 162**,** 1457-1468.

Schruff, M. C., Spielman, M., Tiwari, S., Adams, S., Fenby, N. & Scott, R. J. 2006. The AUXIN RESPONSE FACTOR 2 gene of Arabidopsis links auxin signalling, cell division, and the size of seeds and other organs. *Development,* 133**,** 251-261.

Sekine, M., Ito, M., Uemukai, K., Maeda, Y., Nakagami, H. & Shinmyo, A. 1999. Isolation and characterization of the E2F-like gene in plants. *Febs Letters,* 460**,** 117-122.

Selenko, P., Sprangers, R., Stier, G., Buhler, D., Fischer, U. & Sattler, M. 2001. SMN Tudor domain structure and its interaction with the Sm proteins. *Nature Structural Biology,* 8**,** 27-31.

Sessions, A., Nemhauser, J. L., Mccoll, A., Roe, J. L., Feldmann, K. A. & Zambryski, P. C. 1997. ETTIN patterns the Arabidopsis floral meristem and reproductive organs. *Development,* 124**,** 4481-91.

Shaikhali, J., Barajas-Lopez, J. D., Otvos, K., Kremnev, D., Garcia, A. S., Srivastava, V., Wingsle, G., Bako, L. & Strand, A. 2012. The CRYPTOCHROME1-Dependent Response to Excess Light Is Mediated through the Transcriptional Activators ZINC FINGER PROTEIN EXPRESSED IN INFLORESCENCE MERISTEM LIKE1 and ZML2 in Arabidopsis. *Plant Cell,* 24**,** 3009-3025.

Sharpe, J., Ahlgren, U., Perry, P., Hill, B., Ross, A., Hecksher-Sorensen, J., Baldock, R. & Davidson, D. 2002. Optical projection tomography as a tool for 3D microscopy and gene expression studies. *Science,* 296**,** 541-545.

Shi, J. Q., Li, R. Y., Qiu, D., Jiang, C. C., Long, Y., Morgan, C., Bancroft, I., Zhao, J. Y. & Meng, J. L. 2009. Unraveling the Complex Trait of Crop Yield With Quantitative Trait Loci Mapping in Brassica napus. *Genetics,* 182**,** 851-861.

Shi, J. Q., Zhan, J. P., Yang, Y. H., Ye, J., Huang, S. M., Li, R. Y., Wang, X. F., Liu, G. H. & Wang, H. Z. 2015. Linkage and regional association analysis reveal two new tightly-linked major-QTLs for pod number and seed number per pod in rapeseed (Brassica napus L.). *Scientific Reports,* 5**,** 18.

Shikata, M., Matsuda, Y., Ando, K., Nishii, A., Takemura, M., Yokota, A. & Kohchi, T. 2004. Characterization of Arabidopsis ZIM, a member of a novel plant-specific GATA factor gene family. *Journal of Experimental Botany,* 55**,** 631-639.

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

Simonini, S., Bencivenga, S., Trick, M. & Ostergaarda, L. 2017. Auxin-Induced Modulation of ETTIN Activity Orchestrates Gene Expression in Arabidopsis. *Plant Cell,* 29**,** 1864-1882.

Simonini, S., Deb, J., Moubayidin, L., Stephenson, P., Valluru, M., Freire-Rios, A., Sorefan, K., Weijers, D., Friml, J. & Ostergaard, L. 2016. A noncanonical auxin-sensing mechanism is required for organ morphogenesis in Arabidopsis. *Genes & Development,* 30**,** 2286-2296.

Skoog, F. & Miller, C. O. 1957. Chemical regulation of growth and organ formation in plant tissues cultured in vitro. *Symposia of the Society for Experimental Biology,* 11.

Smertenko, A. & Bozhkov, P. V. 2014. Somatic embryogenesis: life and death processes during apicalbasal patterning. *Journal of Experimental Botany,* 65**,** 1343-1360.

Smyth, D. R., Bowman, J. L. & Meyerowitz, E. M. 1990. EARLY FLOWER DEVELOPMENT IN ARABIDOPSIS. *Plant Cell,* 2**,** 755-767.

Sorefan, K., Girin, T., Liljegren, S. J., Ljung, K., Robles, P., Galvan-Ampudia, C. S., Offringa, R., Friml, J., Yanofsky, M. F. & Ostergaard, L. 2009. A regulated auxin minimum is required for seed dispersal in Arabidopsis. *Nature,* 459**,** 583-U114.

Sornay, E., Forzani, C., Forero-Vargas, M., Dewitte, W. & Murray, J. a. H. 2015. Activation of CYCD7;1 in the central cell and early endosperm overcomes cell-cycle arrest in the Arabidopsis female gametophyte, and promotes early endosperm and embryo development. *Plant Journal,* 84**,** 41-55.

Sotelo-Silveira, M., Cucinotta, M., Chauvin, A. L., Montes, R. a. C., Colombo, L., Marsch-Martinez, N. & De Folter, S. 2013. Cytochrome P450 CYP78A9 Is Involved in Arabidopsis Reproductive Development. *Plant Physiology,* 162**,** 779-799.

Sozzani, R., Maggio, C., Varotto, S., Canova, S., Bergounioux, C., Albani, D. & Cella, R. 2006. Interplay between Arabidopsis activating factors E2Fb and E2Fa in cell cycle progression and development. *Plant Physiology,* 140**,** 1355-1366.

Steinfort, U., Fukai, S., Trevaskis, B., Glassop, D., Chan, A. & Dreccer, M. F. 2017a. Vernalisation and photoperiod sensitivity in wheat: The response of floret fertility and grain number is affected by vernalisation status. *Field Crops Research,* 203**,** 243-255.

Steinfort, U., Trevaskis, B., Shu, F. K., Bell, K. L. & Dreccer, M. F. 2017b. Vernalisation and photoperiod sensitivity in wheat: Impact on canopy development and yield components. *Field Crops Research,* 201**,** 108-121.

Stephenson, P., Baker, D., Girin, T., Perez, A., Amoah, S., King, G. J. & Ostergaard, L. 2010. A rich TILLING resource for studying gene function in Brassica rapa. *Bmc Plant Biology,* 10**,** 10.

Su, C., Zhang, C., Tecle, A., Fu, X., He, J., Song, J., Zhang, W., Sun, X., Ren, Y., Silvennoinen, O., Yao, Z., Yang, X., Wei, M. & Yang, J. 2015. Tudor staphylococcal nuclease (Tudor-SN), a novel regulator facilitating G1/S phase transition, acting as a co-activator of E2F-1 in cell cycle regulation. *J Biol Chem,* 290**,** 7208-20.

Sundaresan, V., Springer, P., Volpe, T., Haward, S., Jones, J. D. G., Dean, C., Ma, H. & Martienssen, R. 1995. PATTERNS OF GENE-ACTION IN PLANT DEVELOPMENT REVEALED BY ENHANCER TRAP AND GENE TRAP TRANSPOSABLE ELEMENTS. *Genes & Development,* 9**,** 1797-1810.

Sundberg, E. & Ostergaard, L. 2009. Distinct and Dynamic Auxin Activities During Reproductive Development. *Cold Spring Harbor Perspectives in Biology,* 1.

Sundstrom, J. F., Vaculova, A., Smertenko, A. P., Savenkov, E. I., Golovko, A., Minina, E., Tiwari, B. S., Rodriguez-Nieto, S., Zamyatnin, A. A., Valineva, T., Saarikettu, J., Frilander, M. J., Suarez, M. F., Zavialov, A., Stahl, U., Hussey, P. J., Silvennoinen, O., Sundberg, E., Zhivotovsky, B. & Bozhkov, P. V. 2009. Tudor staphylococcal nuclease is an evolutionarily conserved component of the programmed cell death degradome. *Nature Cell Biology,* 11**,** 1347-U198.

Swain, S., Myers, Z. A., Siriwardana, C. L. & Holt, B. F. 2017. The multifaceted roles of NUCLEAR FACTOR-Y in Arabidopsis thaliana development and stress responses. *Biochimica Et Biophysica Acta-Gene Regulatory Mechanisms,* 1860**,** 636-644.

Swain, S. M. & Singh, D. P. 2005. Tall tales from sly dwarves: novel functions of gibberellins in plant development. *Trends in Plant Science,* 10**,** 123-129.

Taniguchi, M., Sasaki, N., Tsuge, T., Aoyama, T. & Oka, A. 2007. ARR1 directly activates cytokinin response genes that encode proteins with diverse regulatory functions. *Plant and Cell Physiology,* 48**,** 263-277.

Tiwari, S. B., Hagen, G. & Guilfoyle, T. J. 2004. Aux/IAA proteins contain a potent transcriptional repression domain. *Plant Cell,* 16**,** 533-543.

Tsiatsiani, L., Van Breusegem, F., Gallois, P., Zavialov, A., Lam, E. & Bozhkov, P. V. 2011. Metacaspases. *Cell Death and Differentiation,* 18**,** 1279-1288.

Tsugama, D., Liu, S. K. & Takano, T. 2012. A bZIP Protein, VIP1, Is a Regulator of Osmosensory Signaling in Arabidopsis. *Plant Physiology,* 159**,** 144-155.

U, N. 1935. Genome analysis in Brassica with special reference to the experimental formation of B. napus and peculiar mode of fertilization. *Jap J Bot,* 7**,** 389-452.

Uemukai, K., Iwakawa, H., Kosugi, S., De Jager, S., Kato, K., Kondorosi, E., Murray, J. a. H., Ito, M., Shinmyo, A. & Sekine, M. 2005. Transcriptional activation of tobacco E2F is repressed by co-transfection with the retinoblastoma-related protein: cyclin D expression overcomes this repressor activity. *Plant Molecular Biology,* 57**,** 83-100.

Van Gelderen, K., Van Rongen, M., Liu, A. A., Otten, A. & Offringa, R. 2016. An INDEHISCENT-Controlled Auxin Response Specifies the Separation Layer in Early Arabidopsis Fruit. *Molecular Plant,* 9**,** 857-869.

Van Hautegem, T., Waters, A. J., Goodrich, J. & Nowack, M. K. 2015. Only in dying, life: programmed cell death during plant development. *Trends in Plant Science,* 20**,** 102-113.

Vandepoele, K., Raes, J., De Veylder, L., Rouze, P., Rombauts, S. & Inze, D. 2002. Genome-wide analysis of core cell cycle genes in Arabidopsis. *Plant Cell,* 14**,** 903-916.

Vandepoele, K., Vlieghe, K., Florquin, K., Hennig, L., Beemster, G. T. S., Gruissem, W., Van De Peer, Y., Inze, D. & De Veylder, L. 2005. Genome-wide identification of potential plant E2F target genes. *Plant Physiology,* 139**,** 316-328.

Vlachonasios, K. E., Thomashow, M. F. & Triezenberg, S. J. 2003. Disruption mutations of ADA2b and GCN5 transcriptional adaptor genes dramatically affect Arabidopsis growth, development, and gene expression. *Plant Cell,* 15**,** 626-638.

Vlieghe, K., Boudolf, V., Beemster, G. T., Maes, S., Magyar, Z., Atanassova, A., De Almeida Engler, J., De Groodt, R., Inze, D. & De Veylder, L. 2005. The DP-E2F-like gene DEL1 controls the endocycle in Arabidopsis thaliana. *Curr Biol,* 15**,** 59-63.

Von Groll, U., Berger, D. & Altmann, T. 2002. The subtilisin-like serine protease SDD1 mediates cell-to-cell signaling during arabidopsis stomatal development. *Plant Cell,* 14**,** 1527-1539.

Vrablova, M., Vrabl, D., Hronkova, M., Kubasek, J. & Santrucek, J. 2017. Stomatal function, density and pattern, and CO2 assimilation in Arabidopsis thaliana tmm1 and sdd1-1 mutants. *Plant Biology,* 19**,** 689-701.

Wallace, J. G., Bradbury, P. J., Zhang, N. Y., Gibon, Y., Stitt, M. & Buckler, E. S. 2014. Association Mapping across Numerous Traits Reveals Patterns of Functional Variation in Maize. *Plos Genetics,* 10.

Walton, L. J., Kurepin, L. V., Yeung, E. C., Shah, S., Emery, R. J. N., Reid, D. M. & Pharis, R. P. 2012. Ethylene involvement in silique and seed development of canola, Brassica napus L. *Plant Physiology and Biochemistry,* 58**,** 142-150.

Wang, S., Gu, Y. N., Zebell, S. G., Anderson, L. K., Wang, W., Mohan, R. & Dong, X. N. 2014. A Noncanonical Role for the CKI-RB-E2F Cell-Cycle Signaling Pathway in Plant Effector-Triggered Immunity. *Cell Host & Microbe,* 16**,** 787-794.

Wang, X. D., Chen, L., Wang, A. N., Wang, H., Tian, J. H., Zhao, X. P., Chao, H. B., Zhao, Y. J., Zhao, W. G., Xiang, J., Gan, J. P. & Li, M. T. 2016. Quantitative trait loci analysis and genome-wide comparison for silique related traits in Brassica napus. *Bmc Plant Biology,* 16**,** 15.

Weimer, A. K., Nowack, M. K., Bouyer, D., Zhao, X., Harashima, H., Naseer, S., De Winter, F., Dissmeyer, N., Geldner, N. & Schnittger, A. 2012. RETINOBLASTOMA RELATED1 Regulates Asymmetric Cell Divisions in Arabidopsis. *Plant Cell,* 24**,** 4083-4095.

Weingartner, M., Pelayo, H. R., Binarova, P., Zwerger, K., Melikant, B., De La Torre, C., Heberle-Bors, E. & Bogre, L. 2003. A plant cyclin B2 is degraded early in mitosis and its ectopic expression shortens G2-phase and alleviates the DNA-damage checkpoint. *Journal of Cell Science,* 116**,** 487-498.

Wells, R., Trick, M., Soumpourou, E., Clissold, L., Morgan, C., Werner, P., Gibbard, C., Clarke, M., Jennaway, R. & Bancroft, I. 2014. The control of seed oil polyunsaturate content in the polyploid crop species Brassica napus. *Molecular Breeding,* 33**,** 349-362.

Werner, T., Motyka, V., Laucou, V., Smets, R., Van Onckelen, H. & Schmulling, T. 2003. Cytokinin-deficient transgenic Arabidopsis plants show multiple developmental alterations indicating opposite functions of cytokinins in the regulation of shoot and root meristem activity. *Plant Cell,* 15**,** 2532-2550.

Wu, C. L., Zukerberg, L. R., Ngwu, C., Harlow, E. & Lees, J. A. 1995. In vivo association of E2F and DP family proteins. *Mol Cell Biol,* 15**,** 2536-46.

Wu, H., Mori, A., Jiang, X. S., Wang, Y. X. & Yang, M. 2006. The INDEHISCENT protein regulates unequal cell divisions in Arabidopsis fruit. *Planta,* 224**,** 971-979.

Wu, Z. L., Zheng, S. S. & Yu, Q. 2009. The E2F family and the role of E2F1 in apoptosis. *International Journal of Biochemistry & Cell Biology,* 41**,** 2389-2397.

Wynn, S., Twining, S. & Carter, R. 2016. AHDB Cereals and Oilseeds Final Harvest Summary.

Xiang, D. Q., Venglat, P., Tibiche, C., Yang, H., Risseeuw, E., Cao, Y. G., Babic, V., Cloutier, M., Keller, W., Wang, E., Selvaraj, G. & Datla, R. 2011. Genome-Wide Analysis Reveals Gene Expression and Metabolic Network Dynamics during Embryo Development in Arabidopsis. *Plant Physiology,* 156**,** 346-356.

Xie, Z. D., Lee, E., Lucas, J. R., Morohashi, K., Li, D. M., Murray, J. a. H., Sack, F. D. & Grotewold, E. 2010. Regulation of Cell Proliferation in the Stomatal Lineage by the Arabidopsis MYB FOUR LIPS via Direct Targeting of Core Cell Cycle Genes. *Plant Cell,* 22**,** 2306-2321.

Yadegari, R. & Drews, G. N. 2004. Female gametophyte development. *Plant Cell,* 16**,** S133-S141.

Yan, C., Yan, Z., Wang, Y., Yan, X. & Han, Y. 2014a. Tudor-SN, a component of stress granules, regulates growth under salt stress by modulating GA20ox3 mRNA levels in Arabidopsis. *J Exp Bot,* 65**,** 5933-44.

Yan, L. F., Cheng, X., Jia, R. L., Qin, Q. Q., Guan, L. P., Du, H. & Hou, S. W. 2014b. New phenotypic characteristics of three tmm alleles in Arabidopsis thaliana. *Plant Cell Reports,* 33**,** 719-731.

Yang, N., Lu, Y. L., Yang, X. H., Huang, J., Zhou, Y., Ali, F., Wen, W. W., Liu, J., Li, J. S. & Yan, J. B. 2014. Genome Wide Association Studies Using a New Nonparametric Model Reveal the Genetic Architecture of 17 Agronomic Traits in an Enlarged Maize Association Panel. *Plos Genetics,* 10**,** 14.

Yang, P., Shu, C., Chen, L., Xu, J. S., Wu, J. S. & Liu, K. D. 2012. Identification of a major QTL for silique length and seed weight in oilseed rape (Brassica napus L.). *Theoretical and Applied Genetics,* 125**,** 285-296.

Yang, W. C., Shi, D. Q. & Chen, Y. H. 2010. Female Gametophyte Development in Flowering Plants. *In:* Merchant, S., Briggs, W. R. & Ort, D. (eds.) *Annual Review of Plant Biology, Vol 61.* Palo Alto: Annual Reviews.

Yoshida, H., Nagata, M., Saito, K., Wang, K. L. C. & Ecker, J. R. 2005. Arabidopsis ETO1 specifically interacts with and negatively regulates type 2 1-aminocyclopropane-1-carboxylate synthases. *Bmc Plant Biology,* 5**,** 13.

Zhang, L. W., Yang, G. S., Liu, P. W., Hong, D. F., Li, S. P. & He, Q. B. 2011. Genetic and correlation analysis of silique-traits in Brassica napus L. by quantitative trait locus mapping. *Theoretical and Applied Genetics,* 122**,** 21-31.

Zhang, Y. L., Tessaro, M. J., Lassner, M. & Li, X. 2003. Knockout analysis of Arabidopsis transcription factors TGA2, TGA5, and TGA6 reveals their redundant and essential roles in systemic acquired resistance. *Plant Cell,* 15**,** 2647-2653.

Zhao, H., Wu, D., Kong, F., Lin, K., Zhang, H. & Li, G. 2016. The Arabidopsis thaliana Nuclear Factor Y Transcription Factors. *Front Plant Sci,* 7**,** 2045.

Zhao, Y. D. 2010. Auxin Biosynthesis and Its Role in Plant Development. *Annual Review of Plant Biology, Vol 61,* 61**,** 49-64.

Zhao, Y. D., Dai, X. H., Blackwell, H. E., Schreiber, S. L. & Chory, J. 2003. SIR1, an upstream component in auxin signaling identified by chemical genetics. *Science,* 301**,** 1107-1110.

Zheng, N., Fraenkel, E., Pabo, C. O. & Pavletich, N. P. 1999. Structural basis of DNA recognition by the heterodimeric cell cycle transcription factor E2F-DP. *Genes Dev,* 13**,** 666-74.

Zhong, R., Richardson, E. A. & Ye, Z. H. 2007. The MYB46 transcription factor is a direct target of SND1 and regulates secondary wall biosynthesis in Arabidopsis. *Plant Cell,* 19**,** 2776-2792.

# 9 Appendix

Appendix Table 1. STRUCTURE output, Q-values

| Type | Accession | Q1 | Q2 | Q3 | Q4 |
|---|---|---|---|---|---|
| Chinese | A34_Chu | 0.666 | 0.002 | 0.332 | 0.001 |
| Chinese | NIN | 0.644 | 0.004 | 0.351 | 0.001 |
| CHINESE | A188_Zho | 0.607 | 0.004 | 0.389 | 0 |
| Chinese | A187_Xia | 0.509 | 0.225 | 0.266 | 0 |
| Chinese | A144_She | 0.45 | 0.242 | 0.302 | 0.005 |
| Fodder rape | IB106_MoaR | 0.419 | 0.002 | 0.578 | 0 |
| Forage rape | IB50_DwaE | 0.295 | 0.002 | 0.701 | 0.001 |
| leafy vegetable | R_IB40_CouN | 0.402 | 0.003 | 0.594 | 0.001 |
| Siberian kale | D146_SIBB | 0.266 | 0.275 | 0.457 | 0.001 |
| Siberian Kale | IB15_BraS | 0.001 | 0.001 | 0.998 | 0 |
| Siberian Kale | IB65_GroGS | 0 | 0.001 | 0.999 | 0 |
| Spring OSR | D83_KAR | 0.146 | 0.77 | 0.08 | 0.003 |
| Spring OSR | IB110_N01D_1330 | 0.105 | 0.612 | 0.283 | 0 |
| Spring OSR | A18_Bro | 0.082 | 0.889 | 0.028 | 0.001 |
| Spring OSR | IB111_N02D_1952 | 0.059 | 0.805 | 0.136 | 0.001 |
| Spring OSR | D108_MON | 0.018 | 0.75 | 0.231 | 0.001 |
| Spring OSR | A135_Reg | 0.002 | 0.957 | 0.041 | 0.001 |
| Spring OSR | A155_Ste | 0.001 | 0.993 | 0.006 | 0 |
| Spring OSR | D134_RAPCR | 0.001 | 0.185 | 0.813 | 0 |
| Spring OSR | D173_TOP | 0.001 | 0.036 | 0.963 | 0.001 |
| Spring OSR | D181_WESD | 0.001 | 0.012 | 0.986 | 0.001 |
| Spring OSR | D31_CES | 0.001 | 0.006 | 0.993 | 0 |
| Spring OSR | D69_HAN | 0.001 | 0.003 | 0.995 | 0.001 |
| Spring OSR | R_A48_Dra | 0.001 | 0 | 0.998 | 0 |
| Spring OSR | IB49_Dup | 0 | 0.001 | 0.999 | 0 |
| Spring OSR | IB53_Erg | 0 | 0.001 | 0.999 | 0 |
| Swede | D182_WILR | 0.998 | 0.001 | 0.001 | 0 |
| Swede | IB141_SenNZ | 0.997 | 0.001 | 0.001 | 0.001 |
| Swede | D170_TIN | 0.994 | 0.002 | 0.004 | 0 |
| Swede | R_IB78_JauCV | 0.98 | 0.003 | 0.016 | 0.001 |
| Swede | IB73_Hug | 0.97 | 0.002 | 0.028 | 0 |
| Swede | IB184_Yor | 0.889 | 0.019 | 0.091 | 0 |
| synthetic | D128_Q10 | 0.036 | 0.254 | 0.427 | 0.283 |
| Winter fodder | IB121_Palu | 0.127 | 0.001 | 0.872 | 0 |
| Winter forage rape | IB23_Can | 0.255 | 0.002 | 0.742 | 0.001 |
| Winter OSR | IB164_Tai | 0.84 | 0.007 | 0.152 | 0.001 |
| Winter OSR | D7_APExGIN | 0.186 | 0.014 | 0.798 | 0.002 |
| Winter OSR | R_IB177_Vic | 0.177 | 0.003 | 0.807 | 0.013 |
| Winter OSR | IB150_SlaS | 0.135 | 0.001 | 0.864 | 0.001 |
| Winter OSR | D55_EUR | 0.103 | 0.006 | 0.89 | 0.001 |
| Winter OSR | IB151_SloK | 0.005 | 0.005 | 0.99 | 0.001 |
| Winter OSR | IB143_ShaxWin | 0.003 | 0.073 | 0.923 | 0.001 |
| Winter OSR | IB22_Cab | 0.003 | 0.009 | 0.953 | 0.034 |
| Winter OSR | A25_Cap | 0.002 | 0.209 | 0.789 | 0 |

| | | | | | |
|---|---|---|---|---|---|
| Winter OSR | D117_NOR | 0.002 | 0.002 | 0.996 | 0.001 |
| Winter OSR | D127_PRI | 0.002 | 0.001 | 0.996 | 0.001 |
| Winter OSR | A11_Bie | 0.001 | 0.996 | 0.003 | 0 |
| Winter OSR | A136_Roc | 0.001 | 0.996 | 0.003 | 0 |
| Winter OSR | A168_TEM | 0.001 | 0.959 | 0.04 | 0 |
| Winter OSR | A176_Ver | 0.001 | 0.943 | 0.056 | 0 |
| Winter OSR | A42_Dar | 0.001 | 0.878 | 0.12 | 0.001 |
| Winter OSR | D10_BAL | 0.001 | 0.686 | 0.313 | 0 |
| Winter OSR | D102_MAT | 0.001 | 0.272 | 0.727 | 0 |
| Winter OSR | D137_ROCxLIZ | 0.001 | 0.179 | 0.819 | 0.001 |
| Winter OSR | D26_CAPxMOH | 0.001 | 0.012 | 0.984 | 0.003 |
| Winter OSR | D39_COR | 0.001 | 0.006 | 0.993 | 0 |
| Winter OSR | D4_AMBxCOM | 0.001 | 0.005 | 0.994 | 0.001 |
| Winter OSR | D46_DIP | 0.001 | 0.004 | 0.994 | 0 |
| Winter OSR | D74_HURxNAV | 0.001 | 0.003 | 0.995 | 0.001 |
| Winter OSR | D75_INCxCON | 0.001 | 0.003 | 0.995 | 0.001 |
| Winter OSR | IB100_Maj | 0.001 | 0.002 | 0.997 | 0.001 |
| Winter OSR | IB125_POH285B | 0.001 | 0.002 | 0.997 | 0 |
| Winter OSR | IB130_Qui | 0.001 | 0.002 | 0.996 | 0.001 |
| Winter OSR | IB133_Ram | 0.001 | 0.001 | 0.998 | 0.001 |
| Winter OSR | IB16_Bra | 0.001 | 0.001 | 0.998 | 0 |
| Winter OSR | IB180_Vis | 0.001 | 0.001 | 0.998 | 0 |
| Winter OSR | IB21_Cab | 0.001 | 0.001 | 0.998 | 0 |
| Winter OSR | IB24_CanxCou | 0.001 | 0.001 | 0.998 | 0 |
| Winter OSR | IB29_Cat | 0.001 | 0.001 | 0.998 | 0 |
| Winter OSR | IB45_Dim | 0.001 | 0.001 | 0.997 | 0.001 |
| Winter OSR | IB57_Exc | 0.001 | 0.001 | 0.997 | 0.001 |
| Winter OSR | IB58_Exp | 0.001 | 0.001 | 0.997 | 0.001 |
| Winter OSR | IB60_Fla | 0.001 | 0.001 | 0.997 | 0.001 |
| Winter OSR | IB86_LemM | 0.001 | 0.001 | 0.997 | 0 |
| Winter OSR | TAP | 0.001 | 0 | 0.998 | 0 |
| Winter OSR | D131_Raf | 0 | 0.998 | 0.001 | 0 |
| Winter OSR | D6_APE | 0 | 0.998 | 0.001 | 0 |
| Winter OSR | D79_JETN | 0 | 0.17 | 0.83 | 0 |
| Winter OSR | IB02_AbuN | 0 | 0.001 | 0.999 | 0 |
| Winter OSR | IB139_Sam | 0 | 0.001 | 0.999 | 0 |
| Winter OSR | IB28_Cas | 0 | 0.001 | 0.999 | 0 |
| Winter OSR | IB37_ColxNic | 0 | 0.001 | 0.999 | 0 |
| Winter OSR | IB70_HanxGas | 0 | 0.001 | 0.998 | 0.001 |
| Winter OSR | IB91_LicxExp | 0 | 0.001 | 0.998 | 0 |
| Winter OSR | IB99_MadxRec | 0 | 0 | 0.999 | 0 |
| Winter OSR | IB120_Palm | 0.001 | 0.002 | 0.997 | 0 |
| Winter OSR/Swede | IB169_TeqxAra | 0.997 | 0.001 | 0.001 | 0 |

Appendix Table 2. Trait files used for the GWAS

| Accession | Valve Length 2014 | Pod Length 2014 | Pod Length 2015 | A_JCVI_18795 | C_ES940427 |
|---|---|---|---|---|---|
| A11_Bie | 5.6770 | 6.8310 | 54.5695 | 0.6418 | 0.3827 |
| A135_Reg | 5.5100 | 6.8490 | 63.6734 | 1.4180 | 0.5919 |
| A136_Roc | 6.6720 | 8.0040 | 70.0048 | 0.9557 | 0.9499 |
| A144_She | 6.4920 | 8.1250 | 54.8622 | 0.8358 | 0.6230 |
| A155_Ste | | | 63.1746 | 1.9476 | 0.3613 |
| A168_TEM | 6.3550 | 7.6160 | 72.3157 | 0.6522 | 0.6807 |
| A176_Ver | 5.1650 | 6.2440 | 70.7597 | 1.3270 | 0.2308 |
| A18_Bro | 8.0010 | 9.3820 | 85.5028 | 0.9555 | 0.8159 |
| A187_Xia | 6.4490 | | 93.4234 | 1.3290 | 1.8863 |
| A188_Zho | 11.4010 | | | 1.2551 | 1.4190 |
| A25_Cap | 7.4150 | 8.9100 | 60.1057 | 0.7500 | 0.5218 |
| A34_Chu | 4.1500 | 5.5950 | 54.6220 | 1.0602 | 1.1065 |
| A42_Dar | 8.0460 | 9.3270 | 73.0027 | 0.5379 | 0.8733 |
| D10_BAL | 5.7290 | 6.8490 | 70.7508 | 0.5996 | 0.2888 |
| D102_MAT | 5.4980 | 6.7760 | 64.3709 | 0.5104 | 1.3698 |
| D108_MON | 6.4000 | 7.9100 | 60.6148 | 0.6158 | 0.9283 |
| D117_NOR | 5.8120 | 7.1650 | 60.3133 | 0.7912 | 0.4718 |
| D127_PRI | 5.1170 | | 63.8817 | 0.7651 | 0.5636 |
| D128_Q10 | 3.6840 | 5.5020 | | 1.8743 | 0.0000 |
| D131_Raf | 7.7370 | 9.3830 | 82.7682 | 1.2684 | 0.7942 |
| D134_RAPCR | 6.5220 | 7.6930 | 67.5172 | 1.0369 | 0.3607 |
| D137_ROCxLIZ | 6.3980 | 8.2100 | 71.0014 | 0.9419 | 0.0786 |
| D146_SIBB | 6.3970 | 7.5720 | | 0.4695 | 0.8574 |
| D170_TIN | 5.3370 | 6.5750 | 54.4174 | 0.9556 | 0.6878 |
| D173_TOP | 6.1770 | 7.4920 | 67.5844 | 0.6198 | 1.2128 |
| D181_WESD | 6.0780 | 7.3320 | 61.3488 | 0.6022 | 0.7070 |
| D182_WILR | 5.4140 | 6.4600 | 49.6583 | 1.5113 | 0.0928 |
| D26_CAPxMOH | 6.2670 | 8.0470 | | 0.8583 | 0.5971 |
| D31_CES | 5.4720 | 6.6200 | 55.8776 | 0.9545 | 0.1811 |
| D39_COR | | | 61.1469 | 1.4869 | 0.0913 |
| D4_AMBxCOM | 7.3980 | 8.9490 | | 0.4873 | 0.3698 |
| D46_DIP | | | 48.2270 | 0.7754 | 0.5395 |
| D55_EUR | | | 74.8155 | 0.7983 | 1.5785 |
| D6_APE | 5.9220 | 7.3060 | 64.6359 | 1.2587 | 0.6369 |
| D69_HAN | 5.0670 | 6.1610 | 56.0973 | 1.7155 | 0.0796 |
| D7_APExGIN | | | 71.1022 | 0.7783 | 0.6318 |
| D74_HURxNAV | 5.2840 | 6.6120 | 61.7517 | 1.6495 | 0.2152 |
| D75_INCxCON | 5.8940 | 7.0760 | 69.6457 | 1.0190 | 0.4254 |
| D79_JETN | 8.1570 | 9.6340 | 81.0650 | 0.7123 | 1.1894 |
| D83_KAR | 6.6510 | 8.0300 | 61.7555 | 0.6746 | 0.8893 |
| IB02_AbuN | 5.7960 | 6.9880 | 70.8661 | 0.8483 | 0.7588 |
| IB100_Maj | 6.0730 | 7.6630 | 61.3696 | 1.3971 | 0.4666 |
| IB106_MoaR | 5.7330 | 7.3070 | 61.8976 | 1.1673 | 0.5623 |
| IB110_N01D_1330 | 6.0150 | 7.4940 | 69.0453 | 0.9040 | 0.6004 |

| | | | | | |
|---|---|---|---|---|---|
| IB111_N02D_1952 | 5.7250 | 6.6710 | 57.1327 | 0.9368 | 0.0889 |
| IB120_Palm | | | 67.2071 | 1.1085 | 0.9916 |
| IB121_Palu | 6.4200 | 7.7450 | 72.7571 | 0.9506 | 0.6039 |
| IB125_POH285B | 6.9290 | 8.1450 | 69.3997 | 1.0042 | 1.2226 |
| IB130_Qui | 3.9460 | 5.3070 | 49.8847 | 1.5208 | 0.0000 |
| IB133_Ram | 6.4640 | 7.7620 | 71.7303 | 0.6142 | 0.4273 |
| IB139_Sam | 6.5080 | 7.7100 | 65.4426 | 1.1390 | 0.6934 |
| IB141_SenNZ | 5.1870 | 6.2640 | 52.6978 | 1.9879 | 0.0000 |
| IB143_ShaxWin | 5.6780 | 7.0110 | 67.8736 | 1.1912 | 1.4394 |
| IB15_BraS | 5.8260 | 6.8520 | 61.3928 | 0.9053 | 0.5967 |
| IB150_SlaS | 5.9840 | 8.1780 | 68.0986 | 2.0433 | 1.4733 |
| IB151_SloK | 5.5760 | 6.9490 | 62.5047 | 0.8269 | 0.9205 |
| IB16_Bra | 6.9840 | 9.2110 | | 0.5695 | 0.9509 |
| IB164_Tai | 4.4520 | 5.8330 | 55.2963 | 0.8689 | 1.0881 |
| IB169_TeqxAra | 5.2730 | 7.0360 | 54.6353 | 1.1083 | 1.1566 |
| IB180_Vis | 6.6050 | 7.5860 | 65.7227 | 0.9666 | 1.1680 |
| IB184_Yor | 5.1530 | 7.0240 | 59.1577 | 2.0286 | 0.5405 |
| IB21_Cab | 5.4030 | 6.5540 | 64.2909 | 0.8153 | 0.3583 |
| IB22_Cab | 6.6840 | 8.2210 | 71.1606 | 1.4949 | 1.2764 |
| IB23_Can | 4.4480 | 5.5860 | 62.0082 | 1.2573 | 0.7348 |
| IB24_CanxCou | 6.8110 | 8.2680 | 75.4113 | 0.8772 | 0.5782 |
| IB28_Cas | 4.4060 | 5.5600 | 55.3411 | 0.8247 | 0.4303 |
| IB29_Cat | 6.8330 | 8.6160 | 72.0958 | 0.6854 | 0.8583 |
| IB37_ColxNic | 7.1870 | 8.5530 | 69.8543 | 0.5252 | 1.9489 |
| IB45_Dim | 6.8320 | 8.4490 | 86.0600 | 0.7380 | 1.4121 |
| IB49_Dup | 4.8420 | 5.7790 | 60.2413 | 0.9767 | 0.5560 |
| IB50_DwaE | 4.2470 | 5.0870 | | 2.2334 | 0.2825 |
| IB53_Erg | 5.4900 | 6.7470 | 61.8122 | 0.9592 | 0.7007 |
| IB57_Exc | 5.2960 | 6.2370 | 71.0590 | 0.8647 | 0.6016 |
| IB58_Exp | 6.1680 | 7.5940 | 68.0654 | 1.3212 | 0.5303 |
| IB60_Fla | 4.6170 | 5.6190 | 58.7632 | 1.0465 | 0.8047 |
| IB65_GroGS | 7.0660 | 8.2720 | 74.0306 | 0.7201 | 0.8198 |
| IB70_HanxGas | 7.9290 | 9.3090 | 82.8907 | 0.7621 | 1.0293 |
| IB73_Hug | 6.6680 | 8.1010 | 67.3483 | 0.8299 | 0.9623 |
| IB86_LemM | 5.6980 | 6.8970 | 67.8904 | 0.7594 | 0.3658 |
| IB91_LicxExp | 7.1540 | 8.3650 | 78.1210 | 0.7497 | 1.0432 |
| IB99_MadxRec | 6.5650 | 7.9180 | 78.9913 | 0.7818 | 0.6527 |
| NIN | 4.3750 | 6.0200 | 75.6731 | 1.9445 | 1.1274 |
| R_A48_Dra | 6.4290 | 7.7620 | 73.7274 | 0.3684 | 0.5980 |
| R_IB177_Vic | 6.5340 | 7.7820 | 62.8550 | 0.6675 | 0.2322 |
| R_IB40_CouN | 7.2370 | 8.5970 | 82.9304 | 0.8825 | 1.8419 |
| R_IB78_JauCV | | | 57.4881 | 0.9548 | 0.4429 |
| TAP | 6.2080 | | 75.0725 | 1.3503 | 0.4336 |

Appendix Table 2 continued

| Accession | AreaxSpP | Seed Area | Seeds per Pod | Pod Strength | Beak Length | Pod Width |
|---|---|---|---|---|---|---|
| A11_Bie | 118.52 | 6.66 | 17.79 | 7.66 | 1.15 | 0.42 |
| A135_Reg | 150.08 | 5.64 | 26.62 | 9.26 | 1.34 | 0.44 |
| A136_Roc | 128.07 | 5.49 | 23.34 | 3.04 | 1.33 | 0.43 |
| A144_She | 147.27 | 6.39 | 23.05 | 8.42 | 1.63 | 0.50 |
| A155_Ste | 144.87 | 5.91 | 24.50 | 10.12 | | |
| A168_TEM | 159.16 | 6.11 | 26.04 | 14.54 | 1.26 | 0.46 |
| A176_Ver | 119.91 | 5.85 | 20.49 | 2.47 | 1.08 | 0.35 |
| A18_Bro | 137.69 | 5.19 | 26.54 | 21.01 | 1.38 | 0.36 |
| A187_Xia | 140.25 | 8.50 | 16.50 | 8.76 | | 0.57 |
| A188_Zho | 157.44 | 7.18 | 21.94 | | 1.50 | 0.55 |
| A25_Cap | 162.02 | 6.80 | 23.82 | 6.56 | 1.45 | 0.47 |
| A34_Chu | 79.84 | 6.71 | 11.90 | 4.72 | 1.28 | 0.52 |
| A42_Dar | 135.76 | 5.55 | 24.47 | 7.63 | 1.12 | 0.41 |
| D10_BAL | 103.69 | 6.50 | 15.95 | 2.24 | 1.28 | 0.40 |
| D102_MAT | 121.88 | 5.91 | 20.63 | 1.97 | 1.51 | 0.41 |
| D108_MON | 132.24 | 5.33 | 24.81 | 6.21 | 1.35 | 0.54 |
| D117_NOR | 103.13 | 5.81 | 17.75 | 4.39 | | 0.40 |
| D127_PRI | 125.92 | 5.87 | 21.44 | 6.27 | 1.82 | 0.39 |
| D128_Q10 | | 7.36 | 5.70 | 5.73 | 1.65 | 0.51 |
| D131_Raf | 165.94 | 6.92 | 23.99 | 14.15 | 1.17 | 0.46 |
| D134_RAPCR | 147.56 | 5.29 | 27.90 | 7.58 | 1.81 | 0.42 |
| D137_ROCxLIZ | 138.90 | 5.90 | 23.55 | 4.79 | | 0.45 |
| D146_SIBB | | 5.03 | 23.12 | 7.08 | 1.18 | 0.35 |
| D170_TIN | | 5.03 | 20.45 | 7.93 | 1.24 | 0.36 |
| D173_TOP | 147.74 | 6.43 | 22.97 | 8.68 | 1.32 | 0.50 |
| D181_WESD | 174.19 | 5.57 | 31.29 | 7.82 | 1.25 | 0.42 |
| D182_WILR | | 5.13 | 19.90 | 2.46 | 1.05 | 0.41 |
| D26_CAPxMOH | 140.86 | 6.15 | 22.90 | 5.98 | 1.78 | 0.42 |
| D31_CES | 149.88 | 5.88 | 25.49 | 24.16 | 1.15 | 0.50 |
| D39_COR | 40.09 | 6.54 | 6.13 | 4.43 | | |
| D4_AMBxCOM | 153.59 | 6.63 | 23.17 | 9.14 | 1.55 | 0.46 |
| D46_DIP | 110.09 | 6.51 | 16.90 | 6.73 | 1.17 | 0.38 |
| D55_EUR | 180.57 | 6.57 | 27.50 | 21.60 | | |
| D6_APE | 126.07 | 6.11 | 20.65 | 9.14 | 1.38 | 0.46 |
| D69_HAN | 125.30 | 6.81 | 18.40 | 3.95 | 1.09 | 0.43 |
| D7_APExGIN | 81.45 | 8.57 | 9.50 | 7.82 | 1.37 | 0.44 |
| D74_HURxNAV | 127.35 | 5.28 | 24.12 | 5.72 | 1.33 | 0.41 |
| D75_INCxCON | 153.17 | 5.90 | 25.97 | 4.16 | 1.18 | 0.44 |
| D79_JETN | 141.74 | 5.85 | 24.25 | 7.84 | 1.48 | 0.41 |
| D83_KAR | 139.06 | 5.03 | 27.63 | 2.85 | 1.38 | 0.35 |
| IB02_AbuN | 155.38 | 5.97 | 26.04 | 7.70 | 1.19 | |
| IB100_Maj | 112.28 | 5.71 | 19.65 | 13.52 | 1.59 | 0.42 |
| IB106_MoaR | | 5.51 | 12.72 | 4.11 | 1.57 | 0.40 |
| IB110_N01D_1330 | 110.01 | 5.17 | 21.29 | 4.41 | 1.48 | 0.41 |

| | | | | | | |
|---|---|---|---|---|---|---|
| IB111_N02D_1952 | 102.41 | 6.02 | 17.02 | 6.95 | 0.95 | 0.44 |
| IB120_Palm | 117.36 | 6.03 | 19.45 | 5.77 | | |
| IB121_Palu | | 4.26 | 27.87 | 2.08 | 1.33 | 0.37 |
| IB125_POH285B | 127.59 | 5.59 | 22.82 | 5.17 | 1.22 | 0.39 |
| IB130_Qui | 53.28 | 6.97 | 7.64 | 4.40 | 1.36 | 0.41 |
| IB133_Ram | 107.55 | 5.73 | 18.77 | 5.66 | 1.30 | 0.36 |
| IB139_Sam | 116.92 | 5.27 | 22.17 | 4.29 | 1.20 | 0.37 |
| IB141_SenNZ | | 5.83 | 13.90 | 7.20 | 1.08 | 0.40 |
| IB143_ShaxWin | 160.37 | 7.04 | 22.78 | 3.89 | 1.33 | 0.47 |
| IB15_BraS | | 6.78 | 15.30 | 4.07 | 1.03 | 0.43 |
| IB150_SlaS | 127.20 | 6.38 | 19.95 | 7.23 | 2.19 | 0.47 |
| IB151_SloK | 116.35 | 5.70 | 20.42 | 5.89 | 1.37 | 0.41 |
| IB16_Bra | 159.16 | 6.03 | 26.39 | 8.78 | 2.23 | 0.42 |
| IB164_Tai | | 6.02 | 21.24 | 5.58 | 1.38 | 0.50 |
| IB169_TeqxAra | | 5.74 | 10.41 | 2.31 | 1.76 | 0.41 |
| IB180_Vis | 148.49 | 5.27 | 28.16 | 14.07 | 0.98 | 0.34 |
| IB184_Yor | | 4.39 | 20.30 | 5.48 | 1.87 | 0.36 |
| IB21_Cab | 142.13 | 5.72 | 24.84 | 7.10 | 1.15 | 0.46 |
| IB22_Cab | 162.22 | 6.53 | 24.85 | 12.05 | 1.54 | 0.50 |
| IB23_Can | | 6.38 | 7.87 | 1.86 | 1.14 | 0.38 |
| IB24_CanxCou | 142.22 | 6.04 | 23.55 | 5.11 | 1.46 | 0.41 |
| IB28_Cas | 92.49 | 7.48 | 12.37 | 7.80 | 1.15 | 0.43 |
| IB29_Cat | 145.35 | 6.40 | 22.70 | 9.81 | 1.78 | 0.45 |
| IB37_ColxNic | 162.91 | 6.10 | 26.72 | 11.00 | 1.37 | 0.40 |
| IB45_Dim | 169.46 | 6.32 | 26.80 | 8.46 | 1.62 | 0.39 |
| IB49_Dup | 139.67 | 5.55 | 25.17 | 7.62 | 0.94 | 0.40 |
| IB50_DwaE | | 5.39 | 7.34 | 3.95 | 0.84 | 0.32 |
| IB53_Erg | 117.15 | 4.84 | 24.22 | 5.85 | 1.26 | 0.33 |
| IB57_Exc | 107.96 | 5.98 | 18.06 | 15.85 | 0.94 | 0.40 |
| IB58_Exp | 165.89 | 7.58 | 21.90 | 8.49 | 1.43 | 0.42 |
| IB60_Fla | 109.10 | 7.98 | 13.68 | 7.71 | 1.00 | 0.44 |
| IB65_GroGS | | 6.65 | 19.95 | 8.31 | 1.21 | 0.48 |
| IB70_HanxGas | 155.80 | 6.08 | 25.62 | 11.59 | 1.38 | 0.44 |
| IB73_Hug | | 6.27 | 20.78 | 3.37 | 1.43 | 0.48 |
| IB86_LemM | 115.16 | 6.21 | 18.55 | 7.36 | 1.20 | 0.40 |
| IB91_LicxExp | 179.11 | 6.47 | 27.70 | 6.78 | 1.21 | 0.43 |
| IB99_MadxRec | 161.33 | 6.40 | 25.20 | 8.09 | 1.35 | 0.45 |
| NIN | 92.95 | 7.20 | 12.91 | 4.12 | 1.65 | 0.45 |
| R_A48_Dra | 114.92 | 5.39 | 21.34 | 8.29 | 1.33 | 0.34 |
| R_IB177_Vic | 95.87 | 7.13 | 13.45 | 6.03 | 1.25 | 0.37 |
| R_IB40_CouN | | 6.65 | 18.84 | 10.91 | 1.36 | |
| R_IB78_JauCV | | 5.10 | 22.58 | 3.92 | | 0.45 |
| TAP | | 6.35 | 21.88 | 4.42 | | 0.34 |

Appendix Table 3. Primers used in Chapter 5

| Primer Name | Primer Sequence | Description |
|---|---|---|
| ***B. na* primers** | | |
| 2F (Marlene) | GTCTTCTCAGATCGATTCCT | F to sequence BnaA3.DPb |
| A3_gDNA-R | ATAGTCACAGGCGTTGGAAT | R to sequence BnaA3.DPb |
| F12_B.na | TGCTTCTCAATTAGCTTCTGGC | F to sequence cDNA BnaA3DPb |
| B.naDPbA3qR3 | ACGGTCGCATTTGGAC | R to sequence cDNA BnaA3.DPb |
| R6_Bna | GTTCTTTACGGGAGAATCTTCAT | R at beginning of BnaA3.DPb to sequence promoter |
| B.na_03 | CTTCGTCGCCGTTTCTTC | F within gene upstream of DPb to sequence promoter |
| R_NinPromSeq | CATCATCGACACATTAACTGC | R in Nin promoter to sequence DPb promoter |
| ***B. rapa* primers** | | |
| B.rapa_1 | GGTGTTACTCTTCCATTTATCCTC | F sequencing primer to amplify TILLING line JI32202b and JI41223b |
| B.rapa_2 | GTCTTCACGCTTGACTTTATG | R sequencing primer to amplify TILLING line JI32202b and JI41223b |
| B.rapa_3 | CAGTTTAGCTTGAAAGGTCTCA | F sequencing primer to amplify TILLING line JI31875 and JI30971b |
| B.rapa_4 | GGAAGACCTCTCCACAGAATC | R sequencing primer to amplify TILLING line JI31875 and JI30971b |

Appendix Table 4. Primers used in Chapter 6

| Primer Name | Primer Sequence | Description |
|---|---|---|
| **Sequencing primers** | | |
| DPa_LP | GAATGACTCTCCGCAATCAAG | For sequencing of SALK_075546 |
| DPa_RP | ATGGGAATCATCAACCAGTTG | For sequencing of SALK_075546 |
| DPb_LP | TTTTGACTTCGAATCAGCCAC | For sequencing of SALK_066506c |
| DPb_RP | ACACAGATTGACCTGAAACCG | For sequencing of SALK_066506c |
| E2Fa_LP | TGAAACGGAGGTTTTGTTGAG | For sequencing of SALK_104099 |
| E2Fa_RP | TAGTCCCTTTGATGGTTCGTG | For sequencing of SALK_104099 |
| E2Fb_LP | TTGGATTCCTTCCATTTGATG | For sequencing of SALK_103138C |
| E2Fb_RP | GTGCCTTTACAGCTATCAGCG | For sequencing of SALK_103138C |
| E2Fc_LP | AGCCATCATGATATCATGAAGG | For sequencing of SALK_073459 |
| E2Fc_RP | AGAGAAGAGAGGAAACAGCGG | For sequencing of SALK_073459 |
| TSN1_LP | CCCAAAGCAAAAGTTGACTTG | For sequencing of SALK_045179C |
| TSN1_RP | CATTTGGCCGCACTATAACAC | For sequencing of SALK_045179C |
| TSN2_LP | TTTGTTACAAAATGGTTTGTATTGC | For sequencing of SALK_062222C |
| TSN2_RP | AGCCTTCACTCTCCCTTTGAG | For sequencing of SALK_062222C |
| Lbb1.3 | ATTTTGCCGATTTCGGAAC | Used with RP sequencing primers, amplifies if TDNA is present |
| **Yeast-2 hybrid primers** | | |
| E2FC-Y2H-F | GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGCC GCGACATCAAAC | |
| E2FC_Y2H_R3 | GGGGACCACTTTGTACAAGAAAGCTGGGTTCAGCTGTT GAAGTTGCTCCATAA | |
| 2DPb-Y2H-F | GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGACAA CTACTGGGTCTAATTCT | |
| 2DPb-Y2H-R | GGGGACCACTTTGTACAAGAAAGCTGGGTTCAATTCTC CGGCTTCAT | |
| E2FA-Y2H-F | GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGTCCG GTGTCGTACGATCTTC | |
| 2E2FA-Y2H-R | GGGGACCACTTTGTACAAGAAAGCTGGGTTCATCTCGG GGTTGAGTCAACA | |
| 2E2FB-Y2H-F | GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGTCTG AAGAAGTACCTCAACA | |
| E2FB-Y2H-R | GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAGCTAC CTGTAGGTGATCTC | |
| Dpa-Y2H-F | GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAGTA TGGAGATGGAGTTGTTTG | |
| Dpa-Y2H-R | GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAGCGA GTATCAATGGATCC | |
| Tudor1-Y2H-F | GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGCA ACGGGGGCTGAG | |
| Tudor1_R2_Y2H_SN1,2,3 | GGGGACCACTTTGTACAAGAAAGCTGGGTTTAGAACCG ACCACCTCCAAGTAC | |
| Tudor1_R_Y2H_missing24b | GGGGACCACTTTGTACAAGAAAGCTGGGTTTAGTAAGC ATGGTTCAAACTTCC | |
| Tudor2_F_Y2H_T, SN5 | GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAGG ATTCCCGAAGCTCATATTC | for truncated version of TSN2 |
| Tudor2_Y2H_R2 | GGGGACCACTTTGTACAAGAAAGCTGGGTTTACCCGCG ACCCGGTTT | |

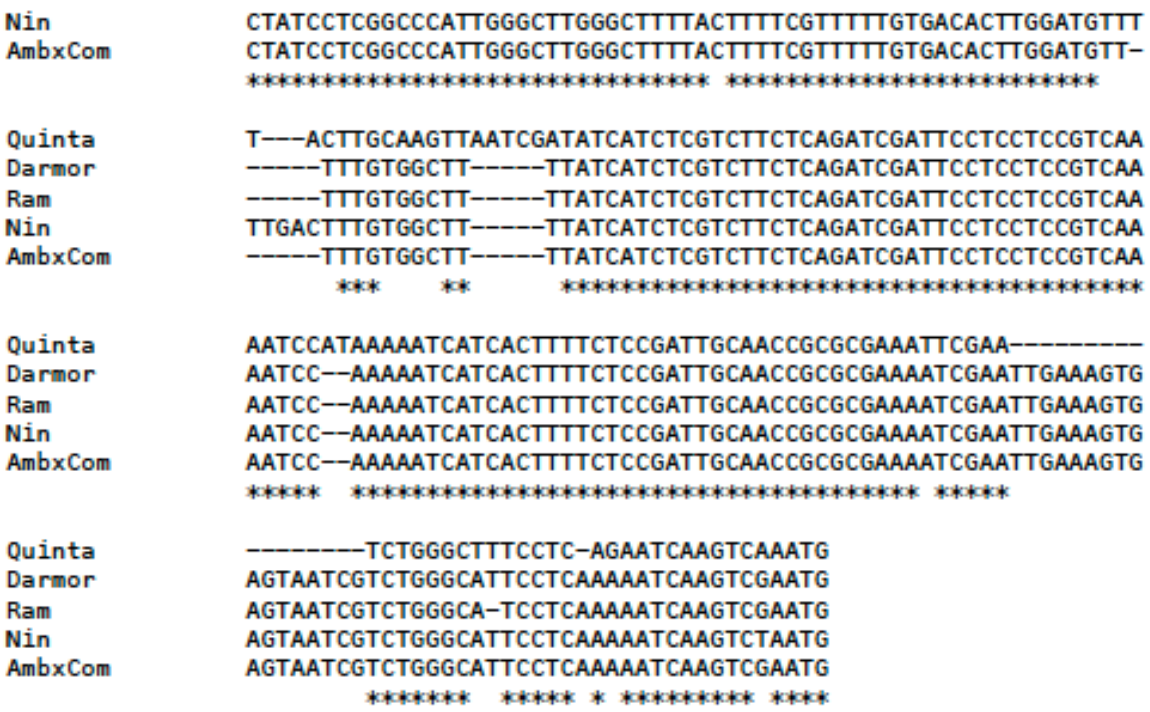
Table 4 continued

| Primer Name | Primer Sequence | Description |
|---|---|---|
| **qPCR primers** | | |
| AtDPa-q1 | AACCCTCACGCAGTAGTC | For qPCR |
| AtDPa-q2 | GCGAGTATCAATGGATCC | For qPCR |
| DPb_F1 | CCTTGTCCAGACTCGTCCTC | For qPCR |
| DPb_R1 | GTTCCGACCGTTTGGTTG | For qPCR |
| E2FA_1F | CTGATCCAGATGAAGCGGC | For qPCR |
| E2FA_1R | CCTTCAAATTCGCTGACGAG | For qPCR |
| F1A.th.E2FBqPCR | GACCAAATCAGAGAATCACAAG | For qPCR |
| R1A.th.E2FBqPCR | CCAGCCTCATCAGGATCT | For qPCR |
| E2Fc_F1 | CTGGAAGGGTGCTGACAATC | For qPCR |
| E2Fc_R3 | GGATCTGAGAGCTTCTTGTCG | For qPCR |
| T1_F11 | AAGGTAGTGGAAGTGGTGAGTG | For qPCR |
| T1_R9 | GGAGATCTGATACTCGAAAGACAG | For qPCR |
| Tudor2_F1 | GCTGGCCTGGCAAAAATG | For qPCR |
| Tudor2_R1 | AGAATATGAGCTTCGGGAATCCT | For qPCR |
| AtMC4-F2 | CGACATGAATCTGATTACTGATG | For qPCR |
| AtMC4_R2 | CAGATTCATCTTCGTCCTCG | For qPCR |
| PP2A_F2 | ACTGCATCTAAAGACAGAGTTCC | qPCR Housekeeping primer |
| PP2A_R2 | CCAAGCATGGCCGTATCATGT | qPCR Housekeeping primer |
| E1Fa_F | GTCGATTCTGGAAAGTCGACC | qPCR Housekeeping primer |
| E1Fa_R | AATGTCAATGGTGATACCACG | qPCR Housekeeping primer |
| **Cloning primers** | | |
| pE2Fb FWD | GGAAGACGGGGAGaaaggggttcttctattgttgtctc | Used to amplify E2Fb for the GUS reporter construct |
| pE2Fb REV | GGAAGACGGCATTaagaacgaatctcgataaaaaatcccc | Used to amplify E2Fb for the GUS reporter construct |

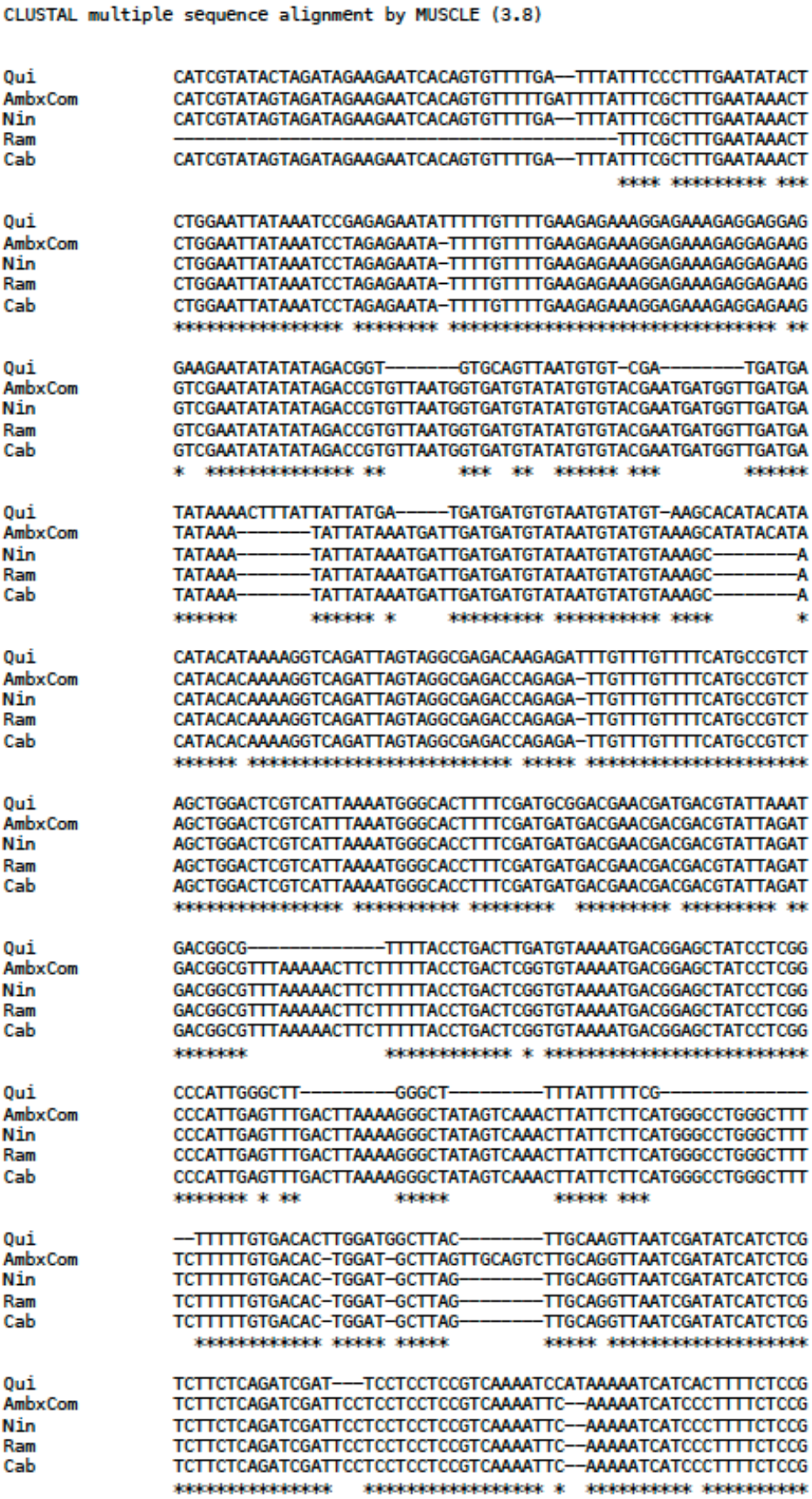
Appendix Fig.1. Alignment of the *B.naA3.DPb* promoters done using ClustalW multiple sequence alignment using Muscle. Final ATG is the beginning of the gene.

```
CLUSTAL multiple sequence alignment by MUSCLE (3.8)


Quinta          CGTATACTAGATAGAAGAATCACAGTGTTTTGATTTATTT----CCCTTTGAATATACTC
Darmor          CGTATACTAGATAGAAGAATCACAGTGTTTTGATTTATTTCCGGCCCTCTGAATATACTC
Ram             CGTATACTAGATAGAAGAATCACAGTGTTTTGATTTATTTCCGGCCCTCTGAATATACTC
Nin             CGTATACTAGATAGAAGAATCACAGTGTTTTGATTTATTT----CCCTTTGAATATACTC
AmbxCom         CGTATAGTAGACAGAAGAATCACAGTGTTTTGATTTATTT----CCCTCTGAATATACTC
                ****** **** ****************************    **** ***********

Quinta          TGGAATTATAAATCCGAGAGAATA--TTTTTGTTTTGAAGAGAAAGGAGAAAGAGGAGGA
Darmor          TGG-ATTATAAATCCTAGAGAATA-TTTTTTGTTTTGAAGAGAAAGGAGAAAGAGGAGGA
Ram             TGG-ATTATAAATCCTAGAGAATA-TTTTTTGTTTTGAAGAGAAAGGAGAAAGAGGAGGA
Nin             TGGAATTATAAATCCGAGAGAATA--TTTTTGTTTTGAAGAGAAAGGAGAAAGAGGAGGA
AmbxCom         TGGAATTATAAATCCTAGAGAATATTTTTTTTGTTTTGAAGAGAAAGGAGAAAGAGGAGGA
                *** *********** ********  ************************************

Quinta          GGAAGAATATAT-------------------------------------------------
Darmor          GGATGGATATAT-------------------------------------------------
Ram             GGATGGATATAT-------------------------------------------------
Nin             GGAAGAATATATATAGACGGTAGGGGTGGGCACTTTACCCGATATCCGAAGTGGCACCCG
AmbxCom         GGAAGAATATAT-------------------------------------------------
                *** * ******

Quinta          -------------------------------------------------------------
Darmor          -------------------------------------------------------------
Ram             -------------------------------------------------------------
Nin             AACCCGATCCGAAAAACCCTAACCGAAATCCGAACCGAAGTAGCAAAATACCCGAACGGG
AmbxCom         -------------------------------------------------------------

Quinta          -------------------------------------------------------------
Darmor          -------------------------------------------------------------
Ram             -------------------------------------------------------------
Nin             TATTGAATAAGGAGAGATTGGATACCCGAACCCGAACGGATAATATCCGAACCCGAATGA
AmbxCom         -------------------------------------------------------------

Quinta          -------------------------------------------------------------
Darmor          -------------------------------------------------------------
Ram             -------------------------------------------------------------
Nin             ATATCCGAAAATAACCGAACATATGTATAATTAACCTTATATTTCTAGTTTATATCTCTC
AmbxCom         -------------------------------------------------------------

Quinta          -------------------------------------------------------------
Darmor          -------------------------------------------------------------
Ram             -------------------------------------------------------------
Nin             ATTTTATATAAAATATTTATATTGATACTACACATACTTTAAGTTCATGTGATATAAATA
AmbxCom         -------------------------------------------------------------

Quinta          -------------------------------------------------------------
Darmor          -------------------------------------------------------------
Ram             -------------------------------------------------------------
Nin             CAATTACGGAAAAAAATGATTTGCTACTCATTTAAAATGCATGTCAAATTTTTTATTTCA
AmbxCom         -------------------------------------------------------------

Quinta          -------------------------------------------------------------
Darmor          -------------------------------------------------------------
Ram             -------------------------------------------------------------
Nin             AAAATTAACAAAAAGTTATATCCAAATTTTTTTTAAAAAAACTAAATTAATGCCTTTTTA
AmbxCom         -------------------------------------------------------------
```

```
Quinta          ------------------------------------------------------------
Darmor          ------------------------------------------------------------
Ram             ------------------------------------------------------------
Nin             GTTTTAAAATATTATATCCAAATCTATTAACCATTCAATCTACTAAAAATAAAAAATTAT
AmbxCom         ------------------------------------------------------------

Quinta          ------------------------------------------------------------
Darmor          ------------------------------------------------------------
Ram             ------------------------------------------------------------
Nin             TTAATTGAAAATTATATTTTAAATACAAGAAACTTGAGAAATGAAAATTTTAATTTTTTT
AmbxCom         ------------------------------------------------------------

Quinta          ------------------------------------------------------------
Darmor          ------------------------------------------------------------
Ram             ------------------------------------------------------------
Nin             TTTCAAAATCTAAATATCCGAACCCGATCCGAAATAACCGAATCCGAACTAAAAATACCC
AmbxCom         ------------------------------------------------------------

Quinta          ------------------------------------------------------------
Darmor          ------------------------------------------------------------
Ram             ------------------------------------------------------------
Nin             GAACCCGATCCGAAGTACAGAAATACCCGAACGGGTTCTACACCTCTATACCGAAATACC
AmbxCom         ------------------------------------------------------------

Quinta          ----------------------------------------------------------ATAG
Darmor          ----------------------------------------------------------ATAG
Ram             ----------------------------------------------------------ATAG
Nin             CGAAAATCCGAAATACCCGATCCGAACCCGAACGGGTACCCGAACGCCCACCCCTAATAG
AmbxCom         ----------------------------------------------------------ATAG
                                                                           ****

Quinta          ACGGTGTGCAGTTAATGTGTCGATGATGATATAAAACTTTATTATTATGATGATGATGTG
Darmor          ACCGTGTGCAGTTAATGTGTCGATGATGATATAAAACTTTATTATTATGATGATGATGTG
Ram             ACCGTGTGCAGTTAATGTGTCGATGATGATATAAAACTTTATTATTATGATGATGATGTG
Nin             ACGGTGTGCAGTTAATGTGTCGATGATGATATAAAACTTTATTATTATGATGATGATGTG
AmbxCom         ACGGTGTGCAGTTAATGTGTCGATGATGATATAAAACTTTATTATTATGATGATGATGTG
                ** ************************************************************

Quinta          TAATGTATGTAAGCACATACATACATACATAAAAGGTCAGATTAGTAGGCGAGACAAGAG
Darmor          TAATGTATGTAAGC----ACATACATACATGAAAGGTCAGATTAGTAGGCGAGACAAGAG
Ram             TAATGTATGTAAGC----ACATACATACATGAAAGGTCAGATTAGTAGGCGAGACAAGAG
Nin             TAATGTATGTAAGCACATACATACATACATAAAAGGTCAGATTAGTAGGCGAGACAAGAG
AmbxCom         TAATGTATGTAAGCACATACATACATACATAAAAGGTCAGATTAGTAGGCGAGACAAGAG
                ************** ********** ******************************

Quinta          ATTTGTTTGTTTTCATGCCGTCTAGCTGGACTCGTCATTAAAATGGGCACTTTTCGATGC
Darmor          AGTTGTTTGTTTTCATGCCGTCTAGCTGGACTCGTCATTAAAATGGGCACTTTTCGATGC
Ram             AGTTGTTTGTTTTCATGCCGTCTAGCTGGACTCGTCATTAAAATGGGCACTTTTCGATGC
Nin             ATTTGTTTGTTTTCATGCCGTCTAGCTGGACTCGTCATTAAAATGGGCACTTTTCGATGC
AmbxCom         ATTTGTTTGTTTTCATGCCGTCTAGCTGGACTCGTCATTAAAATGGGCACTTTTCGATGC
                * ************************************************************

Quinta          GGACGAACGATGACGTATTAAATGACGGCGTTTTACCTGACTTGATGTAAAATGACGGAG
Darmor          GGACGAACGATGACGTATTAAATGACGGCGTTTTACCTGACTTGATGTAAAATGACGGAG
Ram             GGACGAACGATGACGTATTAAATGACGGCGTTTTACCTGACTTGATGTAAAATGACGGAG
Nin             GGACGAACGATGACGTATTAAATGACGGCGTTTTACCTGACTTGATGTAAAATGACGGAG
AmbxCom         GGACGAACGATGACGTATTAAATGACGGCGTTTTACCTGACTTGATGTAAAATGACGGAG
                ************************************************************

Quinta          CTATCCTCGGCCCATTGGGCTTGGGCTTTTATTTTTCGTTTTTGTGACACTTGGATGGCT
Darmor          CTATCCTCGGCCCATTGGGCTTGGGCTTTTACTTTTCGTTTTTGTGACACTTGGATGTT-
Ram             CTATCCTCGGCCCATTGGGCTTGGGCTTTTACTTTTCGTTTTTGTGACACTTGGATGTT-
```


Nin CTATCCTCGGCCCATTGGGCTTGGGCTTTTACTTTTCGTTTTTGTGACACTTGGATGTTT
AmbxCom CTATCCTCGGCCCATTGGGCTTGGGCTTTTACTTTTCGTTTTTGTGACACTTGGATGTT-
Quinta T---ACTTGCAAGTTAATCGATATCATCTCGTCTTCTCAGATCGATTCCTCCTCCGTCAA
Darmor -----TTTGTGGCTT-----TTATCATCTCGTCTTCTCAGATCGATTCCTCCTCCGTCAA
Ram -----TTTGTGGCTT-----TTATCATCTCGTCTTCTCAGATCGATTCCTCCTCCGTCAA
Nin TTGACTTTGTGGCTT-----TTATCATCTCGTCTTCTCAGATCGATTCCTCCTCCGTCAA
AmbxCom -----TTTGTGGCTT-----TTATCATCTCGTCTTCTCAGATCGATTCCTCCTCCGTCAA
Quinta AATCCATAAAAATCATCACTTTTCTCCGATTGCAACCGCGCGAAATTCGAA---------
Darmor AATCC--AAAAATCATCACTTTTCTCCGATTGCAACCGCGCGAAAATCGAATTGAAAGTG
Ram AATCC--AAAAATCATCACTTTTCTCCGATTGCAACCGCGCGAAAATCGAATTGAAAGTG
Nin AATCC--AAAAATCATCACTTTTCTCCGATTGCAACCGCGCGAAAATCGAATTGAAAGTG
AmbxCom AATCC--AAAAATCATCACTTTTCTCCGATTGCAACCGCGCGAAAATCGAATTGAAAGTG
Quinta --------TCTGGGCTTTCCTC-AGAATCAAGTCAAATG
Darmor AGTAATCGTCTGGGCATTCCTCAAAAATCAAGTCGAATG
Ram AGTAATCGTCTGGGCA-TCCTCAAAAATCAAGTCGAATG
Nin AGTAATCGTCTGGGCATTCCTCAAAAATCAAGTCTAATG
AmbxCom AGTAATCGTCTGGGCATTCCTCAAAAATCAAGTCGAATG

Appendix Fig.2. Alignment of the *BnaC3.Dpb* promoters done using ClustalW multiple sequence alignment using Muscle. Final ATG is the beginning of the gene. The CAT at the start is the ATG of



the upstream gene which runs in reverse. Ram and AmbxCom were not fully sequenced.

```
Qui             ATTGCAACCGCGCGAAATTCGA----------------ATCTGGGCTTTCCTC-AGAATC
AmbxCom         ATTG------------------------------------------------------------
Nin             ATTGCATCC-CGCGAAATTCGAATTAAAGTCAGCAATCATCTGGGC-TTCCTCGAAAATC
Ram             ATTGCATCC-CGCGAAATTCGAATTAAAGTCAGCAATCATCTGGGCTTTCCTCGAAAATC
Cab             ATTGCATCC-CGCGAAATTCGAATTAAAGTCAGCAATCATCTGGGCTTTCCTCGAAAATC
                ****

Qui             AAGTCAAATG
AmbxCom         ----------
Nin             AAGTCGAATG
Ram             AAGTCGAATG
Cab             AAGTCGAATG
```

Appendix. ImageJ Macros

ImageJ macros were used to partially automate processes. They were used to measure *B. napus* and *B. oleracea* pod lengths, seed size in *Arabidopsis* and rosette areas of *Arabidopsis*.

Pod length macro:

Setting the scale depends on the pictures and was newly set every time for a new batch of images. To measure pods, a picture was taken from above where they were lined up along a white line (made using micropore). The macro then cropped the picture along that line to ensure the pedicels were not measured too. It was frequently checked that the cropped area correctly cropped only pedicels.

An example image:


Macro:

```
makeLine(1307, 724, 1405, 724);
run("Set Scale...", "distance=98.167 known=1 pixel=1 unit=unit");
//setTool("rectangle");
make
```

```
Rectangle(150, 687, 3219, 1008);
run("Crop");
run("Split Channels");
setAutoThreshold("Default");
//run("Threshold...");
setThreshold(0, 70);
setOption("BlackBackground", false);
run("Convert to Mask");
run("Fill Holes");
run("Dilate");
run("Erode");
run("Skeletonize");
run("Set Measurements...", "  bounding redirect=None decimal=3");
run("Analyze Particles...", "size=0.04-Infinity show=[Overlay Outlines] display");
```

Seed size macro:

```
setAutoThreshold("Default");
//run("Threshold...");
setThreshold(0, 148);
setOption("BlackBackground", false);
run("Convert to Mask");
```

Rosette area macro:

```
setAutoThreshold("Default dark");
//run("Threshold...");
setThreshold(141, 255);
setOption("BlackBackground", false);
run("Convert to Mask");
run("Analyze Particles...", "size=100-Infinity show=[Overlay Outlines] display");
selectWindow("20151028_111828-1.jpg (green)");
run("Duplicate...", "title=[20151028_111828-1.jpg (green)-2]");
```

Appendix Table 5. Silique lengths *e2f tsn* double mutants

| Line | MeanLengthNorm | StDev | StErr | Significance | Plants analysed |
|---|---|---|---|---|---|
| **Col-0** | 1.000 | 0.036 | 0.011 | / | 10 |
| ***e2fa$^{+/-}$ tsn1$^{-/-}$*** | 0.635 | 0.024 | 0.008 | p<0.001 | 8 |
| ***e2fa$^{-/-}$ tsn1$^{-/-}$*** | 0.930 | 0.091 | 0.032 | 0.001 | 8 |
| ***e2fb$^{+/-}$ tsn1$^{-/-}$*** | 0.569 | 0.041 | 0.015 | p<0.001 | 7 |
| ***e2fb$^{-/-}$ tsn1$^{+/-}$*** | 0.948 | 0.084 | 0.037 | 0.301 | 5 |
| ***e2fb$^{-/-}$ tsn1$^{-/-}$*** | 0.964 | 0.094 | 0.036 | 0.420 | 7 |
| ***e2fa$^{-/-}$ tsn2$^{-/-}$*** | 0.990 | 0.043 | 0.014 | 0.652 | 10 |
| ***e2fb$^{-/-}$ tsn2$^{+/-}$*** | 1.127 | | | 0.049 | 1 |
| ***e2fb$^{-/-}$ tsn2$^{-/-}$*** | 0.994 | 0.057 | 0.017 | 0.809 | 11 |
| ***e2fc$^{-/-}$ tsn2$^{-/-}$*** | 0.961 | | | 0.537 | 1 |

Appendix Table 6. Seeds per silique of *e2f tsn* double mutants

| Line | Seed per Silique (Norm to Col-0) | StDev | StErr | Significance |
|---|---|---|---|---|
| **Col-0** | 1.000 | 5.300 | 1.039 | / |
| ***e2fa$^{+/-}$ tsn1$^{-/-}$*** | 0.280 | 0.056 | 0.012 | <.001 |
| ***e2fa$^{-/-}$ tsn1$^{-/-}$*** | 1.009 | 0.082 | 0.009 | 0.78 |
| ***e2fb$^{+/-}$ tsn1$^{-/-}$*** | 0.220 | 0.033 | 0.008 | <.001 |
| ***e2fb$^{-/-}$ tsn1$^{+/-}$*** | 0.878 | 0.123 | 0.014 | <.001 |
| ***e2fb$^{-/-}$ tsn1$^{-/-}$*** | 0.954 | 0.065 | 0.006 | 0.134 |
| ***e2fa$^{-/-}$ tsn2$^{-/-}$*** | 0.973 | 0.058 | 0.013 | |
| ***e2fb$^{-/-}$ tsn2$^{+/-}$*** | | | | |
| ***e2fb$^{-/-}$ tsn2$^{-/-}$*** | 1.007 | 0.011 | 0.008 | 0.93643 |
| ***e2fc$^{-/-}$ tsn2$^{-/-}$*** | | | | |

Appendix Table 7. Silique lengths of overexpressing plants

| Line | Silique Length (Norm to Col-0) | StDev | StErr |
|---|---|---|---|
| **Col-0** | 100.000 | 4.268 | 1.350 |
| **HA-E2Fb/Dpa** | 99.612 | 4.034 | 1.276 |
| **HA-E2FbRBR/Dpa** | 92.920 | 4.822 | 1.607 |
| **DPb oe** | 96.925 | 4.461 | 1.821 |
| **E2FC oe** | 93.460 | 5.057 | 2.261 |

Appendix Table 8. Silique lengths of crosses. First genotype is the maternal plants, second is the pollen donor

| Line | Mean Length (cm) | StDev | StErr | Significance |
|---|---|---|---|---|
| **Col-0** | 1.488 | 0.127 | 0.020 | |
| **e2fc$^{-/-}$ tsn1$^{-/-}$** | 1.547 | 0.067 | 0.021 | |
| **e2fc$^{-/-}$ tsn1$^{-/-}$ x Col-0** | 1.172 | 0.027 | 0.012 | <0.001 |
| **Col-0 x e2fc$^{-/-}$ tsn1$^{-/-}$** | 1.622 | 0.315 | 0.129 | |
| **tsn1$^{-/-}$ tsn2 $^{-/-}$** | 0.979 | 0.067 | 0.015 | <0.001 |
| **tsn1$^{-/-}$ tsn2$^{-/-}$ x Col-0** | 1.086 | 0.056 | 0.023 | <0.001 |
| **Col-0 x tsn1$^{-/-}$ tsn2$^{-/-}$** | 1.655 | 0.267 | 0.134 | |

Appendix Table 9. Plant height

| Line | Plant Height (Norm to Col-0) | StDev | StErr |
|---|---|---|---|
| **Col-0** | 1.000 | 0.099 | 0.027 |
| ***e2fa*$^{-/-}$** | 1.122 | 0.183 | 0.058 |
| ***e2fb*$^{-/-}$** | 1.120 | 0.136 | 0.043 |
| ***e2fc*$^{-/-}$** | 1.056 | 0.180 | 0.052 |
| ***tsn1*$^{-/-}$** | 1.078 | 0.206 | 0.065 |
| ***tsn2*$^{-/-}$** | 1.084 | 0.093 | 0.033 |
| ***dpb*$^{-/-}$** | 0.956 | 0.098 | 0.031 |
| ***e2fa*$^{+/-}$ *e2fb*$^{+/-}$** | 0.980 | 0.074 | 0.033 |
| ***e2fa*$^{-/-}$ *e2fb*$^{-/-}$** | 0.935 | 0.116 | 0.039 |
| ***e2fa*$^{+/-}$ *e2fc*$^{+/-}$** | 1.147 | 0.264 | 0.108 |
| ***e2fb*$^{+/-}$ *e2fc*$^{+/-}$** | 0.938 | | |
| ***e2fb*$^{-/-}$ *e2fc*$^{-/-}$** | 0.980 | 0.115 | 0.038 |
| ***e2fa*$^{+/-}$ *tsn1*$^{+/-}$** | 1.097 | 0.121 | 0.040 |
| ***e2fa*$^{+/-}$ *tsn1*$^{-/-}$** | 1.036 | 0.135 | 0.061 |
| ***e2fa*$^{-/-}$ *tsn1*$^{-/-}$** | 0.903 | 0.157 | 0.056 |
| ***e2fb*$^{+/-}$ *tsn1*$^{+/-}$** | 1.160 | 0.149 | 0.047 |
| ***e2fb*$^{+/-}$ *tsn1*$^{-/-}$** | 1.006 | 0.102 | 0.042 |
| ***e2fb*$^{-/-}$ *tsn1*$^{+/-}$** | 0.916 | 0.178 | 0.079 |
| ***e2fb*$^{-/-}$ *tsn1*$^{-/-}$** | 1.007 | 0.094 | 0.036 |
| ***e2fc*$^{+/-}$ *tsn1*$^{+/-}$** | 0.945 | 0.184 | 0.082 |
| ***e2fc*$^{+/-}$ *tsn1*$^{-/-}$** | 0.842 | 0.128 | 0.074 |
| ***e2fa*$^{+/-}$ *tsn2*$^{+/-}$** | 1.024 | 0.107 | 0.036 |
| ***e2fc*$^{+/-}$ *tsn2*$^{+/-}$** | 1.164 | 0.187 | 0.076 |
| ***tsn1*$^{+/-}$ *tsn2*$^{+/-}$** | 1.124 | 0.206 | 0.073 |
| ***tsn1*$^{+/-}$ *tsn2*$^{-/-}$** | 1.050 | 0.104 | 0.074 |
| ***tsn1*$^{-/-}$ *tsn2*$^{+/-}$** | 1.036 | | |
| ***tsn1*$^{-/-}$ *tsn2*$^{-/-}$** | 0.920 | 0.005 | 0.003 |
| ***e2fa*$^{+/-}$*e2fb*$^{-/-}$*e2fc*$^{+/-}$** | 0.954 | 0.132 | 0.042 |
| ***e2fa*$^{+/-}$*e2fb*$^{+/-}$*tsn1*$^{+/-}$** | 0.937 | 0.186 | 0.083 |
| ***e2fb*$^{+/-}$*e2fc*$^{+/-}$*tsn1*$^{+/-}$** | 0.958 | 0.159 | 0.092 |
| ***e2fa*$^{+/-}$*e2fb*$^{-/-}$*tsn1*$^{+/-}$** | 0.883 | 0.118 | 0.044 |
| ***e2fb*$^{-/-}$*e2fc*$^{+/-}$*tsn1*$^{+/-}$** | 1.001 | 0.085 | 0.060 |
| ***e2fa*$^{+/-}$*e2fb*$^{+/-}$*tsn2*$^{+/-}$** | 0.930 | 0.083 | 0.059 |
| ***e2fb*$^{+/-}$*e2fc*$^{+/-}$*tsn2*$^{+/-}$** | 1.057 | 0.107 | 0.034 |
| ***e2fa*$^{+/-}$*tsn1*$^{-/-}$*tsn2*$^{+/-}$** | 1.011 | | |
| ***e2fb*$^{+/-}$*tsn1*$^{-/-}$*tsn2*$^{+/-}$** | 0.919 | | |
| ***e2fa*$^{+/-}$*e2fb*$^{+/-}$*tsn*$^{+/-}$*tsn2*$^{+/-}$** | 0.965 | | |
| ***e2fb*$^{+/-}$*e2fc*$^{+/-}$*tsn1*$^{+/-}$*tsn2*$^{+/-}$** | 1.140 | 0.111 | 0.064 |