

University of Nottingham
UK | CHINA | MALAYSIA

ALTERING MALE FERTILITY BY MANIPULATING PROTEASES INVOLVED IN POLLEN DEVELOPMENT

MUHAMAD FAHMI YUNUS (BSc, MSc)

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

2018

ABSTRACT

Male sterility and controlling fertility are considered as valuable traits that can be exploited in agriculture to improve crop productivity by selective breeding and hybrid development. Therefore, understanding the processes of anther and pollen development will not only contribute to the understanding of this process, but also to the development of new varieties for crop production. The formation of viable pollen relies on a specific cell layer in the anther, the tapetum, which is critical for pollen wall development and subsequently goes through programmed cell death (PCD) to facilitate deposition of pollen wall materials and viable pollen formation. Tapetum PCD requires the activity of specific proteases, which are associated with proteolytic breakdown. Seven protease genes and one F-Box protein that showed altered expression patterns in *Arabidopsis* mutants involved in pollen development have been analysed. These target genes consisted of six Aspartyl-protease (AP) genes (AT1G25510, AT2G23945, AT4G30030, AT5G10760, AT3G20015 and AT5G19120); one cysteine proteinase gene (AT2G21430) and an F-box protein, AT3G16210 gene. Defects in pollen development were seen associated with misexpression of AT2G23945. The work presented in this thesis shows that AT2G23945 is expressed in the anther and when mutated has impaired fertility due to impaired tapetum development. AT2G23945 T-DNA insertion- SALK 135523, showed defective pollen and anther development after analysis of fertility and pollen development using DAPI, Auramine O, ethidium bromide acridine orange, Alexander's staining and microscopy. To gain insight into the expression pattern of the AT2G23945 gene, transgenic plants carrying the AT2G23945pro:GUS and the AT2G23945pro:AT2G23945:GFP translational fusion were generated. Characterization of the AT2G23945-GUS activity in transgenic Arabidopsis plants indicated different regions of regulatory control in the upstream region. Young silique base is the only region of GUS expression for the construct directed by the short promoter. Constructs with a medium and long promoter resulted in expression in anther tissues as expected. No specific subcellular localization of pUBQ10:AT2G23945:YFP could be determined with expression in the plasma membrane and nucleus, which is identical to the control pUBQ10::YFP construct. Further analysis is needed for transgenic plants harbouring pGKGWG:AT2G23945:GFP.

ACKNOWLEDGEMENTS

In the name of God, the Most Gracious and the Most Merciful. Foremost, I would like to express my deep and sincere gratitude to my supervisor, Professor Dr. Zoe. A. Wilson for the continuous support of my PhD study and research. Her wide knowledge, understanding, encouraging and personal guidance have provided a good basis for the present thesis and have been of great value for me.

The financial support from the Ministry of Higher Education, Malaysia for me to undertake this study is gratefully acknowledged. I would like to thank Dr. Alison Ferguson, Dr. Jose Fernandez-Gomez and Dr. Ivana Ferjentsikova for their encouragement, insightful comments and suggestions. Special thanks to Dr. Alison for technical assistance with qRT-PCR and pGKGWG:AT2G23945:GFP additional analysis. I thank my fellow labmates in A47 Laboratory for the stimulating discussions, and for all the fun we have had in the last four years. I am also grateful to the staff of Plant and Crop Sciences Division and Glass House Team for providing the facilities to carry out the study. Last but not the least, I offer my regards and blessings to all of those who supported me in any respect during the completion of the project.

Dedicated to:

My dearest parents

Yunus Jamaludin
Zawiah Basnun

and

My Siblings

Muhamad Aqqat Yunus
Muhammad Ehsan Sabri Yunus

TABLE OF CONTENTS

TABLE OF CONTENTS	Page
ABSTRACT	i
ACKNOWLEDGEMENTS	ii
TABLE OF CONTENTS	iv
LIST OF TABLES	ix
LIST OF FIGURES	xi
ABBREVIATIONS	xxxi
CHAPTER 1: INTRODUCTION	
1.1 *Arabidopsis thaliana*	1
1.2 Reproductive System of *A. thaliana*	2
1.2.1 Flower	2
1.2.2 Anther	4
1.2.3 *Arabidopsis* pollen development	8
1.2.4 Tapetum	11
1.3 Programmed Cell Death	15
1.4 Regulatory Networks Involved in Plant Tapetum Development	20
1.5 Tapetum Transcription Factors Network	24
1.6 Proteases	34
1.6.1 Aspartyl proteases	36
1.6.2 Cysteine protease	39
1.7 F-Box Protein	40
1.8 Aims and Objectives	44
CHAPTER 2: MATERIALS AND METHODS	
2.1 Analysis of Tapetum Associated Genes	46
2.1.1 Plant materials and plant growth conditions.	46
2.1.2 Primers design.	46
2.1.3 RNA extraction and cDNA synthesis.	47
2.1.4 Genes cloning and Phusion PCR.	47
2.1.5 MicroElute Cycle-Pure spin protocol.	49
2.1.6 Overexpression vector, pUBQ10 vector with YFP-tagged protein	49
2.1.7 Gibson Assembly cloning reaction	50
2.1.8 Plasmid colony PCR	51
2.1.9 Plasmid extraction and sequencing	51
2.1.10 Agrobacterium electroporation	52
2.1.11 Transient expression of seven proteases genes in *Nicotiana benthamiana*	53

2.1.12 Observation of the Subcellular Localization of GFP and YFP-tagged Proteins 53
2.1.13 Detection of epidermis cell death by trypan blue staining 53
2.1.14 *Agrobacterium* mediated floral dip transformation 54
2.1.15 Putative transformants screening 55
2.1.16 Sucrose method for crude DNA extraction from plants 55
2.1.17 Putative transformants genotyping 56
2.1.18 Qualitative and quantitative RT–PCR (qRT–PCR) analysis of expression 56
2.1.19 Phenotypic analysis 58
2.1.20 Alexander's staining of pollen 58
2.1.21 Microscopy condition 59
2.1.22 Statistical analysis 60
2.2 Gene Manipulation Analysis of AT2G23945
2.2.1 Genotyping of SALK plants 60
2.2.2 Quantitative and qualitative PCR (qRT–PCR) analysis of AT2G23945 expression 61
2.2.3 Auramine O staining of pollen exine 62
2.2.4 Anther transverse sectioning study 62
2.2.5 Cell wall and lignification stain for endothecium thickening study 63
2.2.6 Developing pollen viability study 63
2.2.7 Cloning of pCR8/GW/TOPO AT2G23945RNAi using Gateway® system 64
2.2.8 Mach1™-T1R chemically competent *E. coli* transformation protocol 65
2.2.9 Gateway® LR clonase reaction of pK7GW1WG2:AT2G23945RNAi 65
2.2.10 Plasmid Colony PCR of pK7GW1WG2:AT2G23945RNAi 66
2.2.11 Putative transformants screening 66
2.2.12 Putative pK7GW1WG2:AT2G23945RNAi transformants genotyping. 66
2.3 Transcriptional and Translational Fusion of AT2G23945
2.3.1 Leaf DNA extraction protocol 67
2.3.2 Cloning of pCR8/GW/TOPO AT2G23945Pro entry clones with three different size of insertions. 67
2.3.3 pCR8/GW/TOPO AT2G23945Pro colony PCR and sequencing 68

2.3.4 Gateway® LR Clonase Reaction of pGWB3:AT2G23945PRO:GUS 69
2.3.5 pGWB3:AT2G23945PRO:GUS colony PCR and sequencing 69
2.3.6 Putative pGWB3:AT2G23945PRO:GUS transformants genotyping with three different size of insertions. 69
2.3.7 GUS staining method 70
2.3.8 Cloning of pCR8/GW/TOPO AT2G23945 translational fusion entry clones with three different size of insertions 71
2.3.9 pCR8/GW/TOPO AT2G23945 translational fusion colony PCR and sequencing 71
2.3.10 Gateway® LR Clonase Reaction of pGKGWG:AT2G23945:GFP 72
2.3.11 Putative pGKGWG:AT2G23945:GFP transformants genotyping with three different size of insertions. 72
2.3.12 qRT–PCR expression analysis 73

CHAPTER 3: ANALYSIS OF TAPETUM ASSOCIATED GENES
3.1 Introduction 74
3.2 Results 76
3.2.1 Candidate Gene Analysis 76
3.2.1.1 AT2G23945 Aspartyl Protease 76
3.2.1.2 AT1G25510 Aspartyl Protease 79
3.2.1.3 AT4G30030 Aspartyl Protease 80
3.2.1.4 AT5G10760 Aspartyl Protease 81
3.2.1.5 AT3G20015 Aspartyl Protease 83
3.2.1.6 AT5G19120 Aspartyl Protease 84
3.2.1.7 AT2G21430 Cysteine proteinase 85
3.2.1.8 AT3G16210 F-Box Protein 86
3.2.2 Cloning of targeted genes 87
3.2.3 Analysis of transient expression of the selected transgenes 90
3.2.4 Cell lysis analysis and trypan blue staining 93
3.2.5 Analysis of Protease Gene Function Through Overexpression 96
3.2.5.1 AT2G23945 Overexpression Plants 96
3.2.5.2 AT1G25510 Overexpression Plants 101

3.2.5.3 AT4G30030 Overexpression Plants 107
3.2.5.4 AT5G10760 Overexpression Plants 111
3.2.5.5 AT3G20015 Overexpression Plants 118
3.2.5.6 AT5G19120 Overexpression Plants 121
3.2.5.7 AT2G21430 Overexpression Plants 125
3.2.5.8 AT3G16210 Overexpression Plants 128
3.3 Discussion 134

Chapter 4: GENE MANIPULATION ANALYSIS OF AT2G23945
4.1 Introduction 140
4.2 Results 144
4.2.1 Expression analysis of AT2G23945 in different plant tissues using qRT-PCR. 144
4.2.2 Alteration of AT2G23945 gene expression using SALK T-DNA insertion lines 147
4.2.2.1 SALK 135523 148
4.2.2.2 SALK 134919 162
4.2.3 Manipulation of AT2G23945 gene expression by RNAi 167
4.3 Discussion 184

CHAPTER 5: TRANSCRIPTIONAL AND TRANSLATIONAL FUSION OF AT2G23945 GENE
5.1 Introduction 193
5.2 Results 197
5.2.1 Transcriptional analysis of PGWB3:AT2G23945PRO:GUS (-437 bp, -980 bp, -1572 bp) 197
5.2.2 Overexpression vector, pUBQ10: AT2G23945: YFP 214
5.2.3 Transformation using translational protein fusions for AT2G23945:GFP 215
5.3 Discussion 222

CHAPTER 6: GENERAL DISCUSSION
6.1 Summary of Main Findings 229
6.2 Role of AT2G23945 as a putative protease in PCD 234

6.3 Future research 235
6.3.1 Determination of Insertion Copies 235
6.3.2 Characterisation of the role of AT2G23945 in PCD 236
6.3.3 Subcellular localisation of the proteases 236
6.3.4 Terminal-deoxynucleotidyl transferase-mediated nick end labelling (TUNEL) assay 237
6.3.5 Post-translational protein modifications 238
6.3.6 Enzymatic characteristics of AT2G23945 239
6.3.7 Related Genes and PCD Indicator Genes Analysis 241
6.3.8 Transmission electron microscopy (TEM) 241
6.3.9 Reactive Active Oxygen (ROS) analysis 242
6.4 Conclusion 243

REFERENCES 244
APPENDICES
Appendix 1: List of primers used in this work. 261
Appendix 2: List of destination vectors generated in this work 264
Appendix 3: Full length sequences of the transcripts 265
Appendix 4: Preparation of Phosphate-buffered saline (PBS) 280
Appendix 5: Preparation of GUS substrate and buffers 281

LIST OF TABLES

Table 1-1. Advantages of *A. thaliana* for genome analysis based on Meinke *et al.*, 1998; Somerville and Koornneef, 2002. 3

Table 1-2. Summary of the key events during flower, anther and pollen development. 6

Table 1-3. Summary of proteases and F-Box protein which involved in plant developments. 43

Table 2-1: List of primers, PCR extension period and optimized annealing temperatures for Phusion reaction of seven selected proteases genes and one F-Box protein. 48

Table 2-2: List of primers, PCR extension period and optimized annealing temperatures for qRT-PCR of seven selected proteases genes and one F-Box protein. 58

Table 2-3. List of primers, PCR extension period and optimized annealing temperature for AT2G23945pro Phusion PCR. 68

Table 2-4. List of primers, PCR extension period and optimized annealing temperature for pCR8/GW/TOPO AT2G23945pro colony PCR. 68

Table 2-5: List of primers, PCR extension period and optimized annealing temperature for AT2G23945pro Phusion PCR. 71

Table 2-6. List of primers, PCR extension period and optimized annealing temperature for pCR8/GW/TOPO: AT2G23945 colony PCR. 72

Table 3-1. Summary of the analysis of fertility in the AT2G23945OE plants. 100

Table 3-2. Summary of fertility analysis in the AT1G25510OE line. 107

Table 3-3. Summary of fertility analysis in four independent T_1 AT4G30030OE plants and WT-Col. 110

Table 3-4. Summary of fertility analysis in five independent T_1 AT5G10760OE plants and WT-Col. 117

Table 3-5. Summary of fertility analysis in four independent T_1 AT3G20015OE plants and WT-Col. 121

Table 3-6. Summary of fertility analysis in three independent T_1 AT5G19120OE plants and WT-Col. 124

Table 3-7. Summary of fertility analysis in five independent T_1 AT2G21430OE plants and WT-Col. 128

Table 3-8. Summary of fertility analysis in five independent T_1 AT3G16210OE plants and WT-Col. 133

Table 3-9. Summary of fertility analysis in six different protease and one F-Box protein overexpression T_1 lines. 139

Table 4-1. Summary of the analysis of fertility in the SALK 135523 T-DNA insertion line. 162

Table 4-2. Summary of the analysis of fertility in the SALK 134919 T-DNA insertion line. 166

Table 4-3: Summary of the analysis of fertility in the AT2G23945RNAi (T_1) transgenic plants. 177

Table 4-4: Summary of the analysis of fertility in the AT2G23945RNAi (T_2) lines. 183

Table 4-5. Summary of fertility analysis in T_1 plants of SALK 135523, SALK 134919, AT2G23945OE and AT2G23945RNAi. 189

Table 5-1. Summary of the GUS expression in different tissues of plants harbouring the three different pGWB3:AT2G23945pro:GUS promoter fragments. 213

LIST OF FIGURES

Figure 1-1. Schematic diagram of *A. thaliana* flower. 4

Figure 1-2. *A. thaliana* anther development and dehiscence diagram. 5

Figure 1-3. Diagram of an anther showing the four cell layers surrounding the sporogenous cells. 8

Figure 1-4. The development of *Arabidopsis* reproductive system switches between microsporogenesis and microgametogenesis stages. 9

Figure 1-5. Schematic representation of pollen wall layer of Arabidopsis. 10

Figure 1-6. Diagram represents pollen maturation process from developing microspores and mature pollen. 13

Figure 1-7 Proposed tapetal PCD pathway. 21

Figure 1-8. Transcription factors regulatory network during tapetum formation, tapetum development and pollen wall development in *Arabidopsis*. 25

Figure 1-9. Tapetum transcription factors regulatory network for anther development, including pollen development stage in *A. thaliana* as proposed by Ferguson *et al.*, (2016). 28

Figure 1-10. Tapetum PCD and pollen formation regulatory pathways in *Arabidopsis* that have a high similarity with tapetum PCD and pollen formation in monocot rice regulatory pathway. 34

Figure 1-11. Aspartyl proteases activity: It is suggested that a tetrahedral intermediate is formed prior to the peptide bond cleavage. 37

Figure 1-12. Diagram of the F-box protein. 42

Figure 3-1. Level of expression of AT2G23945 in different tissues of *Arabidopsis*. 77

Figure 3-2. Floral microarrays data for expression of the AT2G23945 gene in multiple transcription factor mutants. 78

Figure 3-3. The expression of AT2G23945 gene in the *ams* mutant anther and wild type control. 78

Figure 3-4. Expression analysis of AT1G25510 in different Arabidopsis tissues. 79

Figure 3-5. The expression of AT1G25510 gene in the *ms1* and *ams* mutants and wild-type buds. 80

Figure 3-6. Expression analysis of AT4G30030 in different *Arabidopsis* tissues. 81

Figure 3-7. Expression of the AT4G30030 gene in buds of the *ms1* and *ams* mutants and wild-type. 81

Figure 3-8. Expression analysis of the AT5G10760 gene in different *Arabidopsis* tissues. 82

Figure 3-9. Expression of AT5G10760 gene in the *ms1* and *ams* mutant and wild-type buds. 82

Figure 3-10. Expression analysis of AT3G20015 in different *Arabidopsis* tissues. 83

Figure 3-11. Expression of the AT3G20015 gene in the *ms1* and *ams* mutant and wild-type buds. 83

Figure 3-12. Level of expression of AT5G19120 in different tissues of *Arabidopsis*. 84

Figure 3-13. Expression of the AT5G19120 gene in the *ms1* and *ams* mutant and wild-type buds. 84

Figure 3-14. Level of expression of AT2G21430 in different tissues of *Arabidopsis*. 85

Figure 3-15. Expression of the AT2G21430 gene in the *ms1* and *ams* mutant and wild-type buds. 85

Figure 3-16. Expression of the AT3G16210 gene in the *ms1* and *ams* mutants and wild-type buds. 86

Figure 3-17: Colony PCR screening of insert and vector sequence from destination vectors containing putative proteases involved with tapetum PCD. 88

Figure 3-18. Colony PCR screening of insert and vector sequence from *Agrobacterium* GV3101 containing putative proteases. 89

Figure 3-19. Localisation of transiently expressed YFP fusions for various putative proteases after *Agrobacterium* infiltration of *N. benthamiana*; GV3101::pUBQ10:Proteases Genes:YFP and GV3101::pUBQ10:YFP as a control. 91

Figure 3-20. Localisation of transiently expressed YFP fusions for five different proteases after infiltration; GV3101::pUBQ10:Proteases Genes: YFP in stomatal guard cells. 92

Figure 3-21. Ectopic expression and trypan blue staining of different protease genes; GV3101::pUBQ10:Proteases:YFP and GV3101::pUBQ10:YFP as a control. 94

Figure 3-22. Ectopic expression and trypan blue staining of different protease genes; GV3101::pUBQ10:Protease:YFP and GV3101::pUBQ10:YFP as a control. 95

Figure 3-23. The schematic map of pUBQ10:AT2G23945:YFP overexpression vector. 96

Figure 3-24. The presence of the insertion in twelve T_1 plants was confirmed by PCR analysis. 97

Figure 3-25. Expression pattern of AT2G23945 gene in six T_1 independent transgenic plants harbouring pUBQ10:AT2G23945:YFP. 98

Figure 3-26. Characterization of T_1 AT2G23945OE transgenic plants with phenotype. 99

Figure 3-27. Mean silique length (cm), indicating fertility, between six different overexpression transgenic plants OE2, OE3, OE6, OE7, OE9, OE11) and the WT-Col control after 6 weeks. 100

Figure 3-28. Schematic map of pUBQ10:AT1G25510:YFP overexpression vector. 101

Figure 3-29. PCR screening for overexpression transgene in six T_1 plants. 101

Figure 3-30. qRT-PCR expression analysis of AT1G25510 gene in four independent T_1 over expression transgenic lines harbouring the pUBQ10:AT1G25510:YFP construct. 102

Figure 3-31. Characterization of T_1 AT1G25510OE transgenic plants with different phenotypes. 103

Figure 3-32. Pollen viability assessment using Alexander's staining for all four AT1G25510OE plants. 104

Figure 3-33. Silique length (cm) of six different overexpression transgenic plants and the WT-Col after 6 weeks. 105

Figure 3-34. Analysis of pollen exine structure using Auramine O staining for OE7 and OE8 plants 106

Figure 3-35. Schematic map of pUBQ10:AT4G30030:YFP overexpression vector. 107

Figure 3-36. PCR to confirm presence of the transgene in eight T_1 plants. 108

Figure 3-37. qRT-PCR expression of AT4G30030 in four independent T_1 over expression transgenic lines harbouring the pUBQ10:AT4G30030:YFP construct. 108

Figure 3-38. Characterization of four T_1 AT4G30030OE transgenic plants with different phenotype. 109

Figure 3-39. Phenotyping analysis of silique length (cm) between four different overexpression transgenic plants and the WT-Col control after 6 weeks. 110

Figure 3-40. The schematic map of pUBQ10:AT5G10760:YFP overexpression vector. 111

Figure 3-41. The presence of the insertion in five T_1 plants was confirmed by PCR analysis. 111

Figure 3-42. qRT-PCR expression of AT5G10760 gene in five independent T_1 over expression transgenic lines harbouring the pUBQ10: AT5G10760:YFP construct. 112

Figure 3-43. Characterization of five T_1 AT5G10760OE plants with different phenotype. 113

Figure 3-44. Pollen viability assessment using Alexander's staining for all five AT5G10760OE plants. 114

Figure 3-45. Alexander's staining of pollen from OE3 plant compared to the WT-Col (A). 115

Figure 3-46. Phenotyping analysis of silique length (cm) between in five different overexpression lines and the WT-Col control after 6 weeks. 116

Figure 3-47. Pollen's exine layer assessment for OE3 plant using Auramine O Staining. 117

Figure 3-48. The schematic map of pUBQ10:AT3G20015:YFP overexpression vector. 118

Figure 3-49. PCR confirmation of the presence of pUBQ10:AT3G20015:YFP in 13 T_1 plants. 118

Figure 3-50. Expression pattern of AT3G20015 gene in inflorescence of five independent over expression transgenic plants harbouring pUBQ10:AT3G20015:YFP. 119

Figure 3-51. Images of six week-old plants for four T_1 AT3G20015OE plants. 120

Figure 3-52. Phenotyping analysis of silique length (cm) in four different T_1 AT3G20015OE lines and the WT-Col after 6 weeks. 120

Figure 3-53. The schematic map of pUBQ10:AT5G19120:YFP overexpression vector. 121

Figure 3-54. The presence of the insertion in eight T_1 plants was confirmed by PCR analysis. 122

Figure 3-55. Expression pattern of AT5G19120 gene in inflorescence of three independent over expression transgenic lines harbouring pUBQ10:AT5G19120:YFP. 122

Figure 3-56. Images of three T_1 AT5G19120OE plants at six weeks. 123

Figure 3-57. Phenotyping analysis of silique length (cm) of four different T_1 AT5G19120OE plants and the WT-Col after 6 weeks. 124

Figure 3-58. Schematic map of pUBQ10:AT2G21430:YFP overexpression vector. 125

Figure 3-59. The presence of the insertion in three T1 plants was confirmed by PCR analysis. 125

Figure 3-60. Expression pattern of AT2G21430 gene in inflorescence of five T_1 independent over expression transgenic lines harbouring pUBQ10:AT2G21430:YFP 126

Figure 3-61. Characterization of five T_1 AT2G21430OE plants. 127

Figure 3-62. Phenotyping analysis of silique length (cm) between four different AT2G21430OE plants and the WT-Col after 6 weeks. 127

Figure 3-63. Schematic map of pGWB5:AT3G16210:GFP overexpression vector. 129

Figure 3-64: Amplification of the pCR8/GW/TOPO entry clone with AT3G16210 gene. 129

Figure 3-65: Screening for AT3G16210 destination vector. 130

Figure 3-66: PCR amplification of pGWB5:AT3G16210:GFP (~1.4kbp band) in eight T1 plants. 130

Figure 3-67. Expression pattern of AT3G16210 gene in inflorescence of five independent over expression transgenic plants harbouring pUBQ10:AT3G16210:YFP. 131

Figure 3-68. Images of six weeks old plants of five T_1 AT3G16210OE transgenic lines. 132

Figure 3-69. Phenotypic analysis of silique length (cm) in five different AT3G16210OE lines and WT-Col after 6 weeks. 136

Figure 4-1. Amino acid sequence of AT2G23945 as provided by MEROPS database. 141

Figure 4-5. Mechanism involved in RNAi interference. 143

Figure 4-6. Multiple stages of buds that were used for the anther studies. 144

Figure 4-7. qRT-PCR analysis of expression of a) *PP2A3* Housekeeping gene and b) AT2G23945, in staged buds. 146

Figure 4-8. Comparison of AT2G23945 expression by qRT-PCR in WT-Col in different tissues. 146

Figure 4-9. Schematic map of AT2G23945 gene with T-DNA insertions. 148

Figure 4-10. PCR genotyping of three SALK 135523 T-DNA insertion plants in three different panels where A: SALK 135523 P1 (P1), B: SALK 135523 P2 (P2). and C: SALK 135523 P3 (P3). 149

Figure 4-11. Expression analysis of AT2G23945 in SALK 135523 plants using RT-PCR. A: *PP2A3* expression analysis. B: AT2G23945 expression. 150

Figure 4-12. Expression pattern of AT2G23945 gene in the three homozygous SALK 135523 plants. 151

Figure 4-13. Phenotypic characterization of T-DNA insertion plants, showing the main stem of WT-Col and the three homozygous SALK 135523 plants (P1-3)(B-D). 152

Figure 4-14. Normal dehiscence was observed for all three SALK 135523 plants. 153

Figure 4-15. Pollen viability assessment using Alexander's staining for all three SALK 135523 plants. 154

Figure 4-16. Mean silique length (cm), indicating fertility, between the three overexpression SALK 135523 plants (P1-3) and the WT-Col control after 6 weeks. 154

Figure 4-17. Seed development of different SALK 135523 plants fertile silique and WT-Col plant fertile silique. 155

Figure 4-18. Pollen anlaysis of exine structure using Auramine O staining for three SALK 135523 plants. 156

Figure 4-19. Phenotypic analysis of anther of the WT-Col and SALK 135523 plants with equivalent stages of pollen development. 158

Figure 4-20. Enlarged images of bicellular and mature anthers from Figure 4-18. 159

Figure 4-21. Analysis of lignification of the WT-Col and SALK 135523 (P1-P3 plants) anthers using ethidium bromide acridine orange staining. 160

Figure 4-22. DAPI staining of nuclear development of pollen from the WT-Col and SALK 135523 plants. 161

Figure 4-23. Genotyping of SALK 134919 T-DNA insertion plants (P1-P8). 163

Figure 4-24. Expression analysis of AT2G23945 in buds from multiple SALK 134919 plants by RT-PCR. 163

Figure 4-25. qRT-PCR analysis of AT2G23945 in buds from three homozygous SALK 134919 plants. 164

Figure 4-26. Main stem of three different SALK 134919 plants (P1-P3) and WT-Col control of 6 weeks old plants. 165

Figure 4-27. Mean silique length (cm), indicating fertility, between the three overexpression SALK 134919 plants (P1, P5 and P6) and the WT-Col control after 6 weeks. 166

Figure 4-28. Colony PCR of insert and vector sequences from pCR8/GW/TOPO entry clone which contains 102bp AT2G23945RNAi target insert. 167

Figure 4-29. Schematic map of pK7GW1WG2:AT2G23945RNAi construct for generating RNAi mutant plants. 168

Figure 4-30. Colony PCR of pK7GWIWG2(I): AT2G23945RNAi destination vector. 169

Figure 4-31. Colony PCR of GV3101 containing pK7GWIWG2(I): AT2G23945RNAi colonies. 169

Figure 4-32. Selection of putative transgenic plants and PCR screening of transgenic plants. 170

Figure 4-33. Characterization of AT2G23945RNAi transgenic main stems with different severities of phenotype. 172

Figure 4-34. Characterization of AT2G23945RNAi inflorescence of six different T_1 AT2G23945RNAi transgenic plants (B-G) as comparison with the WT- Col (A) from eight weeks old plants. 173

Figure 4-35. Pollen viability assessment using Alexander's staining of anthers from closed buds of seven weeks old plants. 175

Figure 4-36. Alexander's staining of pollen from the two different plants from the plants that showed highest level of severity of pollen development (C-D: P4; E-F: P7) as 176

compared to the WT-Col (A and B) where percentage of abnormal pollens outnumbered normal pollens.

Figure 4-37. Mean silique length (cm), indicating fertility between six AT2G23945RNAi plants (P1, P2, P3, P4, P5 and P7) and the WT-Col control after 8 weeks. 177

Figure 4-38. qRT-PCR analysis of AT2G23945 in inflorescence from 21 AT2G23945RNAi mutant plants. 180

Figure 4-39. Mean silique length (mm), indicating fertility, between nineteen plants and the WT-Col controls 181

4-40. Mean silique length upon removal of stressed siliques (mm), indicating fertility, between the nineteen plants and the WT-Col controls. 182

Figure 5-1. The schematic map of pCR8/GW/TOPO:AT2G23945Pro entry clone with three different promoter fragments. 198

Figure 5-2. Colony PCR of insert and vector sequences from pCR8/GW/TOPO entry clones for the three different sizes of AT2G23945 promoter. 199

Figure 5-3. The schematic map of pGWB3:AT2G23945PRO:GUS destination vector with three different promoter fragments. 200

Figure 5-4. Colony PCR of insert and vector sequences from pGWB3:AT2G23945PRO:GUS destination vectors with three different sizes of the putative AT2G23945 promoter. 201

Figure 5-5. Colony PCR of insert and vector sequence from GV3101 containing pGWB3:AT2G23945PRO:GUS colonies. 201

Figure 5-6. PCR amplification of the three different length promoter fusions of pGWB3:AT2G23945PRO:GUS transgene in *Arabidopsis* lines. 203

Figure 5-7. AT2G23945 GUS expression patterns were analysed using young and mature siliques from different plants harbouring different sizes of promoter. 204

Figure 5-8. Plantlets-specific GUS expression patterns were analysed using 3-week-old transgenic plantlets 205

screened using MS medium supplemented with Kanamycin (50 ng/µl).

Figure 5-9. Internode-specific GUS expression patterns were analysed from different plants harbouring different sizes of promoter. 205

Figure 5-10. GUS stains in inflorescence from different plants harbouring different sizes of promoter insertions. 207

Figure 5-11. GUS staining in buds from different plants harbouring different size of promoter insertions from the AT2G23945 gene. 208

Figure 5-12. GUS staining fusions anthers harbouring various sizes of AT2G23945 promoter fusions during pollen mitosis. 209

Figure 5-13. GUS staining in anther from different buds stages of plants harbouring the medium promoter fragment (AT2G23945pro-980:GUS). 211

Figure 5-14. No GUS activity was detected in pollen during all stages of pollen development. 212

Figure 5-15. Figure 5-15. Sub-cellular localization of pUBQ10:AT2G23945:YFP and pUBQ10::YFP 4-days post infiltration of *N. benthamiana* leaves. 214

Figure 5-16. Colony PCR of insert and vector sequences from pCR8/GW/TOPO entry clones with three different sizes of putative AT2G23945 translational fusion transcript. 216

Figure 5-17. The schematic map of pGKGWG:AT2G23945:GFP destination vector. 217

Figure 5-18. Colony PCR of insert and vector sequences from pGKGWG:AT2G23945:GFP destination vectors with three different sizes of promoter driving the AT2G23945 translational fusion. 218

Figure 5-19. Colony PCR of insert and vector sequence from GV3101 containing pGKGWG:AT2G23945:GFP colonies. 218

Figure 5-20. PCR amplification of the three different lengths of promoters translational fusions for 219

pGWB3:AT2G23945PRO:GUS transgene in *Arabidopsis* lines.

Figure 5-21. Localization of pGKGWG:AT2G23945:GFP. Green fluorescence detected only in tapetal cells region of anther from plant carrying 2354 bp transcript insertion (P3). 220

Figure 5-22. RT-PCR Expression analysis of AT2G23945-GFP transgene in transgenic plants. 221

LIST OF ABBREVIATIONS

%	Percentage
°C	celcius
µg	microgram
µl	microliter
µM	micromolar
ABA	Abscisic acid
AMS	*Aborted Microspores*
ANOVA	Analysis of variance
ASPG1	*ASPARTIC PROTEASE IN GUARD CELL 1*
BI-1	*Bax inhibitor-1*
bHLH	basic helix-loop-helix
CaMV35S	Cauliflower Mosaic Virus 35S
cDNA	Complementary DNA
CDR1	*Constitutive Disease Resistance1*
cm	Centimetre
CTAB	Cetyl trimethylammonium bromide
Cys	Cysteine
DAPI	4’,6-diamidino-2-phenylindole
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DYT1	*DYSFUNCTIONAL TAPETUM1*
EDTA	Ethylenediaminetetraacetic acid
EAT1	*ETERNAL TAPETUM 1*
ER	Endoplasmic Reticulum

EtBr	Ethidium bromide
g	Gram
GA	Gibberellin
gDNA	Genomic DNA
GFP	Green Fluorescent Protein
GUS	β-Glucuronidase
H_2O_2	Hydrogen peroxide
LB medium	Luria-Bertani medium
LD	Lethal Dose
Min	Minute
ml	milliliter
MMC	Microspore Mother Cell
MS	Murashige and Skoog
mRNA	Messenger Ribonucleic Acid
MS1	*Male Sterility1*
MYB	*Myeloblastosis*
NASC	Nottingham Arabidopsis Stock Centre
NO	Nitrogen Oxide
OD	Optical Density
PARP	Poly ADP-ribose Polymerase
PBS	Phosphate-buffered saline
PCD	Programmed Cell Death
PCR	Polymerase Chain Reaction
pUBQ10	Plasmid Ubiqutin- 10
PHD	Plant Homeodomain
PP2A3	Protein phosphatase 2A subunit 3
qRT-PCR	quantitative RT-PCR

RNA	Ribonucleic acid
RNAi	RNA interference
ROS	Reactive Oxygen Species
RT-PCR	Reverse Transcription-Polymerase Chain Reaction
Sec	Seconds
SD	Standard Deviation
SE	Standard Error
TAIR	The Arabidopsis Information Resource
TBE	Tris-Borate-EDTA-buffer
TDF1	*TAPETAL DEVELOPMENT AND FUNCTION 1*
TSS	Transcription Start Site
TUNEL	Terminal Deoxynucleotidyl Transferase–Mediated dUTP Nick-End Labeling
UTR	Untranslated Region
WT-Col	Wild-type Columbia, Col-0
YFP	Yellow Fluorescent Protein
VPE	Vacuolar processing enzyme

CHAPTER 1: INTRODUCTION

1.1 *Arabidopsis thaliana*

As a model plant for plant biotechnology, *A. thaliana* provides the opportunity to understanding the way that functional genomics can affect commercial crop production. Even though it has no direct economic value, this small annual weed which originated from central Eurasian region has many advantages for genome analysis (Table 1-1) (Azpiroz-Leehan and Feldmann, 1997; Meinke *et al.*, 1998). *A. thaliana* was the first plant genome to be sequenced; the sequenced regions are around 135-megabases and extend into the centromere regions. The genome is organized into five chromosomes that contain approximately 27,655 genes (Cheng *et al.*, 2017). Many different accessions or ecotypes are available from different part of the world, these are used to identify natural variation in traits of importance. Researchers have shifted the focus of *Arabidopsis* study towards identifying the function of every gene encoded in the genome.

Plant genetic engineering has become an easy and rapid aspect of the process of gene characterization. Moreover, *A. thaliana* is easily transformed by floral dipping the flowers with *Agrobacterium* culture (Clough and Bent, 1998), which carried a gene of interest in a selected plasmid vector. Then, selection for the presence of the transgenic in the succeeding generation can be conducted.

1.2 Reproductive System of *A. thaliana*

1.2.1 Flower

A. thaliana provides a model system for floral architecture analysis (Figure 1-1) (Irish, 2010). The plant produces two types of meristems, the floral meristems as well as the vegetative meristems. The floral meristem possesses an outer whorl of four green sepals; these leaf-like organs cover the flower bud during their development. In alternate positions to the sepals are an inner whorl containing four white petals. Internal to these are six stamens, each comprised of a filament and an anther which produces pollen. At the centre is the female organ or gynoecium that contains the ovules, which form the silique upon fertilization with the pollen. The vegetative meristem differs from floral meristems; vegetative shoot apical meristems produce leaves and axillary buds. Moreover, the flower has a limited life span due to its high energy requirments and due to ecological reasons (Irish, 2010). Rogers (2005) suggested the death of specific tissues is regulated to ensure the correct timing of floral progression, in which specific tissues die during flower development, whereas others remain until the flower dies. It is proposed by Rogers (2005) that tissues that degenerate and die early, such as tapetal cells and the synergid cells, provide nutrients to the growing tissues.

Table 1-1 Advantages of *A. thaliana* for genome analysis based on Meinke *et al.*, 1998; Somerville and Koornneef, 2002.

Advantage	Characteristics
Short generation time	The life cycle is completed within six to eight weeks, starting from germination of seed concludes with maturation of the seeds in the siliques.
Small size	1) Flowers are 2 mm long, self-pollinate once bud opens, and can be crossed after emasculation by applying pollen to the stigma surface. 2) Matured seeds are ~0.5 mm in length. Seeds can be kept as germplasm stocks for a long time. 3) Leaves are covered with small unicellular trichomes which have served as a model for morphogenesis and cellular differentiation studies. 4) Mature plants are around 15 to 20 cm in height and can reach up to 10 cm in diameter. 6) Simple root structure. 7) Seeds can be grown in petri dish plates in laboratory conditions and in pots in the glasshouse.
Large number of off spring	Self-fertilizing diploid plant that produces several hundred siliques with more than 5000 total seeds. It can facilitate the production of a vast numbers of offspring even from single transgenic plant.

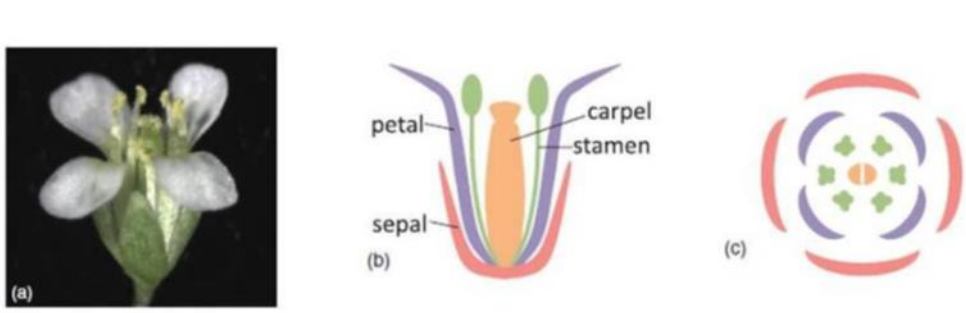


Figure 1-1 Schematic diagram of *A. thaliana* flower (Irish, 2010). (a) Mature flower of Arabidopsis during anthesis. (b) Floral diagram of a vertical section of a mature flower. (c) Floral diagram of a horizontal section showing the location of floral organs. Organ types are coloured as in (b).

1.2.2 Anther

Anther development in *A. thaliana* has been divided into 14 specific developmental stages based upon morphological characteristics that can be visualised by light microscopy (Sanders *et al.*, 1999). Anther development initiates with the emergence of the stamen primordia in the third whorl of the floral meristem and concludes with the release of pollen grains at dehiscence (Goldberg *et al*., 1993). Within the stamen primordium cell-specification and differentiation events give rise to mature anther cell types and generate the morphology of the whole stamen. In *A. thaliana*, the anther has a four-lobed structure (Sanders *et al.*, 1999). Vascular and connective tissues attach these four lobes to the central core. Details of anther development are presented in Figure 1-2 and Table 1.2.

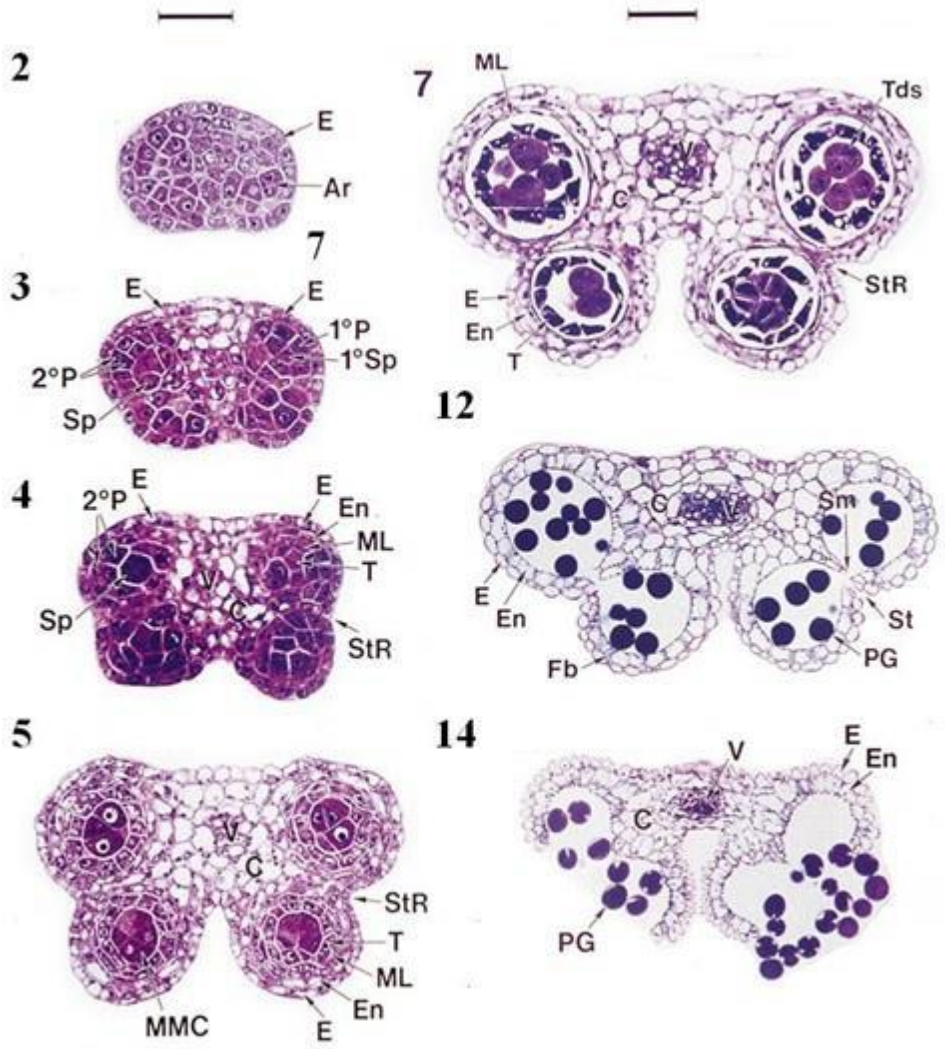


Figure 1-2. *A. thaliana* anther development and dehiscence diagram (Sanders *et al.*, 1999). The schematic diagram shows transverse sections of *A. thaliana* anthers at different stages. Numbering corresponds to the anther stage in Table 1-2. Bar: Stages 1-4=25 mm; Stages 5-8=25 mm.

Keys: Ar: Archesporial cells E: Epidermis, En: Endothecium, ML: Middle layer, T: tapetum, MMC: Microspore mother cell, MC: Meiotic cell, MSp: Microspores, PG: Pollen grain, Sm: Septum, St: Stomium, StR: Stomium region, Tds: Tetrads.

Table 1-2 Summary of the key events during flower, anther and pollen development (Sanders *et al.*, 1999; Ariizumi and Toriyama, 2011).

Anther stage	Major events and morphological markers	Flower stage	Pollen stage
1	Establishment of rounded stamen primordial.	5	
2	Archesporial cells emerge in four "corners" of L2 layer. Rounded primordial become more oval.		
3	L2 Archesporial cells divide periclinally (horizontally) to give rise to outer primary (1°) parietal cells and inner primary (1°) sporogenous cells. Further division of 1° parietal lineage generates the 2° parietal layer which consists of endothecium, middle layer and tapetum. These parietal layers lie under epidermal cells that originated from the L1 layer. Further division of the 1° sporogenous cell generate sporogenous cells, which later develop into microsporocytes.	7	
4	Commencement of a clear four-lobed pattern for anther and vascular region.	8	
5	Four anther locules formed. The four somatic layers, epidermis, endothecium, middle layer and tapetum are present (Figure 1-3). Tapetum has distinct dense cytoplasm and is binucleate. Microsporocytes develop to become microspore mother cells (MMC). This middle layer is proposed to provide nutrition to the developing MMC.	9	3
6	MMC undergo meiosis whilst middle layer degrades. Tapetum becomes vacuolated with increased metabolism. Overall, the anther undergoes an increase in size.		
7	Meiosis ends. Microspores are surrounded by the callose wall. Middle layer debris present in the locule.		4

	Programmed cell death (PCD) believed to commence from this stage. Formation of the pollen exine wall begins.		
8	Dissolution of callose wall. Tetrads degenerate, and individual microspores are released from the tetrads.	10	5
9	Further expansion and growth of the anther. Sporopollenin is actively synthesised and deposited onto the developing pollen to form the exine; microspores become vacuolated.		6-7
10	Tapetum degradation starts. Critical stage to support pollen formation.	11–12	
11	Start of pollen mitotic divisions. Tapetum degeneration initiated. Secondary thickening of endothecium layer and connective cells. Starting point for regeneration of septum cell and differentiation of stomium.		8–9
12	The exine wall of pollen develops during the bicellular pollen stage and continues until Mitosis II. Tapetum degeneration completed. Septum break down, anther become bilocular with tricellular pollen grains.		10
13	Coordinated with pollen development there is a general increase in anther size, degeneration of several cell layers and breakage along the stomium.	13–14	
14	Following dehiscence, the stamen senesces and falls off the plant. Contraction of cells and collapse of anther.	15–16	
15	Senescence of flower.	17	

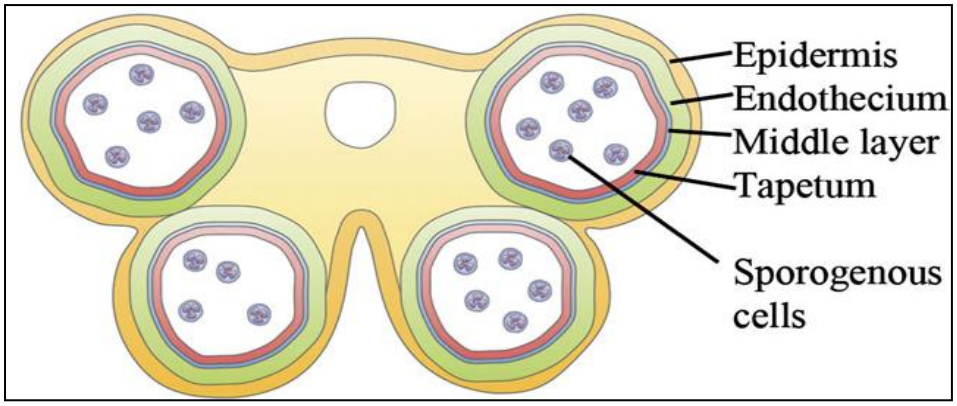


Figure 1-3 Diagram of an anther showing the four cell layers surrounding the sporogenous cells (Wilson *et al.*, 2011).

1.2.3 *Arabidopsis* pollen development

Pollen is the biological protector preserving pollen cytoplasm, which contains the two male sperms, from various harsh conditions, both physically and chemically. Pollen is formed in the anther locules; it is formed by sporogenous cells at the centre of each locule surrounded by four somatic cell layers (Figure 1-3) (Wilson *et al.*, 2011). Although the structures and patterns of pollen exine structure are variable, there are conserved mechanisms for pollen development process across the plant kingdom (Ariizumi and Toriyama, 2011).

Arabidopsis reproduction undergoes two sequential stages, microsporogenesis and microgametogenesis (Figure 1-4). Microsporogenesis occurs within the locule of each anther lobe. The pollen mother cells enter meiosis and generate haploid microspores, morphogenesis, and meiotic events constitute phase one of anther development (Goldberg *et al*., 1993; Borg and Twell, 2011).

Microgametogenesis is a developmental process occurring in the free haploid microspores, which undergo double mitosis. During Mitosis I microsporocytes asymmetrically divide and differentiate to develop bicellular pollen which contains one generative cell and one vegetative cell. Upon mitosis II, the generative cells in bicellular pollen divide into two functional sperm cells to become tricellular pollen. This mature pollen contains a thick exine wall. Prior to dehiscence, the mature pollen possess of two male gametes and a single larger vegetative tube cell which will developed to create pollen tube (Borg and Twell, 2011).

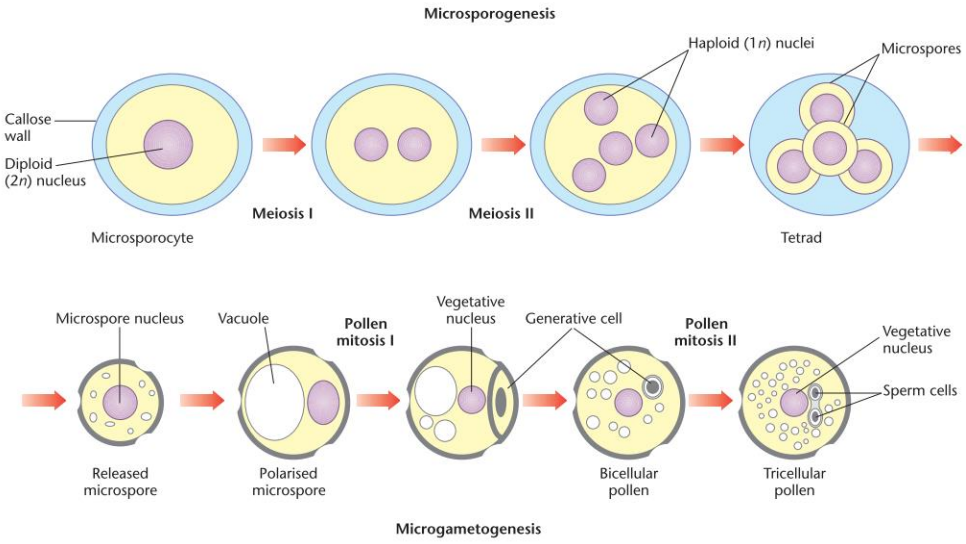


Figure 1-4. The development of *Arabidopsis* reproductive system switches between microsporogenesis and microgametogenesis stages (Borg and Twell, 2011).

Generally, the pollen wall consists of two principal layers, the exine and beneath this the intine (Figure 1-5). The exine is divided into two main parts, the sexine which is the outer most layer and the nexine, the layer bordering the intine. Tectum, bacula, and tryphine make-up the sexine layer. For intine, it build-up with cellulose,

pectin and proteins (Chang *et al.*, 2012). Different from exine, this layer is thought to be involved in pollen tube emergence and germination once the pollen attached to the stigma (Edlund *et al.*, 2004).

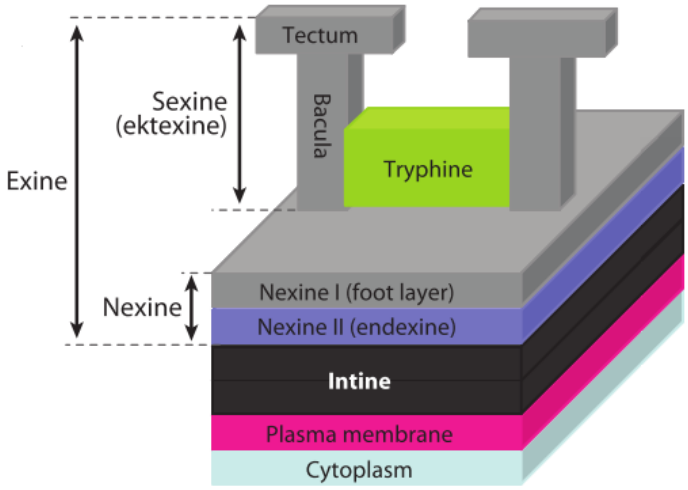


Figure 1-5. Schematic representation of pollen wall layer of *Arabidopsis*. It consists of two main layers, the outermost layer exine and intine. Exine includes sexine and nexine layer with. tectum, bacula, and tryphine make-up the sexine layer. The innermost layer of intine, located adjacent to the plasma membrane (Ariizumi and Toriyama, 2011).

Pollen lands on the stigma, hydrates, and forms a pollen tube, which delivers the sperm cells into the developing ovules (Updegraff *et al.,* 2009). For this reason, successful fertilization depends on the correct formation of the exine which is needed for the integrity of the pollen. Moreover, the exine also prevents loss of water and mechanical damage caused by environmental conditions during the transfer to the stigmatic surface e.g, Ultraviolet light, chemicals,

extreme pH and temperature and microbial infections (Ariizumi and Toriyama, 2011).

1.2.4 Tapetum

The tapetum has been shown to have an important role in determining fertility in plants through the production of the pollen wall, and altering tapetum Programmed Cell Death (PCD) can cause male sterility. As the innermost layer of the four sporophytic anther cell layers, the tapetum interacts directly with the developing gametophytes. It is binucleate by nature. During the metabolically active stage of the tapetum, the cells are rich with mitochondria, ribosomes, rough and smooth endoplasmic reticulum (ER) and Golgi bodies (Ariizumi and Toriyama, 2011; Quilichini et al., 2014). The tapetum cell layer remains in the anther until early tricellular pollen stage. After which point the tapetum undergoes PCD, the tricellular pollen contains a mature pollen wall, following the successful deposition of materials from the tapetum (Quilichini *et al.,* 2014) (See Section 1.3). This is regulated PCD, thus making the tapetum cells a model for understanding the complexity of plant PCD (Goldberg *et al*., 1993; Hsieh and Huang, 2007). Ariizumi and Toriyama, (2011) proposed that three events are required for viable pollen development; callose wall formation and degradation, primexine synthesis, and sporopollenin production, in which tapetum PCD is critical for this process.

Post-meiosis the tapetum secrets callase (β-1,3-glucanase), an enzyme involved in dissolution of β-1,3-glucan or the polymers of

callose wall to facilitate microspore release from the tetrads (Hird *et al.*, 1993). The callose wall inhibits the early degeneration of the developing microspores and provides protection for the tetrad cells (Figure 1-6). Furthermore, the callose layer facilitates primexine formation and provides a template for exine formation (Hird *et al.*, 1993; Dong *et al.*, 2005; Xu *et al.*, 2014). In addition, the callose wall provides glucose to the developing microspores development (Xu *et al.*, 2014). Correct breakdown of the callose is essential for pollen development. The GAMYB, *myb33/myb65* double mutant has aberrant tapetum enlargement that causes defective synthesis of callase, with the callose remaining in the locule after meiosis (Millar and Gublera, 2005). Silencing of the *Callose synthase 5, CalS5* caused abnormal exine wall and aberrant callose wall deposition onto the locule during meiosis (Dong *et al.*, 2005). Moreover, Shi *et al.*, (2015) revealed that *TRISTETRAPROLINE (AtTTP)*-OE lines displayed reduced fertility due to defects in callose wall and sporopollenin deposition.

During early stages of pollen maturation, all materials synthesised by the tapetum are mobilised to the developing microspores to form primexine, which consists of pro-sexine and nexine. Lipidic precursors build-up onto primexine during meiosis (Ariizumi and Toriyama, 2011). Once microspores reach maturity, the primexine and pro-sexine develop into exine and sexine respectively (Figure 1-6). Impairment of this process can cause aberrant pollen as shown by mutation of *NO PRIMEXINE AND PLASMA MEMBRANE UNDULATION (NPU),* which shows disrupted callose formation, absence of primexine deposition, defects in plasma membrane undulation and abnormal sporopollenin deposition (Chang *et al.*, 2012). Interestingly, low level of *CalS5* expression has been

observed in the *NPU* mutant. The authors speculated that *NPU* also is involved in sugar transportation onto the microspores and inhibition of the *CalS5* gene occurred due to the high sugar level in the microspores that blocks primexine formation (Chang *et al.*, 2012).

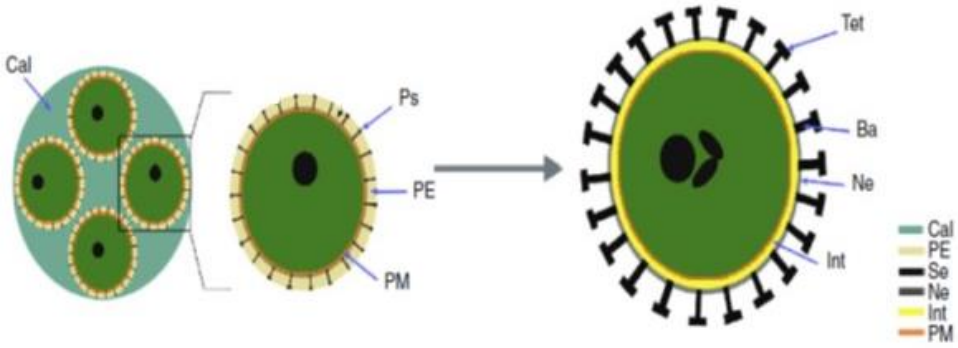


Figure 1-6. Diagram represents pollen maturation process from developing microspores and mature pollen. The left side showing developing microspores contains of primexine (PE) layer, pro-sexine (Ps), plasma membrane (PM), and a single microspore nucleus (n) surrounded by the temporary layer of callose (Cal). Meanwhile, the right side shown the representative of mature pollen where mature exine, derived for PE consists of tectum (Tet), bacula (Ba) and nexine (Ne), and intine layer (Int). Alongside pollen layer development, the nucleus with two sperm cells (oval shape) and vegetative nucleus (round shape) can be seen in mature pollen. Figure was modified from (Lou *et al.*, 2014).

Ariizumi and Toriyama (2011) reported that the tapetum is responsible for the synthesis and deposition of the sporopollenin. Alongside this, the tapetum degenerates to deposit tryphine materials for the pollen coat formation and in pollen hydration (Chang *et al.*, 2012). Tryphine serves to help attach pollen to the stigma surface or pollinator and in pollen hydration (Hernández-Pinzón *et al.*, 1999).

It is important to note that deposition of the sporopollenin and tryphine materials occurred at different times but are released into the locule for pollen wall formation. The sporopollenin precursors are transported whilst the tapetum is still intact, while the tryphine materials are released as the tapetum undergoes cell death (Quilichini *et al*., 2015). Defects in both processes due to disruption during tapetal PCD has been reported to cause male sterility.

For example, *MS2* gene encodes for a fatty acyl reductase which converts fatty acids to fatty alcohols during the sporopollenin synthesis. Mutation of the gene shows impaired sporopollenin deposition with non-viable pollen due to the lack of pollen wall (Aarts *et al.*, 1997; Xu *et al.*, 2010). Furthermore, absence of *DEFECTIVE IN EXINE FORMATION 1 (DEX1)* leads to incorrect deposition of sporopollenin to the plasma membrane instead of developing microspores. For this reason, there is a lack and delayed production of primexine, caus abnormalities for sporopollenin deposition with thin exine wall pollen generated (Paxson-sowders *et al.*, 2001). Additionally, silencing of the tapetum specific transcript *TETRAKETIDE a-PYRONE REDUCTASE1 (TKPR1)* (Grienenberger *et al.*, 2010) resulted in ruptured microspores with thin exine walls with an absence of tecta and baculae. This is due to impaired sporopollenin building block development, fatty acid derivatives known as hydroxylated a-pyrone.

1.3 Programmed Cell Death

PCD, or 'cellular suicide', can be defined as a rapid mechanism by eukaryotic cells during development for the overall maintenance of the cell populations. This natural process is highly conserved in multicellular organisms and is regulated by a complex biochemical pathway as a response to extracellular and intracellular environmental signals (Drew *et al.,* 2000; Chichkova *et al.*, 2012). Once triggered, this irreversible event activates a cascade of events that lead to PCD. PCD is necessary for development, and also to ensure the survival of cells by the selected removal of unwanted and damaged cells (Woltering, 2004; Turk and Stoka, 2007; Van Durme and Nowack, 2016).

Most of the known PCD processes in animals involve apoptosis. Hallmark characterstics of animal PCD are; changing mitochondrial permeability, plasma membrane blebbing, high intracellular calcium concentrations, DNA fragmentation, chromatin condensation, plasma membrane dissolution and disintegration of mitochondrial and cytoskeletal. The end products of this process, the apoptotic bodies, engulfed by phagocytes followed by degradation by lysosomal enzymes (Woltering, 2004; Van Doorn *et al.*, 2011).

In animal system, the mitochondria play a crucial role in apoptotic PCD with Cytochrome C released by mitochondria into the cytoplasm to activate cysteine aspartate-specific proteases (caspases). Caspases are synthesized and reside as inactive precursors (procaspases/zymogens) in the cytosol prior to activation of the

caspases by the removal of the N-terminal peptide (Woltering, 2004). Apoptosome or caspase-activating complexes are assembled in the cytoplasm driven by the release of the cytochrome C, often triggered by the extrinsic death receptors or intrinsic pathways. This apoptosome is the initiator of the death programme which causes DNA fragmentation and cell shrinkage (Turk and Stoka, 2007; Reape and McCabe, 2008).

In animals a cascade of caspases are involved, leading to PCD, caspases are members of the cysteine protease family that have specific cleavage activity (Chichkova *et al.*, 2012). The have been shown to cleave the proteolytic domain at specific recognition of Asp (D) adjacent residue sites (Wotering, 2004). They can be divided into initiator and executioners, and function to cleave substrates in cellular repair, or by directly degrading essential proteins and DNA (Ranganath and Nagashree, 2001). Caspases can be inhibited by artificial peptides that have similar recognition site consists of electrophiles like adeldehydes, nitriles or ketons at the C-terminal which interacts directly with cysteine (Woltering, 2004).

Plant apoptosis like-PCD has similar characteristics with animal apoptosis, such as DNA fragmentation, nucleus condensation as well as cytoplasmic condensation (Van Durme and Nowack, 2016), however does not include phagocytosis due to the rigid plant cell wall. Plant apoptosis like-PCD is triggered by various biotic and abiotic stresses, and development cues, e.g. reproductive and plant vegetative development (Trobacher *et al.*, 2006; Van Durme and Nowack, 2016).

Caspase-like activity is seen in plant PCD, acting in a similar way to caspases during developmental processes (Woltering, 2004, 2010; Bosch and Franklin-Tong, 2007). However, genome analysis has failed to identify orthologues of caspases in *A. thaliana*. Meaning that there is diversity in the enzymes that mimics caspase-like enzymes which cleave at the Asp specific site (Vartapetian *et al.*, 2011). However, plant metacaspases have been reported which resemble animal caspases, displaying cleavage specificity after the arginine (R) and lysine (K) residue, not after Aspartate (D) residues like caspases (Chichkova *et al.*, 2012). Nine *Arabidopsis* metacaspases *AtMC1-AtMC9,* which are involved in cell death activity have been reported (Fagundes *et al.*, 2015). *AtMC1* acts as a positive regulator in the hypersensitive response (HR) whilst *AtMC2* acts as an inhibitor in cell death activity (Coll *et al.*, 2010). *In vivo* analysis has shown that *AtSerpin1* inhibits cell death activity of *AtMC1* (Lema Asqui et al.y, 2017). Nevertheless, metacaspases are not inhibited by the caspase inhibitors and the regulation of metacaspases in tapetum PCD remains vague (Bosch *et al.*, 2008).

Over recent years, reports have indicated the involvement of caspase-like proteases. Limited number of Aspartyl proteases involve in plant-microbe interactions with caspase-like activity has been reported. For example, *Bcl-2-associated athanogene 6 (BAG6)* has been shown to be involved in defense against *Botrytis cinerea*. *BAG6* forms a heterodimer protein complex with *C2 GRAM domain protein (BAGP1)* and an aspartyl protease *APCB1*. *APCB1*-mediated *BAG6* is responsible for proteolysis of the targeted substrate of the pathogen as part of the plant defense mechanism. *In vivo* analysis of *B. cinerea* indicated that *BAG6* cleaved in caspase-1-like-dependent manner. The cleavage activity occurred at a specific

caspase-1-like site, downstream of the *BAG* domain using recombinant human caspase-1. Moreover, pepstatin has been shown to inhibit the aspartyl-protease activity of the protein complex (Li *et al.*, 2016).

Caspase-3 activity has been reported in Papaver. Studies in Papaver by Thomas and Franklin-Tong, (2004) showed an increase in cytosolic calcium concentrations due to self-incompatibility (SI) between pollen and the stigma and style. Cytochrome C was released into the cytosol and stimulated caspase-3-like enzymes (DEVDase) activity, which in turn cleaved the poly (ADP-ribose) polymerase (PARP), a classic substrate for caspase- 3 enzymes. In addition, the activity of caspase-3 like enzymes could be halted by the caspase-3 inhibitor I peptide (Ac-DEVD-CHO), where it blocked DNA fragmentation of pollen-tube. The identification of DEVDase activity in Papaver was the first report of the existence of mammalian caspase 3-activity that mediates PCD.

Further studies by Bosch and Franklin-Tong (2007) demonstrated that the protease exhibits activity at caspase cleavage sites by targeting natural caspase substrates such as poly ADP-ribose polymerase (PARP). At the same time, cytosolic acidification rapidly increases once the SI is induced, creating an optimum *in vivo* pH environment for caspase activities (Woltering, 2004).

Some PCD processes have been reported during tapetum degeneration. Tapetum PCD occurs once the tapetal cells have sufficient metabolites for nursing the pollen walls (Vizcay-Barrena

and Wilson, 2006). The absence of the mitochondria and microfilament marks the end of tapetum PCD (Papini *et al.,* 1999). In the anther, the middle cell layers and septum also undergo PCD degradation after tapetum degeneration, followed by stomium opening to allow the pollen dehiscence (Ge *et al.*, 2005).

Another key point is the presence of *BCL2*-associated X protein (*BAX*) family protein, which play a role as a regulator of apoptosis in animal systems (Ge *et al.,* 2005). There are no *BAX* genes reported in plant genomes (Van Aken and Van Breusegem, 2015). However, mammalian anti-apoptotic protein *BAX* protein, *Bax inhibitor-1 (BI)* is conserved in *Arabidopsis* and rice, and acts as an inhibitor of PCD in plants (Parish and Li, 2010). Several studies have confirmed similarities between plant PCD with their mammalian apoptosis. Kawanabe *et al.*, (2006) revealed that overexpression of *Bax inhibitor-1* (*AtBI-1*), an ortholog for human *Bax inhibitor-1* (*BI-1*) exhibited male sterility due to delayed tapetum PCD processes.

In contrast to PCD, another cell death event, necrosis occurs as a passive slow dying process in response for extrinsic factors such as stress and toxic substances. It starts with early loss of plasma membrane integrity, protoplast shrinkage, and continues until the degradation of the organelle and degeneration of DNA. No signal transduction pathway activation is needed for necrosis, since it is not linked to any developmental process and no energy is required during the process (Drew *et al*., 2000; Van Breusegem and Dat, 2006; Parish and Li, 2010).

1.4 Regulatory Networks Involved in Plant Tapetum Development

In general, PCD is classified into two group; development PCD (dPCD) and environment PCD (ePCD) (Van Durme and Nowack, 2016). Tapetum PCD is a dPCD regulated cell death process, which can provide insight into the molecular mechanisms involved in plant PCD. The timing of tapetum development is highly regulated; early or delayed tapetum degeneration frequently has detrimental effects on pollen development.

Male sterile lines are critical for hybrid plant breeding, since it has potential for improving crop yields via manipulation of pollen development for F1 hybrid seed production. More importantly, this kind of manipulation can be used to minimize the production cost by lowering the labour cost and can be manipulated to produce more environmental friendly genetic manipulation crops (Wilson *et al.*, 2011). Therefore, understanding the processes of pollen development will not only contribute greatly to the knowledge of plant biology, but also to the development of new male sterile varieties for hybrid plant breeding in the future.

In the proposed tapetum pathway, Parish and Li, (2010) suggested that multiple signals such as plant hormones, developmental and defence cues essential to activate tapetal-associated genes for PCD induction (Figure 1-7). It is important to note that in animal system, caspase acts downstream of all PCD signals (See Section 1.3).

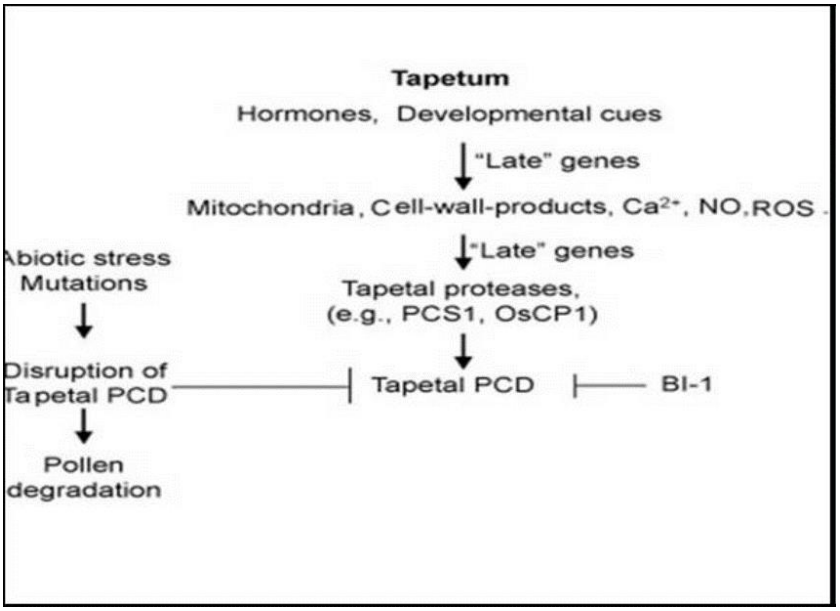


Figure 1-7 Proposed tapetal PCD pathway. Multiple signals triggered the tapetum PCD process as part of the plant development process (Parish and Li, 2010).

Research has been ongoing to reveal the signal transduction pathways involved in plant PCD. It is a complex signal transduction process where multiple external signals must be interpreted by cells. The responsive cells balance the signals before proceeding with the irreversible PCD process. Cells may response by producing different compounds that lead to proteolysis. However endogenous protease inhibitors be induced as a mean for ensuring plant survival (Trobacher *et al.*, 2006).

In the context of tapetum PCD, certain types of plant hormones are believed to be involved, for example gibberellin (GA) deficient mutants exhibit defects in anther development. For example, development of bicellular pollen is arrested in the *ga1-3* mutant (Cheng *et al.*, 2004). Furthermore, two GAMYB homologues,

myb33/myb65 double mutant exhibited male sterility due to failure of GA signal induction (Millar and Gublera, 2005). Rice GA-deficient lipid metabolic mutant, *cyp703a3* showed defects in exine formation and delayed in tapetal PCD (Aya *et al.*, 2009). Moreover, Yang *et al.*, (2013) revealed that an *ARF* transcription factor, *AUXIN RESPONSE FACTOR17* (*ARF17*) plays a role in regulating *CalS5* (See Section 1.2.4). Abnormal exine layer was seen due to impaired primexine and callose synthesis, leading to male sterility in *arf17* mutant. In addition, Auxin is synthesized in the tapetum and defective middle cell layer leads to abnormal auxin distribution and signalling process in anther cell layers as shown by the absence of *RECEPTOR-LIKE PROTEIN KINASE2 (RPK2)* (Cecchetti *et al.*, 2017).

Mitochondria has shown that it has a role for PCD in animal system in releasing Cytochrome C into cells. Nevertheless, their role in plant PCD is unclear. Reape and McCabe, (2008) suggested a high level of ROS in the cell may induce mitochondrial permeability contributing to a high level of Cytochrome C in the cell. In addition, Turk and Stoka, (2007) suggested that mitochondria's membrane permeability may changes due to multiple reasons, such as stimuli from the heat shock, DNA fragmentation and lack of nutrients.

Reactive oxygen species (ROS) such as Nitric oxide (NO) and Hydrogen peroxide (H_2O_2) appear to trigger activation of plant PCD signals in response to the different environmental conditions. ROS and H_2O_2 have also been implicated in signalling pathway targeting mitochondrial and chloroplasts (Bosch *et al.*, 2008; Van Durme and Nowack, 2016).

In plants, ROS production has been speculated to be produced by various organelles, e.g: mitochondria, chloroplasts, apoplast and peroxisomes (Xie *et al.*, 2014). They are naturally produced in plants at low levels, however at high concentrations they act like cellular scavengers and are lethal to cells. ROS activity is tightly regulated by the existence of natural ROS scavengers. Injuries and stresses that disrupt cell homeostasis accelerates the production of ROS (Rodríguez-Serrano *et al.*, 2012). The disintegration of cellular components lead to PCD once the ROS production exceeds the activity of ROS scavengers (Van Breusegem and Dat, 2006; Xie *et al.*, 2014).

Production of ROS and NO, that cause lethality to the cells, was reported *in vitro in* barley microspore and embryogenic suspension culture treated by cold and starvation stress treatments. Interestingly, the cells underwent a caspase-3-like associated PCD. Treatment with caspase-3 inhibitor, DEVD-CHO, significantly reduced cell death (Rodríguez-Serrano *et al.*, 2012). ROS are also involved in the regulation of the pollen-pistil interactions and the growth of the pollen tube PCD during SI in Papaver. Wilkins *et al.*, (2011) suggested that the cascade of increased calcium, followed by high concentrations of ROS and NO in cells induced PCD. Both processes occurred upstream of DEVDase/caspase-3-like activity in Papaver. In addition *TDF1* reported to have redox genes and calcium signalling genes as their possible downstream targets during tapetum PCD (Li *et al.*, 2017).

A NADPH oxidase encoding gene in *Arabidopsis*, *RESPIRATORYBURST OXIDASE HOMOLOG (RBOH),* has been shown to be involved in the production of O_2, which is then converted to toxic H_2O_2. Tapetal dysfunction occurred as a result of *RBOH* gene manipulation, thus cause male sterility due to altered DNA fragmentation (Xie *et al.*, 2014). Recently, *a36* and *a39*, two mutants of putative glycosylphosphatidylinositol (GPI)-anchored aspartic protease genes *Arabidopsis* have been characterised. Mutants exhibited reduced fertility with pollen from both mutants undergoing apoptosis-like PCD. ROS level is high in pollen of both mutants (Gao *et al.*, 2017). In rice, Yi *et al.*, 2016 reported that the *Defective Tapetum Cell Death 1 (DTC1)* rice mutant exhibits male sterility with late tapetal degradation. This transcription factor is thought to regulate the ROS level in cell, specifically the inhibition of ROS-scavenging activity that affects the tapetal degradation.

1.5 Tapetum Transcription Factors Network

Since the tapetum is important tissues for pollen development, it has been frequently targeted for the engineering of male sterility in plants. Liu and Fan, (2013) proposed a complex signal transduction network involving multiple transcription factors with the downstream target genes for *A. thaliana* (Figure 1-8). The transcription factors are grouped into two groups which are: 1) those expressed during tapetum differentiation; *NOZZLE (NZZ)/ SPOROCYTELESS (SPL), TAPETUM DETERMINANT1 (TPD1)* and *EXCESS MICROSPOROCYTES1/ EXTRA SPOROGENOUS CELLS (EMS1/EXS),* and 2) those expressed during tapetum development; *DYSFUNCTIONAL TAPETUM1 (DYT1)*, *TAPETAL DEVELOPMENT AND*

FUNCTION 1 (TDF1) or *MYB35*, *Aborted Microspores (AMS)*, *MYB80/MYB103/MS188* and *Male Sterility1 (MS1)*.

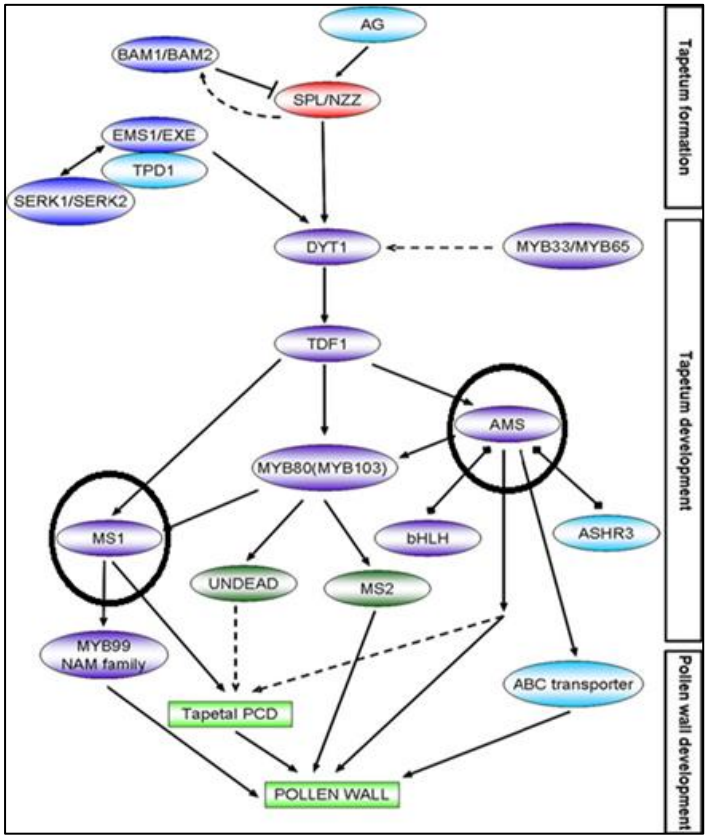


Figure 1-8. Transcription factors regulatory network during tapetum formation, tapetum development and pollen wall development in Arabidopsis (Liu and Fan, 2013). *AMS* and *MS1* are marked with black circles (Discussed in Chapter 3). Keys: Solid arrow: Direct regulation; Dashed arrow: Putative function; Lines ending with a square: Protein-protein interactions; lines ending with a bar: Repression.

NZZ/SPL has been identified as an early regulator, which is expressed during early stage of anther cell differentiation. The *nzz/spl* mutant exhibits reduced or total absence of nucellus cells in the gynoecium, thus the pollen sac fails to develop and failure of MMC stage also causes female sterility (Schiefthaler *et al.*, 1999). In

a typical wild-type anther, the archesporial cells give rise to the outer primary (1°) parietal cell and inner primary (1°) sporogenous cells, but in the *nzz* mutant, this event does not occur resulting in abnormal primary (1°) parietal cells and anthers that lack the endothecium, middle layer and tapetum. This transcription factor is a master regulator for vegetative growth anther cell differentiation, floral organ formation and stamen development (Schiefthaler *et al.*, 1999; Balasubramanian and Schneitz, 2000).

The downstream target of *NZZ/SPL* is *TPD1,* mutants of *tpd1* exhibits male sterility due to the abnormal degradation of the tapetum (Yang *et al.*, 2003).

EMS/EXS mutants display defects in tapetum development (Canales *et al.*, 2002, Zhao *et al.*, 2002). *EXS* regulates the archesporial periclinal divisions in the L2 layer of the anther. Analysis of the *exs* mutant showed that it has extra meiocytes, that progressed through meiosis and developed into microspores, but lacked the tapetum and middle cell layers and therefore did not fully develop resulting in male sterility (Zhao *et al.*, 2002).

The tapetum development pathway continues from the tapetum development stage to tapetum formation stage with the expression of *DYT1. DYT1* encodes for a basic helix-loop-helix (*bHLH*) transcription factor, necessary for normal tapetal differentiation and microspores development. The *dyt1* mutant does not undergo normal meiotic cytokinensis; during meiosis, a thin callose wall is observed surrounding the meiocytes and the tapetum has less dense

cytoplasm with abnormally enlarged vacuoles (Zhang *et al.*, 2006). *DYT1* has been speculated to regulate up to 1000 genes required for pollen exine formation, callose wall formation, normal cytokinensis, lipid and peptide transportation, and phenylpropanoid secretion (Zhang *et al.*, 2006). DYT1 protein formed heterodimers with two *GA-MYB* homologues*,* MYB33 and MYB65 to regulate tapetum PCD-related genes*. myb33* and *myb65* single mutant did not showed any impaired tapetum development. However, *myb33/myb65* double mutants showed tapetal aberrations that occurred before meiosis (Millar and Gublera, 2005; Zhang *et al.*, 2006). In addition, DYT1 also form a protein complex with bHLH010/089/091 (Figure 1-8) (Zhu *et al.*, 2015).

Another key regulator, *TDF1 (MYB35)* R2R3 MYB transcription factor is a downstream target of *DYT1*. A *tdf1* mutant exhibits a male sterile phenotype with a lack of pollen grains and disrupted tapetum development with excessive periclinal divided tapetum and increased vacuoles. In this case, the mutant shows late degeneration of the middle layer and tapetum, with both layers still surrounding the microspores after the tetrad stage. Interestingly, this gene shares similarity with *DYT1* with both transcription factors expressed during the same anther development stages (Zhu *et al.*, 2008). This tapetum specific transcription factor is reported to regulate genes related to lipid metabolism, calcium signalling, redox protein, cell wall formation, and E3 ubiquitin ligases in cell degradation (Li *et al.*, 2017).

AMS encodes for a *bHLH* transcription factor in the *MYC* family. When mutagenised it results in sterile plants with enlarged tapetum layer,

abortion of pollen and short siliques that are devoid of seeds (Sorensen *et al.*, 2003). *AMS* is highly expressed in the tapetum and recent study based on mathematical modelling suggested that *AMS* has a dual function. Biphasic or dual peak expression is reported for *AMS,* with initial expression during plant meiosis and a second peak reported during mitosis (Figure 1-9) (Ferguson *et al.*, 2016).

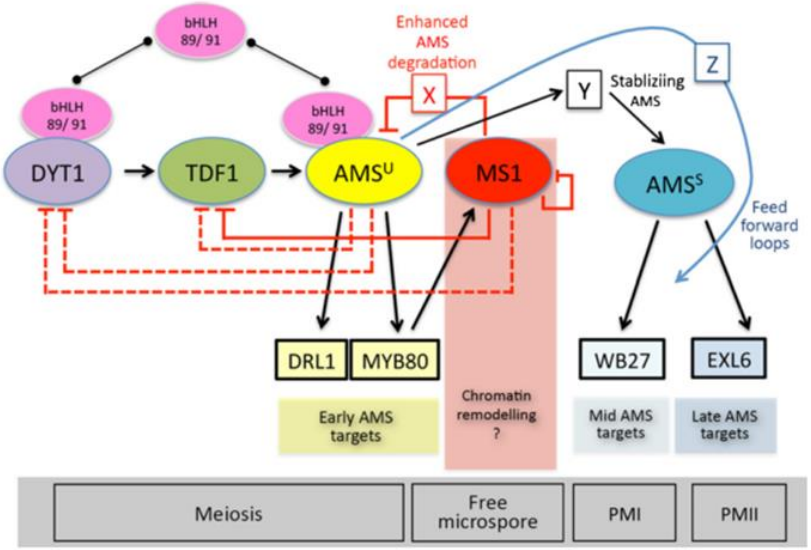


Figure 1-9. Tapetum transcription factors regulatory network for anther development, including pollen development stage in *A. thaliana* as proposed by Ferguson *et al*., (2016). The authors classified the complex network based on pollen development process. Keys: AMS^U: Unstable AMS protein; AMS^S : Stable AMS protein: X,Y and Z: Unknown elements; Solid black arrow: Direct regulation; Dotted lines; lesser role in protein regulation network; lines terminating with a bar: Negative regulation; lines terminating with a circle: Protein-protein interactions.

AMS transcription factor is predicted to act as a key regulator of pollen mitosis 1 by relaxation of the chromatin structure (Sorensen *et al.*, 2003). *AMS* regulates the expression of multiple genes

involved in pollen wall formation, tapetum development, lipid formation and lipid synthesis, substrate oxidation, methyl modification and Cys proteases (Xu *et al.*, 2010, 2014).

AMS forms a complex with other bHLH transcription factors *bHLHs 10, 89 and 91,* which are involved in the tapetum development. These three proteins form a complex with one another and with themselves, creating homodimers and heterodimers. *bHLHs 10, 89 and 91* triple mutants exhibit defect pollen development. However the single mutants of do not show defective pollen phenotypes (Zhu *et al.*, 2015).

Research has also been carried out on multiple downstream targets of *AMS* that showed and defective pollen phenotypes (Xu *et al.*, 2014). For example, *Adenosine Tri-phosphate-binding Cassette* Protein, *WBC27* defects showed aberrant sporopollenin deposition patterns (Xu *et al.*, 2010; Dou *et al.*, 2011). Furthermore, defects of nexine layer formation has been reported by the silencing *TRANSPOSABLE ELEMENT SILENCING VIA AT-HOOK (TEK)* or *AHL16* (Lou *et al.*, 2014). Moreover, GA-deficient lipid metabolic mutant *CYTOCHROME P450, CYP703* showed altered expression in *ams* mutant (Xu *et al.*, 2010). In addition, *AMS* is a key regulator for *CYP98A8* and its paralog *CYP98A9* where both genes involved in phenolic compounds production (Xu *et al.*, 2014).

Impaired pollen coat (tryphine) composition has been reported by the *AMS* target *Extracellular Lipase (EXL),* which is essential for pollen hydration process. In addition, *GRP17* encodes for a pollen

coat protein which shows comparable phenotypic defects like *exl4* mutant (Updegraff *et al.,* 2009). Meanwhile, defective on *EXL5* and *EXL6* expression resulting in abnormal exine, baculum, tectum and irregular tryphine deposition (Xu *et al.*, 2014).

Another key transcription factor that play a master role in tapetum development is *MYB80 (MYB103/MS113)*, the R2R3 MYB family characterised by the existence of R2R3 domain with one or two imperfect tandem repeats (Li *et al.*, 1999). Based on in-situ hybridisation analysis, it is highly expressed in tapetum and middle layer. Notably, premature degeneration of tapetum, partial callose degeneration, smooth pollen coat and reduced pollen cytoplasm has been reported for the *myb80* mutant; overexpression caused the lethality (Higginson *et al.*, 2003; Zhang *et al.*, 2007). Overall, *MYB80* has been reported to regulate lipid and sporopollenin synthesis, lipid synthesis and lipid transportation, callase formation and trichome development (Higginson *et al*., 2003; Zhang et al., 2007; Phan et al., 2011). EMSA analysis has indicated that *MYB80* has a dual function as an activator and repressor for its downstream targets. Three downstream targets of *MYB80*, a glyoxal oxidase (*GLOX1*), a pectin methylesterase (*VANGUARD1*), and an A1 aspartic protease (*UNDEAD*) has been identified, *UNDEAD* was upregulated, while *GLOX1* and *VANGUARD1* were downregulated (Phan *et al.*, 2011).

Equally important, *AMS* has an ability to form a complex with another element like *MYB80* in regulating *CYP703A2* which involved in biosynthesis of exine materials. *cyp703a2* mutants has impaired exine materials biosynthesis with lack of functional pollen as reported by Xiong *et al.*, (2016). In addition, *Anther-specific*

protein6 (A6) (Hird *et al.*, 1993) that play a role in callase enzyme production is also regulated by *MYB80* (Zhang *et al.*, 2007).

Another master regulator essential for normal pollen development is the *MS1* gene, which encodes for a PHD-finger transcription factor. The *MS1* gene plays an important role in microspores maturation (Wilson *et al.*, 2001). Silencing of *MS1* changes tapetum secretion of sporopollenin and pollen coat, leading to aberrant pollen (Wilson *et al.*, 2001; Ito and Shinozaki, 2002). Whereas, overexpression *MS1* causes mature pollen degradation with excess exine precursors deposition, as compare to early collapsed of microspores in *ms1* knockout mutants (Yang *et al.*, 2007). As discussed above, tight regulation system of the *MS1* is essential for the functional pollen development,

MS1 regulates diverse genes involved in the tapetal development, sporopollenin biosynthesis, pollen wall formation, pollen coat synthesis, pollen cytosolic components, the production of lipid bodies and Cysteine protease associated genes (Ito *et al.*, 2007; Yang *et al.*, 2007). It may activate protease genes that control the timing of tapetum PCD. *MS1* regulates tapetal PCD by suppressing PCD related genes during its expression (Vizcay-Barrena and Wilson, 2006).

In normal condition, *MS1* expression has lower expression before tapetum degradation stage, indicating that *MS1* silencing is essential for normal pollen development (Yang *et al.*, 2007). Whereas, in *ms1* mutant, the *MS1* gene is still expressed suggesting that self-

regulatory feed-back or negative regulation by the gene. As proposed by Ito and Shinozaki, (2002) *MS1* may act visa regulation of transcription of downstream genes via chromatin-mediated regulation. Chromatin modifications also occurred through formation of a complex transcription factors with other proteins (Yang *et al.*, 2007). *MS1* protein has also been linked in enhancing *AMS* degeneration, via a as yet uncharacteristic process (Figure 1-8) (Ferguson *et al.*, 2016).

Detailed tapetum PCD studies by Vizcay-Barrena and Wilson (2006) revealed the tapetum became abnormally vacuolated and granulated after meiosis. However absence of TUNEL positive signals in the *ms1* tapetum suggested that normal PCD does not occur in the mutant with abnormal necrotic breakdown of the tapetum observed (Vizcay-Barrena and Wilson, 2006).

MS1 has been shown to coordinate the expression *FACELESS POLLEN-1 (FLP1)* where *flp1* mutant showed smooth pollen surface due to defects of tryphine and sporopollenin synthesis (Ariizumi *et al.*, 2003). In that case, it caused distorted signal for pollen hydration and arrested pollen germination. *MYB99* is proposed as a direct target of the *MS1* which is believed to regulate genes associated with pollen wall biosynthesis (Ito *et al.*, 2007; Yang *et al.*, 2007); however, single *myb99* mutants do not exhibit fertility defects.

The number of researchers studying tapetum PCD in *Arabidopsis*, rice, and barley has grown considerably owing to their importance

in plant reproduction. Rice flower architecture has a distinct appearance compared to *Arabidopsis* with the appearance of asymmetric lemma and palea in the outer whorl and two lodicules in the second whorl (Zhang and Wilson, 2009). In contrast, rice reproductive organs such as stamens and carpels share similarity with *Arabidopsis*. Likewise, many of the pollen formation regulatory pathways in *Arabidopsis* have high similarity with the monocot rice regulatory pathway. Thus, to some degree the study of regulatory networks in rice can be used as a reference in *Arabidopsis* study (Figure 1-10).

Briefly, rice share similarities in pollen development and particularly tapetum development, for example *TAPETUM DEGENERATION RETARDATION (TDR),* an orthologue of Arabidopsis *AMS*, (Li *et al.*, 2006), *ETERNAL TAPETUM 1* (Niu *et al.*, 2013) and *PERSISTANT TAPETAL CELL1 (PTC1)* (Li *et al.*, 2011), counterparts of *Arabidopsis MS1*, *TAPETAL DEVELOPMENT and FUNCTION1 (TDF1)* respectively (Cai *et al.*, 2015). Ablation of these tapetal transcription factors leads to male sterility.

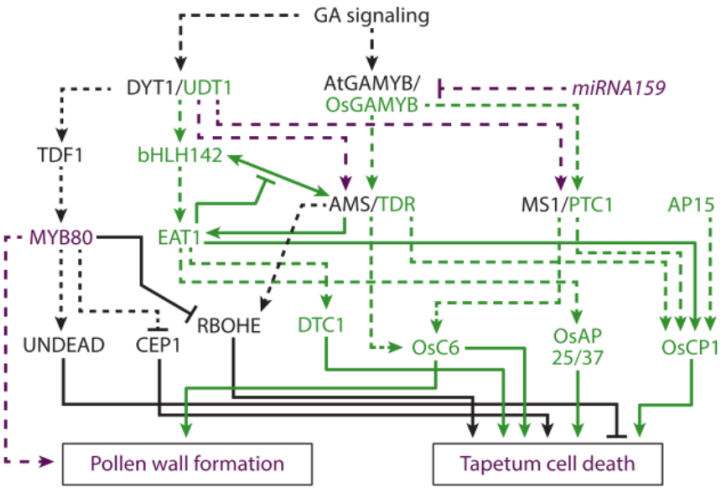


Figure 1-10. Tapetum PCD and pollen formation regulatory pathways in *Arabidopsis* that have a high similarity with tapetum PCD and pollen formation in monocot rice regulatory pathway (Daneva *et al.*, 2016). Black elements: *Arabidopsis* element; Green elements: Rice elements; Purple elements: Common elements; Solid lines: Protein-protein interactions; Dashed lines: Transcriptional regulation; Lines ended with arrow: Positive regulation; Lines ended with bar: Negative regulation.

1.6 Proteases

Analysis of the TAIR database (retrieved on 4th April 2018) of the *A. thaliana* genome, identified 607 proteases which represents 2.2% of the 27,655 coding genes. The high numbers of proteases involved in plant development reflects the important of these proteins in controlling biological and developmental events in plants. Plants have strict protein regulation by increasing or decreasing the number of proteins in response to different conditions and developmental stage. Proteases serve as hydrolysis enzymes involved in intracellular degradation of peptide substrates across organisms (Lee *et al.*, 2004). Many of the plant proteases, that have been identified are regulated or co-regulated by transcription factors

(See Section 1.4). In the proposed tapetal PCD pathway (Figure 1-7), proteases may act to trigger or delay tapetum PCD process (Parish and Li, 2010).

Proteases are separated into several classes based on the reactive residues at the active site of the proteolysis process. Recently, they have been classified into 4 main groups; cysteine proteases (C), serine proteases (S), aspartic proteases (A), and metalloproteases (M), where each groups divided into different sub-groups (Verma *et al., 2016)*. They are also classified based on the nature of their nucleophilic attack, endopeptidase or exopeptidase. Endopeptidases, in which an activated water molecule nucleophilic attack of the peptide bonds reactions takes place within the amino acid molecules with aspartic, glutamic and metallopeptidases as examples. On the contrary, exopeptidases in which the cleavage of the peptide bonds reactions takes place at the terminal side of amino acid molecules with Cysteine, threonine and serine as examples. The end product upon the proteolysis reaction of the peptide chain is the single amino acid (Rawlings and Barrett, 2013).

The involvement of caspase-like enzymes in plant PCD is the subject of much debate. It is known that proteases are linked with the PCD signal transduction process, however, less is known about the precise mechanism and their exact role during plant PCD. At the same time, several cysteine proteases have been reported as having altered expression in different transcription mutants, whether these proteases are associated with a sub-group of cysteine proteases, caspases remains unknown. Recent evidence has linked different type of proteases which exhibit cellular changes, with plant PCD.

1.6.1 Aspartyl Proteases

Named after Aspartate residues, Aspartyl protease (Enzyme Commission number, EC 3.4.23) are proteolytic enzymes involved in protein processing, development and degradation in broad range of organisms. This protein family is widely conserved in humans, animals, plants, yeast, fungi and viruses (Artal-Sanz and Tavernarakis, 2005), where plants make the largest part of the family (Gao *et al.*, 2017). The aspartic group of *A. thaliana* has been expanded to 69 genes (Takahashi *et al.*, 2008). Importantly, the signature characteristics of the group are the presence of the conserved two aspartate (D) residues located at their active conserved region of Asp-Thr/Ser-Gly-Thr/Ser motifs and also a Tyr residue (Tyr75/Y75). All reported aspartic enzymes are endopeptidases (Rawlings and Barrett, 2013).

Based on the HIV-1 protease as an aspartic model (Motwani *et al.*, 2015), the researchers reported that standard proteolysis mechanisms relies on the Asp 25 and Asp 215, the two ligands residues located at the catalytic centre or bilobed mature protein, one at each lobe (Figure 1-11). These Asp pairs donated the protons to the water molecules, thus activating the water molecules. As a result, the nucleophilic attack by catalytic water molecules starts at the substrate peptide bonds (Faro and Gal, 2005; Wlodawer *et al.*, 2013). Proteases activity can be halted by pepstasin A (Wlodawer *et al.,* 2013).

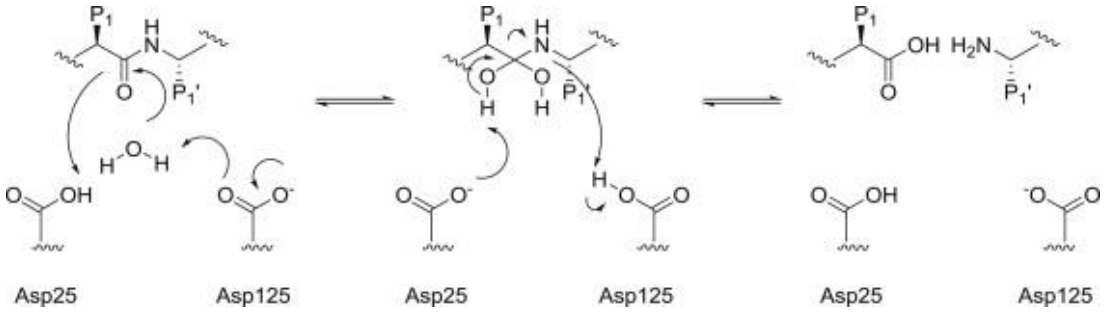


Figure 1-11. Aspartyl proteases activity: It is suggested that a tetrahedral intermediate is formed prior to the peptide bond cleavage. Asp215 may serves as proton donor to the substrates and the Asp25 regains the negative charge from the carboxyl anion through the carbonyl group (Motwani *et al.*, 2015).

Takahashi *et al.*, (2008) speculated that the aspartic proteases showed expression in different plant tissues and grouped them into two different subfamilies; 1) Pepsin-type aspartic proteases (PAP) subfamily and 2) Nepenthesin-like aspartic proteases (NAP) subfamily. The different features between these two groups are the number of Cys residues in NAP which has approximately 12 residues for every molecule, twice compare to that of the PAP. Along this, NAP has approximately 20-30 residues of plant-specific insert (PSI) compared to about 100 residues of PSI in PAPs. A recent update by Rawlings and Barrett in (2013), assigned the eukaryotic aspartic protease into several clans. The clans divided into five different clans which consist of Clan AA, Clan AC, Clan AD, Clan AD, Clan AE and Clan AF. For each clan, the protein folds differently from one another in three-dimensional structures.

In the last decade, many researchers have reported aspartic proteases involvement in different plant development process, with contrary functions of the proteases due to the different substrates in various tissues (Gao *et al.*, 2017). Furthermore, enhanced

expression of the apoplastic aspartic protease, *Constitutive disease resistance (CDR1),* increased the resistance of plants to virulent *Pseudomonas syringae* (Xia *et al.*, 2004). A report on its orthologue, *Oryza sativa CDR (OsCDR1)* relvealed identical characters (Prasad *et al.*, 2009). Moreover, mutation of *PROMOTION OF CELL SURVIVAL 1 (PCS1*) has been reported to have detrimental effects on reproduction development with excessive cell death of developing embryos and abnormality of male and female gametophytes (Ge *et al.*, 2005a). Furthermore, aspartyl-protease *APCB1* as reported by Li *et al.*, (2016) is essential for plant immunity and PCD, and mediating basal resistance processing by *BAG6 gene*. Additionally, an Arabidopsis gene, *ASPARTIC PROTEASE IN GUARD CELL 1 (ASPG1)* is involved in abscisic acid (ABA) drought-signalling pathways and localised in guard cells. Decreased water loss, due to the increased ABA sensitivity, was reported for overexpression *ASPG1-OE* (Yao *et al.*, 2012). In addition, mutants in an aspartyl-protease gene, *EDS1-DEPENDENT 1 (AED1)* is needed for systemic acquired resistance (Breitenbach *et al.*, 2014).

Other proteases are also involved in plant PCD, *UNDEAD* regulates the timing of the tapetal PCD where early tapetal PCD reported in *undead* mutant. Localised in mitochondria, *UNDEAD* may function as an inhibitor for PCD process or as proteolysis of apoptosis-inducing PCD protein, where it inhibits the destruction of tapetal mitochondria. Furthermore, TUNEL analysis was conducted and exhibited positive signal indicating that DNA cleavage takes place (Phan *et al.*, 2011). A recently published report by Gao *et al.*, (2017) revealed that over expression of *A36* and *A39* Aspartic proteases exhibited an upsurge of unviable pollen, leading to a reduction in plant fertility.

In addition, Niu *et al*., (2013) reported that *OsAP25* and *OsAP37* plays a cell death function when overexpressed in plants and ectopically in yeast. It is also thought that rice *EAT1* downstream target genes, rice aspartic protease *OsAP65,* promoting cell death where *osap65* mutants exhibited impaired male gametophyte and defects in pollen tube development (Huang *et al.*, 2013a).

1.6.2 Cysteine protease

Cysteine (EC 3.4.22), also known as thiol proteases, are found in a diverse range of organisms. The catalytic process occurs at the C-Terminal of polypeptide chains, hence they are exopeptidase (Otto and Schirmeister, 1997). Interestingly, the degraded amino acids upon enzymatic reaction can be used for new protein synthesis (Lee et al., 2004). Association of Cysteine proteases with plant PCD remains unclear, although a group of Cysteine called Caspases (Discussed in section 1.3) are well-studied in animal system.

Another group of cysteine proteases, papain-like cysteine proteases, which cleave at specific recognition sites of asparagine (N) carboxyl group. Notably, *CEP1* is a papain-like cysteine protease that has been reported to be involved in tapetum PCD (Phan *et al.*, 2011). It is identified as one of the downstream targets of *MYB80* and is expressed in the tapetum and the *cep1* mutant showed delayed tapetum degradation, causing defects in pollen formation. Whereas, overexpression plants exhibited premature tapetum degradation leading to abnormal pollen maturation. Importantly, TUNEL assay

showed that *cep1* mutant exhibited a hallmark PCD character with DNA fragmentation occurred in the tapetum that undergo delayed degradation (Zhang *et al.*, 2014). In addition, *Cysteine Protease 51*, *CP51* has been shown to be necessary for pollen exine development (Yang *et al.*, 2014). *CP51-RNAi* plants with reduced expression caused premature degradation of tapetum and abnormal pollen abortion.

The rice *TDR* gene is also proposed as the regulator for the Cysteine protease namely *OsCP1,* and protease inhibitor *OsC6*; both genes were positively regulated by TDR during tapetum PCD. Rice proteases gene *OsCP1* is highly expressed in the tapetum and developing pollen (Li *et al.*, 2006). Silencing of *OsCP1* exhibited male sterility with defective pollen (Lee *et al.*, 2004), whilst *OsC6* mutant showed reduced fertility caused by abnormality in Ubisch bodies and pollen exine development (Zhang *et al.*, 2010). A summary of proteases which are involved in plant development is presented in Table 1-3.

1.7 F-Box Protein

The F-box proteins contain at least one F-box motif, and members of this protein structural motif contain a conserved 40–50 amino acid F-box motif. The name originated from the human cyclin F, the first protein identified in the complex, and conserved in many genomes of eukaryotes. Nearly 700 F-box proteins have been identified in the *Arabidopsis* genome and act as a site of protein-protein interaction. In general, F-box proteins function in many cellular mechanisms during plant developmental, physiological processes and regulating

the availability of transcription factors for initiating or maintaining transcription (Chae *et al.*, 2008; Kim *et al.*, 2012).

F-box proteins are elements of a complex known as SCF, composed of SKP1, CULLIN, and F-box proteins and RBX1 proteins (Figure 1-12). These proteins are key factors in mobilizing substrates for degradation due to their ability to interact with SKP1 via a N-terminal F-box and with the substrate via a C-terminal protein - protein interaction domain. Furthermore, this domain is rich with WD repeats and leucine-rich repeats (LRR) motifs (Kipreos and Pagano, 2000; Xiao and Jang, 2000). Substrates degradation by ubiquitin-conjugating enzymes is facilitated by the SCF complexes by transferring the ubiquitin onto substrates. Then, these poly-ubiquitinated substrates degraded by the proteosome (Kipreos and Pagano, 2000). The majority of targeted proteins are broken down by the 26S proteosome, which consists of 20S proteosome and one or two 19S regulatory complexes (Woltering, 2010).

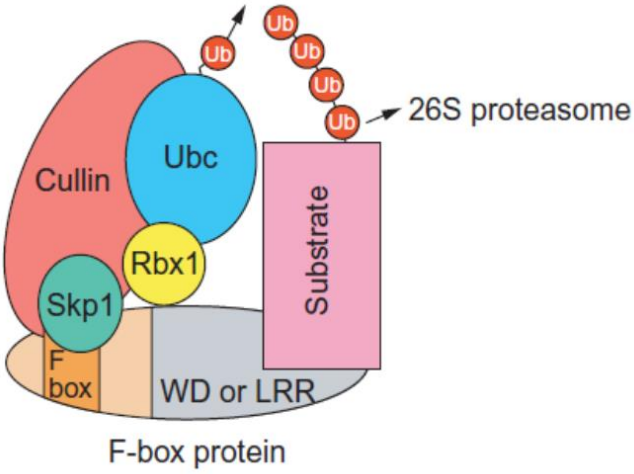


Figure 1-12. Diagram of the F-box protein. The F-box protein interacts to the SCF complex via interaction between the F-box and Skp1 (Kipreos and Pagano, 2000).

F-box proteins play important roles in plant development. Nevertheless, there is little information on involvement of the F-Box protein in *Arabidopsis* PCD. An *Arabidopsis* F-box protein called *FBL17* has been identified as involved in the cell cycle regulation of pollen development as stated by Gusti *et al.*, (2009). The authors stated that the *fbl17* mutant results in an arrest of pollen development during mitosis II which leads to defects in embryo development and abnormal endosperm development. For rice, Anther Development F-box (ADF) protein, a downstream target of *TDR*, an orthologue of Arabidopsis *AMS,* showed delayed tapetum PCD and aborted pollen development on *ADF-RNAi* lines. Suppression of the F-Box protein lead to altered tapetum degeneration (Li *et al.*, 2014).

Table 1-3. Summary of proteases and F-Box protein which involved in plant developments.

Proteases and F-box	Brief description	Reference
Aspartyl Protease		
CDR1	OE plants showed increased resistance of plants to virulent *P. syringae.*	Xia *et al*., (2004).
PCS1	Involve in anti-PCD process. Mutants exhibit excessive cell death of developing embryos, and abnormality of male and female gametophytes.	Ge *et al.*, (2005).
OsCDR1	OE plants showed increased resistance of plants to virulent *P. syringae*.	Prasad *et al.,* (2009).
UNDEAD	Involve in anti-PCD process. Mutant exhibits early tapetal PCD.	Phan *et al.,* (2011).
ASPG1	Involve in (ABA) drought-signalling.	Yao *et al.,* (2012).
OsAP25 and OsAP37	Involve in pro-PCD process. Overexpression caused cell death in plants and in yeast.	Niu *et al.,* (2013).
OsAP65	Mutant exhibits impaired male gametophyte and defects in pollen tube development.	Huang *et al.,* (2013).
AED1	Essential for systemic acquired resistance.	Breitenbach *et al.,* (2014).
APCB1	Involved in plant immunity and PCD and mediating basal resistance processing by BAG6 gene.	Li *et al.*, (2016).
A36 and A39	Involve in anti-PCD process. Mutants exhibited PCD in pollen and female gametophyte.	Gao *et al.,* (2017).
Cysteine Protease		
OsCP1	Mutant exhibits male sterility with defective pollen.	Lee *et al.*, (2004).
OsC6	Mutant exhibits abnormality in ubisch bodies and pollen exine development.	Zhang *et al.,* (2010).
CEP1	Involve in pro-PCD process. Mutant shows delayed tapetum degradation.	Zhang *et al.,* (2014).

CP51	Involve in anti-PCD process. Mutant exhibits premature degradation of tapetum.	Yang *et al.,* (2014).
F-Box Protein		
FBL17	Mutant defects on embryo development and abnormal endosperm development.	Gusti *et al.,* (2009).
OsADF	Involve in pro-PCD process. Mutants exhibits delayed tapetum PCD and aborted pollen development.	Li *et al.,* (2014).

1.8 Aims and Objectives

Two approaches can be used to establish the link between phenotype and gene sequence, these involve forward and reverse genetics. Reverse genetics can be defined as phenotypic changes using specific engineered gene sequences to generate targeted gene mutants (Kleinboelting *et al.*, 2016). Whereas, forward genetics begins as mutant screens to identify phenotypic changes prior to investigation of alteration of the gene sequences (Krysan *et al*., 1999). Several strategies have been implemented to interfere with the anther and pollen development by genetic engineering to enable the analysis of pollen development (Wang *et al.*, 2012).

Proteases that showed differential expression in *ams* and *ms1* male sterile mutants that have defects in tapetum PCD were identified (Pearce *et al.*, 2015). Their expression was determined using bioinformatic tools and those expressed specifically during pollen development were selected for analysis by reverse genetic approaches to establish their role in pollen development. Proteases had been considered to play a central role in protein degradation by cleavage of the specific substrates. These proteases and F-Box

protein are speculated to hydrolyse an apoptosis-inducing protein or anti-apoptosis-inducing protein.

This study focuses on analysing the process of tapetum degeneration and pollen formation, to determine the effects of seven selected proteases and a F-box protein on plant development and their potential roles associated with reproduction and tapetum PCD. These target genes consists of six different Aspartyl proteases genes (AT1G25510, AT2G23945, AT4G30030, AT5G10760, AT3G20015 and AT5G19120); one cysteine proteinase gene (AT2G21430) and a single F-box protein encoded gene, AT3G16210 gene, which have been previously shown to have altered expression in male sterile mutants associated with tapetum development, namely *ams* and *ms1* (Pearce *et al.*, 2015). This project aimed to uncover the biological functions of the selected genes. The specific objectives project were:

i. To analyse the effects of over-expression of genes encoding seven different proteases and an F-box protein that are expressed specifically in the tapetum (Chapter 3).

ii. To study the effects of overexpressing and silencing AT2G23945 Aspartyl-protease gene on tapetum degradation process and its expression profiles (Chapter 4).

iii. To study the transcriptional and translational expression patterns of AT2G23945 gene (Chapter 5).

CHAPTER 2: MATERIALS AND METHODS

2.1 Analysis of Tapetum Associated Genes

2.1.1 Plant materials and plant growth conditions

Growth conditions consisted of a 16/8- hour photoperiod at 20±2 °C. Plants were grown in 9 cm plastic pots filled with Levington® M3 pot and bedding compost. *Arabidopsis* (*Col*) seeds and all insertion lines were obtained from Nottingham *Arabidopsis* Stock Centre (NASC) and were sown at a density 4-5 seeds per pot. Plants were watered until the siliques had dried completely and then the plants allowed to dry slowly on dry benches for maximum viable seed production. Seeds were harvested after the whole plants had dried out and were stored in moisture-porous paper bags in a dry atmosphere at room temperature.

2.1.2 Primer design

Primer sequences were designed using the Primer 3 software in Biology Workbench version 3.2, accessible from the internet (http://biowb.sdsc.edu/CGI/BW.cgi). Primers were selected according to the following criteria: GC content > 40%, optimal annealing temperature (Tm) ~ 55 - 65^{o}C, primer length 22 - 28 bases, and sequences specific to the gene of interest only. All

primers sequences are listed in Appendix 1 and full-length cDNA of all proteases genes are listed in Appendix 3.

2.1.3 RNA extraction and cDNA synthesis

RNA was extracted from intact inflorescences. The protocol for RNA extraction was provided by the QIAGEN RNeasy Kit. RNA quality and quantity were tested on a NanoDrop fluorospectrometer. The RNA samples were stored at -80°C. Total RNA will be used to obtain the first strand cDNA by reverse transcription (RT). The cDNA Synthesis, with final concentration of 2.0 µg was conducted following the protocol of cDNA Synthesis with SuperScript® III RT' by Thermo Fisher Scientific Inc.

2.1.4 Genes cloning and Phusion PCR

Thermo Scientific Phusion High-Fidelity DNA Polymerases was used for gene cloning purposes that required high fidelity. The product of this cloning technique is blunt end DNA. The PCR reaction mixture solution contained 2.5 µl of 5X GC buffer (contain 1.5 mM $MgCl_2$), 0.2 mM dNTPs, 0.25 µm of protease genes specific forward and reverse primer, 0.4 µl DMSO, 0.25 µl of Phusion polymerase (2U/µl) and 200ng of cDNA template. Sigma Molecular water was added to a final volume of 10 µl. Primers used in this experiment are given in Table 2-1 and primer sequences are shown in Appendix 1. List of destination vectors generated in this work are in Appendix 2. Full-length transcripts for seven proteases and one F-Box protein are listed in Appendix 3.

The standard thermal profile of the Phusion reaction that has been used throughout the study was: initial denaturation at 98 °C for 3 min, followed by 35 cycles at 98 °C for 15 sec, annealing at different temperatures depending on the primers for 30 sec, and extension at 72 °C at different time length depending on the fragment sizes. Details of the extension time are provided in every section. The final extension was 72 °C for 4 min followed by cooling to 10 °C. PCR products were separated by electrophoresis through 1.0 % (w/v) agarose gels in 0.5X TBE buffer (Tris-Borate-EDTA-buffer (5X) Sigma) at 100 V for 40 min. The gel was stained with Ethidium Bromide (EtBr) (0.2 ng/µl) and was documented using GeneSnap gel doc system (SynGene, Cambridge).

Table 2-1: List of primers, PCR extension period and optimized annealing temperatures for Phusion reaction of seven selected proteases genes and one F-Box protein. These genes were identified as having a high level of expression during flower development and altered expression in the two male sterile mutants, *ams* and *ms1*. PCR extension period used during amplification was two minutes with Optimized Tm of 60 °C. Optimized annealing condition at 60 °C for 30 sec, and extension at 72 °C for a minute and 30 seconds.

Gene	Forward Primer (5'-3')	Reverse Primer (5'-3')	Fragment length (bp)
AT1G25510	AT1G25510_Gibson_ Forward_Primer	AT1G25510_Gibson_ Reverse_Primer	1491
AT2G23945	AT2G23945_Gibson_ Forward_Primer	AT2G23945_Gibson_ Reverse_Primer	1416
AT4G30030	AT4G30030_Gibson_ Forward_Primer	AT4G30030_Gibson_ Reverse_Primer	1314
AT5G10760	AT5G10760_Gibson_ Forward_Primer	AT5G10760_Gibson_ Reverse_Primer	1434
AT3G20015	AT3G20015_Gibson_ Forward_Primer	AT3G20015_Gibson_ Reverse_Primer	1452
AT5G19120	AT5G19120_Gibson_ Forward_Primer	AT5G19120_Gibson_ Reverse_Primer	1200
AT2G21430	AT2G21430_Gibson_ Forward_Primer	AT2G21430_Gibson_ Reverse_Primer	1125
AT3G16210	At3g16210_CDS_F	At3g16210_CDS_NS_R	1084

2.1.5 MicroElute cycle-pure spin protocol

Following the Phusion PCR cloning, the blunt end PCR products were produced. MicroElute Cycle-Pure Spin has been conducted to purify the PCR products by removing all the Phusion DNA Polymerase. The PCR product need to be purifies to prevent the Phusion DNA Polymerase disrupting the whole PCR product. MicroElute® DNA Clean Up protocol was provided by the Omega Bio-tek, Inc Company.

2.1.6 Overexpression vector, pUBQ10 vector with YFP-tagged protein

Plasmid Ubiqutin- 10 promoter (pUBQ10) developed by manipulating the promoter region of the *Arabidopsis* AT4G05320 gene with additional YFP tagged protein at the C-Terminal fusion (Grefen *et al.*, 2010). For transient expression, this gateway compatible vector was developed to enhance the gene expression with longer expression time. In addition, it consists *BAR* gene which generates the transgenic plants resistance to Basta® herbicide application. Thus, simple transgenic screening process devoid of laborious screening using medium supplemented with antibiotics. p19 plasmid was used to avoid post-transcriptional gene silencing.

2.1.7 Gibson assembly cloning reaction

The Gibson Assembly® Master Mix developed by New England Biolabs was used to facilitate the cloning of amplified PCR products into a pUBC plasmid vector that contains pUBQ10 promoter. The protocol was provided by New England Biolabs ®Inc. The Gibson Assembly® Master Mix Cloning reaction contained: 200 ng of fresh PCR product, 10 µl of Gibson Assembly Master Mix and 70 ng linearized pUBC plasmid vector; the final volume of the reaction mixture was 20 µl. Linearization of plasmid was conducted using *XhoI* restriction enzyme at *XhoI/XhoI* cutting site. This mixture was mixed gently and incubated for an hour at 50 °C. The cloning reaction was then placed on ice and followed by transformation into DB 3.1 competent cells.

The protocol for DB 3.1 transformation was based on the protocol suggested by New England Biolabs ®Inc with some modifications. First, 5 µl of the Gibson Assembly® Master Mix Cloning reaction was added into a vial of DB 3.1 competent cells and gently mixed. The mixture was incubated on ice for 2 minutes. Then, the cells were heat-shocked for 90 seconds at 42°C without shaking, followed by immediate transfer to ice. Next, 250 µl of room temperature S.O.C. medium was added. The tube was shaken horizontally (200 rpm) at 37°C for 1 hour. Then 100 µl from the transformed cells were spread on a prewarmed LB medium supplemented with 50 ng/ul Kanamycin antibiotic and was incubated overnight at 37°C.

2.1.8 Plasmid colony PCR

After overnight selection colonies were picked for PCR analysis. The standard thermal profile of the reaction was programmed as follows. The standard PCR reaction mixture contains 5 µl of REDTaq® DNA Polymerase, 0.3 µm of pUBQ10 vector specific forward primer, 0.3 µm of gene specific reverse primer and 100 ng of cDNA templates. Sigma Molecular water was added to a final volume of 10 µl. The standard PCR reaction throughout the study was programmed as follows: initial denaturation at 94 °C for 2 min, followed by 35 cycles at 94 °C for 30 sec, annealing at different temperatures depending on the primers °C for 30 sec, and extension at 72 °C for different time depending on the experiment. The final extension was at 72 °C for 10 min followed by cooling to 10 °C. PCR products were separated by electrophoresis as in section 2.5. The successful inserted colonies were chosen and transferred to a fresh liquid LB medium supplemented with 50 ng/ul Kanamycin antibiotics. The samples were incubated at 37 °C for overnight with continuous shaken at 200 rotations per minute (rpm).

2.1.9 Plasmid extraction and sequencing

Single colony culture needs to be extracted. Two basic principles of plasmids extraction are the lysis of the plasmids cells and purification of the plasmid DNA. The lysis needs to be done to separate the plasmid DNA from the other components. The protocol was based on the protocol provided by the GenElute™ Plasmid MiniPrep Kit – Sigma. Plasmid DNA was sent to Eurofins Scientific

Company (UK) for sequencing. Then, the sequencing result was BLAST (Basic Local Alignment Search Tool) which can be accessed through the Internet (http://www.ncbi.nlm.nih.gov/) for confirmation of amplification direction. Details on primers used for sequencing purposes are given in every analysis section.

2.1.10 Agrobacterium electroporation

Agrobacterium-mediated transformation system was used for this technique. The plasmid vectors that carried the insertion were transformed into *Agrobacterium tumefaciencs* strain GV3101 using electroporation. The electroporation cuvet (4.5 ml polystyrene Kartell cuvette), was chilled on ice before starting of the experiment; 150 ng of the vector plasmid with inserted cDNA was added to 40 µl of *A. tumefaciens* strain GV3101 competent cells in an Eppendorf tube and gently mixed. The mixture was inoculated into the chilled chamber and electroporation conducted (MicroPulser, with 6 m/s pulse). Next, 1 ml of LB medium was added to the chamber; the mixture was left at 28 °C for 3 hours before plating on LB medium supplemented with Gentamycin (30 ng/µl), Rifampicin (25 ng/µl), and Kanamycin/Spectinomycin (50 ng/µl) antibiotics depending on the experiments conducted. The plate was incubated at 28 °C for 2-3 days.

2.1.11 Transient expression of seven proteases genes in *Nicotiana benthamiana*

Transient expression analysis were conducted in *N. benthamiana* leaves based on (Voinnet *et al.*, 2003).

2.1.12 Observation of the subcellular localization of GFP and YFP-tagged proteins

The inoculated regions of the leaves were observed through laser scanning confocal microscopy after four days post-inoculation. The inoculated region was excised into 1 cm x 1 cm sections. The samples were mounted in water and covered by cover slip. Then, the fluorescence of the expressed proteins inspected by Leica SP2 confocal laser scanning microscope (Leica Microsystems, Wetzlar, Germany) with the fluorescence excitation at 488 nm with an argon laser and YFP excitation collection between 520 and 560 nm. For GFP the fluorescence excitation at 488 nm with an argon laser and GFP excitation collection between 500 and 515 nm. Images were processed using Leica SP2 Image Analysis Software.

2.1.13 Detection of epidermis cell death by Trypan Blue staining

The infected leaves were dissected from the plant five and ten days post-infectious. Trypan Blue Solution, 0.4% (Thermo Fisher Scientific Inc) was used for detection of cell death. This dye exclusion test stained blue in the permeable dead cells; on the other hand, the

impermeable viable cells would not take up the dye. The stock solution (0.4%) was diluted with ethanol (96 %; 1:3 v/v) for preparation of working solution. Then, the infected leaves were transferred into 50 ml eppendorf tube and fully immersed with the working staining solution. Next, the tube was placed in boiled water for 10 minutes. The stained infected leaves were left overnight in a tube at room temperature in a staining solution. The next morning, the leaves were de-stained 3 times (or more) with PBS solution (Appendix 4). If necessary, the leaves can be immersed with 95% ethanol overnight to ensure the chlorophyll was completely removed from the leaves. The leaves can be kept in the PBS solution for several days.

2.1.14 *Agrobacterium* mediated floral dip transformation

After verification by PCR, the *A. tumefaciens* harbouring the genes of interest was used to transform *A. thaliana* plants. A single colony of *A. tumefaciens* strain GV3101 with the gene of interest was inoculated into 5 ml of LB supplemented with Kanamycin/Spectinomycin (50 ng/µl), Gentamycin (30 ng/µl) and Rifampicin (25 ng/µl) in a 30 ml tube. The cultures were grown at 28 °C for 24 hours with constant shaking and were used as a liquid starter culture. Then, the entire 5 ml of the liquid starter culture was used to inoculate 100 ml of LB supplemented with the same concentration of antibiotics in two 250 ml Erlenmeyer flask. After that, the cultures were grown for approximately 6-8 hours at 28 °C shaking until the OD reached = 0.8-1.2. Next, 5 g sucrose (2 % (w/v) and 50 µl silvet (200 µl/L) was added to the transformation culture. Floral dip was performed based on Clough and Bent, (1998);

the *Arabidopsis* inflorescences were dipped into the culture for 90 seconds. The plants were covered with a cellophane bag (Dillewijn Zwapak) and placed in the glass house with 16/8- hour photoperiod at 20±2 °C. A week later, the bag was opened to allow the flowering to finish and seed to dry prior to collection.

2.1.15 Putative transformants screening

The collected seeds were germinated in plastic trays filled with Levington® M3 and bedding compost to screen the transformants. Resistant transformants were identified by spraying BASTA® four times, in which the plants possess of vector conferring resistance to the BASTA® will survive. The survived plants were transferred to 9 cm plastic pots filled with Levington® M3 and bedding compost. The plants were grown in the glass house with growth conditions consisted of a 16/8- hour photoperiod at 20±2 °C.

2.1.16 Sucrose method for crude DNA extraction from plants

Young leaves (2-3mm diameter) from plants were placed directly into 100 µl of Sucrose Buffer on ice. Sucrose buffer contains 50 mM of Tris-Cl (pH 7.5), 300 mM of NaCl, 300 mM of sucrose and 7.25 ml sterilized distilled water. The leaves are then crushed with a yellow pipette. The samples are then heated to 99–100°C for 10 min by using PCR machine and then briefly spun at 2000–6000 g for 5 sec. Samples are placed on ice until PCR or stored at -20 °C. This protocol was developed by Berendzen *et al*., (2005).

2.1.17 Putative transformants genotyping

The thermal profile of the reaction was programmed as follows. The PCR reaction mixture contains 5 µl of REDTaq® DNA Polymerase, 0.3 µm of specific forward and reverse primer and 50 ng of gDNA templates. Sigma Molecular water was added to a final volume of 10 µl. The reaction was programmed as follows: initial denaturation at 94 °C for 3 min, followed by 30 cycles at 94 °C for 30 sec, annealing was set at different temperatures depending on the experiments for 30 sec, and extension at 72 °C for at the different length depending on the experiment. The final extension was at 72 °C for 10 min followed by cooling to 10 °C. PCR products were separated by electrophoresis as in Section 2.1.4.

2.1.18 Qualitative and quantitative RT–PCR (qRT–PCR) analysis of expression

RNA extraction and CDNA synthesis were performed based on the protocol in Section 2.1.3. cDNA with the concentration of 2.0 µg was used for expression analysis. The thermal profile of the reaction was programmed as follows. The PCR reaction mixture contains 5 µl of REDTaq® DNA Polymerase 0.3 µm of specific forward expression primer, 0.3 µm of specific reverse primer, 0.5 µl of cDNA template. Primers used for House-keeping amplifications were Protein phosphatase 2A subunit 3 (*PP2A3)* Forward and Reverse primer. Sigma Molecular water was added to a final volume of 10 µl.

The thermal profile of the reaction was programmed as follows: initial denaturation at 94 °C for 2 min, followed by 30 cycles at 94 °C for 30 sec, annealing at different temperature depending on the experiment °C for 30 sec, and extension at 72 °C at different temperature depending on the experiment. The final extension was at 72 °C for 10 min followed by cooling to 10 °C. PCR products were separated by electrophoresis as mentioned in Section 2.1.4.

Furthermore, qRT-PCR analyses were performed in a 384-plate set-up using the Light Cycler 480 real-time PCR system (Roche Applied Science). Each qRT-PCR Reaction contained SYBR Green QPCR Master Mix (Fermentas) in a final volume of 9.2 µl with 0.22 µM of the gene specific forward primer and 0.22 µM of gene specific reverse primer and 34 ng of cDNA. Refer to Appendix 1 for primers sequence details.

Reactions were set up using the Maxima SYBR Green qPCR Master Mix in a final volume of 9 µL, containing 0.2 µL of cDNA and 0.2 µL of the appropriate primers (Table 2-2). PCR cycling conditions for amplification were 95°C for 10 min, then 40 cycles of 95°C for 30 s, 58°C for 1 min and 72°C for 1 min. All samples were run in tri-replicate. Relative expression levels were determined in comparison to PP2A3 using the $2^{(DCt)}$ analysis method (Livak and Schmittgen, 2001) and relative expression level of each gene in the line was set as 1.

Table 2-2: List of primers, PCR extension period and optimized annealing temperatures for qRT-PCR of seven selected proteases genes and one F-Box protein.

Gene	**Forward Primer (5'-3')**	**Reverse Primer (5'-3')**	**Optimized Tm**
PP2A3	PP2A3_1355F	PP2A3_1677R	60 °C
AT1G25510	1g25510_1041F	1g25510_1312R	61 °C
AT2G23945	2g23945_486F	2g23945_715R	60 °C
AT4G30030	4g30030_602F	4g30030_885R	61 °C
AT5G10760	5g10760_1165F	5g10760_1354R	61 °C
AT3G20015	3g20015_437F	3g20015_731R	60 °C
AT5G19120	5g19120_186F	5g19120_440R	61 °C
AT2G21430	SALK_095569_F	AT2g21430 EXP R	60 °C
At3g16210	At3g16210_454F	At3g16210_667R	60 °C

2.1.19 Phenotypic analysis

Phenotyping was conducted with 6-weeks old plant visually. Parameters taken were total no of silques on main stem, length of silique along with percentage of fertility.

2.1.20 Alexander's staining of pollen

Alexander staining (1969) is a well-known technique for assessing the viability of the pollen. Alexander's staining technique has an ability to discriminate the aborted and non-aborted pollen grains by staining the cell wall and intact cytoplasm. For viability test one hundred pollen grains were counted from every plant.

The buds were dissected under the stereomicroscope and were placed on a microscopic slide. The buds were cut in the bottom using the fine needle to remove the sepals and petals. The stamen containing anther and filament were put on a slide. Two drops of the Alexander's staining solution applied on the anther. A cover slip was placed on the sample and a soft pressure was applied on the cover slip to ensure the pollens were dissected from the anthers and to eliminate the bubble. The slides were incubated for 30 minutes at room temperature to make sure the pollen fully stained.

Alexander staining containing 10 ml 96% ethanol, 10 mg Malachite green (1 ml of 1% solution in 96% ethanol), 50 ml distilled water, 25 ml glycerol, 5 g phenol, 5 gm chloral hydrate, 50 mg acid fuchsin (5 ml of 1% solution in water), 5 mg Orange G (0.5 ml of 1% solution in water), Glacial Acetic acid to the final concentration of 4%. Red colour indicates viable pollen while the green colour indicates non-viable pollen (Alexander, 1969).

2.1.21 Microscopy condition

Individual flowers, buds, gynoecium and stamen were viewed by using Leica ZEISS stereomicroscope (Leica Microsysytems Limited, Switzerland) and the photos were taken with the Canon Powershot A620 camera and software. Pollen microscopy slides were inspected under a Leica DMRB microscope (Leica Microsysytems Limited, Switzerland) the photos were taken with Leica Application Suite Ver. 1.6.0. YFP was observed using confocal microscopy (TCS SP2; Leica

Microsystems (UK) Ltd, Milton Keynes, UK; 514 nm excitation wavelength, emission collection: 520–600 nm).

2.1.22 Statistical analysis

Data of silique length were obtained by the combined silique length of all siliques, sterile and non-sterile, divided by total number of siliques on the main stem. Data of silique length were subjected to analysis of variance (ANOVA) using GenStat 17th Edition statistical analysis (©VSN International Ltd, Hemel Hempstead, UK). The data were transformed if the data did not meet normal distribution assumptions. If the treatments were significant, h

2.2 Gene Manipulation Analysis of AT2G23945

2.2.1 Genotyping of SALK plants

All insertion lines were obtained from NASC. In short, pROK2 vector consists of CaMV 35S promoter and neomycin phosphotransferase II gene (NPT II) encoded for kanamycin antibiotics resistant activity was used for SALK line generations (Baulcombe *et al.*, 1986).

The PCR was divided into two parts by using a pair of gene specific primer and a pair of insertion specific primer. In gene specific primer, a pair of SALK 134917 gene-specific and right primer were designed to amplify the region with the T-DNA insertion region, where the

homozygous SALK lines will not produce any band and the heterozygous and wild-type will produce a band. For insertion specific primer, the T-DNA left border specific primer (LBb1.3) as suggested by the SALK's website and right primer of the was used, which in case of homozygosity and heterozygosity will produced a band of the predicted size (~805 bp) and the wild-type will not produce any band.

The thermal profile of the reaction was programmed as follows. The PCR reaction mixture contains 5 µl of REDTaq® DNA Polymerase, 0.3 µm of specific left primer/ LBb1.3, 0.3 µm of right primer and 50 ng of gDNA templates from both SALK lines. Sigma Molecular water was added to a final volume of 10 µl.

The reaction was programmed as follows: initial denaturation at 94 °C for 3 min, followed by 30 cycles at 94 °C for 30 sec, annealing at 58 °C for 30 sec, and extension at 72 °C for 40 sec. The final extension was at 72 °C for 10 min followed by cooling to 10 °C. PCR products were separated by electrophoresis as in section 2.1.5.

2.2.2 Quantitative and qualitative PCR (qRT–PCR) analysis of AT2G23945 expression

The PCR was split into two parts. Primers used for House-keeping amplifications were *PP2A3* Forward primer and *PP2A3* Reverse primer (28 cycles). Primers used for AT2G23945 expression were AT2G23945 forward expression primer (2g23945_486F) and AT2G23945 reverse expression primer (2g23945_715R) (28 cycles).

The qRT-PCR reaction and thermal profile of the reaction was programmed as mentioned in Section 2.1.18.

2.2.3 Auramine O staining of pollen exine

Anthers were dissected and place on a microscopy slide. Next, 30 µL of 0.001% auramine O staining solution, diluted in 50 mM Tris-HCl, pH 7.5 were applied for staining purposes. The slides were covered with a coverslip. The anthers were knocked gently to disperse the pollen from the anther. The pollen grains were incubated in a staining solution for five minutes for hydration purposes. The slides were observed using confocal microscopy 100 x magnifications with Excitation at 490 nM emission was collected at 520-560 nm.

2.2.4 Anther transverse sectioning study

Complete inflorescences from 6-weeks old plants were collected from SALK 135523 and WT-Col. The buds were dissected using a stereo-microscope and separated according to different stages of buds. Buds were vacuum infiltrated for 30 minutes in fresh fixative solution, consists of 4% paraformaldehyde in PBS (pH 7.2), 0.1% Triton X-100, 0.1%Tween 20. Then, buds were incubated in fresh fixative solutions over night at 4°C and followed by double washed with PBS (pH 7.2) at room temperature. Fixed tissues were dehydrated in ethanol series (30%, 50%, 70%, 90% and twice with 100% (v/v)) for at least 1 h in each. Resin infiltration were conducted using Technovit® 7100 from Electron Microscopy Sciences Company. The sections of 8 µm were prepared using

BRIGHT 5040 microtome and the sections were stained with 0.05%toulidine blue for 30 sec.

2.2.5 Cell wall and lignification stain for endothecium thickening study

Fresh open flowers were washed briefly using 1X PBS (pH 7.2) containing 2% (v/v) Tween 20 at room temperature. Next, the flowers were rinsed with 1X PBS without Tween 20. After that, the flowers were stained with 0.01% (w/v) acridine orange for 1 hour at room temperature followed by double washed with 1X PBS for 10 minutes each time. Subsequently, the flowers were stained with 0.00005% (0.5 µg/ml) EtBr for 1 hour at room temperature. Then, the flowers were washed twice with 1X PBS for 10 minutes each time. Next, the flowers were dissected in 10% (v/v) glycerol on a slide and cover with a coverslip. The slides were observed soon using confocal microscopy. EtBr stained lignified cell wall with red fluorescence by 590 nm excitation, with 590-620 collection emission; acridine orange stained non-lignified cell wall with green fluorescence by 520 nm excitation and emission was collected at 480-520.

2.2.6 Developing pollen viability study

To focus on the study of pollen development, DAPI (4',6-diamidino-2-phenylindole) staining was used to test pollen viability. Pollen Isolation Buffers consists of 100 mM NaPO4, pH 7.5, 1mM of EDTA, 0.1 % (v/v) Triton X-100 with 1 mg/ml Of DAPI was prepared. The

working concentration for this fluorescent was set to 3 µg/ml. The dilution was carried out using 50% glycerol. The buds were dissected under the stereomicroscope and were placed on a microscopic slide. Two drops of the DAPI staining solution was applied on the anther. A cover slip was placed on the sample and a soft pressure was applied on the cover slip to ensure the pollens were pulled out from the anthers and to eliminate the bubble. The slides were incubated overnight in a 4 °C fridge to ensure the pollens were fully stained. The next day, the photos of the pollens were taken with a fluorescent microscope.

2.2.7 Cloning of pCR8/GW/TOPO AT2G23945RNAi using Gateway® system

Gene cloning and Phusion PCR was conducted as mentioned in Section 2.1.4. Amplification of AT2G23945RNAi Phusion PCR fragment (102 bp) using RNAiF2g23945 forward primer and RNAiR2g23945 Reverse primer (YFP_NY-R). PCR extension period was 20s and optimized annealing temperature was 60 °C. The pCR8/GW/TOPO TA Cloning® Kit cloning of *Taq* polymerase was used to facilitate the cloning of amplified PCR products into a plasmid vector. The protocol use based on the protocol suggested by Invitrogen™ Corporation. The primer used for sequencing analysis was pCR8 GW 2.1 R.

2.2.8 Mach1™-T1R chemically competent *E. coli* transformation protocol

The protocol used was based on the protocol suggested by Invitrogen™ Corporation. Then 100 µl from the transformed cells were spread on a prewarmed LB medium with Spectinomycin antibiotic (50 ng/ul) and was incubated overnight at 37°C.

2.2.9 Gateway® LR clonase reaction of pK7GW1WG2:AT2G23945RNAi

Amplification of the fragment from the pCR8/GW/TOPO AT2G23945RNAi was performed to minimise the error for the LR reaction since the entry clone and the destination vector shared similar antibiotic resistance markers. The fragment of pCR8/GW/TOPO AT2G23945RNAi that includes attL region was amplified by employing PCR. The amplification of a fragment (443 bp) from pCR8/GW/TOPO AT2G23945RNAi was conducted using ROligo_New_CompF forward and ROligo_New_CompR reverse primer. The PCR extension period was 20s and optimized annealing temperature was 62°C.

Clones were transfered from a fragment of pCR®8/GW/TOPO® entry vector to a pK7GW1WG2 (VIB-UGent Center for Plant Systems Biology, Gent, Belgium), a Gateway® compatible destination vector for expression of the gene of interest. The LR Clonase reaction was l based on the protocol suggested by Invitrogen™ Corporation. The primer used for sequencing analysis was PK7 F2 and PH7GW1 R1.

2.2.10 Plasmid Colony PCR of pK7GW1WG2:AT2G23945RNAi

For plasmid colony PCR of pK7GW1WG2:AT2G23945RNAi destination vector, different primers were used to amplify different orientation of transcripts. The amplifications were conducted targeting two regions consisting of sense (443bp) and anti-sense strand (475bp). Sense strand amplification using RNAiF2g23945 with PK7 F2 vector specific forward primer. Anti-sense strand amplification RNAiF2g23945 with PH7GW1 R1 vector specific reverse primer. For both set of primers, PCR extension period was set at 30s with optimized annealing temperatures of 62 °C.

2.2.11 Putative transformants screening

The putatively transformed seeds (T_1) were collected and screened on MS medium containing Spectinomycin (50 ng/µl) (See section 2.1.15).

2.2.12 Putative pK7GW1WG2:AT2G23945RNAi transformants genotyping

The gDNA were extracted using the Sucrose technique (Section 2.1.16). Details of the programs as in Section 2.1.17. The positive control was the extracted plasmid from the same constructs that were used for transformation and the negative controls were WT-Col

gDNA and water. Sense strand amplification was conducted using RNAiF2g23945 with PK7 F2 while anti-sense strand amplification was conducted using RNAiF2g23945 with PK7 R2. For both set of primers, the PCR extension period was set at 30s and optimized annealing temperatures for plant genotyping was 62 °C .

2.3 Transcriptional and Translational Fusion of AT2G23945

2.3.1 Leaf DNA extraction protocol

The DNAs were extracted by using ISOLATE II Genomic DNA Kit (Bioline).

2.3.2 Cloning of pCR8/GW/TOPO AT2G23945Pro entry clones with three different size of insertions

For putative promoters cloning and Phusion PCR details, see Section 2.1.4. Different primers were used to generate different length of promoters (Table 2-3). For Cloning protocols see Section 2.2.10 and Section 2.2.11. Full-length sequence of three different promoter lengths are listed in Appendix 3.

Table 2-3. List of primers, PCR extension period and optimized annealing temperature for AT2G23945pro Phusion PCR.

Insertion Length (bp)	Forward Primer (5'-3')	Reverse Primer (5'-3')	PCR Extension Period	Optimized Tm
437	2G23945PRO F4	2G23945Pro R1	1min 30s	62 °C
980	2G23945PRo F5	2G23945PRo R1	1min 30s	62 °C
1572	2G23945PRO F1	2G23945Pro R1	1min 30s	62 °C

2.3.3 pCR8/GW/TOPO AT2G23945Pro colony PCR and sequencing

Details for the plasmid colony PCR available in Section 2.1.8. Different primers were used to amplify different lengths of insertions for plasmid colony PCR (Table 2-4). Multiplications were conducted by liquid culture supplemented with Spectinomycin (50 ng/µl) antibiotic. The primer used for sequencing analysis was pCR8 GW 2.1 R.

Table 2-4. List of primers, PCR extension period and optimized annealing temperature for pCR8/GW/TOPO AT2G23945pro colony PCR.

Insertion Length (bp)	Forward Primer (5'-3')	Reverse Primer (5'-3')	PCR extension period	Optimized Tm
437	2G23945PRO F4	pCR8 GW 2.1 R	2min	58°C
980	2G23945PRo F5	pCR8 GW 2.1 R	2min	58°C
1572	M13F (-20)	2G23945Pro R1	2min	58°C

2.3.4 Gateway® LR Clonase Reaction of pGWB3:AT2G23945PRO:GUS

Clones were transfer from a fragment of pCR8/GW/TOPO AT2G23945Pro entry vector to a pGWB3:AT2G23945PRO:GUS (Invitrogen™ Corporation), a Gateway® compatible destination vector for GUS expression. The LR Clonase reaction was based on the protocol suggested by Invitrogen™ Corporation.

2.3.5 pGWB3:AT2G23945PRO:GUS colony PCR and sequencing

Please refer to Section 2.11 for colony PCR details. For destination vector colony PCR, 2G23945PRO F4 forward and GUSpGWB3_138R reverse primer were used to amplify different lengths of constructs. PCR extension period was set at 40s and optimized annealing temperatures for LR reaction colony PCR was set at 58 °C. Multiplications were conducted by liquid culture supplemented with Kanamycin (50 ng/µl) antibiotic. The primers used for sequencing analysis were GUSpGWB3_138R for -437 bp promoter insertion, 2G23945Pro F5 and GUSpGWB3_138R for -980 bp promoter insertion, 2G23945Pro F1 and GUSpGWB3_138R for -1572 bp.

2.3.6 Putative pGWB3:AT2G23945PRO:GUS transformants genotyping with three different size of insertions.

The gDNAS were extracted using the Sucrose technique (Section 2.1.16). The genotyping of the plants was conducted by using the

same primers and conditions as in Section 2.3.5. Details of the programs as in Section 2.1.17. The positive control was the extracted plasmid from the same constructs that were used for transformation and the negative controls were WT-Col gDNA and water.

2.3.7 GUS staining method

The GUS solution was prepared based on (Jefferson *et al.,* 1987) (Appendix 5). For staining, intact tissue from different part of plants including inflorescence, leaf, silique, internodes, plantlets were placed in 500 µl Eppendorf tube. Then, the tissues were immersed with 150 µl of GUS substrate solution and incubated for 30 hours at 37°C.

After an overnight incubation, the tissues were clarified to remove the chlorophyll. The tissues were soaked with acidified methanol at 55°C for overnight followed by soaking with 7% NaOH in 60% EtOH at room temperature for 20 minutes. Then, the clarification continued with the soaking with different concentration of ethanol (40%, 20 % and 10%) at room temperature for 20 minutes each. Finally, the tissues were soaked in 50% Glycerol for preservation. The tissues were observed under the microscope as described in Section 2.1.21.

2.3.8 Cloning of pCR8/GW/TOPO AT2G23945 translational fusion entry clones with three different size of insertions

For Phusion PCR details, see Section 2.1.4. Different primers were used to generate different lengths of transcripts (Table 2-5). Full-length of insertions for three translational fusion constructs are listed in Appendix 3.

Table 2-5: List of primers, PCR extension period and optimized annealing temperature for AT2G23945pro Phusion PCR.

Insertion Length (bp)	Forward Primer (5'-3')	Reverse Primer (5'-3')	PCR Extension Period	Optimized Tm
1811	2G23945PRO F4	2G23945 TF R1	3min	58°C
2345	2G23945PRo F5	2G23945 TF R1	3min	58 °C
2946	2G23945PRO F1	2G23945 TF R1	4min	58°C

2.3.9 pCR8/GW/TOPO AT2G23945 translational fusion colony PCR and sequencing

Different primers were used to amplify different length of insertions for plasmid colony PCR (Table 2-6). Multiplications were conducted by liquid culture supplemented with Spectinomycin (50 ng/µl) antibiotic. The primers used for sequencing analysis were M13F(-20) for 1811bp insertion and pCR8 GW 2.1 R for 2345bp and 2946bp insertions.

Table 2-6. List of primers, PCR extension period and optimized annealing temperature for pCR8/GW/TOPO: AT2G23945 colony PCR.

Insertion Length (bp)	Forward Primer (5'-3')	Reverse Primer (5'-3')	PCR Extension Period	Optimized Tm
1811	M13F (-20)	2G23945Pro R1	1min 30s	58°C
2345	2g23945_976F	pCR8 GW 2.1 R	1min	58 °C
2946	2g23945_976F	pCR8 GW 2.1 R	1min	58 °C

2.3.10 Gateway® LR Clonase Reaction of pGKGWG:AT2G23945:GFP

For LR reaction details, see Section 2.2.12. For pGKGWG:AT2G23945:GFP destination vector colony PCR, 2g23945_486F forward and YFP_NY-R reverse primer were used to amplify different constructs. The PCR extension period was set at 1min 30s and optimized annealing temperature was set at 58°C. Multiplications were conducted by liquid culture supplemented with Kanamycin (50 ng/µl) antibiotic. The Primers that was used for sequencing analysis were YFP_NY-R and 2g23945_715R for 1811bp insertion construct, 2G23945Pro F5, 2g23945_715R and YFP_NY-R for 2345bp insertion construct, 2G23945Pro F1, 2g23945_715R and YFP_NY-R for 2946bp insertion construct.

2.3.11 Putative pGKGWG:AT2G23945:GFP transformants genotyping with three different size of insertions.

The genotyping of the plants was conducted by using the same primers and conditions as in Section 2.3.12. Details of the programs as in Section 2.1.17.

2.3.12 qRT–PCR expression analysis

The RT-PCR was split into two parts, PCR A and PCR B. Primers used for House-keeping amplifications (PCR A) were *PP2A3* Forward primer and *PP2A3* Reverse primer while expression analysis conducted using AT2G23945 Forward Expression Primer (2g23945_976F) coupled with YFP_NY reverse primer (35 cycles). See 2.1.18 for details of the PCR contents and RT-PCR conditions.

CHAPTER 3: ANALYSIS OF TAPETUM ASSOCIATED GENES

3.1 Introduction

Tapetum development and its subsequent degeneration, plays a key part of the pollen development process, thus it is important to identify which genes are activated or repressed during the process. Tapetum degeneration has been associated with PCD (Van Durme and Nowack, 2016) and a number of transcriptional regulators have been reported as involved in *Arabidopsis* and rice (See Section 1.2.4). Two important players in this pathway are *ABORTED MALE STERILITY (AMS)* (Sorensen *et al.*, 2003; Xu *et al.*, 2010, 2014) and *MALE STERILITY1* (*MS1)* (Ito and Shinozaki, 2002; Wilson *et al*., 2001; Yang *et al.*, 2007) (See Section 1.4). Mutants in these genes show altered tapetum development and changes to tapetum PCD.

Microarray investigation by Pearce *et al.*, (2015) revealed a number of genes that showed altered expressions pattern in the *ams* and *ms1* mutants compared to WT. As part of an integrated regulatory networks, alteration of one transcription factors may lead to alterations of other elements in the related regulatory network (Tang *et al.*, 2007). Deep understanding of *MS1* and *AMS* gene networks will provide critical information about the interactions of these transcription factors and their possible downstream targets, including proteases that may provide insight into tapetum PCD-specific pathways.

Reverse genetic approaches can be used to reveal the biological function and aid characterisation of the targeted gene in these pathways. Overexpression of a target gene, using a promoter of a highly expressed gene (Prelich, 2012), is frequently adopted to investigate gene function, since this approach has the advantage of not being influenced by any potential redundancy in the target gene.

A number of genes have been identified as having altered expression in the *ams* and *ms1* mutants, which have altered tapetum PCD (Pearce *et al.*, 2015) (https://www.cpib.ac.uk/tools-resources/resources/flowernet/). Proteases have been shown to be key drivers of PCD (Parish and Li, 2010); the role of aspartyl proteases, cysteine proteases and F-Box protein in plant PCD is discussed in greater details in Section 1.6. A subset of these differentially expressed genes are putative proteases and therefore may be associated with tapetum PCD. Six different Aspartyl proteases genes (AT1G25510, AT2G23945, AT4G30030, AT5G10760, AT3G20015 and AT5G19120); one cysteine proteinase gene (AT2G21430) and a single F-box protein encoded gene, AT3G16210 gene, show specific expression changes in the *ams* and *ms1* mutants (Pearce *et al.*, 2015). Nevertheless, it is unclear whether these genes have a role in pollen development and tapetum PCD, and whether they are directly, or indirectly regulated by AMS and MS1. This chapter describes the over-expression of these genes in stable *Arabidopsis* transgenics and phenotyping of any associated changes in fertility and development.

3.2 Results

3.2.1 Candidate Gene Analysis

The expression pattern for each of the proteases was extracted from the Electronic Fluorescent Pictograph (e-FP) browser (Winter *et al.*, 2007) and from the Flowernet expression profiles (Pearce *et al.*, 2015) (https://www.cpib.ac.uk/tools-resources/resources/flowernet/).

3.2.1.1 AT2G23945 Aspartyl Protease

AT2G23945 gene has been selected since it encodes for a aspartyl-protease and shows high specificity level of expression during flower development (Figure 3-1) based on the Electronic Fluorescent Pictograph (e-FP) browser developed by Winter *et al.*, (2007). Furthermore, the AT2G23945 appeared to be associated with pollen development and expressed in the tapetum (Pearce *et al.*, 2015)

Several plant-specific regulators have been extensively investigated for tapetum degeneration, male sterility and PCD. Two transcription factors that have a key role in this process are *AMS* and *MS1*, both are involved in pollen development in *Arabidopsis* (See Section 1.5 and 3.1). Figure 3-2 shows microarray analysis of the gene expression across various samples and the interaction with various transcription factors. The gene is down regulated in mutants of four

major tapetum transcription factor, *MS1, AMS, DYT1* and *MYB80*. For *ms1* mutant, the gene is highly expressed in young buds compared to old buds. Furthermore, the gene is severely down regulated in the *myb80* mutant and *ms1* old stage mutant (Pearce *et al.*, 2015). However, it is still unclear whether this gene is directly or indirectly regulated by these transcription factors.

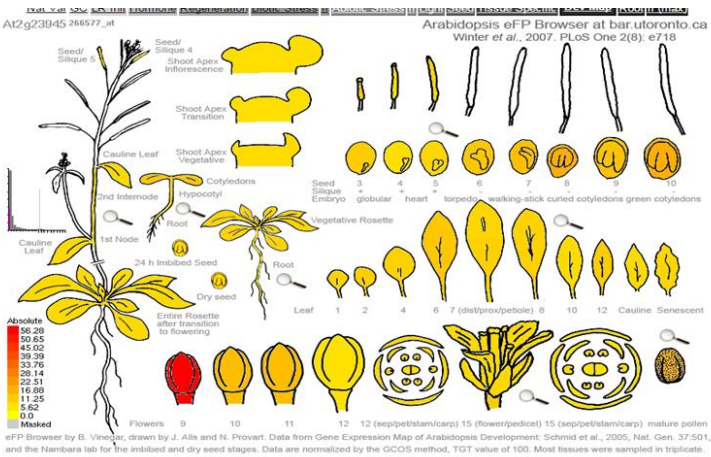


Figure 3-1. Level of expression of AT2G23945 in different tissues of *Arabidopsis*. Figure from http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi (Winter *et al*., 2007).

Based on Flowernet data (Figure 3-3), the *AMS* transcription factor appears to down regulate the expression level of AT2G23945 as shown in the *ams* mutant. This gene is highly expressed in the wild-type plant particularly during the Meiosis stage. However, gene expression is slightly lower during mitosis stage I compared to the previous stage in wild-type plant while this gene recorded its highest expression in *ams* mutant. Furthermore, this gene reached the lowest level of expression in wild-type and *ams* mutant during the bicellular stage of pollen development. Interestingly, the gene shows slightly increased expression during post-mitosis II as shown in the wild type plant and *ams* mutant (Figure 3-3).

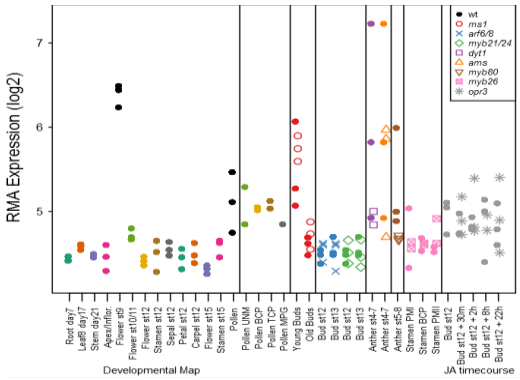


Figure 3-2.Floral microarrays data for expression of the AT2G23945 gene in multiple transcription factor mutants (Pearce *et al.*, 2015). This plot can be accessed online at https://www.cpib.ac.uk/tools-resources/resources/flowernet/.

Key: Mei: Meiosis, PM I: Post Meiotic I, BCP: Bicellular pollen, PM II: Post Meiotic II; Sample 1: Mature Flowers; Sample 7: Immature buds/inflorescence meristem.

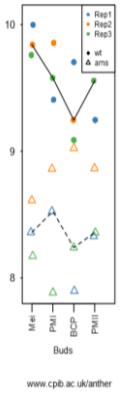


Figure 3-3. The expression of AT2G23945 gene in the *ams* mutant anther and wild type control(Pearce *et al.*, 2015). This plot can be accessed online at https://www.cpib.ac.uk/tools-resources/resources/flowernet/.

Key: Mei: Meiosis, PM I: Post Meiotic I, BCP: Bicellular pollen, PM II: Post Meiotic II.

3.2.1.2 AT1G25510 Aspartyl Protease

AT1G25510 shows high levels of expression during active embryo and flower development (Figure 3-4) (Pearce *et al.*, 2015). Figure 3-5 shows the microarray analysis of gene expression across various floral samples and in the *ams* and *ms1* mutants. The gene shows higher levels of expression in the young buds of the *ms1* mutant compared to WT-control (Figure 3-5). Conversely, the gene expression is higher in WT-control compared to the *ms1* mutant in older buds. Similarly, the *AMS* transcription factor may be involved in the repression of this gene, as indicated by the increased expression in the *ams* mutant plot, with an exception of bicellular stage where the expression levels are comparable with the WT-control (Pearce *et al.*, 2015).

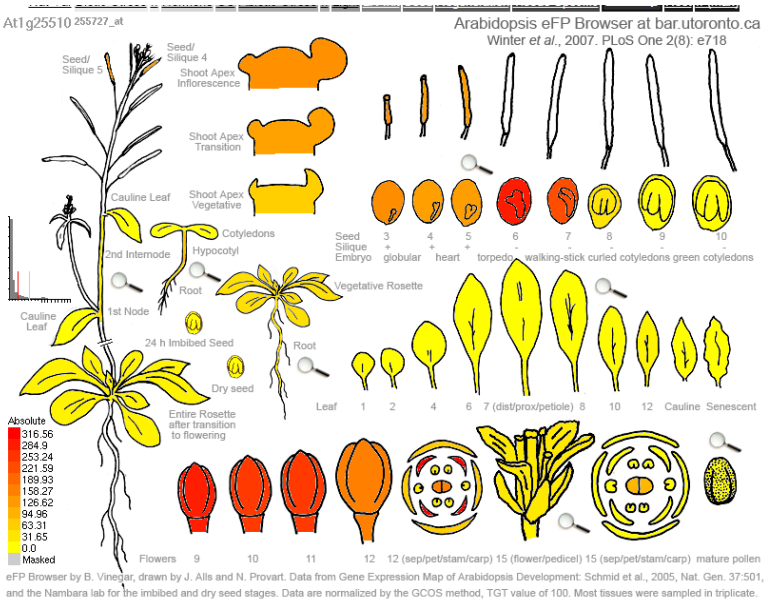


Figure 3-4. Expression analysis of AT1G25510 in different *Arabidopsis* tissues. Figure from http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi (Winter *et al*., 2007).

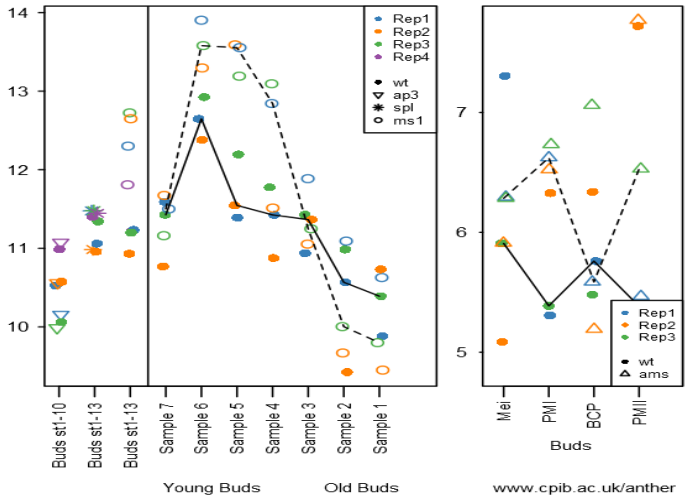


Figure 3-5. The expression of AT1G25510 gene in the *ms1* and *ams* mutants and wild-type buds (Pearce *et al.*, 2015). These plots can be accessed online at (https://www.cpib.ac.uk/tools-resources/resources/flowernet/).

Key: Mei: Meiosis, PM1: Post Meiotic 1, BCP: Bicellular pollen, PM11: Post Meiotic 2; Sample 1: Mature Flowers; Sample 7: Immature buds/inflorescence meristem.

3.2.1.3 AT4G30030 Aspartyl Protease

At4G30030 shows high levels of expression during flower stage 9 (Figure 3-6). The gene has similar expression for most samples tested in both WT-control and the *ms1* and *ams* mutant samples (Figure 3-7). The gene appears to be maximal in anthers at the meiosis stage in both WT-control and the *ams* mutant.

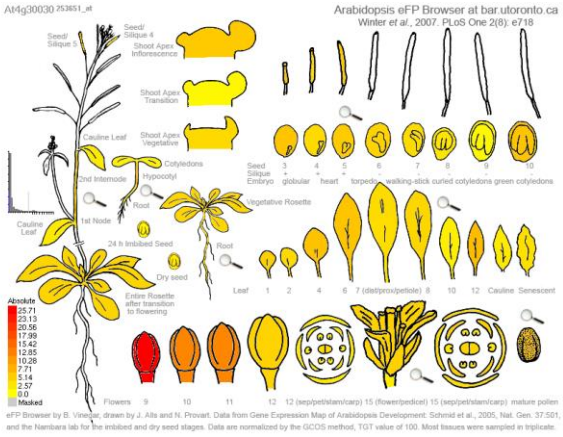


Figure 3-6. Expression analysis of AT4G30030 in different *Arabidopsis* tissues. Figure from http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi (Winter *et al*., 2007).

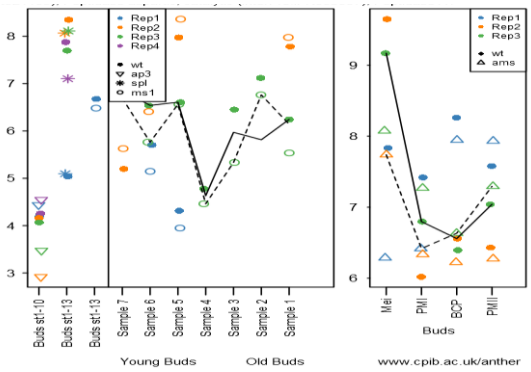


Figure 3-7. Expression of the AT4G30030 gene in buds of the *ms1* and *ams* mutants and wild-type (Pearce *et al.*, 2015). These plots can be accessed online at https://www.cpib.ac.uk/tools-resources/resources/flowernet/.

Key: Mei: Meiosis, PM1: Post Meiotic 1, BCP: Bicellular pollen, PM11: Post Meiotic 2; Sample 1: Mature Flowers; Sample 7: Immature buds/inflorescence meristem.

3.2.1.4 AT5G10760 Aspartyl Protease

High levels of AT5G10760 expression are seen in leaves and mature flower sepals (Figure 3-8); a similar pattern of expression is seen in the *ms1* and *ams* mutants compared to the WT-control (Figure 3-9). Expression is greatest in buds at stage 10 and the meiosis stage.

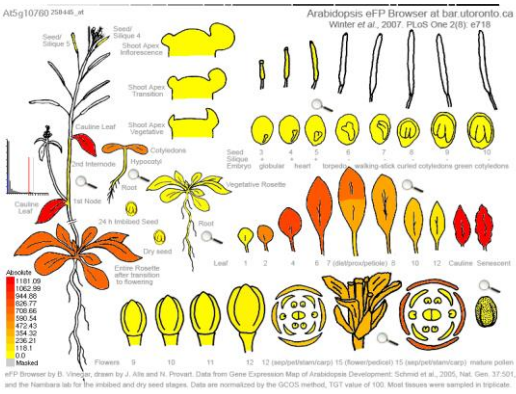


Figure 3-8. Expression analysis of the AT5G10760 gene in different *Arabidopsis* tissues. Figure from http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi (Winter *et al*., 2007).

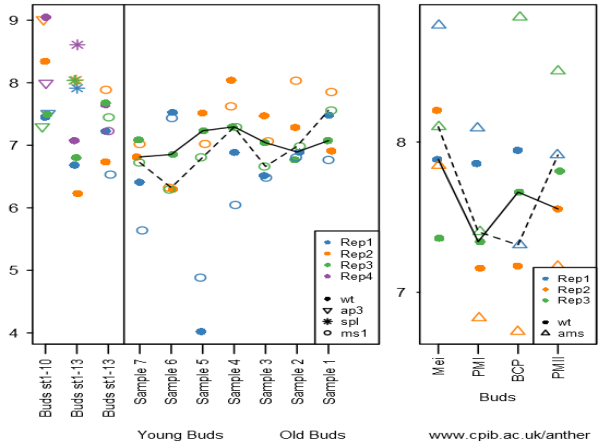


Figure 3-9. Expression of AT5G10760 gene in the *ms1* and *ams* mutant and wild-type buds (Pearce *et al.*, 2015). These plots can be accessed online at https://www.cpib.ac.uk/tools-resources/resources/flowernet/.

Key: Mei: Meiosis, PM1: Post Meiotic 1, BCP: Bicellular pollen, PM11: Post Meiotic 2; Sample 1: Mature Flowers; Sample 7: Immature buds/inflorescence meristem.

3.2.1.5 AT3G20015 Aspartyl Protease

AT3G20015 shows high levels of expression in the shoot apex, inflorescence and embryo, and moderate levels of expression during flower development (Figure 3-10). Similar patterns of expression are seen in the *ms1* mutant and WT buds, however expression is increased in *ams1* mutant buds compared to the WT (Figure 3-11). This suggests that the AMS may be involved in directly or indirectly repressing At3G20015 expression.

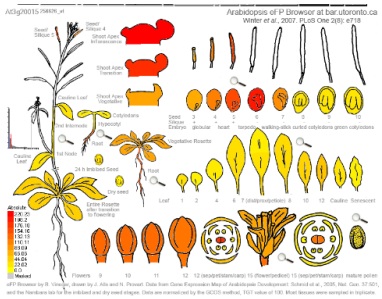


Figure 3-10. Expression analysis of AT3G20015 in different *Arabidopsis* tissues. Figure from http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi (Winter *et al*., 2007).

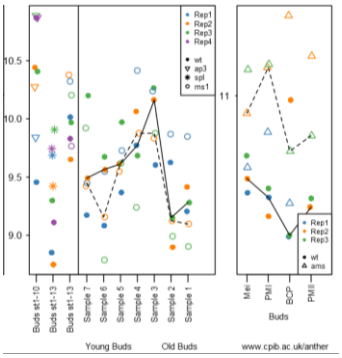


Figure 3-11. Expression of the AT3G20015 gene in the *ms1* and *ams* mutant and wild-type buds (Pearce *et al.*, 2015). These plots can be accessed online at https://www.cpib.ac.uk/tools-resources/resources/flowernet/).

Mei: Meiosis, PM1: Post Meiotic 1, BCP: Bicellular pollen, PM11: Post Meiotic 2; Sample 1: Mature Flowers; Sample 7: Immature buds/inflorescence meristem.

3.2.1.6 AT5G19120 Aspartyl Protease

AT5G19120 shows a high level of expression in many organs, including flowers, leaves and the vegetative rosette (Figure 3-12). Expression appears highest during the later (stage 13) stages of flower development and appears to be consistent in wild type and the *ams* and *ms1* mutants (Figure 3-13).

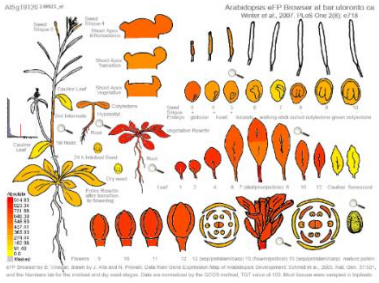


Figure 3-12. Level of expression of AT5G19120 in different tissues of *Arabidopsis*. Figure from http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi (Winter *et al*., 2007).

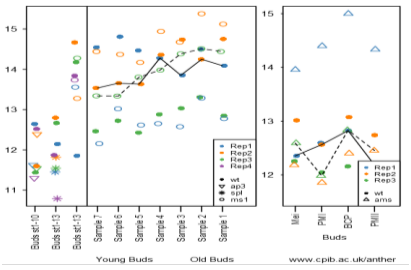


Figure 3-13. Expression of the AT5G19120 gene in the *ms1* and *ams* mutant and wild-type buds (Pearce *et al.*, 2015). These plots can be accessed online at https://www.cpib.ac.uk/tools-resources/resources/flowernet/.

Key: Mei: Meiosis, PM1: Post Meiotic 1, BCP: Bicellular pollen, PM11: Post Meiotic 2; Sample 1: Mature Flowers; Sample 7: Immature buds/inflorescence meristem.

3.2.1.7 AT2G21430 Cysteine proteinase

AT2G21430 shows a high level of expression in young buds, anthers, embryos and young siliques (Figure 3-14). Expression is generally higher in *ms1* mutant buds but reduced in *ams* buds (Figure 3-15). This suggests that AMS may be involved in directly or indirectly upregulating AT2G21430 expression.

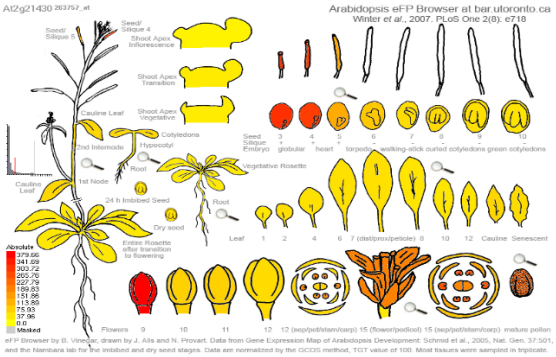


Figure 3-14. Level of expression of AT2G21430 in different tissues of *Arabidopsis*. Figure from http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi (Winter *et al*., 2007).

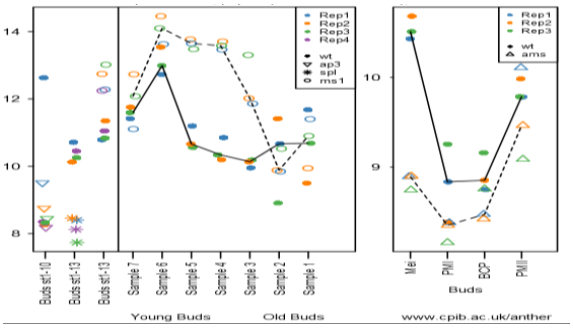


Figure 3-15. Expression of the AT2G21430 gene in the *ms1* and *ams* mutant and wild-type buds (Pearce *et al.*, 2015). These plots can be accessed online at https://www.cpib.ac.uk/tools-resources/resources/flowernet/.

Key: Mei: Meiosis, PM1: Post Meiotic 1, BCP: Bicellular pollen, PM11: Post Meiotic 2; Sample 1: Mature Flowers; Sample 7: Immature buds/inflorescence meristem.

3.2.1.8 AT3G16210 F-Box Protein

No information is available regarding expression of AT3G16210 from Electronic Fluorescent Pictograph (e-FP) browser (Winter *et al.*, 2007). However, Flowernet indicates that expression is highest around pollen mitosis I and is reduced in the *ams* mutant and increased in the *ms1* mutant (Figure 3-16). This suggests that AMS may be involved in directly or indirectly upregulating AT3G16210 expression, whilst MS1 may be involved in directly or indirectly repressing AT3G16210 expression.

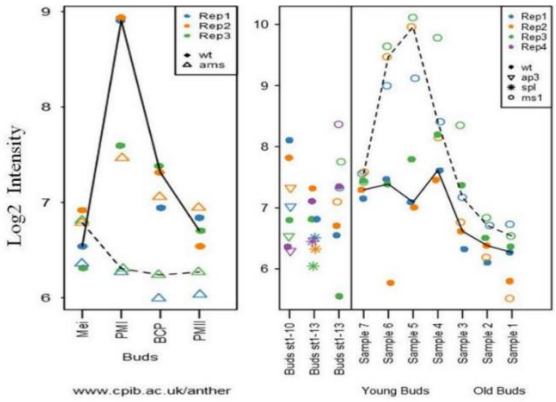


Figure 3-16. Expression of the AT3G16210 gene in the *ms1* and *ams* mutants and wild-type buds (Pearce *et al.*, 2015). These plots can be accessed online at https://www.cpib.ac.uk/tools-resources/resources/flowernet/.

Key: Mei: Meiosis, PM1: Post Meiotic 1, BCP: Bicellular pollen, PM11: Post Meiotic 2; Sample 1: Mature Flowers; Sample 7: Immature buds/inflorescence meristem.

3.2.2 Cloning of targeted genes

The eight candidate genes with expression during anther development and altered expression in the *ms1* and/or *ams1* mutants, were analysed for their potential roles during tapetum degradation by overexpression analysis. The preparation of the overexpression constructs were conducted by two approaches, the seven proteases constructs (AT1G25510, AT2G23945, AT4G30030, AT5G10760, AT3G20015, AT5G19120, AT2G21430) were prepared using Gibson Assembly (Section 2.1.7), and the single F-Box protein (AT3G16210) construct was prepared by Gateway cloning (Invitrogen) (Section 2.3.4). Amplification of the full-length cDNA was performed using gene specific forward and reverse primers as listed in Table 2-1 (See Section 2.1.4).

Amplifications generated products of the predicted sizes of between ~1.29 kbp to ~1.74 kbp and were cloned into the pUBC destination vector and confirmed by PCR screening (Figure 3-17). The selected plasmids were then transformed into *Agrobacterium* (GV3101) by electroporation, verified by PCR (Figure 3-18) and transformed into WT-Col *Arabidopsis*.

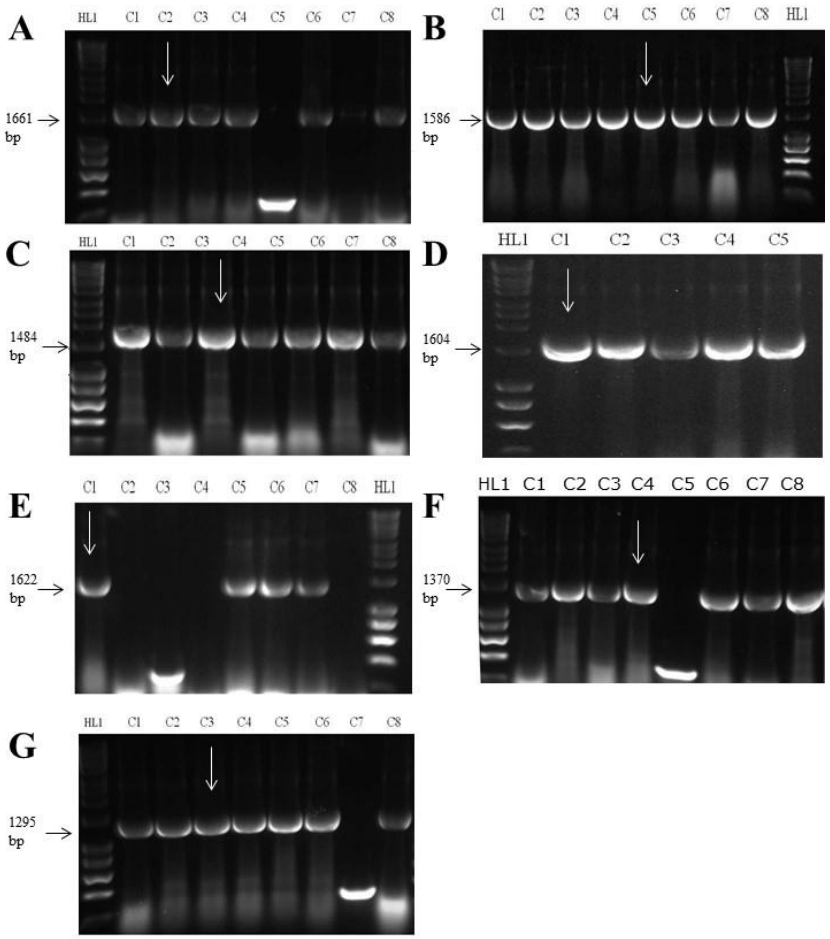


Figure 3-17: Colony PCR screening of insert and vector sequence from destination vectors containing putative proteases involved with tapetum PCD, A) AT1G25510, B) AT2G23945, C) AT4G30030, D) AT5G10760, E) AT3G20015, F) AT5G19120 and G) AT2G21430. The amplifications were conducted using pUBQ10 forward primer and gene specific reverse primers (Table 2-2). The arrows points to the specific colonies used for further analysis. C1-C8: Different colonies of the destination vector, HL1: HyperLadder 1.

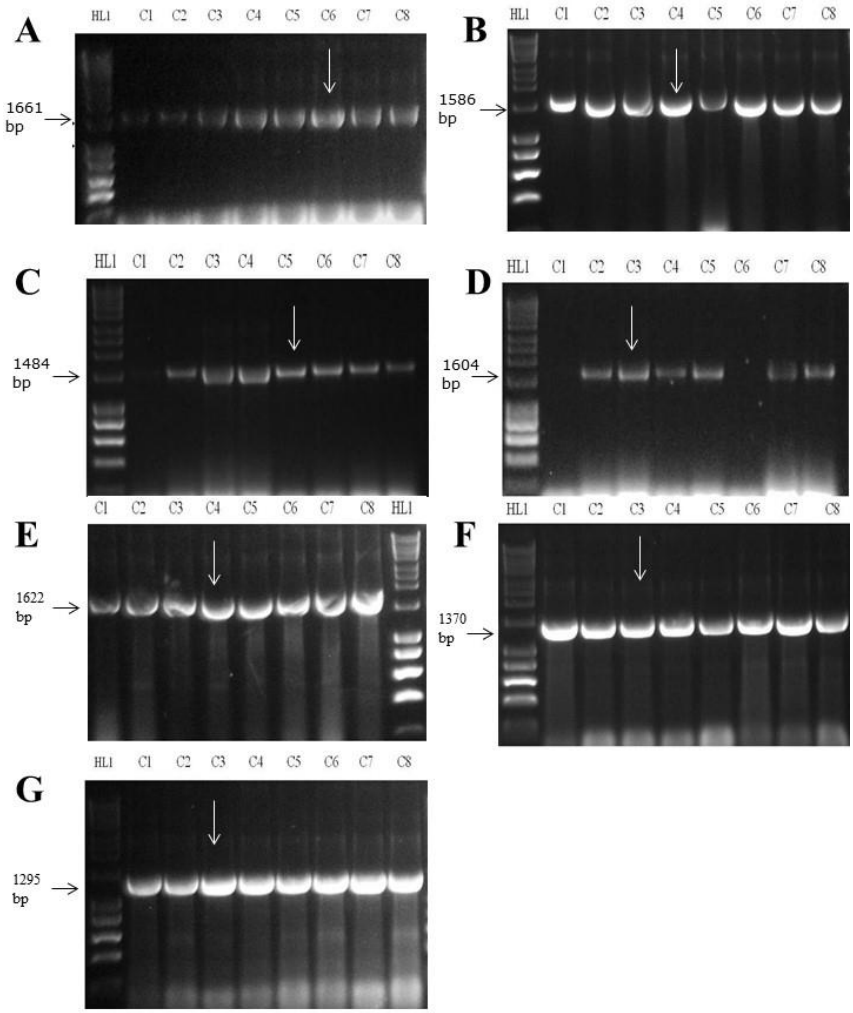


Figure 3-18. Colony PCR screening of insert and vector sequence from *Agrobacterium* GV3101 containing putative proteases, A) AT1G25510, B) AT2G23945, C) AT4G30030, D) AT5G10760, E) AT3G20015, F) AT5G19120 and G) AT2G21430. The amplifications were conducted using pUBQ10 forward primer and a gene specific reverse primer. The arrows point to the specific colony used for plant transformation. C1-C8: Different colonies of the destination vector, HL1: HyperLadder 1.

3.2.3 Analysis of transient expression of the selected transgenes

The overexpression vector also can be used for ectopic expression to uncover the protein subcellular-localisation (Prelich, 2012). The overexpression gene constructs (Section 3.3.1), containing the pUBQ10 promoter and a C-terminal YFP tagged protein to enable visualisation of the transgene, alongside the empty starting vector, were introduced into *A. tumefaciens* and used to infiltrate into *N. benthamiana* epidermal cells for transient expression. The empty vector; GV3101::pUBQ10:YFP construct showed general expression in the nucleus and plasma membrane (Figure 3-19). The YFP tagged fusion proteins were also distributed in the nucleus and membrane for all the protease genes investigated (Figure 3-19). Interestingly, for several proteases genes (AT1G25510, AT4G30030, AT5G19120 and AT2G21430), localization could be detected in stomatal guard cells as shown in Figure 3-20, whereas this was not seen for the wild type controls. Analysis of further AT2G23945 transient expression is presented in Section 5.2.2 and Section 5.2.3.

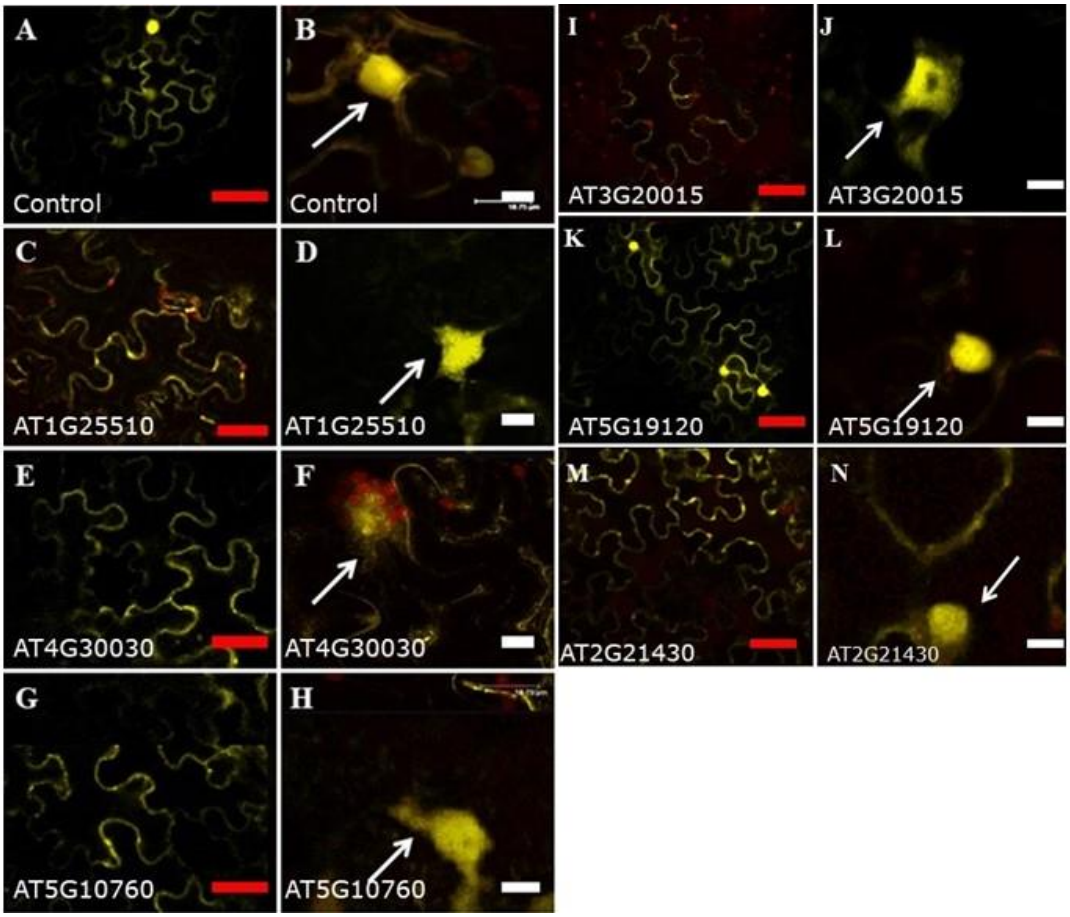


Figure 3-19. Localisation of transiently expressed YFP fusions for various putative proteases after *Agrobacterium* infiltration of *N. benthamiana*; GV3101::pUBQ10:Proteases Genes:YFP and GV3101::pUBQ10:YFP as a control. Pictures were taken 4-days post inoculation. Representative transient expression from three biological replicates are shown for each sample. Control (A and B), AT1G25510 (C and D), AT4G30030 (E and F), AT5G10760 (G and H), AT3G20015 (I and J), AT5G19120 (K and L) and AT2G21430 (M and N). Arrow points to *N. benthamiana* nuclei. Red bar: 50 µm and White Bar: 10 µm.

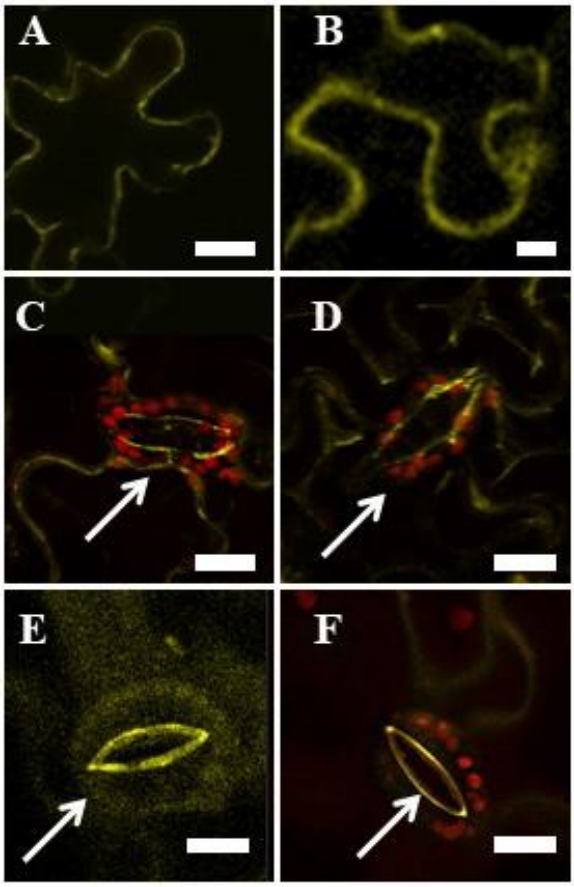


Figure 3-20. Localisation of transiently expressed YFP fusions for five different proteases after infiltration; GV3101::pUBQ10:Proteases Genes: YFP in stomatal guard cells. Pictures were taken 4-days post inoculation. Representative transient expression from three biological replicates are shown for each sample. Arrows indicate *N. benthamiana* stomatal guard cells and infiltration of C) AT1G25510, D) AT4G30030, E) AT5G19120 and F) AT2G21430. Control GV3101::pUBQ10:YFP construct (A and B) did not show any expression in stomatal guard cells. Bar: 20 µm.

3.2.4 Cell lysis analysis and trypan blue staining

The impact of the proteases on cell viability was determined by infiltration of the seven constructs (GV3101::pUBQ10:Proteases:YFP) into *N. benthamiana* leaves and subsequent cell viability assessment using trypan blue staining. Inviable cells stained blue due to the uptake of trypan blue, whereas viable cells with intact impermeable membranes did not take up the dye. Ectopic expression of the proteases and a control construct were investigated 5-days and 10-days post infiltration. Five days post infiltration, the control GV3101::pUBQ10:YFP construct and the protease genes did not appear to cause any ectopic cell death in the *N. tabacum* leaves (Figure 3-21). All infiltrated leaf areas started to show chlorosis 5-days after infiltration. However, 10-days post infiltration, all constructs (control GV3101::pUBQ10:YFP and the with seven proteases constructs) caused cell death (Figure 3-22). All infiltrated leaf areas exhibited intense regions of chlorosis 6 to 7-days after infection and showed strong trypan blue staining.

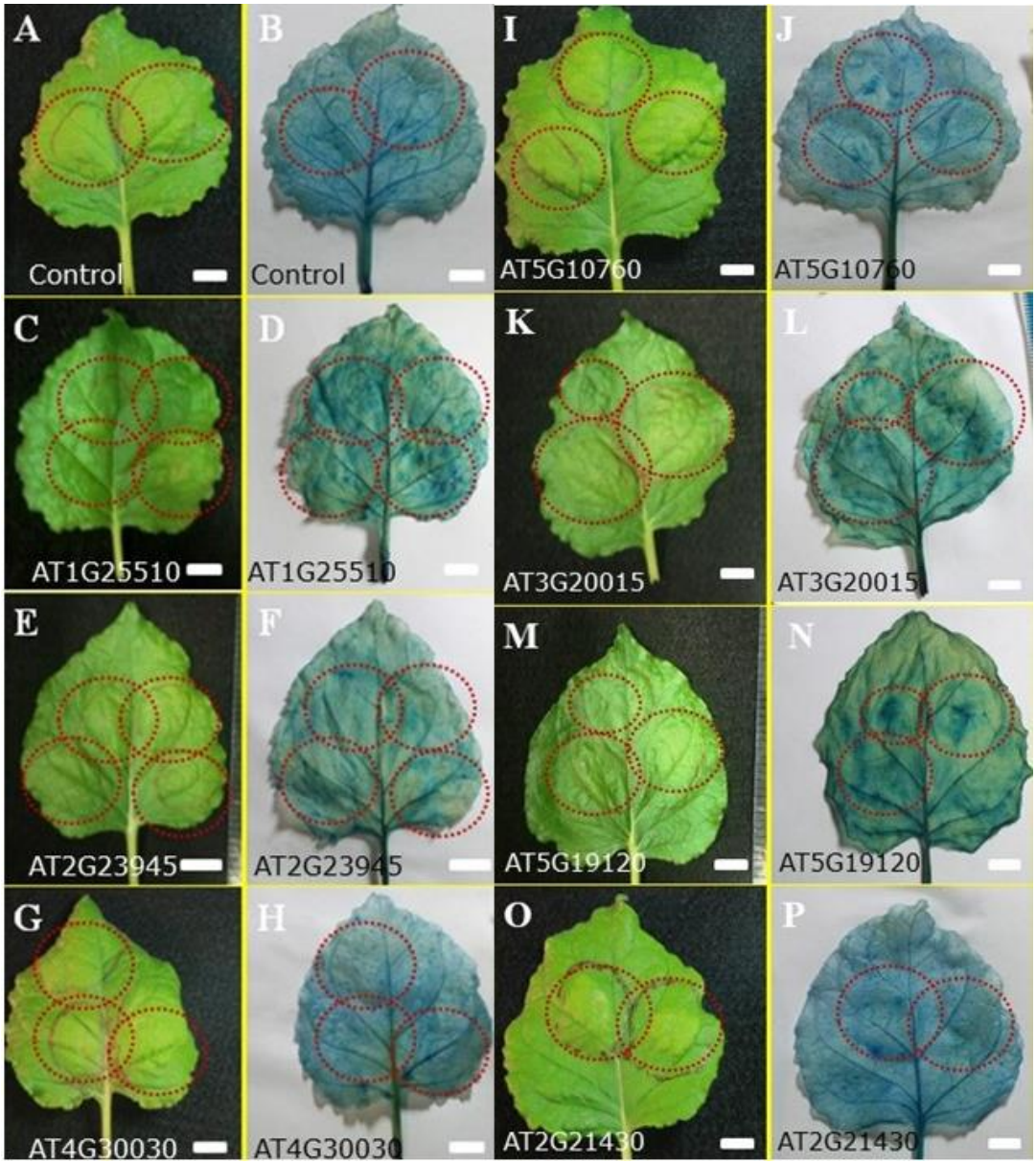


Figure 3-21. Ectopic expression and trypan blue staining of different protease genes; GV3101::pUBQ10:Proteases:YFP and GV3101::pUBQ10:YFP as a control. Pictures were taken 5-days post infiltration. Representative trypan blue staining from three biological replicates are shown for each sample. Control (A and B), AT1G25510 (C and D), AT2G23945 (E and F), AT4G30030 (G and H), AT5G10760 (I and J), AT3G20015 (K and L), AT5G19120 (M and N) and AT2G21430 (O and P). Control GV3101::pUBQ10:YFP construct with general ectopic expression which did not cause any cell death. Red-dashed circle marks the infiltration site. White Bar: 1 cm.

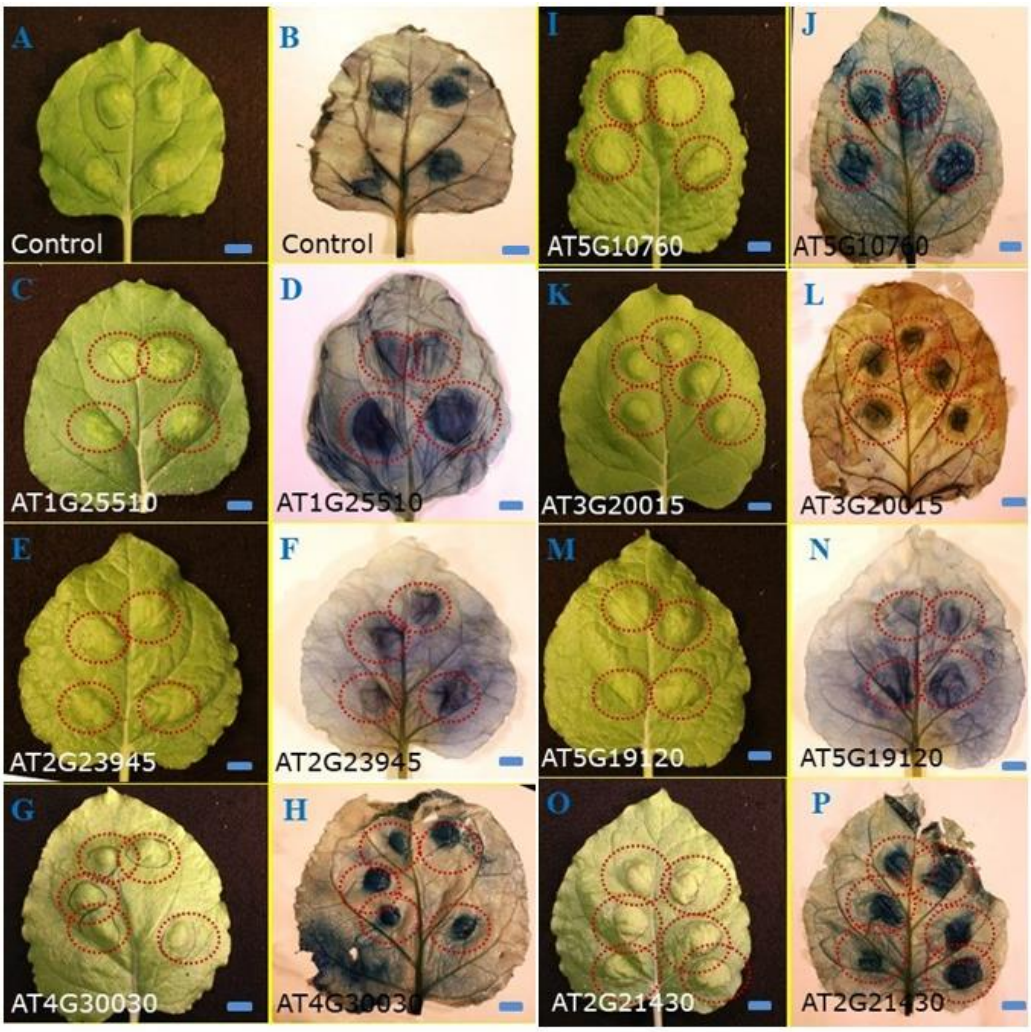


Figure 3-22. Ectopic expression and trypan blue staining of different protease genes; GV3101::pUBQ10:Protease:YFP and GV3101::pUBQ10:YFP as a control. Pictures were taken 10-days post infiltration. Representative trypan blue staining from three biological replicates are shown for each sample. Control (A and B), AT1G25510 (C and D), AT2G23945 (E and F), AT4G04460 (G and H), AT4G30030 (I and J), AT5G10760 (K and L), AT3G20015 (M and N), AT5G19120 (O and P) and AT2G21430 (Q and R). Control GV3101::pUBQ10:YFP construct with general ectopic expression which caused cell death. Red-dashed circle marks the infiltration site. White Bar: 1 cm.

3.2.5 Analysis of Protease Gene Function Through Overexpression

The impact of overexpression of the eight putative proteases was determined on tapetum PCD and male fertility. Putatively transformed seeds (T_1) were screened using BASTA ® treatment and PCR to confirm the integration of the transgenes. PCR amplifications were conducted targeting a region consisting of the YFP-tagged protein and the CDS; PCR analysis confirmed multiple positive transgenic plants. The fertility of these transgenic lines was scored from the main stem of six weeks-old plants.

3.2.5.1 AT2G23945 Overexpression Plants

To test the ability of AT2G23945 gene in altering tapetum development, analysis of the AT2G23945 gene using overexpression vector of pUBQ10::YFP was conducted (See section 3.3.1). The schematic map of the vector is presented in Figure 3-23.

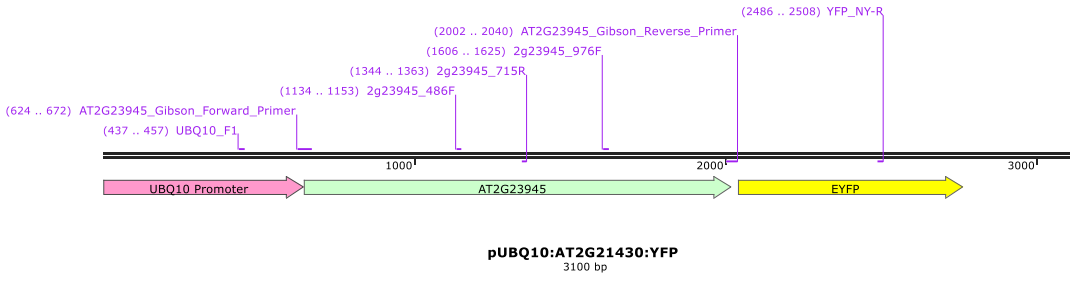


Figure 3-23. The schematic map of pUBQ10:AT2G23945:YFP overexpression vector. The pink arrow indicates the UBQ10 promoter, green arrow indicates the insertional gene and the yellow arrow indicates the YFP protein tag. The locations of the primers are shown on the map.

The putatively transformed seeds were collected and BASTA® screened (Section 2.1.15). Selective putative transgenic plants were transferred to soil and screened by PCR for of the transgene. PCR analysis showed 12 transgenic plants carried the pUBQ10:AT2G23945:YFP insertion construct with the size of ~1375 bp (Figure 3-24).

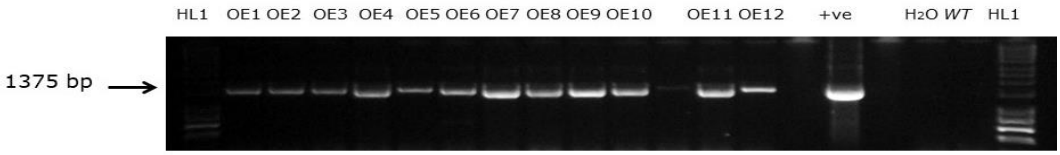


Figure 3-24: The presence of the insertion in twelve T_1 plants was confirmed by PCR analysis. Amplification of pUBQ10:AT2G21430:YFP fragment (1375 bp) using gene specific Forward primer (2g23945_486F) and pUBC YFP Reverse primer (YFP_NY-R). OE1-OE12: Twelve different AT2G23945OE plants, +ve: pUBQ10:AT2G23945:YFP vector as positive control WT: WT-*Col* genomic DNA as negative control, H_20: Water as a negative control, HL1: HyperLadder 1.

Expression analysis of AT2G23945 gene was conducted using inflorescence tissues in six AT2G23945OE plants using qRT-PCR to analyse the effects of overexpressing the gene driven by the pUBQ10 promoter. cDNA expression analysis was conducted using *PP2A3* Housekeeping primers and AT2G23945 forward (2g23945_486F) and reverse expression primers (2g23945_715R). As expected, qRT-PCR revealed that the level of AT2G23945 expression in all

AT2G23945OE plants lines were increased (Figure 3-25) from 46 to 142-fold, compared to the WT-Col.

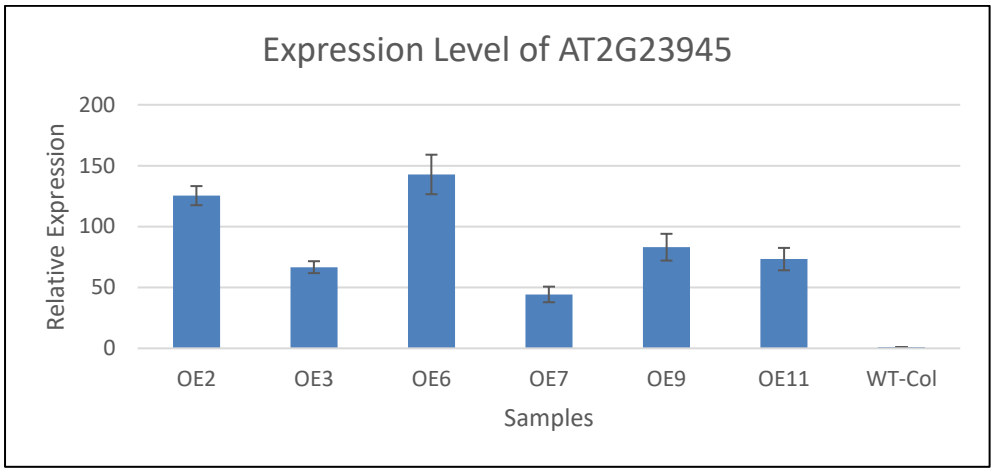


Figure 3-25. Expression pattern of AT2G23945 gene using inflorescence tissues in six T_1 independent transgenic plants harbouring pUBQ10:AT2G23945:YFP. Bar represents a mean value (n=3) of AT2G23945 expression relative to the Housekeeping *PP2A3* gene. The relative expression level of each gene in the line was set as 1. Error bars represent standard deviation. OE2, OE3, OE6, OE7, OE9 and OE11: Six different homozygous plants of AT2G23945OE.

Twelve T_1 independent transgenic plants were obtained, and six were analysed for phenotypic changes during 6-weeks growth (Figure 3-26). The AT2G23945OE transgenic lines showed partial sterility compared to the WT-Col (Figure 3-26). AT2G23945OE plants showed a decrease of the total number of siliques on the main stem. The sterility percentage rates varied for the six AT2G23945OE plants, 18.2 %, 22.2 %, 16.7 %, 12.5 %, 18.8 %, and 22.2 % for OE2, OE3, OE6, OE7, OE9, and OE11 respectively. All showed reduced fertility compared to the wild-type control, which exhibited ~10 % sterility (Table 3-1). The initiation of flowering for AT2G23945OE plants was comparable to the WT-Col.

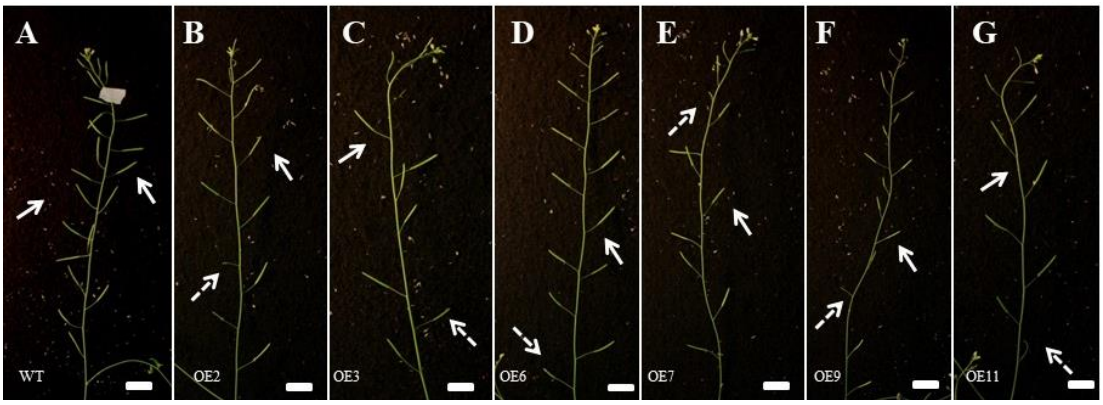


Figure 3-26. Characterization of T_1 AT2G21430OE transgenic plants with phenotype. Main stem of WT-Col as a control (A) and six independent AT2G21430OE plants (B-G). Images were taken of the six weeks old plants. White arrow: Fertile silique. Dashed- arrow: Sterile silique. Bar = 1 cm.

Based on the GenStat statistical analysis (See section 2.1.22), overexpression of the AT2G23945 gene in all six transgenic plants showed significant effect on mean of silique length when compare to the WT-Col; no significant differences were observed between the different transgenic lines (Figure 3-27).

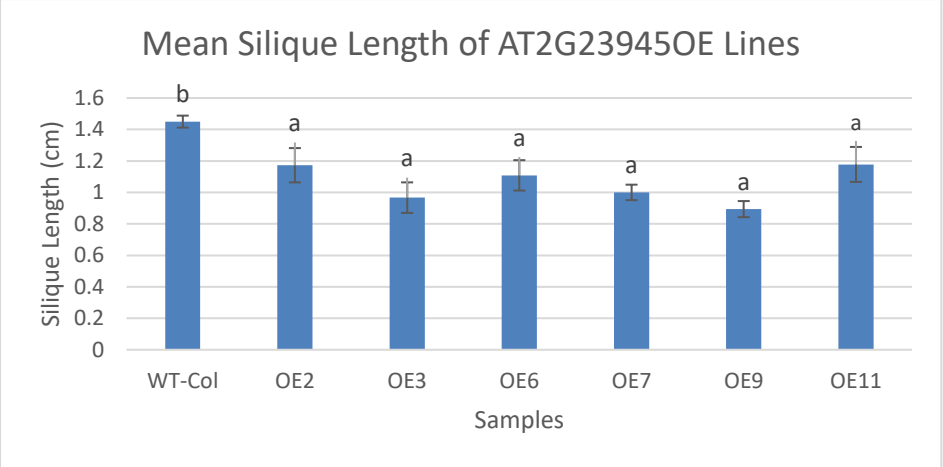


Figure 3-27. Mean silique length (cm), indicating fertility, between six independent overexpression transgenic plants OE2, OE3, OE6, OE7, OE9, OE11) and the WT-Col control after 6 weeks. Mean silique length was obtained for all siliques, sterile and non-sterile by dividing total length by total number of siliques on the main stem. Identical letter signifies no significant difference where $P<0.05$; ANOVA, Fisher's Least Significant Difference test. Error bars represent standard error.

Table 3-1. Summary of the analysis of fertility in the AT2G23945OE plants. The total number of fully developed siliques and percentage of sterility on main stem were determined alongside transcript level changes in six independent T_1 AT2G23945OE lines and WT-Col.

Sample	**Total no. of siliques**	**Percentage of sterility on main stem (%)**	**Transcript Level Changes**
Wild-type Col	20	10	1
OE2	11	18.2	125 X
OE3	9	22.2	65 X
OE6	12	16.7	142 X
OE7	8	12.5	43 X
OE9	16	18.8	81 X
OE11	9	22.2	72 X

3.2.5.2 AT1G25510 Overexpression Plants

PCR analysis showed four transgenic plants contained the pUBQ10:AT1G25510:YFP insertion construct (1109 bp; Figure 3-28 and 3-29).

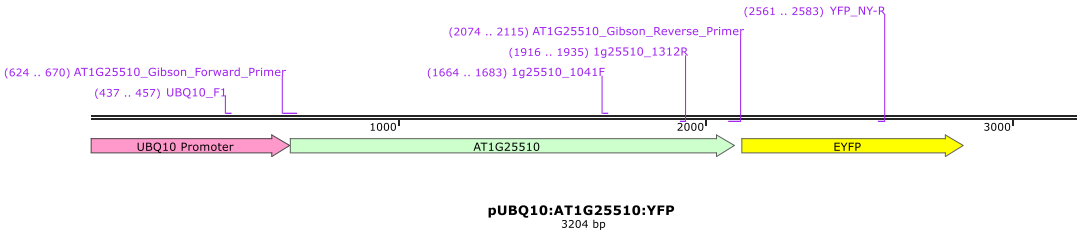


Figure 3-28. Schematic map of pUBQ10:AT1G25510:YFP overexpression vector. The pink arrow indicates the UBQ10 promoter, green arrow indicates the insertional gene and the yellow arrow indicates the YFP protein tag. The locations of the primers are shown on the map.

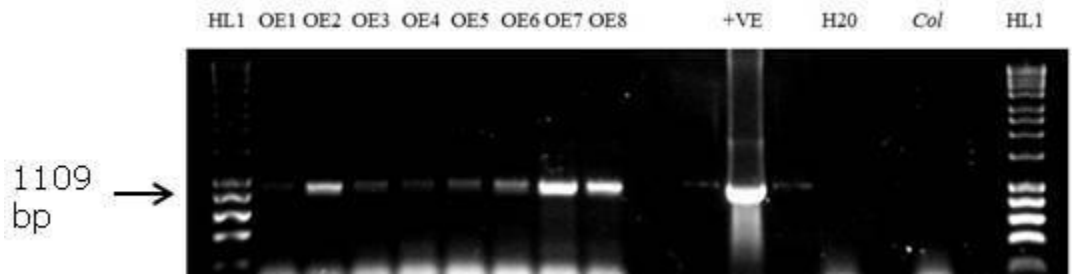


Figure 3-29. PCR screening for overexpression transgene in six T_1 plants. Amplification of pUBQ10:AT1G25510:YFP (1109 bp) using gene specific Forward primer (1g25510_1041F) and pUBC YFP Reverse primer (YFP_NY-R), indicating integration of the gene into the plant chromosome. OE1-OE8: Twelve different AT2G23945OE plants, +ve: pUBQ10:AT1G25510:YFP vector as positive control WT: WT-*Col* genomic DNA as negative control, H_2O: Water as negative control, HL1: HyperLadder 1.

qRT-PCR expression analysis was conducted in six independent AT1G25510OE lines (1g25510_1041F and 1g25510_1312R); expression was normalised using *PP2A3* Housekeeping gene expression. As expected, AT1G25510 expression increased in all AT1G25510OE lines, varying from 29 to 430-fold compared to the WT-Col (Figure 3-30, Table 3-2).

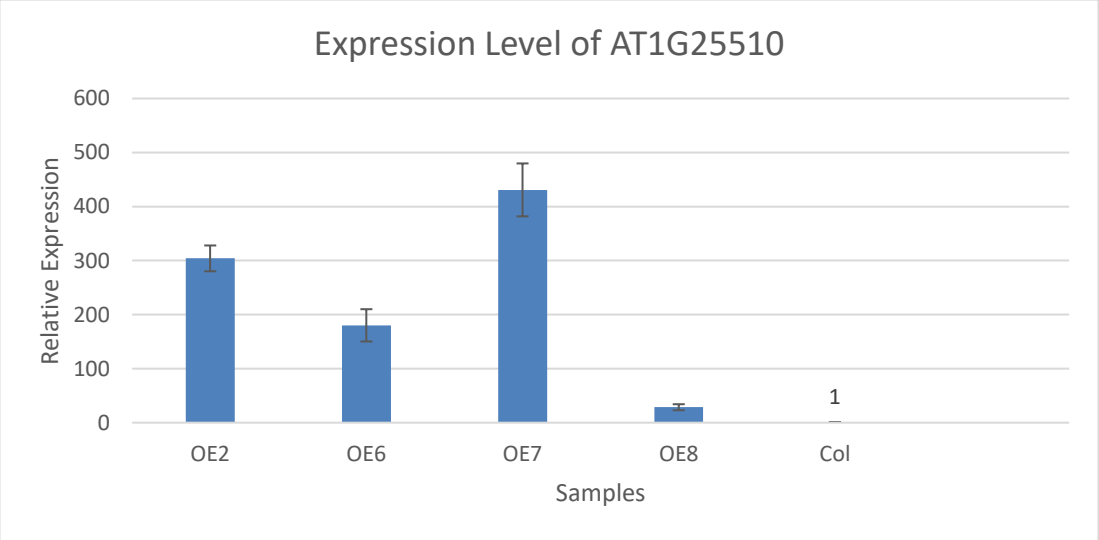


Figure 3-30. qRT-PCR expression analysis of AT1G25510 gene in four independent T_1 over expression transgenic lines harbouring the pUBQ10:AT1G25510:YFP construct. Bar represents a mean value (n=3) of AT1G25510 expression, normalised using *PP2A3* expression. Relative expression level is compared to WT-*Col*. Error bars represent standard deviation. OE2, OE6, OE7 and OE8: Four different transgenic plants of AT1G25510OE.

Four T_1 independent transgenic lines were analysed for phenotypic changes during 6-weeks growth. Variability in fertility was seen for the different lines as compared to the WT-Col (Figure 3-31). Sterility varied from 0 %, 7%, 46.4%, and 41.4%, for OE2, OE6, OE7, and OE8 respectively. All plants showed reduced fertility, except for OE2, which exhibited 100% fully developed siliques, however, OE2 plant showed decreased numbers of siliques on the main stem, with just

18 siliques compared to 28 in WT-Col (Table 3-2); this was the only AT1G25510OE line to show this.

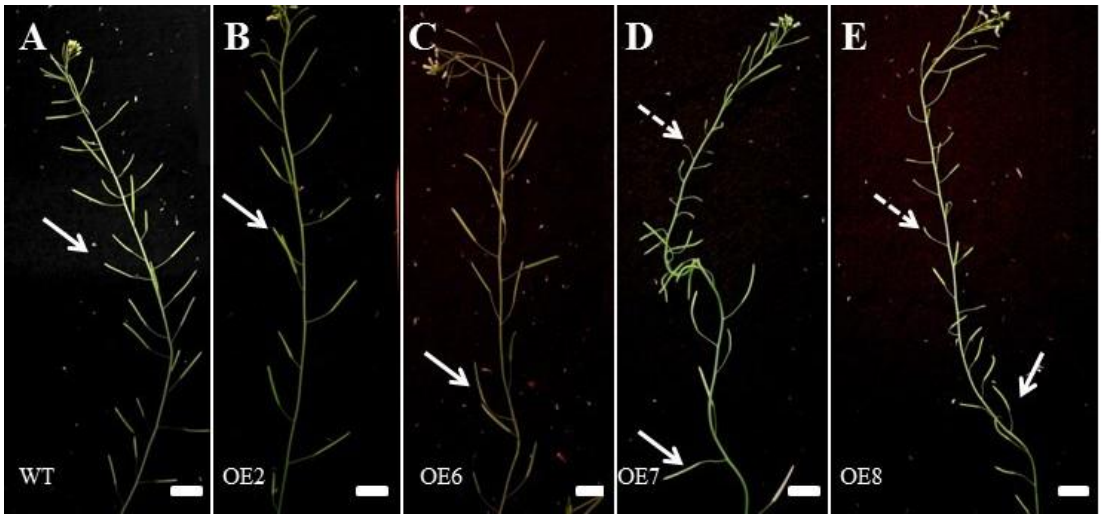


Figure 3-31. Characterization of T_1 AT1G25510OE transgenic plants with different phenotypes. Main stem of WT-*Col* as a control (A) and four independent AT2G21430OE plants (B-E). Images were taken of the six weeks old plants. White arrow: Fertile siliques. Dashed-arrow: Sterile siliques. Bar = 1 cm.

Alexander's stain (Alexander, 1969) was used to assess pollen viability on whole anther squashes; red stained pollen indicates the pollen is viable, whilst non-viable pollen was green/unstained. Viable and non-viable pollen were seen in all AT1G25510OE plants. OE2 and OE6 recorded 10% non-viable pollen, whereas ~50% of pollen from OE7 and OE8 plants were inviable. In comparison, almost all pollen from WT-Col stained red and were viable (Figure 3-32). Ectopic overexpression of AT1G25510 appeared to negatively impact on pollen fertility, although no direct correlation between the level of expression and extent of viability was apparent, however an impact on fertility was evident at higher AT1G25510 expression levels.

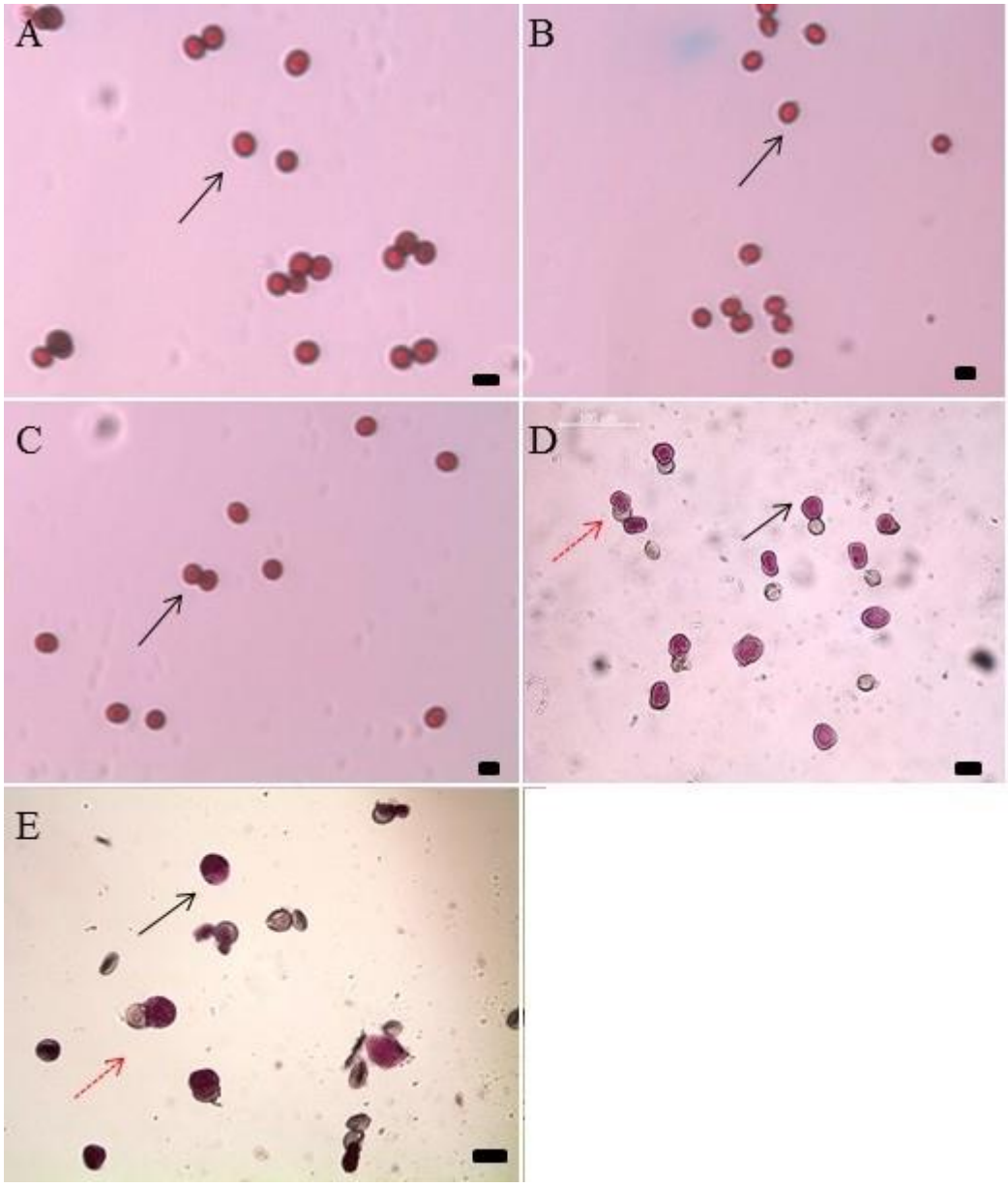


Figure 3-32. Pollen viability assessment using Alexander's staining for all four AT1G25510OE plants where one hundred pollen grains were counted from every independent plant. Red stained pollen indicates viable pollen, while the non-viable pollen were unstained. Pollen from A: WT-Col, B:OE2; C:OE6; D:OE7; E:OE8 plant. Pollen from two transgenic plants OE2 (B) and OE6 (C) appeared identical to the WT-Col pollen (A) whilst OE7 (D) and OE8 (E)showed abnormal pollen grains. Solid arrows indicate viable pollen grains which are stained red while non-viable pollen is shown by dashed arrows. Bar = 10 μm.

Based on the statistical analysis (See Section 2.1.22), overexpression of the AT1G25510 gene in OE7 and OE8 showed a significantly reduced mean silique length compared to the WT-Col, whilst OE2 and OE6 did not differ significantly (Figure 3-33). The flowers, stamens, filaments and carpels of AT1G25510 OE plants were of equivalent size and growth characteristics to the WT-Col. Furthermore, pollen dehiscence appeared to occur normally during for all AT1G25510OE plants.

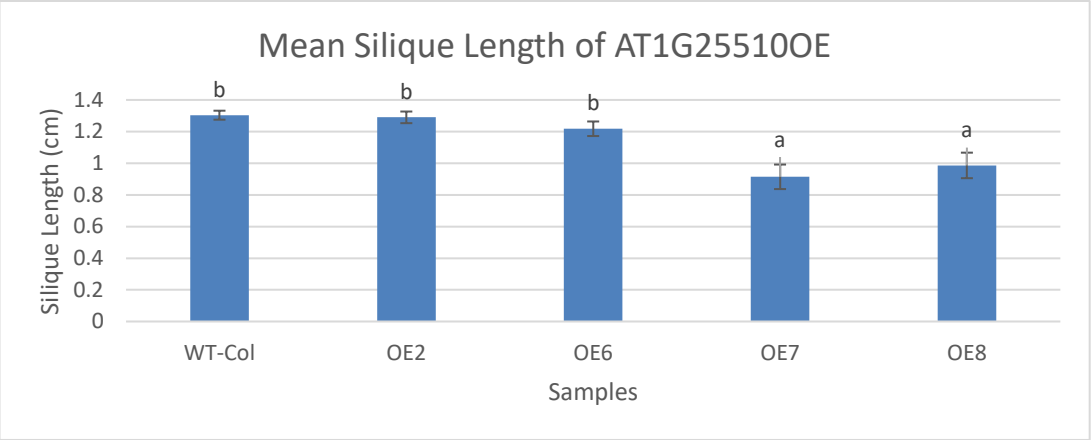


Figure 3-33. Silique length (cm) of six independent overexpression transgenic plants and the WT-Col after 6 weeks. Mean silique length was obtained for all siliques, sterile and non-sterile by dividing total length by total number of siliques on the main stem. Identical letter signifies no significant difference where $P<0.05$; ANOVA, Fisher's Least Significant Difference test. Error bars represent standard error.

Two transgenic plants of AT1G25510OE, OE7 and OE8 which exhibited the most severe reduction in fertility were chosen for further investigation. OE7 and OE8 pollen were tested by Auramine O staining (Kommareddi *et al.*, 1984) to check the nature of their pollen wall (exine). Abnormal pollen from these two lines (OE7 and OE8) had disrupted surfaces, with shrunken exine (Figure 3-34B-F); in contrast, normal WT-Col pollen was round and of regular shape

(Figure 3-34A). The shrunken exine structure seen in the OE lines suggests that the manipulation of AT1G25510 expression impacts on pollen wall development.

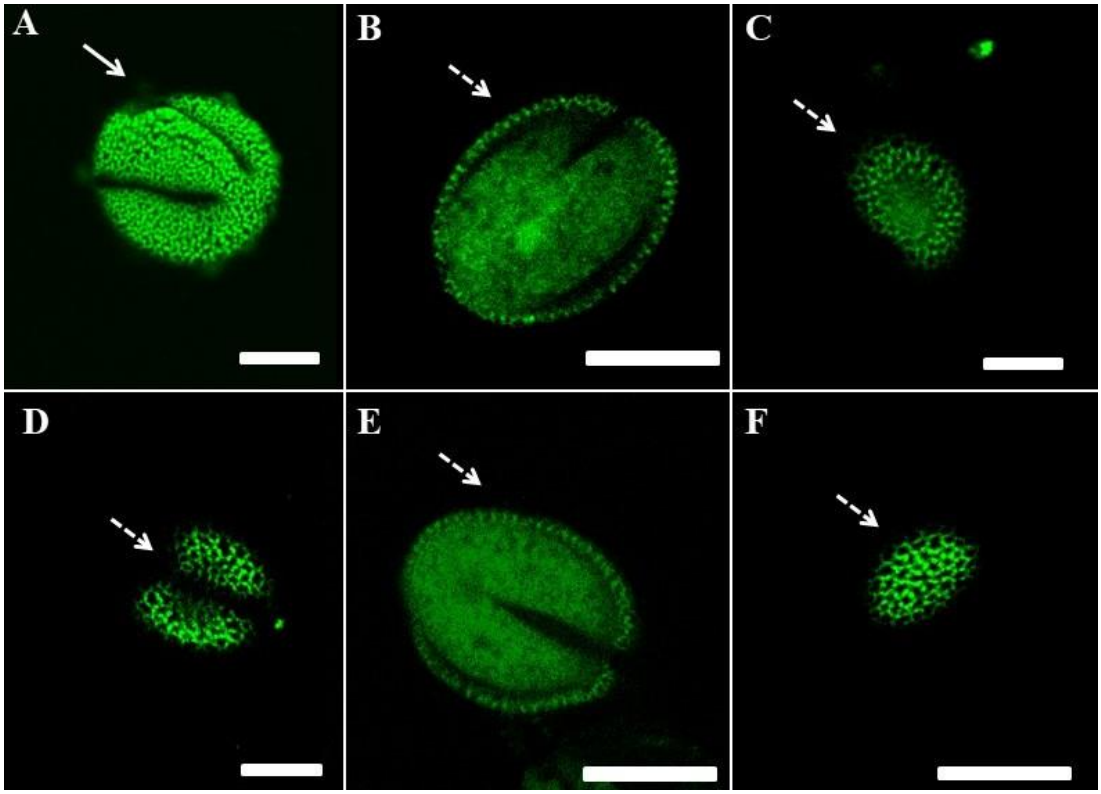


Figure 3-34. Analysis of pollen exine structure using Auramine O staining for OE7 and OE8 plants. Representative Auramine O staining from three biological replicates are shown for each sample. A: WT-Col; B and C: OE7; D-F: OE8. B-F showed abnormal structure of pollen meanwhile A represents normal structure of pollen. Solid arrow pointed at normal pollen, dashed arrow pointed at abnormal pollen. Bar: 10 μm.

Table 3-2. Summary of fertility analysis in four independent T_1 AT1G25510OE plants and WT-Col. The total number of siliques, percentage of sterility on main stem and pollen viability were determined alongside transcript level changes in four AT1G25510OE plants and WT-Col.

Sample	Total no. of siliques	Percentage of sterility on main stem (%)	Pollen Viability (%)	Transcript Level Changes
WT-Col	28	0	~100	1
OE2	18	0	~100	304x
OE6	28	7.0	~100	180x
OE7	28	46.4	~50	430x
OE8	29	41.4	~50	29x

3.2.5.3 AT4G30030 Overexpression Plants

PCR analysis of pUBQ10:AT4G30030:YFP plants identified eight transgenic plants carrying the transgene (1484 bp; Figure 3-35 and 3-36).

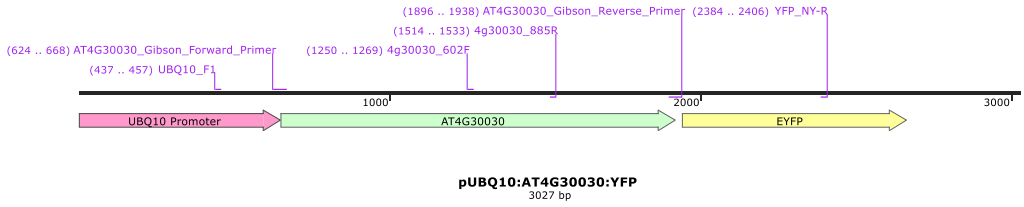


Figure 3-35. Schematic map of pUBQ10:AT4G30030:YFP overexpression vector. The pink arrow indicates the UBQ10 promoter, green arrow indicates the insertional gene and the yellow arrow indicates the YFP protein tag. The locations of the primers are shown on the map.

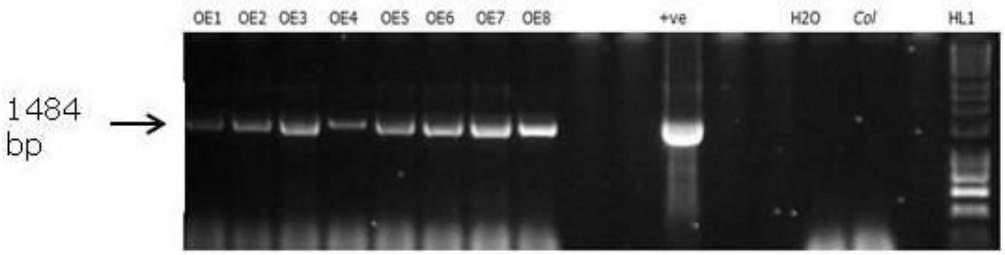


Figure 3-36. PCR to confirm presence of the transgene in eight T_1 plants. Amplification of pUBQ10:AT4G30030:YFP fragment (1484 bp) using gene specific (AT4G30030_Gibson_Reverse_Primer) and vector (UBQ10_F1) primers. OE1-OE8: Eight different AT4G30030 plants, +ve: pUBQ10: AT4G30030:YFP vector as positive control WT: WT-Col genomic DNA as negative control, H_2O: Water as a negative control, HL1: HyperLadder 1.

qRT-PCR expression analysis in four independent AT4G30030OE lines revealed increased AT4G30030 expression in all AT4G30030OE lines, from 20 to 101-fold compared to the WT-Col (Table 3-3, Figure 3-37).

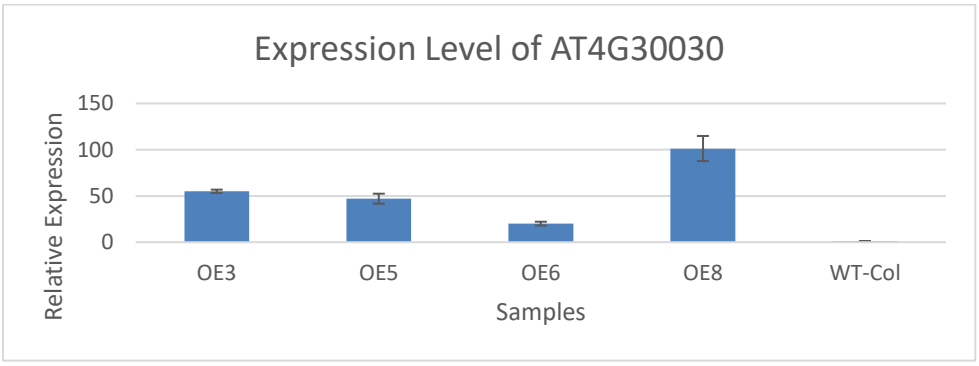


Figure 3-37. qRT-PCR expression of AT4G30030 in four independent T_1 over expression transgenic lines harbouring the pUBQ10:AT4G30030:YFP construct. Bar represents a mean value (n=3) of the AT4G30030 normalised using *PP2A3* expression. Relative expression level is compared to WT control. Error bars represent standard deviation. OE3, OE5, OE6 and OE8: Four different transgenic plants of AT1G25510OE.

Four T_1 independent transgenic lines were analysed for phenotypic changes over 6-weeks growth. Variability in fertility was seen for the different lines as compared to the WT-Col (Figure 3-38). The plants showed a partial sterility phenotype compared to the WT-Col (Figure 3-38), with percentage sterility varying from 9 %, 16.7 %, 23 %, and 7 % for OE3, OE5, OE6 and OE8 respectively, compared to the WT-Col, which exhibited 100% fully developed siliques. All transgenic plants exhibited a reduced number of siliques on the main stem (11, 12, 13, and 14 for OE3, OE5, OE6 and OE8 respectively compared to 21 siliques in WT-Col, Table 3-3). Based on the statistical analysis (See Section 2.1.22), overexpression of the AT4G30030 gene did not have any significant effect on mean of silique length compared to the WT-Col (Figure 3-39). Nevertheless, a previous investigation on the T-DNA insertion GK-024G06, did not show sterility defects although 90% reduction of gene activity was recorded.

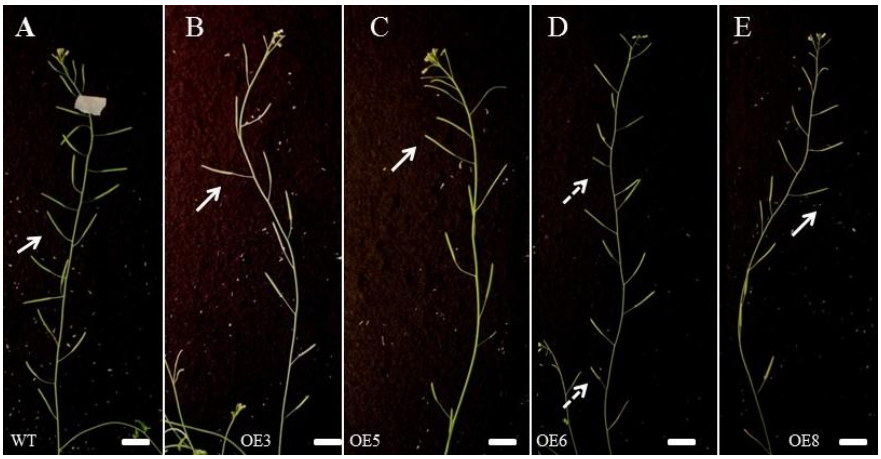


Figure 3-38. Characterization of four independent T_1 AT4G30030OE transgenic plants with different phenotype. Main stem of WT-*Col* as a control (A) and four different AT4G30030OE transgenic plants (B-E). Images were taken of the six weeks old plants. White arrow: Fertile silique. Dashed- arrow: Sterile silique. Bar = 1 cm.

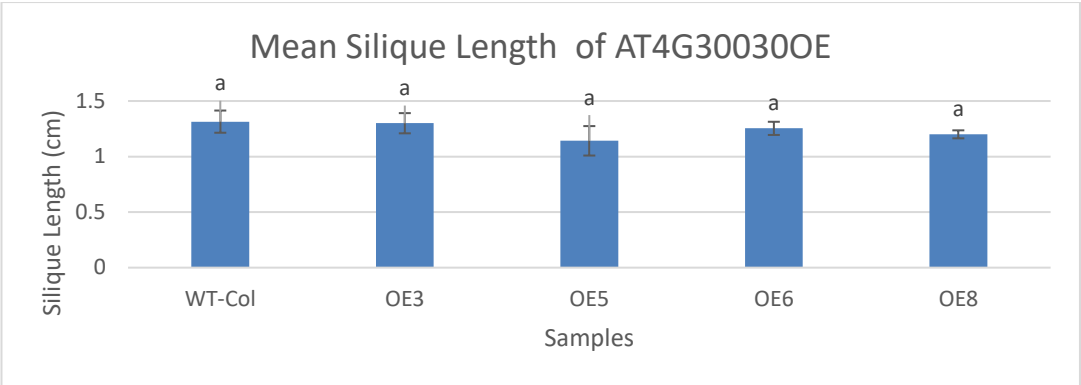


Figure 3-39. Phenotyping analysis of silique length (cm) between four different overexpression transgenic plants and the WT-Col control after 6 weeks. Mean silique length was obtained for all siliques, sterile and non-sterile by dividing total length by total number of siliques on the main stem. Identical letter signifies no significant difference where $P<0.05$; ANOVA, Fisher's Least Significant Difference test. Error bars represent standard error.

Table 3-3. Summary of fertility analysis in four T_1 AT4G30030OE plants and WT-Col. The total number of siliques, percentage of sterility on main stem were determined alongside transcript level changes in four T_1 AT4G30030OE plants and WT-Col.

Sample	**Total no. of fully developed siliques**	**Percentage of sterility on main stem (%)**	**Transcript Level Changes**
WT-Col	21	0	1
OE3	11	9	55x
OE5	12	16.7	47x
OE6	13	23	20x
OE8	14	7	101x

3.2.5.4 AT5G10760 Overexpression Plants

PCR analysis identified six transgenic plants carrying the pUBQ10:AT5G10760:YFP construct (Figures 3-40 and 3-41).

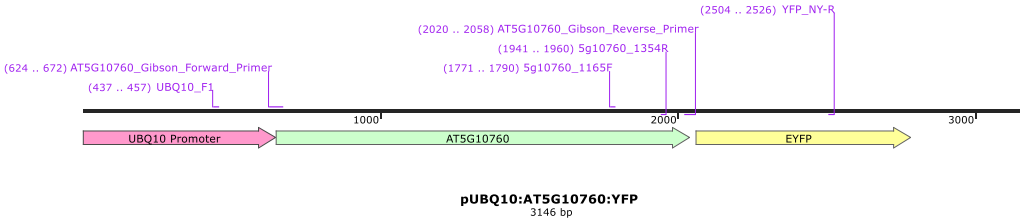


Figure 3-40. The schematic map of pUBQ10:AT5G10760:YFP overexpression vector. The pink arrow indicates the UBQ10 promoter, green arrow indicates the insertional gene and the yellow arrow indicates the YFP protein tag. The locations of the primers are shown on the map

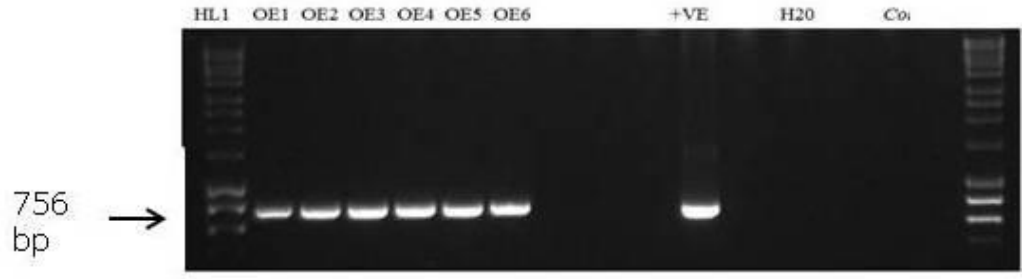


Figure 3-41. The presence of the insertion in five T_1 plants was confirmed by PCR analysis. Amplification of pUBQ10: AT5G10760:YFP (756 bp) using gene specific (5g10760_1165F) and pUBC YFP (YFP_NY-R) primers. OE1-OE6: Six different AT5G10760OE plants, Eight different AT4G30030 plants, +ve: pUBQ10:AT5G10760:YFP vector as positive control WT: WT-*Col* genomic DNA as negative control, H_2O: Water as a negative control, HL1: HyperLadder 1.

qRT-PCR expression analysis revealed increased AT5G10760 expression in all AT5G10760OE lines (Figure 3-42), varying from 150 to 360-fold, compared to the WT-Col (Table 3-4).

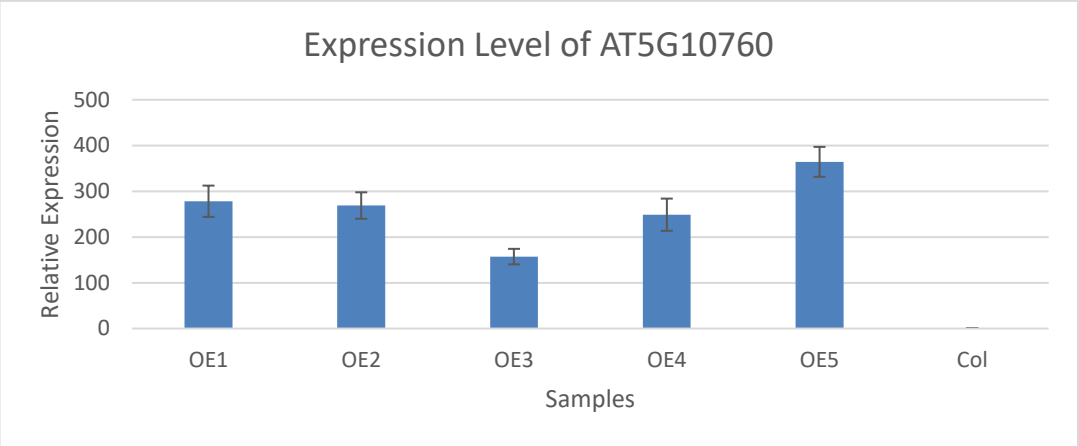


Figure 3-42. qRT-PCR expression of AT5G10760 gene in five independent T_1 over expression transgenic lines harbouring the pUBQ10: AT5G10760:YFP construct. Bar represents a mean value (n=3) of the AT5G10760 normalised using *PP2A3* expression. Relative expression level is compared to WT-*Col*. Error bars represent standard deviation. OE1, OE2, OE3, OE4, and OE5: Five different transgenic plants of AT5G10760OE.

Five T_1 independent transgenic lines were analysed for phenotypic changes over 6-weeks growth. Variability in fertility was seen for the different lines as compared to the WT-Col (Figure 3-43), full fertility was observed for OE1, OE2 and OE4, whereas 95 % and 10.7 % sterility were recorded for OE3 and OE5 respectively compared to WT-Col which exhibited 100% fully developed siliques (Table 3-4). The OE3 plant also showed decreased number of siliques on the main stem, with just 20 siliques compared to 28 siliques in WT-*Col* (Table 3-4).

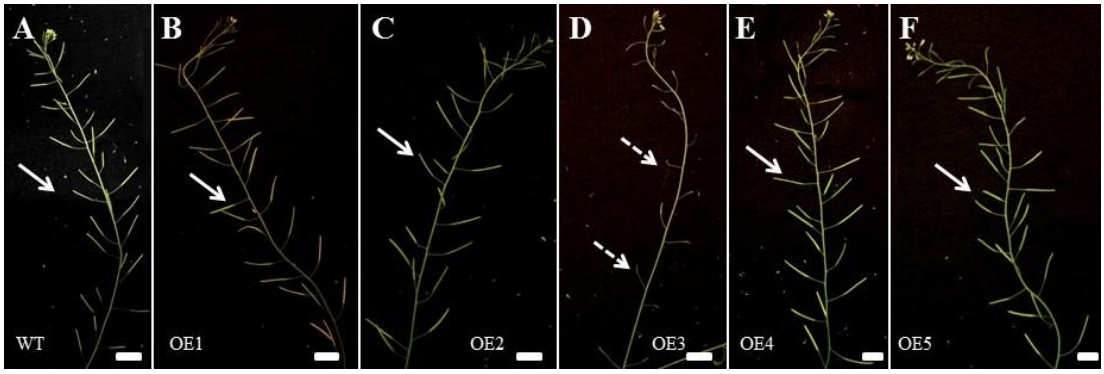


Figure 3-43. Characterization of five independent T_1 AT5G10760OE plants with different phenotype. Main stem of WT-*Col* as a control (A) and five different AT5G10760OE transgenic plants (B-F). Images were taken of the six weeks old plants. White arrow: Fertile silique. Dashed- arrow: Sterile silique. Bar = 1 cm.

Approximately 50% of pollen in OE3 plant was inviable (Figure 3-44D and 3-45), whereas the rest of the transgenic plants (OE1, OE2, OE4 and OE5) and WT-Col were viable (Figure 3-44B, C, E, F and A respectively).

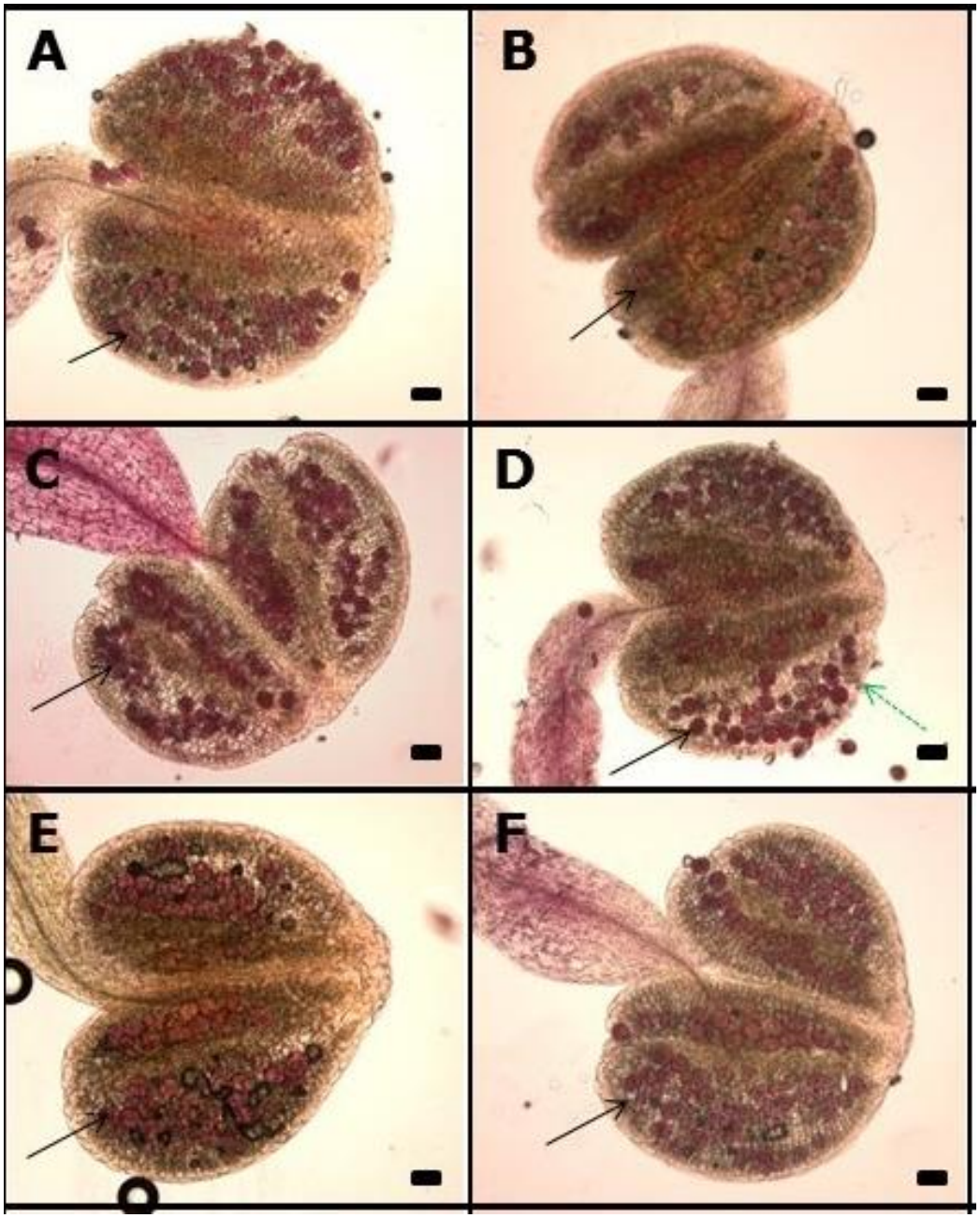


Figure 3-44. Pollen viability assessment using Alexander's staining for all five AT5G10760OE plants where one hundred pollen grains were counted from every independent plant. Red stained pollen indicates the viable pollen. Pollen from A: WT-Col, B:OE1; C:OE2; D:OE3; E:OE4; F:OE5 plant. OE3 (D) was the only line showing abnormal pollen. Black arrows indicate viable pollen grains which are stained red while non-viable pollens were pointed by the green arrow. Bar: 10 µm.

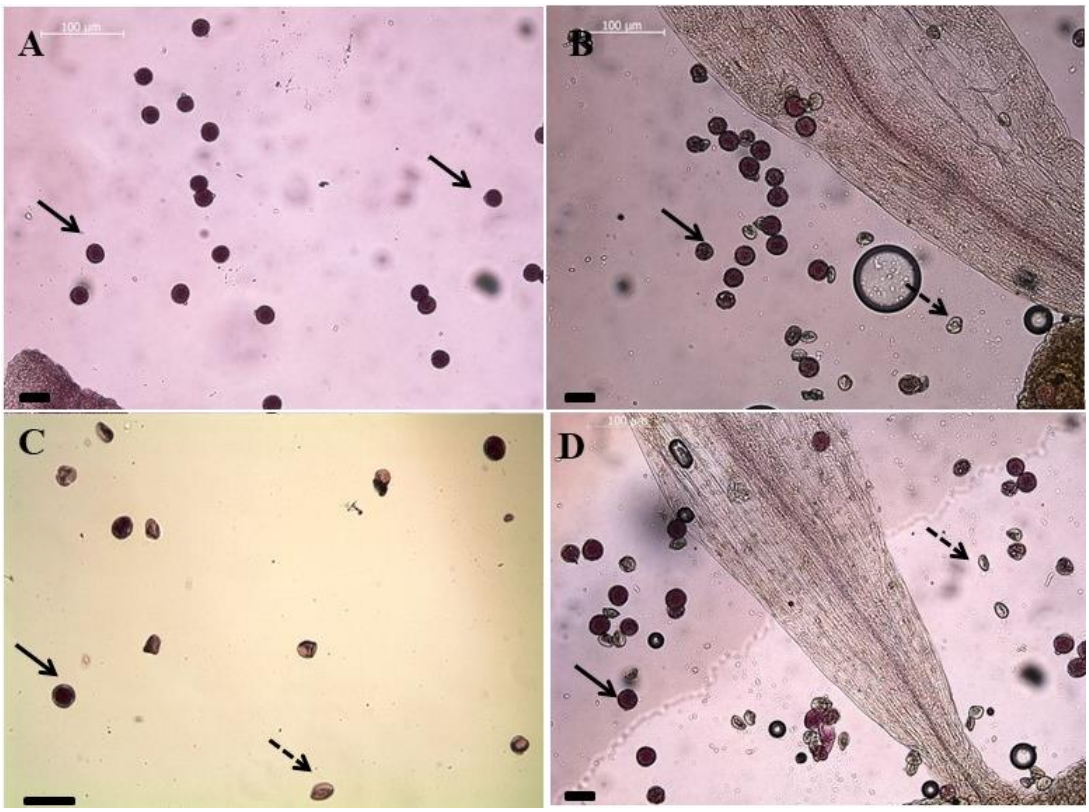


Figure 3-45. Alexander's staining of pollen from OE3 plant compared to the WT-Col (A). Red stained pollen indicates the pollen are viable. (B-D) some pollen appeared abnormal to the WT-Col pollen. Solid arrows indicate viable pollen grains, non-viable pollen are indicated by dashed arrow. Bar: 10 µm.

Based on the statistical analysis (See Section 2.1.22), OE3 was the only AT5G10760OE line that showed significant effect on mean of silique length when compare to WT-Col (Figure 3-46). The flowers, stamens, filaments and carpels of AT5G10760 OE plants were of equivalent size and growth characteristics to the WT-Col. Pollen dehiscence also occurred normally for all AT5G10760 OE lines.

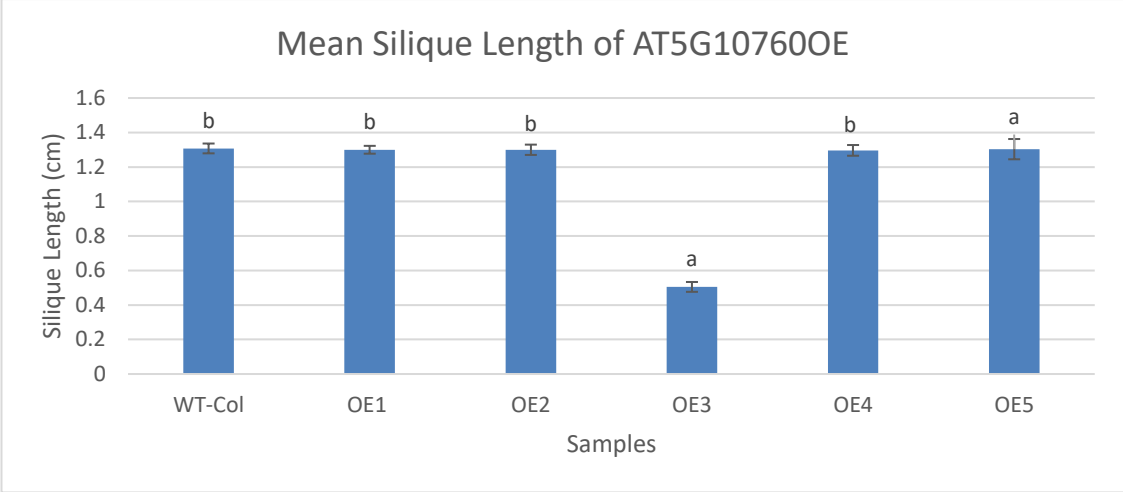


Figure 3-46. Phenotyping analysis of silique length (cm) between in five different overexpression lines and the WT-Col control after 6 weeks. Mean silique length was obtained for all siliques, sterile and non-sterile by dividing total length by total number of siliques on the main stem. Identical letter signifies no significant difference where $P<0.05$; ANOVA, Fisher's Least Significant Difference test. Error bars represent standard error.

The OE3 pollen were tested with the Auramine O staining (Kommareddi *et al.*, 1984) to check the nature of their pollen wall (exine). The abnormal pollen from this single plant had disrupted surfaces, with shrunken exine (Figure 3-47B and C); in contrast, normal WT-Col pollen was round and of regular shape (Figure 3-47A).

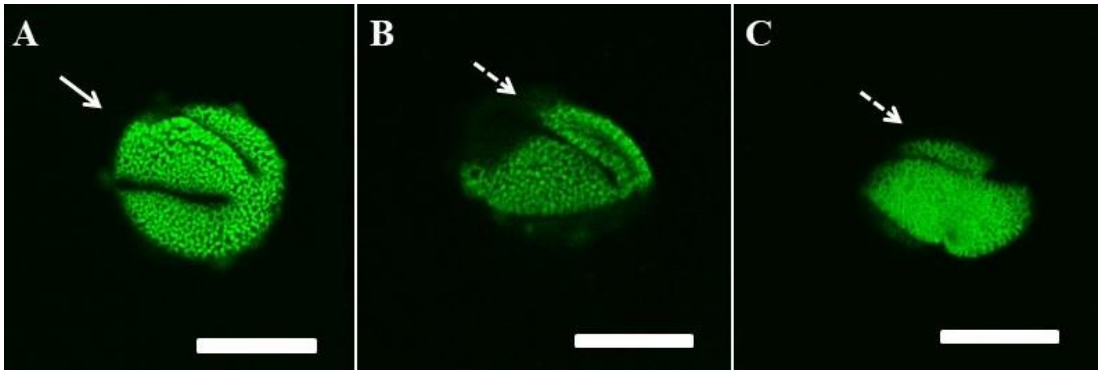


Figure 3-47. Pollen's exine layer assessment for OE3 plant using Auramine O Staining. Representative Auramine O staining from three biological replicates are shown for each sample. A: Control; B and C: OE3. B and C show abnormal structure of pollen, whereas control (A) has normal pollen. Solid arrow pointed at normal pollen, dashed arrow pointed at abnormal pollen. Bar: 10 μm.

Table 3-4. Summary of fertility analysis in five independent T_1 AT5G10760OE plants and WT-Col. The total number of siliques, percentage of sterility on main stem, pollen viability percentage were determined alongside transcript level changes in five T_1 AT5G10760OE plants and WT-Col.

Sample	**Total no. of siliques**	**Percentage of sterility on main stem (%)**	**Pollen Viability (%)**	**Transcript Level Changes**
WT-Col	28	0	~100	1
OE1	26	0	~100	278x
OE2	26	0	~100	268x
OE3	20	95	~50	157x
OE4	28	0	~100	248x
OE5	28	10.7	~100	364x

3.2.5.5 AT3G20015 Overexpression Plants

PCR analysis identified 13 transgenic plants carrying the pUBQ10:AT3G20015:YFP construct (1622bp Figure 3-48 and 3-49).

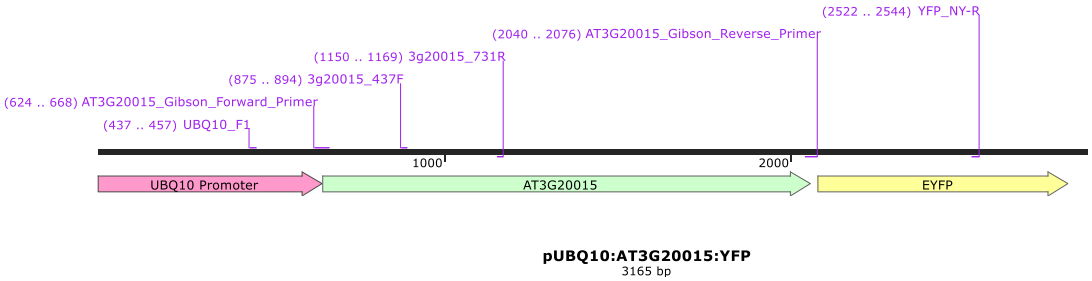


Figure 3-48. The schematic map of pUBQ10:AT3G20015:YFP overexpression vector. The pink arrow indicates the UBQ10 promoter, green arrow indicates the insertional gene and the yellow arrow indicates the YFP protein tag. The locations of the primers are shown on the map.

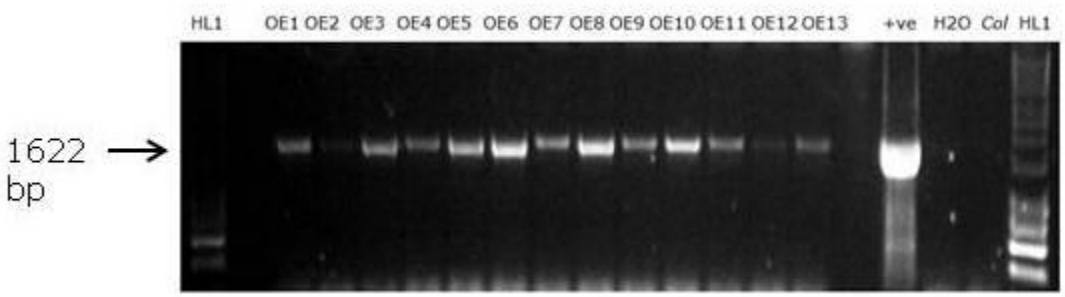


Figure 3-49. PCR confirmation of the presence of pUBQ10:AT3G20015:YFP in 13 T_1 plants. Amplification of pUBQ10:AT3G20015:YFP (1622 bp), using pUBC vector Forward (UBQ10_F1) and gene specific Reverse primer (AT3G20015_Gibson_Reverse_Primer). OE1-OE13: Different AT3G20015OE plants, +ve: pUBQ10:AT3G20015:YFP vector as positive control WT: WT Col genomic DNA as negative control, H_2O: Water as a negative control, HL1: HyperLadder 1.

qRT-PCR expression analysis indicated increased AT3G20015 expression in all AT3G20015OE lines (Figure 3-50), this increase varied from 13 to 39-fold, compared to the WT-Col (Table 3-5).

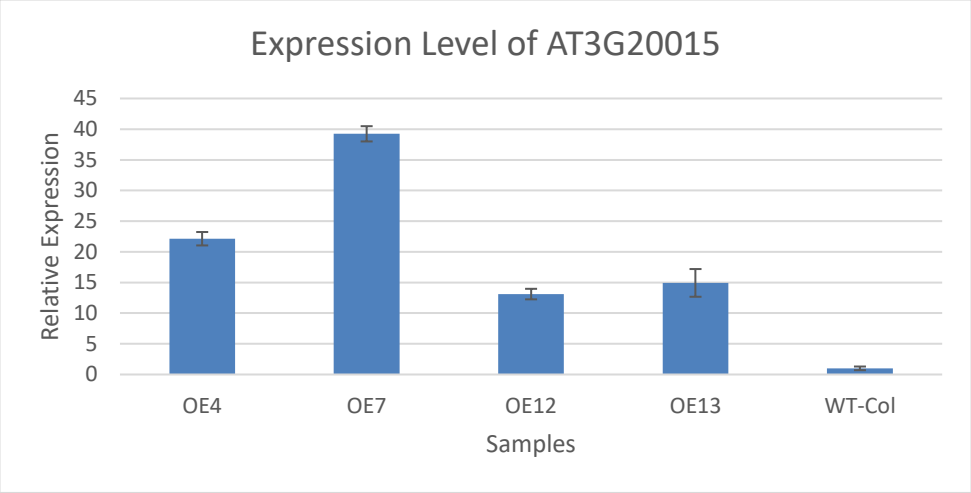


Figure 3-50. Expression pattern of AT3G20015 gene in inflorescence of five independent over expression transgenic plants harbouring pUBQ10:AT3G20015:YFP. Bar represents a mean value (n=3) of the AT3G20015 cDNA normalised using *PP2A3* expression. The relative expression level of each gene in the line was set as 1. Error bars represent standard deviation.

Four T_1 independent transgenic lines were analysed for phenotypic changes over 6-weeks growth. Variability in fertility was seen for the different lines as compared to the WT-Col (Figure 3-51), the percentage sterility varied from 10 %, 16.7 %, 18.2 % and 15.4 % for OE4, OE7, OE12, and OE13 respectively, compared to the WT-Col, which exhibited 100% fully developed siliques (Table 3-5). AT3G20015OE plants showed decreased number of siliques on main stems compared to WT-Col, except for OE4 which showed comparable fertility to WT-Col.



Figure 3-51. Images of six week old plants for four independent T_1 AT3G20015OE plants. Main stem of WT-Col as a control (A) and four different AT3G20015OE lines (B-E). White arrow: Fertile silique. Dashed- arrow: Sterile silique. Bar = 1 cm.

Based on the statistical analysis (See Section 2.1.22), overexpression of the AT3G20015 gene did not result in any significant effect on mean of silique length when compared to WT-Col (Figure 3-52).

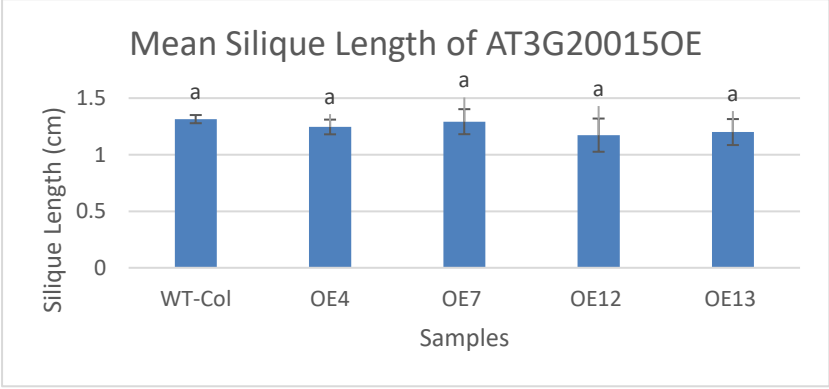


Figure 3-52. Phenotyping analysis of silique length (cm) in four independent T_1 AT3G20015OE lines and the WT-Col after 6 weeks. Mean silique length was obtained for all siliques, sterile and non-sterile by dividing total length by total number of siliques on the main stem. Identical letter signifies no significant difference where $P<0.05$; ANOVA, Fisher's Least Significant Difference test. Error bars represent standard error.

Table 3-5. Summary of fertility analysis in four independent T_1 AT3G20015OE plants and WT-Col. The total number of siliques, percentage of sterility on main stem were determined alongside transcript level changes in four T_1 AT3G20015OE plants and WT-Col.

Sample	**Total no. of siliques**	**Percentage of sterility on main stem** (%)	**Transcript Level Changes**
WT-Col	21	0	1
OE4	20	10	22x
OE7	12	16.7	39x
OE12	11	18.2	13x
OE13	13	15.4	15x

3.2.5.6 AT5G19120 Overexpression Plants

PCR analysis identified eight transgenic plants carrying the pUBQ10:AT5G19120:YFP construct (832 bp; Figure 3-53 and 3-54).

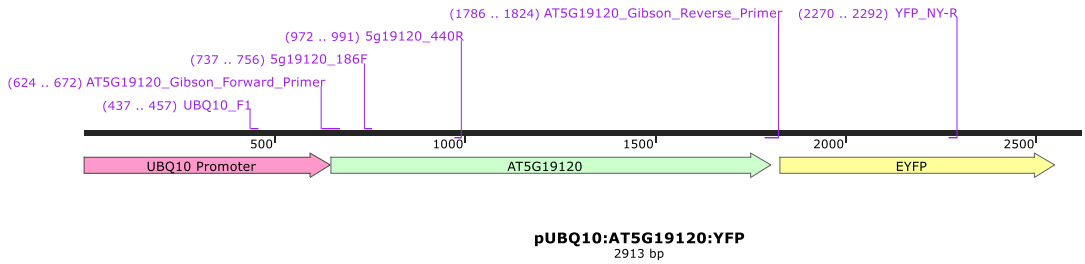


Figure 3-53. The schematic map of pUBQ10:AT5G19120:YFP overexpression vector. The pink arrow indicates the UBQ10 promoter, green arrow indicates the insertional gene and the yellow arrow indicates the YFP protein tag. The locations of the primers are shown on the map.

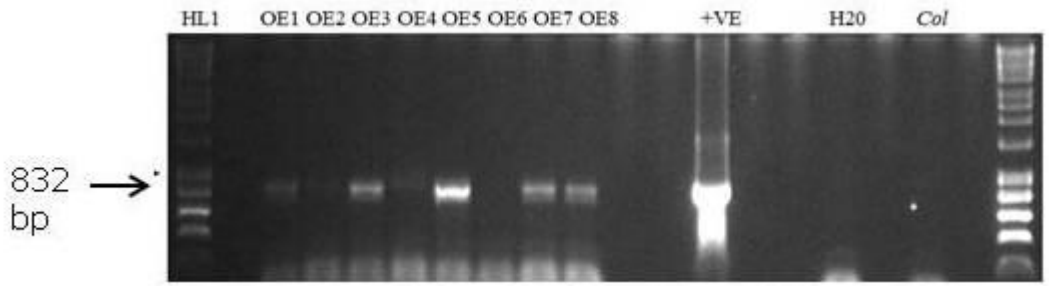


Figure 3-54. The presence of the insertion in eight T_1 plants was confirmed by PCR analysis. Amplification of pUBQ10:AT5G19120:YFP (832 bp), using gene specific Forward primer (SALK_014938 RP) and pUBC YFP Reverse primer (YFP_NY-R). OE1-OE8: eight different AT5G19120OE plants, +ve: pUBQ10:AT5G19120:YFP vector as positive control WT: WT-Col genomic DNA as negative control, H_2O: Water as a negative control, HL1: HyperLadder 1.

qRT-PCR expression analysis in three independent AT5G19120OE lines revealed increased AT5G19120 expression (Figure 3-55), which varied from increases of 25 to 44-fold, compared to the WT-Col (Table 3-6).

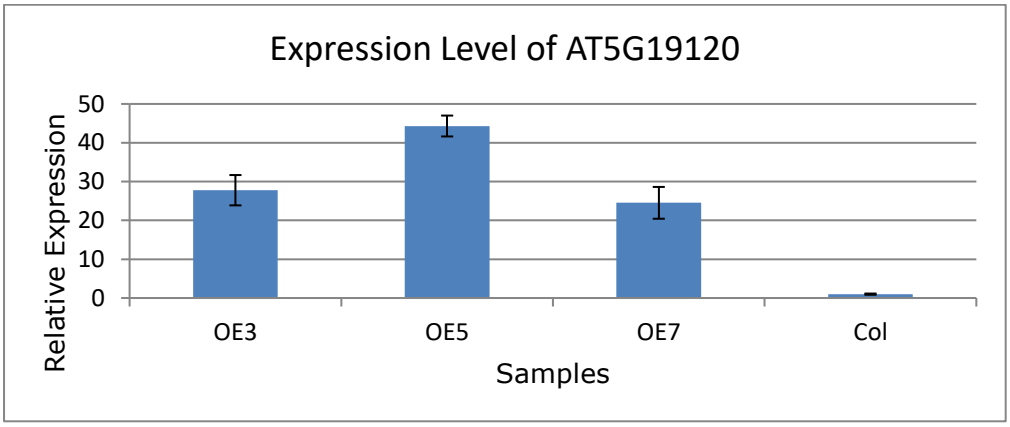


Figure 3-55. Expression pattern of AT5G19120 gene in inflorescence of three independent over expression transgenic lines harbouring pUBQ10:AT5G19120:YFP. Bar represents a mean value (n=3) of the AT5G19120 normalised using *PP2A3* expression. The relative expression level of each gene in the line was set as 1. Error bars represent standard deviation. OE3, OE5 and OE7: Three different transgenic plants of AT5G19120OE.

Three T_1 independent transgenic lines were analysed for phenotypic changes over 6-weeks growth. Variability in fertility was seen for the different lines as compared to the WT-Col (Figure 3-56). AT5G19120OE plants of OE3 and OE5, and WT-Col showed full fertility in terms of total number of siliques on main stem to WT-Col, whereas OE7 exhibited 9 % sterility (Table 3-6). All transgenic plants exhibited reduced number of siliques on the main stem, 21, 17, and 11 for OE3, OE5, and OE7 respectively compared to 28 siliques in WT-Col.

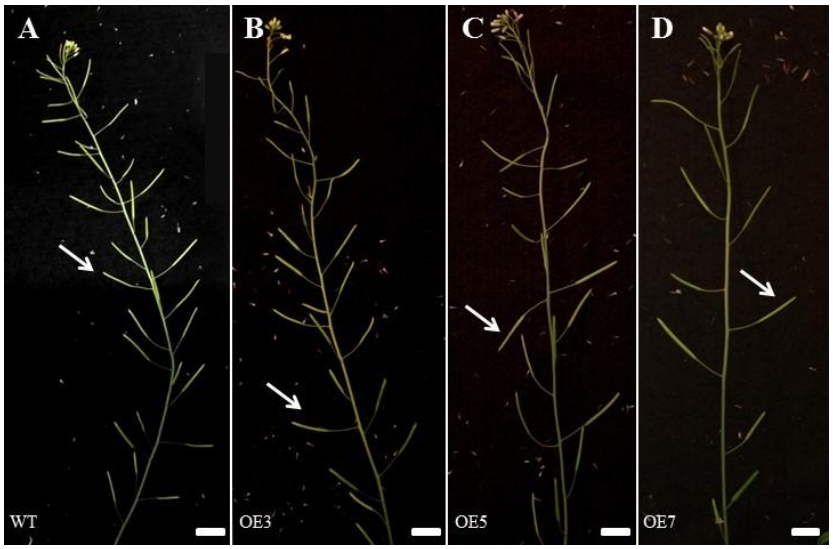


Figure 3-56. Images of three independent T_1 AT5G19120OE plants at six weeks. Main stem of WT-Col as a control (A) and three different AT5G19120OE Plants (B-D). White arrow: Fertile silique. Dashed- arrow: Sterile silique. Bar = 1 cm.

Based on the statistical analysis (See Section 2.1.22), overexpression of the AT5G19120 gene did not show any significant effect on mean of silique length when compared to WT-Col (Figure 3-57).

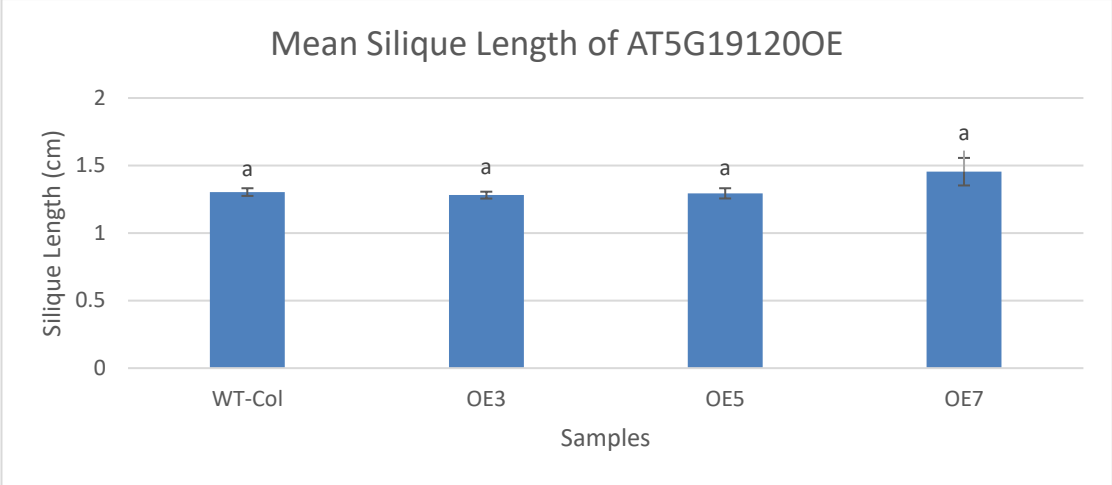


Figure 3-57. Phenotyping analysis of silique length (cm) of four independent T_1 AT5G19120OE plants and the WT-Col after 6 weeks. Mean silique length was obtained for all siliques, sterile and non-sterile by dividing total length by total number of siliques on the main stem. Identical letter signifies no significant difference where $P<0.05$; ANOVA, Fisher's Least Significant Difference test. Error bars represent standard error.

Table 3-6. Summary of fertility analysis in three independent T_1 AT5G19120OE plants and WT-Col. The total number of siliques, percentage of sterility on main stem were determined alongside transcript level changes in three T_1 AT5G19120OE plants and WT-*Col*.

Sample	**Total no. of siliques**	**Percentage of sterility on main stem**	**Transcript Level Changes**
WT-Col	28	0	1
OE3	21	0	28x
OE5	17	0	44x
OE7	11	9	25x

3.2.5.7 AT2G21430 Overexpression Plants

PCR analysis identified four transgenic plants carrying the pUBQ10: AT2G21430:YFP insertion construct (1295 bp; Figure 3-58 and 3-59).

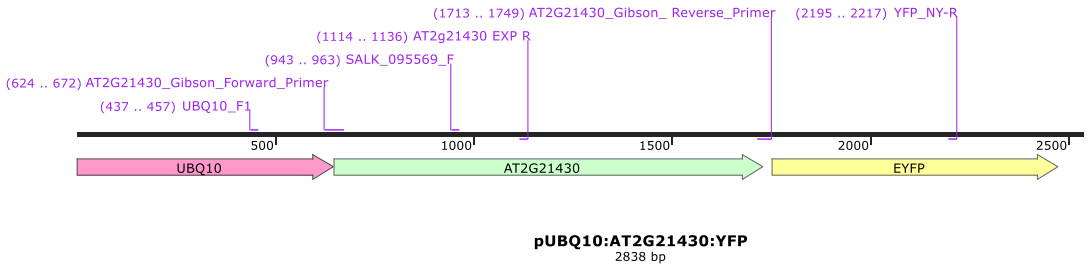


Figure 3-58. Schematic map of pUBQ10:AT2G21430:YFP overexpression vector. The pink arrow indicates the UBQ10 promoter, green arrow indicates the insertional gene and the yellow arrow indicates the YFP protein tag. The locations of the primers are shown on the map.

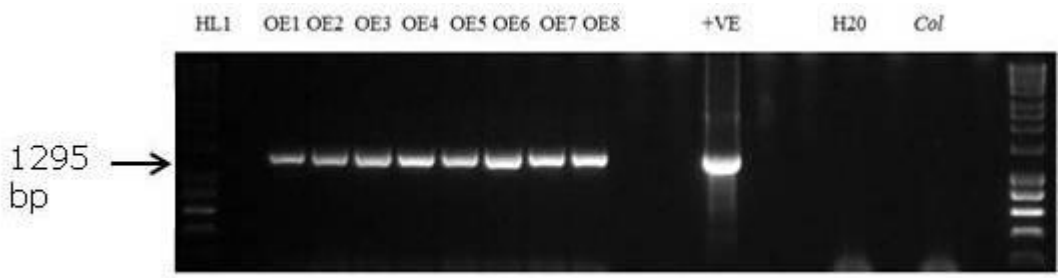


Figure 3-59. The presence of the insertion in three T1 plants was confirmed by PCR analysis. Amplification of pUBQ10:AT2G21430:YFP (~1295 bp), using pUBC YFP Forward primer (UBQ10_F1) and gene specific Reverse primer (AT2G21430_Gibson_Reverse_Primer), indicating integration of the gene into the plant chromosome. OE1-OE8: eight different AT2G21430OE plants, +ve: pUBQ10:AT2G21430:YFP vector as positive control, WT: WT-Col genomic DNA as negative control, H_2O: Water as a negative control, HL1: HyperLadder 1.

qRT-PCR expression analysis revealed increased AT2G21430 expression in all AT2G21430OE lines (Figure 3-60), which varied from 7 to 39-fold, compared to the WT-Col (Table 3-7).

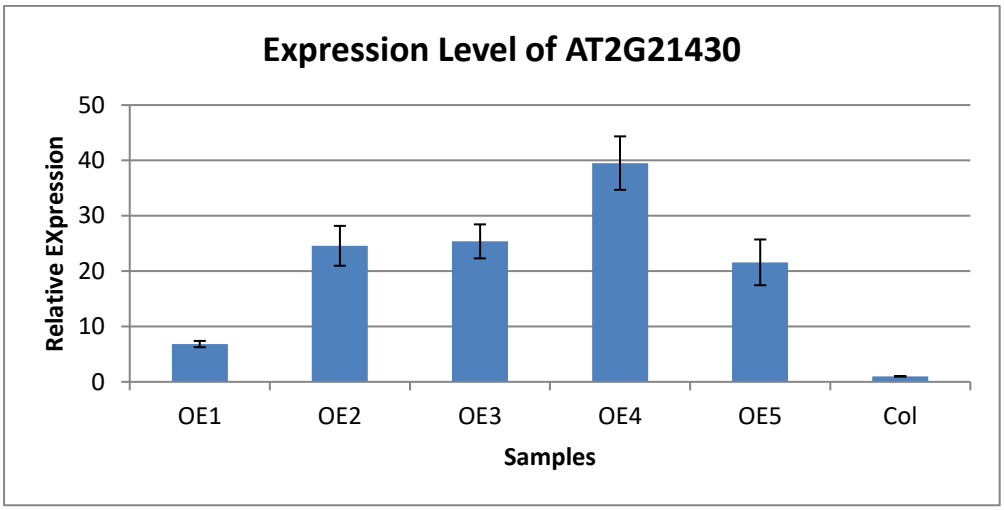


Figure 3-60. Expression pattern of AT2G21430 gene in inflorescence of five independent T_1 over expression transgenic lines harbouring pUBQ10:AT2G21430:YFP. Bar represents a mean value (n=3) of AT1G25510 normalised using *PP2A3* expression. The relative expression level of each gene in the line was set as 1. Error bars represent standard deviation. OE1, OE2, OE3, OE4 and OE5: Five different transgenic plants of AT2G21430OE.

Five T_1 independent transgenic lines were analysed for phenotypic changes over 6-weeks growth. Variability in fertility was seen for the different lines as compared to the WT-Col (Figure 3-61), with ~100% fertility obtained for three AT2G21430OE plants, OE2, OE4, OE5, and WT-Col. However, slight reductions of sterility were observed for OE1 and OE3, which exhibited 9% and 7.7% sterility, respectively. All transgenic plants showed slight decreased number of siliques on main stem, 22 siliques for OE1, OE4 and OE5 whilst 26 siliques were recorded for OE2 and OE3 compared to 28 siliques in WT-Col (Table 3-7).

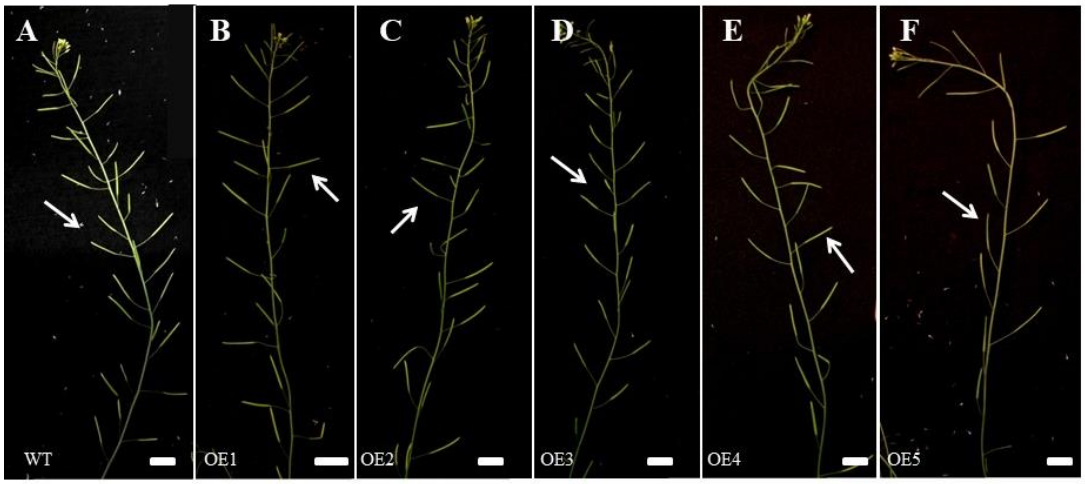


Figure 3-61. Characterization of five independent T_1 AT2G21430OE plants. Main stem of WT-Col as a control (A) and five different AT2G21430OE plants (B-F). Images were taken of the six weeks old plants. White arrow: Fertile silique. Dashed- arrow: Sterile silique. Bar = 1 cm.

Based on the statistical analysis (See Section 2.1.22), overexpression of the AT2G21430 gene in OE2, OE4 and OE5 showed a significant effect on mean of silique length when compare to WT-Col, whilst OE1 and OE3 did not differ significantly (Figure 3-62).

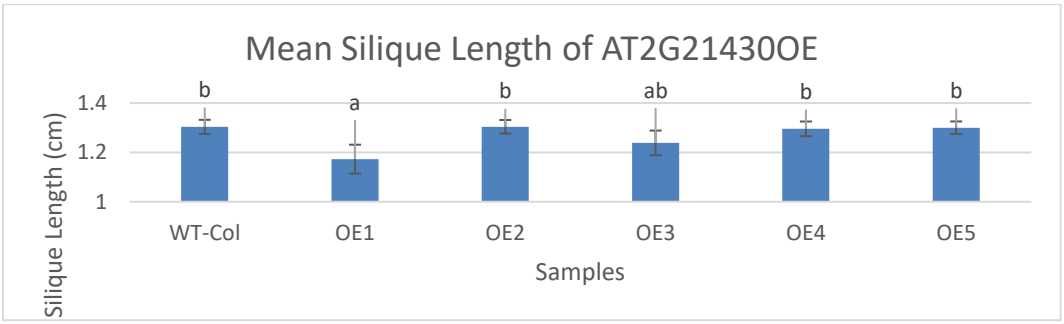


Figure 3-62. Phenotyping analysis of silique length (cm) between four independent AT2G21430OE plants and the WT-Col after 6 weeks. Mean silique length was obtained for all siliques, sterile and non-sterile by dividing total length by total number of siliques on the main stem. Identical letter signifies no significant difference where $P<0.05$; ANOVA, Fisher's Least Significant Difference test. Error bars represent standard error.

Table 3-7. Summary of fertility analysis in five T_1 AT2G21430OE plants and WT-Col. The total number of fully developed siliques, percentage of sterility on main stem were determined alongside transcript level changes in three T1 AT2G21430OE plants and WT-Col.

Sample	**Total no. of siliques**	**Percentage of sterility on main stem**	**Transcript Level Changes**
WT-Col	28	0	1
OE1	22	9	6x
OE2	26	0	25x
OE3	26	7.7	25x
OE4	22	0	45x
OE5	22	0	22x

3.2.5.8 AT3G16210 Overexpression Plants

Analysis of the AT1G25510 gene was conducted by overexpression using a pGWB5:AT1G25510:GFP construct (Figure 3-63). Amplification of AT3G16210 generated a ~1.2 kbp product (Figure 3-64); pCR8/GW/TOPO clone 4 (C4) was confirmed by sequencing and then used for further analysis.

Amplification of AT3G16210 gene construct generated a product of ~1.2 kbp (Figure 3-65). pCR8/GW/TOPO clone 4 (C4) was used for further analysis after sequencing confirmation.

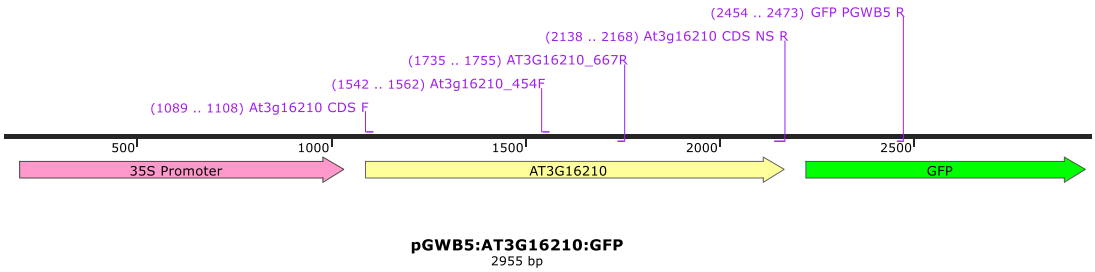


Figure 3-63. Schematic map of pGWB5:AT3G16210:GFP overexpression vector. The pink arrow indicates the 35S promoter, green arrow indicates the gene insertion, green arrow indicates the GFP protein tag. The locations of the primers are shown on the map.

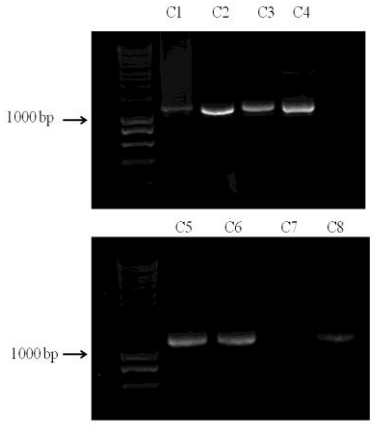


Figure 3-64: Amplification of the pCR8/GW/TOPO entry clone with AT3G16210 gene. Gene amplification of ~1.2 kbp band from 8 colonies PCR using AT3G16210 CDS forward primer and pCR8/GW/TOPO 2.1 reverse primer. C1-C8= Different colonies of pCR8/GW/TOPO:AT3G16210 vector; Ladder: HyperLadder 1

The pGWB5:AT3G16210:GFP construct was amplified (~1.4 kbp size of band (Figure 3-57) and clone 4 (C4) confirmed by sequencing. The pGWB5:AT3G16210:GFP construct was transformed into *Agrobacterium* using electroporation, verified by PCR, and transformed into WT-Col.

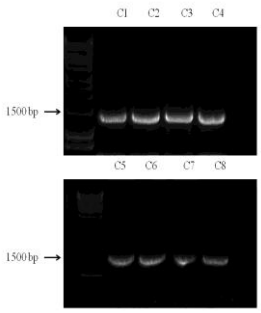


Figure 3-65: Screening for AT3G16210 destination vector. Each amplified PCR product was cloned into pGWB5 vector by a LR reaction. Gene amplification using AT3G16210 forward specific primer and PGWB5 reverse primer with the expected size of band was around ~1.4 kbp. C4 was chosen to further analysis. Note: C1-C8= Different colonies of pGWB5:AT3G16210:GFP vector; Ladder: HyperLadder 1.

The putatively transformed seeds were screened on MS medium containing Kanamycin (50 ng/µl). PCR analysis confirmed that all 8 transgenic plants carried the AT3G16210 gene (~1.4 kbp, Figure 3-66).

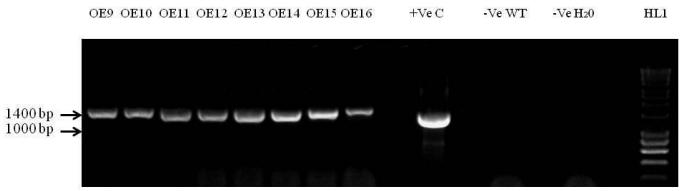


Figure 3-66: PCR amplification of pGWB5:AT3G16210:GFP (~1.4kbp band) in eight T_1 plants. Amplification using gene specific Forward primer and pGWB5: GFP Reverse primer. OE9-OE16: different transgenic plants, +Ve C: pGWB5:AT3G16210:GFP vector as positive control WT, -Ve WT: WT-Col genomic DNA as negative control, -Ve H_2O: Water as a negative control, HL1: HyperLadder 1.

qRT-PCR expression analysis in five independent AT3G16210OE lines showed increased AT3G16210 expression in all lines (Figure 3-67), which varied from 203 to 271-fold, compared to the WT-Col (Table 3-8).

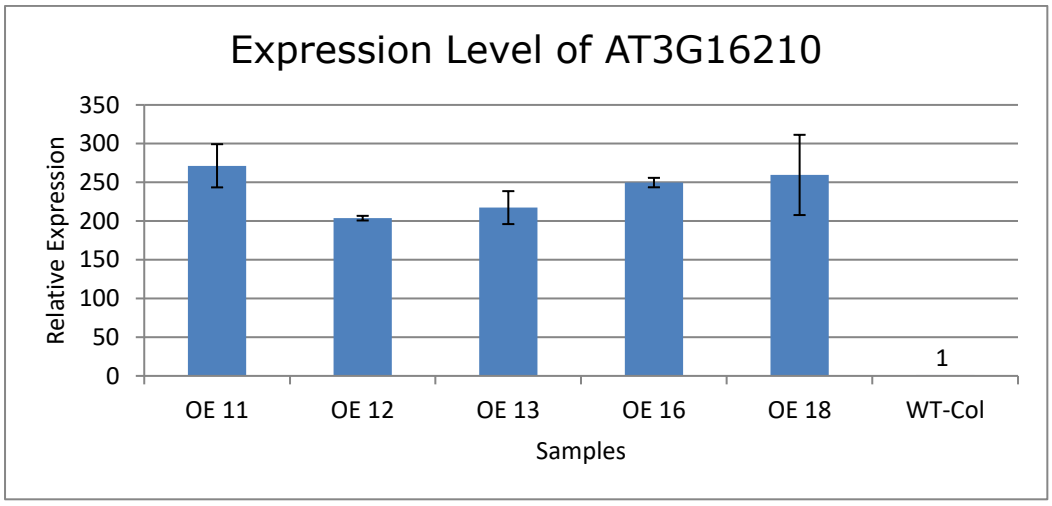


Figure 3-67. Expression pattern of AT3G16210 gene in inflorescence of five independent over expression transgenic plants harbouring pUBQ10:AT3G16210:YFP. Bar represents a mean value (n=3) of the AT3G16210 normalised using *PP2A3* expression. The relative expression level of each gene in the line was set as 1. Error bars represent standard deviation.

Four T_1 independent transgenic lines were analysed for phenotypic changes over 6-weeks growth. Variability in fertility was seen for the different lines as compared to the WT-Col (Figure 3-68), which varied from 10 %, 0 %, 0 %, 11 % and 0%, for OE11, OE12, OE13, OE16 and OE18 respectively. However, all plants showed reduced fertility except for OE2. All transgenic plants exhibited reduced number of siliques on the main stem, (20, 20, 22, 19 and 23 for OE11, OE12, OE13, OE16 and OE18 respectively compared to 28 siliques in WT-Col; Table 3-8).



Figure 3-68. Images of six weeks old plants of five independent T_1 AT3G16210OE transgenic lines. Main stem of WT-Col control (A) and five different AT3G16210OE plants (B-E). White arrow: Fertile silique. Dashed- arrow: Sterile silique. Bar = 1 cm.

Based on the statistical analysis (Section 2.1.22), overexpression of the AT3G16210 gene in OE12 and OE16 showed a significant effect on mean silique length compared to WT-Col, whilst OE11, OE13 and OE18 did not differ significantly (Figure 3-69).

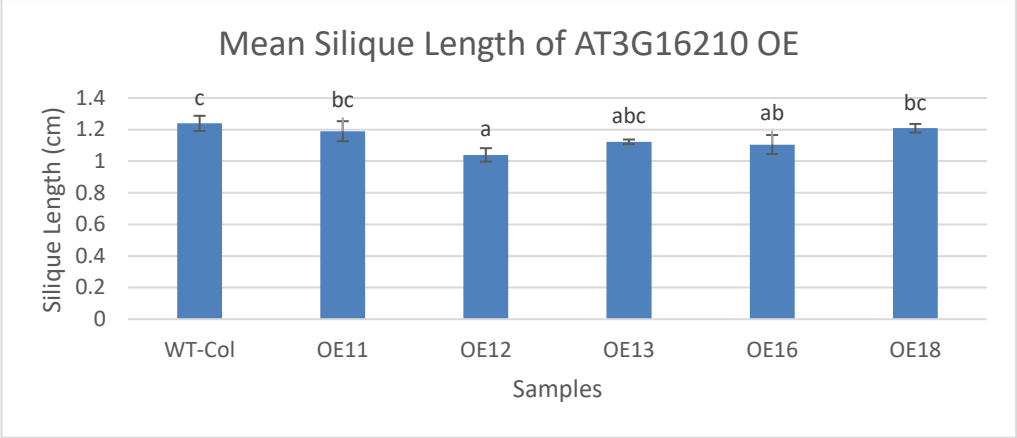


Figure 3-69. Phenotypic analysis of silique length (cm) in five independent AT3G16210OE lines and WT-Col after 6 weeks. Mean silique length was obtained for all siliques, sterile and non-sterile by dividing total length by total number of siliques on the main stem. Identical letter signifies no significant difference where $P<0.05$; ANOVA, Fisher's Least Significant Difference test. Error bars represent standard error.

Table 3-8. Summary of fertility analysis in five independent T_1 AT3G16210OE plants and WT-Col. The total number of fully developed siliques, percentage of sterility on main stem were determined alongside transcript level changes in five T_1 AT3G16210OE plants and WT-*Col*.

Sample	**Total no. of siliques**	**Percentage of sterility on main stem**	**Transcript Level Changes**
WT-Col	28	10	1
OE11	20	10	271x
OE12	20	0	203x
OE13	22	0	217x
OE16	19	11	249x
OE18	23	0	259x

3.3 Discussion

Overexpression was chosen as a reverse genetics strategy to determine the impact of ectopic increased expression of the protease genes. The results were varied, with different severity of sterility phenotypes recorded between constructs and in independent overexpression lines.

No strong sterility phenotypic changes were observed except for AT2G23945 and AT4G30030 transgenic plants. For both lines, reduction of total number of siliques have been observed in all overexpression mutant lines. In term of vegetative growth, all transgenic plants were comparable to the parental WT-Col. Flowers, stamens, filaments and carpels of all transgenic plants were of equivalent size and growth characteristics to the WT-Col; pollen dehiscence occurred normally for all transgenic plants. Increased target gene expression was detected in the positive transgenic lines for all genes tested, with very high increases recorded for AT3G16210 (203 to 271-fold increase). For most of the overexpression plants generated, the expression level did not show a direct correlation with the level of sterility with the majority of the transgenic plants displaying a weak sterility phenotype (Table 3-9).

AT2G23945 overexpression plants driven by pUBC vector showed partial sterility, ranging from 12.5% to 22.2%. Furthermore, qRT-PCR revealed a very high level of expression, between 43 to 125-fold (Table 3-1). The fertility level of all overexpression mutants was

reduced compared to the WT-Col even though the level of sterility and expression pattern were variable.

AT4G30030 Aspartyl-Protease overexpression line also exhibited significant changes in both silique number and fertility Table 3-9. AT4G30030OE plants showed a decreased total number of siliques on the main stem (11, 12, 13, and 14 for OE3, OE5, OE6 and OE8 respectively compared to 21 siliques in WT-Col). Additionally, the sterility percentage rates varied for the four AT4G30030OE plants, 9 %, 16.7 %, 23 %, and 7 % for OE3, OE5, OE6 and OE8 respectively, whereas, WT-Col which exhibited full fertility (Table 3-3).

Considerable phenotypic variation has been previously reported among transgenic plants even though they have been generated from the same construct and maintained under identical growing condition (Butaye and Cammue, 2005). Two out of four AT1G25510OE plants and a single plant out of five from AT5G10760OE plants showed a strong sterility phenotype. The main stem sterility for AT1G25510OE plants ranging from 41.4 % to 46.4 % with half of the pollen exhibited abnormalities. On the other hand, AT5G10760OE plants showed 95 % sterility with just 50 % pollen viability recorded. This may be due to the transgenes inserting into different region in the genome. Pronounced variability among the transformants sterility phenotypes may cause conflicting results (see Section 6.1).

Table 3-9. Summary of fertility analysis in six different protease and one F-Box protein overexpression T_1 lines. Effects on number of siliques on main stem, effects on sterility on main stem were determined alongside transcript level changes in seven different T_1 OE plants.

Gene	**Transcript Level Changes**	**Effects on number of Siliques on main stem**	**Effects on sterility on main stem**
AT2G23945	43 to 125-fold	Reduced number for all six OE plants	Partial sterility for all six OE plants
AT1G25510	10 to 420-fold	Weak effects	Weak sterility phenotype
AT4G30030	20 to 101-fold	Reduced number for all four OE plants	Partial sterility for all four OE plants
AT5G10760	150 to 360-fold	Weak effects	Weak sterility phenotype
AT3G20015	13 to 39-fold	Weak effects	Partial sterility for all four OE plants
AT5G19120	25 to 44-fold	Weak effects	Weak sterility phenotype
AT2G21430	7 to 39-fold	Weak effects	Weak sterility phenotype
AT3G16210	203 to 271-fold	Reduced number for all five OE plants	Weak sterility phenotype

Interestingly, the Gene Ontology Consortium (http://geneontology.org/) suggested that aspartyl-protease AT2G23945 is closely related to AT4G30030 and AT4G30040 (See Section 4.1), which exhibited fertility effects upon gene manipulation trough SALK T-DNA insertion and RNAi techniques (See Section 4.2). Even though reverse genetic study has become well established, it still facing several limitations, particularly a lack of visible phenotypic changes (Prelich, 2012). Schubert *et al*., 2004 attributed this phenomenon to the RNA sensing mechanism, a defense mechanism at the cellular level by plant, which occurred once the expression level exceeds certain threshold level. It is part of the adaptation by the cell for certain level of expression, meant to eliminate any harmful effects by the excessive transcriptional activity.

The overexpression plants that do not produce any phenotypic changes, masks the actual effects of the gene and make it difficult to analyse its biological function. This outcome can be explained in several ways. High expression level can result in the disruption of potential protein complexes, which compete for the functional protein complex and other factors, which are necessary for normal gene functioning. It may also trigger unrelated responses, that cause disruption to the normal function and reaction balance, and change the protein structure required for normal pathway operation (Prelich, 2012).

According to Schubert *et al.*, (2004), each gene has its own specific minimum and maximum threshold levels that can cause phenotypic changes. Balance of expression is activated by specific cellular RNA mechanisms including degradation of another pathway, after

detection of abnormal transcription activity from the introduced gene. In fact, the plants need to maintain its normal function and severity of the phenotypes can be avoided (Lloyd and Meinke, 2012).

Complex gene expression regulation is disrupted by the presence of the endogenous and introduced genes that can contribute to the phenotypic variation. To overcome the disruption, cells can activate feedback loops to maintain the equilibrium, by activating degeneration in certain pathway, or by co-suppression of the certain genes until it reached homeostatic stage of gene interactions (Singh, 1998). As part of the stress response, genes are speculated to be activated to maintain the normal development of transgenic plants (Singh, 1998).

The YFP tag was used to help enable visualization of the transgene and was used as a reporter of expression. Similar localization of all six proteases and the control construct were observed in the membrane and the nucleus, thus it was not possible to determine localization of expression. Interestingly, for four proteases genes (AT1G25510, AT4G30030, AT5G19120 and AT2G21430) localization occurred in stomatal guard cells, whereas this was not seen for the wild type controls. It may give suggest that these proteases may undergo hydrolysis activity in the stomatal region.

Loss of membrane integrity is part of the apoptotic-like feature in animal system. Hence, tissue lysis and Trypan Blue staining analysis were conducted as part of the effort to link the proteases activity with the PCD and overcome any issues of potential redundancy

(Elmore, 2007). Loss of membrane integrity leads to the cell death in epidermal cells as shown by the absorption of the trypan blue staining (Grosse-Holz *et al.*, 2017). However, no immediate cell death phenotypes were observed. Cell death was observed from the control and all proteases genes constructs after 10 days, which may be due to the excessive lysis activity by the proteases after 6-7 days post-inoculation, or from the inoculation process persee. This suggests that these proteases are not causing cell death, however they may be involved in protection from PCD and cellular breakdown, since a slight reduced level of trypsin blues staining was seen in some lines after 10 days compared to the WT-Col. However, it is also important to note that the T-DNA integrated into *N. benthamiana* cells can take several days for the protein to be expressed and functional protein expression may not have been achieved and therefore potential activity and PCD may not have been shown (Grosse-Holz *et al.*, 2017).

CHAPTER 4: GENE MANIPULATION ANALYSIS OF AT2G23945

4.1 Introduction

Over-expression of AT2G23945 has shown promising results as discussed in Section 3.2.5.1. This chapter focuses on further functional characterization of the Aspartyl-protease AT2G23945 by employing several reverse genetic techniques; SALK T-DNA insertion, and RNAi downregulation.

General gene information was obtained through The Arabidopsis Information Resource's (TAIR) website: (http://www.arabidopsis.org/). This gene is located on chromosome 2, and is organized into two exons, separated by an intron, and 3' untranslated region (Figure 4-8, Section 4.2.2) with molecular weight of 51352 Da. Further information about the gene was accessed by MEROPS peptidase public database (http://merops.sanger.ac.uk/). The protein (A01.A39) protein ID has conserved aspartic residues of Aspartic and Thyronine amino acid in the sequence (Figure 4-1), details of aspartic proteases has been discussed in section 1.6.1.

The *Arabidopsis* AT2G23945 gene was examined to find related homologs and orthologs by using public bioinformatics tools, particularly Basic Local Alignment Search Tool (BLAST®) (https://blast.ncbi.nlm.nih.gov/smartblast/smartBlast.cgi). A putative aspartic protease, AT2G35615, with 39% amino acid similarity to AT2G23945 was identified, however there is no

information on this gene in Flowernet database (http://www.cpib.ac.uk/anther) indicating that it may not be involved in pollen development. However the Gene Ontology Consortium (http://geneontology.org/) suggests that the aspartyl-protease genes AT4G30030 and AT4G30040 are closely related to AT2G23945 gene but that it is not related to AT2G35615. In addition, Gramene open-source website (http://gramene.org/) suggests that these three genes AT4G30030, AT4G30040 and AT2G35615 are close homologs of the AT2G23945 gene. Aspartic proteinase CDR1-like and putative aspartic protease AT2G35615 in *Glycine max* are the ortholog candidates as examined by BLAST® analysis.

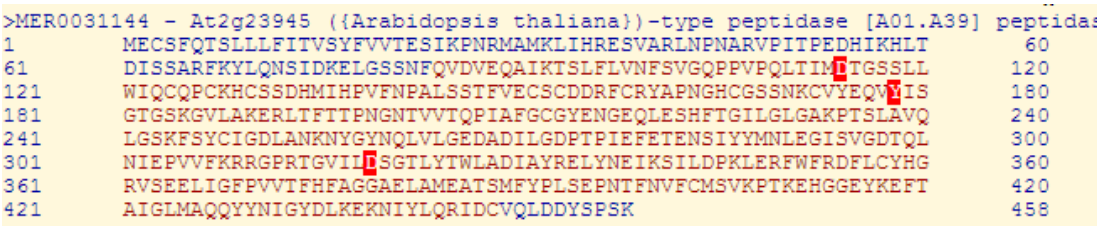


Figure 4-1: Amino acid sequence of AT2G23945 as provided by MEROPS database. The conserved Aspartic (D) residues located at amino acid no. 114 and 319 while the conserved Thyrosine (Y) located at the amino acid no. 178, conserved residues indicated in red-highlighted letters. The peptidase unit started at the amino acid 84 to 448.

For reverse genetics analysis, transfer DNA (T-DNA) insertion provides a method to study the effects of disruption of the gene of interest in *Arabidopsis*. It remains one of the most powerful tools for reverse genetics study in *Arabidopsis* for functional genomic studies (Gao and Zhao, 2013). The insertion occurs randomly and often produces null alleles or knockout mutants and homozygous knockout mutant characterized by gain-of-function or loss-of

function mutations for genes tested (Krysan *et al.,* 1999; Kleinboelting *et al.,* 2016).

T-DNA insertion information is available from the The Salk Institute Genome Analysis Laboratory (SIGnAL) website (http://signal.salk.edu/), which means that mutants can be identified *in silico* for functional analysis from the Nottingham Arabidopsis Stock Centre (NASC) and Arabidopsis Biological Resource Center (ABRC) (Kleinboelting *et al., 2016)*.

The vectors used to generate SALK lines carry distinct tag sequence which acts as a marker for mutant isolation. By conducting PCR with specific primer from the insertion sequence and targeted transcript, the foreign DNA insertions are easily identified. In addition, the manipulation of vectors by adding new elements such as a constitutive promoter, CaMV 35S, and reporter genes like GUS and GFP enables localisation of expression (Feldmann, 1991; Azpiroz-Leehan and Feldmann, 1997; Sanchez-serrano and Walker, 2014).

RNA interference was first reported and characterized in *C .elegans* by Fire *et al.*, (1998). This post-transcriptional interference provides an alternative for reverse genetic screening with specific and inheritable dsRNA. It is used to suppress expression by introduction of experimentally long double-stranded RNAs (dsRNAs), driven by a constitutive promoter. An enzyme called dicer acts as an RNAse to cleave dsRNAs into 21-23 bp small interfering RNAs (siRNAs). siRNAs consist of the passenger strand and the guide strand. RNA-Induced Silencing Complex (RISC) is an association of different

proteins that comprise of Dicer, TRBP, and Argonaute 2 (Ago2). Then, the guide strand enters the RISC while the passenger strands degraded. The guide strand, as the name implies, guide the Ago2 RISC to cleave the target endogenous mRNA that have sequence similarity (Qiu *et al*., 2005; Peek and Behlke, 2007). The level of specific endogenous mRNA reduced, targeted by the excessive dsRNA in the cells (Chuang and Meyerowitz, 2000). The biochemical pathway in RNA interference is presented in Figure 4-5 (Peek and Behlke, 2007).

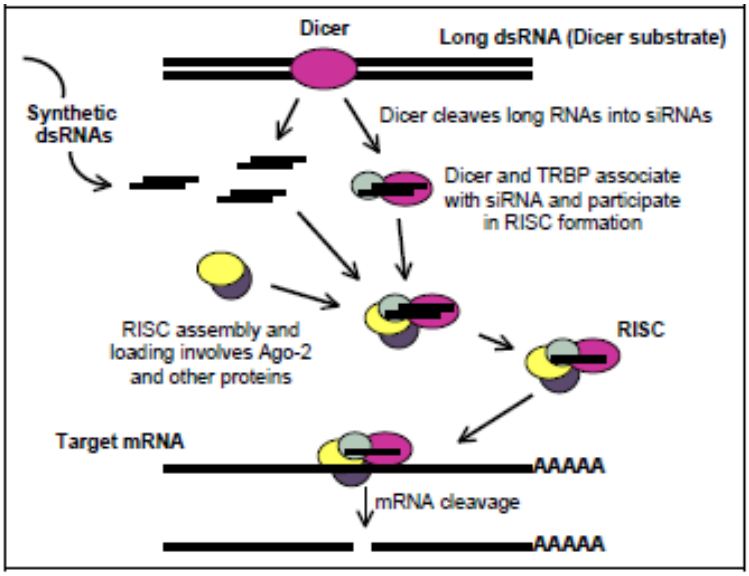


Figure 4-5. Mechanism involved in RNAi interference (Peek and Behlke, 2007).

This chapter describes study of overexpressing and silencing of AT2G23945 Aspartyl-protease gene on tapetum degradation process and its expression profiles.

4.2 Results

4.2.1 Expression analysis of AT2G23945 in different plant tissues using qRT-PCR.

Expression of AT2G23945 was analysed in staged buds by qRT-PCR. Buds were numbered starting with the immature buds: B1 to the mature buds B4 (Figure 4-6). The different stages of buds were compared to defined morphological landmarks of anther (Sanders *et al.*, 1999) and pollen development (Sebastian Bednarek's Personal Website: https://biochem.wisc.edu/labs/bednarek/dapi-pollen-development-research-bednarek-lab). The stages can be described as: B1: early anther development to tetrad stage (Anther Stage 1–6); B2: polarised microspore stage (Anther Stage 7– 10); B3: bicellular and tricellular microspore stage (Anther Stage 11–14); B4: mature pollen stage (Anther Stage 12–15) (see Figure 4-22).

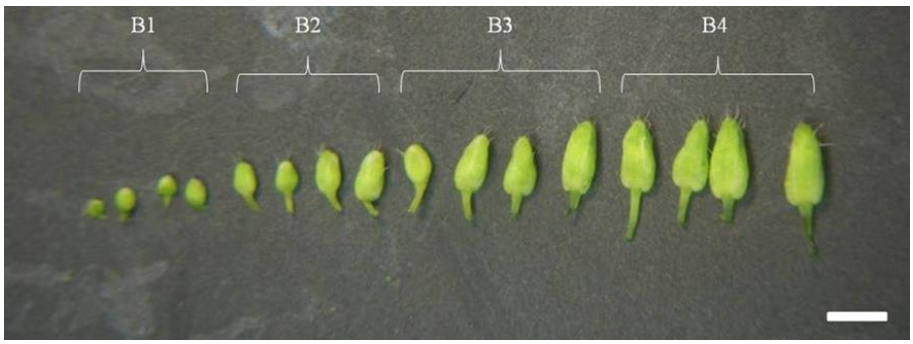


Figure 4-6. Multiple stages of buds that were used for the anther studies. B1: early anther development to tetrad stage; B2: polarised microspore stage; B3: bicellular and tricellular microspore stage; B4: mature pollen stage. Bar: 1mm

Control qRT-PCR was carried out on the staged cDNA using primers for the housekeeping gene *PP2A3*; these generated a band of the expected size (350 bp) (Figure 4-7A), and primers for AT2G23945 expression analysis 2g23945_486F and 2g23945_715R (Appendix 1) (Figure 4-7B). RT-PCR (Figure 4-7) and qRT-PCR analysis (Figure 4-8) using RNA extracted from different stages of bud development and other organs from WT-Col plants, revealed that high levels of AT2G23945 gene expression was detected in the B4 buds (mature pollen stage) and a low level of expression was detected in B1 (early anther development to tetrad stage) buds. No significant expression was detected in the other tissues tested (Figure 4-7B). This was confirmed by qRT-PCR, which showed that B4 (mature pollen) stage buds have the highest AT2G23945 activity, three times higher than B1. Moreover, a low level of expression was detected in the rest of the tissues investigated (Figure 4-8).

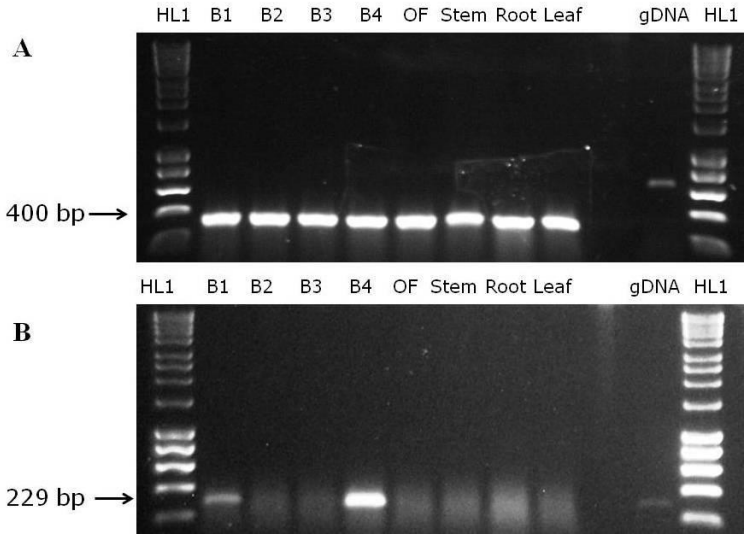


Figure 4-7. qRT-PCR analysis of expression of a) *PP2A3* Housekeeping gene and b) AT2G23945, in staged buds. Primers used (A) were *PP2A3* Forward primer and PP2A3 Reverse. Primers used for AT2G23945 expression (B) were 2g23945_486F and 2g23945_715R (28 cycles).

Lanes: B1: early anther development to tetrad stage; B2: polarised microspore stage; B3: bicellular and tricellular microspore stage; B4: mature pollen stage OF: Open flower, gDNA: WT-Col genomic DNA, act as a control, HL1: HyperLadder 1.

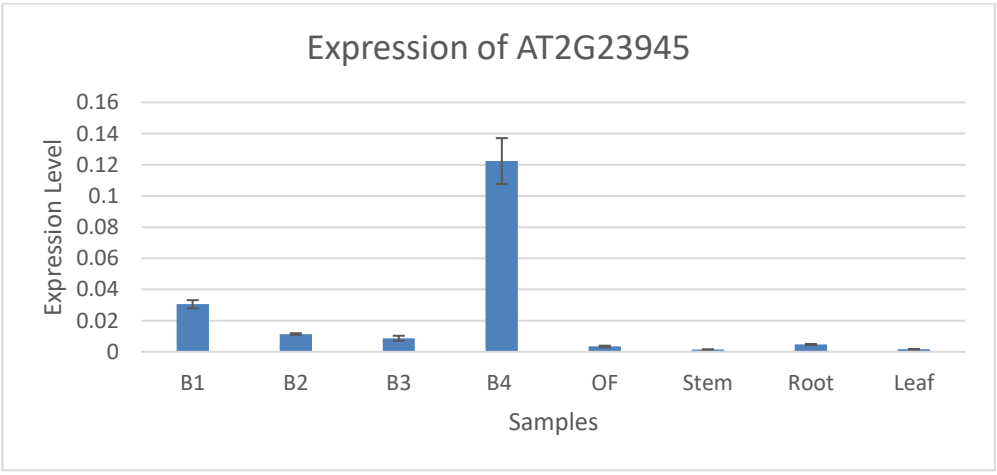


Figure 4-8. Comparison of AT2G23945 expression by qRT-PCR in WT-Col in different tissues. Bar represents a mean value (n=3) of the *AT2G23945* normalising using *PP2A3* expression. Error bars represent standard deviation.

Keys: B1: early anther development to tetrad stage; B2: polarised microspore stage; B3: bicellular and tricellular microspore stage; B4: mature pollen stage OF: Open flower.

4.2.2 Alteration of AT2G23945 gene expression using SALK T-DNA insertion lines

To explore the function of the AT2G23945 gene, three T-DNA lines which carry insertions in the promoter region or close to the transcription start site (TSS) were analysed. These were identified from the SALK T-DNA insertion population (Alonso *et al.*, 2003); SALK 135523 is located in the promoter region at -83bp upstream of TSS, while SALK 134917 and SALK 134919 are located in an Exon 1 downstream of the TSS +16bp and +51bp, respectively (Figure 4-9). All plants derived from the T-DNA insertion lines originated from the same original batches (provided by NASC) of in *Col* background.

Confirmed SALK 135523 and SALK 134919 insertion mutants were screened for defects that effect plant fertility. No insertion was detected for SALK 134917 (data not shown) and it is believed that the SALK lines have lost the insertion; no further analysis was conducted for this line. For further analysis of SALK 135523 and SALK 134919, three different plants were analysed from both lines. Two set of primers, SALK_134917 LP coupled with SALK_134917 RP for gene specific primers, and LBb1.3 left border primers and SALK_134917 RP for insertion specific primers were used to genotype the lines for presence of the T-DNA insertions.

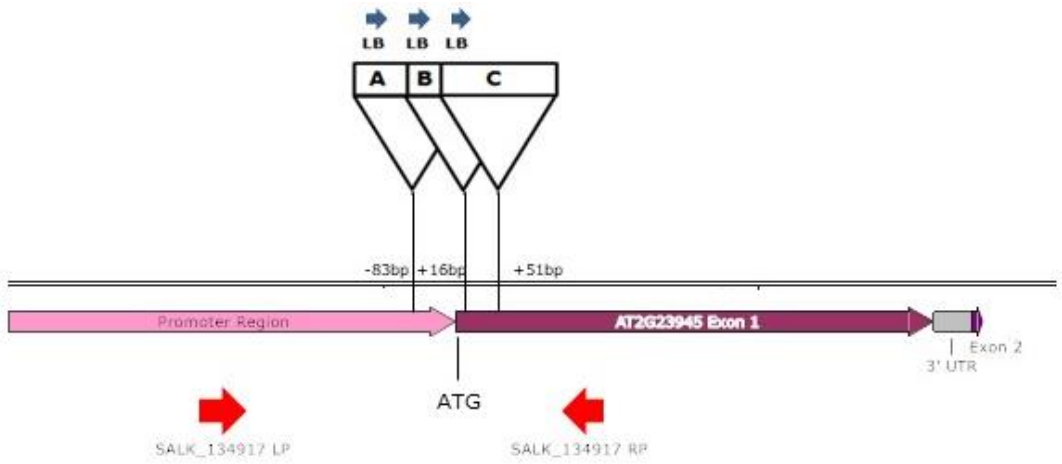


Figure 4-9. Schematic map of AT2G23945 gene with T-DNA insertions. Exons are highlighted by the maroon arrow; promoter region is indicated by pink arrow while 3' UTR region is highlighted in grey. Red arrows indicate the SALK insertion primers. LBb1.3 left border primers are shown indicating the insertion orientation (blue arrows). Location of the T-DNA insertions were shown in this map: A) SALK 135523 where the T-DNA insertion started -83bp upstream of the start codon; B) SALK 134917 where the T-DNA insertion started at +16bp downstream of start codon; C) SALK 134919 where the T-DNA insertion started at +51bp downstream of start codon. SALK_134917 LP was used as left primer and SALK_134917 RP was used as right primer for PCR.

4.2.2.1 SALK 135523

Three homozygous SALK 135523 lines, which carry a T-DNA insert in AT2G23945, were genotyped by PCR amplification (Figure 4-10). Amplification products were detected using insert and gene specific primers for the homozygous mutants but not for gene specific primers (Figure 4-10).

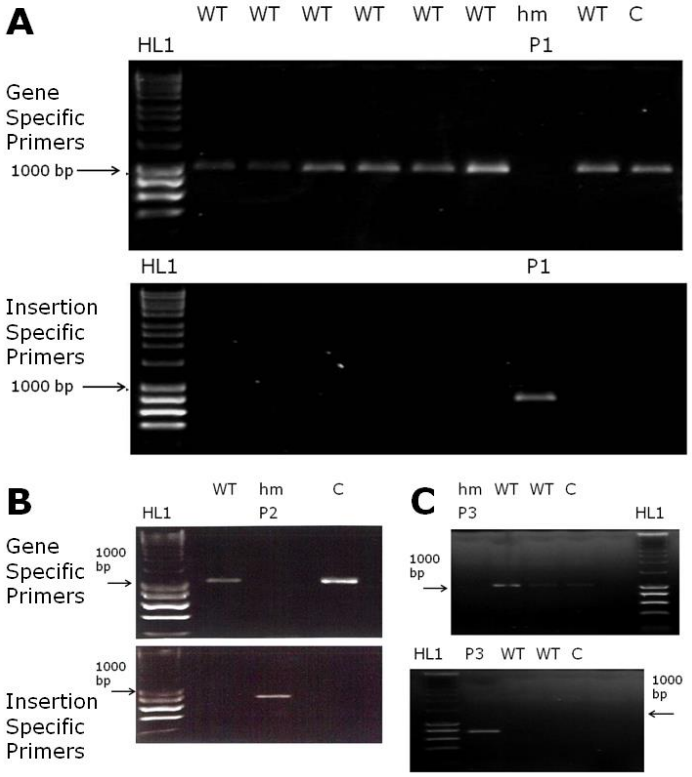


Figure 4-10. PCR genotyping of three SALK 135523 T-DNA insertion plants in three different panels where A: SALK 135523 P1 (P1), B: SALK 135523 P2 (P2). and C: SALK 135523 P3 (P3). Gene specific primers used were SALK_134917 LP and SALK_134917 RP; Insertion specific primers refer to SALK_134917 RP and LBb1.3 left border primer (see Figure 4.8). hm: homozygous for T-DNA insertion; WT: Mutant that are Wild-type with no insertion; C: WT-Col genomic DNA, HL1: HyperLadder 1.

Three SALK 135523 plants (P1, P2 and P3) were screened by RT-PCR to examine the expression of AT2G23945 gene in comparison to the Housekeeping *PP2A3* gene (Figure 4-11A). qRT-PCR expression analysis using AT2G23945 forward and reverse primers generated a band of approximately ~230 bp in all SALK plants (Figure 4-11B). Surprisingly, the three SALK 135523 lines showed brighter bands compared to the control wild type bands, suggesting increased levels of AT2G23945 expression.

qRT-PCR confirmed the RT-PCR results and the effects of disruption of the gene by the T-DNA insertion. Interestingly, the level of AT2G23945 expression was increased in all the SALK plants (Figure 4-12; Table 4-1). The expression of the gene varied and was three, seven and six times higher for P1, P2 and P3 respectively, compared to the WT-Col.

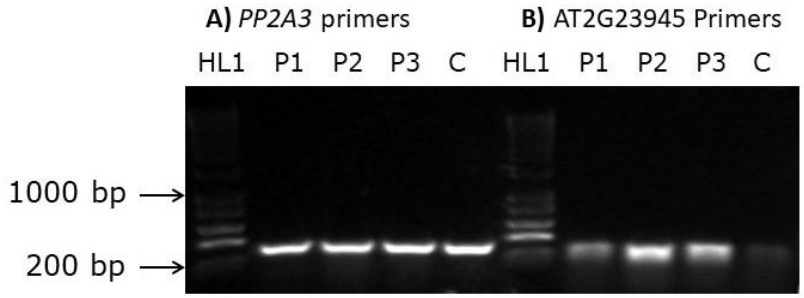


Figure 4-11. Expression analysis of AT2G23945 in SALK 135523 plants using RT-PCR. A: *PP2A3* expression analysis. B: AT2G23945 expression. P1-P3: Three different homozygous plants of SALK 135523, C: WT-Col genomic DNA, HL1: HyperLadder 1.

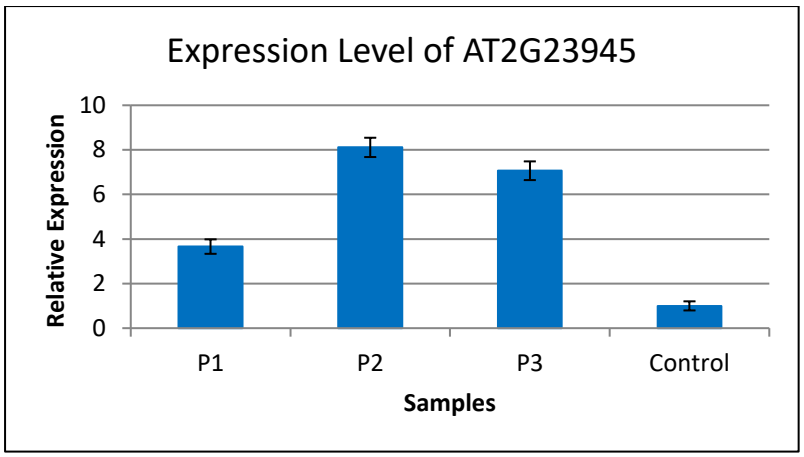


Figure 4-12. Expression pattern of AT2G23945 gene in the three homozygous SALK 135523 plants. Bar represents a mean value (n=3) of the AT2G23945 cDNA relative to the Housekeeping *PP2A3* cDNA. The relative expression level of each gene in the line was set as 1. Error bars represent standard deviation. P1-P3: Three different homozygous plants of SALK 135523.

Phenotypic analysis of these lines identified a reduction in fertility and delayed flowering compared to the WT-Col (Figure 4-13). The levels of sterility were similar for all 3 SALK plants, 64 %, 69.2 %, and 61.9 % for P1, P2, and P3 respectively; whereas only 10 % sterility was seen in the wild-type controls (Table 4-1). The initiation of flowering was also delayed for all SALK 135523 plants compared to WT-Col. The flowers, stamens, filaments and carpels of SALK 135523 plants were of equivalent size and growth characteristics to the wild type. Furthermore, pollen dehiscence occurred normally during flower opening for all homozygous SALK 135523 plants (Figure 4-14).

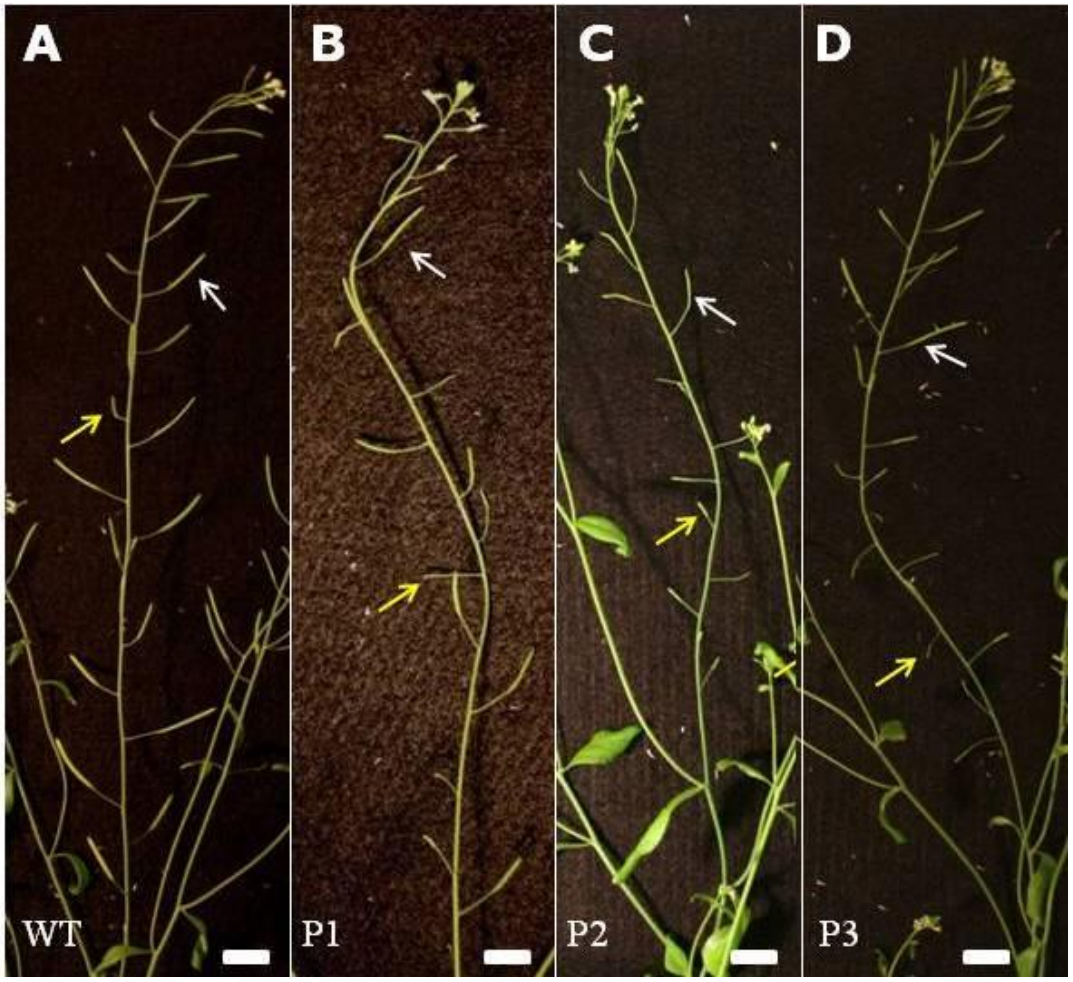


Figure 4-13. Phenotypic characterization of T-DNA insertion plants, showing the main stem of WT-Col and the three homozygous SALK 135523 plants (P1-3)(B-D). Images were taken of 6-week old plants. The SALK plants)(B-D) showed reduced fertility as indicated by an increased number of small siliques, compared to the control (A). White arrow: Fertile siliques. Yellow arrow: Sterile siliques. Bar = 1 cm.

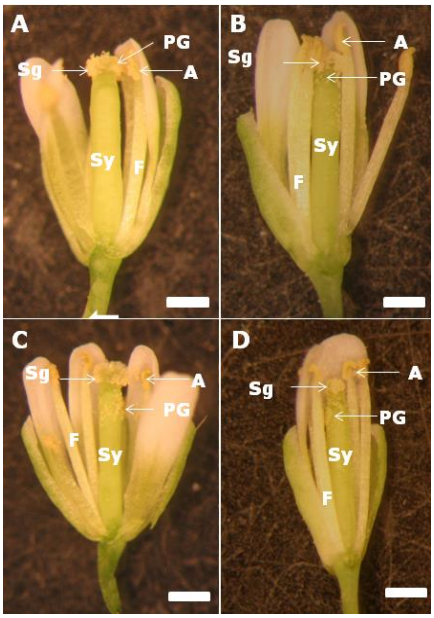


Figure 4-14. Normal dehiscence was observed for all three SALK 135523 plants. Representative flower from three biological replicates are shown for each sample. Two or three sepals were removed prior to taking images. A: WT-Col as a control. B: P1 flower; C: P2 flower; D: P3 flower. Flower containing dehisced anthers at flower opening. Bar= 1 mm.

Note: Sg: Stigma. F: Filament. PG: Pollen Grain. A: Anther. Sy: Style. Arrow: Anther attached to the stigma. Bar: 100 µm.

Alexander's stain (Alexander, 1969) was used to assess pollen viability on whole anther squashes; red stained pollen indicates the pollen is viable, whilst non-viable pollen was green/unstained. One hundred pollen grains were counted for pollen viability counts from each SALK 135523 plant. Approximately 60% of pollen in the three SALK 135523 plants were inviable (Figure 4-15 B-D), whereas almost all pollen from WT-Col samples were viable (Figure 4-15A). This suggests that the reduction in fertility in the SALK 135523 line was at least partly due to a reduction in pollen viability (Table 4-1).

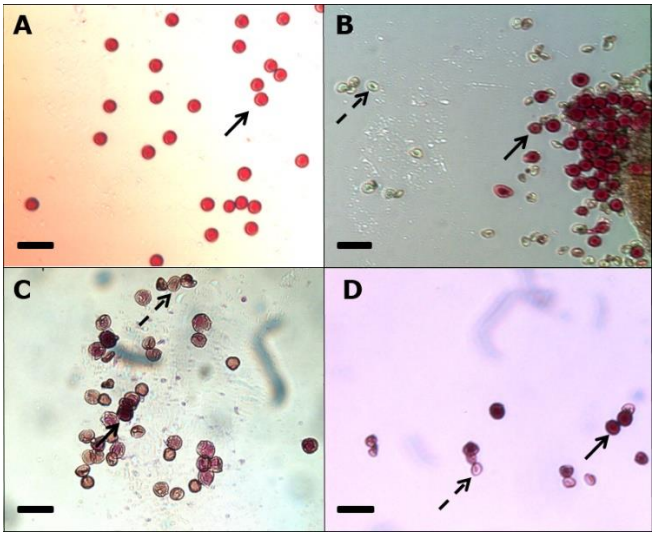


Figure 4-15. Pollen viability assessment using Alexander's staining for all three SALK 135523 plants where one hundred pollen grains were counted from every independent plant. Pollen from A: WT-Col, B: P1 line; C: P2 line; D: P3 line. Solid arrows indicate viable pollen grains which are stained red while non-viable pollen is shown by dashed arrows. Bar = 20 µm.

Based on the statistical analysis (See section 2.1.22), overexpression of the AT2G23945 gene in all three selected SALK 135523 showed significant effect on mean of silique length when compare to the WT-Col (Figure 4-16).

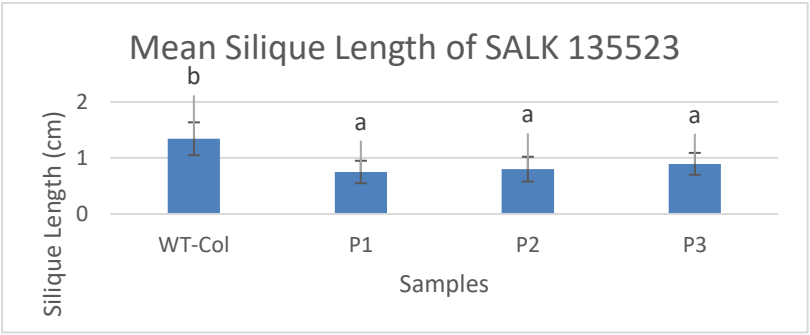


Figure 4-16. Mean silique length (cm), indicating fertility, between the three overexpression SALK 135523 plants (P1-3) and the WT-Col control after 6 weeks. Mean silique length was obtained for all siliques, sterile and non-sterile by dividing total length by total number of siliques on the main stem. Identical letter signifies no significant difference where $P<0.05$; ANOVA, Fisher's Least Significant Difference test. Error bars represent standard error.

Opening of the fertile siliques indicated that embryos of the T-DNA insertion lines and control developed to become normal seeds and all normal siliques were filled with seeds (Figure 4-17).

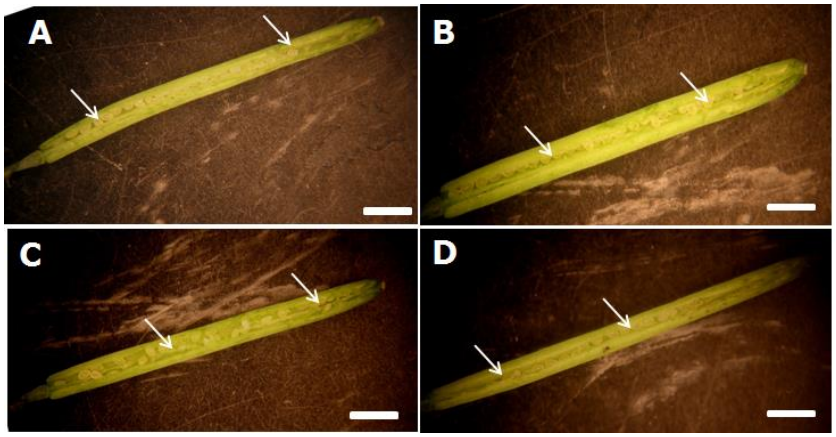


Figure 4-17. Seed development of different SALK 135523 plants fertile silique and WT-Col plant fertile silique. Representative silique from three biological replicates are shown for each sample. Comparable seed development of all SALK 135523 plants and WT-Col upon dissection was observed. Fertile Silique of A: wild-type Col as a control. B: P1 silique; C: P2 silique; D: P3 silique. Solid arrow: Normal seed. Bar: 200 μm.

The SALK 135523 pollen were tested with the Auramine O staining (Kommareddi *et al.*, 1984) to check the nature of their exine pollen wall. The abnormal pollen from all three SALK 135523 plants had disrupted surfaces, with shrunken exine (Figure 4-18B-D); in contrast, normal wild type pollen was round and of regular shape (Figure 4-18A).

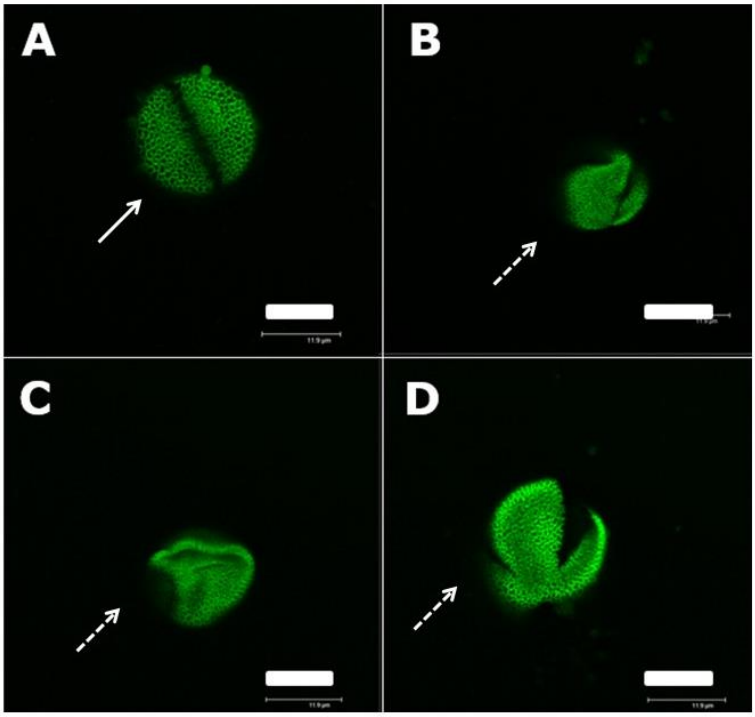


Figure 4-18. Pollen anlaysis of exine structure using Auramine O staining for three SALK 135523 plants. Representative pollen stained with Auramine O from three biological replicates are shown for each sample. A: WT-Col; B: P1; C: P2; D: P3. B, C and D shows abnormal structure pollen exine compared to WT-Col (A). Solid arrow pointed at normal pollen, dashed arrow pointed at abnormal pollen. Bar: 10 µm

Anther development shows that the SALK 135523 and control plants have identical anther and pollen development during their early stages, until stage 10 of anther development (Figure 4-19). Microspores mother cells (MMC) developed normally during early anther development (Figure 4-19A and F), then, the MMC developed to become meiotic cells (MC), the tapetum cells and middle layer could be be clearly seen at this stage, no differences were seen between wild type and the mutant (Figure 4-19B and G). The middle layers gradually degenerated and tapetum vacuolation occurred

normally in both control and SALK 135523 lines (Figure 4-19C and H).

Significant differences between the mutant anthers and control started during anther stage 11 (bicellular anther stage, B3 buds). The control tapetum cells started degeneration, with small developing pollen grains evident (Figure 4-19D). In contrast, the vacuolated tapetum remained in the SALK 135523 anthers and microspore development was delayed (Figure 4-19I). When the anther reached maturity, several important events occurred in control anthers. The anther endothecium cells develop a lignin cellulose layer of secondary thickening, and degeneration of the septum and stomium breakage, and anther dehydration mechanically forced the mature anther to dehisce prior to pollination and fertilization (Figure 4-19E). By contrast, delayed tapetum degeneration was seen in the SALK 135523 lines, pollen grains appeared abnormal; delayed endothecium thickening was seen and no septum and stomium breakdown was observed, even during the late anther stage development (Figure 4-19J). Higher magnification of bicellular and mature pollen stages are presented in Figure 4-20.

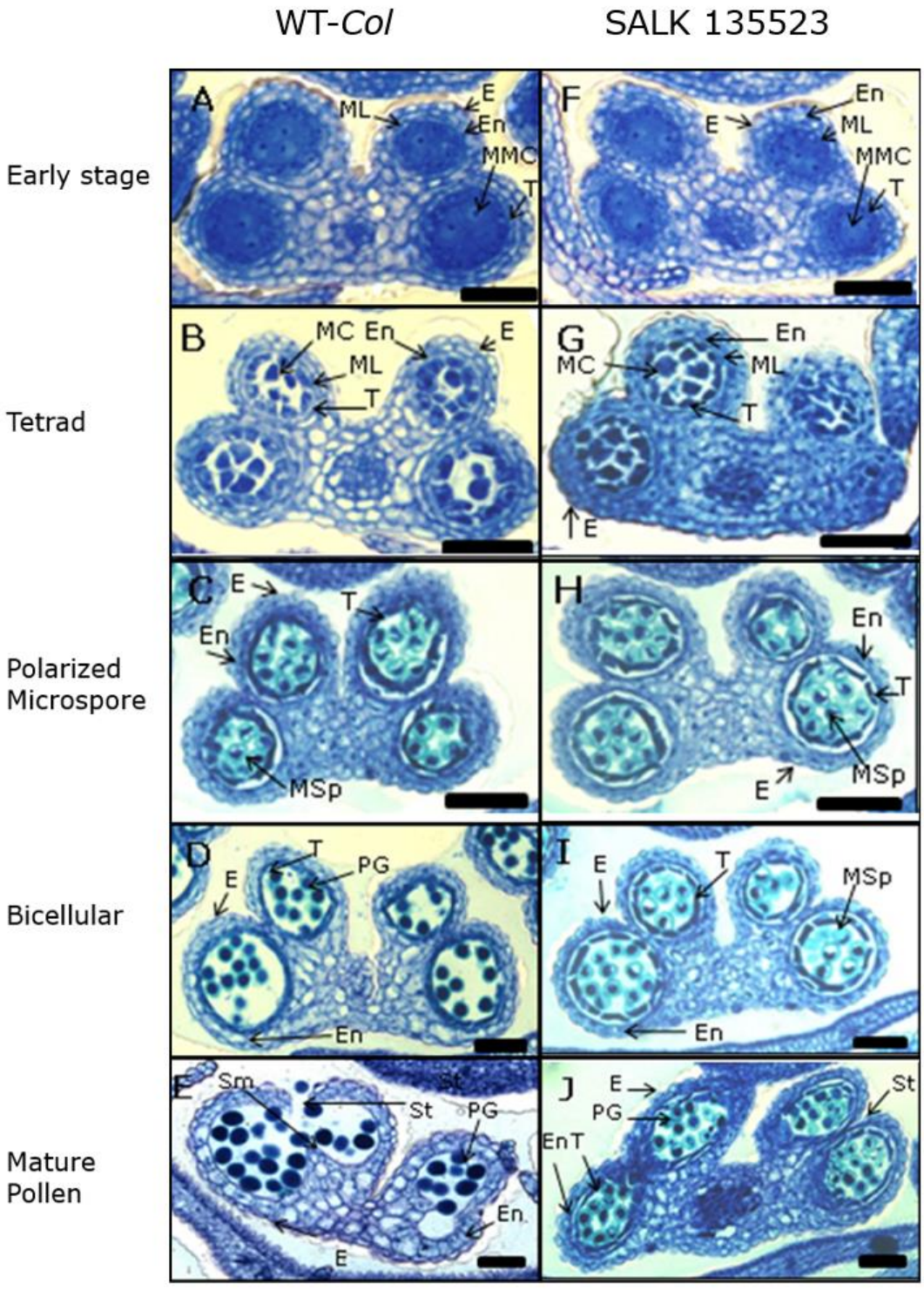


Figure 4-19. Phenotypic analysis of anther of the WT-Col and SALK 135523 plants with equivalent stages of pollen development. Representative anther sectioning from three biological replicates are shown for each sample. Left panel (A-E): WT-Col anthers. Right panel (F-J): SALK 135523 anthers. Bar: 50 µm.

Keys: E: Epidermis, En: Endothecium, ML: Middle layer, T: tapetum, MMC: Microspore mother cell, MC: Meiotic cell, MSp: Microspores, PG: Pollen grain, Sm: Septum, St: Stomium.

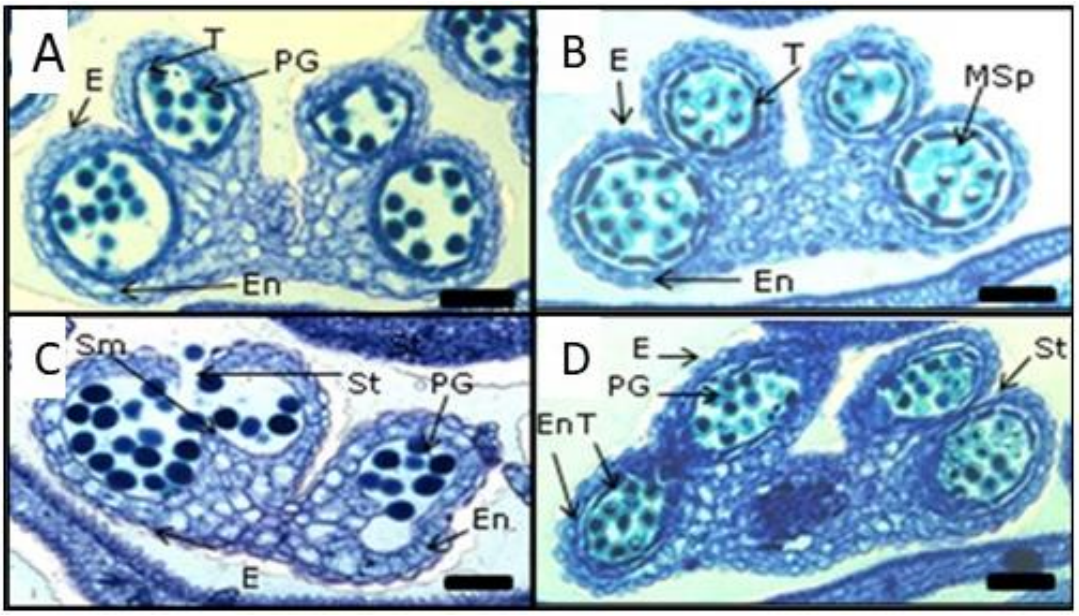


Figure 4-20. Enlarged images of bicellular and mature anthers from Figure 4-18. A and C: WT-Col. B and D: SALK 135523. A and B: Bicellular stage; C and D: Mature pollen stage. Bar: 50 μm.

Acridine orange/Ethidium bromide staining indicated that endothecium thickening occurred normally prior to dehiscence for the SALK 135523 homozygous plants (Figure 4-21).

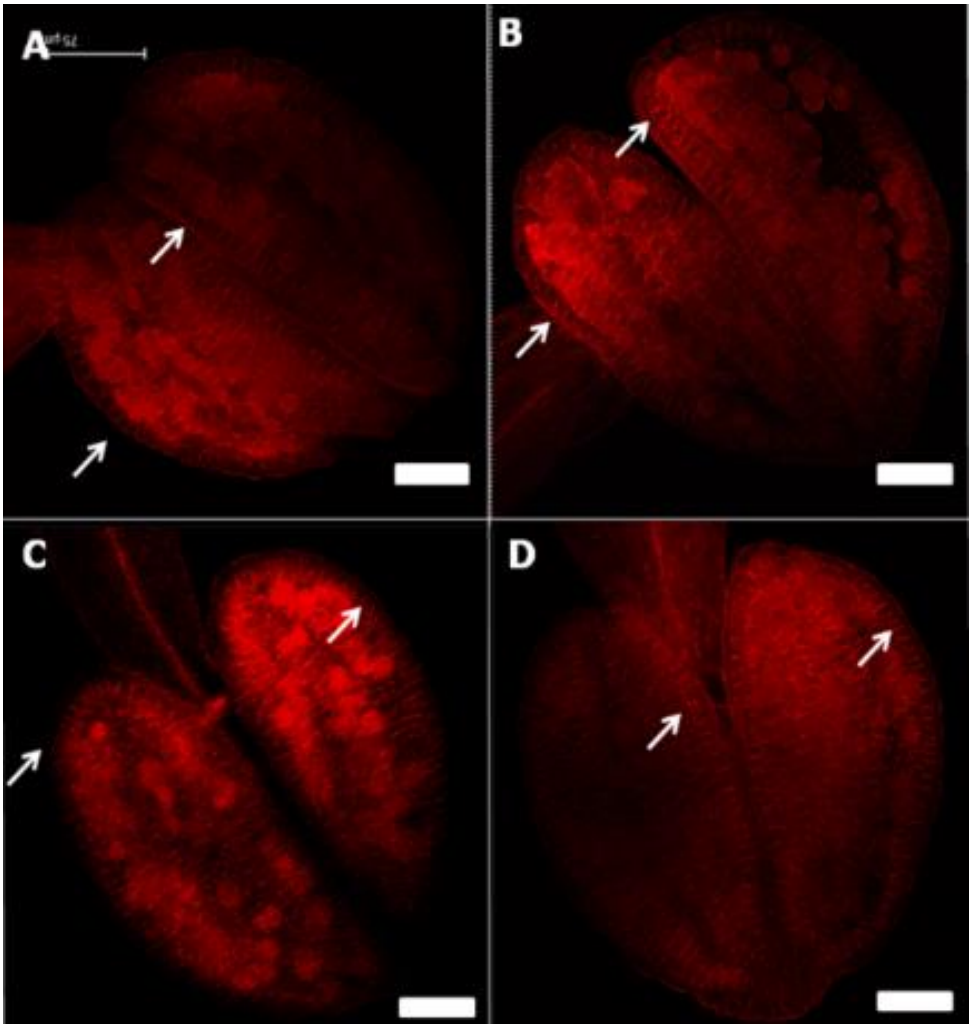


Figure 4-21. Analysis of lignification of the WT-Col and SALK 135523 (P1-P3 plants) anthers using ethidium bromide acridine orange staining. A) WT-Col B) P1; C) P2; D) P3. Representative pollen stained with ethidium bromide/acridine orange from three biological replicates are shown for each sample. Comparable endothecium development in all SALK 135523 plants and WT-Col was observed. Arrows indicate normal lignified endothecium thickening. Bar: 50 µm.

SALK 135523 and control plant have an identical anther development during their early stage until the bicellular stage (B3 buds) of pollen development (Figure 4-22). Normal meiosis occurs to produce a tetrad (Figure 4-22A and G). DAPI (4’,6-diamidino-2-phenylindole) staining was employed to study the progression of the pollen development. Normal meiotic development was seen for both

WT-Col (Figure 4-22A-C) and SALK 125523 (Figure 4-22G-I) pollen in early pollen development between tetrad and polarized microspore stage; the nucleus can be seen throughout these stages. Subsequently, during the bicellular stage in the nucleus divided to form a single generative cell (bright circle) and one vegetative tube cell (faint circle) (Figure 4-22D) indicating mitosis I occurred normally. After that, in the wild-type two bright spots were seen during tricellular stage, indicating normal generative cell division to produce two male gametes during Mitosis II stage (Figure 4-22E). In contrast, the SALK 135523 pollen abnormality started to occur during bicellular stage and afterwards (Figure 4-22J-L). The SALK 135523 plants were semi-sterile, where some of the pollen managed to develop normally and some of the pollen progression halted at bicellular stage.

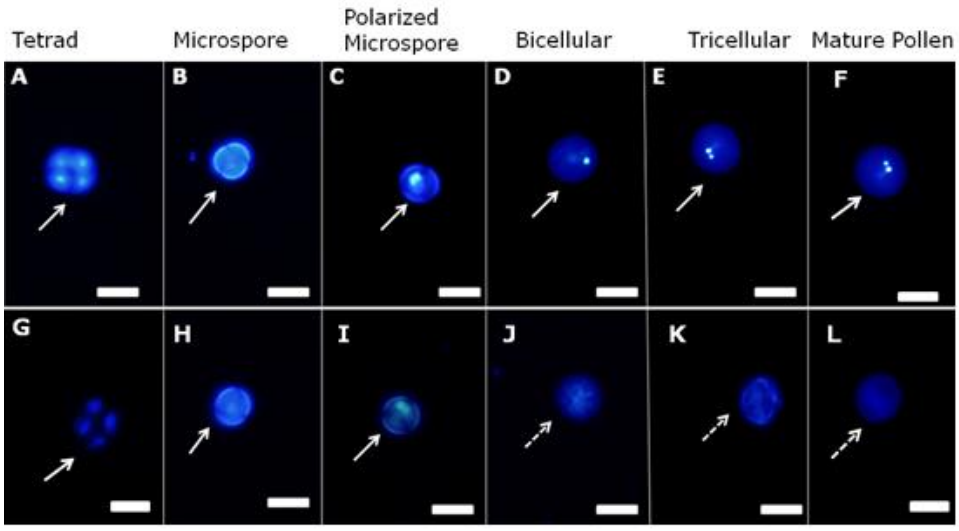


Figure 4-22. DAPI staining of nuclear development of pollen from the WT-Col and SALK 135523 plants. The SALK 135523 pollen resemble the WT-Col until bicellular stage. Representative DAPI staining from three biological replicates are shown for each sample Top panel (A-F): Control; Bottom panel (G-L): SALK 135523. Solid arrow point out normal pollen, while the dashed arrow indicates abnormal pollen. Bar: 10 µm.

Table 4-1. Summary of the analysis of fertility in the SALK 135523 T-DNA insertion lines. The total number of fully developed siliques, percentage of sterility on main stem, pollen viability percentage were determined alongside transcript level changes in the three SALK 135523 T-DNA insertion lines and WT-Col.

Sample	**Total no. of siliques**	**Percentage of sterility on main stem (%)**	**Pollen Viability (%)**	**Transcript Level Changes**
Wild-type *Col* (control)	21	10	95	1
SALK Plant 1 (P1)	14	64	60	3 X
SALK Plant 2 (P2)	13	69.2	60	7 X
SALK Plant 3 (P3)	21	61.9	60	6 X

4.2.2.2 SALK 134919

To confirm the observed defects were associated with altered expression of AT2G23945, an independent investigation of SALK 134919 line was conducted. It carried a T-DNA insert in an Exon 1 at downstream of TSS +51bp, whereas T-DNA insert for SALK 135523 is in the promoter region at -83bp upstream of TSS. Eight SALK 134919 homozygous plants were genotyped using a mixture of Gene specific and insertion specific primers (SALK_134917 LP, SALK_134917 RP, and LBb1.3 left border primers) which successfully produced bands (~650 bp) that proved the insertion in the SALK plant's genome (Figure 4-23). Three homozygous plants were examined for phenotypic changes and compared to SALK 135523.

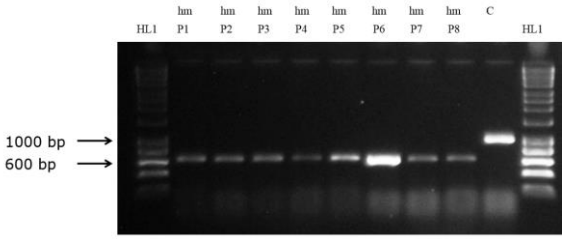


Figure 4-23. Genotyping of SALK 134919 T-DNA insertion plants (P1-P8). SALK insertions were confirmed by the PCR. The primers used were SALK_134917 LP, SALK_134917 RP and LBb1.3 left border primer. hm: homozygous for insertion, P1-P8: Eight different homozygous plants of SALK 134919.WT: Mutant that are Wild-type with no insertion, C: WT-Col genomic DNA, HL1: HyperLadder 1.

Buds from three SALK plants (P1, P5 and P6) were screened by RT-PCR to examine the expression of the AT2G23945 gene. Expression analysis using Housekeeping *PP2A3* primers amplified a band of the correct size band of ~350 bp (Figure 4-24A). Expression analysis using AT2G23945 forward and reverse expression primers amplified 230 bp band in all SALK plants (Figure 4-24B). The three SALK 135523 bands were brighter compared to the wild-type indicating increased expression in the transgenic lines.

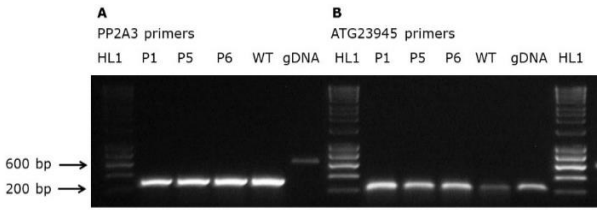


Figure 4-24. Expression analysis of AT2G23945 in buds from multiple SALK 134919 plants by RT-PCR. A: *PP2A3*. B: AT2G23945. P1, P5 and P6: Three different homozygous plants of SALK 134919, C: Wild-type *Col* genomic DNA, HL1: HyperLadder 1.

qRT-PCR expression analysis of AT2G23945 gene in buds from the three SALK 134919 plants confirmed the results of RT-PCR (Figure 4-24) that the level of AT2G23945 expression in all SALK plants was

increased. Expression was increased by approximately 20, 16 and 13-fold for P1, P5 and P6 respectively, compared to the WT-Col (Figure 4-25).

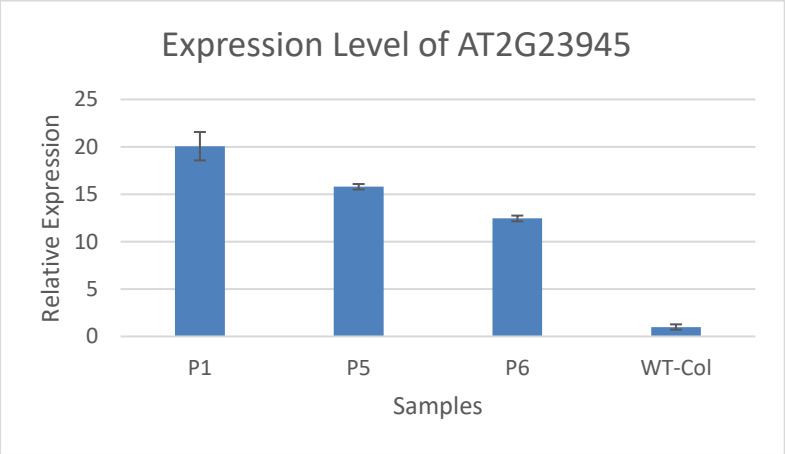


Figure 4-25. qRT-PCR analysis of AT2G23945 in buds from three homozygous SALK 134919 plants. Bars represents a mean value (n=3) of the *AT2G23945* relative to the Housekeeping *PP2A3* gene. The relative expression level of each gene in the line was set as 1. P1, P5 and P6: Three different homozygous plants of SALK 134919. Error bars represent standard deviation.

Phenotypic analysis of three independent SALK 134919 plants confirmed a partial fertility phenotype compared to the WT-Col (Figure 4-26) as previously observed in the SALK 135523 transgenic lines. SALK 134919 showed defects in fertility in terms of percentage sterility on the main stem. The sterility rates were almost identical for the 3 SALK plants, 27.2 %, 28.5 %, and 25 % for P1, P5, and P6 respectively (Figure 4-26B-D); whereas only 3 % (Figure 4-26A) sterility was seen for wild-type controls (Table 4-2). Initiation of flowering occurred at a similar stage for the three SALK plants compared to the wild-type *Col*.

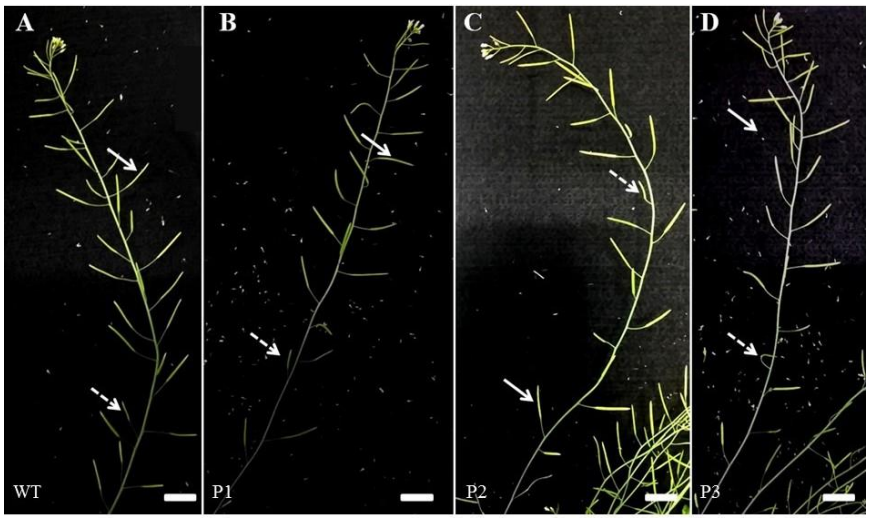


Figure 4-26. Main stem of three different SALK 134919 plants (P1-P3) and WT-Col control of 6 weeks old plants. A: WT-Col control. B: P1; C: P5; D: P6. The three SALK plants appeared phenotypically normal bud with reduced fertility compared to the control (A). White arrow: Fertile silique. Dashed arrow: Sterile silique. Bar = 1 cm.

Based on the statistical analysis (See Section 2.1.22), overexpression of the AT2G23945 gene in all three selected SALK 134919 showed significant effect on mean silique length when compare to the WT-Col. In addition, all SALK 134919 plants did not differ significantly in terms of mean of silique length when compare to one another (Figure 4-27).

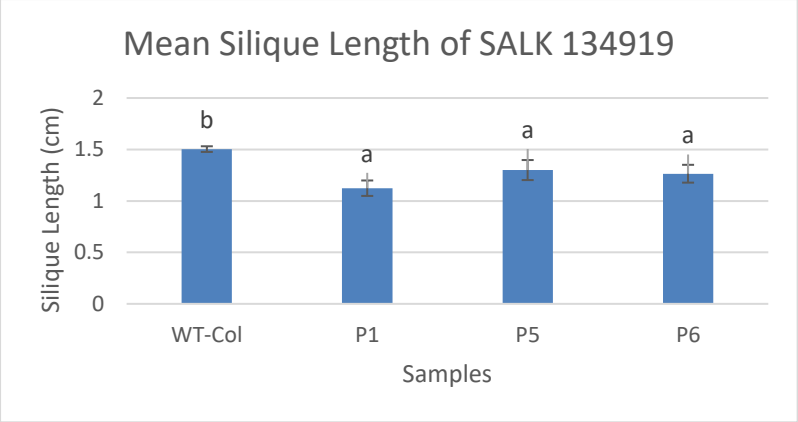


Figure 4-27. Mean silique length (cm), indicating fertility, between the three overexpression SALK 134919 plants (P1, P5 and P6) and the WT-Col control after 6 weeks. Mean silique length was obtained for all siliques, sterile and non-sterile by dividing total length by total number of siliques on the main stem. Identical letter signifies no significant difference where $P<0.05$; ANOVA, Fisher's Least Significant Difference test. Error bars represent standard error.

Table 4-2. Summary of the analysis of fertility in the SALK 134919 T-DNA insertion line. The total number of fully developed siliques and percentage of sterility on main stem were determined alongside transcript level changes in the three SALK 134919 T-DNA insertion lines and WT-Col.

Sample	**Total no. of siliques**	**Percentage of sterility on main stem (%)**	**Transcript Level Changes**
Wild-type *Col* (control)	28	3	1
SALK Plant 1 (P1)	25	27.2	20 X
SALK Plant 5 (P5)	21	28.5	16 X
SALK Plant 6 (P6)	24	25	13 X

4.2.3 Manipulation of AT2G23945 gene expression by RNAi

To knock-down the expression level of the gene, RNAi was employed. The non-conserved region between 85-187 bp upstream of TSS was chosen as a RNAi target. This 102bp conserved region was selected using Gramene open-source website (http://gramene.org/). The amplification of the fragment was performed using RNAiF2g23945 and RNAiR2g23945 (Appendix 1).

The 102 bp AT2G23945RNAi target fragment was cloned into the pCR8/GW/TOPO and confirmed by PCR RNAiF2g23945 and pCR8 GW 2.1 R vector specific primer (Figure 4-28). pCR8/GW/TOPO clone 6 (C6) was used for further analysis after confirmation by sequencing.

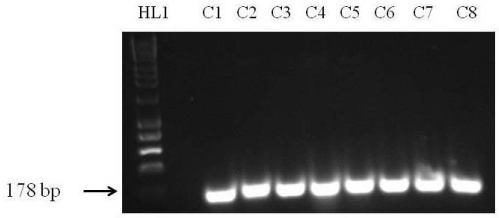


Figure 4-28. Colony PCR of insert and vector sequences from pCR8/GW/TOPO entry clone which contains 102bp AT2G23945RNAi target insert. Fragment amplification generated a 178bp fragment RNAiF2g23945 and pCR8 GW 2.1 R reverse primer vector specific primer. C1-C8: Different colonies of the entry clone, HL1: HyperLadder 1.

The fragment of pCR8/GW/TOPO AT2G23945RNAi vector that includes attL region and the insert was amplified by PCR, (ROligo_New_CompF forward primer and ROligo_New_CompR reverse primer), and the 443bp insert was purified and used for the

LR reaction. This technique was performed to avoid multiplication of entry clone and destination vector in the same *E. coli* host cells since both vectors confer resistance to Spectinomycin.

Subsequently, the isolated fragments were transferred into the Gateway-compatible vector pK7GWIWG2(I). This destination vector consists of sense strand and anti-sense strand, linked by an intron, and control by the 35S CaMV promoter; this enables dsRNA expression in plants. The schematic map of pK7GWIWG2(I):AT2G23945RNAi hairpin construct consists of CaMV35S::$_{85}$AT2G23945$_{102}$::INTRON:$_{102}$AT2G23945$_{85}$ and primers location are shown in Figure 4-29. Amplification of different segments of the construct confirmed the clone identity (Figure 4-30); sequencing of pK7GWIWG2(I):AT2G23945RNAi clone 2 (C2) was used to confirm the sequence. Then, the pK7GWIWG2(I):AT2G23945RNAi plasmids were transformed into *Agrobacterium* using electroporation. After verification by PCR (Figure 4-31), GV3101 clone 2 (C2) were transformed into WT-Col *Arabidopsis*.

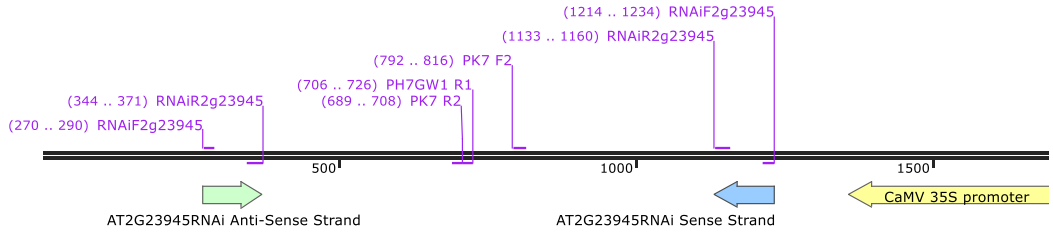


Figure 4-29. Schematic map of pK7GW1WG2:AT2G23945RNAi construct for generating RNAi mutant plants. CaMV 35S promoter indicated with the yellow arrow was used for overexpression purposes. Orientation and location of the sense-strand and anti-sense strand are indicated by the blue arrow and green yellow respectively. Furthermore, both insertional strands linked with an intron region. Location of the primers are shown.

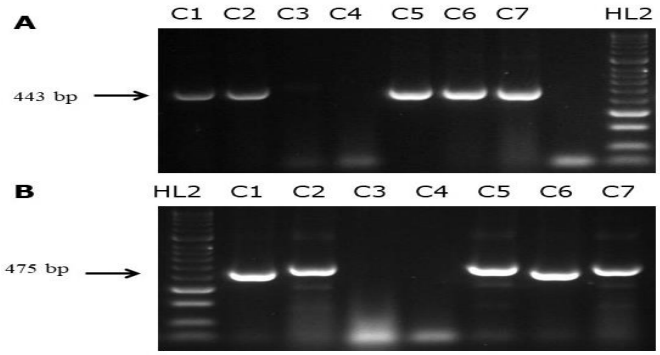


Figure 4-30. Colony PCR of pK7GWIWG2(I): AT2G23945RNAi destination vector. The amplifications were conducted targeting two regions consisting of (A) sense (443bp) and (B) anti-sense strand (475bp). (A) Sense strand amplification using RNAiF2g23945 with PK7 F2 vector specific forward primer. B) Anti-sense strand amplification RNAiF2g23945 with PH7GW1 R1 vector specific reverse primer. C1-C7: Different colonies of the destination vector, HL2: HyperLadder 2.

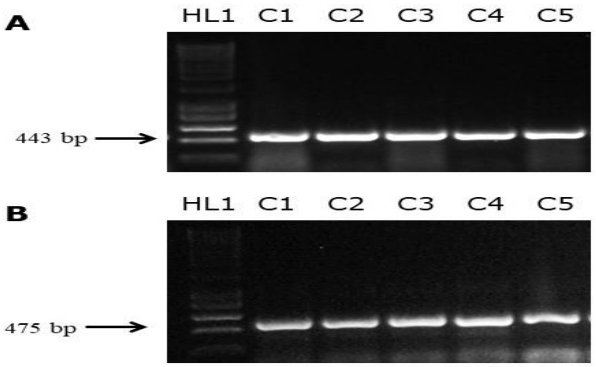


Figure 4-31. Colony PCR of GV3101 containing pK7GWIWG2(I): AT2G23945RNAi colonies. The amplifications were conducted targeting two regions consisting of (A) sense (443 bp) and (B) anti-sense strand (475 bp). A) Sense strand amplification using RNAiF2g23945 with PK7 F2 vector specific forward primer. B) Anti-sense strand amplification RNAiF2g23945 with PH7GW1 R1 vector specific reverse primer. C1-C5: Different colonies of the destination vector, HL1: HyperLadder 1.

The putatively transformed seeds (T_1) from all constructs were collected and screened using MS medium (Murashige and Skoog, 1962) containing Spectinomycin (50 ng/µl). Putative transgenic plants were transferred to soil and screened by PCR to confirm the integration of the transgene, sense and anti-sense strands (463 bp, Figure 4-32A and 458 bp, Figure 4-32B respectively). PCR analysis showed that seven putative transgenic plants were isolated.

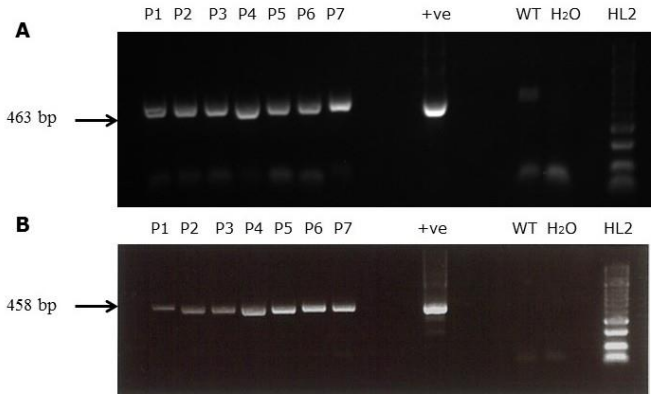


Figure 4-32. Selection of putative transgenic plants and PCR screening of transgenic plants. The presence of the insertions in putative transformants was confirmed by PCR analysis. Amplification of fragments with the right size of (A) 463 bp and (B) 458 bp. (A) sense strand amplification using RNAiF2g23945 with PK7 F2. (B) Anti-sense strand amplification using RNAiF2g23945 with PK7 R2. P1-P7: Seven different AT2G23945RNAi plants, +ve: pK7GWIWG2(I):AT2G23945RNAi vector as positive control WT: WT-Col genomic DNA as negative control, H_20: Water as a negative control, HL2: HyperLadder 2.

Six independent T_1 AT2G23945RNAi transformants (Figure 4-33) were analysed for phenotypic changes after 8-weeks. Flowers, stamens, filaments and carpels of transgenic plants were of equivalent size and growth characteristics to the WT-Col, however reductions in fertility were observed.

Interestingly, these plants showed different levels of sterility, five out of six transgenic plants P1-P5 (Figure 4-33B-F) appeared phenotypically different with reduced numbers of siliques compared to the WT-Col (Figure 4-33A). Meanwhile P7 (Figure 4-33G) had an increased number of siliques compared to the WT-Col plant Table 4-4. Interestingly, all transformant plants showed a longer life span. In addition, three plants showed stunted growth with two, eight and six siliques recorded for P1, P2 and P4 respectively.

The inflorescences (Figure 4-34) generally appeared phenotypically normal except for P1 (Figure 4-34B) and P4 (Figure 4-34E) which showed reduced bud number in all inflorescences of the plants. The WT-Col senesced after eight weeks growth, while all transformants continued to grow for up to 10 weeks. Pollen dehiscence occurred normally during flower opening for all transgenic plants.

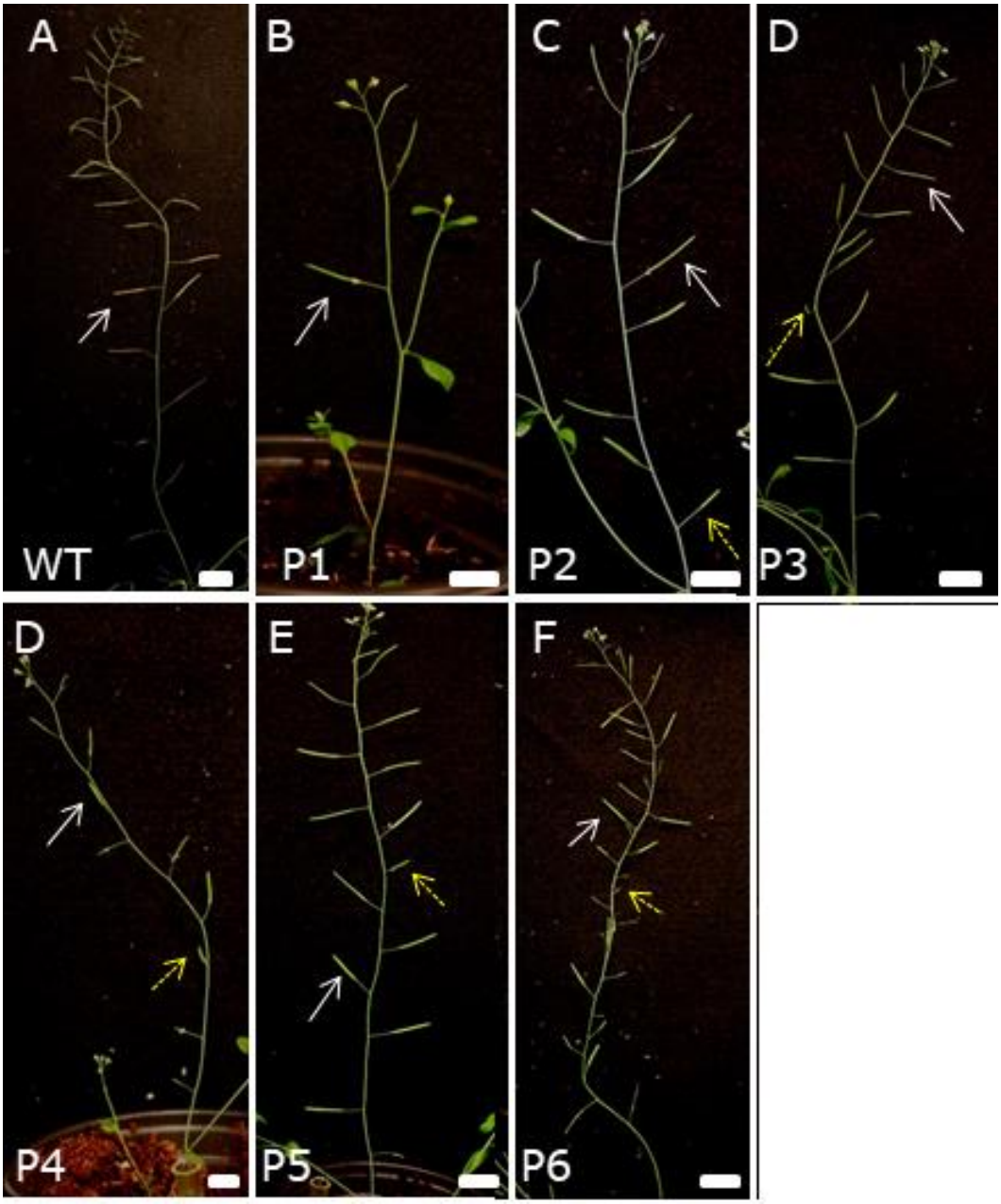


Figure 4-33. Characterization of AT2G23945RNAi transgenic main stems with different severities of phenotype. Main stem images of WT-Col as a control (A) and six independent AT2G23945RNAi plants (B-G) were taken of the eight weeks old plants. White arrow: Fertile silique. Dashed-Yellow arrow: Sterile silique. Bar = 1 cm.

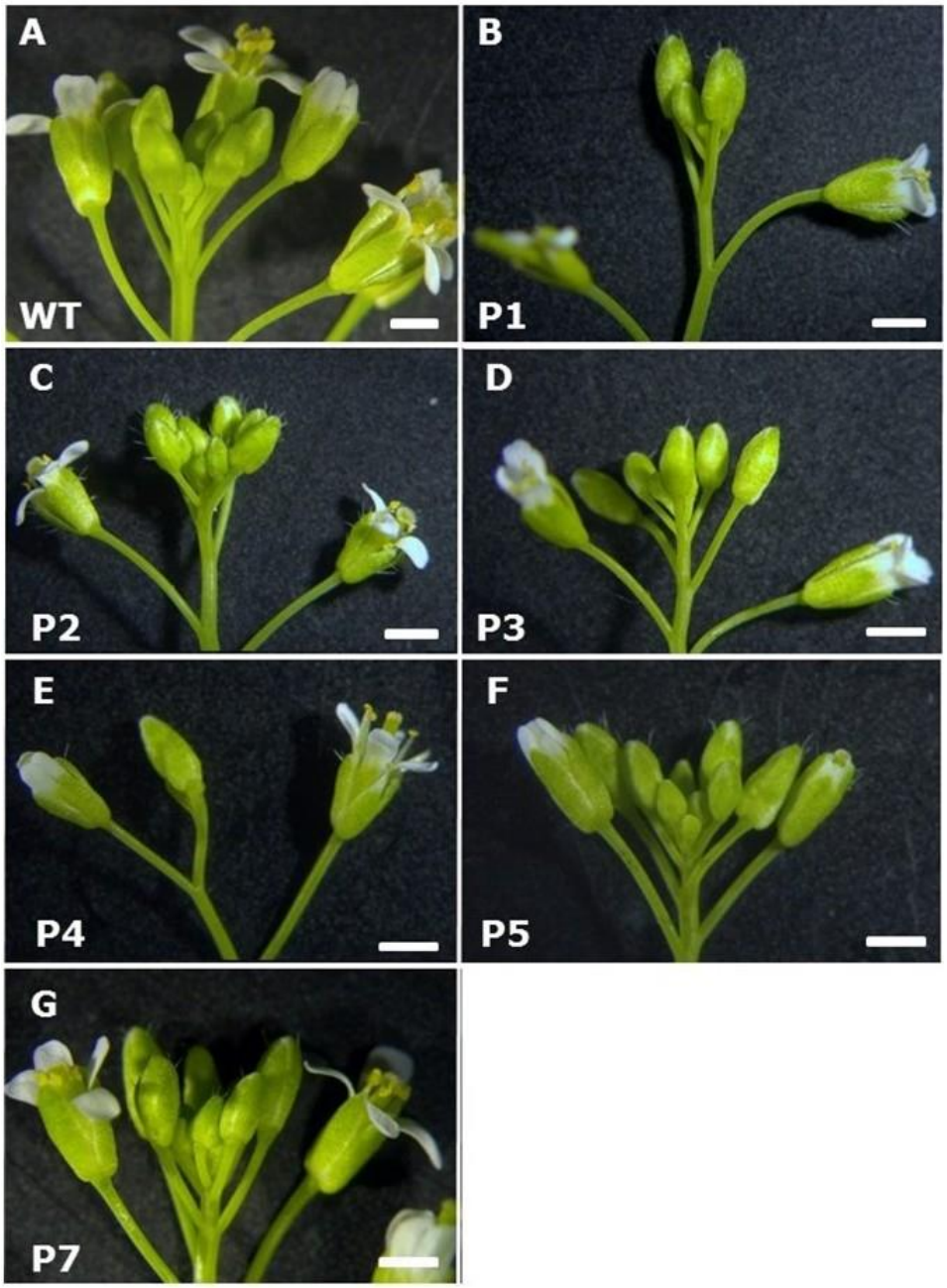


Figure 4-34. Characterization of AT2G23945RNAi inflorescence of six independent T_1 AT2G23945RNAi transgenic plants (B-G) as compared to WT- Col (A) from eight weeks old plants. Abnormal inflorescences with low numbers of buds were observed in all P1 (B) and P4 (E) inflorescences.

Pollen viability analysis from the pK7GWIWG2(I): AT2G23945RNAi T_1 transformant plants showed high viability in all transgenic lines except for P4 (E) and P7(G) (Figure 4-35). In P4 and P7 lines approximately 10% and 40% of pollen were viable, respectively (Table 4-3; Figure 4-36). Almost all pollen from WT-Col samples were fully viable (Figure 4-36A and B). Thus, non-viable pollen indicates that the AT2G23945RNAi plants produced a change of pollen fertility compared to the WT-Col plant.

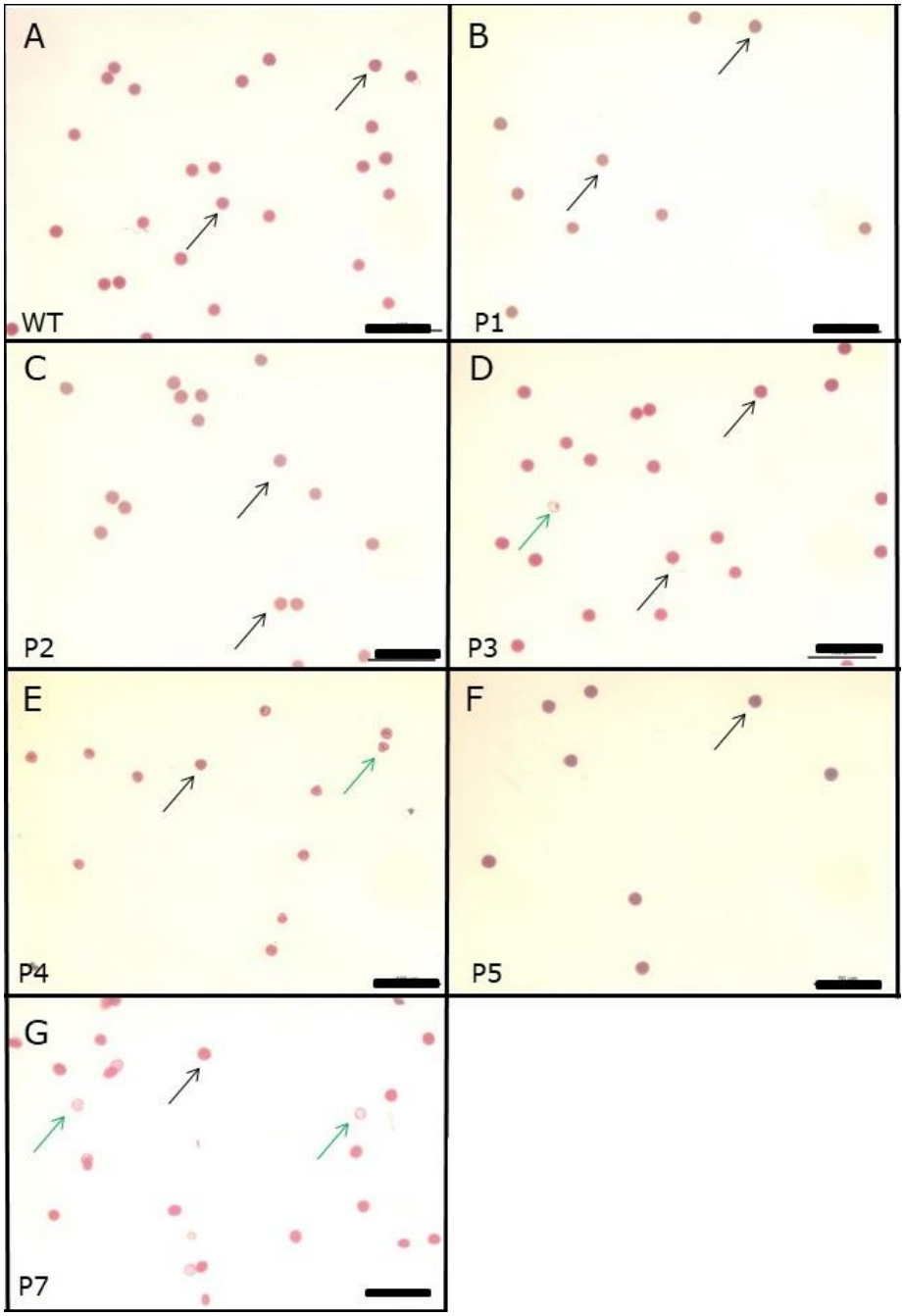


Figure 4-35. Pollen viability assessment using Alexander's staining of anthers from closed buds of seven weeks old plants where one hundred pollen grains were counted from every independent plant. All RNAi plants showed different level of pollen viability. Two out of six transformants showed high level of severity with pollen lethality of more than 60%. Black arrows indicate viable pollen grains which are stained red while non-viable pollens are indicated by the green arrow. A: WT-Col control; B-G: Pollen from pK7GWIWG2(I): AT2G23945RNAi lines. Bar: 50 µm.

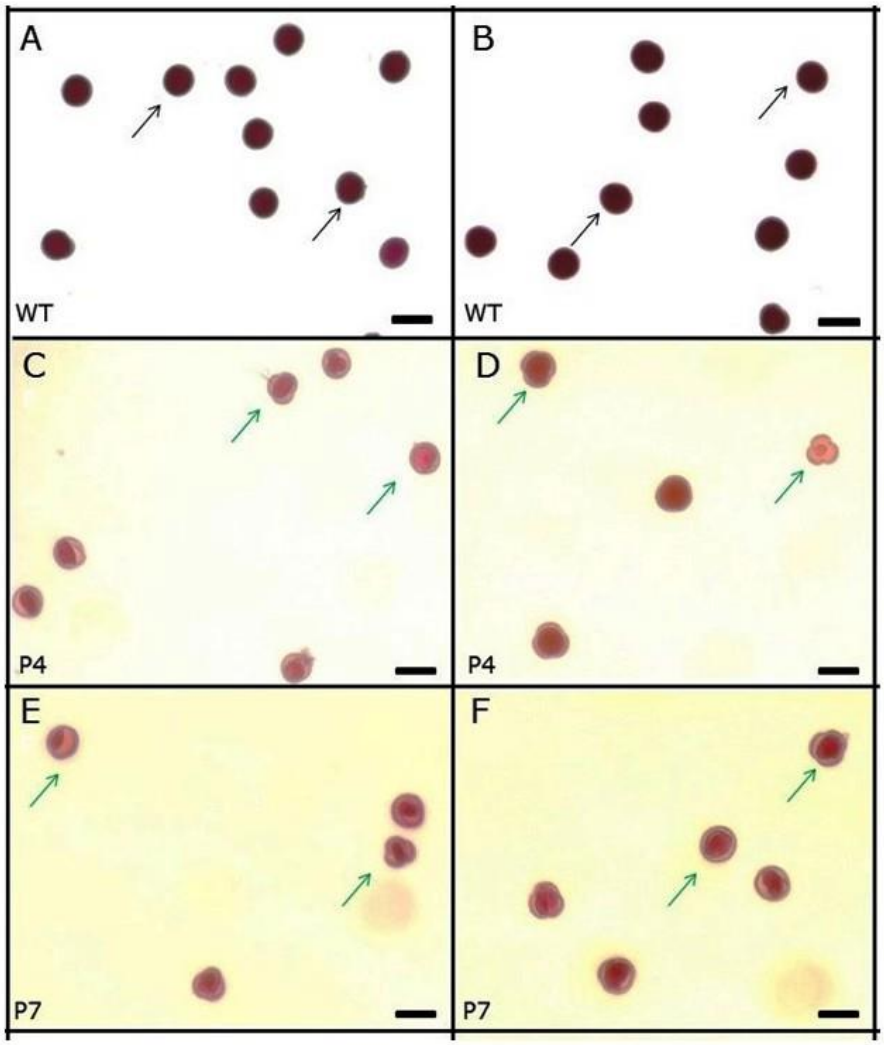


Figure 4-36. Alexander's staining of pollen from the two different plants from the plants that showed highest level of severity of pollen development (C-D: P4; E-F: P7) as compared to the WT-Col (A and B) where percentage of abnormal pollens outnumbered normal pollens. All images are at the same magnification (40x). Black arrows indicate viable pollen grains which are stained red while non-viable pollens are indicated by the green arrow. Pollens from the plant harbouring pK7GWIWG2(I): AT2G23945RNAi. Bar: 10 µm

Based on the statistical analysis (See Section 2.1.22), mean silique length was reduced for all the AT2G23945RNAi lines, but when compared to the WT-Col the P4 and P7 lines differed significantly compared to WT-Col (Figure 4-37). Reduction of total number

siliques has been observed on P4 main stem, while the total number of siliques on P7 main stem was slightly increased compared to the WT-Col.

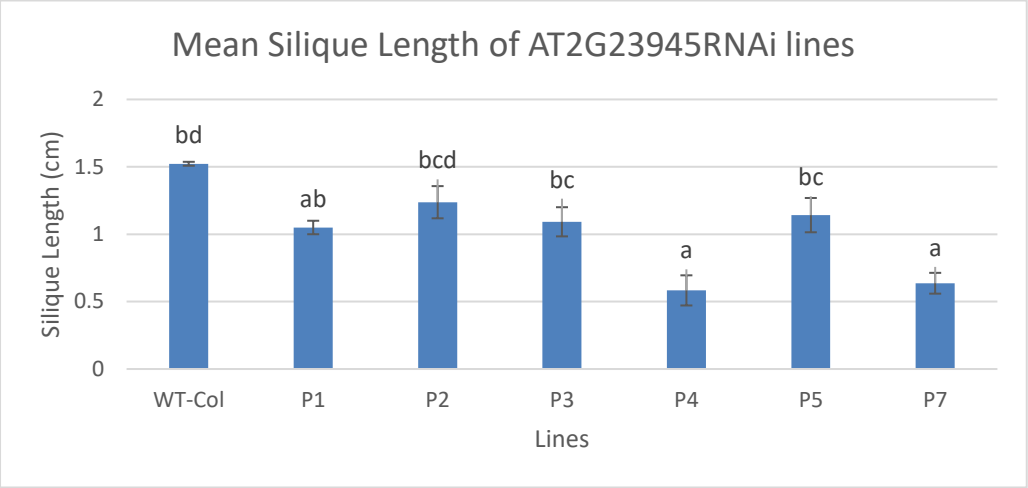


Figure 4-37. Mean silique length (cm), indicating fertility between six AT2G23945RNAi plants (P1, P2, P3, P4, P5 and P7) and the WT-Col control after 8 weeks. Mean silique length was obtained for all siliques, sterile and non-sterile by dividing total length by total number of siliques on the main stem. Identical letter signifies no significant difference where $P<0.05$; ANOVA, Fisher's Least Significant Difference test. Error bars represent standard error.

Table 4-3: Summary of the analysis of fertility in the independent AT2G23945RNAi (T_1) transgenic plants. The total number of fully developed siliques, percentage of sterility on main stem, and pollen viability were determined in six T_1 AT2G23945RNAi lines and WT-Col.

Samples (Independent plants)	**Total no. of siliques**	**Percentage of sterility on main stem (%)**	**Pollen Viability (%)**
Wild-type Col	22	4.5	100
P1	2	0	>95
P2	8	12.5	>95
P3	13	15.4	>95
P4	6	50	~10
P5	12	16.7	>95
P7	28	50	~40

Differences in seed set and fertility were observed between the various T_1 RNAi lines however due to the limited numbers of flowers and time constraints, correlation between phenotypes and changing level of endogenous mRNA in RNAi mutants could not be analysed for T_1 RNAi lines. In brief, Line P4 is the most severely affected line in terms of sterility rate, length of the silique, reduced number of buds' set and pollen viability.

Analysis of the expression profile and sterility effects of T_2 RNAi lines have been continued by Dr Alison Ferguson. However, unfortunately two trays of plants underwent water deficient (P1, P2, P3, P4, P5 and P7), but one tray was well-watered (P6). Thus, it is necessary to prepare and analyse these data separately by considering relevant WT-Col in the same tray.

The general expression observed was that RNAi resulted in gene reduction, however 7 out of 21 mutants have enhanced expression of the gene, indicating unusual pattern of expression in some of the T_2 generation (Figure 4-38, Table 4-4). For well-watered mutant plants (RNAi P6 24-26). Unusual expression level were recorded where the expression level were higher compared to the WT-Col 11. Expression levels were increased by approximately 58%, 30% and 68% for RNAi P6 24, RNAi P6 25 and RNAi P6 26 respectively, compared to the WT-Col (Figure 4-38, Table 4-4). There was no correlation observed between level of expression and sterility effects, the sterility percentage for all three plants were higher compared to the WT-Col 27. No significant difference was observed

for mean of silique length for both with and without stress siliques (Figure 4-39, Figure 4-40 and Table 4-4).

For T_2 generation, RNAi P3 7-9 plants under stress condition showed a strong sterility effect. Expression reductions of 81%, 61.1%, 73.7% were recorded for RNAi P3 7, RNAi P3 8, and RNAi P3 9 respectively (Figure 4-38, Table 4-4). Mean silique length was strongly reduced for RNAi P3 9 where it recorded 100% sterility and the shortest siliques that have been recorded, both with and without stress. Mean of silique length for P3 9 was significantly different compared to WT-Col 10, both with and without stressed silique (Figure 4-39, Figure 4-40 and Table 4-4).

T_1 generation of P4 line showed the strongest sterility effects, but with no expression analysis, it is not possible to link the sterility effects with expression pattern. Three plants were generated for T_2 generation under stress condition, RNAi P4 12-14 showed a strong reduction of gene expression, reduction of 87%, 79% and 85% for RNAi P4 12, RNAi P4 13 and RNAi P4 14 respectively (Figure 4-38, Table 4-4).

For analysis of main stem with stressed silique, the WT-Col 10 has a strong sterility effect with 50% of the siliques on main stem sterile. Fertility levels for the three mutants were lower compared to the WT-Col 10, with 30%, 5% and 25% for RNAi P4 12, RNAi P4 13, and RNAi P4 14 respectively (Figure 4-39, Table 4-4). For mean of silique length without stressed siliques, RNAi P4 12 and RNAi P4 13 showed significantly different length while RNAi P4 14 was not significantly

different compared to WT-Col 11. If removal of stress siliques was considered, RNAi P4 12 showed slightly higher 9% sterility rate compared to 0% sterility rate for WT-Col 10 (Figure 4-40, Table 4-4).

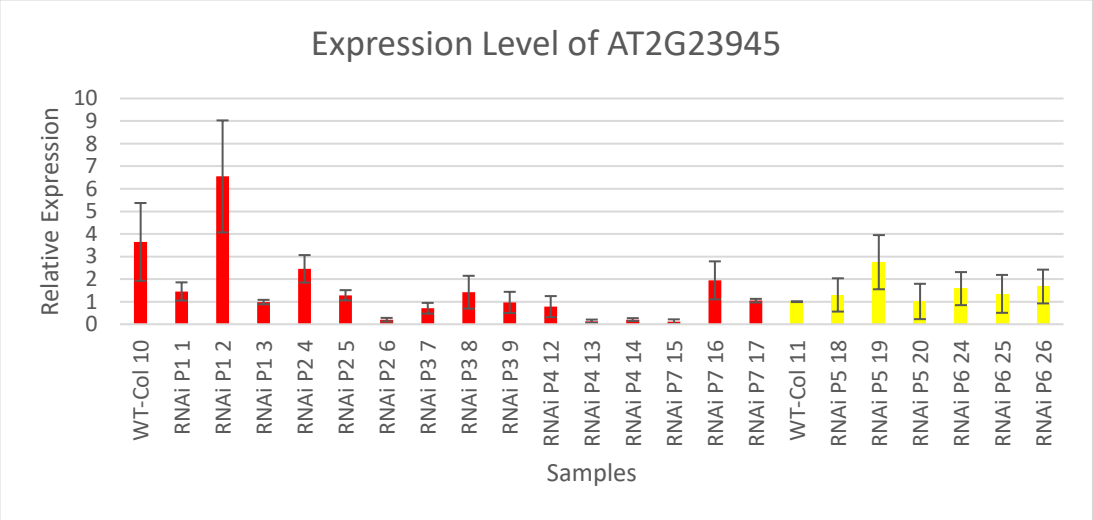


Figure 4-38. qRT-PCR analysis of AT2G23945 in inflorescence from 21 AT2G23945RNAi mutant plants. Bars represents a mean value (n=3) of the AT2G23945 relative to the Housekeeping *PP2A3* gene. RNAi P1 1 – RNAi P6 26: Twenty-one different plants of AT2G23945RNAi mutant plants. Mutant Plants were compared with the relevant Col-WT. Red bar: Mutant plants under water stress condition (RNAi P1 1-3, RNAI P2 4-6, RNAi P3 7-9, RNAi P4 12-14, RNAi P7 15-17) were compared with WT-Col 10; Yellow bar; Mutant plants under normal condition (RNAi P5 18-20, RNAi P 6 24-26) were compared with WT-Col 11. Error bars represent standard deviation.

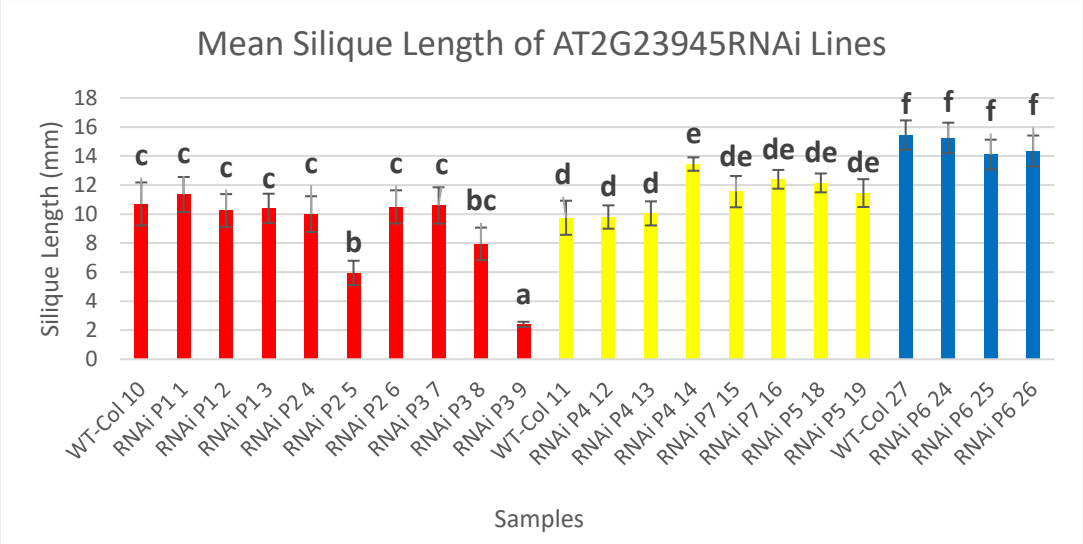
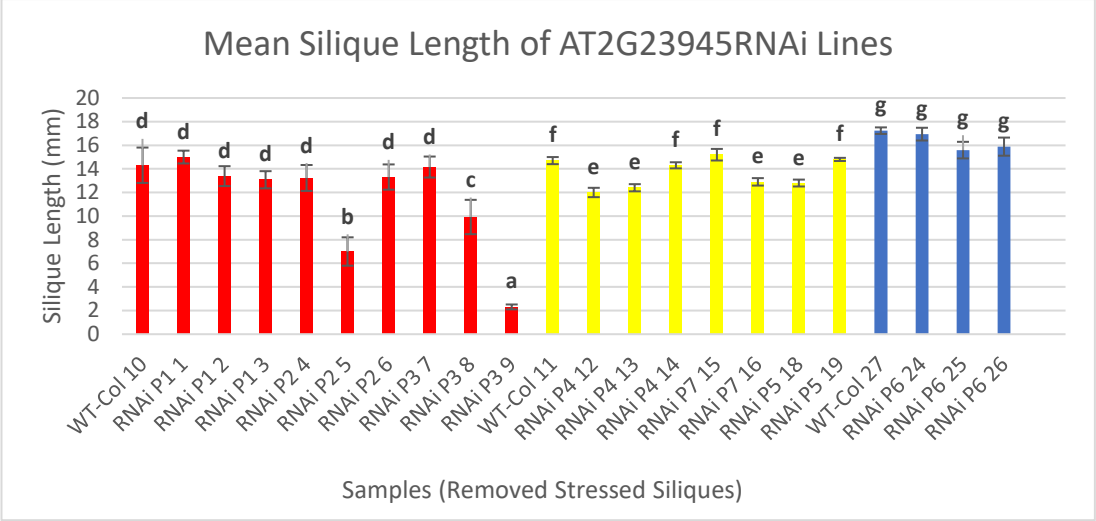


Figure 4-39. Mean silique length (mm), indicating fertility, between nineteen plants and the WT-Col controls. Mean silique length was obtained for all siliques, sterile and non-sterile by dividing total length by total number of siliques on the main stem. Mutant Plants were compared with the relevant Col-WT. Red bar: Mutant plants (RNAi P1 1-3, RNAI P2 4-6, RNAi P3 7-9 were compared with WT-Col 10; Yellow bar; Mutant plants (RNAi P4 12-14, RNAi P5 18-19, RNAi P7 15-16) were compared with WT-Col 11; Blue bar; Mutant plants (RNAi P6 24-26) were compared with WT-Col 27. Identical letter within the same colour group signifies no significant difference where $P<0.05$; ANOVA, Fisher's Least Significant Difference test. Error bars represent standard error.



4-40. Mean silique length without stressed siliques (mm), indicating fertility, between the nineteen plants and the WT-Col controls. Mean silique length was obtained for all siliques, sterile and non-sterile by dividing total length by total number of siliques on the main stem. Plants were compared with the relevant Col-WT. Red bar: Mutant plants (RNAi P1 1-3, RNAI P2 4-6, RNAi P3 7-9 were compared with the relevant WT-Col 10; Yellow bar; Mutant plants (RNAi P4 12-14, RNAi P5 18-19, RNAi P7 15-16) were compared with WT-Col 11; Blue bar; Mutant plants (RNAi P6 24-26) were compared with WT-Col 27. Identical letter within the same colour group signifies no significant difference where $P<0.05$; ANOVA, Fisher's Least Significant Difference test. Error bars represent standard error.

Table 4-4. Summary of the analysis of fertility in the AT2G23945RNAi (T_2) lines. Mean silique length (with and without stressed siliques), percentage of sterility on main stem (with and without stressed siliques), were determined alongside transcript level changes (%) in nineteen AT2G23945RNAi lines. Means followed by the same alphabet are not significantly different based on ANOVA, Fisher's Least Significant Difference test ($p < 0.05$). S.E: Standard error.↑ increased expression; ↓ decreased expression.

	Samples	Mean silique length (mm) ± S.E	Sterility on main stem (%)	Mean silique length (removed stressed silique) (mm) ± S.E	Sterility on main stem (removed stressed silique) (%)	Relative change (%)
Early Water Stressed Condition	WT-Col 10	10.7±1.48c	50	14.31±1.5d	23.1	0
	RNAi P1 1	11.35±1.21c	40	15±0.54d	7.7	60.3% ↓
	RNAi P1 2	10.25±1.13c	40	13.38±0.84d	7.7	79.5% ↑
	RNAi P1 3	10.4±1c	45	13.08±0.73d	13.4	-73.2% ↓
	RNAi P2 4	10±1.23c	50	13.23±1.1d	23.1	-32.9% ↓
	RNAi P2 5	5.95±0.84b	85	7±1.2b	76.9	-64.9% ↓
	RNAi P2 6	10.5±1.14c	45	13.308±1.07d	15.4	-95% ↓
	RNAi P3 7	10.6±1.25c	40	14.15±0.89d	7.7	-81% ↓
	RNAi P3 8	7.95±1.12bc	65	9.92±1.45c	46.2	-61.1% ↓
	RNAi P3 9	2.4±0.18a	100	2.31±0.21a	100	-73.7% ↓
	WT-Col 11	9.75±1.17d	50	14.7±0.3f	0	0
	RNAi P4 12	9.8±0.8d	30	12±0.4e	9	-22% ↓
	RNAi P4 13	10.05±0.83d	25	12.4±0.31e	0	-87% ↓
	RNAi P4 14	13.45±0.46e	5	14.3±0.26f	0	-79% ↓
	RNAi P7 15	11.55±1.08de	30	15.2±0.5f	0	-85% ↓
	RNAi P7 16	12.4±0.65de	15	12.9±0.31e	0	94% ↑
	RNAi P5 18	12.15±0.65de	10	12.8±0.29e	0	29% ↑
	RNAi P5 19	11.45±0.96de	35	14.8±0.13f	0	175% ↑
Normal Condition	WT-Col 27	15.45±1f	15	17.24±0.28g	0	0
	RNAi P6 24	15.25±1.05f	20	16.94±0.54g	5.9	58% ↑
	RNAi P6 25	14.1±1.03f	20	15.59±0.70g	5.9	30% ↑
	RNAi P6 26	14.35±1.06f	20	15.88±0.77g	5.9	68% ↑

4.3 Discussion

Extensive research of the putative downstream targets of *AMS* and *MS1* transcription factors has shown that mutants of these target genes have impaired pollen development that cause male sterility. Manipulation of the downstream targets of *AMS* such as *AHL16* (Alves-Ferreira *et al.*, 2007), *EXL4* and *GRP17* (Updegraff *et al.,* 2009), WBC27 (Xu *et al.*, 2010; Dou *et al.*, 2011), *TEK* (Lou *et al.*, 2014), *EXL5* and *EXL6* (Xu *et al.*, 2014) has severe impacts on fertility. Also, a downstream target of the *MS1* transcription factor, (*flp1*) gene has been reported by Ariizumi *et al.*, (2003) To arrest during pollen germination leading to male sterility.

AT2G23945 functional analysis was conducted through manipulation of expression levels using SALK T-DNA insertion lines, overexpression (See section 3.2.5.1), and RNA interference lines. Two allelic forms of AT2G23945 were isolated and these two SALK T-DNA insertions were analysed, which carried insertions in the promoter and exon regions, which led to increased AT2G23945 expression. Different severity of phenotypes was observed with the altered levels of expression (Table 4-5). SALK 135523, where the insertion is in the promoter region, resulted in more severe phenotypes with the expression increased 3 to 7-fold as shown by these SALK T-DNA insertion mutant plants. Furthermore, SALK 134919 plants which carried an insertion in the exon region displayed mild phenotypes even though the mRNA accumulation was slightly higher compared to the SALK 135523 (Table 4-5). The sterility on the main stem ranged from 61.9% to 64.0%, and 25.0% to 28.5% for SALK 135523 and SALK 134919 respectively. The

general expected impact from T-DNA insertion lines is gene silencing, however these results indicate that both mutants have enhanced expression of the gene. Integration of T-DNA into genomic regions can induce mutations and prevent normal gene expression (Koncz *et al.*, 1992; Alonso *et al.*, 2003). Furthermore, Ülker *et al.*, (2008) speculated that T-DNA insertions were also able to increase the endogenous gene or caused the gain-of-function, depending on integration region. Hence, the increased expression observed in the SALK insertion lines is likely to be a consequence of disruption of the promoter, or an enhancer element, by the T-DNA.

Importantly, the insertion site of SALK 135523 is located at -83bp upstream of TSS, and it consists of no DNA-binding elements that are vital for initiation of transcription processes (See Chapter 5). Plus, the nearest DNA-motif binding site located at -283bp upstream of TSS for *MYB80* and *MYB99*. It is possible that the phenotypic change has no correlation with the failure or alteration of the transcription initiation process and expression of AT2G23945. In contrast with the other results, for example T-DNA insertion in promoter region of *gl1* mutant that is speculated to disrupt the enhancer element which is required for trichomes development along with gene silencing (Oppenheimer *et al.*, 1991). In brief, SALK 135523 overexpression mutants have the most severe phenotype of the lines analysed. Delayed tapetum breakdown was observed in anther trans-section study in SALK 135523 (Figure 4-8), which may be the cause of the shrunken (Figure 4-7), and non-viable pollen (Figure 4-4). These impaired fertility SALK plants can be distinguished from the wild type by the lacked of fertile siliques.

For AT2G23945RNAi plants, different degree of phenotypes were observed, with some plants showing a strong phenotype with severe sterility in the T_1 generation. The sterility on the main stem reached 50% and five out of six AT2G23945RNAi plant appeared phenotypically different with reduced numbers of siliques compared to the WT-Col. In addition, the RNAi plants had extended (10-weeks) lifespan with longer flowering time, which is normal for some male-sterile mutants (Yang *et al.*, 2007). Unfortunately, due to the time constraints and limited numbers of flowers available in the T_1 generation, the expression analysis for T_1 generation was not conducted. The continued analysis of the T_2 is summarized in Table 4-4.

Due to unexpected growth problems, two sets of data were analysed: 1) stressed, water deficient during early development (RNAi P1 1 - RNAi P5 19); and 2) Normal condition, well-watered (RNAi P6 24-26). Analysis of T_2 generations were disrupted by the water-deficiency during early development. Although there were changes in expression in these RNAi lines, overall, there was no clear correlation between the changes of expression pattern and the sterility effects, however these data need to be repeated.

It is still unclear whether the early stress condition caused the broad variability of gene expression pattern. Taken together, limited data from normal condition development process made it difficult to make a conclusion. Thus, it is necessary to have further experiments with more replicates to clarify the analysis.

The water deficite appeared to have a significant impact on pollen development. It has been reported before that the water stress may induce or inhibit tapetum PCD where it triggers signals that interfere the tapetum PCD process (Parish *et al.*, 2012).

Inconsistent sterility rate and siliques length from the same lines have been observed in some of the T_2 generation (Table 4-4). Tapetal development is sensitive to water stress and interuption of the tapetum PCD process or pollen developmentally regulated networks (Jin *et al.*, 2013) causes severe effects. Impact of water deficient in rice and wheat, where the stress disturbed the sugar transportation process in floral organs, reducing starch accumulation in the anther leading to tapetal starvation (Storme and Geelen, 2014).

Expression levels of the transgenes were highly variable even within the same parental lines, which may be occurring due to the segregation in the T_2 generation (Table 4-4). It is frequently seen in some transgenic plants after several generations change in phenotypes due to the segregation during reproductive cells development (Gelvin, 2003). The T_2 generation appeared different from the parental plants, but the phenotype did not pass from one generation to the next. This high variable may be due to the weak and unstable expression of the gene during the T_1 generation (Feldmann et al., 1997). Stable inheritance and continued expression for the subsequent generation also depends on location of the insertion in chromatin (discussed in Section 3.3). In addition, no information on copy number of constructs integrated in the genome is available.

Altered tapetum development and PCD significantly impacts on pollen development (See Section 1.4). Taylor *et al.*, (1998) reported tapetal specific genes involved in tapetum degeneration, *Male Sterility7 (MS7)* and *Male Sterility8 (MS8)*. Early tapetum PCD in both mutants occurred due to tapetal cytoplasm disintegration during the microspore stage. Furthermore, enhanced tapetum PCD in *GUS-negative (GNE)* mutants, *gne1* and *gne2* mutants have been observed due to the deficiencies of middle layer and tapetal cells formation in the early anther development (Sorensen *et al.*, 2002). The *fat tapetum* mutant exhibited abnormally enlarge tapetum and middle layer, where it failed to degenerate (Sanders *et al.*, 1999).

Table 4-5. Summary of fertility analysis in T_1 plants of SALK 135523, SALK 134919, AT2G23945OE and AT2G23945RNAi. Effects on number of siliques on main stem, effects on sterility on main stem were determined alongside transcript level changes in four different AT2G23945 gene manipulation plants.

Sample	**Transcript Level Changes**	**Effects on sterility on main stem**	**Percentage of sterility on main stem (%)**
SALK 135523 (T_1)	3 to 7-fold	Reduced	61.9 – 64.0
SALK 134919 (T_1)	13 to 20-fold	Reduced	25.0 - 28.5
AT2G23945OE (T_1)	43 to 125-fold	Reduced	12.5 – 22.2
AT2g23945RNAi (T_1)	Not available	Reduced	0 - 50
AT2g23945RNAi (T_2)	Varied	Varied	0-100

The overexpression of the plants through overexpression vector exhibited milder sterility phenotype compared to the SALK 135523 and SALK 134919 insertion lines (See section 3.2.5.1). To confirm whether the AT2G23945 gene is expressed ectopically in other tissues, there is a need to conduct the analysis of expression level in different tissues on transgenic plants generated by overexpression vector.

The overexpression lines may have activated or repressed other genes in other tissues which are involved in development programs. Shi *et al.*, (2015) suggested that overexpression of one gene also caused certain gene to be downregulated due to the altered

equilibrium in signal transduction complex. Overexpression of *GLABROUS1 (GL1)* displayed altered trichome development due to the changed of signal transduction during gene enhancement (Oppenheimer *et al.*, 1991). On the other hand, overexpression of *MS1* gene also showed that the tapetum transcript was ectopically enhanced in leaf tissues, which may cause the stunted growth of the overexpression plants due to unusual signal development (Yang *et al.*, 2007). Additionally, in the *AtTTP*-OE line, *CalS5* and *A6* needed for normal callose dissolution have been downregulated which may cause defective pollen development (Shi *et al.*, 2015).

Another key point is the expression threshold level that determines the function of the gene and is associated with defective pollen development (Zhang *et al.*, 2014). In *CEP1*, mRNA expression levels ranging from less than 50% downregulation, to up to 200% upregulation of the gene caused partial reduction of fertility. Moreover, overexpression of the *AGAMOUS (AG)* gene did not show any phenotypic change because it may not reach the minimum threshold of expression, whilst the silencing of the gene caused abnormal anther and filament development (Liu *et al.*, 2009).

Several proteases associated with male sterility have been reported in *Arabidopsis* and rice. *Arabidopsis UNDEAD* encodes for an *A1* aspartic protease which displayed enhanced tapetum degradation due to abnormal vacuolated tapetum which caused aborted microspores (Phan *et al.*, 2011). It is also reported that rice aspartic protease *osap65* mutants exhibited impaired male gametophyte and defects in pollen tube development (Huang *et al.*, 2013a). Whereas, Zhang *et al.*, (2014) stated that *CEP1* was involved directly with cell

degradation in the tapetum. The mutants displaying late tapetum degradation while enhanced expression exhibited early tapetum PCD.

Cells naturally have pro and anti-PCD components to coordinate development through homeostatic equilibrium (Ge *et al.*, 2005) and understanding plant PCD networks mediated by the various proteases remains challenging. Multiple function of proteases in PCD process may also occur because of different substrates in different tissues (Gao *et al.*, 2017).

Increased aspartyl-protease *CDR1* protein in leaves during pathogen attacks triggered the defense mechanism by activating the peptide signal system related to defense. Interestingly, rice *CDR1* orthologue *OsCDR1* shared similar characteristics with *AtCDR1* (Prasad *et al.*, 2009). Huang *et al.*, 2013 suggested that *OsAP65* aspartyl-protease hydrolyze specific substrates during PCD process, and the degraded substrates processed by cells to produce substance needed for pollen growth and pollen tube development. Thus, silencing of *OsAP65* caused the disruption of PCD and lack of essential substances for pollen growth. Additionally, early tapetum PCD in *CP51*-RNAi plants showed that *CP51* involved in proteolysis of the substrate and mediates tapetum PCD (Yang *et al.*, 2014). *In vitro* analysis showed that *A36* and *A39* displayed a proteolytic activity with pollen from both mutants displaying apoptosis-like PCD activity. *A36* and *A39* genes are speculated to serve as an activator of survival signal, or as inhibitor of the PCD process during natural degradation (Gao *et al.*, 2017).

In wild-type, *PCS1* is needed for cellular function and metabolic activity. Inhibition of anther dehiscence occurred after *PCS1* was ectopically expressed in the anther by activating an anti-PCD process, where it advocates cell survival during embryogenesis and gametogenesis process. *PCS1* may therefore activate certain peptide or unknown survival factors or degrading pro-PCD peptide in cells (Ge *et al.*, 2005b).

It may be too early to speculate on the main function of AT2G23945 on the tapetum PCD, however the tapetum appeared to not breakdown in the SALK lines and persist suggesting impaired PCD processes. A number of other proteases, *PCS1* (Ge *et al.,* 2005), *UNDEAD* (Phan *et al.,* 2011), *A36* and *A39* (Gao *et al.*, 2017) are negative regulators of PCD. However only limited information of aspartyl-protease substrate involved in plant apoptosis like-PCD has been reported (Li *et al.*, 2016). Neverthekess the lack of apparent tapetum breakdown in the overexpression lines suggests a possible repressor-like role for AT2g23945 in tapetum PCD

CHAPTER 5: TRANSCRIPTIONAL AND TRANSLATIONAL FUSION OF AT2G23945 GENE

5.1 Introduction

Understanding the regulation of gene expression is an important feature of characterising gene function and network analysis, nevertheless this can be challenging and requires practical validation of potential regulatory sequences. A gene promoter functions to promote transcription in a spatial and temporal manner. The promoter region is located upstream of the transcription start codon of the gene and is a site for the pre-initiation complex for gene regulation (Smale, 2001). Prior to protein synthesis a number of steps are required pre-initiation of transcription, transcription, post-transcription, pre-translation, translation and post- translational processes (Molina and Grotewold, 2005).

In general, a eukaryote promoter region consists of a TATA box, which provides a binding site for specific binding protein (TBP) and TBP-associated factors (TAFs) to form a pre-initiation complex. Subsequently, RNA polymerase II is responsible for the highly regulated transcription process (Molina and Grotewold, 2005). The TATA box is rich in T and A nucleotides and is generally located 25-35bp upstream of the start codon. Transcriptional regulation is dependent upon the macromolecule complex that leads to RNA polymerase II attaching to the transcription initiation site (Smale, 2001). Specific transcription factors control the transcription process by forming a complex with RNA polymerase II. This protein can act

as an activator or repressor, controlling the spatial and temporal expression of the gene. Transcription factors recognize specific DNA-binding domains and thus control expression in a regulated manner (Smale and Kadonaga, 2003).

The transcription process commences in the nucleus with the synthesis of mRNA. The mRNA is then transported to the ribosomes, which are responsible for translational. Protein dynamics involve complex protein secretory pathways (Singh, 1998). In eukaryotic cells, the endomembrane system is an interlinked system that connect multiple organelles in the cytoplasm of cells. The mature peptide travels to Rough Endoplasmic Reticulum (RER) where proteins are synthesized. Subsequently, Golgi apparatus process the protein before it is transported to its destined compartment. In plant cells, a protein can be transported into several different locations to conduct its biological function. The amount of expression can be estimated according to the level of transcript expression, however this does not always correlate with protein level (Held *et al.*, 2008). The specificity of protein localization is also a key factor in the control and regulation of biological functions (Li *et al.,* 2006).

Reporter gene fusions, that can be visualised by various approaches that have been used to monitor gene and protein expression patterns. In general, two types of gene fusions are used, transcriptional and translational fusions. Transcriptional fusions consist of the putative gene specific promoter fused to a reporter. No localisation or gene sequences are included; therefore, expression is based upon the regulation by the promoter and the transcription only. In contrast, translational fusions include both

promoter and gene sequences, and therefore reflect regulation by any gene sequences e.g. intron sequences, combined with post-transcriptional regulation and indicate the expression patterns of the protein of interest rather than the transcript. Fluorescence tagged proteins such as the Green Fluorescent Protein (GFP) and Yellow Fluorescent Protein (YFP), and β-glucuronidase (GUS) proteins are frequently used as reporter to enable expression visualization (Vitha, 2012).

The use of the GUS reporter gene was first reported by Jefferson (1987) and remains one of the main methods for gene expression analysis. β-glucuronidase (GUS) encoded by the *uidA* gene of *E. coli* has many advantages, including being a robust enzyme, simple assays, highly efficient detections, easy quantification, high sensitivity for abundance and moderate expression of localization of the gene in specific time and tissues (Jefferson, 1987). Low levels of GUS activity in plants also reduces the possibilities of false positive during analysis (Jefferson, 1989; Vitha, 2012). However, dynamic temporal and spatial analysis of GUS expression is difficult in living tissues. There is also the problem that GUS signal may leak into adjacent cells from the site of synthesis (Miki and McHugh, 2004).

To overcome these problems and enable live imaging, fluorescence tags have been used. GFP was first reported by Prasher *et al.*, in 1992. It was isolated from crystal jellyfish (*Aequorea Victoria*), the protein is 238 amino acids in length and ~27 kDa (Prasher *et al.*, 1992). Since then large numbers of variants that produce different fluorescent signals have been described, a frequently used one is the yellow fluorescent protein (YFP) (Ormo *et al.*, 1996). These

reporters allow live imaging of transcripts and proteins, enabling the visualization of protein localisation and transport within different organelles and between cells. Three-dimensional (3D) visualisation under laser confocal microscope can be used since the GFP is stable under high temperatures (Haseloff, 1998). In addition, they can be used to investigate the response of the protein to external/environmental stress, internal stimuli, developmental stimuli and different mutant backgrounds (Brandizzi *et al.,* 2002). These fluorescence proteins can also be used to investigate the interactions of the gene with other regulatory elements, protein-nucleic acid interactions and protein-protein interactions. This can also be analysed in ectopic systems, by transient expression of the gene construct, in other species and tissues, such as *N. benthamiana* leaves (Koo *et al.*, 2007).

In short, the aim for this chapter was to study the transcriptional and translational expression patterns of AT2G23945 gene.

5.2 Results

5.2.1 Transcriptional analysis of PGWB3:AT2G23945PRO:GUS (-437 bp, -980 bp, -1572 bp).

Transcriptional fusions of the upstream region of AT2G23945 were generated in the pGWB3 plasmid containing the GUS protein at the C-terminal end (Section 2.3.4). Three different lengths of upstream sequences were amplified using different primer combinations to identify the region required for expression. The promoter fragments are identified as short (-437bp promoter fragment), medium(-980bp promoter fragment) and long promoter (-1572bp promoter fragment), all promoter regions were located upstream of transcription start site (TSS). One putative binding site for *MYB80* located at -193bp to -208bp upstream of TSS, and two putative binding site for *MYB99* in positive and negative strand at -195bp to -208bp and -1258bp to -1271bp respectively. The location of the putative binding site for *MYB80* and *MYB99* are shown in Figure 5-1.

The putative promoter fragments were cloned into the Gateway entry vector pCR8/GW/TOPO and confirmed by colony PCR screening using the forward specific primer and pCR8/GW/TOPO 2.1 reverse primer. The schematic map of pCR8/GW/TOPO:AT2G23945Pro entry clone with three different promoter fragments is shown in Figure 5-1. Amplifications were of the expected sizes (Figure 5-2; 513 bp, 1056 bp and 1648 bp) for short, medium and long promoter fragments respectively. The successful amplifications were conducted using primers 2G23945PRO F4 and pCR8 GW 2.1 R,

2G23945PRo F5 and pCR8 GW 2.1 R, M13F (-20) and 2G23945Pro R1 for short, medium and long promoter fragments respectively. pCR8/GW/TOPO clone 5 (C5) for the short promoter sequence, clone 14 (C14) for the medium promoter, clone 18 (C18) for long promoter fragments were confirmed by sequencing and selected for further analysis.

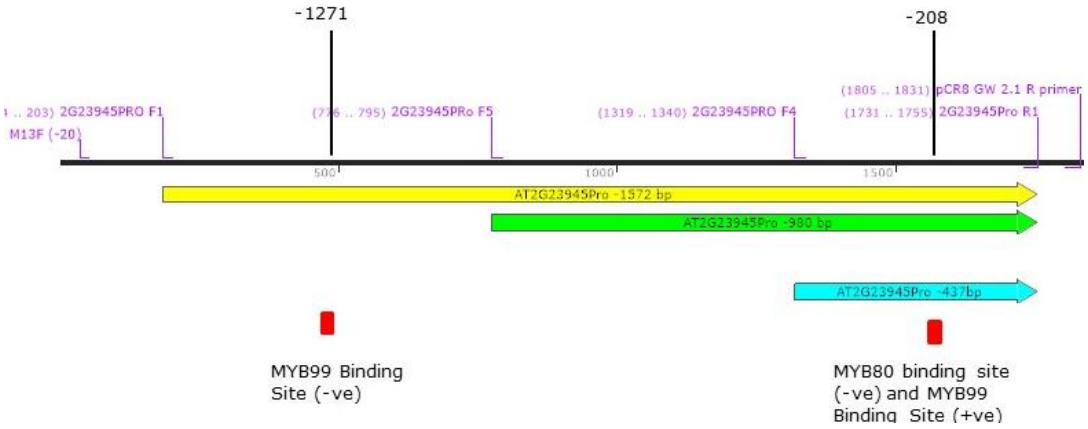


Figure 5-1. The schematic map of pCR8/GW/TOPO:AT2G23945Pro entry clone with three different promoter fragments. The yellow arrow indicates the AT2G23945 long promoter fragment (1572bp), green arrow indicates the AT2G23945 medium promoter fragment (980bp), and the light blue arrow indicates the AT2G23945 short promoter fragment (437bp). The locations of the primers are shown on the map. The presence of two putative binding sites for *MYB80* located at -193bp to -208bp upstream of the TSS, and *MYB99* in positive and negative strand at -195bp to -208bp and -1258bp to -1271bp respectively. Binding Site Prediction developed by Center for Bioinformatics, Peking University, Plant Transcriptional Regulatory Map. (http://plantregmap.cbi.pku.edu.cn/binding_site_prediction.php).

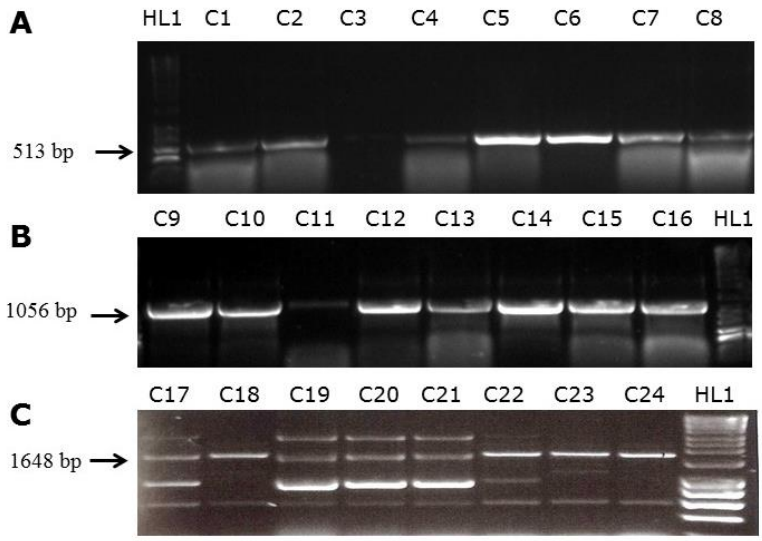


Figure 5-2. Colony PCR of insert and vector sequences from pCR8/GW/TOPO entry clones for the three different sizes of AT2G23945 promoter. Amplifications generated fragments of (A) 513 bp, from clones containing a short insertion fragment (B) 1056 bp, from clones containing a medium insertion fragment and (C) 980 bp, from clones containing long insertion fragment. C1-C8: Different colonies of the entry clone for short promoter fragments; C9-C16: different colonies of the entry clone for medium promoter fragments; C17-C24: different colonies of the entry clone for long promoter fragments, HL1: HyperLadder 1.

After confirmation by sequencing, constructs were transferred into Gateway binary vector pGWB3. The schematic map of pGWB3:AT2G23945PRO:GUS destination vector with three different promoter fragments is shown in Figure 5-3. Amplification of three different sizes of pGWB3:AT2G23945PRO:GUS constructs confirmed the clone identity (Figure 5-4). Amplifications generated fragments of (A) 648 bp for all constructs using 2G23945PRO F4and GUSpGWB3_138R. Sequencing of pGWB3:AT2G23945PRO:GUS clone 1 (C1), clone 3 (C3) and clone (C5) for short, medium and long promoter fragments respectively, were conducted to confirm the sequence identity. The pGWB3:AT2G23945PRO:GUS plasmids

were transformed into *Agrobacterium* using electroporation. After verification by PCR (Figure 5-5), GV3101 clone 1 (C1), clone 12 (C12) and clone 14 (C14) for short, medium and long promoter fragments respectively, were transformed into WT-Col (Section 2.1.11).

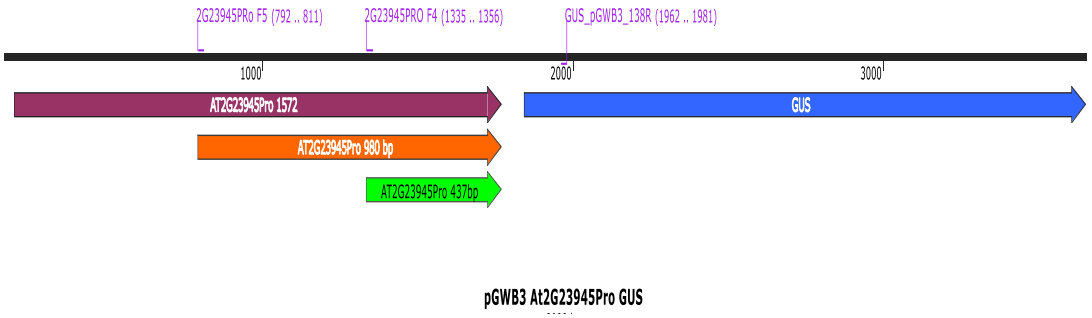


Figure 5-3. The schematic map of pGWB3:AT2G23945PRO:GUS destination vector with three different promoter fragments. The maroon arrow indicates the AT2G23945 long promoter fragment (AT2G23945pro-1572:GUS), orange arrow indicates the AT2G23945 medium promoter fragment (AT2G23945pro-980:GUS), the green arrow indicates the AT2G23945 short promoter fragment (AT2G23945pro-437:GUS), and blue arrow indicates the GUS protein tag. The locations of the primers are shown on the map.

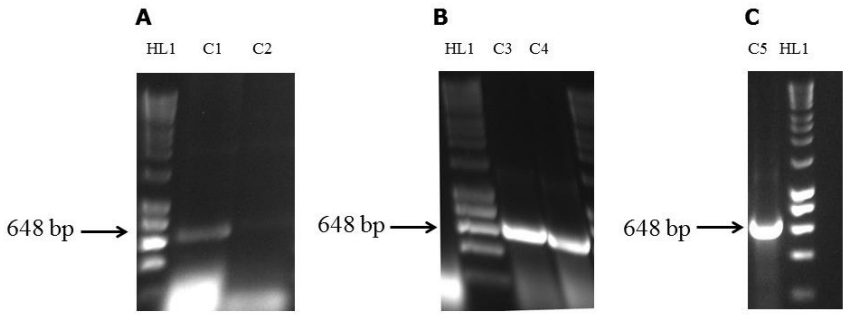


Figure 5-4. Colony PCR of insert and vector sequences from pGWB3:AT2G23945PRO:GUS destination vectors with three different sizes of the putative AT2G23945 promoter. Amplification of the destination vector with the right fragment size of ~648bp. The amplifications were conducted using the same promoter specific forward primer (2G23945PRO F4) coupled with the pGWB3 GUS reverse primer. pGWB3:AT2G23945PRO:GUS with similar size of promoter fragments; (A) short promoter, (B) medium promoter, (C) long promoter. C1-C5: Different entry clones, HL1: HyperLadder 1.

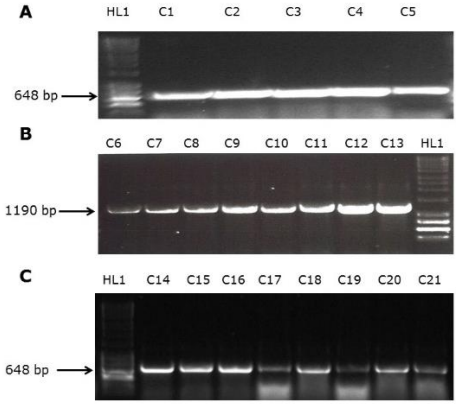


Figure 5-5. Colony PCR of insert and vector sequence from GV3101 containing pGWB3:AT2G23945PRO:GUS colonies. Amplification of pGWB3:AT2G23945PRO:GUS transcripts with the correct fragment size of 648bp (A and C) and 1190bp (B). For A and C the amplifications were conducted using a same set of promoter specific forward primer (2G23945PRO F4) coupled with pGWB3 GUS reverse primer. Meanwhile, 2G23945PRO F5 forward primer coupled with pGWB3 GUS reverse primer were used for (B). pGWB3:AT2G23945PRO:GUS with different sizes of promoter fragments; (A) short promoter, (B) medium promoter, (C) long promoter. C1-C21: Different destination vectors, HL1: HyperLadder 1.

The putatively transformed seeds (T_1) from all constructs were collected and screened using MS medium (Murashige and Skoog, 1962) containing Kanamycin (50 ng/µl). Putative transgenic plants were transferred to soil and screened by PCR to confirm the integration of the transgene. PCR analysis showed that three out of eight putative transgenic plants carried a segment of AT2G23945 short promoter construct with the size of 648 bp (Figure 5-6A), whilst other putative transgenic plants generated faint bands. The amplification was generated using the promoter specific forward primer (2G23945PRO F4) coupled with pGWB3 GUS reverse primer. PCR analysis showed that all putative transgenic plants which carried medium and long promoter construct, generated a segment of 1190 bp, indicating successful transformation (Figure 5-6B and 5-6C). The amplification was generated using the forward primer (2G23945PRO F5) coupled with pGWB3 GUS reverse primer.

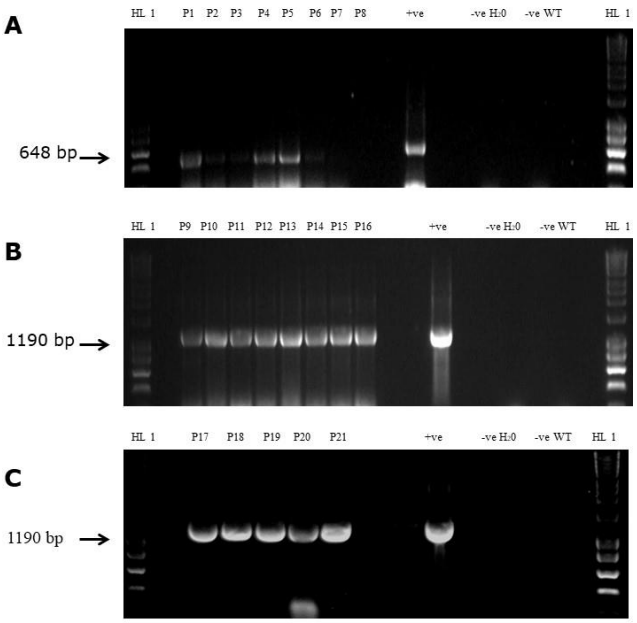


Figure 5-6. PCR amplification of the three different length promoter fusions of pGWB3:AT2G23945PRO:GUS transgene in *Arabidopsis* lines. Amplification of fragments with the correct size of 648bp (A) and 1190 bp (B and C), indicated integration of the promoter construct into the plant genome. For (A), promoter specific forward primer (2G23945PRO F4) coupled with pGWB3 GUS reverse primer was used for amplification. (B and C) the amplifications were performed using forward primer (2G23945PRO F5) coupled with pGWB3 GUS reverse primer. Different transgenic plants with (A): short promoter insertion (P1 to P8); (B): medium promoter insertion (P9 to P16); (C): long promoter insertion (P17 to P21); +Ve : pGWB3:AT2G23945PRO:GUS vector as positive control WT from each lines, -Ve H_2O: Water negative control, -Ve WT: WT-Col genomic DNA negative control, HL1: HyperLadder 1.

Three independent transgenic lines for each length of promoter fragment were investigated by GUS assays in reproductive and vegetative tissues. The GUS signal varied depending on the length of the promoter. Interestingly, vegetative GUS activity was visualised in the base of the silique during young silique development when driven by the short 437 bp promoter fragment;

and no GUS activity was detected in the wild type control plants at any developmental stage (Figure 5-7).

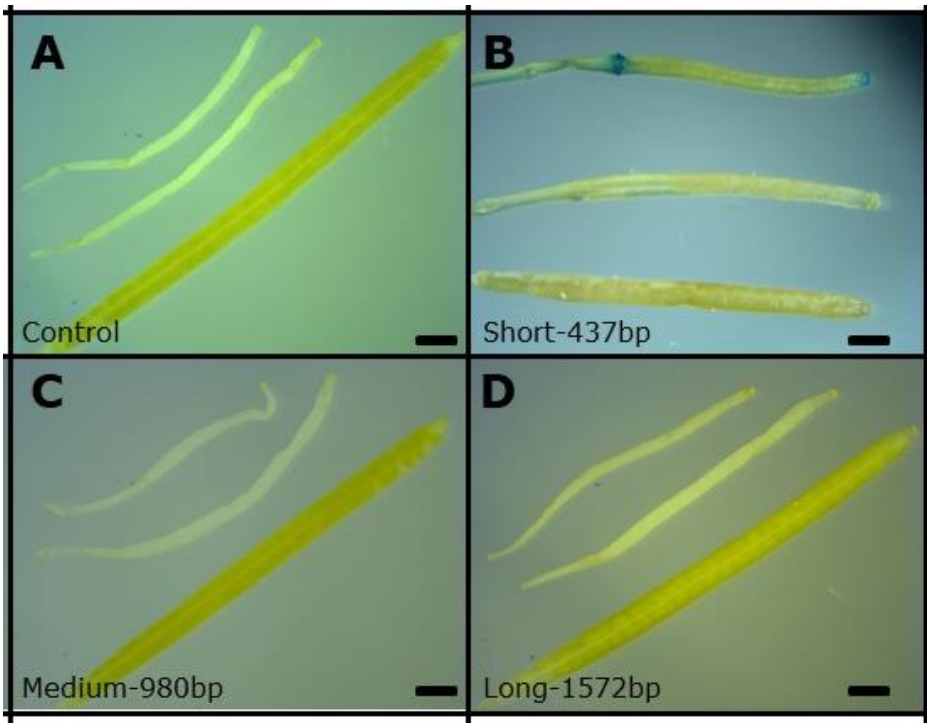


Figure 5-7. AT2G23945 GUS expression patterns were analysed using young and mature siliques from different plants harbouring different sizes of promoter. GUS activities was presence in the young silique base from plants containing short promoter fragment. Representative GUS staining for three biological replicates are shown, A: Siliques of WT-Col control; B, C and D: Siliques from the plants harbouring short promoter (AT2G23945pro-437:GUS); medium promoter (AT2G23945pro-980:GUS); and long promoter (AT2G23945pro-1572:GUS) respectively. Bar: 1 mm.

For plantlets analysis, three-week-old plantlets screened on MS medium supplemented with Kanamycin (50 ng/µl) were used for GUS staining where none of the plantlets showed any GUS expression. Moreover, none of the internodes observed any GUS expression (Figure 5-8).

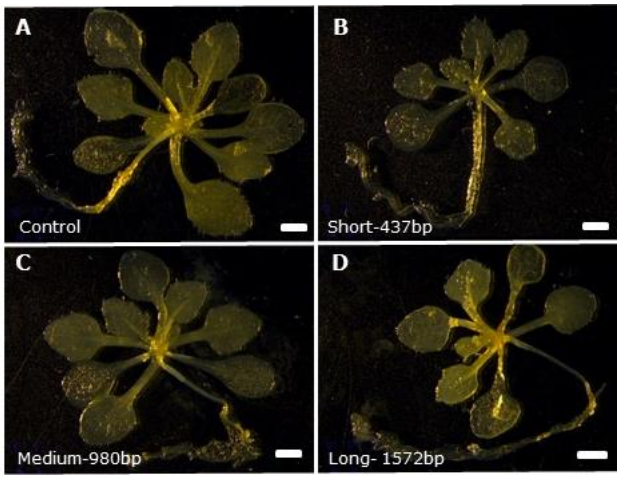


Figure 5-8. Plantlets-specific GUS expression patterns were analysed using 3-week-old transgenic plantlets screened using MS medium supplemented with Kanamycin (50 ng/µl). GUS activity was absence for all plantlets investigated. Representative GUS staining from three biological replicates are shown for each sample, A: WT-Col control; B, C and D: Plantlets harbouring the short (AT2G23945pro-437:GUS); medium (AT2G23945pro-980:GUS); and long promoter (AT2G23945pro-1572:GUS) respectively. Bar: 1 mm.

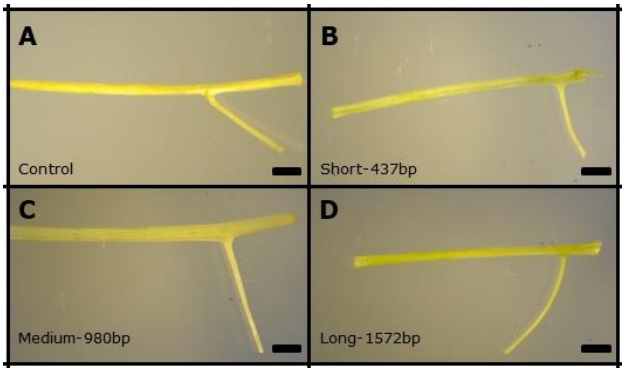


Figure 5-9. Internode-specific GUS expression patterns were analysed from different plants harbouring different sizes of promoter. No GUS staining was seen in internodes from wild-type plant and transgenic plants. Photos were taken using stems from mature transgenic plants. Representative of GUS staining from three biological replicates for every sample, A: Stem of WT-Col as a control; B, C and D: Stem from the plant harbouring short promoter (AT2G23945pro-437:GUS); medium promoter (AT2G23945pro-980:GUS); long promoter (AT2G23945pro-1572:GUS) respectively. Bar: 1 mm.

Furthermore, GUS signals were detected in anthers of plants containing constructs regulated by the medium and long promoter insertions. GUS staining images from different T_1 inflorescences are shown in Figure 5-10, representative GUS staining pictures of different bud stages based upon size are shown in Figure 5-11.

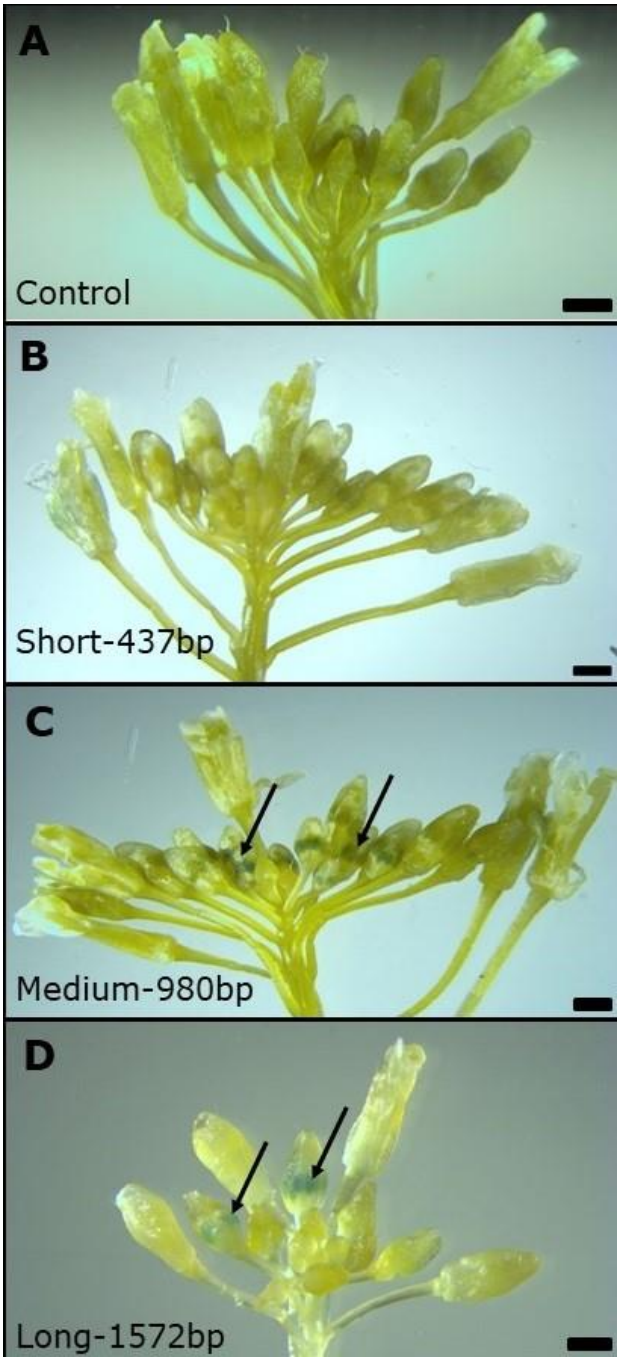


Figure 5-10. GUS staining of the inflorescences from different plants harbouring different sizes of promoter insertions. Representative sample of GUS staining from three biological replicates for every sample, A: WT-Col control; B, C and D: Inflorescences from the short (AT2G23945pro-437:GUS); medium (AT2G23945pro-980:GUS); and long promoters (AT2G23945pro-1572:GUS) respectively. Black arrows point at positive GUS signals. Bar: 1 mm

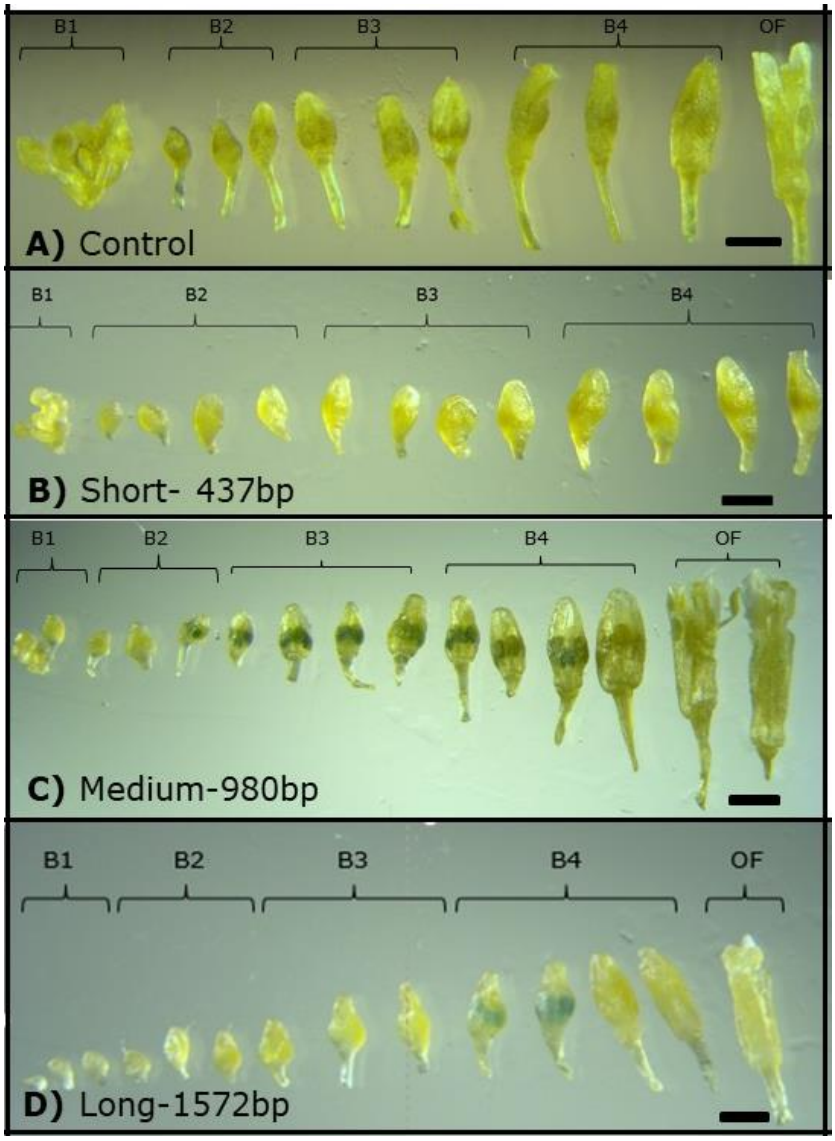


Figure 5-11. GUS staining in buds from different plants harbouring different size of promoter:GUS insertions from the AT2G23945 gene. Representative GUS staining from three biological replicates for every sample, A: WT-Col as a control. B, C and D: Buds from the plant harbouring short promoter (AT2G23945pro-437:GUS), medium promoter (AT2G23945pro-980:GUS); long promoter (AT2G23945pro-1572:GUS) respectively. Bar: 1 mm.

Closer examination throughout anther development showed that GUS staining driven by medium promoter insertion, was localised in the anther tapetum (Figure 5-12C and Figure 5-12E). Moreover, the GUS gene driven by the long promoter insertion, revealed expression during the late stage of tapetum development between Mitosis I to Mitosis II, and was undetectable for the rest of anther development (Figure 5-12D and Figure 5-12F). No GUS expression was observed in WT-Col and in anthers from plants carrying the short promoter (Figure 5-12A and Figure 5-12B respectively).

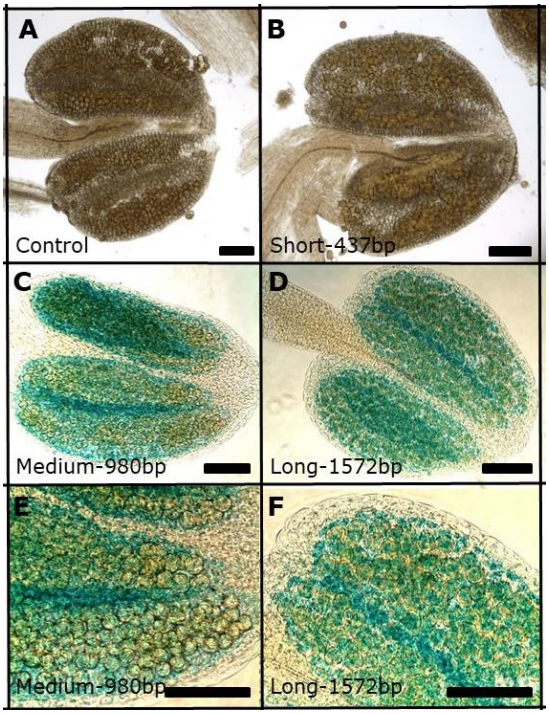


Figure 5-12. GUS staining of anthers harbouring various sizes of AT2G23945 promoter:GUS fusions during pollen mitosis. Representative GUS staining from three biological replicates. A: Anther of WT-Col. B: Anther from short promoter fragment (AT2G23945pro-437:GUS). C and E: Anther from a plant harbouring medium promoter fragment (AT2G23945pro-980:GUS). D and F: Anther from the plant harbouring long promoter fragment (AT2G23945pro-1572:GUS). Bar: 50µm.

The medium length promoter showed the stronger GUS expression and therefore the buds were dissected and analysed in detail. GUS expression was restricted to the tapetum regions and could be detected throughout tapetum development. The GUS signal was undetectable in early anther stages, during the pre-meiotic stage. The GUS signal was clearly visible during the post-tetrad stage of anther development where it showed strongest level of expression once the tapetal cells were clearly defined (Figure 5-13A - 5-13D). Low GUS signal was seen during mitosis I to mitosis II (Figure 5-13E and Figure 5-13F); and became relatively weaker during the late stage of anther development, when the tapetum breakdown occurred (Figure 5-13G and Figure 5-13H), GUS signal was undetectable when buds were about to open (Figure 5-13I and Figure 5-13J). GUS expression was not detected in pollen for all plants (Figure 5-14). Summary of the GUS expression in different tissues of plants harbouring three different pGWB3:AT2G23945pro:GUS promoter segments is presented in Table 5-1.

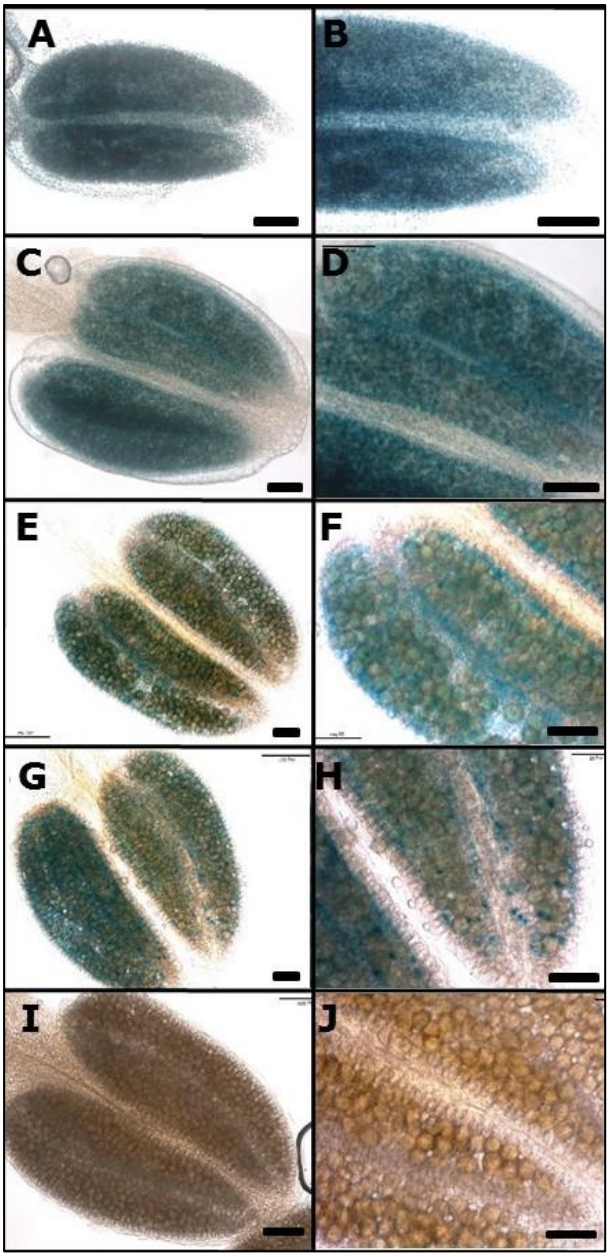


Figure 5-13. GUS staining in anther from different buds stages of plants harbouring the medium promoter fragment (AT2G23945pro-980:GUS). Right panel is the magnification of left panel. Representative GUS staining from three biological replicates for every sample, A and C: Anthers during the tetrad stage. B and D: Higher magnification of A and C respectively. E: Anther undergo mitosis I and mitosis II. F: Higher magnification of E. G: Anther during the tapetum breakdown. H: Higher magnification of G. I and J: Anther from the old closed-buds immediately prior to pollen dehisce. J: Higher magnification of I. Bar: 50μm.

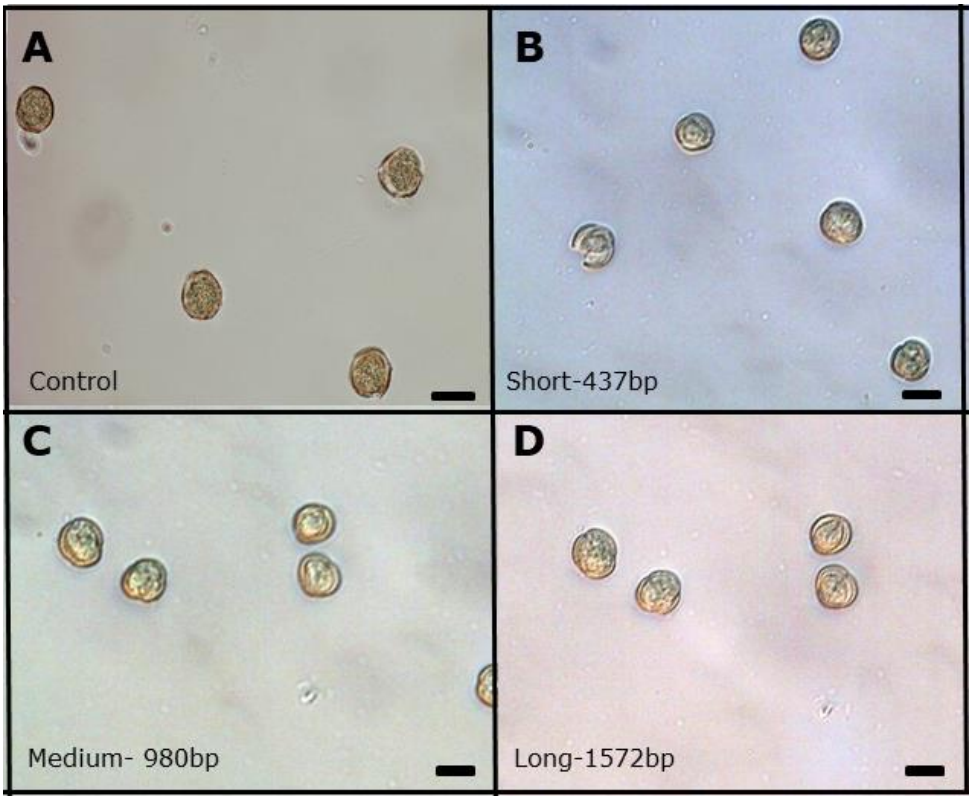


Figure 5-14. No GUS activity was detected in pollen during all stages of pollen development. Representative images are shown using pollen from late stage of pollen development (B4). A: Pollen from WT-Col control; B, C and D: Pollen from the plants harbouring short (AT2G23945pro-437:GUS); medium fragment (AT2G23945pro-980:GUS); and long promoter (AT2G23945pro-1572:GUS) respectively. Bar: 10 µm.

Table 5-1. Summary of the GUS expression in different tissues of plants harbouring the three different pGWB3:AT2G23945pro:GUS promoter fragments. X= Positive GUS signal; -: Negative GUS signal. Short promoter (AT2G23945pro-437:GUS); medium promoter (AT2G23945pro-980:GUS), long promoter (AT2G23945pro-1572:GUS).

TISSUE/STAGE	Short promoter	Medium promoter	Long promoter
Pre-meiosis anther	-	-	-
Tetrad stage anther	-	X	-
Mitosis stage anther	-	X	X
Late development anther	-	-	-
Three-week-old *In vitro* plantlets	-	-	-
Base of young silique	X	-	-
Mature silique	-	-	-
Leaf	-	-	-
Internode	-	-	-

5.2.2 Overexpression vector, pUBQ10: AT2G23945: YFP

The impact of overexpression of the AT2G23945 gene was determined using the pUBQ10 promoter and a C-terminal YFP, enabling visualisation of the transgene. This was transferred, alongside the empty pUBQ10:YFP vector into *A. tumefaciens* and transformed into *N.benthamiana* plants. The control vector of pUBQ10:YFP construct, exhibited expression in the nucleus and membrane (Figure 5-15A and Figure 5-15B) which is identical to the localization of AT2G23945-YFP tagged protein (Figure 5-15C and Figure 5-15D). The high level of expression seen in the control constructs means it is possible to determine the correct localization of AT2G23945 gene from the transient expression analysis in *N. benthamiana*.

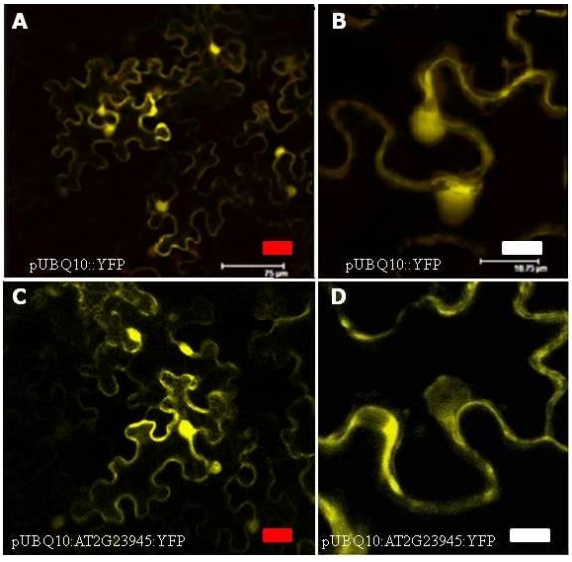


Figure 5-15. Sub-cellular localization of pUBQ10:AT2G23945:YFP and pUBQ10::YFP 4-days post infiltration of *N. benthamiana* leaves. Top panels (A and B) showed the YFP expression on membrane and nucleus for the control construct (pUBQ10:YFP) whilst the bottom panels (C and D) showed equivalent expression on the membrane and nucleus for the AT2G23945 gene. Red bar: 50 µm and White Bar: 10 µm.

5.2.3 Transformation using translational protein fusions for AT2G23945:GFP

To further determine the characteristics of AT2G23945 protein expression, transgenic *Arabidopsis* plants were generated containing the AT2G23945-GFP construct under the control of the native AT2G23945 promoter. Translational fusions were regulated using three different lengths of AT2G23945 promoter, 1811bp (short transcript), 2385bp (medium transcript) and 2946bp (long transcript). In brief, the smallest promoter consisted of the 437bp intergenic region between AT2G23945 and the flanking gene AT2G23950. Two constructs of 2385bp and 2946bp fragments were prepared which extended further and included part of the AT2G23950.

The three AT2G23945 transcripts were cloned into the Gateway entry vector pCR8/GW/TOPO and confirmed by PCR screening using forward specific primers (See Section 2.3.3 and 2.35) and pCR8/GW/TOPO 2.1 reverse primer. Details of the primers are available in Appendix 1. Amplification of entry clone generated the expected product of 965 bp (Figure 5-16A), 493 bp (Figure 5-16B) and 493 bp (Figure 5-16C) for the 1811 bp, 2385 bp, 2946 bp insertion transcripts respectively. pCR8/GW/TOPO clone 8 (C8), clone 11 (C11), clone 18 (C18) for short, medium and long transcript insertions respectively, were confirmed by sequencing and used for further analysis.

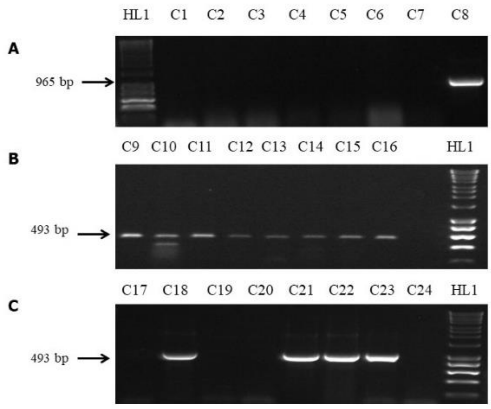


Figure 5-16. Colony PCR of insert and vector sequences from pCR8/GW/TOPO entry clones with three different sizes of putative AT2G23945 translational fusion transcript. Amplifications generated fragments of (A) 965 bp (B) 493 bp (C) 493 bp. Different colonies of the entry clone for; (A) C1-C8: short translational fusion; (B) C9-C16: medium translational fusion; (C) C17-C24: long translational fusion, HL1: HyperLadder 1.

These constructs were introduced into the Gateway binary vector pGKGWG which contains a C-terminal GFP tagged protein enabling visualisation of the protein fluorescence by confocal microscopy. The schematic map of pGKGWG:AT2G23945:GFP destination vector with primer locations is shown in Figure 5-17. Amplification of three different lengths of pGKGWG:AT2G23945:GFP constructs occurred successfully with the desired size of bands using a same promoter specific forward primer (2g23945_486F) coupled with YFP_NY-R reverse primer (Figure 5-18). Primers are listed in Section 2.3.12 and details are shown in Appendix 1. Sequencing of pGKGWG:AT2G23945:GFP clone 1 (C1), clone 9 (C9), clone 17 (C17) were used confirmed insert orientation and sequence. These pGKGWG:AT2G23945:GFP vectors were transformed into *Agrobacterium* using electroporation together with pSoup plasmid (Hellens *et al*., 2000). After verification by PCR (Figure 5-19), clone 1 (C1), clone 9 (C9), and clone 17 (C17) GV3101 from short,

medium and long insertion fragments were transformed into WT-Col.

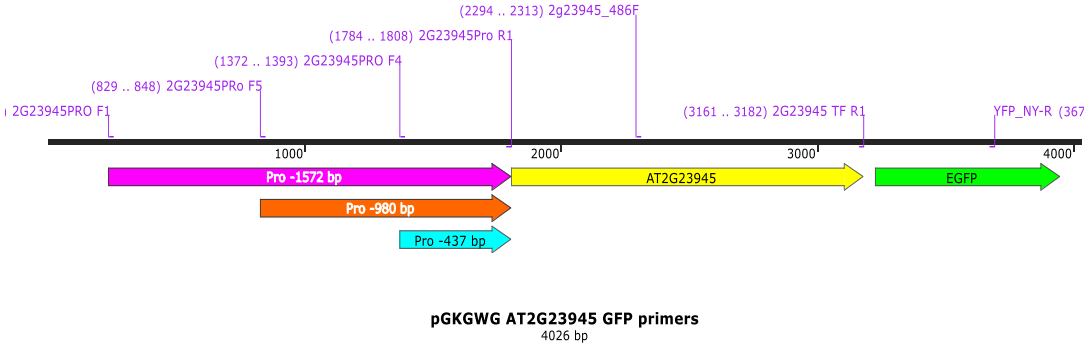


Figure 5-17. The schematic map of pGKGWG:AT2G23945:GFP destination vector. The pink arrow indicates the AT2G23945 long translational fusion transcript, the maroon indicates the AT2G23945 medium translational fusion transcript, the blue arrow indicates the AT2G23945 short translational fusion transcript, and the green arrow indicates the YFP protein tag. The locations of the primers are shown.

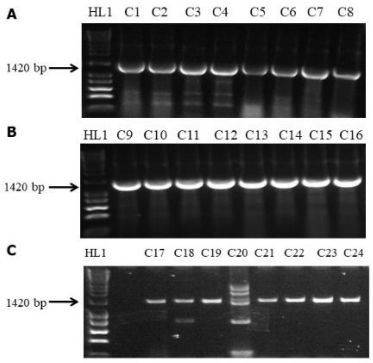


Figure 5-18. Colony PCR of insert and vector sequences from pGKGWG:AT2G23945:GFP destination vectors with three different sizes of promoter driving the AT2G23945 translational fusion. Amplification of destination vector with the expected fragment size of 1420bp. The amplifications were conducted using the promoter specific forward primer (2g23945_486F) coupled with YFP_NY-R reverse primer. pGKGWG:AT2G23945:GFP with different promoters sizes; (A) short, (B) medium, (C) long. Different colonies of the entry clones; C1-C8: short, C9-C16: medium, C17-C24: long, HL1: HyperLadder 1.

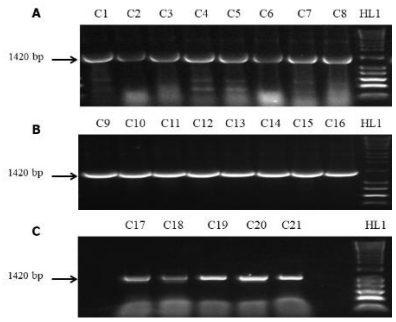


Figure 5-19. Colony PCR of insert and vector sequence from GV3101 containing pGKGWG:AT2G23945:GFP colonies. Amplification of the pGKGWG:AT2G23945:GFP transcripts of ~1420bp (A, B and C) where the amplifications were conducted using a same set of promoter specific forward primer (2g23945_486F) coupled with YFP_NY-R reverse primer; short translational fusion transcript (A), medium translational fusion transcript (B), long translational fusion transcript (C). Different colonies of the entry clone for; A) C1-C8: short promoter fragments, B) C6-C13: medium promoter fragments; C) C17-C21: long promoter fragments, HL1: HyperLadder 1.

PCR analysis showed that five out of eight putative transgenic plants carried the AT2G23945 short promoter translational fusion construct, all six putative transgenic plants tested carried the medium promoter translational fusion construct and seven out of eight carried the AT2G23945 long promoter translational fusion construct (Figure 5-20). Meanwhile, other putative transgenic plants generated faint bands or failed to generate any band. The amplifications were conducted using the same set of gene specific forward primer (2g23945_976F) coupled with YFP_NY-R reverse primer.

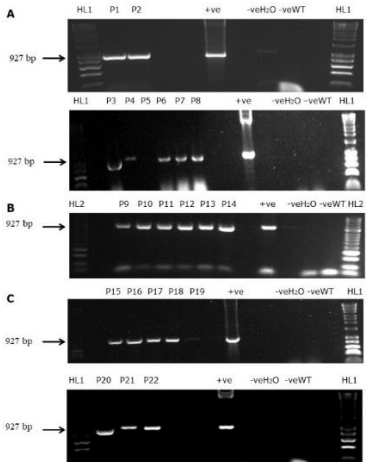


Figure 5-20. PCR amplification of the three different lengths of promoters translational fusions for pGWB3:AT2G23945PRO:GUS transgene in *Arabidopsis* lines. Amplification of fragments with the correct size of 927bp (A, B and C), indicated integration of the transgene into the plant genome. The amplifications were conducted using gene specific forward primer (2g23945_976F) coupled with the YFP_NY-R reverse primer. Different transgenic plants harbouring pGKGWG:AT2G23945:GFP with; (A) P1 to P8: short; (B) P9 to P14: medium; and (C) P15 to P22: long translational fusions; +Ve : pGKGWG:AT2G23945:GFP vector as positive control from each lines, -Ve H_2O: Water negative control, -Ve WT: WT-Col genomic DNA negative control, HL1: HyperLadder 1, HL2: HyperLadder II.

GFP expression patterns were analysed using four independent transgenic plants carrying protein fusions regulated by three different lengths of promoters fused with GFP, three flowers from each biological replicate were tested. GFP expression were almost undetectable for all transgenic plants which carried short and long transgene investigated, except for pGKGWG:AT2G23945:GFP (Medium-2385 bp) (Figure 5-21D-E). Closer examination showed that only one transgenic plant out of four transgenic plants with the medium construct had clear GFP expression in the tapetum region during the free microspore stage. Furthermore, faint GFP expression was observed in the degraded tapetum layers post-mitosis (Figure 5-21E). No GFP signal was detected for WT-Col.

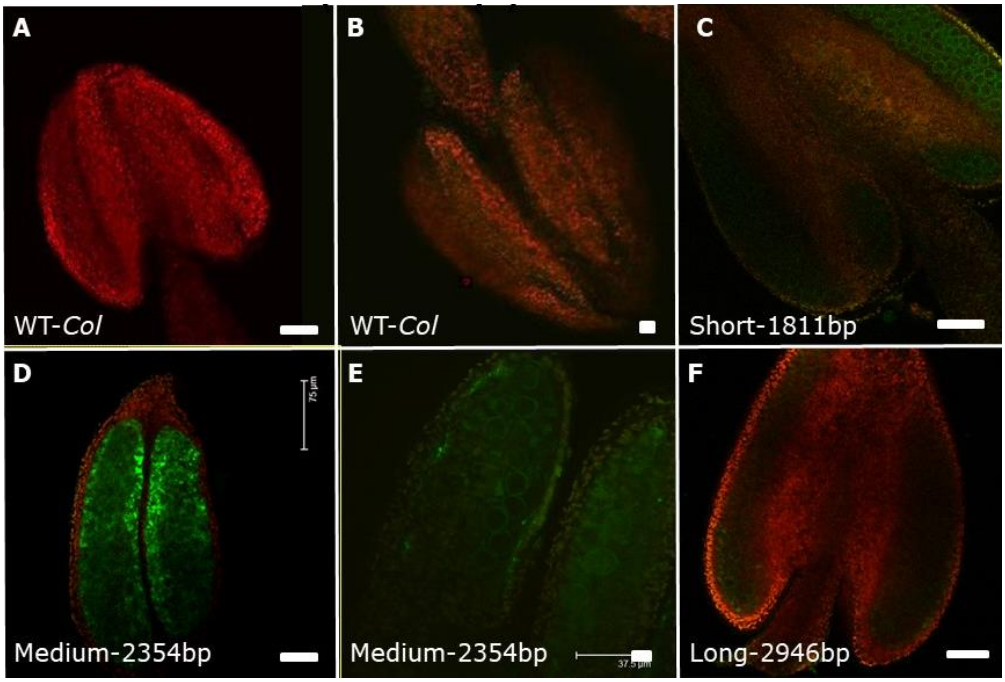


Figure 5-21. Localization of pGKGWG:AT2G23945:GFP. Green fluorescence detected only in tapetal cells region of anther from plant carrying 2354 bp transcript insertion (P3). Anther from the plant harbouring pGKGWG:AT2G23945:GFP carrying; C) Short construct -1811bp; D and E) medium construct -2354 bp; and F) long construct -2964 bp along with WT-Col of A and B. A and D: Free microspore stage upon meiosis. B, C, E and F: Anther during the late stage of anther development. Bar: 10 µm.

Four transgenic plants from each line (short, medium, and long) generated from three different constructs were analysed for transgene expression by RT-PCR. The experiment was conducted to examine the transcript of the GFP-tagged genes due to the failure of the visual detection for plants carried the short and long translational fusion transgene, and the limited expression with only a single plant showing a positive signal for the medium construct. cDNA transcript analysis using the *PP2A3* housekeeping gene successfully amplified a product of ~350 bp (Figure 5-22 A,C and E). Expression analysis using AT2G23945 Forward Primer (2g23945_976F) coupled with YFP_NY reverse primer, successfully produced an amplification

product from the AT2G23945-GFP transcript (band ~927bp; Figure 5-22 B and D). This indicated that the translational transgene fusion transcript was detected in three out of four transgenic plants generated from the short and medium translational fusion constructs. No transcript was detected from plants carrying the pGKGWG:AT2G23945:GFP long transgene (Figure 5-23F).

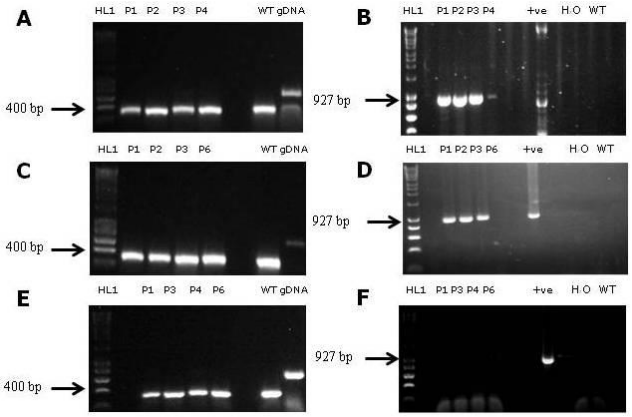


Figure 5-22. RT-PCR Expression analysis of AT2G23945-GFP transgene in transgenic plants. A, C and E: cDNA expression analysis of *PP2A3* House-keeping gene. B, D and F: expression analysis of AT2G23945 (2g23945_976F) and YFP_NY primer, A and B: plant harbouring pGKGWG:AT2G23945:GFP consist of 1811 bp construct; C and D; 2385 bp construct; E and F: 2946 bp construct.

5.3 Discussion

To gain insight into the expression pattern of the AT2G23945 gene, transgenic plants harbouring AT2G23945pro:GUS and the AT2G23945:GFP translational fusion were generated. In brief, the flanking gene of AT2G23950 is located -437bp upstream of AT2G23945 TSS which mean some of the DNA binding-motifs may located in the transcribed region of AT2G23950 gene. In addition, bioinformatic analysis indicated two putative binding sites for *MYB80* located at -193bp to -208bp upstream of the TSS, and *MYB99* in the positive and negative strand at -195bp to -208bp and -1258bp to -1271bp respectively. Therefore, three different lengths of promoter were used to investigate the gene regulation process of AT2G23945.

Characterization of the AT2G23945-GUS activity in transgenic *Arabidopsis* plants indicated different regions of regulatory control in the upstream region. Young silique base is the only region with GUS expression for the construct directed by the short promoter. This signal was not observed in plants harbouring the medium or long promoter constructs. Expression was targeted to the anther tapetum in the constructs driven by the medium and longer promoters. In the medium promoter construct, GUS activity corresponded to the onset of tapetum formation until tapetum breakdown (anther stage 5-anther stage 12; Sanders *et al.*, 1999). Whereas, the long promoter restricted GUS activity to the tapetum region during Mitosis I to Mitosis II stage (anther stage 5-anther stage 12; Sanders *et al.*, 1999). Expression in the anther region is expected since the gene showed altered expression in different transcription factors related to tapetum development. In summary, the analysis suggests that

AT2G23945 is expressed during tapetum development with different expression patterns seen but requires a region of greater than 437bp upstream of the TSS, and the short length of promoter was not sufficient to drive the expression in the tapetum tissues.

Regulatory domains that are essential for tapetum expression may be located -437bp upstream of the TSS. However, the two putative binding motifs for *MYB80,* located at -193bp to -208bp upstream of TSS, and *MYB99,* in positive and negative strand at -195bp to -208bp and -1258bp to -1271bp respectively, were not sufficient to drive anther expression. AT2G23945 is downregulated in *ams* and *ms1* mutants, however there is no AMS and MS1 DNA binding-motifs are present in the 3000bp upstream of the TSS. This suggests additional regulatory sequence further upstream are needed for tapetum expression. Expression analysis of *DYT1* indicated the shortest length for the promoter capable to drive the GUS expression in the tapetum region was 550bp. Promoters ranging between 91bp to 463bp were insufficient to drive expression. Hence, the authors speculated that the crucial motifs binding site for the gene to perform its normal function to be in the region between -463bp to -550bp of promoter region, which consists of *Cis*-element of MYB, WRKY, and bZIP family protein, but another unknown elements may also be necessary for normal expression of the gene (Gu *et al.*, 2014). DYT1pro-GUS expression commenced during the establishment of the anther four locules where the tapetal cells start to differentiate, expression in the tapetum peaked during the meiosis stage and reduced after the tetrad formed (Gu *et al.*, 2014).

Medium and long promoter constructs were sufficient for the expression in anther tissue. Furthermore, the addition of the additional 592bp directed the expression to more specific stages of tapetum development restricting expression to mitotic stages. This suggests that other elements than MYB80 or MYB99, are involved in the regulation of the AT2G23945 transcription process and that transcriptional repression or post-transcriptional processes are limiting expression to specific stages in tapetum development. In addition, insertion of a large transgene can alter the normal configuration of the chromatin and therefore can change the GUS expression pattern (Millar and Gublera, 2005).

Research with different lengths of promoter regions have been previously used to show the complexity of regulation. For example, *MYB103* (*MYB80*) has been shown to be the regulator of anther and trichomes development. MYB103pro-GUS with a 1.0kbp promoter (Higginson *et al*., 2003) and a 3.2kbp promoter (Li *et al.*, 1999) showed similar promoter activity in developing anthers, with no GUS signal detected in trichome tissues. Interestingly, Higginson *et al.*, (2003) revealed that GUS expression was seen in trichomes using a 3.6kbp transgene which included a 6bp region along with transcribed region and 3’ UTR. This indicated that the regulatory sequence was located downstream of the TSS, not in the promoter region.

In the same way, the *MYBGL1* promoter regions contains *MYB* DNA-binding motif involved in stipules and trichomes development. Early studies showed that the putative promoter region directed GUS signal to stipules, not trichomes (Oppenheimer *et al.*, 1991). Subsequently, a promoter construct consisting of the transcribed

region along with 3’ UTR successfully directed the GUS signal to leaf primordia and trichomes (Larkin *et al.*, 1993). The authors speculated that the transcribed region contained an enhancer for normal trichomes development.

It is unclear whether AT2G23945 is regulated by MYB80. However, it is clear that additional motifs other that the MYB80 and MYB99 are needed to regulate expression Comparatively, the AT2G23945GUS expression shared similar GUS expression pattern with *MYB80 (MYB103)*. ATMYB103pro-GUS expression was observed in developing anthers, appeared during anther meiosis stage with decreased expression observed once the anther reached maturity (Li *et al.*, 1999), which is similar to the AT2G23945pro-GUS expression driven by the medium promoter insertion.

In addition, transcription factor may form another complex with other elements to bind to the promoter region for transcription regulation and ongoing study is necessary to understand the full regulation process of the gene (Gu *et al.*, 2014).

MS1pro-GUS signal is observed in the tapetal cells during the late pollen meiosis stage (Wilson *et al.*, 2001). Furthermore, investigation of AMSpro-GUS promoter fusion showed that GUS was detected during the post meiosis stage until mitosis I stage. The GUS activity was visualised mainly in the tapetal cells, developing microspores with low expression seen in the anther locule (Sorensen *et al.*, 2003).

A number of proteases have been shown to be expressed specifically in the anther and to regulate tapetum PCD*. PCS1* gene is an aspartic protease expressed during anthers and pollen development (Ge *et al.*, 2005a). *UNDEAD,* encoded for an aspartic-protease and regulated by *MYB80,* showed similar expression with *MYB80* during anther meiosis stage to tapetum degradation stage (Phan *et al.*, 2011). Current research on A36 and A39 Aspartic proteases, A36pro-GUS revealed the GUS activity have been specifically in tricellular pollen, whereas and A39pro-GUS activity has been detected in pollen during tetrad to bicellular stage (Gao *et al.*, 2017).

To compliment the GUS assay analysis and enable live imaging, dynamic temporal and spatial analysis, fluorescence tags have been employed. For protein subcellular localization, two types of the constructs have been tested which carried fluorescence-tagged protein. The first construct developed with transgenes driven by the overexpression pUBQ10 promoter with a YFP-tagged protein, and the second construct control by the native promoter, containing the gene and GFP tag. Visualization of YFP and GFP-tagged protein can help the identification of the specific cellular localization to give an idea about protein function and protein interactions (Li *et al.*, 2006). In addition, there is a need to include gene sequence along with promoter segment for protein localisation analysis. The investigation may give an idea whether modifications occurred during transcription, post-transcription, pre-translation, translation or post-translational processes (Molina and Grotewold, 2005).

No specific subcellular localization of pUBQ10:AT2G23945:YFP could be seen with expression in the plasma membrane and nucleus,

which is identical to the control pUBQ10::YFP construct. For this reason, transgenic plants harbouring pGKGWG:AT2G23945:GFP with different length of promoter.

Transient expression using overexpression vector and *N. benthamiana* leaves has several advantages over generating new transgenic lines. It is less laborious and can be performed in shorter period with less cost required. However, it also has several disadvantages. The gene inserted is driven by a strong constitutive promoter that might important in protein expression due to abundance of the protein in the cells. The use of over expression constructs can also lead to suppression of expression (Grefen *et al.*, 2010). A further concern for ectopic expression may be localised or be in an altered form which impacts on protein and disrupts the multiprotein complexes (Tian *et al.*, 2004).

The GFP expression patterns were not consistent; GFP expression was only detected during active tapetum development from transgenic plants harbouring medium transcript (2385bp) construct in one single plant (P3). Thus, it is hard to provide a concrete analysis of GFP expression since only a single plant showed GFP expression. The transgenes have been confirmed to be fused in-frame with GFP with the absence of any stop codon. Transcripts of the transgenes were detected using RT-PCR for short and medium transcripts, indicating expression is occurring, but no transcript was detected from the long promoter constructs (Figure 5-23). GFP transformants analysis is necessary to confirm the expression and the possibility that GFP expression was silenced due to the post-

translational modification for short and medium transcript insertions.

Faint fluorescence has been observed from the transformants generated by the medium transgene, but no other signal was detected from the other lines. Lack of positive fluorescence may be due to the protein conformation masking fluorescence. In some cases, the fluorescence protein can be degraded over time as suggested by Casey *et al.*, (2000) which may cause the undetectable RT-PCR signal from the plant harbouring long transcript (2946bp) construct. Furthermore, silencing of the GFP may be occurring, or the fluorescence protein may not be folded properly, as observed in other fusions by Toth *et al.*, (2007) and Casey *et al.*, (2000).

To get the fluorescence signal, the location of the fluorescence proteins within tissues must be reachable by the confocal laser. Otherwise, the masking of the protein can disrupt fluorescence protein detection (Tian *et al.*, 2004). Another concern, C-terminal and N-terminal fusion may result in different localization due to the large size of tagged protein (~27kDa) which may block protein targeting and disrupt protein interactions. Nelson *et al.,* (2007) suggest using both N- and C- terminal orientation of tagged-protein to increase the possibility of positive signals. Another possibility is the weak native promoter of AT2G23945 may cause difficult detection of fluorescence protein (Li *et al.,* 2006; Nelson *et al.,* 2007). To overcome the absence of fluorescence protein detection, immunofluorescence could be employed to indicate GFP subcellular localisation and enhance the GFP signal (Swenson *et al.*, 2007).

CHAPTER 6: GENERAL DISCUSSION

6.1 Summary of Main Findings

The primary aim of this research was to determine the effects of the manipulation of selected proteases and determine whether they are involved in tapetum and pollen development. Reverse genetics and transgene analysis were utilised to explore the effects of altered expression of seven proteases and one F-Box protein. Overexpression analysis differs from the knockout analysis as phenotypes can be generated in situations where knocking out of a single gene does not cause phenotypic changes due to redundancy issues (Prelich, 2012). The selected proteases and F-box gene are likely to exhibit redundancy, since they are members of large gene families, therefore the approach of overexpression and ectopic infiltration into *N.benthamiana* leaves was selected as an method for gene characterisation. For a complete analysis of gene function, however there is a need for a null mutant, but if redundancy is occurring it may be necessary to create knockouts in multiple genes, to obtain phenotypic changes. Such analyses can then be used to investigate the gene network (Cutler and McCourt, 2005).

Of the eight genes investigated, only manipulation of AT2G23945 and AT4G30330 exhibited reproducible changes in fertility, suggesting a potential role in pollen development. No apparent phenotypic changes were observed for the AT3G20015, AT5G19120, AT2G21430 and AT3G16210 overexpression lines (See Section 3.3.4 and Chapter 4).

However, two out of four AT1G25510 overexpression transformants showed severe sterility effects with 46.4% and 41.4% sterility observed on main stem (lines OE7 and OE8, respectively, see Chapter 3.2.5.1). Also, one out five AT5G10760OE transformants, exhibited 95% percentage of sterility on main stem for OE3, however the remaining transformants showed identical fertility to WT-Col (See Chapter 3.2.5.3). For both lines, no clear correlation between percentage of sterility and gene expression level could be confirmed. Inconsistency of T_1 transformants phenotypes were also observed in AT2G23945, AT1G25510 and AT5G10760 transgenic lines with varied levels of transgene expression.

The observed variation may happen due to various reasons, but is frequently due to positional effects of the transgene, or multiple insertional events that potentially have caused chromosomal arrangements (Butaye and Cammue, 2005). Furthermore, the T_1 generation traits are still segregating, so analysis in subsequent generations is required to confirm phenotypes (Nacry *et al.*, 1998; Tax and Vernon, 2001). Positional effects of transgene insertions, and whether this is in heterochromatin or euchromatin regions, can explain the wide range of expression levels and methylation related gene silencing frequently observed (Day *et al.*, 2000). Methylation-associated silencing can cause varied levels of expression where the insertions are in the heterochromatin regions (Schubert *et al.*, 2004). For example, *AGAMOUS-LIKE18* (*AGL18*) silencing has been shown to be due to the positional effects of the transgene that caused chromosomal arrangement and gene deletions (Xing and Zachgo, 2007).

A comprehensive study including the T-DNA insertions, gene overexpression through ectopic promoter, and RNAi silencing have been conducted for the AT2G23945 (see Chapter 4). Two T-DNA insertions lines were analysed for AT2G23945, SALK 135523 and SALK 134919, which unexpectedly resulted in increased expression of the AT2G23945 gene, rather than reduction of expression. SALK 135523 showed defective of pollen and anther development by microscopy using DAPI, Auramine O, ethidium bromide acridine orange and Alexander's staining (See Section 4.2.2). Low pollen viability was observed with more than half of the pollen appearing shrunken and inviable, a delay in the breakdown of the tapetum was also observed in SALK 135523. Overexpression of the AT2G23945 gene, in three independent lines, which showed different levels of increased expression (SALK 135523, SALK 134919 and AT2G23945OE), resulted in significant reduction in mean silique length due to a lack of seed set (See Section 4.2.2 and 4.2.3). Since the two AT2G23945 SALK insertional mutants showed increased expression rather than null mutations, RNAi was used as an approach to knockdown gene expression. For AT2G23945RNAi transformants, different degrees of severity of phenotypes were observed with some plants showing severe sterility in the T_1 generation (See Chapter 4.2.4). These AT2G23945RNAi plants showed different levels of sterility where five out of six transgenic plants P1-P5 exhibited reduced numbers of siliques, whereas P7 had an increased number of siliques compared to the WT-Col plant (Table 4-4, Figure 4-38). The inflorescences appeared phenotypically normal except for P1 and P4 which showed reduced bud number all inflorescences (Figure 4-39). In addition, pollen viability was not affected from the T_1 transformant plants in all transgenic lines,

except for P4 and P7 where approximately 10% and 40% of pollen were non-viable, respectively (Table4-4; Figure 4-41).

As indicated, reductions in fertility were observed in the RNAi lines (Section 4.2.4), although the presence of the RNAi transgenes were confirmed, analysis of expression in these T_1 lines was not possible due to the limited amounts of transgenic floral material and the time constraints. Ongoing analysis in T_2 lines showed high variability in gene expression patterns, male sterility and silique length. It is still unclear whether the reduction in sterility rate in RNAi plants has a direct link on water stress, or high variability of gene expression pattern.

The comparative analysis of gene expression patterns were investigated using Bioinformatic tools, qRT-PCR, GUS fusion assay, and GFP fusion analysis. The results varied depending on the sensitivity of the detection methods. Information from e-FP browser public access software (See Chapter 4.1) revealed that high expression level only in flower stage 9, equivalent to pre-meiosis stage (P M Sanders *et al.*, 1999) of anther development (Winter *et al*., 2007). Furthermore, RT-PCR tissues analysis revealed that AT2G23945 expression was high during the late stage of the anther development with minimal expression detected in other tissues (See Chapter 4.2.1, Figure 4-7). Based on the promoter-GUS examinations (See Chapter 5.2.1), AT2G23945 promoter activity was shown to be active in anthers, with GUS primarily detected specifically in the tapetum throughout tapetum development, and to a lesser degree using the short promoter constructs in the young silique base. These data demonstrated that uncharacterised

regulatory sequence further upstream of TSS are needed for anther and tapetum expression. Moreover, GFP signal detection was not detected by confocal microscopy for most of the AT2G23945-GFP lines investigated. Limited GFP expression was detected during the free microspore stage and tapetum degeneration process (See Chapter 5.2.3), however further investigation is needed to confirm if this signal is genuine.

A transcription factor may bind to a specific cis-regulatory element but in many cases, the transcription factor may not function alone (Zhang, 2003). In some cases, two transcription factors may form homodimers or heterodimers for binding process. Recruitment of the complex by the gene promoter which gives more option for greater DNA targeted sequence and increase the probability of a specific recognition site. Precise arrangement of the protein DNA binding motif and the promoter binding site is necessary for the transcription initiation process (Singh, 1998). Differences in the observed expression patterns in the different length promoter fusions may be due to the absence of key regulatory motifs in the cloned regions, or to the different binding of a regulatory factor due to altered interactions between different protein subunits or changed protein conformation (Prouse and Campbell, 2012).

A number of tapetum transcription factors have also shown similar differences in expression. Low expression was detected for *MS1* by employing RT-PCR and promoter GUS fusions, with signal detected in closed flower buds. The promoter fusion showed highest MS1pro-GUS signal in tapetum cells post meiosis, and to a lesser extent in the locule and pollen grains (Wilson *et al.*, 2001). Repeated RT-PCR

analysis and use of functional translational fusion lines (Yang *et al.*, 2007) indicated that MS1 expression was for a much narrower period of time initiated during microspore release.

6.2 Role of AT2G23945 as a putative protease in PCD

The putative aspartyl protease AT2G23945, is expressed specifically in the anther and overexpression of AT2G23945 resulted in impaired tapetum degeneration (sections 4.2.2.1). A large number of proteases are present in the genome and likely to have diverse functions, partly due to the different substrates in different tissues. In the tapetum, *CEP1* (Zhang *et al*., 2014), *OsAP25* and *OsAP37* have been shown to have pro-PCD roles (Niu *et al.*, 2013), whilst *PCS1* (Ge *et al.*, 2005), *UNDEAD* (Phan *et al*., 2011), *CP51* (Yang *et al.,* 2014) and *A36* and *A39* (Gao *et al.*, 2017) exhibit anti-PCD roles. The AT2G23945 aspartyl-protease did not show significant cell death upon infiltration into tobacco leaves (section 3.2.4), this suggests that it is not acting as an executioner protease that is inducing ectopic cell death. Nevertheless, this lack of activity could also be due to AT2G23945 functioning as part of a complex with key proteins missing from the infiltrated tobacco leaves, or that AT2G23945 was not expressed in a functional form. There is also the potential that AT2G23945 acts in an anti-apoptotic role in the anther to maintain tapetum development rather than inducing tapetum PCD. It may be that under stress conditions, such as high temperature or oxidative stress, that infiltration with AT2G23945 would cause resilience and prevents tapetum PCD. Altered tapetum development significantly impacts on pollen development as has been shown by overexpression SALK 135523 TDNA-Insertion line (See Chapter

4.2.21). Thus, it is tempting to speculate that the delayed tapetum degeneration may be due to the anti-PCD character of AT2G23945. To confirm whether AT2G23945 plays a role in tapetum PCD, its biochemical properties and activity during pollen development needs to be identified.

6.3 Future Research

6.3.1 Determination of Insertion Copies

Several things need to be considered in transformation studies, but a key factor is whether the alteration of the expression and phenotypic changes are due to a single transgene integration. Further analysis should be carried out to confirm whether Mendelian inheritance is seen for the T_1 population. Single recessive sporophytic mutation can be confirmed through reciprocal crosses with wild-type plant or Southern blot analysis (Lou *et al.*, 2014; Zhang *et al.*, 2014).

Pollen from tapetum mutants/overexpression lines could be crossed with the wild type; if the mutant phenotype is due to a single transgene insertion, the F_1 plants generation will follow Mendelian inheritance with production of a fully fertile F1 generation, which if selfed will segregate in a 3:1 ratio (Fertile:Sterile), confirming that the mutation is due to a single insertion. Independent confirmation can be conducted through Southern blot analysis, using *NPTII* as a probe (Nagaya *et al.*, 2005). Additionally, the exact integration position can be mapped to the genome afterwards by inverse

polymerase chain reaction (IPCR) and sequence analysis (Nagaya *et al.*, 2005).

6.3.2 Characterisation of the role of AT2G23945 in PCD

To help characterise the potential role that AT2G23945 plays in tapetum PCD an *in vivo* investigation could be performed to identify its potential role in cell death activity in yeast, which would complement the cell lysis study and trypan blue staining (See Chapter 3.2.4). This approach has been reported previously for *EAT1* regulated proteases, where the construct was transformed into a PCD-deficient yeast strain, *Δyca1* (Niu *et al.*, 2013). The authors reported that overexpression of *OsAP25* and *OsAP37* in deficient-PCD yeast serve as a promoter of cell death activity while empty *Δyca1* yeast showed no impairment (Niu *et al.*, 2013). A similar approach could be adopted for AT2G23945 to analyse the potential function in PCD.

6.3.3 Subcellular localisation of the proteases

Additional work is required to confirm the localisation of the mature protein as mentioned in Section 5.2.3. The confocal analysis of the GFP fusion did not establish a defined localisation for the protein since GFP signal was not consistent between transgenic lines. An alternative approach that could be conducted in the future is by immunolocalisation using an anti-GFP antibody hybridised to tissue sections. To provide greater insight into the transcript expression profile, *in situ* hybridization can also be conducted using gene-

specific antisense probes. Precise expression pattern of tapetum specific transcript in the tapetum has been shown in *EXL5* and *EXL6* are (Xu *et al.*, 2014); and *TKPR1* and *TKPR2* (Grienenberger *et al.*, 2010).

6.3.4 Terminal-deoxynucleotidyl transferase- mediated nick end labelling (TUNEL) assay

Increased expression of AT2G23945 (SALK 135523, See Section 4.2.2) resulted in altered delayed tapetum breakdown and abnormal pollen. This delayed tapetum breakdown is likely to be due to altered PCD, however needs to be confirmed in future work. This could be done by using a Terminal-deoxynucleotidyl transferase- mediated nick end labelling (TUNEL) assay to confirm DNA fragmentation, which is seen in apoptosis-like cell death programmes during tapetum PCD. This is an enzymatic technique incorporating fluorescein-12-dUTP into the 3'-hydroxyl-ends of fragmented DNA (Gavrieli *et al*., 1992), which is typically used for cell-specific analysis of PCD. This approach is used extensively to study tapetum PCD, for example cleavage of DNA occurred in the tapetum of the *undead* mutant and its master regulator *myb80* mutant (Phan *et al.*, 2011), and the rice transcription factor, *eat1* mutant (Niu et al., 2013). TUNEL positive signals in the sperm and vegetative nuclei have also been described by Gao *et al.,* (2017) in *Arabidopsis a36 a39* double mutants.

6.3.5 Post-translational protein modifications

A distinct feature of aspartic proteases is the presence of a plant-specific domain of around 100 amino acids inserted into the C-terminal domain. Under native conditions, most of the zymogen-derived proteins like Aspartic proteases are not in thermodynamically stable folded forms. In these cases, the prosegment plays a distinct role in stabilising the protein and inhibiting enzyme reactions. The prosegment and plant specific insert (PSI) are removed from the precursor through proteolytic maturation of the inactive proenzyme to become an active form of protease which possess two Asp residues (Simões and Faro, 2004). The enzyme displays an active form in acidic pH upon protonation of carboxylate side chains of the enzyme, triggering conformation changes, thus opening the active site for catalytic activity of the targeted peptide bond (Szecsi, 1992).

Western blotting can be performed to show proteolysis as a consequence of enzyme maturation, which produces a mature enzyme for the executor phase. Zhang *et al.*, (2014) showed that the 38-kD *CEP1* Pro-enzyme transformed into an active executor mature 26-kD enzyme during the pollen meiosis stage. *CEP1* expression is greatest during post meiosis stage, as shown by the qRT-PCR results, which correlate with the enzymatic reaction analysis. Western blotting can also be used to investigate if post-translational modifications are occurring. The aspartic protease, *PCS1* probed by anti-FLAG-M2 monoclonal antibody, and A36-GPF and A39-GFP probed with anti-GFP antibody (Ge *et al.*, 2005; Gao *et al.*, 2017) exhibited a larger size than that expected due to post-

translational glycosylation. Furthermore, western blotting of cysteine protease *RD21,* using anti-RD21 antibodies, revealed that post translational-modification occurred with the absence of the granulin domain in the purified protein (Yamada *et al.*, 2001). These approaches could be used in future work to determine any functional modifications occurring to the AT2G23945 protease.

6.3.6 Enzymatic characteristics of AT2G23945

Future work could also be conducted to confirm whether AT2G23945 as an executor in the tapetum PCD pathway. The hydrolysis activity of specific aspartyl-proteases has been previously studied using different substrates. *CDR1* displayed protease activity against labelled fluorescein isothiocyanate-hybovine serum albumin (FITC-BSA) (Xia *et al.*, 2004). Aspartic protease *PCS1* has been shown to hydrolyze a casein substrate (Ge *et al.*, 2005). Furthermore, *OsAP25* and *OsAP37* displayed positive activity using the mammalian caspase substrate, FITC-VAD-fmk, but no activity against caspase-1-9 (Niu *et al.*, 2013). The authors suggested that these two aspartyl-protease may activate another caspase-like protease to carry out the hydrolysis process, rather than acting directly as executors in PCD. In addition, recent i*n vitro* analysis of *A36* and *A39* indicated aspartic proteolytic activity against FITC-labeled casein (Gao *et al.*, 2017). *In vitro* proteolysis analysis, of these two Asp residues, resulted in defective proteolysis activity (Gao *et al.*, 2017). The importance of aspartyl protease activity can be identified through catalytic site-directed mutagenesis which specifically cleaves the targeted substrate (Li *et al.*, 2016). Mutations at the conserved D180 and D379 aspartic residues of *ASPG1* resulted in

deficient protease activity (Yao *et al.*, 2012). Aspartic- *CDR1* site mutation D108N and D319N reduced the aspartic activity, but with no apparent phenotype changes (Xia *et al.*, 2004).

Pepstatin, an inhibitor of protease activity can also be used to test the activity and role of proteases. *In vitro* aspartyl-protease activity of *ASPG1* was reduced after pepstatin treatment (Yao *et al.*, 2012); the same was also seen for *A36* and *A39* (Gao *et al.*, 2017). Moreover, addition of pepstatin into the infiltration culture reduced the *CDR1* anti pathogen activity and caused enhanced pathogen attack on *N. benthamiana* leaves (Xia *et al.*, 2004). Furthermore, treatment of the *OsAP25* or *OsAP37* infiltration cultures with pepstatin successfully inhibited the aspartyl-protease activity and reduced the cell death on inoculated leaves (Niu *et al.*, 2013).

Multiple proteases inhibitor can be used to investigate protease activity using inhibition assays, for example the papain-like cysteine protease inhibitor E-64, the cysteine protease inhibitor leupeptin (Zhang *et al.*, 2014). Additionally, to test whether any caspases related activity occurs in AT2G23945, the caspase inhibitors Z-VAD-fmk and Ac-YVAD-cmk could be tested, since both inhibitors weekly inhibited *CEP1* lytic activity. In addition, assays can be conducted in different temperature and pH to determine the properties of putative proteases.

6.3.7 Related Genes and PCD Indicator Genes Analysis

Silencing of the AT2G23945 gene may cause certain genes to have altered expression which may give an insight into other gene involvement in tapetal PCD. The AT2G35615, AT4G30030 and AT4G30040 aspartic protease genes are thought to be closely related to the AT2G23945 gene (See Chapter 4.1). Further analysis of these potential family members may help determine if the lack of observed phenotypes are due to gene redundancy. Crosses and genetic analysis can also be used to determine if the, protease tapetal PCD marker genes, *UNDEAD* (Phan *et al.*, 2011) and *CEP1* (Zhang *et al*., 2014) are part of the same pathways. In addition to the genes mentioned above, animal PCD indicator genes could be tested on AT2G23945 gene-deficient mutant. As discussed in Section 1.3, analysis of the expression level of *ARABIDOPSIS BAX INHIBITOR* 1 (*AtBI-1*) may help confirm the level of PCD pathway conservation between animals and plants.

6.3.8 Transmission electron microscopy (TEM)

Degradation of mitochondria, condensed cytoplasm, vacuole enlargement and undefined organelles and ER enlargement are landmark events for animal PCD (See Section 1.3). Hence, ultrastructural observations of tapetum and pollen development through Transmission Electron Microscopy (TEM) analysis could be conducted to study the differences between mutant and wild-type during tapetal cell and pollen development (See Chapter 1.2.3 and Chapter 1.2.4). This could help characterise the tapetum associated

defects of AT2G23945 and establish whether these defects are due to altered PCD (Niu *et al.*, 2013; Zhang *et al.*, 2014).

6.3.9 Reactive Active Oxygen (ROS) Analysis

As discussed in Chapter 1.3, ROS, particularly H_2O_2, play an important role in triggering caspases-driven animal PCD (Van Breusegem and Dat, 2006; Xie *et al.*, 2014). Interestingly, some of the plant PCD processes show high accumulation of ROS in intercellular fluid during the cell death process, with cell lysis occurred once the homeostasis between ROS-producing and ROS scavenging is disrupted. High levels of ROS are seen in pollen grains of aspartyl-protease *a36* and *a39* mutants during PCD (Gao *et al.*, 2017), H_2O_2 accumulation in rice aspartyl-protease *OsCDR1* overexpression plant as a signal response to pathogen attack and cell hydrolysis activation (Achard *et al.*, 2004). Analysis of ROS levels in the AT2G23945 mutant and Overexpression lines would be interesting future work.

6.4 Conclusions

In conclusion, the data presented in this study provides valuable insight into the role of the putative aspartyl-protease, AT2G23945, in tapetum PCD and pollen development. AT2G23945 expression is necessary for normal tapetum development, PCD and pollen development, with alterations of tapetum degeneration observed in the overexpression lines. These defects are likely to be due to altered tapetum PCD and to be linked to the altered aspartyl protease expression. Nevertheless, there is a need to confirm the biochemical role of AT2G23945 in the tapetum to help understanding of plant PCD processes and pollen development.

REFERENCES

Aarts, M. G. *et al.* (1997) 'The Arabidopsis MALE STERILITY 2 protein shares similarity with reductases in elongation/condensation complexes.', *The Plant Journal*, 12(3), pp. 615–623. doi: 10.1046/j.1365-313X.1997.00615.x.

Aarts, M. G. M. *et al.* (1997) 'The Arabidopsis MALE STERILITY 2 protein shares similarity with reductases in elongation/condensation complexes', *Plant Journal*, 12(3), pp. 615–623. doi: 10.1046/j.1365-313X.1997.d01-8.x.

Achard, P. *et al.* (2004) 'Modulation of floral development by a gibberellin-regulatedmicroRNA', *Development*, 131(14), pp. 3357–3365. doi: 10.1242/dev.01206.

Van Aken, O. and Van Breusegem, F. (2015) 'Licensed to Kill: Mitochondria, Chloroplasts, and Cell Death', *Trends in Plant Science*. Elsevier Ltd, 20(11), pp. 754–766. doi: 10.1016/j.tplants.2015.08.002.

Alexander, M. P. (1969) 'DIFFERENTIAL STAINING OF ABORTED AND NONABORTED POLLEN', *Stain Technology*, 44, pp. 117–122.

Alonso, J. M. *et al.* (2003) 'Genome-Wide Insertional Mutagenesis of Arabidopsis thaliana', *Science*, 301(5633), pp. 653–657. doi: 10.1126/science.1086391.

Alves-Ferreira, M. *et al.* (2007) 'Global Expression Profiling Applied to the Analysis of Arabidopsis Stamen Development', *Plant Physiology*, 145(3), pp. 747–762. doi: 10.1104/pp.107.104422.

Ariizumi, T. *et al.* (2003) 'A novel male-sterile mutant of Arabidopsis thaliana, faceless pollen-1, produces pollen with a smooth surface and an acetolysis-sensitive exine', *Plant Molecular Biology*, 53(1–2), pp. 107–116. doi: 10.1023/B:PLAN.0000009269.97773.70.

Ariizumi, T. and Toriyama, K. (2011) 'Genetic regulation of sporopollenin synthesis and pollen exine development.', *Annual review of plant biology*, 62, pp. 437–460. doi: 10.1146/annurev-arplant-042809-112312.

Artal-Sanz, M. and Tavernarakis, N. (2005) 'Proteolytic mechanisms in necrotic cell death and neurodegeneration', *FEBS Letters*. Wiley-Blackwell, 579(15), pp. 3287–3296. doi: 10.1016/j.febslet.2005.03.052.

Aya, K. *et al.* (2009) 'Gibberellin Modulates Anther Development in Rice via the Transcriptional Regulation of GAMYB', *the Plant Cell Online*, 21(5), pp. 1453–1472. doi: 10.1105/tpc.108.062935.

Azpiroz-Leehan, R. and Feldmann, K. A. (1997) 'T-DNA insertion mutagenesis in Arabidopsis: Going

back and forth', *Trends in Genetics*, 13(4), pp. 152–156. doi: 10.1016/S0168-9525(97)01094-9.

Balasubramanian, S. and Schneitz, K. (2000) 'NOZZLE regulates proximal-distal pattern formation, cell proliferation and early sporogenesis during ovule development in Arabidopsis thaliana', *Development*, 127(19), pp. 4227–4238. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10976054%5Cnhttp://dev.biologists.org/content/127/19/4227.full.pdf.

Baulcombe, D. C. . *et al.* (1986) 'Expression of biologically active viral satellite RNA from the nuclear genome of transformed plants', *Nature*, 321(February), pp. 446–449. doi: 10.1038/321446a0.

Berendzen, K. *et al.* (2005) 'A rapid and versatile combined DNA/RNA extraction protocol and its application to the analysis of a novel DNA marker set polymorphic between Arabidopsis thaliana ecotypes Col-0 and Landsberg erecta', *Plant Methods*. BioMed Central, 1(1). doi: 10.1186/1746-4811-1-4.

Borg, M. and Twell, D. (2011) 'Pollen: Structure and Development', *eLS*, (October 2011). doi: 10.1002/9780470015902.a0002039.pub2.

Bosch, M. *et al.* (2008) 'Initiation of Programmed Cell Death in Self-Incompatibility: Role for Cytoskeleton Modifications and Several Caspase-Like Activities', *Molecular Plant*. Elsevier BV, 1(6), pp. 879–887. doi: 10.1093/mp/ssn053.

Bosch, M. and Franklin-Tong, V. E. (2007) 'Temporal and spatial activation of caspase-like enzymes induced by self-incompatibility in Papaver pollen', *Proceedings of the National Academy of Sciences*, 104(46), pp. 18327–18332. doi: 10.1073/pnas.0705826104.

Brandizzi, F., Fricker, M. and Hawes, C. (2002) 'A greener world: The revolution in plant bioimaging', *Nature Reviews Molecular Cell Biology*, 3(7), pp. 520–530. doi: 10.1038/nrm861.

Breitenbach, H. H. *et al.* (2014) 'Contrasting Roles of the Apoplastic Aspartyl Protease APOPLASTIC, ENHANCED DISEASE SUSCEPTIBILITY1-DEPENDENT1 and LEGUME LECTIN-LIKE PROTEIN1 in Arabidopsis Systemic Acquired Resistance.', *Plant physiology*, 165(2), pp. 791–809. doi: 10.1104/pp.114.239665.

Van Breusegem, F. and Dat, J. F. (2006) 'Reactive Oxygen Species in Plant Cell Death', *Plant Physiology*, 141(2), pp. 384–390. doi: 10.1104/pp.106.078295.

Butaye, K. M. J. and Cammue, B. (2005) 'Approaches to minimize variation of transgene expression in plants', *Molecular Breeding*, 16(1), pp. 79–91. doi: 10.1007/s11032-005-4929-9.

Cai, C. F. *et al.* (2015) 'The functional analysis of OsTDF1 reveals a conserved genetic pathway for

tapetal development between rice and Arabidopsis', *Science Bulletin*. Science China Press, 60(12), pp. 1073–1082. doi: 10.1007/s11434-015-0810-3.

Canales, C. *et al.* (2002) 'EXS, a putative LRR receptor kinase, regulates male germline cell number and tapetal identity and promotes seed development in Arabidopsis', *Current Biology*, 12(20), pp. 1718–1727. doi: 10.1016/S0960-9822(02)01151-X.

Casey, J. L. *et al.* (2000) 'Green fluorescent antibodies: novel in vitro tools', *Protein Engineering*, 13(6), pp. 445–452. doi: 10.1093/protein/13.6.445.

Cecchetti, V. *et al.* (2017) 'An auxin maximum in the middle layer controls stamen development and pollen maturation in Arabidopsis', *New Phytologist*, 213(3), pp. 1194–1207. doi: 10.1111/nph.14207.

Chae, E. *et al.* (2008) 'An Arabidopsis F-box protein acts as a transcriptional co-factor to regulate floral development', *Development*, 135(7), pp. 1235–1245. doi: 10.1242/dev.015842.

Chang, H.-S. *et al.* (2012) 'NO PRIMEXINE AND PLASMA MEMBRANE UNDULATION Is Essential for Primexine Deposition and Plasma Membrane Undulation during Microsporogenesis in Arabidopsis', *Plant Physiology*, 158(1), pp. 264–272. doi: 10.1104/pp.111.184853.

Cheng, C. Y. *et al.* (2017) 'Araport11: a complete reannotation of the Arabidopsis thaliana reference genome', *The Plant Journal*, 89(4), pp. 789–804. doi: 10.1111/tpj.13415.

Cheng, H. *et al.* (2004) 'Gibberellin regulates Arabidopsis floral development via suppression of DELLA protein function.', *Development (Cambridge, England)*, 131(5), pp. 1055–1064. doi: 10.1242/dev.00992.

Chichkova, N. V *et al.* (2012) 'Plant phytaspases and animal caspases: structurally unrelated death proteases with a common role and specificity', *Physiologia Plantarum*. Wiley-Blackwell, 145(1), pp. 77–84. doi: 10.1111/j.1399-3054.2011.01560.x.

Chuang, C.-F. and Meyerowitz, E. M. (2000) 'Specific and heritable genetic interference by double-stranded RNA in Arabidopsis thaliana', *Proceedings of the National Academy of Sciences*, 97(9), pp. 4985–4990. doi: 10.1073/pnas.060034297.

Clough, S. J. and Bent, A. F. (1998) 'Floral dip: A simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana', *Plant Journal*, 16(6), pp. 735–743. doi: 10.1046/j.1365-313X.1998.00343.x.

Coll, N. S. *et al.* (2010) 'Arabidopsis type I metacaspases control cell death', *Science*, 330, pp. 1393–1397. doi: 10.1126/science.1194980.

Cutler, S. and McCourt, P. (2005) 'Dude, where's my phenotype? Dealing with redundancy in signaling networks.', *Plant physiology*, 138(June), pp. 558–559. doi: 10.1104/pp.104.900152.

Daneva, A. *et al.* (2016) 'Functions and Regulation of Programmed Cell Death in Plant Development', *Annual Review of Cell and Developmental Biology*, 32(1), pp. 441–468. doi: 10.1146/annurev-cellbio-111315-124915.

Day, C. D. *et al.* (2000) 'Transgene integration into the same chromosome location can produce alleles that express at a predictable level, or alleles that are differentially silenced', *Genes and Development*, 14(22), pp. 2869–2880. doi: 10.1101/gad.849600.

Dong, X. *et al.* (2005) 'Callose synthase (CalS5) is required for exine formation during microgametogenesis and for pollen viability in Arabidopsis', *Plant Journal*, 42(3), pp. 315–328. doi: 10.1111/j.1365-313X.2005.02379.x.

Van Doorn, W. G. *et al.* (2011) 'Morphological classification of plant cell deaths', *Cell Death and Differentiation*. Nature Publishing Group, 18(8), pp. 1241–1246. doi: 10.1038/cdd.2011.36.

Dou, X. Y. *et al.* (2011) 'WBC27, an adenosine tri-phosphate-binding cassette protein, controls pollen wall formation and patterning in arabidopsis', *Journal of Integrative Plant Biology*, 53(1), pp. 74–88. doi: 10.1111/j.1744-7909.2010.01010.x.

Drew, M. C., He, C. J. and Morgan, P. W. (2000) 'Programmed cell death and aerenchyma formation in roots', *Trends in Plant Science*, 5(3), pp. 123–127. doi: 10.1016/s1360-1385(00)01570-3.

Van Durme, M. and Nowack K., M. (2016) 'Mechanisms of developmentally controlled cell death in plants', *Current Opinion in Plant Biology*. Elsevier BV, 29, pp. 29–37. doi: 10.1016/j.pbi.2015.10.013.

Van Durme, M. and Nowack, M. (2016) 'Mechanisms of developmentally controlled cell death in plants', *Current Opinion in Plant Biology*. Elsevier Ltd, 29(February), pp. 29–37. doi: 10.1016/j.pbi.2015.10.013.

Edlund, A. F., Swanson, R. and Preuss, D. (2004) 'Pollen and Stigma Structure and Function: The Role of Diversity in Pollination', *The Plant Cell*, 16, pp. S84–S97. doi: 10.1105/tpc.015800.Pollen.

Elmore, S. (2007) 'Apoptosis: a review of programmed cell death.', *Toxicologic pathology*, 35(4), pp. 495–516. doi: 10.1080/01926230701320337.

Fagundes, D. *et al.* (2015) 'Caspases in plants: metacaspase gene family in plant stress responses', *Functional and Integrative Genomics*, 15(6), pp. 639–649. doi: 10.1007/s10142-015-0459-7.

Faro, C. and Gal, S. (2005) 'Aspartic proteinase content of the Arabidopsis genome.', *Current protein*

& peptide science, 6(6), pp. 493–500. doi: 10.2174/138920305774933268.

Feldmann, K. A. (1991) 'T-DNA insertion mutagenesis in Arabidopsis: mutational spectrum', *The Plant Journal*, 1(1), pp. 71–82. doi: 10.1111/j.1365-313X.1991.00071.x.

Ferguson, A. C. *et al.* (2016) 'Biphasic regulation of the transcription factor ABORTED MICROSPORES (AMS) is essential for tapetum and pollen development in Arabidopsis', *New Phytologist*, 213(2), pp. 778–790. doi: 10.1111/nph.14200.

Fire, A. *et al.* (1998) 'Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans', *Nature*. Macmillan Magazines Ltd., 391, p. 806. Available at: http://dx.doi.org/10.1038/35888.

Gao, H. *et al.* (2017) 'Two Membrane-Anchored Aspartic Proteases Contribute to Pollen and Ovule Development', *Plant Physiology*, 173(1), pp. 219–239. doi: 10.1104/pp.16.01719.

Gao, Y. and Zhao, Y. (2013) 'Epigenetic suppression of T-DNA insertion mutants in arabidopsis', *Molecular Plant*. Â© The Authors. All rights reserved., 6(2), pp. 539–545. doi: 10.1093/mp/sss093.

Gavrieli Y, Y, S. and Ben-Sasson SA. (1992) 'Identification of programmed cell death in situ via specific labeling of nuclear DNA fragmentation.', *The Journal of Cell Biology*, 119(3), pp. 493–501. doi: 10.1083/jcb.119.3.493.

Ge, X. *et al.* (2005a) 'An Arabidopsis aspartic protease functions as an anti-cell-death component in reproduction and embryogenesis.', *EMBO reports*, 6(3), pp. 282–288. doi: 10.1038/sj.embor.7400357.

Ge, X. *et al.* (2005b) 'An Arabidopsis aspartic protease functions as an anti-cell-death component in reproduction and embryogenesis', *EMBO reports*. Wiley-Blackwell, 6(3), pp. 282–288. doi: 10.1038/sj.embor.7400357.

Goldberg, R. B., Beals, T. P. and Sanders, P. M. (1993) 'Anther Development: Basic Principles and Practical Applications', *The Plant Cell*, 5(10), pp. 1217–1229. doi: 10.1105/tpc.5.10.1217.

Greenberg, J. T. (1996) 'Programmed cell death: a way of life for plants.', *Proceedings of the National Academy of Sciences*, 93(22), pp. 12094–12097. doi: 10.1073/pnas.93.22.12094.

Grefen, C. *et al.* (2010) 'A ubiquitin-10 promoter-based vector set for fluorescent protein tagging facilitates temporal stability and native protein distribution in transient and stable expression studies', *Plant Journal*, 64(2), pp. 355–365. doi: 10.1111/j.1365-313X.2010.04322.x.

Grienenberger, E. *et al.* (2010) 'Analysis of *TETRAKETIDE α-PYRONE REDUCTASE* Function in

Arabidopsis thaliana Reveals a Previously Unknown, but Conserved, Biochemical Pathway in Sporopollenin Monomer Biosynthesis', *The Plant Cell*, 22(12), pp. 4067–4083. doi: 10.1105/tpc.110.080036.

Grosse-Holz, F. *et al.* (2017) 'The transcriptome, extracellular proteome and active secretome of agroinfiltrated Nicotiana benthamiana uncover a large, diverse protease repertoire', *Plant Biotechnology Journal*, pp. 1068–1084. doi: 10.1111/pbi.12852.

Gu, J. N. *et al.* (2014) 'DYT1 directly regulates the expression of TDF1 for tapetum development and pollen wall formation in Arabidopsis', *Plant Journal*, 80(6), pp. 1005–1013. doi: 10.1111/tpj.12694.

Gusti, A. *et al.* (2009) 'The Arabidopsis thaliana F-Box Protein FBL17 Is Essential for Progression through the Second Mitosis during Pollen Development', *PLoS ONE*, 4(3). doi: 10.1371/journal.pone.0004780.

Haseloff KR, J. S. (1998) 'The Uses of GFP in Plants', *Green Fluorescent Protein: Properties, Applications and Protocols*, p. 40. Available at: http://www.plantsci.cam.ac.uk/Haseloff/laboratory/labdocs/LabPDFs/Haseloff97a.pdf.

Held, M. A., Boulaflous, A. and Brandizzi, F. (2008) 'Advances in fluorescent protein-based imaging for the analysis of plant Endomembranes', *PLANT PHYSIOLOGY*. American Society of Plant Biologists (ASPB), 147(4), pp. 1469–1481. doi: 10.1104/pp.108.120147.

Hellens, R. P. *et al.* (2000) 'pGreen: A versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation', *Plant Molecular Biology*, 42(6), pp. 819–832. doi: 10.1023/A:1006496308160.

Hernández-Pinzón, I. *et al.* (1999) 'Composition and role of tapetal lipid bodies in the biogenesis of the pollen coat of Brassica napus', *Planta*, 208(4), pp. 588–598. doi: 10.1007/s004250050597.

Higginson, T., Li, S. F. and Parish, R. W. (2003) 'AtMYB103 regulates tapetum and trichome development in Arabidopsis thaliana', *The Plant Journal*. Wiley-Blackwell, 35(2), pp. 177–192. doi: 10.1046/j.1365-313x.2003.01791.x.

Hird, D. L. *et al.* (1993) 'The anther-specific protein encoded by the Brassica napus and Arabidopsis thaliana A6 gene displays similarity to beta-1,3-glucanases.', *The Plant journal : for cell and molecular biology*, pp. 1023–33. doi: 10.1046/j.1365-313X.1993.04061023.x.

Hsieh, K. and Huang, H. C. A. (2007) 'Tapetosomes in Brassica Tapetum Accumulate Endoplasmic Reticulum-Derived Flavonoids and Alkanes for Delivery to the Pollen Surface', *The Plant Cell*, 19(2), pp. 582–596. doi: 10.1105/tpc.106.049049.

Huang, J. *et al.* (2013a) 'OsAP65, a rice aspartic protease, is essential for male fertility and plays a role in pollen germination and pollen tube growth', *Journal of Experimental Botany*. Oxford University Press (OUP), 64(11), pp. 3351–3360. doi: 10.1093/jxb/ert173.

Huang, J. *et al.* (2013b) 'OsAP65, a rice aspartic protease, is essential for male fertility and plays a role in pollen germination and pollen tube growth', *Journal of Experimental Botany*, 64(11), pp. 3351–3360. doi: 10.1093/jxb/ert173.

Irish, V. F. (2010) 'The flowering of Arabidopsis flower development', *The Plant Journal*, 61(6), pp. 1014–1028. doi: 10.1111/j.1365-313x.2009.04065.x.

Ito, T. *et al.* (2007) 'Arabidopsis MALE STERILITY1 encodes a PHD-type transcription factor and regulates pollen and tapetum development.', *The Plant cell*, 19(11), pp. 3549–62. doi: 10.1105/tpc.107.054536.

Ito, T. and Shinozaki, K. (2002) 'The MALE STERILITY1 gene of Arabidopsis, encoding a nuclear protein with a PHD-finger motif, is expressed in tapetal cells and is required for pollen maturation.', *Plant & cell physiology*, 43(11), pp. 1285–1292. doi: 10.1093/pcp/pcf154.

Jefferson, R. A. (1987) 'Assaying chimaeric genes in plants : the GUS gene fusion system', *Plant Molecular Biology Reporter*, pp. 387–405.

Jefferson, R. A. (1989) 'The GUS reporter gene system', *Nature*. Nature Publishing Group, 342(6251), pp. 837–838. doi: 10.1038/342837a0.

Jefferson, R. A., Kavanagh, T. A. and Bevan, M. W. (1987) 'GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants.', *The EMBO journal*, 6(13), pp. 3901–7. doi: 10.1073/pnas.1411926112.

Kawanabe, T. *et al.* (2006) 'Abolition of the tapetum suicide program ruins microsporogenesis', *Plant and Cell Physiology*, 47(6), pp. 784–787. doi: 10.1093/pcp/pcj039.

Kim, Y. Y. *et al.* (2012) 'A novel F-box protein represses endothecial secondary wall thickening for anther dehiscence in Arabidopsis thaliana', *Journal of Plant Physiology*, 169(2), pp. 212–216. doi: 10.1016/j.jplph.2011.09.006.

Kipreos, E. T. and Pagano, M. (2000) 'The F-box protein family', *Genome Biology*, 1(5), pp. 1–7.

Kleinboelting, N., Huep, G. and Weisshaar, B. (2016) 'Enhancing the GABI-Kat *Arabidopsis thaliana* T-DNA Insertion Mutant Database by Incorporating Araport11 Annotation', *Plant and Cell Physiology*, 58(December 2016), p. pcw205. doi: 10.1093/pcp/pcw205.

Kommareddi, S. *et al.* (1984) 'Nontuberculous mycobacterial infections: Comparison of the fluorescent auramine-o and Ziehl-Neelsen techniques in tissue diagnosis', *Human Pathology*. Elsevier, 15(11), pp. 1085–1089. doi: 10.1016/S0046-8177(84)80253-1.

Koncz, C. *et al.* (1992) 'T-DNA insertional mutagenesis in Arabidopsis', *Plant Molecular Biology*, 20(5), pp. 963–976. doi: 10.1007/BF00027166.

Koo, J. *et al.* (2007) 'A GUS/luciferase fusion reporter for plant gene trapping and for assay of promoter activity with luciferin-dependent control of the reporter protein stability', *Plant and Cell Physiology*, 48(8), pp. 1121–1131. doi: 10.1093/pcp/pcm081.

Krysan, P. J., Young, J. C. and Sussman, M. R. (1999) 'T-DNA as an Insertional Mutagen in Arabidopsis', *the Plant Cell*, 11(12), pp. 2283–2290. doi: 10.1105/tpc.11.12.2283.

Larkin, J. C. *et al.* (1993) 'Arabidopsis GLABROUS1 Gene Requires Downstream Sequences for Function', *the Plant Cell Online*, 5(12), pp. 1739–1748. doi: 10.1105/tpc.5.12.1739.

Lee, S. *et al.* (2004) 'Isolation and characterization of a rice cysteine protease gene , OsCP1 , using T-DNA gene-trap system', *Plant Molecular Biology*, 54, pp. 755–765.

Lema Asqui, S. *et al.* (2017) 'AtSERPIN1 is an inhibitor of the metacaspase AtMC1-mediated cell death and autocatalytic processing in planta', *New Phytologist*. doi: 10.1111/nph.14446.

Li, D.-D. *et al.* (2017) 'Gene Regulatory Network for Tapetum Development in Arabidopsis thaliana', *Frontiers in Plant Science*, 8(September), pp. 1–14. doi: 10.3389/fpls.2017.01559.

Li, H. *et al.* (2011) 'PERSISTENT TAPETAL CELL1 Encodes a PHD-Finger Protein That Is Required for Tapetal Cell Death and Pollen Development in Rice', *Plant Physiology*, 156(2), pp. 615–630. doi: 10.1104/pp.111.175760.

Li, L. *et al.* (2014) 'An anther development F-box (ADF) protein regulated by tapetum degeneration retardation (TDR) controls rice anther development', *Planta*. Springer, 241(1), pp. 157–166. doi: 10.1007/s00425-014-2160-9.

Li, N. *et al.* (2006) 'The Rice Tapetum Degeneration Retardation Gene Is Required for Tapetum Degradation and Anther Development', *The Plant Cell*, 18(11), pp. 2999–3014. doi: 10.1105/tpc.106.044107.

Li, S., Ehrhardt, D. W. and Rhee, S. Y. (2006) 'Systematic Analysis of Arabidopsis Organelles and a Protein Localization Database for Facilitating Fluorescent Tagging of Full-Length Arabidopsis Proteins', *Plant Physiology*, 141(2), pp. 527–539. doi: 10.1104/pp.106.078881.

Li, S. F., Higginson, T. and Parish, R. W. (1999) 'A novel MYB-Related gene from Arabidopsis thaliana expressed in developing anthers', *Plant and Cell Physiology*. Oxford University Press (OUP), 40(3), pp. 343–347. doi: 10.1093/oxfordjournals.pcp.a029548.

Li, Y. *et al.* (2016) 'Aspartyl protease mediated cleavage of AtBAG6 is necessary for autophagy and fungal resistance in plants', *The Plant Cell*, 28(January), p. TPC2015–00626–RA. doi: 10.1105/tpc.15.00626.

Liu, L. and Fan, X. (2013) 'Tapetum: regulation and role in sporopollenin biosynthesis in Arabidopsis', *Plant Molecular Biology*. Springer, 83(3), pp. 165–175. doi: 10.1007/s11103-013-0085-5.

Liu, X. *et al.* (2009) 'The SPOROCYTELESS/NOZZLE gene is involved in controlling stamen identity in Arabidopsis.', *Plant physiology*, 151(November), pp. 1401–1411. doi: 10.1104/pp.109.145896.

Lloyd, J. and Meinke, D. (2012) 'A Comprehensive Dataset of Genes with a Loss-of-Function Mutant Phenotype in Arabidopsis', *Plant Physiology*, 158(3), pp. 1115–1129. doi: 10.1104/pp.111.192393.

Lou, Y. *et al.* (2014) 'The tapetal AHL family protein TEK determines nexine formation in the pollen wall', *Nature Communications*. Nature Publishing Group, 5(May), pp. 1–9. doi: 10.1038/ncomms4855.

Meinke, D. W. *et al.* (1998) 'Arabidopsis thaliana: A Model Plant for Genome Analysis', *Science*, 282(5389), pp. 662–682. doi: 10.1126/science.282.5389.662.

Miki, B. and McHugh, S. (2004) 'Selectable marker genes in transgenic plants: Applications, alternatives and biosafety', *Journal of Biotechnology*, 107(3), pp. 193–232. doi: 10.1016/j.jbiotec.2003.10.011.

Millar, A. A. and Gublera, F. (2005) 'The Arabidopsis GAMYB-Like Genes, MYB33 and MYB65, Are MicroRNA-Regulated Genes That Redundantly Facilitate Anther Development', *the Plant Cell*, 17(3), pp. 705–721. doi: 10.1105/tpc.104.027920.

Molina, C. and Grotewold, E. (2005) 'Genome wide analysis of Arabidopsis core promoters.', *BMC genomics*, 6(1), p. 25. doi: 10.1186/1471-2164-6-25.

Motwani, H. V. *et al.* (2015) 'Aspartic protease inhibitors containing tertiary alcohol transition-state mimics', *European Journal of Medicinal Chemistry*. Elsevier Masson SAS, 90, pp. 462–490. doi: 10.1016/j.ejmech.2014.11.036.

Nacry, P. *et al.* (1998) 'Major chromosomal rearrangements induced by T-DNA transformation in Arabidopsis', *Genetics*, 149(2), pp. 641–650.

Nagaya, S. *et al.* (2005) 'Expression of randomly integrated single complete copy transgenes does not

vary in Arabidopsis thaliana', *Plant and Cell Physiology*, 46(3), pp. 438–444. doi: 10.1093/pcp/pci039.

Nelson, B. K., Cai, X. and Nebenführ, A. (2007) 'A multicolored set of in vivo organelle markers for co-localization studies in Arabidopsis and other plants', *Plant Journal*, 51(6), pp. 1126–1136. doi: 10.1111/j.1365-313X.2007.03212.x.

Nelson K., B., Cai, X. and Nebenführ, A. (2007) 'A multicolored set of in vivo organelle markers for co-localization studies in Arabidopsis and other plants', *The Plant Journal*. Wiley-Blackwell, 51(6), pp. 1126–1136. doi: 10.1111/j.1365-313x.2007.03212.x.

Niu, N. *et al.* (2013) 'EAT1 promotes tapetal cell death by regulating aspartic proteases during male reproductive development in rice.', *Nature communications*. Nature Publishing Group, 4, p. 1445. doi: 10.1038/ncomms2396.

Oppenheimer, D. G. *et al.* (1991) 'A myb gene required for leaf trichome differentiation in Arabidopsis is expressed in stipules', *Cell*, 67(3), pp. 483–493. doi: 10.1016/0092-8674(91)90523-2.

Ormo, M. *et al.* (1996) 'Crystal Structure of the Aequorea victoria Green Fluorescent Protein', *Science*, 273(5280), pp. 1392–1395. doi: 10.1126/science.273.5280.1392.

Otto, H.-H. and Schirmeister, T. (1997) 'Cysteine Proteases and Their Inhibitors', *Chemical Reviews*, 97(1), pp. 133–172. doi: 10.1021/cr950025u.

Papini, A., Mosti, S. and Brighigna, L. (1999) 'Programmed-cell-death events during tapetum development of angiosperms', *Protoplasma*, 207(3–4), pp. 213–221. doi: 10.1007/bf01283002.

Parish, R. and Li, S. (2010) 'Death of a tapetum: A programme of developmental altruism', *Plant Science*, 178(2), pp. 73–89. doi: 10.1016/j.plantsci.2009.11.001.

Paxson-sowders, D. M. *et al.* (2001) 'DEX1, a novel plant protein, is required for exine pattern formation during pollen development in Arabidopsis', *Plant Physiology*, 127(4), pp. 1739–1749. doi: 10.1104/pp.010517.is.

Pearce, S. *et al.* (2015) 'FlowerNet: a gene expression correlation network for anther and pollen development.', *Plant physiology*, 167(4), pp. 1717–30. doi: 10.1104/pp.114.253807.

Peek, A. S. and Behlke, M. a (2007) 'Design of active small interfering RNAs.', *Current Opinion in Molecular Therapeutics*, 9(2), pp. 110–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17458163.

Phan, H. *et al.* (2011) 'The MYB80 Transcription Factor Is Required for Pollen Development and the Regulation of Tapetal Programmed Cell Death in Arabidopsis thaliana', *The Plant Cell*, 23(6), pp. 2209–

2224. doi: 10.1105/tpc.110.082651.

Prasad, B. D. *et al.* (2009) 'Overexpression of rice (Oryza sativa L.) OsCDR1 leads to constitutive activation of defense responses in rice and Arabidopsis.', *Molecular plant-microbe interactions : MPMI*, 22(12), pp. 1635–1644. doi: 10.1094/MPMI-22-12-1635.

Prasher, D. C. *et al.* (1992) 'Primary structure of the Aequorea victoria green-fluorescent protein', *Gene*, pp. 229–233. doi: 10.1016/0378-1119(92)90691-H.

Prelich, G. (2012) 'Gene overexpression: Uses, mechanisms, and interpretation', *Genetics*, 190(3), pp. 841–854. doi: 10.1534/genetics.111.136911.

Prouse, M. B. and Campbell, M. M. (2012) 'The interaction between MYB proteins and their target DNA binding sites', *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*. Elsevier B.V., 1819(1), pp. 67–77. doi: 10.1016/j.bbagrm.2011.10.010.

Qiu, S., Adema, C. M. and Lane, T. (2005) 'A computational study of off-target effects of RNA interference', *Nucleic Acids Research*, 33(6), pp. 1834–1847. doi: 10.1093/nar/gki324.

Quilichini, T. D., Douglas, C. J. and Lacey Samuels, A. (2014) 'New views of tapetum ultrastructure and pollen exine development in Arabidopsis thaliana', *Annals of Botany*, 114(6), pp. 1189–1201. doi: 10.1093/aob/mcu042.

Quilichini, T. D., Grienenberger, E. and Douglas, C. J. (2015) 'The biosynthesis, composition and assembly of the outer pollen wall: A tough case to crack', *Phytochemistry*. Elsevier Ltd, 113(June), pp. 170–182. doi: 10.1016/j.phytochem.2014.05.002.

Ranganath, R. M. and Nagashree, R. N. (2001) 'Role of programmed cell death in development', *International Review of Cytology*, 202, pp. 159–242. doi: 10.1016/s0074-7696(01)02005-8.

Rawlings, N. D. and Barrett, A. J. (2013) 'Introduction: Aspartic and Glutamic Peptidases and Their Clans', *Handbook of Proteolytic Enzymes*, 1, pp. 3–19. doi: 10.1016/B978-0-12-382219-2.00001-6.

Reape, T. J. and McCabe, P. F. (2008) 'Apoptotic-like programmed cell death in plants', *New Phytologist*. Wiley-Blackwell, 180(1), pp. 13–26. doi: 10.1111/j.1469-8137.2008.02549.x.

Rodríguez-Serrano, M. *et al.* (2012) 'NO, ROS, and cell death associated with caspase-like activity increase in stress-induced microspore embryogenesis of barley', *Journal of Experimental Botany*, 63(5), pp. 2007–2024. doi: 10.1093/jxb/err400.

Rogers, H. J. (2005) 'Programmed Cell Death in Floral Organs: How and Why do Flowers Die?', *Annals of Botany*, 97(3), pp. 309–315. doi: 10.1093/aob/mcj051.

Sanchez-serrano, J. J. and Walker, J. M. (2014) *Arabidopsis Protocols IN Series Editor*. Part III. Edited by J. S. Sanchez-Serrano, Jose J. Humana Press. doi: 10.1007/978-1-4939-2450-9.

Sanders, P. M. *et al.* (1999) 'Anther developmental defects in Arabidopsis thaliana male-sterile mutants', *Sexual Plant Reproduction*. Springer, 11(6), pp. 297–322. doi: 10.1007/s004970050158.

Sanders, P. M. *et al.* (1999) 'Anther developmental defects in Arabidopsis thaliana male-sterile mutants', *Sexual Plant Reproduction*, 11(6), pp. 297–322. doi: 10.1007/s004970050158.

Schiefthaler, U. *et al.* (1999) 'Molecular analysis of NOZZLE, a gene involved in pattern formation and early sporogenesis during sex organ development in Arabidopsis thaliana.', *Proceedings of the National Academy of Sciences of the United States of America*, 96(20), pp. 11664–11669. doi: 10.1073/pnas.96.20.11664.

Schubert, D. *et al.* (2004) 'Silencing in Arabidopsis T-DNA Transformants: The Predominant Role of a Gene-Specific RNA Sensing Mechanism versus Position Effects', *The Plant Cell*, 16(10), pp. 2561–2572. doi: 10.1105/tpc.104.024547.

Shi, Z. H. *et al.* (2015) 'Overexpression of AtTTP affects ARF17 expression and leads to male sterility in Arabidopsis', *PLoS ONE*, 10(3), pp. 1–14. doi: 10.1371/journal.pone.0117317.

Simões, I. and Faro, C. (2004) 'Structure and function of plant aspartic proteinases', *European Journal of Biochemistry*. Wiley-Blackwell, 271(11), pp. 2067–2075. doi: 10.1111/j.1432-1033.2004.04136.x.

Singh, K. B. (1998) 'Transcriptional Regulation in Plants: The Importance of Combinatorial Control', *Plant Physiology*, 118(4), pp. 1111–1120. doi: 10.1104/pp.118.4.1111.

Smale, S. T. (2001) 'Core promoters: Active contributors to combinatorial gene regulation', *Genes and Development*, 15(19), pp. 2503–2508. doi: 10.1101/gad.937701.

Smale, S. T. and Kadonaga, J. T. (2003) 'The RNA polymerase II core promoter.', *Annual review of biochemistry*, 72(1), pp. 449–79. doi: 10.1146/annurev.biochem.72.121801.161520.

Somerville, C. and Koornneef, M. (2002) 'A fortunate choice: the history of Arabidopsis as a model plant', *Nature Reviews Genetics*, 3(11), pp. 883–889. doi: 10.1038/nrg927.

Sorensen, A. *et al.* (2002) 'A novel extinction screen in Arabidopsis thaliana identifies mutant plants defective in early microsporangial development', *Plant Journal*, 29(5), pp. 581–594. doi: 10.1046/j.0960-7412.2001.01240.x.

Sorensen, A. M. *et al.* (2003) 'The Arabidopsis ABORTED MICROSPORES (AMS) gene encodes a MYC class transcription factor', *The Plant Journal*, 33(2), pp. 413–423. doi: 10.1046/j.1365-

313x.2003.01644.x.

Swenson, E. S. *et al.* (2007) 'Limitations of Green Fluorescent Protein as a Cell Lineage Marker', *Stem Cells*, 25(10), pp. 2593–2600. doi: 10.1634/stemcells.2007-0241.

Szecsi, P. B. (1992) 'The aspartic proteases', *Scandinavian Journal of Clinical and Laboratory Investigation*, 52, pp. 5–22. doi: 10.3109/00365519209104650.

Takahashi, K. *et al.* (2008) 'Widespread tissue expression of nepenthesin-like aspartic protease genes in Arabidopsis thaliana', *Plant Physiology and Biochemistry*, 46(7), pp. 724–729. doi: 10.1016/j.plaphy.2008.04.007.

Tang, G., Galili, G. and Zhuang, X. (2007) 'RNAi and microRNA: Breakthrough technologies for the improvement of plant nutritional value and metabolic engineering', *Metabolomics*, 3(3), pp. 357–369. doi: 10.1007/s11306-007-0073-3.

Tax, F. E. and Vernon, D. M. (2001) 'T-DNA-Associated Duplication/Translocations in Arabidopsis. Implications for Mutant Analysis and Functional Genomics', *Plant Physiology*, 126(4), pp. 1527–1538. doi: 10.1104/pp.126.4.1527.

Taylor, P. E. *et al.* (1998) 'Genetic control of male fertility in Arabidopsis thaliana: Structural analyses of postmeiotic developmental mutants', *Planta*, 205(4), pp. 492–505. doi: 10.1007/s004250050348.

Thomas, S. G. and Franklin-Tong, V. E. (2004) 'Self-incompatibility triggers programmed cell death in Papaver pollen', *Nature*, 429(6989), pp. 305–309. doi: 10.1038/nature02540.

Tian, G.-W. *et al.* (2004) 'High-Throughput Fluorescent Tagging of Full-Length Arabidopsis Gene Products in Planta', *Plant Physiology*, 135(May), pp. 25–38. doi: 10.1104/pp.104.040139.

Toth, Z. E. *et al.* (2007) 'Sensitive detection of GFP utilizing tyramide signal amplification to overcome gene silencing', *Experimental Cell Research*, 313(9), pp. 1943–1950. doi: 10.1016/j.yexcr.2007.02.024.

Trobacher, C. P., Senatore, A. and Greenwood, J. S. (2006) 'Masterminds or minions? Cysteine proteinases in plant programmed cell deathThis review is one of a selection of papers published in the Special Issue on Plant Cell Biology.', *Canadian Journal of Botany*, 84(4), pp. 651–667. doi: 10.1139/b06-038.

Turk, B. and Stoka, V. (2007) 'Protease signalling in cell death: Caspases versus cysteine cathepsins', *FEBS Letters*. Wiley-Blackwell, 581(15), pp. 2761–2767. doi: 10.1016/j.febslet.2007.05.038.

Ülker, B. *et al.* (2008) 'Getting the most out of publicly available T-DNA insertion lines', *Plant Journal*,

56(4), pp. 665–677. doi: 10.1111/j.1365-313X.2008.03608.x.

Updegraff, E. P., Zhao, F. and Preuss, D. (2009) 'The extracellular lipase EXL4 is required for efficient hydration of Arabidopsis pollen', *Sexual Plant Reproduction*, 22(3), pp. 197–204. doi: 10.1007/s00497-009-0104-5.

Vartapetian, A. B. *et al.* (2011) 'A plant alternative to animal caspases: subtilisin-like proteases', *Cell Death and Differentiation*. Nature Publishing Group, 18(8), pp. 1289–1297. doi: 10.1038/cdd.2011.49.

Verma, S., Dixit, R. and Pandey, K. C. (2016) 'Cysteine proteases: Modes of activation and future prospects as pharmacological targets', *Frontiers in Pharmacology*, 7(APR), pp. 1–12. doi: 10.3389/fphar.2016.00107.

Vitha, S. (2012) 'Histochemical localization of β-glucuronidase (GUS) reporter activity in plant tissues', *Imaging*, pp. 1–6. doi: 10.1385/0-89603-302-3:185.

Vizcay-Barrena, G. and Wilson, Z. A. (2006) 'Altered tapetal PCD and pollen wall development in the Arabidopsis ms1 mutant', *Journal of Experimental Botany*, 57(11), pp. 2709–2717. doi: 10.1093/jxb/erl032.

Voinnet, O. *et al.* (2003) 'An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus', *The Plant Journal*, 33(5), pp. 949–956. doi: 10.1046/j.1365-313x.2003.01676.x.

Wang, C. *et al.* (2012) 'ANAP: An Integrated Knowledge Base for Arabidopsis Protein Interaction Network Analysis', *Plant Physiology*, 158(4), pp. 1523–1533. doi: 10.1104/pp.111.192203.

Wilkins, K. A. *et al.* (2011) 'Reactive Oxygen Species and Nitric Oxide Mediate Actin Reorganization and Programmed Cell Death in the Self-Incompatibility Response of Papaver', *Plant Physiology*. American Society of Plant Biologists (ASPB), 156(1), pp. 404–416. doi: 10.1104/pp.110.167510.

Wilson, Z. *et al.* (2001) 'The Arabidopsis MALE STERILITY1 (MS1) gene is a transcriptional regulator of male gametogenesis, with homology to the PHD-finger family of transcription factors', *Plant Journal*, 28(1), pp. 27–39. doi: 10.1046/j.1365-313X.2001.01125.x.

Wilson, Z. A. *et al.* (2001) 'The Arabidopsis MALE STERILITY1 (MS1) gene is a transcriptional regulator of male gametogenesis, with homology to the PHD-finger family of transcription factors', *The Plant Journal*, 28(1), pp. 27–39. doi: 10.1046/j.1365-313x.2001.01125.x.

Wilson, Z. A. *et al.* (2011) 'The final split: the regulation of anther dehiscence', *Journal of Experimental Botany*, 62(5), pp. 1633–1649. doi: 10.1093/jxb/err014.

Winter, D. *et al.* (2007) 'An "Electronic fluorescent Pictograph" Browser for exploring and analyzing large-scale biological data sets', *PLoS ONE*. Edited by I. Baxter. Public Library of Science (PLoS), 2(8), p. e718. doi: 10.1371/journal.pone.0000718.

Wlodawer, A., Gustchina, A. and James, M. N. G. (2013) 'Catalytic Pathways of Aspartic Peptidases', *Handbook of Proteolytic Enzymes*, 1(1982), pp. 19–26. doi: 10.1016/B978-0-12-382219-2.00002-8.

Woltering, E. J. (2004) 'Death proteases come alive', *Trends in Plant Science*. Elsevier BV, 9(10), pp. 469–472. doi: 10.1016/j.tplants.2004.08.001.

Woltering, E. J. (2010) 'Death proteases: alive and kicking', *Trends in Plant Science*. Elsevier BV, 15(4), pp. 185–188. doi: 10.1016/j.tplants.2010.02.001.

Xia, Y. *et al.* (2004) 'An Extracellular Aspartic Protease Functions in *Arabidopsis* Disease Resistance Signaling', *EMBO Journal*, 23(4), pp. 980–988. doi: 10.1038/sj.emboj.7600086.

Xiao, W. and Jang, J.-C. (2000) 'F-box proteins in Arabidopsis', *Trends in Plant Science*, 5(11), pp. 454–457. doi: 10.1016/s1360-1385(00)01769-6.

Xie, H. T. *et al.* (2014) 'Spatiotemporal production of Reactive oxygen species by NADPH Oxidase is critical for Tapetal programmed cell death and pollen development in Arabidopsis', *The Plant Cell*. American Society of Plant Biologists (ASPB), 26(5), pp. 2007–2023. doi: 10.1105/tpc.114.125427.

Xing, S. and Zachgo, S. (2007) 'Pollen lethality: a phenomenon in Arabidopsis RNA interference plants.', *Plant physiology*, 145(2), pp. 330–333. doi: 10.1104/pp.107.103127.

Xiong, S. X. *et al.* (2016) 'The transcription factors MS188 and AMS form a complex to activate the expression of CYP703A2 for sporopollenin biosynthesis in Arabidopsis thaliana', *Plant Journal*, 88(6), pp. 936–946. doi: 10.1111/tpj.13284.

Xu, J. *et al.* (2010) 'The ABORTED MICROSPORES Regulatory Network Is Required for Postmeiotic Male Reproductive Development in Arabidopsis thaliana', *The Plant Cell*, 22(1), pp. 91–107. doi: 10.1105/tpc.109.071803.

Xu, J. *et al.* (2014) 'ABORTED MICROSPORES Acts as a Master Regulator of Pollen Wall Formation in Arabidopsis', *The Plant Cell*, 26(4), pp. 1544–1556. doi: 10.1105/tpc.114.122986.

Yamada, K. *et al.* (2001) 'A slow maturation of a cysteine protease with a granulin domain in the vacuoles of senescing Arabidopsis leaves.', *Plant Physiology*, 127(4), pp. 1626–1634. doi: 10.1104/pp.010551.

Yang, C. *et al.* (2007) 'MALE STERILITY1 Is Required for Tapetal Development and Pollen Wall

Biosynthesis', *The Plant Cell*, 19(11), pp. 3530–3548. doi: 10.1105/tpc.107.054981.

Yang, J. *et al.* (2013) 'AUXIN RESPONSE FACTOR17 Is Essential for Pollen Wall Pattern Formation in Arabidopsis', *Plant Physiology*, 162(2), pp. 720–731. doi: 10.1104/pp.113.214940.

Yang, S. *et al.* (2003) 'TAPETUM DETERMINANT1 Is Required for Cell Specialization in the Arabidopsis Anther', *Plant Cell*, 15(December), pp. 2792–2804. doi: 10.1105/tpc.016618.Our.

Yang, Y. *et al.* (2014) 'Cysteine Protease 51 (CP51), an anther-specific cysteine protease gene, is essential for pollen exine formation in Arabidopsis', *Plant Cell, Tissue and Organ Culture*, 119(2), pp. 383–397. doi: 10.1007/s11240-014-0542-0.

Yao, X. *et al.* (2012) 'Overexpression of the aspartic protease ASPG1 gene confers drought avoidance in Arabidopsis', *Journal of Experimental Botany*, 63(7), pp. 2579–2593. doi: 10.1093/jxb/err433.

Yi, J. *et al.* (2016) 'Defective Tapetum Cell Death 1 (DTC1) Regulates ROS Levels by Binding to Metallothionein during Tapetum Degeneration.', *Plant physiology*, 170(3), pp. 1611–23. doi: 10.1104/pp.15.01561.

Zhang, D. *et al.* (2010) 'OsC6 , Encoding a Lipid Transfer Protein , Is Required for Postmeiotic Anther Development In Rice', 154, pp. 149–162. doi: 10.1104/pp.110.158865.

Zhang, D. *et al.* (2014) 'The cysteine protease CEP1, a key executor involved in tapetal programmed cell death, regulates pollen development in Arabidopsis.', *The Plant cell*, 26(7), pp. 2939–61. doi: 10.1105/tpc.114.127282.

Zhang, D. and Wilson, Z. A. (2009) 'Stamen specification and anther development in rice', *Chinese Science Bulletin*. Springer, 54(14), pp. 2342–2353. doi: 10.1007/s11434-009-0348-3.

Zhang, J. Z. (2003) 'Overexpression analysis of plant transcription factors', *Current Opinion in Plant Biology*, 6(5), pp. 430–440. doi: 10.1016/S1369-5266(03)00081-5.

Zhang, W. *et al.* (2006) 'Regulation of Arabidopsis tapetum development and function by DYSFUNCTIONAL TAPETUM1 (DYT1) encoding a putative bHLH transcription factor.', *Development*, 133(16), pp. 3085–3095. doi: 10.1242/dev.02463.

Zhang, Z. B. *et al.* (2007) 'Transcription factor AtMYB103 is required for anther development by regulating tapetum development, callose dissolution and exine formation in Arabidopsis', *The Plant Journal*. Wiley-Blackwell, 52(3), pp. 528–538. doi: 10.1111/j.1365-313x.2007.03254.x.

Zhao, D.-Z. *et al.* (2002) 'Excess microsporocytes1 gene encodes a putative leucine-rich repeat receptor protein kinase that controls somatic and reproductive cell fates in the Arabidopsis anther',

Genes & Development, 16, pp. 2021–2031. doi: 10.1101/gad.997902.pollen.

Zhu, E. *et al.* (2015) 'The DYT1-interacting proteins bHLH010, bHLH089 and bHLH091 are redundantly required for Arabidopsis anther development and transcriptome', *Plant Journal*, 83(6), pp. 976–990. doi: 10.1111/tpj.12942.

Zhu, J. *et al.* (2008) 'Defective in Tapetal Development and Function 1 is essential for anther development and tapetal function for microspore maturation in Arabidopsis', *Plant Journal*, 55(2), pp. 266–277. doi: 10.1111/j.1365-313X.2008.03500.x.

APPENDICES

Appendix 1

List of primers used in this project.

Primers	Primers Sequence (5'-3')	Purpose
AT1G25510_Gibson_ Forward_Primer	ttttctgattaacactcgagccaccATGTCTCCTA ACTACTCCTTCT	Amplification of AT1G25510 gene
AT1G25510_Gibson_ Reverse_Primer	AGTTGGATATCTCGAGTCGCATTTATTAG ACGAGAATCCAAT	Amplification of AT1G25510 gene
AT2G23945_Gibson_ Forward_Primer	ttttctgattaacactcgagccaccATGGAATGTT CTTTCCAAACTTCA	Amplification of AT2G23945 gene
AT2G23945_Gibson_ Reverse_Primer	AGTTGGATATCTCGAGTCCTTTGATGGG GAATAGTCAC	Amplification of AT2G23945 gene
AT4G30030_Gibson_ Forward_Primer	ttttctgattaacactcgagccaccATGGCCATCA CTCTACTACC	Amplification of AT4G30030 gene
AT4G30030_Gibson_ Reverse_Primer	AGTTGGATATCTCGAGTCAGAATCAATAA TCTCACAATCAGTT	Amplification of AT4G30030 gene
AT5G10760_Gibson_ Forward_Primer	ttttctgattaacactcgagccaccATGAGCATAA TGAGGAATTTCTTG	Amplification of AT5G10760 gene
AT5G10760_Gibson_ Reverse_Primer	AGTTGGATATCTCGAGTCACAACCATTG GGAGCAAAACC	Amplification of AT5G10760 gene
AT3G20015_Gibson_ Forward_Primer	ttttctgattaacactcgagccaccATGCTCCTCC CTCTCTTCTT	Amplification of AT3G20015 gene
AT3G20015_Gibson_ Reverse_Primer	AGTTGGATATCTCGAGTCGCAAACGTTG GGGCCAAAA	Amplification of AT3G20015 gene
AT5G19120_Gibson_ Forward_Primer	ttttctgattaacactcgagccaccATGGCTTCTT CTTCTTGTTTGAAT	Amplification of AT5G19120 gene
AT5G19120_Gibson_ Reverse_Primer	AGTTGGATATCTCGAGTCCAACGAAGTT GAATCAGAGCG	Amplification of AT5G19120 gene
AT2G21430_Gibson_ Forward_Primer	ttttctgattaacactcgagccaccATGGATTATC ATCTTAGGGTTTTG	Amplification of AT2G21430 gene
AT2G21430_Gibson_ Reverse_Primer	AGTTGGATATCTCGAGTCAGAAGTGGTA GCAGCGACG	Amplification of AT2G21430 gene
AT3G16210 gene specific F primer (At3g16210_CDS_F)	ATGTCGAAATTTCTCCCGGA	Amplification of AT3G16210 Gene
AT3G16210 gene specific R primer (At3g16210_CDS_NS_R)	AATAAATATGCGATTAACAATCTTAAGAG TC	Amplification of AT3G16210 Gene
PP2A3 Forward House Keeping Primer (PP2A3_1355F)	TCCGTGAAGCTGCTGCAAAC	House-keeping gene

PP2A3 Reverse House Keeping Primer (PP2A3_1677R)	CACCAAGCATGGCCGTATCA	House-keeping gene
AT1G25510 forward expression primer (1g25510_1041F)	GCGGCGAGTTGCTTCAGATT	AT1G25510 gene expression
AT1G25510 reverse expression primer (1g25510_1312R)	CGGCAAAGCCAACATCTTCC	AT1G25510 gene expression
AT2G23945 forward expression primer (2g23945_486F)	AAACGGACATTGCGGCTCAT	AT2G23945 gene expression
AT2G23945 reverse expression primer (2g23945_715R)	CGGCAAGTGATGTGGGTTTG	AT2G23945 gene expression
AT4G30030 forward expression primer (4g30030_602F)	ATAGCGGCGTGCTCGGTTTA	AT4G30030 gene expression
AT4G30030 reverse expression primer (4g30030_885R)	GTCGATGACCGTGCCTCCTT	AT4G30030 gene expression
AT5G10760 forward expression primer (5g10760_1165F)	ACGTCGGGATACGGGCTTTT	AT5G10760 gene expression
AT5G10760 reverse expression primer (5g10760_1354R)	AGATGGCGGGAAGGTCATCA	AT5G10760 gene expression
AT3G20015 forward expression primer (3g20015_437F)	ACCACCACCACCGTCTCCAT	AT3G20015 gene expression
AT3G20015 reverse expression primer (3g20015_731R)	TCAAAAACCGGGTCGGATTG	AT3G20015 gene expression
AT5G19120 forward expression primer (5g19120_186F)	ACGGGGTCGTTTTCCCTGTT	AT5G19120 gene expression
AT5G19120 reverse expression primer (5g19120_440R)	CGGCGTTTTGGTCTTTACGC	AT5G19120 gene expression
AT2G21430 reverse expression primer (SALK_095569_F)	CTGACTCGTTCTGAGTTTCGG	AT2G21430 gene expression
AT2G21430 reverse expression primer (AT2g21430 EXP R)	CCTGTAGTGCTAAAACTCCAGCA	AT2G21430 gene expression
AT3G16210 forward expression primer (At3g16210_454F)	CGGCTTGATGTTTCCACGGCG	AT3G16210 gene expression
AT3G16210 reverse expression primer (AT3G16210_667R)	CCGGGAGAGGGAGTTTTCGGT	AT3G16210 gene expression
2g23945_976F	TCGGGCACACTCTACACTTG	AT2G23945 gene insertion

pUBQ10 F primer	CGAAGACGATTTTCTGGGTTT	pUBC cloning vector
LBb1.3	ATTTTGCCGATTTCGGAAC	SALK knockout left border primer
SALK_134917 LP (AT2G23945)	TTCATTCGCCATGGAACTATC	SALK 135523 and SALK 134919 left border primer
SALK_134917 RP (AT2G23945)	TCCATGATAGTTAACTGCGGC	SALK 135523 and SALK 134919 right border primer
AT2G23945 RNAi forward primer (RNAiF2g23945)	ATGGCGATGAAGCTGATACAC	Amplification of AT2G23945RNAi Fragment
AT2G23945 RNAi reverse primer (RNAiR2g23945)	AATATCTGTCAGGTGTTTGATATGGTCT	Amplification of AT2G23945RNAi Fragment
ROligo_New_CompF	GTTTTCCCAGTCACGACGTT	Amplification of a pCR8/GW/TOPO fragment
ROligo_New_CompR	GAGCTGCCAGGAAACAGCTA	Amplification of a pCR8/GW/TOPO fragment
PK7_F2	gccaaaatttatgctactctaagga	RNAi vector forward primer
PH7GW1_R1	gcttagcgttcattgaatttg	RNAi vector reverse primer
PK7_R2	ttgatggccatagggttta	RNAi vector reverse primer
2G23945PRO F4	ATGTGTTGGTTTTGGTCTGTTA	AT2G23945 forward primer, short promoter amplification
2G23945PRo F5	caattgagttttgcagGTGA	AT2G23945 forward primer, medium promoter amplification
2G23945PRO F1	TGCGAGTTCTAGCGAAAGAC	AT2G23945 forward primer, long promoter amplification
2G23945Pro R1	TGTTGTTGAAACGGTACAGAAACAT	AT2G23945 reverse primer, all promoter amplification
2G23945 TF R1	CTTTGATGGGGAATAGTCATCG	AT2G23945 reverse primer, translational fusion amplification
pCR8 GW 2.1 R	TCGTTGCAACAAATTGATGAGCAATTA	pCR8 TOPO Cloning
GFP_PGWB5_R	AAGTCGTGCTGCTTCATGTG	GFP pGWB5 Cloning
M13F (-20)	TGTAAAACGACGGCCAG	pCR8/GW/TOPO vector sequencing
YFP_NY R	GGCCATGATATAGACGTTGTGGC	GFP and YFP-tagged protein amplification
GUS pGWB3 reverse primer (GUSpGWB3_138R)	AGCAATTGCCCGGCTTTCTT	GUS Expression Study

Appendix 2

List of destination vectors generated in this work. Keys: Kan, Kanamycin; Bar, Basta; Spec, Spectinomycin.

Name	Gene	Selection bacteria	Selection plant
pGWB5: AT3G16210:GFP	AT3G16210	Kan	Kan
pUBQ10:AT5G10760:YFP	AT5G10760	Kan	Bar
pUBQ10:AT2G21430:YFP	AT2G21430	Kan	Bar
pUBQ10:AT3G20015:YFP	AT3G20015	Kan	Bar
pUBQ10:AT5G19120:YFP	AT5G19120	Kan	Bar
pUBQ10:AT1G25510:YFP	AT1G25510	Kan	Bar
pUBQ10:AT4G30030:YFP	AT4G30030	Kan	Bar
pUBQ10:AT2G23945:YFP	AT2G23945	Kan	Bar
pGWB3: AT2G23945:GUS (-437bp)	AT2G23945	Kan	Kan
pGWB3: AT2G23945:GUS (-980bp)	AT2G23945	Kan	Kan
pGWB3: AT2G23945:GUS (-1572bp)	AT2G23945	Kan	Kan
pGKGWG:AT2G23945:GFP (1811bp)	AT2G23945	Kan	Kan
pGKGWG:AT2G23945:GFP (2354bp)	AT2G23945	Kan	Kan
pGKGWG:AT2G23945:GFP (2946bp)	AT2G23945	Kan	Kan
pK7GWIWG2(I): AT2G23945 RNAi	AT2G23945	Spec	Kan

Appendix 3

Full Length Sequences of the Transcripts

AT1G25510 cDNA full length sequence

Exons: 1; Transcript length: 1,663 bps

ACTGTTAACGGTCGGAATCTAAAAAATGTCTCCTAACTACTCCTTCTTCTTCTTCATCTTC
TTCCTCACTTCACACTCTTCTGTTTTCTCCAGAATCTTACCAGAGACCTCCACAACCACT
ACCTCAATTCTCAATGTCGCTGACTCCATCCACAGAACCAAGTATACGTCGTCGTTTCG
GTTAAACCAGCAAGAAGAACAGACTCACTCTGCTTCTTCTTCTTTCAGCTTGCAGCTTCA
TTCACGAGTTTCCGTCCGAGGAACTGAACACTCAGACTATAAGTCACTCACACTAGCCA
GACTCAACCGTGACACAGCCCGAGTCAAATCTCTCATAACCCGTTTAGATCTAGCCATC
AACAACATATCCAAAGCAGATCTCAAACCCATCAGCACAATGTACACCACTGAGGAACA
AGATATCGAAGCTCCTTTAATCTCCGGCACGACTCAAGGAAGTGGCGAGTACTTCACAC
GCGTCGGTATCGGAAAACCGGCGCGTGAAGTTTATATGGTTCTAGATACTGGAAGTGA
CGTGAACTGGCTACAATGTACTCCTTGCGCTGACTGTTACCACCAAACCGAACCAATCT
TCGAGCCATCGTCGTCTTCTTCTTACGAGCCTCTCTCTTGTGACACTCCACAGTGCAAT
GCCCTTGAAGTCTCAGAATGTAGAAACGCCACGTGTCTTTACGAGGTTTCTTACGGTGA
TGGTTCTTACACCGTTGGAGATTTCGCCACCGAGACTCTCACAATCGGCTCAACTCTAG
TTCAAAACGTCGCCGTTGGTTGTGGACACAGCAACGAAGGGTTGTTCGTCGGCGCCGC
CGGTTTACTTGGACTCGGCGGTGGTTTGTTAGCTCTTCCGTCGCAGCTCAACACGACGT
CGTTCTCGTACTGTCTCGTTGACCGTGACTCAGACTCGGCTTCGACCGTGGATTTCGGG
ACTTCTCTTTCTCCTGACGCCGTCGTGGCTCCGTTGCTACGTAACCACCAGCTCGACAC
GTTTTACTACCTAGGCCTCACCGGAATAAGCGTCGGCGGCGAGTTGCTTCAGATTCCTC
AGTCTTCTTTCGAGATGGACGAGTCAGGCAGCGGTGGGATCATCATCGACTCAGGAAC
CGCCGTGACTCGTCTCCAGACCGAAATCTATAACTCGCTTAGAGACTCGTTCGTGAAAG
GAACGTTGGACTTGGAAAAAGCAGCCGGGGTGGCGATGTTTGACACGTGTTATAACCT
TTCGGCTAAGACGACGGTAGAAGTCCCGACCGTGGCTTTTCATTTCCCCGGCGGGAAG
ATGTTGGCTTTGCCGGCGAAGAACTACATGATTCCGGTTGACTCGGTGGGGACATTTT
GTTTGGCATTTGCACCGACGGCGAGTTCACTAGCGATTATCGGAAACGTGCAGCAACA
AGGGACACGTGTCACTTTCGATTTGGCAAATTCTCTCATTGGATTCTCGTCTAATAAATG
CTAAGCTGTTTTCATCAAAAAAAACAAAAATGCTAAGCTGTTTTTTCTATAAATATATTTG
AGTTACAATGAAAAGGAAAAGTCACTTAAAAGATAGTAATGTTAAGTTAGAGTACTGTA
CTACTGTTTGGTTGTGTTTGTTTCTTGTACTTGTTGTTGCAATGCATAGAGTGAAAAAGA
GTAAAAAAAA

AT2G23945 cDNA full length sequence

Exons: 1 Transcript length: 1,377 bps

ATGGAATGTTCTTTCCAAACTTCACTCTTACTCTTCATCACTGTATCATACTTTG
TAGTAACAGAATCCATTAAACCTAACCGAATGGCGATGAAGCTGATACACCGT
GAATCAGTGGCCCGTCTCAACCCAAATGCTCGAGTTCCGATCACTCCAGAAGA
CCATATCAAACACCTGACAGATATTTCATCAGCCCGGTTTAAGTATCTACAGAA
TTCGATTGACAAAGAGCTAGGCAGCAGCAATTTCCAAGTTGATGTGGAACAAG
CAATCAAAACATCATTGTTTCTTGTGAATTTCTCCGTAGGACAGCCACCGGTGC
CGCAGTTAACTATCATGGATACTGGAAGCTCACTTTTATGGATCCAATGTCAAC
CCTGCAAACACTGCTCATCAGATCATATGATCCATCCTGTCTTCAATCCAGCAT
TGTCTTCAACATTTGTAGAGTGTTCCTGCGATGATCGTTTTTGCAGGTATGCAC
CAAACGGACATTGCGGCTCATCAAACAAGTGTGTTTACGAACAAGTGTATATAA
GCGGCACTGGGTCAAAAGGCGTTCTAGCTAAAGAACGGTTAACCTTCACAACA
CCTAACGGTAACACTGTTGTGACCCAACCTATTGCTTTTGGGTGTGGCTACGA
GAATGGAGAGCAGTTGGAAAGTCATTTCACTGGTATTTTGGGGTTAGGAGCCA
AACCCACATCACTTGCCGTACAGCTTGGGTCTAAATTCTCTTACTGCATTGGTG
ATTTAGCTAACAAAAACTACGGTTACAATCAGCTGGTTCTGGGAGAGGACGCG
GATATTCTCGGCGATCCAACACCCATTGAATTTGAGACAGAGAATAGTATCTAC
TATATGAATCTAGAAGGGATTAGTGTTGGGGATACACAACTCAATATCGAACCA
GTGGTATTCAAAAGAAGAGGACCAAGAACAGGGGTCATACTAGACTCGGGCA
CACTCTACACTTGGCTAGCAGATATAGCCTACAGAGAACTCTACAATGAAATAA
AATCCATTCTGGATCCAAAATTAGAGAGATTCTGGTTTAGAGACTTCTTGTGCT
ACCATGGGAGAGTGAGCGAAGAGCTCATAGGCTTCCCTGTGGTGACATTTCAT
TTTGCAGGCGGAGCTGAGCTGGCTATGGAGGCCACAAGCATGTTTTATCCATT
ATCAGAACCGAATACGTTTAACGTCTTTTGTATGTCAGTGAAGCCAACCAAAGA
ACATGGAGGGGAATATAAGGAATTCACAGCTATAGGGCTAATGGCTCAACAAT
ACTATAATATTGGTTACGACCTCAAAGAAAAAAACATCTATCTCCAGAGAATCG
ACTGTGTTCAACTCGATGACTATTCCCCATCAAAGTAA

AT4G30030 cDNA full length sequence

Exons: 1 Transcript length: 1,275 bps

ATGGCCATCACTCTACTACCATTACGTATTCGTGCCTCCCTTCTTATCATTATAT
TCGCTCTCACTTGTTCAAAGGAATGCACCTCGCATTCCCGACTTACGCTAAGAA
CCAAAACCCAAGAATCATCGAAAATCAAAATCGGTTACTTGCATTCCAAGTCAA
CCCCCGCTTCTCGCCTAGATAACTTGTGGACCGTCTCACATGTCACTCCTATTC
CGAATCCAGCAGCATTCCTTGCAAATATTTCTATCGGCAATCCACCTGTCCCGC
AACTCTTACTCATAGATACAGGTAGTGACCTAACTTGGATCCATTGTCTTCCCT
GCAAATGTTACCCTCAAACAATCCCTTTTTTCCACCCTTCAAGATCTTCCACTTA
CCGAAACGCTTCTTGTGTATCCGCACCACACGCCATGCCCCAAATCTTTCGCGA
CGAAAAGACTGGAAACTGCCAATATCACCTACGCTACCGCGACTTTTCTAACAC
CAGAGGCATTTTAGCTGAGGAAAAACTCACATTTGAAACATCTGACGACGGTTT
AATCTCTAAACAAAACATTGTTTTCGGATGCGGACAAGATAACTCCGGTTTCAC
CAAATATAGCGGCGTGCTCGGTTTAGGTCCAGGTACATTCTCCATAGTAACCC
GCAATTTCGGCTCTAAGTTCTCGTACTGCTTCGGCAGCCTAACCAATCCGACTT
ACCCTCACAACATTCTAATCCTCGGAAACGGCGCCAAAATCGAAGGCGATCCA
ACACCTTTACAAATATTCCAAGACCGTTACTACCTTGACCTCCAAGCAATCTCCT
TTGGAGAGAAACTTCTCGATATCGAACCCGGAACTTTCCAGAGATATCGGTCC
CAAGGAGGCACGGTCATCGACACAGGCTGTTCTCCAACGATTCTAGCCAGAGA
AGCTTACGAAACTCTCTCTGAAGAAATTGATTTTCTTCTAGGAGAAGTTCTCAG
ACGCGTTAAAGATTGGGATCAATACACGACTCCTTGCTACGAAGGGAATCTTA
AACTTGATCTCTACGGCTTTCCCGTCGTCACGTTCCATTTCGCCGGCGGCGCT
GAGCTGGCGTTAGACGTTGAGAGTTTGTTCGTTAGTAGTGAAAGTGGAGATAG
CTTTTGCTTGGCTATGACTATGAACACGTTTGATGATATGAGCGTTATTGGTGC
CATGGCTCAGCAGAATTATAATGTTGGCTATAATCTTCGAACCATGAAGGTTTA
TTTCCAGAGAACTGATTGTGAGATTATTGATTCTTAA

AT5G10760 cDNA full length sequence

Exons: 2 Transcript length: 1,766 bps

GAGTAACACTTTTCTTGTTGTGAGAGTAAGAGGGAGGTTAAAATGAGCATAATGAGGAA
TTTCTTGAGTATGATAATTATGCTTTGCGTATGTCTCAATTGGTGTTTTGCTGAAGGAGC
TGAGAAAAGCGATAGTGGAAAGGTACTTGATTCTTACACCATTCAAGTCAGCTCTCTGT
TCCCTTCATCATCATCTTGTGTCCCTTCTTCAAAAGCATCTAATACCAAGTCGTCGCTGC
GCGTGGTGCACATGCATGGTGCGTGCTCGCATCTTAGCAGCGATGCGAGAGTCGACCA
TGACGAAATCATCAGACGCGATCAAGCACGTGTGGAGTCCATCTACTCAAAATTGTCAA
AGAATAGCGCAAATGAAGTCAGCGAAGCCAAGTCAACGGAGCTACCGGCTAAAAGTGG
AATCACACTCGGATCAGGAAACTACATCGTAACAATCGGAATCGGGACACCGAAGCAC
GATCTATCGCTGGTCTTCGACACCGGAAGTGACCTGACATGGACTCAATGTGAGCCAT
GTCTAGGAAGTTGCTATTCACAAAAGGAACCAAAATTTAACCCTTCTTCGTCTTCTACTT
ACCAAAACGTCTCGTGCTCATCTCCAATGTGTGAAGACGCTGAGAGTTGTTCGGCTTCC
AACTGCGTATACAGTATCGTTTACGGCGATAAGTCGTTTACCCAAGGATTTCTTGCAAA
GGAGAAATTTACCCTAACGAACTCTGATGTTTTGGAGGATGTTTACTTTGGTTGCGGTG
AAAACAACCAAGGACTTTTCGACGGTGTTGCTGGGCTTCTCGGCCTTGGCCCTGGCAA
ACTCTCACTACCGGCGCAAACAACAACGACCTACAACAATATCTTCTCCTACTGTCTACC
TTCTTTCACAAGTAACAGTACTGGCCATCTTACCTTTGGATCCGCTGGAATCTCTGAATC
CGTTAAATTCACACCGATTTCTTCCTTCCCAAGCGCATTCAATTATGGCATCGATATCAT
AGGTATCAGTGTAGGCGACAAGGAATTAGCGATTACTCCAAACTCATTCTCGACCGAAG
GTGCTATAATCGACTCAGGCACAGTGTTTACTCGCCTCCCTACCAAAGTCTACGCGGAA
CTCCGCAGTGTTTTTAAGGAAAAGATGTCGAGTTACAAGAGTACGTCGGGATACGGGC
TTTTTGACACATGTTACGATTTCACTGGCTTGGATACGGTGACATACCCGACGATTGCA
TTCTCCTTCGCCGGTAGCACTGTCGTAGAACTTGATGGATCAGGGATTTCGTTGCCGAT
CAAGATATCACAGGTTTGTTTGGCGTTTGCAGGGAATGATGACCTTCCCGCCATCTTTG
GAAACGTTCAGCAGACGACGCTAGATGTTGTGTACGACGTGGCTGGTGGACGAGTCG
GTTTTGCTCCCAATGGTTGTTAATTAGCTACTTGAACAAAAAGAAGAGGAGAGGGTTTT
GTAATTTGAGTAAGGGTTTGGCTCGTAAATTTAGCTCCTATTTATATATAGTACGTGTGC
TAAGTAAATTGCCCCAAAAATAATTTTGTGTGTGCGCGGCAATTGCTTTCTTATATATAT
ACGACATGTAGATTTGCTTTCTTATATACTTCAATCCTGACTTGGAAGTAAGTTTTTTAG
TAACGTCAAACATGCTTCAACATTTGCTTAATGTACAAGTCCGATTTAAGCTACCAGAAT
CATCTTTTTCCTTGAGTGTAACCTTTTTTACTCGAAAAAGCATAAATTCATTA

AT3G20015 cDNA full length sequence

Exons: 1 Transcript length: 1,892 bps

ATATCGACCCATCACTTCCCAAATCACAAGTCTCAAGTGGTCTCTCTCTCTCTCTCTCTC
AATAAATAAATCTCAAAAAAATCTCATCTTTCAGAGAGAAGCAAGACACCAAAAACTTAG
TCTGACATCAAAACCCAAAATCCAAGAAGAAAGAAAGAGAGAGTCTTCACGCTCACACT
GGCTTTGGCTTTGGCTCTGGCGTTGTAAACAATGCTCCTCCCTCTCTTCTTCTTCTTTCT
CCATCTCCATCTCCATCTCTCTTCTTCTTCTTCAATCTCCTTCCCTGATTTTCAAATCATC
GACGTTTTACAACCTCCTCTCACCGTCACTGCAACTCTTCCCGATTTCAACAACACCCAT
TTCTCCGATGAATCCAGTTCTAAGTACACTCTCCGTCTTCTCCACCGCGACAGATTCCCT
TCAGTCACTTACCGTAACCACCACCACCGTCTCCATGCTCGTATGCGCAGAGACACAGA
CAGAGTCTCCGCCATTCTCCGCCGCATTTCTGGGAAAGTTATACCTTCTTCCGATTCTA
GATACGAGGTTAACGATTTTGGGTCCGATATTGTTTCCGGTATGGATCAAGGAAGTGGT
GAGTATTTCGTCAGAATCGGCGTCGGAAGTCCTCCGAGAGATCAGTACATGGTCATTG
ATTCCGGTAGCGATATGGTTTGGGTTCAATGTCAGCCTTGTAAGCTCTGTTACAAACAA
TCCGACCCGGTTTTTGACCCGGCTAAATCCGGGTCTTACACTGGAGTCTCTTGTGGCTC
TTCCGTCTGTGACCGGATTGAAAACTCCGGTTGTCACTCCGGTGGGTGCCGTTACGAG
GTCATGTACGGAGACGGGTCTTACACGAAGGGAACGTTAGCTCTCGAAACGCTGACGT
TTGCTAAGACGGTGGTTAGAAACGTTGCGATGGGATGTGGTCACCGGAATAGAGGAAT
GTTTATTGGAGCTGCCGGTTTATTGGGTATCGGAGGCGGTTCAATGTCGTTTGTGGGTC
AATTAAGCGGTCAAACCGGTGGTGCGTTCGGTTATTGTTTAGTGAGTCGTGGCACTGA
CTCGACCGGTTCATTGGTATTCGGGCGGGAAGCGTTGCCCGTTGGTGCCTCGTGGGTC
CCGTTGGTCCGGAACCCGCGGGCTCCAAGTTTTTACTACGTTGGGCTCAAGGGCCTTG
GAGTAGGAGGCGTCCGGATTCCGTTACCGGATGGTGTTTTTGACCTGACTGAAACCGG
TGACGGTGGTGTGGTTATGGATACCGGAACAGCCGTCACAAGACTTCCCACGGCGGCT
TACGTCGCTTTTCGAGATGGATTCAAGTCGCAGACAGCGAATTTGCCGCGGGCTTCCG
GTGTATCGATATTCGACACGTGTTATGATCTGTCAGGGTTTGTTTCGGTACGAGTCCCG
ACCGTGTCGTTCTATTTCACGGAAGGTCCGGTTTTGACATTACCGGCGAGGAATTTTCT
GATGCCGGTGGATGATTCGGGAACGTATTGTTTTGCGTTTGCGGCGTCTCCGACGGGT
TTGTCAATAATAGGGAACATACAACAAGAAGGGATTCAAGTATCTTTCGACGGAGCCAA
TGGCTTTGTTGGTTTTGGCCCCAACGTTTGCTAGCGGTTAAGATAGAGAAATTACGATT
TTTGTTTGCTAGTAAGTAAAAAGCTTTTGTCTCAAAACTATATATGTATTTGTCTCTTTTT
TTCTTTCTTATTAGTTTGGTTTGAGTTTTTAATAAATCTGATTAGTATAGAGAGATAACAA
AAATGTATATGTTTTGACCAAATTTGTCTTTTGATTTCTGATTGGTTTGGAATCTATTATT
TTGAGTTACTGTTTGGTTTGTTGTATTGTATGAATGAAGTGAAACACCTTTTAGAGCACC
A

AT5G19120 cDNA full length sequence

Exons: 1 Transcript length: 1,420 bps

CCAGCGTCGGCCCACCGCATCCCAAACCATAACGCTTAATCCTTAAAAACCGT
CCAGATCCATTTCCTTCCTCTAAAACTCATCACACACACTTCTAATGGCTTCTTC
TTCTTGTTTGAATCTCTTCTTCTTCTCTTTTCTCTCCGCTCTAATCATTTCTAAAT
CTCAGATCTCTGATTCCGTTAACGGGGTCGTTTTCCCTGTTGTTAAAGACCTTC
CAACCGGTCAATACCTCGCACAGATTCGACTCGGTGACTCACCGGATCCAGTC
AAGCTCGTAGTCGATCTGGCCGGCTCAATTCTCTGGTTTGATTGTTCCTCAAGA
CATGTTTCCTCGTCAAGGAACCTAATATCCGGTAGCTCTAGCGGCTGCTTAAAG
GCAAAAGTTGGAAATGAACGAGTATCATCATCATCATCATCGCGTAAAGACCAA
AACGCCGATTGCGAATTACTTGTTAAAAACGACGCCTTTGGGATTACGGCGAG
AGGAGAGCTTTTCAGTGACGTCATGTCGGTTGGCTCTGTTACTTCTCCGGGAA
CCGTTGATCTACTTTTCGCTTGTACTCCACCGTGGCTGTTACGTGGGCTCGCTA
GTGGAGCCCAAGGAGTAATGGGCTTAGGAAGGGCCCAAATCTCACTCCCTTCA
CAACTCGCCGCGGAAACTAACGAACGTCGTCGGTTAACTGTTTACCTGTCGCC
GTTGAACGGTGTCGTATCAACGAGTTCGGTGGAGGAGGTTTTTGGAGTGGCTG
CGTCTAGATCGCTGGTTTACACGCCTTTGTTAACCGGTTCGAGTGGCAACTAC
GTAATTAACGTGAAATCGATTAGAGTCAACGGGGAAAAGCTCTCCGTGGAAGG
TCCGTTGGCGGTGGAGCTGAGCACGGTGGTTCCTTACACGATTTTAGAGAGCT
CGATCTACAAGGTGTTCGCGGAAGCTTACGCCAAAGCTGCGGGTGAAGCTACA
TCGGTGCCTCCTGTCGCACCGTTTGGGCTCTGCTTCACTAGTGACGTAGATTTT
CCGGCGGTAGATCTAGCGTTGCAGAGTGAGATGGTGCGGTGGAGGATTCATG
GGAAGAATCTGATGGTGGATGTTGGTGGTGGAGTCCGATGCTCGGGGATTGT
TGACGGCGGGTCAAGTCGAGTCAACCCGATTGTAATGGGTGGGCTACAATTAG
AAGGCTTTATATTGGACTTTGATTTGGGTAACTCAATGATGGGCTTTGGACAAA
GAACACGCTCTGATTCAACTTCGTTGTAATACAGAAACCTTTGTAATTGTAACC
TTCTAATTATCTGGTAGATATCCTCAATTTTGATCTCTTTTGAGTGTACTCTTTTC
TTACAATACCTGCTACAAGCAATCATGCATTTGTACAACTAATAGATCTAATTTA
AGAAACAGCTATTACTTAAATGAAGT

AT2G21430 cDNA full length sequence

Exons: 3 Transcript length: 1,203 bps

AGCTCTGTTTTCTTCTCTCCATAATTTTTTAAAAAGTCATGGATTATCATCTTAG
GGTTTTGTTCTCTGTTTCTCTTATCTTTGTGTTCGTCTCTGTTTCGGTTTGTGGA
GACGAAGATGTGTTGATCAGGCAAGTGGTTGATGAAACAGAGCCGAAAGTTTT
GTCATCGGAGGATCACTTTACTCTCTTCAAGAAGAAATTCGGCAAGGTTTACGG
TTCCATTGAGGAGCATTATTACAGATTCTCTGTTTTCAAAGCAAATCTCCTGCG
AGCGATGCGTCACCAGAAGATGGATCCTTCTGCTCGTCACGGCGTAACGCAGT
TTTCGGATCTGACTCGTTCTGAGTTTCGGAGAAAGCATTTAGGGGTTAAAGGT
GGGTTTAAGCTTCCTAAAGACGCTAATCAAGCCCCGATTCTCCCTACTCAGAAT
CTTCCGGAGGAGTTTGATTGGAGAGATCGCGGCGCGGTTACTCCCGTGAAAAA
CCAGGGATCTTGTGGTTCTTGCTGGAGTTTTAGCACTACAGGTGCCCTTGAAG
GTGCTCACTTCCTCGCTACTGGCAAACTCGTTAGTCTCAGTGAACAACAGCTCG
TCGATTGTGATCACGAGTGTGATCCAGAAGAGGAAGGTTCATGTGATTCTGGT
TGTAACGGTGGGCTTATGAACAGTGCCTTTGAATACACGCTTAAAACCGGAGG
GCTTATGCGAGAGAAAGATTATCCTTACACCGGTACTGATGGAGGGAGCTGCA
AGTTAGACAGGTCTAAGATCGTTGCCTCTGTCTCCAACTTTAGTGTTGTGTCCA
TCAACGAAGATCAAATAGCTGCAAACCTTATCAAAAACGGTCCTCTAGCTGTGG
CGATCAATGCAGCTTATATGCAAACCTACATTGGAGGAGTCTCGTGTCCTTACA
TATGCTCGAGGCGGCTCAACCACGGTGTGTTGTTGGTAGGTTATGGCTCGGCT
GGGTTTTCTCAAGCAAGGTTAAAGGAAAAACCGTATTGGATCATCAAGAACTC
GTGGGGTGAGTCTTGGGGTGAGAATGGTTTCTATAAGATATGTAAAGGTCGTA
ACATTTGCGGTGTTGACAGTTTGGTCTCCACCGTCGCTGCTACCACTTCTTGAG
CTTTTCATTATTCTGCTGCAATGCTATTTAAAGTCTGTAAATAAATGATTAACAA
AAACATATTTCTTGGTGAATTCG

AT3G16210 cDNA full length sequence

Exon: 1 – Transcripts length: 1080 bp

ATGTCGAAATTTCTCCCGGAAGAATTAGCCATTGAGATTCTAGTGAGGCTATCG
ATGAAGGATCTTGCTCGATTCAGATGCGTCTGCAAGACATGGAGAGATCTTAT
CAACGACCCTGGATTCACAGAAACGTACCGAGACATGTCTCCGGCAAAGTTCG
TGTCGTTTTACGACAAGAACTTTTACATGCTTGACGTTGAAGGCAAGCATCCGG
TTATTACAAACAAGCTCGATTTCCCCTTGGATCAGTCAATGATCGATGAATCTA
CATGTGTGCTTCATTGTGATGGGACTTTATGTGTTACCTTAAAGAATCACACTTT
AATGGTTTGGAATCCGTTTTCGAAGCAATTCAAGATCGTACCAAATCCCGGTAT
ATACCAAGATTCAAACATTCTGGGTTTCGGTTATGACCCGGTTCACGACGATTA
CAAAGTCGTAACGTTTATCGACCGGCTTGATGTTTCCACGGCGCATGTTTTCGA
ATTTAGAACCGGTTCATGGGGAGAGAGCCTACGTATTTCTTATCCTGATTGGCA
TTACAGAGACAGGAGAGGTACATTTTTGGACCAATACCTCTATTGGATTGCGTA
CCGGTCTAGTGCTGACCGGTTTATCCTATGCTTTAATCTCTCAACACATGAATA
CCGAAAACTCCCTCTCCCGGTTTACAACCAAGGCGTGACGTCTTCATGGTTAG
GCGTTACAAGTCAAAAACTCTGTATTACAGAGTATGAAATGTGCAAGAAGGAG
ATTCGGATTTCGGTTATGGAGAAAACCGGATCATGGAGCAAGATTATAAGCCT
TTCAATGTCGAGTTTCATCTCTGTACAAGATCGTATCTACGATTATCAAGTCGA
ATTCGTGTCGTTTACTAGAAAGAACGATCTCGTCGTAACATTCACTGGATACAA
CGATCATTTCGAGATGGAACCTGAAGAGCGTACGAAGAAGAAGATGTTCTTGT
ACAAAACCGGAAACGAGAGATCTGAAGAAGTCCGGTTCTGTAACCCTCTAGCC
GGGTTACGGTTTCTCTGTGAATGTGTGGAGACTCTTAAGATTGTTAATCGCATA
TTTATTTGA

AT2G23945 short promoter, 437bp length sequence

Atgtgttggttttggtctgttatccaaatttgtaaaatttgaggtcctgtgaaagcaatttttggtc
aaactctgagattcatgtcttattatcaaaaaatttaccgaccaacaacttaatgatatgttgac
ctgtacttctaattgtttttgttcaatttatttttctaaaaaacataacacaaaacaacaaaagat
gtaaagtacatgatagcttttatatacataaccccaacaaacttttgctatttatagtctcatgat
ctttggttgtttttggatattatttgttgtcatcaattcgttgtgttttgatcttttatattcttcttcat
atcacagtcttgatgctattaccctaacttaatctttatttgaggttacaaaatggtctctgtaatc
ccccaagtccatgtttctgtaccgtttcaacaaca

AT2G23945 medium promoter, 980bp length sequence

caattgagttttgcagGTGAGGAAGCTACACAAGGAAATGAAAGTAGAGGAG
CTAGTAGACCGAGAACTGGGGACAACCTACGATAGAATAGAAGTTGGA
GAGATGCTACAAGTGGCACTGCTCTGCACTCAGTTTCTTCCAGCTCACA
GACCCAAAATGTCTGAAGTAGTTCAGATGCTTGAAGGAGATGGATTAGC
TGAGAGATGGGCTGCTTCACATGACCATTCACATTTCTACCATGCCAAC
ATGTCTTACAGGACTATTACCTCTACTGATGGCAACAACCAAACCAAAC
ATCTGTTTGGCTCCTCAGGATTTGAAGATGAAGATGATAATCAAGCGTT
AGATTCATTCGCCATGGAACTATCTGGTCCAAGGTAGtaaatcttggacaca
gaaagaaacagatataatatccccatgacttcaatttttgtttttgagatgatatagacatgtaa
atggttatcaaacctttgaaaaaattgagattctgggtctaatggttaatattttggatgtatat
attaaggttttttttaaatgtgttggttttggtctgttatccaaatttgtaaaatttgaggtcctgtg
aaagcaatttttggtcaaactctgagattcatgtcttattatcaaaaaatttaccgaccaacaac
ttaatgatatgttgacctgtacttctaattgtttttgttcaatttatttttctaaaaaacataacaca
aaacaacaaaagatgtaaagtacatgatagcttttatatacataaccccaacaaacttttgcta
tttatagtctcatgatctttggttgtttttggatattatttgttgtcatcaattcgttgtgttttgatctt
ttatattcttcttcatatcacagtcttgatgctattaccctaacttaatctttatttgaggttacaaa
atggtctctgtaatcccccaagtccatgtttctgtaccgtttcaacaaca

AT2G23945 long promoter, 1572bp length sequence

TGCGAGTTCTAGCGAAAGACTTCTTGTTTACCCTTACATGTCCAATGGCAGCGTCGCCT
CTAGGCTCAAAGgtagtcaaaaaaatggtttctgataactgcttttagtaaaaaaaaaaatgaaatcacttg
aactgatttttgttactgttacagCTAAGCCAGCGTTGGACTGGAACACAAGGAAGAAGATAGC
GATTGGAGCTGCAAGAGGGTTGTTTTATCTACACGAGCAATGCGATCCCAAGATTATTC
ACCGAGATGTCAAGGCAGCAAACATTCTCCTAGATGAGTATTTTGAAGCAGTTGTTGGG
GATTTTGGACTAGCAAAGCTACTCAACCACGAGGATTCACATGTCACAACCGCGGTTAG
AGGAACTGTTGGTCACATTGCACCTGAGTATCTCTCCACCGGTCAGTCATCTGAGAAAA
CCGATGTCTTTGGGTTCGGTATACTTTTGCTAGAGCTCATCACAGGAATGAGAGCTCTC
GAGTTTGGCAAGTCTGTTAGCCAGAAAGGAGCTATGCTAGAATGGgtaagcaatcactataa
gtttagatcaaacgtataatcaaaagatagtcttaaaactcaattgagttttgcagGTGAGGAAGCTACACA
AGGAAATGAAAGTAGAGGAGCTAGTAGACCGAGAACTGGGGACAACCTACGATAGAAT
AGAAGTTGGAGAGATGCTACAAGTGGCACTGCTCTGCACTCAGTTTCTTCCAGCTCACA
GACCCAAAATGTCTGAAGTAGTTCAGATGCTTGAAGGAGATGGATTAGCTGAGAGATG
GGCTGCTTCACATGACCATTCACATTTCTACCATGCCAACATGTCTTACAGGACTATTAC
CTCTACTGATGGCAACAACCAAACCAAACATCTGTTTGGCTCCTCAGGATTTGAAGATG
AAGATGATAATCAAGCGTTAGATTCATTCGCCATGGAACTATCTGGTCCAAGGTAGtaaa
tcttggacacagaaagaaacagatataatatccccatgacttcaatttttgtttttgagatgatatagacatgtaaat
ggttatcaaacctttgaaaaaattgagattctgggtctaatggttaatattttggatgtatatattaaggttttttttaaa
tgtgttggttttggtctgttatccaaatttgtaaaatttgaggtcctgtgaaagcaatttttggtcaaactctgagattc
atgtcttattatcaaaaaatttaccgaccaacaacttaatgatatgttgacctgtacttctaattgtttttgttcaattta
ttttttctaaaaaacataacacaaaacaacaaaagatgtaaagtacatgatagcttttatatacataaccccaacaa
acttttgctatttatagtctcatgatctttggttgtttttggatattatttgttgtcatcaattcgttgtgttttgatcttttat
attcttcttcatatcacagtcttgatgctattaccctaacttaatctttatttgaggttacaaaatggtctctgtaatccc
ccaagtccatgtttctgtaccgtttcaacaaca

AT2G23945 short translational fusion, 1811bp length sequence

atgtgttggttttggtctgttatccaaatttgtaaaatttgaggtcctgtgaaagcaatttttggtcaaactctgagatt
catgtcttattatcaaaaaatttaccgaccaacaacttaatgatatgttgacctgtacttctaattgtttttgttcaattt
atttttctaaaaaacataacacaaaacaacaaaagatgtaaagtacatgatagcttttatatacataaccccaaca
aacttttgctatttatagtctcatgatctttggttgtttttggatattatttgttgtcatcaattcgttgtgttttgatctttta
tattcttcttcatatcacagtcttgatgctattaccctaacttaatctttatttgaggttacaaaatggtctctgtaatcc
cccaagtccatgtttctgtaccgtttcaacaacaATGGAATGTTCTTTCCAAACTTCACTCTTACTCTT
CATCACTGTATCATACTTTGTAGTAACAGAATCCATTAAACCTAACCGAATGGCGATGAA
GCTGATACACCGTGAATCAGTGGCCCGTCTCAACCCAAATGCTCGAGTTCCGATCACTC
CAGAAGACCATATCAAACACCTGACAGATATTTCATCAGCCCGGTTTAAGTATCTACAG
AATTCGATTGACAAAGAGCTAGGCAGCAGCAATTTCCAAGTTGATGTGGAACAAGCAAT
CAAAACATCATTGTTTCTTGTGAATTTCTCCGTAGGACAGCCACCGGTGCCGCAGTTAA
CTATCATGGATACTGGAAGCTCACTTTTATGGATCCAATGTCAACCCTGCAAACACTGC
TCATCAGATCATATGATCCATCCTGTCTTCAATCCAGCATTGTCTTCAACATTTGTAGAG
TGTTCCTGCGATGATCGTTTTTGCAGGTATGCACCAAACGGACATTGCGGCTCATCAAA
CAAGTGTGTTTACGAACAAGTGTATATAAGCGGCACTGGGTCAAAAGGCGTTCTAGCTA
AAGAACGGTTAACCTTCACAACACCTAACGGTAACACTGTTGTGACCCAACCTATTGCT
TTTGGGTGTGGCTACGAGAATGGAGAGCAGTTGGAAAGTCATTTCACTGGTATTTTGG
GGTTAGGAGCCAAACCCACATCACTTGCCGTACAGCTTGGGTCTAAATTCTCTTACTGC
ATTGGTGATTTAGCTAACAAAAACTACGGTTACAATCAGCTGGTTCTGGGAGAGGACGC
GGATATTCTCGGCGATCCAACACCCATTGAATTTGAGACAGAGAATAGTATCTACTATA
TGAATCTAGAAGGGATTAGTGTTGGGGATACACAACTCAATATCGAACCAGTGGTATTC
AAAAGAAGAGGACCAAGAACAGGGGTCATACTAGACTCGGGCACACTCTACACTTGGC
TAGCAGATATAGCCTACAGAGAACTCTACAATGAAATAAAATCCATTCTGGATCCAAAAT
TAGAGAGATTCTGGTTTAGAGACTTCTTGTGCTACCATGGGAGAGTGAGCGAAGAGCT
CATAGGCTTCCCTGTGGTGACATTTCATTTTGCAGGCGGAGCTGAGCTGGCTATGGAG
GCCACAAGCATGTTTTATCCATTATCAGAACCGAATACGTTTAACGTCTTTTGTATGTCA
GTGAAGCCAACCAAAGAACATGGAGGGGAATATAAGGAATTCACAGCTATAGGGCTAA
TGGCTCAACAATACTATAATATTGGTTACGACCTCAAAGAAAAAAACATCTATCTCCAGA
GAATCGACTGTGTTCAACTCGATGACTATTCCCCATCAAAG

AT2G23945 medium translational fusion, 2354bp length sequence

caattgagttttgcagGTGAGGAAGCTACACAAGGAAATGAAAGTAGAGGAGCTAGTAGACCGAGAAC
TGGGGACAACCTACGATAGAATAGAAGTTGGAGAGATGCTACAAGTGGCACTGCTCTGCACTCAG
TTTCTTCCAGCTCACAGACCCAAAATGTCTGAAGTAGTTCAGATGCTTGAAGGAGATGGATTAGCT
GAGAGATGGGCTGCTTCACATGACCATTCACATTTCTACCATGCCAACATGTCTTACAGGACTATT
ACCTCTACTGATGGCAACAACCAAACCAAACATCTGTTTGGCTCCTCAGGATTTGAAGATGAAGAT
GATAATCAAGCGTTAGATTCATTCGCCATGGAACTATCTGGTCCAAGGTAGtaaatcttggacacagaa
agaaacagatataatatccccatgacttcaatttttgtttttgagatgatatagacatgtaaatggttatcaaacctttgaaaaaat
tgagattctgggtctaatggttaatattttggatgtatatattaaggtttttttaaatgtgttggttttggtctgttatccaaatttgta
aaatttgaggtcctgtgaaagcaatttttggtcaaactctgagattcatgtcttattatcaaaaaatttaccgaccaacaacttaat
gatatgttgacctgtacttctaattgtttttgttcaatttatttttctaaaaaacataacacaaaacaacaaaagatgtaaagtacat
gatagcttttatatacataaccccaacaaactttgctatttatagtctcatgatctttggttgtttttggatattatttgttgtcatcaa
ttcgttgtgttttgatcttttatattcttcttcatatcacagtcttgatgctattaccctaacttaatctttatttgaggttacaaaatggt
ctctgtaatcccccaagtccatgtttctgtaccgtttcaacaacaATGGAATGTTCTTTCCAAACTTCACTCTTACTC
TTCATCACTGTATCATACTTTGTAGTAACAGAATCCATTAAACCTAACCGAATGGCGATGAAGCTG
ATACACCGTGAATCAGTGGCCCGTCTCAACCCAAATGCTCGAGTTCCGATCACTCCAGAAGACCA
TATCAAACACCTGACAGATATTTCATCAGCCCGGTTTAAGTATCTACAGAATTCGATTGACAAAGA
GCTAGGCAGCAGCAATTTCCAAGTTGATGTGGAACAAGCAATCAAAACATCATTGTTTCTTGTGAA
TTTCTCCGTAGGACAGCCACCGGTGCCGCAGTTAACTATCATGGATACTGGAAGCTCACTTTTATG
GATCCAATGTCAACCCTGCAAACACTGCTCATCAGATCATATGATCCATCCTGTCTTCAATCCAGC
ATTGTCTTCAACATTTGTAGAGTGTTCCTGCGATGATCGTTTTTGCAGGTATGCACCAAACGGACA
TTGCGGCTCATCAAACAAGTGTGTTTACGAACAAGTGTATATAAGCGGCACTGGGTCAAAAGGCG
TTCTAGCTAAAGAACGGTTAACCTTCACAACACCTAACGGTAACACTGTTGTGACCCAACCTATTG
CTTTTGGGTGTGGCTACGAGAATGGAGAGCAGTTGGAAAGTCATTTCACTGGTATTTTGGGGTTA
GGAGCCAAACCCACATCACTTGCCGTACAGCTTGGGTCTAAATTCTCTTACTGCATTGGTGATTTA
GCTAACAAAAACTACGGTTACAATCAGCTGGTTCTGGGAGAGGACGCGGATATTCTCGGCGATCC
AACACCCATTGAATTTGAGACAGAGAATAGTATCTACTATATGAATCTAGAAGGGATTAGTGTTGG
GGATACACAACTCAATATCGAACCAGTGGTATTCAAAAGAAGAGGACCAAGAACAGGGGTCATAC
TAGACTCGGGCACACTCTACACTTGGCTAGCAGATATAGCCTACAGAGAACTCTACAATGAAATAA
AATCCATTCTGGATCCAAAATTAGAGAGATTCTGGTTTAGAGACTTCTTGTGCTACCATGGGAGAG
TGAGCGAAGAGCTCATAGGCTTCCCTGTGGTGACATTTCATTTTGCAGGCGGAGCTGAGCTGGCT
ATGGAGGCCACAAGCATGTTTTATCCATTATCAGAACCGAATACGTTTAACGTCTTTTGTATGTCA
GTGAAGCCAACCAAAGAACATGGAGGGGAATATAAGGAATTCACAGCTATAGGGCTAATGGCTCA
ACAATACTATAATATTGGTTACGACCTCAAAGAAAAAAACATCTATCTCCAGAGAATCGACTGTGT
TCAACTCGATGACTATTCCCCATCAAAG

AT2G23945 long translational fusion, 2946bp length sequence

TGCGAGTTCTAGCGAAAGACTTCTTGTTTACCCTTACATGTCCAATGGCAGCGTCGCCT
CTAGGCTCAAAGgtagtcaaaaaaatggtttctgataactgcttttagtaaaaaaaaaaatgaaatcacttg
aactgatttttgttactgttacagCTAAGCCAGCGTTGGACTGGAACACAAGGAAGAAGATAGC
GATTGGAGCTGCAAGAGGGTTGTTTTATCTACACGAGCAATGCGATCCCAAGATTATTC
ACCGAGATGTCAAGGCAGCAAACATTCTCCTAGATGAGTATTTTGAAGCAGTTGTTGGG
GATTTTGGACTAGCAAAGCTACTCAACCACGAGGATTCACATGTCACAACCGCGGTTAG
AGGAACTGTTGGTCACATTGCACCTGAGTATCTCTCCACCGGTCAGTCATCTGAGAAAA
CCGATGTCTTTGGGTTCGGTATACTTTTGCTAGAGCTCATCACAGGAATGAGAGCTCTC
GAGTTTGGCAAGTCTGTTAGCCAGAAAGGAGCTATGCTAGAATGGgtaagcaatcactataa
gtttagatcaaacgtataatcaaaagatagtcttaaaactcaattgagttttgcagGTGAGGAAGCTACACA
AGGAAATGAAAGTAGAGGAGCTAGTAGACCGAGAACTGGGGACAACCTACGATAGAAT
AGAAGTTGGAGAGATGCTACAAGTGGCACTGCTCTGCACTCAGTTTCTTCCAGCTCACA
GACCCAAAATGTCTGAAGTAGTTCAGATGCTTGAAGGAGATGGATTAGCTGAGAGATG
GGCTGCTTCACATGACCATTCACATTTCTACCATGCCAACATGTCTTACAGGACTATTAC
CTCTACTGATGGCAACAACCAAACCAAACATCTGTTTGGCTCCTCAGGATTTGAAGATG
AAGATGATAATCAAGCGTTAGATTCATTCGCCATGGAACTATCTGGTCCAAGGTAGtaaa
tcttggacacagaaagaaacagatataatatccccatgacttcaatttttgtttttgagatgatatagacatgtaaat
ggttatcaaacctttgaaaaaattgagattctgggtctaatggttaatattttggatgtatatattaaggttttttttaaa
tgtgttggttttggtctgttatccaaatttgtaaaatttgaggtcctgtgaaagcaatttttggtcaaactctgagattc
atgtcttattatcaaaaaatttaccgaccaacaacttaatgatatgttgacctgtacttctaattgtttttgttcaattta
tttttctaaaaaacataacacaaaacaacaaaagatgtaaagtacatgatagcttttatatacataaccccaacaa
actttgctatttatagtctcatgatctttggttgtttttggatattatttgttgtcatcaattcgttgtgttttgatcttttat
attcttcttcatatcacagtcttgatgctattaccctaacttaatctttatttgaggttacaaaatggtctctgtaatccc
ccaagtccatgtttctgtaccgtttcaacaacaATGGAATGTTCTTTCCAAACTTCACTCTTACTCTT
CATCACTGTATCATACTTTGTAGTAACAGAATCCATTAAACCTAACCGAATGGCGATGAA
GCTGATACACCGTGAATCAGTGGCCCGTCTCAACCCAAATGCTCGAGTTCCGATCACTC
CAGAAGACCATATCAAACACCTGACAGATATTTCATCAGCCCGGTTTAAGTATCTACAG
AATTCGATTGACAAAGAGCTAGGCAGCAGCAATTTCCAAGTTGATGTGGAACAAGCAAT
CAAAACATCATTGTTTCTTGTGAATTTCTCCGTAGGACAGCCACCGGTGCCGCAGTTAA
CTATCATGGATACTGGAAGCTCACTTTTATGGATCCAATGTCAACCCTGCAAACACTGC
TCATCAGATCATATGATCCATCCTGTCTTCAATCCAGCATTGTCTTCAACATTTGTAGAG
TGTTCCTGCGATGATCGTTTTTGCAGGTATGCACCAAACGGACATTGCGGCTCATCAAA
CAAGTGTGTTTACGAACAAGTGTATATAAGCGGCACTGGGTCAAAAGGCGTTCTAGCTA

AAGAACGGTTAACCTTCACAACACCTAACGGTAACACTGTTGTGACCCAACCTATTGCT
TTTGGGTGTGGCTACGAGAATGGAGAGCAGTTGGAAAGTCATTTCACTGGTATTTTGG
GGTTAGGAGCCAAACCCACATCACTTGCCGTACAGCTTGGGTCTAAATTCTCTTACTGC
ATTGGTGATTTAGCTAACAAAAACTACGGTTACAATCAGCTGGTTCTGGGAGAGGACGC
GGATATTCTCGGCGATCCAACACCCATTGAATTTGAGACAGAGAATAGTATCTACTATA
TGAATCTAGAAGGGATTAGTGTTGGGGATACACAACTCAATATCGAACCAGTGGTATTC
AAAAGAAGAGGACCAAGAACAGGGGTCATACTAGACTCGGGCACACTCTACACTTGGC
TAGCAGATATAGCCTACAGAGAACTCTACAATGAAATAAAATCCATTCTGGATCCAAAAT
TAGAGAGATTCTGGTTTAGAGACTTCTTGTGCTACCATGGGAGAGTGAGCGAAGAGCT
CATAGGCTTCCCTGTGGTGACATTTCATTTTGCAGGCGGAGCTGAGCTGGCTATGGAG
GCCACAAGCATGTTTTATCCATTATCAGAACCGAATACGTTTAACGTCTTTTGTATGTCA
GTGAAGCCAACCAAAGAACATGGAGGGGAATATAAGGAATTCACAGCTATAGGGCTAA
TGGCTCAACAATACTATAATATTGGTTACGACCTCAAAGAAAAAAACATCTATCTCCAGA
GAATCGACTGTGTTCAACTCGATGACTATTCCCCATCAAAG

Appendix 4

Preparation of Phosphate-buffered saline (PBS)

Component	**Concentration**
NaCl	137 mM
KCl	2.7 mM
Na_2HPO_4	10 mM
KH_2PO_4	2 mM

For 1 litre of 1x phosphate-buffered saline dissolve the following in 800 ml of distilled water and adjust the pH to 7.4 with HCl. Add distilled water to 1 litre.

- 8 g of NaCl
- 0.2 g of KCl
- 1.44 g of Na2HPO4
- 0.24 g of KH2PO4

PBS can be sterilised by autoclaving and stored at room temperature.

Appendix 5

Preparation of GUS substrate and buffers

0.1 M Phosphate Buffer
If desired, 1 litre of Phosphate Buffer can be prepared, autoclaved and used when required

Prepare solutions A and B:

Solution A: 0.2M $Na_2HPO_4.2H_2O$ (weigh 3.56g; dissolve in 100 ml of water)
Solution B: 0.2M $NaH_2PO_4.H_2O$ (weigh 2.76g; dissolve in 100 ml of water)

Mix the following:

Solution A:30.5 ml
Solution B:19.5 ml
ddH2O:50.0 ml
TOTAL:100.0 ml

Adjust the pH to 7.0

GUS substrate solution (modified from Lindsey and Wei 2000)

Stock solutions

10 mg/ml (20 mM) X-Gluc (5-bromo-4-chloro-3-indolyl-ß-D-glucuronide) in N,N-Dimethyl-formamide (DMF) (handle with care; can be stored in 100 µl aliquots at -20°C)
33 mg/ml (1 mM) $K_3F_6(CN)_6$ in 0.05 M phosphate buffer pH 7.2.
43 mg/ml (1 mM) $K_4Fe(CN)_6.3H_2O$ in 0.05 M phosphate buffer pH 7.2.
10 mg/ml X-Gluc staining solution (MW 521.8)

Mix in order

4.245 ml 0.05 M PO4 buffer pH 7.2
250µl 33 mg/ml $K_3F_6(CN)6$
250 µl 43mg/ml $K_4Fe(CN)6.3H2O$
250 µl 10 mg/ml X-Gluc
5 µl Triton X-100