MOLECULAR MECHANISMS UNDERPINNING LATERAL ROOT DEVELOPMENT IN *ARABIDOPSIS* AND CROPS

by
CLARE MARGARET CLAYTON

A thesis submitted to
The University of Birmingham
For the degree of
DOCTOR OF PHILOSOPHY

School of Biosciences
The University of Birmingham
September 2019

Abstract

Lateral roots (LR) branch out from the primary root in order to provide essential stability and allow water and nutrient uptake. The threat of climate change and an increasing global population represent crucial challenges to plant biodiversity and food security, and the manipulation of plant root system architecture could offer the potential to produce new and more resilient crop varieties.

The transcription factor *AtMYB93* is known to inhibit the initiation of new LR primordia in the model organism *Arabidopsis*, preventing resource expenditure on unnecessary root growth. It is hypothesised that the transcription factor *AtMYB93* manipulates the suberin barrier layer in the endodermis, however the molecular mechanism through which *AtMYB93* acts is poorly characterised.

Here we begin to answer questions about AtMYB93, its regulatory network, and the phenotypic responses of *AtMYB93* and its S24 clade members - *AtMYB92* and *AtMYB53* - under stress, and have attempted to apply these findings on MYB93 homologues to genetic manipulation of crop root systems.

Acknowledgements

Thank you to my amazing supervisor Juliet Coates, who has been an excellent academic and emotional support, and has provided me with boundless inspiration throughout this project. Thank you to my co-supervisor Daniel Gibbs for his words of wisdom and great ideas, and to my collaborators who have provided me with seeds and have analysed my tissue samples. Thank you to my superb lab and office family who are too numerous to list in their entirety, and for many their time at Birmingham has been far too brief! I have enjoyed working with you immensely and look forward to catching up soon. In particular, I would like to thank my fellow PhD lab colleagues and friends Xulyu, Alex and Fatemah, and the post docs next door Anne-Marie and Mark. Thank you also to Helen for joining me on this project to complete her PhD training mini-project. Finally, thank you to my family and friends who have put up with my unusual working hours, last-minute cancellations and have provided much tea, wine and hugs when needed.

Contents

Chapter 1: Introduction **1**

1.1: Context and significance of the research **2**
1.1.1 Global food security 2
1.1.2 Major problems with plant productivity due to soil quality 4
1.1.3 Sulfur deficiency and its effects on plants 6
1.1.4 Soil salinisation and its effects on plants 8
1.2: The plant root system **11**
1.2.1 Root structure and function 11
1.2.2 Lateral root development 14
1.2.3 Lateral root initiation and patterning 15
1.2.4 Lateral root emergence 16
1.3: Hormonal regulation of root development **19**
1.3.1 The role of hormones in plant growth and development 19
1.3.2 Hormonal regulation of root system architecture 20
1.4: Transcription factors regulating lateral root development **22**
1.4.1 The MYB family transcription factors 22
1.4.2 MYB transcription factors important for regulating root development 23
1.4.3 *At*MYB93 interacts with *At*ARABIDILLO proteins and is restricted to the endodermis 24
1.4.4 Close relatives of *AtMYB93* 28
1.5: Complex biopolymers in plants **30**
1.5.1 MYBs and complex polymers 30
1.5.2 The role of suberin and its potential regulation by *AtMYB93* 31
1.5.3 Suberin and the seed coat 35
1.6: AtMYB93 and environmental triggers **36**
1.6.1 Abiotic factors regulating the *AtMYB93* response 36
1.6.2 Sulfur stress and the *AtMYB93* response 36
1.6.3 Drought and salt stress and the *AtMYB93* response 38
1.7: Real world applications of modulating *AtMYB93* function **39**
1.7.1 Exploiting lateral root structure as a way to improve crops 39
1.7.2 Translating knowledge from *Arabidopsis* to crops 41
1.8: Aims and objectives of the project **42**

Chapter II: Materials and methods **43**

2.1: Species and strains used **44**
2.1.1 Plant species 44
2.1.2 Bacterial strains 45
2.1.3 Yeast strains 45
2.2: Vectors **45**

2.3: Growth conditions **46**
2.3.1 Plant growth conditions 46
2.3.3 Seed sterilisation 47
2.3.2 Bacterial growth conditions 48
2.3.3 Yeast growth conditions 48
2.4: Bioinformatic analysis **48**
2.4.1 Identification of putative *At*MYB93 downstream target genes 48
2.4.2 Promoter analysis of putative *At*MYB93 target genes 50
2.4.3 Creation of phylogenetic trees of MYB93 homologues 51
2.4.4 Gene structure analysis 52
2.5: Nucleic acid isolation and handling **52**
2.5.1 Quick extraction of genomic DNA for genotyping Arabidopsis 52
2.5.2 Extraction of high-quality genomic DNA from *Arabidopsis* 53
2.5.3 Extraction of total RNA from plants 53
2.5.4 Estimation of nucleic acid concentration 54
2.5.5 Primer design 54
2.6: Manipulation of nucleic acids **56**
2.6.1 Polymerase chain reaction (PCR) 56
2.6.2 Reverse transcription PCR (RT-PCR) 56
2.6.3 Agarose gel electrophoresis 57
2.6.4 Semi-quantification of RT-PCR band intensity 58
2.6.5 cDNA synthesis 58
2.6.6 Quantitative RT-PCR (qRT-PCR) 59
2.6.7 qRT-PCR data analysis 59
2.7: Gene cloning **60**
2.7.1 Purification of PCR reactions 60
2.7.2 Extraction of plasmid DNA from bacteria 60
2.7.3 Restriction enzyme digest 60
2.7.4 Single colony PCR 61
2.7.5 Plasmid DNA de-phosphorylation and precipitation 61
2.7.6 Ligation of DNA into plasmids 62
2.7.7 Transformation of plasmid DNA into *E. coli* 62
2.7.8 Transformation of plasmid DNA into *S. cerevisiae* 63
2.8: Yeast one-hybrid assays **63**
2.9: Floral dip and transgenic plant selection **64**
2.9.1 Transformation of plasmid DNA into *A. tumefaciens* 64
2.9.2 *A. tumefaciens*-mediated transformation into *Arabidopsis* using floral dip 64
2.9.3 Selection of transgenic plant lines 65
2.9.4 Selection of mutant plant lines 65
2.10: Phenotyping assays **66**
2.10.1 Root bioassays 66

2.10.2	Germination assays	66
2.10.3	Inductively Coupled Plasma Mass Spectrometry (ICP-MS)	67
2.10.3.1	Preparation and growth of *Arabidopsis* tissue	67
2.10.3.2	Harvest and experimental analysis of tissue samples	67
2.10.3.3	Analysis of elemental data	68
2.11: Proteins		**68**
2.11.1	MG-132 treatment	68
2.11.2	Protein extraction from plants	68
2.11.3	Estimation of protein concentration	69
2.11.4	SDS-PAGE	69
2.12: Western blotting		**69**
2.12.1	Protein transfer	69
2.12.2	Blocking and antibody probing	70
2.12.3	Protein detection and analysis	70
2.12.4	Coomassie staining	70
2.13: Statistical analysis		**71**
2.14: Graphical software used for figure preparation		**71**
Chapter III: Identifying upstream regulators and downstream targets of *At*MYB93		73
3.1: Introduction		**74**
3.2: Bioinformatical analysis to select putative downstream targets of *At*MYB93		**77**
3.2.1	Analysis of microarray expression data after lateral root induction	77
3.2.2	Genes matching the expression profile of *AtMYB93* in response to hormones and stresses	80
3.2.3	Potential *At*MYB93 target genes identified	82
3.2.4	Candidate gene promoter analysis	87
3.3: Strategies for confirming putative *At*MYB93 targets		**92**
3.3.1	Quantitative reverse transcription PCR was used to analyse relative expression patterns	92
3.3.2	Yeast one-hybrid to detect *At*MYB93 protein-candidate gene interactions	97
3.3.3	Targeted chromatin immunoprecipitation techniques	103
3.4: Analysis of RNA-seq data		**108**
3.5: A potential upstream repressor of *AtMYB93*		**111**
3.6: Discussion		**113**
3.6.1	Candidate target genes of *At*MYB93 can be categorised by function	114
3.6.2	Comparing the candidate gene mutants with RNA-seq data from the *Atmyb93* mutant	114
3.6.3	Identities of candidate genes found to match the expression profile of *AtMYB93*	115

3.6.3.1	*AtPME10*: a putative cell wall remodelling enzyme	115
3.6.3.2	*AtCYP708A1*: a gene putatively involved in fatty acid or suberin metabolism	117
3.6.3.3	*AtMYB54*: a transcription factor involved in cell wall biosynthesis	118
3.6.3.4	*AtPDLP4*: an endodermal membrane protein	119
3.6.3.5	*AtAHA4*: a plasma membrane ATPase	120
3.6.3.6	*At4G17480*: a possible fatty acid or suberin metabolism gene	120
3.6.4	Identities of other putative target genes	121
3.6.4.1	*AtCASPL1B1*: a Casparian strip membrane protein	121
3.6.4.2	*AtLTP8*: a gene putatively involved in fatty acid or suberin metabolism	121
3.6.4.3	*AtGPAT6*: a cutin biosynthesis gene	122
3.6.4.4	*AtQRT1*: a putative cell wall remodelling enzyme	122
3.6.4.5	*AtCYS5*: a cystatin proteinase inhibitor	123
3.6.4.6	*AtABCG2*: a suberin biosynthesis gene	123
3.6.4.7	Uncharacterised genes	125
3.6.5	*At*SCARECROW is an upstream repressor of *AtMYB93*	125
3.6.6	Challenges and limitations	126
3.6.6.1	Limitations of assessing expression patterns	126
3.6.6.2	Limitations of promoter analysis	127
3.6.6.3	Experimental limitations of using whole seedling tissue	128
3.6.6.4	Root tissue mass yield optimisation	129
3.6.6.5	Technical issues with qRT-PCR	131
3.6.6.6	Yeast-one-hybrid considerations	131
3.6.6.7	Glasshouse issues	132
3.6.7	Conclusion	133

Chapter IV: Characterisation of mutants related to AtMYB93 signalling 134

4.1: Introduction **135**

4.2: Genotyping of AtMYB93 putative target gene mutants **137**

4.3: Phenotypic characterisation of *At*MYB93 putative target gene mutants **140**

4.3.1	Lateral root density phenotypes of *At*MYB93 putative target gene mutants	140
4.3.2	Elemental analysis of *At*MYB93 putative target gene mutants	145

4.4: Genotyping of single, double and triple MYB mutants **148**

4.5: Phenotypic characterisation of single, double and triple MYB mutants **150**

4.5.1	Germination phenotypes of MYB mutants	150

4.5.2	*AtMYB41* expression in MYB mutants	153
4.5.3	Lateral root density phenotypes of MYB mutants	154
4.5.4	Elemental analysis of MYB mutants	160
4.5.5	Elemental analysis of the *Atmyb93* mutant under sulfur (and salt) stress	162
4.6: Discussion		**164**
4.6.1	Lateral root phenotypes have been characterised in putative downstream targets of *At*MYB93 but did not reveal a phenotype in *Atmyb53* or MYB triple mutants	164
4.6.2	MYB triple mutants have a limited response to germination under salt stress but *AtMYB41* could be functionally redundant	166
4.6.3	Elemental analysis of the ion content of *Arabidopsis* shoot tissue	166
4.6.4	Challenges and limitations	167
4.6.4.1	Genotyping candidate gene mutant lines	167
4.6.4.2	Replicating the *AtMYB93* phenotype	169
4.6.4.3	Failed growth of plants under salt stress for elemental analysis	170
4.7: Conclusion		**171**
Chapter V: MYB93 homologues: evolutionary conservation including in crops		**172**
5.1: Introduction		173
5.2: MYB93 homologues		174
5.2.1	Phylogenetic analysis of MYB93 crop homologues	174
5.2.2	Intron analysis of MYB93 crop homologues	183
5.2.3	Predicted expression of putative MYB93 homologues in cereals	185
5.2.4	RT-PCR expression of putative MYB93 homologues in cereals	186
5.3: Analysis of an AtARABIDILLO over-expressor in wheat		189
5.4: Discussion		192
5.4.1	Phylogenetic analysis identified putative MYB homologues in monocots	192
5.4.2	Transgenic wheat lines are not expressing YFP	194
5.4: Conclusion		196
Chapter VI: General discussion		**197**
6.1: Introduction		198
6.2: Downstream targets of *At*MYB93		198
6.3: Upstream repression of *AtMYB93* by *AtSCARECROW*		200
6.4: Close relatives of *AtMYB93*		201

6.5: MYB93 crop homologues 202
6.6: Final remarks 204

Chapter VII: References **205**

List of tables and figures

List of figures		
1.1	Diagram of mature Arabidopsis and a magnified cross-sectional image of the cellular structure of the root	12
1.2	Auxin distribution in the developing lateral root	17
1.3	Hormonal regulation of lateral root initiation	21
1.4	Phenotypes and possible model of interaction between *At*MYB93 and downstream promoters	26
1.5	*AtMYB93* expression is restricted to a limited number of endodermal cells	27
1.6	Consensus and divergence between published MYB phylogenies	29
1.7	Diagram of the effects of hormones and stresses on suberin deposition and *AtMYB93* activity	33
2.1	Primer design to avoid amplification of gDNA contamination when looking at gene expression levels	55
3.1	Flowchart depicting selection of candidate gene targets of *At*MYB93 through bioinformatic analysis	78
3.2	Heatmap of fold-change in expression of well-known suberin lateral root biosynthesis genes not thought to be *At*MYB93 targets.	79
3.3	Heatmap of fold-change in expression of the 16 genes selected as possible transcriptional targets of AtMYB93 (also shown)	83
3.4	Conserved motifs in the putative *At*MYB93 DNA binding domains of the 16 candidate target genes	91
3.5	qRT-PCR analysis of relative expression of AtMYB93 putative targets in 7-day-old seedlings	93-96
3.6	qRT-PCR analysis of relative *AtMYB93* expression between Col-0, *Atmyb93* mutant and *35S:AtMYB93* over-expressor mutants	98
3.7	Diagram of the yeast one-hybrid interaction procedure	100
3.8	Schematic diagram of restriction enzyme cloning process used to create yeast one-hybrid prey and bait plasmids	101
3.9	Graphical Illustration of the constructs that have been created to generate transgenic Arabidopsis for ChIP-qPCR study	104
3.10	Schematic diagram of restriction enzyme cloning process used to create ChIP-qPCR constructs	106
3.11	Venn diagram comparing genes obtained through bioinformatic analysis vs RNA-seq during lateral root initiation	110
3.12	qRT-PCR analysis of *At*SCR as a possible transcriptional target of *AtMYB93*	112
4.1	Genotyping PCR results of candidate gene mutants	141
4.2	Primary and lateral root phenotypes of putative *At*MYB93 targets of 7-day-old seedlings	142

4.3	Primary and lateral root phenotypes of putative *At*MYB93 targets of 8-day-old seedlings	143
4.4	Primary and lateral root phenotypes of putative *At*MYB93 targets of 9-day-old seedlings	144
4.5	Comparison of the sulfur, sodium and magnesium ion content of *Atmyb93* and mutant candidate gene targets of *At*MYB93 relative to WT in 25-day-old *Arabidopsis* shoot tissue	146
4.6	RT-PCR results of MYB mutants	149
4.7	Germination assays of MYB mutants under (A) salt stress or (B) normal conditions	152
4.8	Semi-quantitative RT-PCR analysis of *AtMYB41* expression in MYB mutants	155
4.9	Primary and lateral root phenotypes of MYB mutants of 7-day-old seedlings	156
4.10	Primary and lateral root phenotypes of MYB mutants of 8-day-old seedlings	157
4.11	Primary and lateral root phenotypes of MYB mutants of 9-day-old seedlings	158
4.12	Comparison of the sulfur, sodium and magnesium ion content of MYB mutants relative to WT in 25-day-old *Arabidopsis* shoot tissue	161
4.13	Comparison of the sulfur ion content in *Atmyb93* mutant tissue relative to WT in 25-day-old *Arabidopsis* shoot tissue grown on medium with or without sulfur	163
5.1	Phylogenetic trees identifying MYB homologues in other species	178-181
5.2	Comparison of MYB homologue gene structure	184
5.3	RT-PCR analysis of putative MYB93 homologues in cereals	187
5.4	Western blot and RT-PCR analysis of *35S::AtARABIDILLO*:YFP in wheat	191
5.5	Alignment of *AtARABIDILLO* RT-PCR primers to wheat sequence	195

List of tables

3.1	Hormone and stress expression profiles for the 16 genes selected as possible transcriptional targets of *At*MYB93 (also shown)	85
3.2	Expression of the 16 putative *At*MYB93 (also shown) target genes in different tissues	86
3.3	Functional categorisation of putative *At*MYB93 targets	88
3.4	Predicted *At*MYB93 DNA binding domain motifs identified from a promoter analysis of the 16 putative *At*MYB93 targets	90
3.5	Yeast one-hybrid progress	102

3.6	Candidate *At*MYB93 target gene expression from RNA-seq during lateral root initiation	110
4.1	Expected vs actual genotyping phenotypes of T-DNA mutants	139
5.1	Putative *AtMYB93* homologues identified in BLAST search	176

List of appendices

Appendix 2.1	Mutant plant lines	I
Appendix 2.2	Recipes	II
Appendix 2.3	Primers	VI
Appendix 3.1	Genes identified from bioinformatic analysis to select candidate downstream targets of *At*MYB93	X
Appendix 4.1	LBb1.3 and LB1 border primers used for SALK and SAIL lines respectively	XVI

List of abbreviations

3-AT – 3-amino-1,2,4-triazole

ABA – Abscisic acid

AD – Activation domain

ARF – Auxin response factor

ATP – Adenosine triphosphate

BSA – Bovine serum albumin

cDNA – Complementary DNA

ChIP – Chromatin immunoprecipitation

CIP – Calf intestinal phosphatase

Col-0 – Columbia

dH_2O – Deionised water

DMSO – Dimethyl sulfoxide

DNA – Deoxyribonucleic acid

EDTA – Ethylenediaminetetraacetic acid

eFP – Electronic Fluorescent Pictograph

ERF – Ethylene response factor

gDNA – Genomic DNA

GSDS – Gene Structure Display Server

ICP-MS – Inductively coupled plasma mass spectrometry

KO – Knock out

LR – Lateral root

M-MLV – Moloney murine leukemia virus

MCS – Multiple cloning site

MS – Murashige & Skoog

NIAB – National Institute of Agricultural Botany

PCR – Polymerase chain reaction

PEG – Polyethelyne glycol

PlantTFDB – Plant Transcription Factor Database

PP60 – Polyphenon-60

PPM – Plant Preservative Mixture

PVPP – Polyvinylpolypyrrolidone

qRT-PCR – Quantitative RT-PCR

RNA – Ribonucleic acid

RNAi – RNA interference

RT – Reverse transcriptase

RT-PCR – Reverse transcription-PCR

SAIL – Syngenta *Arabidopsis* Insertion Library

SCR - SCARECROW

SD – Standard deviation

SHR – SHORT-ROOT

TBS – Tris buffered saline

TBST – Tris buffered saline + Tween 20

T-DNA – Transfer DNA

TF – Transcription factor

UTR – Untranslated region

VLCFA – Very long chain fatty acids

WT – Wild-type

Y1H – Yeast one-hybrid

Y2H – Yeast two-hybrid

Chapter I

INTRODUCTION

1.1 Context and significance of the research

1.1.1 Global food security

Food security can be defined as being when people have consistent access to enough affordable and nutritious food (FAO, 1996). It is an important global issue with up to 800 million, or 11 percent of the population, currently facing undernourishment (FAO *et al.*, 2015). Criticism has been made regarding the way scientists frame the issue of food security when addressing their research questions, since emphasis is often placed on improving crop yield (Helliwell *et al.*, 2017). On average we are already producing enough food to sustain the current global population (FAO; 2017), however, in developing countries in particular, lack of access to an adequate supply of food is the principal issue. While clearly more effort is needed to address this challenge, it is important to continue to research ways to improve crop production in parallel. The global population is continuing to rise and is predicted to reach 9.8 billion by 2050 (UN, 2017). There are many challenges that could hinder our ability to achieve food security with recent estimates suggesting that in order to keep up with the rising population, agricultural output must increase by between 48 and 70 percent (FAO, 2009; 2017).

The Green Revolution is widely credited as being a significant contributing factor in increasing global crop yields and its impacts have been measured for several decades (Evenson & Gollin, 2003; Pingali, 2012). However, since the turn of the millennium, investment into agriculture has dropped significantly and the benefits of the Green Revolution have seemingly plateaued (Herdt, 2010). In fact, in recent decades even the implementation of positive measures such as reducing pollution has actually had a negative impact on nutrient levels in the soil, thus resulting in

lower yields (Lewandowska *et al.*, 2011). For example, advancements in fuel technology over the past several decades have resulted in the successful removal of nearly all sulfur from petrol and diesel (Stanislaus *et al.*, 2010). As a result of this, there is a need for more sulfur-containing fertiliser to be added to soils and some crops are more sensitive to soil sulfur levels than others. *Brassica napus* (oilseed rape) is one such crop that is highly dependent on sulfur and requires much more of the nutrient than many other crops (Berry *et al.*, 2014). Oil seed rape growers in China can expect yields per hectare to achieve just 50% of those seen in Europe. Furthermore, it is estimated that more than a third of soils in China are sulfur deficient, yet fertilisation with sulfur can increase yields by between 7-15% (The Sulphur Institute (2009).

One impending threat to food security is water scarcity and although climate change appears to be exacerbating the issue (Arnell, 2004; IPCC, 2014), there are other significant factors at play. The increasing global population and development of nations means that demand will increase, and figures show that we are already putting immense strain on current water resources (Falkenmark, 1998; as discussed in Ray & Shaw, 2016). Moreover, poor resource management, such as in Kazakhstan (Peachey, 2004), results in water resources that are spread too thinly. The development of cultivars that are more salt-tolerant is one measure that could be taken to help alleviate the issues faced in regions affected by a lack of water security.

In addition to worsening water scarcity from prolonged droughts, climate change could also lead to more frequent and intense flooding events (IPCC, 2014) as well as greater pressure from pests and diseases (Rosenzweig *et al.*, 2001). Climate modeling systems (HadGEM2-ES (Jones *et al.*, 2011) and IPSL-CM5A-LR (Dufresne

et al., 2013)) have predicted shifts in the areas of land that are currently suitable for growing the crops simulated by these models. Although gains in yield production are predicted in some regions, these are expected to be offset by losses in productivity in others. The models predict that there could be reductions in global yield of between 10 and 40 percent for some of our most important crops (*Oryza sativa* (rice), *Triticum aestivum* (wheat) and *Zea mays* (maize)) by 2050 (Müller & Robertson, 2014). This is similarly paralleled with simulations modeling the situation if the Green Revolution never happened, which has been reviewed by Evenson and Gollin (2003). The simulations show that although developed nations would have increased agricultural outputs by up to 4.8% (largely due to increasing food prices incentivising more intensive agricultural production), the developing world would have seen a fall in crop yields of up to 23.5% (Evenson & Gollin, 2003).

In light of these predictions, it is clear that advances must be made in researching ways to produce new varieties of crops capable of generating higher yields, while also being more tolerant of harsher environments and changing climatic conditions.

1.1.2 Major problems with plant productivity due to soil quality

Soil quality is important for crop production and can have impacts on yield and quality (D'Hose *et al.*, 2014; Fan *et al.*, 2012; Mueller *et al.*, 2012). Conditions in the soil can either worsen or help to alleviate issues created by a range of abiotic stresses including drought, salinity, flooding and nutrient deficiency. Biotic factors such as pathogens are also important and those that are soil-based provide specific challenges for plant defence mechanisms in the roots. Plant responses to the

environment are mediated by hormones and this is discussed in the context of this thesis in section 1.3.

Soil nutrient deficiencies account for the largest gaps in global crop yields between maximum potential production and actual recorded yields (Mueller *et al.*, 2012). If efforts were made to balance nutrient conditions in the soil or improve root uptake of nutrients, this could be enough close gaps in up to 73% of areas with the greatest underperforming yields (Mueller *et al.*, 2012). Nitrogen, phosphorus and potassium are classically considered as the three fundamental macronutrients essential for plant growth and development. Nitrogen is a component of nucleic acids, forming the basis of amino acids and is a key component of chlorophyll (Masclaux-Daubresse *et al.*, 2010). Without sufficient nitrogen, plants can suffer chlorosis (yellowing) of the leaves and stunted growth (Masclaux-Daubresse *et al.*, 2010). Phosphorus is necessary for seed germination and root growth, as well as being vital for the formation of nucleic acids, adenosine triphosphate (ATP) for energy transfer and phospholipids for cellular membranes (Hawkesford *et al.*, 2012). Phosphate is particularly important in the early stages of growth and development, and stunted growth or plants taking a long time to mature are typical signs of plants lacking in phosphate (Hawkesford *et al.*, 2012). Potassium is less crucial than the other two macronutrients, but is important for water transport, enzyme activation, plant resilience, and potassium also aids in the efficient use of nitrogen (Hawkesford *et al.*, 2012). When plants are deficient in potassium, they can share a similar appearance to plants deficient in nitrogen, in addition to suffering poor resilience to environmental changes and diseases (Hawkesford *et al.*, 2012). Aside from these three ‘classical’ nutrients, there is a growing belief that sulfur and magnesium are just

as important for plant growth and development as nitrogen, phosphorus and potassium, and that a greater emphasis needs to be placed on these 'forgotten' macronutrients (Anjum *et al*, 2015; Cakmak & Yazici, 2010).

Mueller *et al.* (2012) determined that the second largest gap in global crop yields was down to insufficient water. Plants need water for germination, photosynthesis and transport and both soil structure and quality can have a big impact on how quickly or slowly water travels through the soil and thus its availability to the plant. In combination with unfavourable environmental conditions, this can either lead to soil that is too dry, producing drought conditions, or flooded and waterlogged soils that result in hypoxic conditions. Soil salinity is also a huge issue (Deehan & Taylor, 2002; Rengasamy, 2006; Sairam & Tyagi, 2004), which can not only mimic the conditions of drought but also creates further problems and is much more difficult to resolve than merely a lack of water (section 1.1.4).

This thesis will focus on the problems of salinity and nutrient deficiencies, with specific reference to sulfur.

1.1.3 Sulfur deficiency and its effects on plants

Sulfur is an essential plant nutrient needed to make the amino acids cysteine and methionine (Dijkshoorn & van Wijk, 1967) as well as other compounds like glutathione, an important antioxidant compound involved in stress response (Noctor *et al*., 2002). Sulfur is involved in a number of vital processes for growth and metabolism and without enough sulfur the plant has a reduced ability to cope with abiotic and biotic stresses (Kruse *et al.*, 2007). Once considered a "secondary"

nutrient, it is beginning to be viewed as being potentially as important as the three macronutrients: nitrogen, phosphorus and potassium (Anjum *et al*, 2015).

The levels of nutrients in the soil are critical for healthy plant growth and sulfur, in particular, is regularly replenished through atmospheric deposition. This is generally via the products of air pollution and historically levels of sulfur in the soil have allowed greater crop yields, boosted by sulfuric acid falling as acid rain (Lewandowska & Sirko, 2008; Mazid *et al.*, 2011). However, as levels of pollutants have increasingly diminished in recent decades, nutrient deficiencies in the soil have followed and farmers have observed an upsurge in problems due to sulfur stress (Lewandowska & Sirko, 2008; Mazid *et al.*, 2011). Nutrient availability is a key limiting factor in crop yields (section 1.1.2) (Loomis & Conner, 1992; Mueller *et al*., 2012), yet at the same time the incorrect use of fertilisers is economically and environmentally harmful (Galloway *et al.*, 2003). Consequently, efforts should be made to develop plants with a greater tolerance to poor nutrient availability, or the capacity to increase their nutrient uptake.

Yellowing leaves are signs of both sulfur and nitrogen deficiency, however unlike nitrogen stress which affects older growth, under sulfur stress the yellowing occurs in new, younger leaves (Hawkesford *et al.*, 2012). Plants will also suffer stunted growth and mature plants often remain small. Aside from affecting crop yield, insufficient sulfur can negatively impact produce quality, and this appears particularly important for cereal crops. Studies have shown that sulfur is important for the quality of wheat used in baking (Zhao *et al.*, 1999) as it improves the composition of gluten proteins (Zörb *et al.*, 2009). Sulfur also plays a role in minimising acrylamide formation in cooked wheat products, since wheat grown in sulfur deficient fields has

been shown to contain higher levels of asparagine – the amino acid that reacts with sugars to form acrylamide when exposed to high temperatures during cooking (Muttucumaru *et al.*, 2006). The malting quality of *Hordeum vulgare* (barley) is also enhanced when crops are grown with enough sulfur (Zhao *et al.*, 2006). Members of the *Brassicaceae* family, which includes many crop varieties and *Arabidopsis thaliana* (hereafter referred to as *Arabidopsis*), are naturally high in sulfur and brassica crops deliver an important dietary source of glucosinolates – sulfur containing compounds associated with protection against certain diseases like cancer (Johnson, 2002; Traka, 2013).

1.1.4 Soil salinisation and its effects on plants

Climate change is likely to be a significant driver in generating more frequent and prolonged periods of drought and additionally, soil degradation through physical and chemical processes such as salinisation is also worsening. Salinisation of the soil is a particularly critical concern, as most plants do not tolerate salt (Sairam & Tyagi, 2004). Unfortunately, the issue is deteriorating, in part due to poor irrigation practices, and consequently is leading to the loss of land suitable for growing crops (Deehan & Taylor, 2002). Irrigation water contains dissolved salts that will accumulate in the soil over time, since these salts are left behind when excess water is evaporated (Rengasamy, 2006; Smedema & Shiati, 2002). In areas lacking adequate drainage irrigation can raise the water table enough to create an overlap between the root zone and salinised groundwater (Rengasamy, 2006; Sharma *et al.*, 2017). Despite the concerns over salinisation, there is a lack of data regarding its agricultural impact. It is loosely estimated that between 0.3 and 1.5 million ha of

arable land becomes unusable each year and that the quality of a further 20 to 40 million ha of land is reduced due to salinisation (FAO & ITPS, 2015).

The effect of too much salt in the soil manifests in two different negative impacts on plant growth. Firstly, the increase in uptake of Na^+ and Cl^- ions results in cellular toxicity that disrupts membranes and impairs metabolism (Krasensky & Jonak, 2012). Moreover, an increase in accumulation of these ions has the effect of creating an imbalance in the ion composition of the plant. The structure of Na^+ is similar to K^+ and so an increase in sodium will create competition for potassium in the plant, thereby reducing access to an important nutrient (Grattan & Grieve, 1999). Secondly, if the level of soil salinity becomes severe enough, physiological drought in the plant will follow (Krasensky & Jonak, 2012). The effect of drought is to increase the osmotic potential of the cytoplasm, which in turn reduces osmotic pressure in the roots to make them shrink (Krasensky & Jonak, 2012). Furthermore, soil salinity can indirectly impact plant growth by reducing the ability of the plant to uptake nutrients from the soil. Salinity decreases the solubility of phosphate, thus causing plants to become deficient even when the soil contains a sufficient amount (Hu & Schmidhalter, 2005).

Numerous studies have highlighted the diverse plasticity of the root system in response to salinity, for instance, salt can induce an auxin-mediated tropic response in roots by causing them to grow away from areas of higher salt (Galvan-Ampudia *et al.*, 2013). Data are sometimes contradictory though and responses may depend on the level of salt stress. Wang *et al.* (2009) found that low salt stress (25 mM) resulted in reduced branching of the root system (section 1.2), but that higher salt stress (100 mM and 150 mM) saw an increase in the density and length of root branches, or

lateral roots (LRs – explained in detail in section 1.2). The apparent increase in LR density could be due to the significant reduction in primary root length (section 1.2.1) observed under these conditions, because the group also found that salt stress suppressed initiation (section 1.2.3) of LRs. A more recent study has confirmed this result in maize (Lu *et al.*, 2019). Moderate (100 mM) and high (140 mM) salt stress can create a period of quiescence in the emerged LRs of young *Arabidopsis* seedlings (Duan *et al.*, 2013; Geng *et al.*, 2013). Moreover, larger reductions in relative root length are observed in LRs than are seen in the primary root, indicating that LRs are more strongly affected by salt stress (Duan *et al.*, 2013).

In some crop species salt stress creates a greater root surface area, predominantly brought about via an increase in root hair length and density: root hairs are elongated single epidermal cells that extend into the surrounding medium to increase root surface area (Grierson & Schiefelbein, 2002). A recent study in oilseed rape identified an increase in root surface area of around 20 percent in plants grown under 100 mM salt stress (Arif *et al.*, 2019). Most notable were increases in root hair length and density but also an increase in length of third-order LRs (Arif *et al.*, 2019). In *Manihot esculenta* (cassava) low salt concentrations (10 mM) increased the surface area and number of fibrous rootlets (Cheng *et al.*, 2016).

Plant breeding often neglects the roots in favour of the aboveground parts of the plant and yet there are many aspects of the roots that could offer ways to improve crop yield, resilience and quality (Den Herder *et al.*, 2010). For example, an evaluation of different lines of *Cicer arietinum* (chickpea) demonstrated substantial genetic variation in several root traits, with an increase in root length density and root mass being associated with tolerance to terminal drought stress (Serraj *et al.*, 2004).

The ability of plants like chickpea to adapt the root system to water-limited soil conditions is important for maintaining crop yield (Purushothaman *et al.*, 2017).

1.2 The plant root system

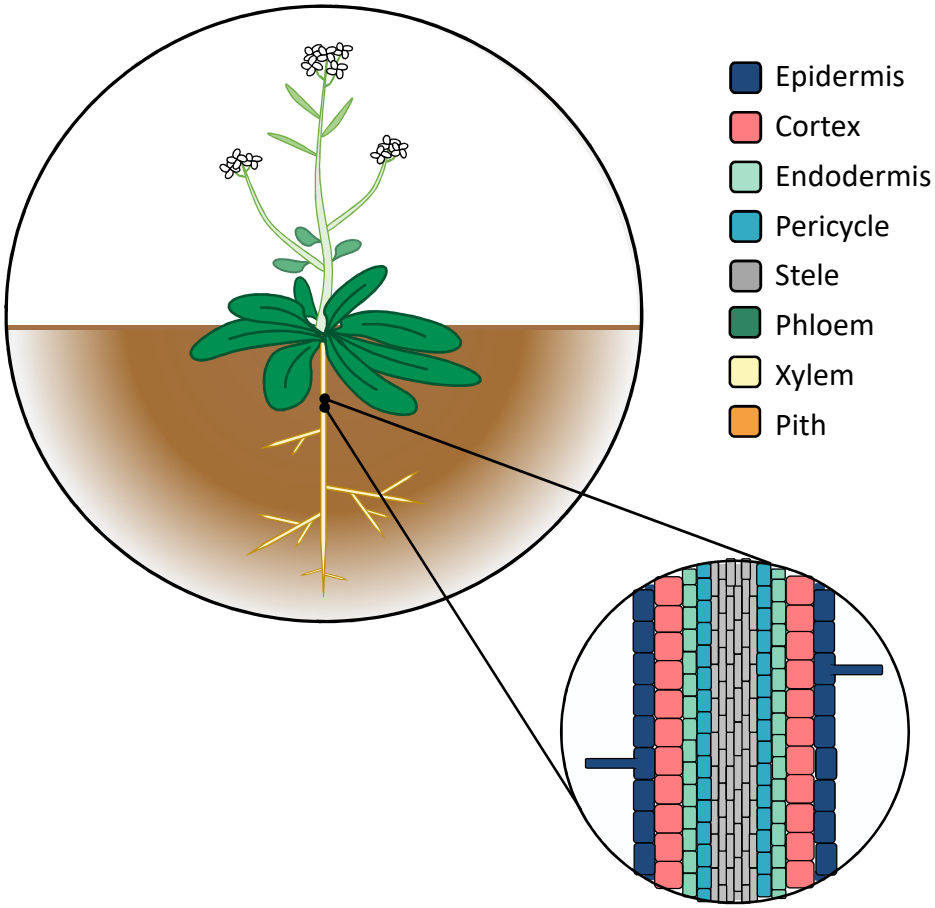
1.2.1 Root structure and function

The principal function of the roots is to act as the sensors and providers of water and inorganic nutrients to the growing parts of the plant. They are also necessary to anchor and support the above ground structure of the plant, they provide a site for nutrient storage, and they allow the plant to sense and respond to abiotic and biotic signals in the soil, such as in response to stress or inter-plant competition (Nibau *et al.*, 2008). Plants have to respond to range of biotic (pests and diseases) and abiotic stresses (physical and chemical) and the roots can often be the first part of the plant to be affected by stress. In order to facilitate these functions, the plant produces a system of branching LRs that originates from the primary root and, in turn, each root is coated in single-celled root hairs which further increase the surface area of the root system (Malamy & Benfey, 1997a). Together, the total structure of these roots makes up the root system architecture of the plant, which can be manipulated by the plant in response to stresses and environmental changes. Modifications to root architecture can include changing the angle and rate of growth or growing more of a different type of root (Rogers & Benfey, 2015).

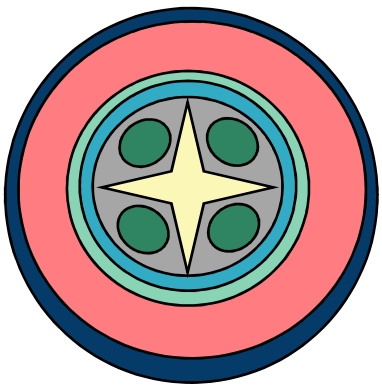
The structure of the dicot root (Figures 1.1A-B) consists of a vascular tissue core, or stele, comprised of the xylem surrounded by the phloem. In monocots (Figure 1.1C), the vascular bundles are arranged in a ring shape encircling the pith – a spongy layer involved in the storage and transportation of nutrients that is

Figure 1.1. Diagram of mature Arabidopsis. (A) a magnified longitudinal cross-section of the cellular structure of the Arabidopsis root, (B) transverse cross-section of a typical dicot and (C) monocot root. *Epidermis (dark blue), cortex (pink), endodermis (mint), pericycle (blue), stele (grey), phloem (green), xylem (pale yellow), pith (orange).*

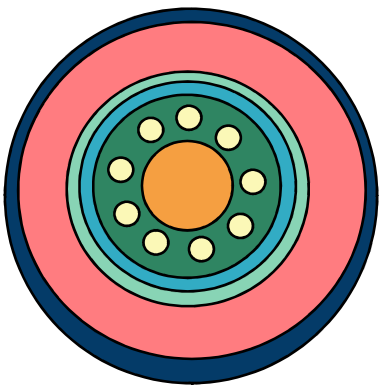
A


Epidermis
Cortex
Endodermis
Pericycle
Stele
Phloem
Xylem
Pith

B



C



restricted to the stem in dicots. In both monocots and dicots, this central vascular structure is then surrounded by the pericycle, followed by the endodermis, cortex and finally the epidermis (Dolan *et al.*, 1993). Root hair cells, which increase the surface area of the root for absorption also form part of the epidermal layer. The first root to emerge from the seed is the radicle and this matures to become the primary root.

The root system architecture of plants can be vastly different both between and within species. In dicotyledonous (dicot) species like *Arabidopsis*, an allorhizic system forms the structure of roots as they develop, whereas monocotyledonous (monocot) species generally have a secondary homorhizic root system (Osmont *et al.*, 2007). The allorhizic system is comprised of a central root, known as a taproot, which typically grows vertically into the soil. LRs branch off from the taproot, followed by subsequent orders of additional branching roots (Nibau *et al.*, 2008; Osmont *et al.*, 2007). The secondary homorhizic root system often resembles a fibrous, mat like structure.

Both dicots and monocots first develop a primary root and the continued importance of this root as the plant ages varies from species to species, as in some dicots it may develop into a substantial storage organ (for example *Daucus carota* (carrot) and *Raphanus raphanistrum* subsp. *sativus* (radish)) (Esau, 1940). Meanwhile, the role of LRs is to increase the surface area of the root system, thereby allowing the plant to explore further into the surrounding environment. Other types of root include adventitious roots, of which the majority are usually found in monocot root systems (Osmont *et al.*, 2007). These are branching roots that develop from the stem at the point where it meets the root and, along with LRs, help contribute towards a shallow root system that can tap into phosphate reserves often located in

the top layers of the soil (Steffens & Rasmussen, 2016). Importantly, the number and position of new LRs is not predetermined and can be altered by the plant according to the availability of water and nutrients (Malamy & Benfey, 1997a). The abilities of LRs in increasing root surface area and facilitating the plasticity of the root system makes them so relevant to study from an agricultural perspective and has led to a large proportion of research being dedicated on uncovering how LRs develop.

Plant growth and development, including LR development, is affected by a wide range of environmental signals and stresses, and plant responses to the environment rely on signal integration by hormones (section 1.3). For example, the hormone abscisic acid (ABA) is heavily implicated in plant stress responses (Bari & Jones, 2009) and, with respect to LRs; ABA appears to have a role in LR development, which is examined in section 1.3.2.

1.2.2 Lateral root development

During LR formation there are four distinguishable stages that have been identified: priming, initiation, patterning and emergence (Lavenus *et al.*, 2013; Malamy & Benfey, 1997b; Péret *et al.*, 2009) and a complex regulatory system controls their development. This is through a combination of environmental, hormonal and protein signaling, some elements of which are poorly understood (Lavenus *et al.*, 2013).

The formation of new roots is different to new tissue growth of other plant organs in that roots originate from within the interior layers of the tissue and thus must safely emerge through surrounding structures without disrupting plant integrity: new LRs form in the pericycle in most vascular plants (Dolan *et al.*, 1993; Péret *et al.*,

2013). The selection of pericycle cells that have the potential to initiate a new LR occurs in the basal meristem through a process known as priming. Not all pericycle cells have the ability to become LRs and there are evolutionary differences, for example in ferns such as *Ceratopteris richardii* new LRs are derived from the endodermis (Hou *et al.*, 2004). In *Arabidopsis* LRs can only form from pericycle cells at the xylem poles (Dubrovsky *et al.*, 2000) and this remains true for the majority of species (De Smet *et al*., 2006). There are exceptions however, as in some species it is the cells overlying the phloem pole that possess the capacity to divide and form new LRs. Examples of phloem pole pericycle-derived LRs include those of *D. carota* as well as several cereal crops (De Smet *et al*., 2006).

Perhaps the best well-known regulatory control of LR development is the role of the plant hormone auxin (Benková & Hejátko, 2009; Overvoorde *et al.*, 2010; Péret *et al.*, 2009) and its role in LR initiation and emergence is discussed below.

1.2.3 Lateral root initiation and patterning

The initiation of a new root is an auxin-dependent step (Casimiro *et al.*, 2001). The stages of LR initiation are visibly distinct and were first described as eight discrete steps by Malamy and Benfey (1997b). The process begins with the asymmetrical division of two cells in the pericycle that have been primed as founder cells (Casimiro *et al.*, 2001). The resultant daughter cells are smaller and will remain between the larger parental founder cells as symmetrical division of these daughter cells continues and the LR primordium is formed (Casimiro *et al.*, 2001; Dubrovsky *et al.*, 2001). Subsequent anticlinal and periclinal divisions result in the primordium having an arched shape (Casimiro *et al.*, 2001; Dubrovsky *et al.*, 2001). Figure 1.2

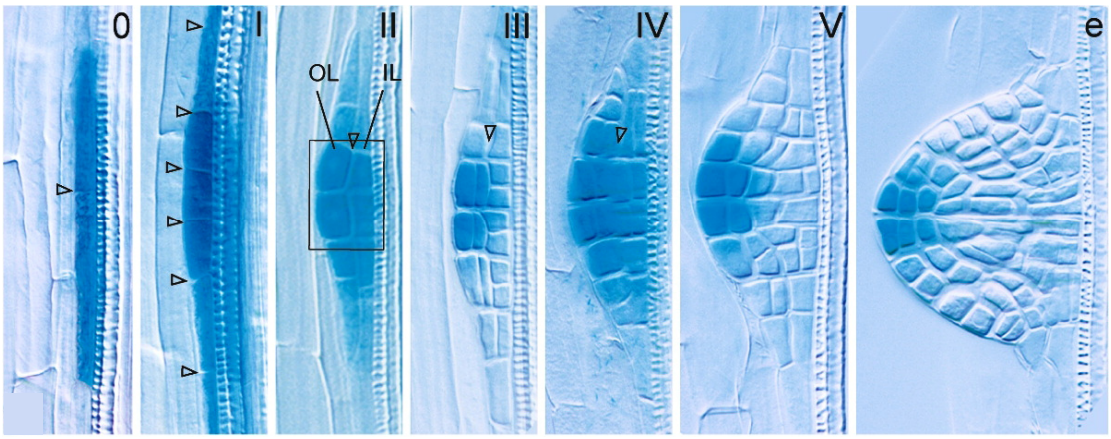
shows what happens to the pattern of auxin activity during LR development using a *DR5::GUS* reporter system (Benková *et al.*, 2003). DR5 activity was initially recorded in all cells of the primordium (stage I), before becoming more centralised as the outer and inner layers form (stage II). A gradient develops as the LR begins to take shape; with most DR5 concentrated at the tip of the primordium (stages III-V) and by stage 'e' the LR has emerged (Benková *et al.*, 2003) (Figure 1.2).

The organisation of cells in the LR primordium is referred to as patterning and begins around stage VI of LR initiation. It involves the expression of several genes that help to define the boundaries of the LR primordium including *AtPUCHI*, an auxin-induced gene encoding am AP2 family transcription factor (TF) (Hirota *et al.*, 2007). AP2 TFs are part of the wider ethylene response factor (ERF) family and there has been recent interest in the role of *AtPUCHI* in root organogenesis. Trinh *et al.* (2019) have established that *AtPUCHI* is responsible for very long chain fatty acid (VLCFA) biosynthesis in LR formation. VLCFAs are components of cellular membranes and protective barrier layers (section 1.5) that influence the properties of these layers during cell proliferation in LR initiation (Bach *et al.*, 2013). The roles of these genes are crucial to develop the highly organised, arched shape derived from cell division during patterning that will eventually arise as the meristem of the new LR.

1.2.4 Lateral root emergence

Auxin is required for new LRs to emerge from the endodermal, cortical and epidermal tissues before reaching the soil (Dolan *et al.*, 1993; Péret *et al.*, 2013). The connectivity of plant cells via their rigid cell walls means that tissues surrounding the LR primordia are mechanically constrained and thus must be broken in order for the

Figure 1.2. Auxin distribution in the developing lateral root. *DR5::GUS expression (in blue) represents auxin activity during LR development. At stage 0 DR5 activity was only detected within the pericycle at the site of lateral root initiation. OL = outer layer of dividing cells, IL = inner layer, e = emergence, central cells outlined. (Figure reproduced from Benková et al., 2003).*


0
I
II
OL
IL
III
IV
V
e

root to emerge. This is facilitated through the interplay between a series of chemical signals and mechanical forces, with the endodermis playing a pivotal role (Stoeckle *et al.*, 2018). Endodermal tissue is highly specialised and its impregnation with hydrophobic substances like the lignin-based Casparian strip (section 1.4.2) and suberin lamellae (section 1.5.2) pose a specific challenge for the LR primordia to navigate (Stoeckle *et al.*, 2018; Vermeer *et al*., 2014). Overlying endodermal cells shrink in size by losing water in order to make room for the developing LR and maintain the integrity of Casparian strip function. The level of volume loss is severe enough that the upper and lower plasma membranes are able to fuse together, creating a channel (Vermeer *et al.*, 2014). Auxin regulated aquaporins such as Plasma membrane Intrinsic Proteins (PIPs) enable this water loss, as shown by Péret *et al.* (2012), who identified that both *At*PIP2;1 over-expressing and *Atpip2;1* mutant plants slowed LR emergence by altering the flow of water and creating a flattened LR primordia instead of the typical dome shape.

Once past the endodermis, cell wall remodelling is necessary to allow the developing LR to pass through the cortex and epidermis (Péret *et al.*, 2009; Péret *et al.*, 2012). *AtLAX3* encodes an auxin influx carrier protein with expression limited to the cortical and epidermal cells overlying LR primordia (Swarup *et al.*, 2008). Auxin induces the upregulation of *AtLAX3,* and a positive feedback loop further accumulates auxin, this, in turn activates a series of cell wall remodelling genes that enable cell separation, allowing the developing LR to pass through and emerge into the soil (Swarup *et al.*, 2008).

1.3 Hormonal regulation of root development

1.3.1 The role of hormones in plant growth and development

Hormones are involved in regulating all areas of plant growth and development. For example, auxin is essential for plant growth and regulates cell elongation, division and differentiation (Ljung, 2013), jasmonates are important for defence against stresses (Bari & Jones, 2009) and ethylene is responsible for flower development, fruit ripening and the abscission of leaves (Lin *et al.*, 2009). Other examples of plant hormones include ABA, brassinosteroids, cytokinins, ethylene, gibberellin and strigolactones. Hormones can perform overlapping functions and for example, while ABA has important functions in developmental processes like germination and stomatal regulation (Christmann *et al.*, 2006), it has also been strongly associated with a role in stress tolerance (Bari & Jones, 2009; Christmann *et al.*, 2006). Generally, hormones are involved in signaling pathways in response to environmental or exogenous signals, however they may also respond to intrinsic, endogenous signals as part of gene networks (Friml, 2003). An example of this is in the auxin response in roots. Roots react to the exogenous signal of gravity, known as gravitropism, which ensures they grow downwards in the direction of gravitational pull and this is mediated by auxin (Chen *et al.*, 1999). At sites of LR initiation, auxin responds to endogenous signals and is transported from the primary root during LR initiation and from the shoot during LR emergence (Benková *et al.*, 2003).

ABA has a role in regulating root development under salt stress (Duan et al., 2013, Julkowska et al., 2014), though the mechanism of this regulation is not conserved between species. A recent study has attempted to clarify the relationship between ABA and auxin in maize root development (Lu et al., 2019). Salt stress

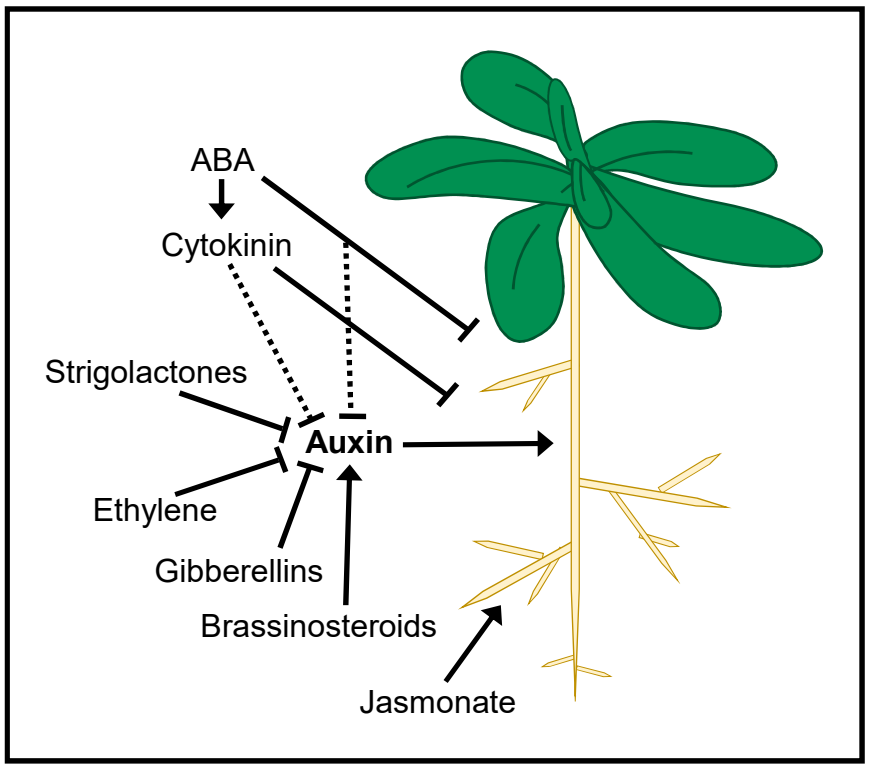
increased ABA in maize roots, particularly at sites of LR initiation and in existing LRs. This accumulation of high concentrations of ABA is thought to perturb auxin distribution, once again via *Zm*PIN1, but instead by disrupting the polar localisation of *Zm*PIN1 (Lu et al., 2019). Hence, initiation is blocked, unlike in *Arabidopsis* where the effects of ABA and salt have been shown on LR emergence and growth but not initiation (section 1.1.4) (Duan et al., 2013, Julkowska et al., 2014). However, Wang *et al.* (2009), demonstrated the role of auxin in salt-induced inhibition of LR initiation (section 1.1.4) and Liang *et al.* (2007) found low levels of ABA promote LR initiation in *M. truncatula*.

1.3.2 Hormonal regulation of root system architecture

Aside from the crucial role of auxin in LR development examined in sections 1.2.3 and 1.2.4, several other hormones are associated with regulating root system architecture and LR development (Figure 1.3), though these hormones can frequently be found acting in conjunction with auxin in these roles. For example, brassinosteroids work alongside auxin to positively regulate LR formation (Bao *et al.*, 2004) and auxin is responsible for promoting primary root elongation via the modulation of another hormone – in this case gibberellin (Fu & Harberd, 2003). Conversely, since auxin is so prolific in influencing LR development, the function of many other hormones and molecules is to interrupt a particular part of the auxin signaling pathway, such as by reducing auxin transport (Lavenus *et al.*, 2013). For example, cytokinins operate to suppress auxin-mediated LR initiation via *At*PIN protein degradation in LR founder cells (Laplaze *et al.*, 2007), hence interrupting the

Figure 1.3. Hormonal regulation of lateral root initiation

Full lines represent known interactions. Dashed lines represent possible interactions. Arrows indicate positive interactions while flat arrow heads indicate negative interactions. (Figure adapted from Verstraeten et al., 2014).


ABA
Cytokinin
Strigolactones
Auxin
Ethylene
Gibberellins
Brassinosteroids
Jasmonate

polar auxin transport that that creates the auxin gradient associated with LR development (Benková *et al.*, 2003).

1.4 Transcription factors regulating lateral root development

1.4.1 The MYB family transcription factors

The effects of plant hormones on LR development are ultimately mediated by changes in gene expression (Péret *et al.*, 2009; Voß *et al.*, 2015) that are fundamental in producing the variation necessary for survival and adaptation in response to environmental signals. TFs bring about these changes in expression and play a principal role modulating transcription, working with other, interacting, proteins to either promote or inhibit the process, thus regulating the rate of mRNA transcribed from DNA.

One abundant group of transcription factors is the MYB family. TF families are generally highly conserved and members of these protein families, including MYBs, can be found throughout many species of animals, plants and yeast (Kranz *et al.*, 1998). MYBs can often be identified through their shared DNA binding domain that is well conserved, yet their protein sequences can vary considerably between species. In plants, an extensive group of MYBs has been identified with a large number of them containing the R2R3-type MYB domain (Romero *et al.*, 1998). These proteins are responsible for a wide range of processes in the plant such as metabolism, stress response and development (Dubos *et al.*, 2010; Stracke *et al.*, 2001). The R2R3 domain structure is characterised by two helix-turn-helix repeats of around 53 amino acids each (Kanei-Ishii *et al.*, 1990).

1.4.2 MYB transcription factors important for regulating root development

A number of MYBs have been implicated in root development. *At*MYB77 is suspected to have a role in regulating LR development under changing environmental conditions (Shin *et al.*, 2007). The gene was found to be involved in auxin response through its interaction with auxin response factors (ARFs), and *Atmyb77* mutants exhibited reductions in LR density compared to wild-type (WT) plants when either treated with auxin or under low nutrient conditions (Shin *et al*., 2007).

Another MYB, *AtMYB36,* is important for Casparian strip formation (Kamiya *et al*., 2015). The Casparian strip is a band of thickened cell wall found on the radial and transverse walls of the root endodermis and is composed primarily of lignin (Naseer *et al.*, 2012) with a layer of suberin surrounding the remainder of the endodermal cell. The presence of suberin, which is discussed in more detail in section 1.5.2, results in a hydrophobic layer that blocks water movement. The main function of the Casparian strip is to prevent water and solutes from entering the stele and instead direct them to pass across the plasma membrane and flow through the cytoplasm of the endodermis (Naseer *et al.*, 2012). This is important for selective absorption and also defence. MYB TFs such as *AtMYB36* are involved in the formation of the Casparian strip and Kamiya *et al.* (2015) uncovered that *AtMYB36* directly and positively regulates the expression of several genes required for Casparian strip development. The group also showed that *Atmyb36* mutants fail to properly develop the structure of the Casparian strip and its function as a barrier is impaired (Kamiya *et al.*, 2015).

AtMYB93 has been identified as another regulator of root development. *At*MYB93 negatively regulates LR formation through a complex, auxin-induced

negative feedback loop that helps to ensure a new root is only produced when needed, thus avoiding unnecessary energetic expenditure (Gibbs *et al.*, 2014; Gibbs & Coates, 2014). In Gibbs *et al*. (2014) *Atmyb93* mutants were shown to have a higher LR density than WT and mutants also developed their LRs more quickly. Two other MYBs have been identified that are closely related to *AtMYB93* and can be found within the same evolutionary clade: *AtMYB92* and *AtMYB53* (Du *et al.*, 2015). *AtMYB92* is highly expressed in the roots and to a lesser extent in the germinating seed, stem, flowers and siliques (Gibbs *et al.*, 2014; Winter *et al*., 2007). Unlike *AtMYB93*, *Atmyb92* mutants were not found to have a LR phenotype, although examination of double mutants revealed *Atmyb92* to enhance the *Atmyb93* mutant phenotype (Gibbs *et al.*, 2014). *AtMYB53* also has the most widespread expression profile of the three genes and while it is once again strongly expressed in the roots, *AtMYB53* is also present in germinating seeds, leaves, stems and flowers (Gibbs *et al.*, 2014; Winter *et al*., 2007). Moreover, as *AtMYB93*, *AtMYB92* and *AtMYB53* are regulated by different hormones (Gibbs *et al.*, 2014; Kranz *et al.*, 1998, Winter *et al*., 2007), this suggests both redundant and unique functions. As of yet, *Atmyb53* mutants have not been characterised and this is something that will be examined in this thesis.

1.4.3 AtMYB93 interacts with AtARABIDILLO proteins and is restricted to the endodermis

*At*ARABIDILLO proteins are *Arabidopsis* armadillo repeat-containing proteins containing an F-box motif that are known to promote root branching (Coates *et al.*, 2006) and have been shown to interact with *At*MYB93 (Gibbs et al., 2014). The exact

mechanism of the interaction is unknown but two possible modes of action for how *At*MYB93 might interact with an ARMADILLO protein have been described by Gibbs & Coates (2014). The first proposes that *At*MYB93 is a direct inhibitor of a LR development gene while the second suggests that *At*MYB93 acts as an activator of a LR inhibition gene (Figure 1.4). This thesis will seek to distinguish these two possibilities.

*At*MYB93 promoter activity can be detected at sites of LR primordia formation (Figure 1.5.A). Its expression is localised to just a small number of cells in the root endodermis (Gibbs *et al*, 2014). The mechanism for the regulation of expression of *At*MYB93 is unknown, however a chromatin immunoprecipitation-DNA microarray (ChIP-chip) showed that *At*MYB93 was a direct target of *At*SCARECROW (*At*SCR) (a TF that is expressed in the endodermis) (Iyer-Pascuzzi *et al.*, 2011), and further confirmation of this was shown in a yeast one-hybrid screen by Sparks *et al.*, (2016). *AtSCR* belongs to the GRAS TF family and has an essential role in regulating the radial organisation of roots by determining cell fate in the endodermis and cortex (Di Laurenzio, *et al.*, 1996; Sparks *et al*., 2016). *At*SHORTROOT (*AtSHR*) is a TF expressed in the stele that is relocated to the nucleus of an adjacent single celled layer during early development and, together with *AtS*CR, forms a complex responsible for the asymmetrical cell division that gives rise to the endodermis (Cui *et al.*, 2007; Nakajima *et al.*, 2001). The binding of *At*SCR to the *At*MYB93 promoter therefore suggests the role of *At*SCR as an upstream regulator of At*MYB93* (Figure 1.5.B). Interestingly, *AtSCR* gene expression was downregulated just prior to *At*MYB93 upregulation in an experiment that recorded changes in gene expression after the induction of LR initiation (J.C. Coates, unpublished observation; Voß *et al.*,

***Figure 1.4. Phenotypes and possible model of interaction between* AtMYB93 and downstream promoters.** *AtMYB93 over-expressor and mutant lines show LR density phenotypes. 1: AtMYB93 is the inhibitor of lateral root genes and downstream targets have a similar phenotype to the AtMYB93 over-expressor, or 2: AtMYB93 activates LR inhibition genes and downstream targets have a similar phenotype to the Atmyb93 mutant. (Created using BioRender).*

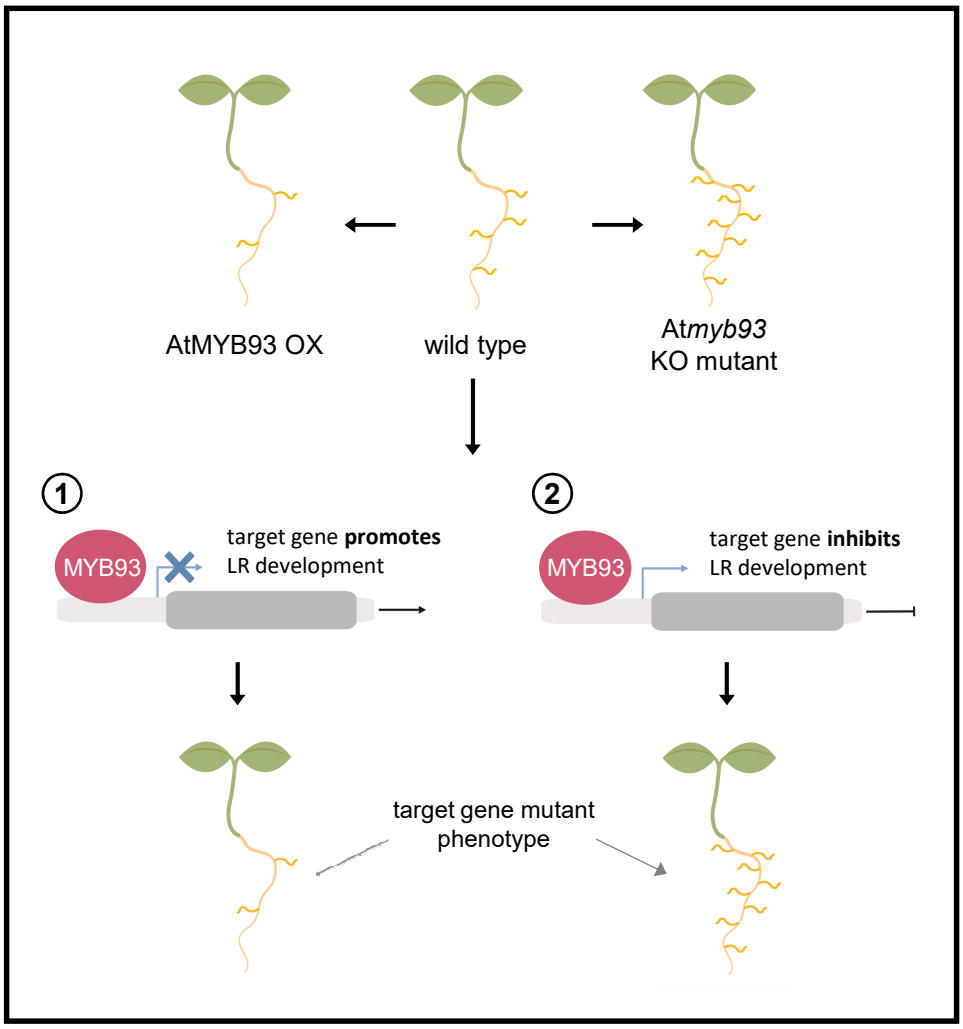
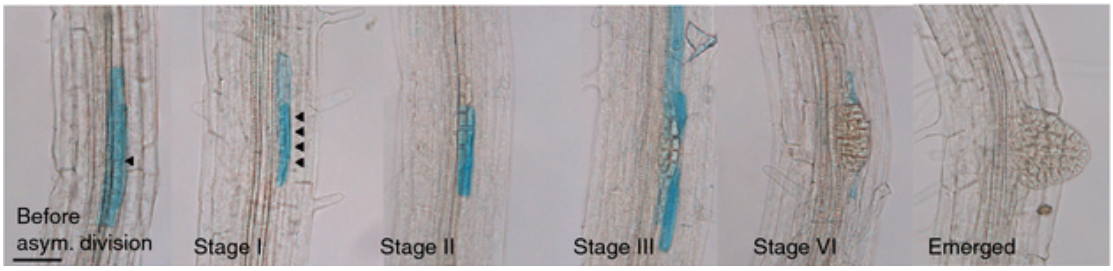
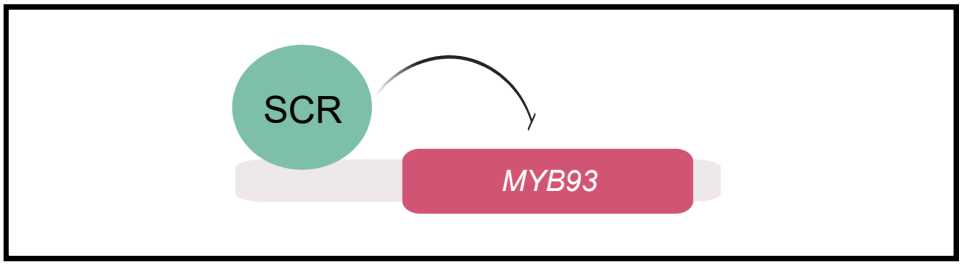

AtMYB93 OX
wild type
Atmyb93
KO mutant
1
MYB93
target gene promotes
LR development
2
MYB93
target gene inhibits
LR development
target gene mutant
phenotype

Figure 1.5. AtMYB93 expression is restricted to a limited number of endodermal cells. *(A) AtMYB93 promoter activity in the endodermal cells overlying Arabidopsis lateral root primordia demonstrated via pAtMYB93::GUS reporter expression. Expression is strongest in the early stages of cell division and after this expression tails off. Arrows indicate early cell divisions. Scale bar: 50 µm. (Figure reproduced from Gibbs et al., 2014). (B) Possible model of interaction between* AtMYB93 and a putative upstream repressor, *AtSCARECROW (AtSCR), in which the AtSCR protein binds to the AtMYB93 promoter and inhibits its expression. (Created using BioRender).*

A


Before
asym. division
Stage I
Stage II
Stage III
Stage VI
Emerged

B


SCR
MYB93

2015). It is therefore possible that *At*SCR is responsible for repressing *At*MYB93 in the majority of endodermal cells, a hypothesis that will be tested in this thesis.

1.4.4 Close relatives of AtMYB93

As covered earlier in section 1.4.2, the closest relatives of *AtMYB93* are *AtMYB92* and At*MYB53*, and together they comprise the S24 clade (Du *et al.*, 2015; Gibbs *et al.*, 2014; Stracke *et al.*, 2001). Both have been implicated in suberin regulation in the root endodermis along with *AtMYB93* (Hu, 2018). *AtMYB93*,

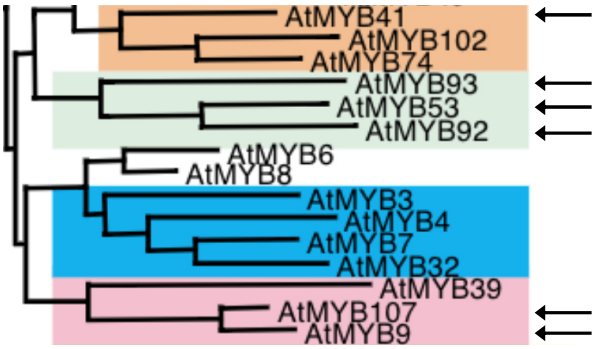
which is the focus of this thesis, is much more specific in its expression pattern and diverged at an earlier date than both *AtMYB92* and *AtMYB53*, which are thought to have evolved more recently (Figures 1.6.A and 1.6.B) (Du *et al.*, 2015; Gibbs *et al.*, 2014; Stracke *et al.*, 2001). Furthermore, *AtMYB93* homologues have been identified in other species such as *Malus domestica* (apple), *Solanum lycopersicum* (tomato) and rice (Du *et al.*, 2015; X. Cao, unpublished; Legay *et al.*, 2016).

AtMYB107 and *AtMYB9* are two genes that are closely related and have been identified are located in a neighbouring clade to that of At*MYB93* (Figures 1.6.A and 1.6.B) (Du *et al.*, 2015, Stracke *et al.*, 2001). These genes are linked to roles in regulating suberin biosynthesis in the seed coat (Gou *et al.*, 2017; Lashbrooke *et al.*, 2016) and in fruit (Legay *et al.*, 2016).

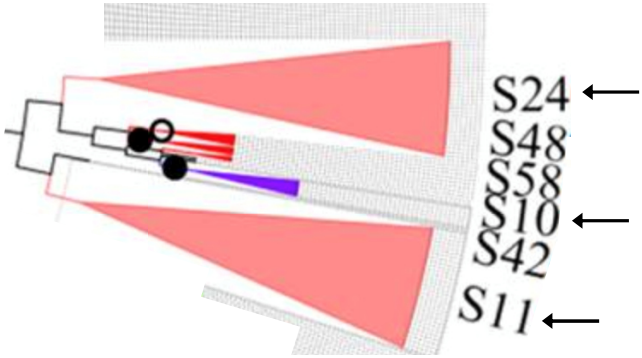
AtMYB41 is another MYB that shares much of its evolutionary history with *AtMYB93* and though it is less closely related to *AtMYB93* than *AtMYB92* and *AtMYB53*, *AtMYB107* and *AtMYB9* are more divergent still (Figures 1.6.A and 1.6.B) (Du *et al.*, 2015; Stracke *et al.*, 2001). *AtMYB41* has been identified as having a likely

Figure 1.6. Consensus and divergence between published MYB phylogenies. *The evolutionary relationship between the S24 clade members: AtMYB93, AtMYB53 and AtMYB93, S10 clade members: AtMYB9 and AtMYB107, and S11 member: AtMYB41 from published phylogenies by (A) Stracke et al. (2001), (B) Du et al. (2015) and (C) Lashbrooke et al. (2016). Arrows highlight the relevant MYBs located in subsections taken from larger published phylogenies. (Figures adapted from Stracke et al. (2001), Du et al. (2015) and Lashbrooke et al. (2016) respectively).*

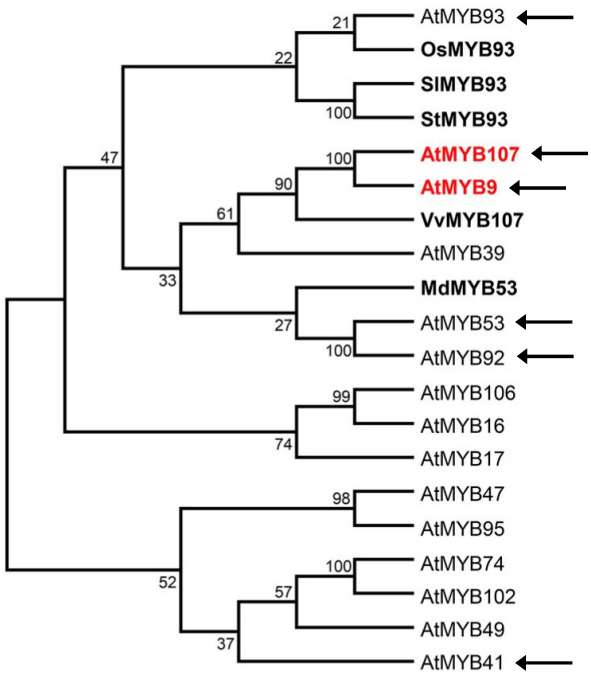
A


AtMYB41
AtMYB102
AtMYB74
AtMYB93
AtMYB53
AtMYB92
AtMYB6
AtMYB8
AtMYB3
AtMYB4
AtMYB7
AtMYB32
AtMYB39
AtMYB107
AtMYB9

B


S24
S48
S58
S10
S42
S11

C


AtMYB93
OsMYB93
SlMYB93
StMYB93
AtMYB107
AtMYB9
VvMYB107
AtMYB39
MdMYB53
AtMYB53
AtMYB92
AtMYB106
AtMYB16
AtMYB17
AtMYB47
AtMYB95
AtMYB74
AtMYB102
AtMYB49
AtMYB41
21
22
100
47
100
90
61
33
27
100
99
74
98
52
100
57
37

role in suberin biosynthesis in the roots and seed coat in response to stress (Kosma *et al.*, 2014).

The phylogeny in Lashbrooke *et al.* (2016) presents a slightly different view of the evolutionary relationships of these MYBs and puts *AtMYB107* and *AtMYB9* in the same clade as *AtMYB53* and *AtMYB92*, with *AtMYB93* sitting outside of these four MYBs and *AtMYB41* in a separate and much more distantly related clade (Figure 1.6.C). However, this phylogeny is considered to be less convincing than the others given the consensus between the other phylogenies, in addition to experimental evidence of the different roles of each of these MYBs.

1.5 Complex biopolymers in plants

1.5.1 MYBs and complex polymers

The role of MYB TFs in biopolymer synthesis is well characterised. Lipid polyesters are a class of polymer important for providing barriers that control the movement of water, nutrients and other molecules throughout the plant. Cutin and suberin are the two major lipid polyesters in plants (Beisson *et al.*, 2012). Cutin is a waxy substance that forms the cuticle, providing a hydrophobic interface between the interior of the plant and its aerial surfaces. *AtMYB106* and *AtMYB16* both directly activate a cutin biosynthesis regulator called WAX INDUCER1/SHINE1 and a number of cutin genes (Oshima *et al.*, 2013), and the *AtMYB93* relative, *AtMYB41*, has been found to negatively regulate cutin biosynthesis in response to osmotic stress, drought and also ABA (Cominelli *et al.*, 2008). The study by Cominelli *et al* (2008) revealed that when *AtMYB41* is overexpressed its phenotype comprises more

permeable leaf surfaces and thus reduced tolerance to desiccation, alongside a downregulation of certain cutin genes, suggesting a disrupted barrier is there.

Lignin has a less similar structure to that of suberin and cutin but is still a water-tight barrier that is rigid and provides some stability to cell walls. It is the primary component of the Casparian strip, as described in section 1.4.2, and an array of MYBs have been recognised as having roles in its biosynthesis. These include *AtMYB46* and *AtMYB83*, two MYBs that are functionally redundant in their transcriptional activation of a number of other downstream TFs identified as regulators of lignin synthesis (Zhong & Ye, 2012). These include *AtMYB58* and *AtMYB63*, which are positive regulators (Zhou *et al.*, 2009), while *AtMYB4* and *AtMYB32* are examples of MYBs that have been shown to negatively regulate lignin biosynthesis genes (Jin *et al*., 2000; Preston *et al.*, 2004). Feng *et al.* (2004) looked at the role of *AtMYB68* in both root development and responses to environmental signals. They determined that expression of *AtMYB68* is limited to root pericycle cells with enriched expression at the site of LR primordia and that *Atmyb68* mutant roots had a two-fold greater lignin content than WT plants. Furthermore, expression of *AtMYB68* was modified in response to higher temperatures, suggesting regulation by environmental signals (Feng *et al.*, 2004).

1.5.2 The role of suberin and its potential regulation by AtMYB93

Suberin is an insoluble lipid polymer composed of polyaliphatic and polyaromatic components that helps provide a protective barrier to the endodermal cell wall (Beisson *et al.*, 2012; Bernards, 2002). Monocots in particular are likely to develop endodermal suberin lamellae, though it is also a common feature amongst

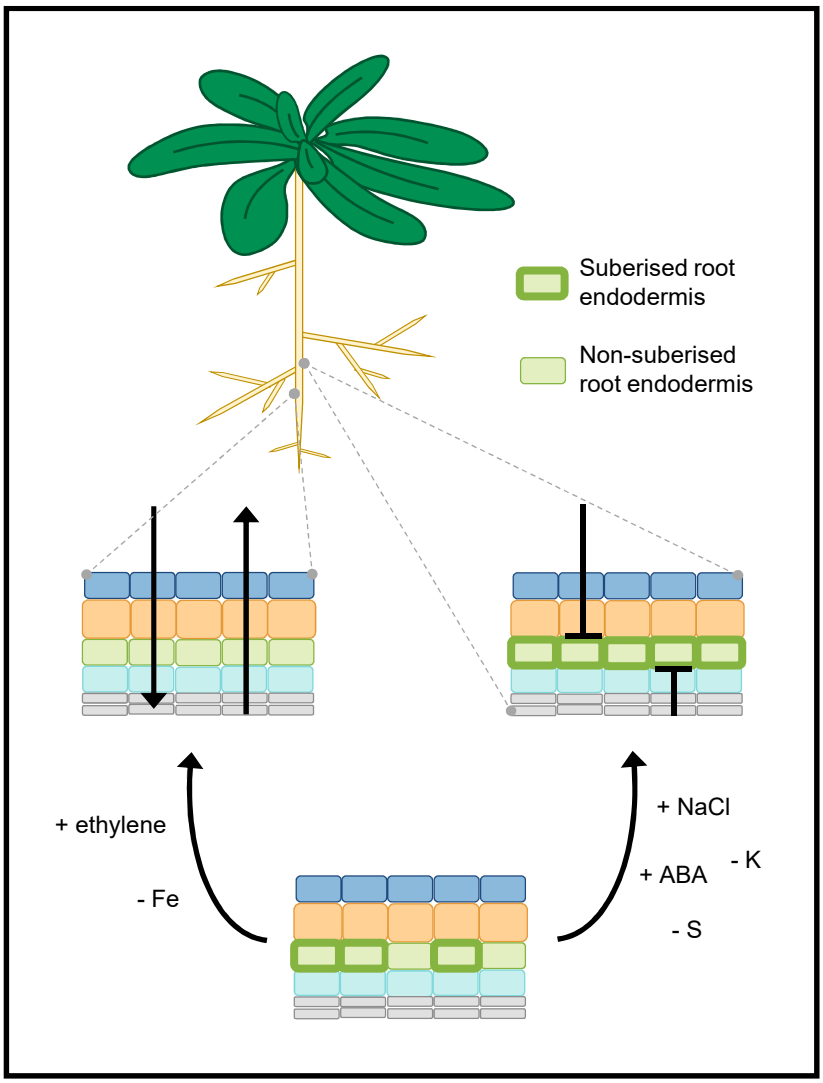
dicots and furthermore, the composition and quantity of suberin varies across species (Pollard *et al.*, 2008). Rather than a discrete layer, suberin forms within the endodermal cells as a secondary cell wall during primary root development (Anderson *et al.*, 2015). This creates a hydrophobic coating that can be switched on or off, thus alternating between active absorption and protective coating, and allowing the plant to react to a range of stresses such as from wounding, pathogens, drought and nutrient stresses (Figure 1.7.A) (Barberon *et al.*, 2016; Baxter *et al.*, 2009, Beisson *et al.*, 2012).

MYBs in the *At*MYB93 clade and in neighbouring clades such as *AtMYB53*, *AtMYB92*, *AtMYB107* and *AtMYB9* have each been implicated as positive regulators of suberin biosynthesis (Gou *et al.*, 2017; Hu, 2018; Lashbrooke *et al.*, 2016). Interestingly, while the above section highlighted *AtMYB41* as a negative regulator of cutin biosynthesis in response to stress (Cominelli *et al.*, 2008), *AtMYB41* is also implicated in suberin biosynthesis in response to similar stresses (Kosma *et al.*, 2014). There is also evidence that *AtMYB93* and *MdMYB93* are also suberin regulators (Hu, 2018; Legay *et al.*, 2016).

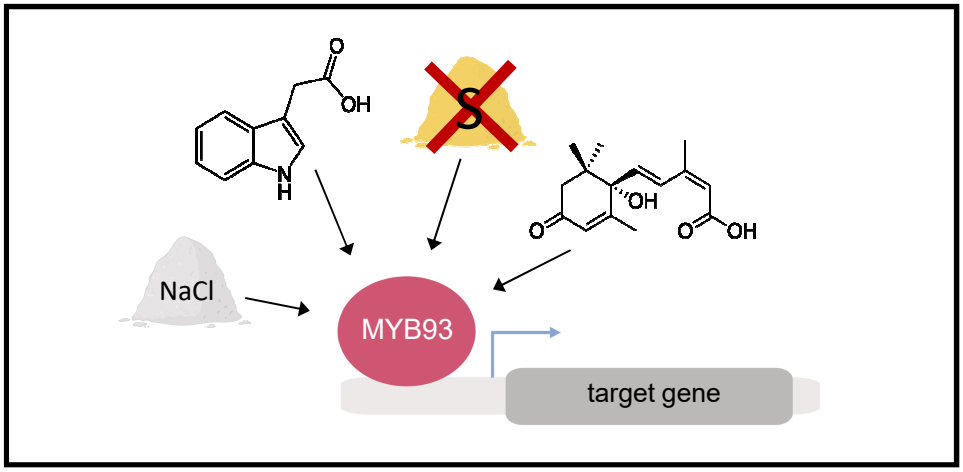
Since the suberin endodermal layer exhibits a high level of plasticity in response to changing conditions, it is likely to be important factor at sites of LR initiation, as the endodermal cells above the pericycle cells from which LR primordia first initiate must be remodelled to accommodate the developing LR primordium (Vermeer *et al.*, 2014). One hypothesis is that *AtMYB93* has an effect on the endodermal suberin layer during the initial development of new LRs either through increasing its biosynthesis, or through manipulation and remodelling. Some suberin biosynthesis genes show a similar, though less restricted, pattern of localisation as

Figure 1.7. Diagram of the effects of hormones and stresses on suberin deposition and AtMYB93 activity. *(A) Suberisation of the root endodermis creates a hydrophobic barrier. Nutrient availability, stress and hormones can augment or remove suberin deposition. (Figure adapted from Barberon et al., 2016). (B) Hormones, nutrient availability and stresses upregulate AtMYB93 expression. From left-right: salt, auxin, sulfur deficiency and ABA. (Created using BioRender).*

A


Suberised root endodermis
Non-suberised root endodermis
+ ethylene
- Fe
+ NaCl
+ ABA
- K
- S

B


OH
N
H
S
OH
O
O
OH
NaCl
MYB93
target gene

*At*MYB93. These include *AtFAR4, AtFAR5* and *AtGPAT5*, which have all been implicated in suberin biosynthesis and are located in the root endodermis (Beisson *et al.*, 2007; Domergue *et al.*, 2010), but as they are upregulated much earlier than *At*MYB93 (Voß *et al*., 2015) they are less likely to be transcriptional targets.

A study by Li *et al.* (2017) found that suberin allows the plant to maintain a diffusion barrier during the later formation of new LRs (during LR emergence through overlying cell layers). Li *et al.* (2017) suggest that suberin may be deposited to fill the gaps created from the disruption of the endodermal cell wall to make way for the emerging LR. This builds on previous research by and Martinka *et al.* (2012) which found targeted deposits of suberin in endodermal cells overlying newly developing LR primordia. Suberin therefore could be allowing the cell wall to be modified while still preserving a level of protection. The plant must also avoid making too much suberin however, as this can result in cells that do not properly expand. *At*MYB93 over-expressor mutants suggest this is the case as they exhibit a phenotype with fewer LR than WT plants (Gibbs *et al*., 2014) and additionally, *At*MYB93 over-expressors sometimes also show stunted growth and whitened leaf surfaces (E. Chapman, unpublished data). Interestingly, ABA upregulates *AtMYB93* expression (Gibbs *et al.*, 2014; Kranz *et al.*, 1998) and ABA may have an early positive role in LR initiation (Liang *et al.*, 2007), but conversely has a later inhibitory role during LR emergence (De Smet *et al.*, 2006; Duan *et al.*, 2013). The inhibitory effect of ABA corresponds with evidence suggesting that ABA increases the suberin layer in roots (Baberon *et al.*, 2016), something also echoed in response to salt stress and nutrient deficiency.

1.5.3 Suberin and the seed coat

Suberin deposition is not limited to the roots, and another region it has been identified in is as a component of the seed coat (Molina *et al.*, 2006). Suberin deposition is located on the exterior side of the seed coat, whereas the interior of the seed coat is thought to be formed primarily from cutin (Molina *et al.*, 2008). There are some differences in the composition of the suberin in roots and of the seed coat, as seed coat suberin is enriched with more long-chain fatty acid and carboxylic acids than that of roots (Beisson *et al.*, 2007; Domergue *et al.*, 2010), though ultimately its biosynthesis in both tissues is regulated by many of the same genes. For example, much like in roots (section 1.5.2) *AtFAR4, AtFAR5* and *ATGPAT5* are involved in suberin biosynthesis in the seed coat (Beisson *et al.*, 2007; Domergue *et al.*, 2010). The main role of suberin in the seed coat is thought to be in maintaining seed dormancy (Bolingue *et al.*, 2010; Fedi *et al.*, 2017; Rodriguez *et al.*, 2015).

Maximising the success of the offspring of a plant depends on germination occurring under favourable conditions and so environmental triggers are often required to break seed dormancy (Bentsink & Koorneef, 2008). In some species requiring a physical trigger to break dormancy such as *Medicago truncatula*, the role of the seed coat is to exclude water (Bolingue *et al.*, 2010). While in other species such as cereals, oxygen sensing may be used and, in this case, dormancy is considered physiological (Gibbs *et al.*, 2011; Rodriguez *et al.*, 2015). Fedi *et al.* (2017) found that *Atawake1*, a mutant allele of *AtABCG20*, which is involved in suberin deposition (Yadav *et al.*, 2014), is required for seeds to enter dormancy. Mutant plants demonstrated perturbed dormancy and increased seed coat

permeability consistent with other mutants with a reduced suberin content in the seed coat, as in the case of *At*GPAT5 (Beisson *et al.*, 2007; Fedi *et al.*, 2017).

The *At*MYB93 interactor and LR promoter *At*ARABIDILLO may also have a role in regulating seed suberin deposition. *Atara1* mutants have a reduced rate of germination when treated with ABA, while *At*ARA1 over-expressor mutants are somewhat insensitive to ABA-mediated inhibition of germination (Moody *et al.*, 2016). These results corroborate the idea of *At*ARABIDILLO inhibiting *At*MYB93 promotion of suberin, with *Atara1* leading to increased *AtMYB93* activity and *At*ARA1 over-expressors causing reduced activity.

1.6 AtMYB93 and environmental triggers

1.6.1 Abiotic factors regulating the AtMYB93 response

The hormones auxin, ABA and salicylic acid modulate *AtMYB93* expression (Chen *et al*., 2012; Gibbs *et al*., 2014; Kranz *et al.*, 1998; Singh *et al*., 2015), as do nutrient stresses such as nitrogen and sulfur starvation (Bielecka *et al.*, 2015; Scheible *et al.*, 2004) and other abiotic stresses such as hypoxia and salt (Dinneny *et al.*, 2008; Mustroph *et al.*, 2009) (Figure 1.7.B); a more detailed description of this is presented in section 3.2.2 of chapter III. Collectively these data show that *At*MYB93 may act as a regulatory hub integrating diverse environmental signals to regulate root development.

1.6.2 Sulfur stress and the AtMYB93 response

Bielecka *et al*. (2015) found that expression of *AtMYB93* increases during sulfur stress, while a study by Dan *et al.* (2007) revealed a phenotype of reduced LR

density in plants grown under sulfur deficiency as a result of auxin-mediated negative regulation in response to sulfur stress. This reduced LR phenotype has also been replicated by the Coates lab in unpublished data collected on young seedlings (R. Etherington, unpublished; B. Hutton, unpublished; N. McMulkin, unpublished; A. Oliver, unpublished). Together these results lead to the generation of two hypotheses about the role of *AtMYB93*. Firstly, we know *At*MYB93 inhibits LR development and is upregulated by sulfur stress and so *AtMYB93* may play a role in reducing LR formation in response to a lack of sulfur. It can be predicted that *Atmyb93* mutants might have more LRs under sulfur stress than WT plants and further, that mutants may have altered responses at both the root- and shoot-level under sulfur deprivation when compared to WT. This is supported by the fact that *AtMYB93* over-expressor mutants produce fewer LR (Gibbs *et al*., 2014). However, this is also in conflict with data from an earlier study by Kutz *et al*. (2002), which found that sulfur starvation leads to increased numbers of LR, although this was in older plants and only used a very small sample size. Secondly, if *Atmyb93* mutants also have less suberin, then this could lead to changes in root- and shoot-level responses to sulfur deficiency. For example, more roots or less suberin could be advantageous to the plant under reduced sulfur conditions.

In keeping with the above hypothesis of leaky roots with less suberin, shoot ion analysis has demonstrated that *Atmyb93* mutant plants are able to take up more sulfur than WT. Conversely, when grown under deficient conditions, proportionately more sulfur is lost by *Atmyb93* mutants than WT plants (Wilkinson, 2018). One caveat of this study is that by growing plants on medium that either contains or lacks sulfur, the experiment fails to mimic ‘real life’ growth conditions. That is, initially it is

likely that soil would be rich in sulfur, only to become scarce over time. It is therefore possible that the initial increase in sulfur uptake seen in *Atmyb93* mutants could be of benefit to the plant if sulfur is later deficient in the soil. Another potential source of variation could be due to differences observed in the dry biomass of tissue samples. While there was no significant difference measured between WT and *Atmyb93* tissue grown under the same conditions (either control or no sulfur), there was a significant decrease in the dry biomass of *Atmyb93* plants grown without sulfur when compared with both *Atmyb93* and WT plants grown on normal medium (Wilkinson, 2018).

1.6.3 Drought and salt stress and the AtMYB93 response

AtMYB93 gene expression is upregulated by salt (Dinneny *et al.*, 2008). Moreover, *AtMYB93* over-expressor mutants appear to be insensitive to salt stress compared to WT plants in *Arabidopsis* as they produce fewer LR under salt-stressed conditions. While over-expressors already have a reduced LR density under normal salt conditions, the addition of salt stress seems to further accentuate this phenotype (E. Chapman, unpublished data). Preliminary data has also indicated differences in Na^+ uptake between WT and *Atmyb93* mutants when grown under salt stress, but not under control conditions (Wilkinson, 2018). Wilkinson (2018) tested the differences in ion uptake of two concentrations of salt (75 mM and 100 mM) and while *Atmyb93* mutants appeared to take up slightly more sodium under the 75 mM salt stress, the increase became more pronounced under medium containing 100 mM of salt. Further research looking at the effects of stress on *AtMYB93* expression and function is needed but could eventually lead to the possibility of providing growers with crop varieties producing a range of phenotypes to best suit growth in an array of soil

conditions. Overall *At*MYB93 manipulation could represent a novel and precise way to facilitate crop improvement that involves manipulating a gene expressed only in a very small number of cells in the root.

1.7 Real world applications of modulating AtMYB93 function

1.7.1 Exploiting lateral root structure as a way to improve crops

The root system of the plant presents a key potential target for crop improvement and an adaptable LR network could help in addressing the challenges presented to future food security (de Dorlodot *et al.*, 2007). A study in maize (*Zea mays*), for example, compared plants with two different LR phenotypes (few but long or many but short) and showed that drought tolerance was improved in plants with a lower LR density (Zhan *et al.*, 2015). A modeling project, also in maize, established that variations in the presence of nutrients in the soil made a significant difference to the optimal level of LR density. Postma *et al.* (2014) used *SimRoot*, a "functional structural plant model that can simulate the growth of maize", to establish that lower densities of LR (<7 cm^{-1}) are optimal to absorb nitrate, whereas higher densities of LR (>9 cm^{-1}) are best for phosphorus.

As these studies show, the ability to manipulate the structure of the LR system could provide new crop varieties able to withstand changing climatic conditions, or even allow them to be grown in previously inhospitable areas. Root phenotypes can exhibit high plasticity in response to soil conditions (Gruber *et al.*, 2013) and such traits are often unlikely to be passed onto subsequent progeny (Maiti *et al.*, 1996; Serraj *et al.*, 2004). Molecular techniques in particular, may be relevant for future crop improvement efforts, as the utilisation of plant breeding techniques to select for

root traits can be difficult and has not traditionally been exploited. Even the breeding teams responsible for crop development during the Green Revolution did not select root architecture as an important focus (Waines & Ehdaie, 2007).

MYB93 is proposed as a good target for modifying the root system because of its specificity. LR development is controlled by a large number of genes (Wachsman *et al.*, 2015) and AtMYB93 appears to be modulating this under certain, specific conditions. The genetic manipulation of more widely expressed genes is more likely to lead to other unwanted changes in the plant, for example auxin is strongly involved in LR development but many genes associated with it have pleiotropic effects. *AtARABIDILLO* has also been ruled out as a suitable target because although both *AtARABIDILLO1* and *AtARABIDILLO2* promote LR development and hence could be targets for LR manipulation (J.C. Coates *et al.*, 2006), the *At*ARABIDILLO1 protein is unstable (Nibau *et al.*, 2011). F-box proteins are often unstable and Nibau *et al.* (2011) uncovered *At*ARABIDILLO1 to likely be degraded by ubiquitin-mediated proteolysis.

Although in many species it is likely that expression of *MYB93* will be restricted to the roots as it is in *Arabidopsis*, *MYB93* homologues have been found to be expressed in the fruit skin of apple (Legay *et al.*, 2016) and tomato (X. Cao, unpublished data). Legay *et al.* (2016) suggest that *Md*MYB93 is involved in suberin deposition in the skins of some apple varieties such as the ‘Cox orange Pippin’, in a process known as russeting. The result of this process is that fruit acquires a rough, brown skin and hence is not considered a desirable trait. While in this case it is theorised that *myb93* mutants could be beneficial to fruit development by reducing the amount of suberin in the fruit epidermis, evidence of *MYB93* expression in

different tissue types in other species should be investigated in order to ensure it remains a suitable candidate to exploit for crop improvement.

1.7.2 Translating knowledge from Arabidopsis to crops

Arabidopsis remains probably the most thoroughly studied species of flowering plant (Koornneef & Meinke, 2010). Its small stature, small genome, rapid generation time, and ease of reproduction through self-pollination make it an excellent model organism for genetic studies (Koornneef & Meinke, 2010). *Arabidopsis* is a member of the *Brassicaceae* family and so is a relevant relative to study the numerous brassica crops that are found within the family such as oilseed rape, *Brassica oleraceae* (e.g. broccoli, cauliflower and kale) and *Brassica napa* (e.g. turnip and bok choy). Yet, studies in *Arabidopsis* have even increased our understanding of aspects of plant biology seemingly more typical of much less closely related species, such as wood formation in trees. The genes involved in secondary thickening of *Arabidopsis* inflorescence stems leading to cambium formation are highly conserved in poplar trees (*Populus* sp.) (Hertzberg *et al*., 2001).

Cereals are some of the most important crop plants; directly, they form the staple of most diets throughout the world and are also used to feed livestock. They are members of the *Poaceae* grass family, a broad collection of monocotyledonous angiosperms (The Plant List, 2013). Wheat, barley and rice are three widely cultivated cereal crops and – along with maize and *Sorghum bicolor* (sorghum) – make up the top five cereals ranked by global production weight (FAOSTAT, 2018). As previously discussed in section 1.2.1, monocot root systems typically develop differently to that of dicots and while much of the progress in uncovering the factors

that determine root system architecture has been made in the dicot and model organism, *Arabidopsis*, the molecular mechanisms behind root branching in monocots is less well understood. Hence, it would be valuable to identify MYB93 homologues in monocots and, in particular, crop monocots like cereals.

1.8 Aims and objectives of the project

*At*MYB93 is a TF involved in a complex network for the regulation of LR development, though the mechanism of action of *At*MYB93 is currently unknown. The principal aims of this work are to: (i) understand how *At*MYB93 functions at the molecular level, (ii) establish how modifying *At*MYB93 affects LR development and formation of the root barrier, and therefore ultimately plant stress resilience, and (iii) uncover ways in which this work could help tackle the issue of global food security.

To address these aims, the project has the following objectives:

(i) To identify the downstream transcriptional targets and upstream regulators of *At*MYB93,

(ii) To characterise the physiological characteristics of *Atmyb93*-clade mutants and mutants in identified regulators/targets identified in objective (i).

(iii) To identify, characterise and ultimately manipulate *MYB93* homologues in crops to determine whether they could provide a promising opportunity for crop improvement.

Each objective is addressed in a subsequent results chapter (Chapters III-V) in this thesis.

Chapter II

MATERIALS AND METHODS

2.1 Species and strains used

2.1.1 Plant species

Arabidopsis thaliana – Col-0 ecotype. For experiments conducted using the following plant lines, seed was obtained from (Hu, 2018): WT, *Atmyb93-1*, *Atmyb92-1*, *Atmyb53-1*, *Atmyb53-2*, *Atmyb92-1/Atmyb93-1*, *Atmyb53-1/Atmyb92-1/Atmyb93-1* and *Atmyb53-2/Atmyb92-1/Atmyb93-1*. Where WT, *Atmyb93-1* and p35S::*AtMYB93* in Col-0 (*AtMYB93* over-expressor) were used, as well as mutants for the putative *At*MYB93 targets, seed used was as in Gibbs *et al*. (2014). These represent two different pools of *Atmyb93-1* seed, which will be referred to as *Atmyb93-1* (Hu) and *Atmyb93-1* (Gibbs) respectively. Lines that were mutant for the putative target genes of *At*MYB93 were identified from the T-DNA express database (SIGnAL). As the site of insertion of T-DNA is random, the location of T-DNA insertions within a chosen gene was optimised to increase the chance of selecting mutant lines that were complete KOs for gene function and additionally for ease of selection and experimentation. Where available, mutant lines were purchased from the European *Arabidopsis* Stock Centre (NASC) in Nottingham, UK. Details of mutant *Arabidopsis* lines can be found in Appendix 2.1.

Triticum aestivum (wheat) – Fielder cultivar for work on *AtARABIDILLO* transgenic lines (NIAB). Chinese Spring cultivar for experiments looking at MYB93 homologue expression.

Hordeum vulgare (barley) – Morex cultivar

Oryza sativa (rice) – Nipponbare cultivar

2.1.2 Bacterial strains

Escherichia coli DH5α – a non-pathogenic strain modified to make it suitable for efficient plasmid transformations. *SupE*44, Δ*lac*U169 (Φ80*lacZ*ΔM15), *hsd*R17, *recA*1, *endA*1, *gyrA*96, *thi*-1, *relA*1.

Agrobacterium tumefaciens GV3101 – disarmed strain retaining the Ti virulence plasmid, pMP90 (pTiC5ΔT-DNA), but missing its T-DNA region (Pappas & Winans, 2003); this has gentamycin and rifampicin resistance genes.

2.1.3 Yeast strains

Saccharomyces cerevisiae Y187 (Cat. 630457, Clontech) – commercially available strain originally developed for yeast two-hybrid (Y2H) selection. *MATα, ura3-52, his3-200, ade2-101, trp1-901, leu2-3, 112, gal4Δ, met–, gal80Δ, MEL1, URA3::GAL1UAS-GAL1TATA-lacZ.*

2.2 Vectors

pCR™-Blunt (Cat. K270020, Invitrogen) – a 3.5 kb vector that allows cloning of any blunt DNA fragment. The insert is cloned into *LacZα-ccdB* gene fusion (*ccdB* is lethal to *E. coli*), thus disrupting its expression and ensuring only positive recombinants can grow after transformation; contains both zeomycin and kanamycin resistance genes for selection.

pGBKT7 (Cat. 630443, Clontech) – a 7.3 kb vector that contains a kanamycin resistance marker for selection and a c-myc epitope tag which allows recognition by an antibody.

pGreen0229 (Hellens *et al*., 2000) – a 4.6 kb vector that is used in conjunction with a pSoup helper plasmid to allow replication in *Agrobacterium tumefaciens*. The plasmid contains a kanamycin resistance gene for selection in bacteria, while the T-DNA insert contains a nos-Bar cassette which confers BASTA® resistance for plant screening. The pSOUP helper plasmid contains a tetracycline resistance gene and a *pSA* replicase gene to act in trans with the *pSA* replication origin in pGreen.

pINT1-HIS3NB (Ouwerkerk & Meijer, 2001) – an 11.7 kb yeast vector that integrates into the *Saccharomyces cerevisiae* Y187 genome and contains both ampicillin and gentamycin (G418) resistance genes. It is made up of a combination of two vectors, pINT1 and pHIS3NB, to make a plasmid with the *HIS* reporter gene that integrates into the yeast genome for yeast one-hybrid (Y1H).

pGADT7 (Cat. 640442, Clontech) – an 8 kb yeast vector which expresses a protein of interest fused to a GAL4 activation domain and driven by the constitutively expressed *ADH1* promoter. It contains a *LEU2* gene which acts as a nutritional marker for selection in yeast.

2.3 Growth conditions

Media recipes can be found in Appendix 2.2.

2.3.1 Plant growth conditions

For *Arabidopsis*, seeds were stored at 4°C for a minimum of 48 h for stratification after imbibition with dH_2O. For the growth of seedlings, sterilised (section 2.3.1.1) and cold-treated seeds were pipetted onto plates of growth medium

using BARKY tips and either kept horizontally (for germination assays or transgenic plant selection), or vertically (for root assays and tissue collection), in a temperature and light controlled growth room (22°C, 16 h light: 8 h dark). For work on mature plants, seeds were either sterilised and cold treated before being placed on growth medium in magenta pots in the growth room, or alternatively sown on soil (4:2:1, compost: vermiculate: perlite) and kept in a glasshouse with supplementary lighting.

In the case of wheat, barley and rice; sterilised seeds (section 2.3.1.1) were put onto plates using sterile forceps, which contained sterile filter paper saturated with dH_2O and were kept in a temperature and light controlled growth room (22°C wheat and barley, 25°C rice, 16 h light: 8 h dark). For gene expression experiments using RT-PCR, wheat and barley were grown until they had reached growth stage 11 of the BBCH scale (Meier, 2001) and rice was grown until stage 13 (Meier, 2001). For confocal microscopy, wheat was grown for 5 days in the dark to encourage root growth. For work on mature wheat, seeds were sown in soil (4:2:1, compost: vermiculate: perlite) in a glasshouse (22°C. 16 h light: 8 h dark).

2.3.1.1 Seed sterilisation

A mixture of 20% (w/v) Parozone Original bleach solution and either *Arabidopsis*, wheat, barley or rice seeds were mixed on a rotary mixer for 10 min. The bleach was replaced with sterile, dH_2O and tubes were mixed again to wash the seeds. This was repeated for a total of three washes. Seeds were finally resuspended in dH_2O.

2.3.2 Bacterial growth conditions

E. coli

All inoculations were performed under aseptic conditions. Plates for colony growth and liquid cultures were grown for ~16 h at 37°C. In the case of liquid cultures, samples were additionally grown on a rotary shaker at ~200 rpm.

A. tumefaciens

All inoculations were performed under aseptic conditions. To prepare cultures for floral dip, single colonies were initially obtained on plates containing rifampicin and kanamycin before being cultured for ~16 h at 28°C in ~10 ml of liquid media on a rotary shaker at ~200 rpm. These mini-cultures were each used to inoculate a larger volume of liquid medium (250 ml) under selection, which were then grown for ~48 h at 28°C on a rotary shaker at ~200 rpm.

2.3.3 Yeast growth conditions

S. cerevisiae

All inoculations were carried out under aseptic conditions using appropriate selection when required. Agar plate cultures were grown inverted for ~48-72 h at 30°C. Liquid cultures were grown for ~16 h at 30°C on a rotary shaker at ~200 rpm.

2.4 Bioinformatic analysis

2.4.1 Identification of putative AtMYB93 downstream target genes

The lateral root induction time series data from Voß *et al*. (2015) was used to compare the expression pattern of *At*MYB93 with that of the rest of the *Arabidopsis*

transcriptome. The fold-change expression database of the experiment was initially searched using the following parameters to filter the data: a non-specific search of all genes and all time points with a p-value cut off of 0.05 and a log2 fold-change cut off of 1.0.

All genes were ranked according to the fold-change difference in their expression at the 15 h time point relative to the 12 h time point, as this is the point at which *At*MYB93 is significantly upregulated. The database was then searched for significant differences in fold-change expression in the time points both immediately before and after this time point (12 h vs. 9 h and 18 h vs. 15 h respectively). After the initial search, the $\log_2$ fold-change cut off was reduced to below 1.0 to see if there were any smaller but still significant differences in the fold-change expression of *At*MYB93. As further described in section 3.2.3 of chapter III, the 48 h vs. 45 h time point was also used to compare with the other genes to add an additional element by which to filter the longlist candidate genes, since this is when *At*MYB93 is significantly downregulated.

Analysis of differential expression data from published experiments, which can be viewed at the EMBL-EBI expression atlas database (Petryszak *et al.*, 2016), was used to determine the effects of a range of different variables on the regulation of expression of each gene. Data on the following variables were included where available: ABA, auxin, hypoxia, iron application, iron deprivation, pathogen challenge, phosphate deprivation, nitrogen deprivation, salt, drought stress and other notable factors identified for each gene that promote up or downregulation.

Expression data from the *Arabidopsis* Electronic Fluorescent Pictograph (eFP) browser 2.0 (Winter *et al*., 2007) were used to view the level of gene expression that

has been observed in published experiments and the location of gene expression, either in the roots, seeds or other organs.

Two databases were used to look for existing information on each gene – The Arabidopsis Information Resource (TAIR) (Berardini *et al*., 2015) and NCBI BLAST® (Altschul *et al.*, 1990). Information that was recorded included gene descriptions, where the gene is expressed, conserved domains and putative functions. Synonymous gene names were noted where applicable as these sometimes provided further clues to the potential function of a gene.

2.4.2 Promoter analysis of putative AtMYB93 target genes

The first step in promoter analysis was to determine the size of potential promoter regions upstream of each gene of interest. The maximum size a promoter was estimated to be was 2 kb upstream of the start codon, and the presence of the nearest upstream gene was used to reduce this estimated promoter size if existing. Analysis was performed by entering nucleic acid sequences in FASTA format into the Plant Transcription Factor Database (PlantTFDB) (Jin *et al.*, 2017). An output file was produced for each putative promoter region that contained the following data: sequence motif matched between input sequence and TF, gene family of TF, gene name of TF, strand and nucleotide location of motif, p-value and q-value (a measure of significance that uses the false discovery rate to give the proportion of significant results that have incorrectly been defined as significant, i.e. are false positives). The data were then analysed for the presence of *At*MYB93 binding motifs in the output file. Sequence alignment was performed using MEGA version 7 (Kumar *et al.*, 2016)

using both the MUSCLE (Edgar, 2004) and Clustal Omega (ClustalO) (Sievers *et al.*, 2011) alignment methods with their default parameters.

2.4.3 Creation of phylogenetic trees of MYB93 homologues

Protein sequences of putative MYB93 homologues from plant species were collected, and pBLAST® searches were carried out against the *Arabidopsis At*MYB93 peptide sequence. The newest versions of genome sequencing data available from the Phytozome database version 12.1 (Goodstein *et al.*, 2012) were initially used to gather putative homologues through the website's pBLAST® function, and the top four hits were chosen for each species. The same pBLAST® searches were then carried out using the NCBI database and again the four top hits were recorded. The exception to this was wheat, where the top 6 hits were selected from each database due to the species being hexaploid. The identifier names given to genes differ between Phytozome and NCBI and so the lists were cross-checked for duplicate results and deleted. Additionally, pBLAST® searches were performed against two rice protein sequences purported to be MYB93 homologues (Cao, unpublished). These were LOC_Os02g51799.1 and LOC_Os06g11780.1. Many of the hits from the rice BLAST® search also matched the *Arabidopsis* search and so these duplicates were also removed.

Alignments and trees were produced using SeaView version 4.7 (Gouy *et al.*, 2010) and sequences were aligned using both the MUSCLE and ClustalO alignment methods. Protein sequences for the following *Arabidopsis* genes were included in the alignment: *AtMYB93*, *AtMYB92*, *AtMYB53*, *AtMYB9*, *AtMYB107*, *AtMYB41*, *AtMYB36* and *AtMYB75*. Alignments were optimised by hand to remove incorrectly

annotated sequence and correct obvious misalignments. Trees were created using both BioNJ distance analysis (Gascuel, 1997) and PhyML (Guindon *et al*., 2010) algorithms. Bootstrapping was carried out with 100 replicates and options were kept to their default settings within SeaView.

2.4.4 Gene structure analysis

The following species were used in the gene structure analysis, which assessed the conserved intron and exon structure of a range of putative homologues of relevant species identified in the phylogenetic analysis. These were: *Arabidopsis*, *A. trichopoda*, *B. rapa*, *B. oleracea*, wheat, barley and rice. Genomic DNA (gDNA) and complementary DNA (cDNA) transcript sequences of putative MYB93 homologues were obtained for these species and diagrams depicting gene structure were created using the Gene Structure Display Server (GSDS) version 2.0 (Hu *et al*., 2015).

2.5 Nucleic acid isolation and handling

Buffer recipes can be found in Appendix 2.2.

2.5.1 Quick extraction of genomic DNA for genotyping Arabidopsis

A small amount of *Arabidopsis* tissue (a leaf disc of approximately 0.25 cm^2 or two seven-day-old whole seedlings) was macerated in 40 µl of genotyping extraction buffer and then heated to 95°C for 10 min. Samples were put on ice for 5 min before adding 40 µl of genotyping dilution buffer and centrifuging for 30 sec at 13,000 G. A 1 µl sample was then used as a PCR reaction template.

2.5.2 Extraction of high-quality genomic DNA from Arabidopsis

An ISOLATE II Plant DNA Kit (BIoline) was used to extract DNA from *Arabidopsis* tissue. The kit was used according to the manufacturer's instructions. Up to 100 mg tissue frozen in liquid nitrogen was homogenized with a pestle and mortar and then transferred to a 1.5 ml Eppendorf tube where a lysis buffer was added to help break open cells. This was followed by the addition of an RNase solution and an incubation step. Samples were filtered through an initial column and the DNA containing supernatant was then bound to a second column for washing. DNA samples was then be eluted into nuclease-free water and stored at -20°C before use.

2.5.3 Extraction of total RNA from plants

RNA was extracted primarily using an ISOLATE II RNA Plant Kit (Bioline) according to the manufacturer's instructions. Firstly, up to 100 mg of tissue frozen in liquid nitrogen was homogenized with a pestle and mortar. Tissue was transferred to a 1.5 ml Eppendorf tube and a lysis buffer containing guanidine thiocyanate and ß-mercaptoethanol was used to deactivate RNases. Samples were then filtered through a column and ethanol was added to adjust RNA binding conditions. Once the RNA had been bound to a silica membrane any gDNA contamination was digested with DNase. Washing steps removed other unwanted molecules like salts before the RNA was finally eluted into nuclease-free water. Alternatively, a *Quick*-RNA Miniprep Kit (Zymo) was sometimes used for RNA extraction. This uses a similar extraction protocol but has the option of homogenizing tissue using a bead beater and an additional RNA clean-up step using PVPP. Samples were stored at -80°C.

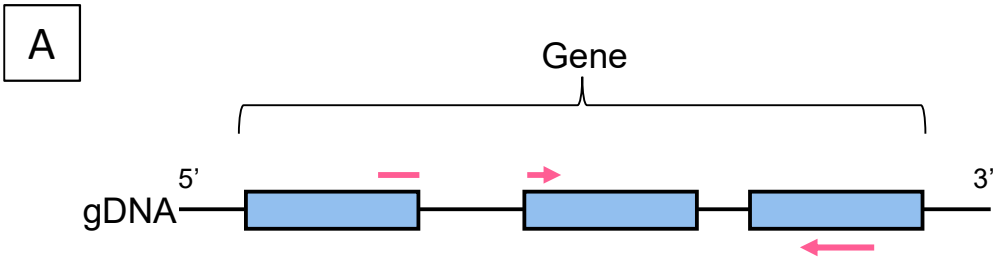
2.5.4 Estimation of nucleic acid concentration

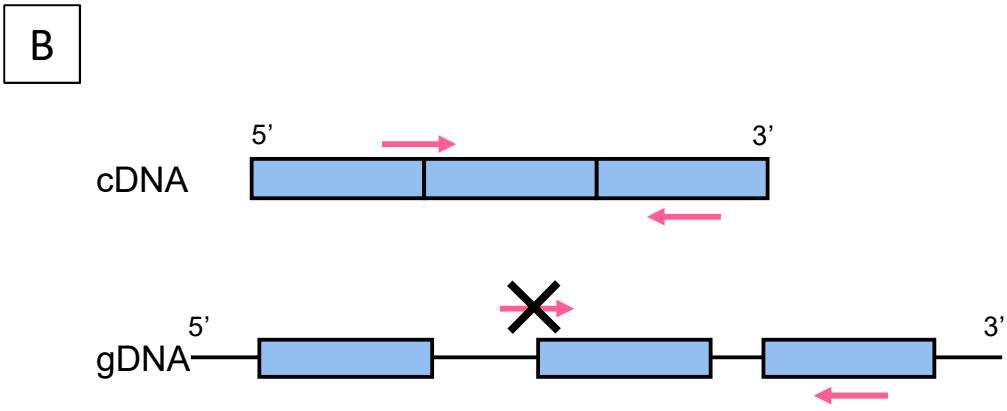
Concentrations of DNA and RNA were quantified using an ND-1000 NanoDrop spectrophotometer (Labtech International) that measures the absorbance of nucleic acid molecules in a 1 µl sample at a wavelength of 260 nm. The ratio of absorbance at 260/280nm and at 260/230nm is used to identify contaminants such as protein and phenol. Preferred values were around 1.8 for DNA and 2.0 for RNA for 260/280 nm, and between 2.0-2.2 for 260/230 nm (REF).

2.5.5 Primer design

Unless otherwise stated, primers were designed to be as close as possible within the following parameters: a predicted melting temperature of between 59-61°C, a GC content of around 40-60%, no more than three nucleotide repeats in a row, three out of the last five base pairs at the 3' end as either a C or G. Amplicon size was dependent on the intended use of the primers. For qRT-PCR, primers were designed for use with SYBR Green and the resultant amplicon length was between 180 and 200 bp. For RT-PCR and PCR, amplicon length was designed to be around 500 bp. For the amplification of cDNA where it was not necessary to obtain the full-length gene, one primer for each pair was designed to span an intron/exon junction in the sequence to avoid amplification of any remaining gDNA (Figure 2.1). Finally, primers for genotyping were designed using the T-DNA Primer Design feature developed by SIGnAL. A list of primers can be found in Appendix 2.2.

Figure 2.1. Primer design to avoid amplification of gDNA contamination when looking at gene expression levels. *(A) Primers designed using CDS sequence so that one spans two exons with >6 bases of complementary sequence on each exon for successful annealing. (B) cDNA is amplified but gDNA is not. The 5' primer is unable to bind gDNA due to the presence of the intron between the first and second exon. Introns denoted by black lines and exons as blue boxes. Primers depicted by pink arrows.*


A
Gene
5’
3’
gDNA


B
5’
3’
cDNA
5’
3’
gDNA

2.6 Manipulation of nucleic acids

2.6.1 Polymerase chain reaction (PCR)

PCR reactions were used to amplify gDNA or plasmid DNA and were set up and run using the manufacturer's instructions, unless otherwise stated. The initial concentration of primers added to each reaction was 10 $pmol.\mu l^{-1}$ and primer annealing temperatures were determined using an online calculator from Applied Biosystems.

In general, *Taq* polymerase (New England Biolabs) was used (e.g. for genotyping and single colony PCR). For the purposes of cloning, either a VELOCITY DNA Polymerase kit (Bioline) or Phusion® High-Fidelity DNA Polymerase (New England BioLabs) were used because of their proof-reading function, which improves accuracy, following manufacturer's instructions. If required, extension times were increased to at least double the recommended time to improve amplification efficiency.

2.6.2 Reverse transcription PCR (RT-PCR)

MyTaq™ One-Step RT-PCR (Bioline) was used to synthesise cDNA and carry out the PCR reaction in the same tube. This was used for all applications except qRT-PCR where a two-step reaction was used (a two-step reaction gives more control over cDNA synthesis conditions and was used when creating cDNA template for quantitative RT-PCR). The kit was used according to the manufacturer's instructions including making the suggested recommendations for the amplification of larger fragments; the extension time was increased from 30 $s.kb^{-1}$ to 1 $min.kb^{-1}$ and the RT time was doubled to 40 min. The annealing temperatures of primers were determined

using an online calculator from Applied Biosystems. For experiments looking at semi-quantitative differences in gene expression, the cycle number was optimised and amended accordingly to allow differences in the amount of amplified product to be observed.

2.6.3 Agarose gel electrophoresis

Electrophoresis was used to separate and visualise DNA and RNA. Gels were made using 0.8-1.0% (w/v) of agarose (Bioline) in 1x TBE, with gel concentration depending on DNA or RNA fragment size. When visualising RNA samples qualitatively or obtaining semi-quantitive data with DNA or cDNA, 2 µl in 100 ml of GelRed® nucleic acid stain 10,000x stock (Cat. 41010, Biotium) was added directly to the melted gel after it had cooled to around 60°C. For visualising DNA or cDNA qualitatively, 3 µl in 500 µl of GelRed® was added to 6x DNA loading dye (Cat. B7025S, New England Biolabs). DNA or RNA samples were mixed in 6x DNA loading dye to a final concentration of 1x loading dye before being loaded onto a horizontal agarose gel in a tank of 1x TBE buffer. Depending on the size of product to be visualised, a 1 kb or 100 bp ladder (both New England Biolabs) was used to estimate DNA size and concentration. Generally, gels were run at between 80-120 V for between 20 min and 1 h. Running gels at lower voltages for longer helped resolve smaller fragments in a mixture, or those that were similar in size. Gels were viewed using a ChemiDoc^{TM} MP Imaging System (BioRad).

2.6.4 Semi-quantification of RT-PCR band intensity

Analysis of RT-PCR-generated DNA band intensity on gels was completed using Fiji (Image J) version 1.50e (Schindelin *et al*., 2015). Images were first converted to 8-bit grey-scale. The “Rectangular Selection” tool was then used to outline the first band, and this was set as the ‘First Lane’ by selecting this option through the gel analysis menu. Subsequent bands were set by moving the rectangle over each of them and selecting the “Next Lane” option. Once all bands to be analysed were set, the “Plot Lanes” option was selected to produce profile plots of each band. The “Straight Line” tool was used to close off the area of each profile peak and the “Wand” tool was used to measure the area of each peak. The value of each peak as a percentage of the total area of all peaks could also be calculated by selecting the “Label Peaks” option.

The next step was to calculate the relative pixel density (i.e. the intensity) of each band. One of the samples, generally the one with the largest peak area, was chosen as a standard by which to compare the other bands to. The following calculation was used to calculate relative density: percentage peak area of sample / percentage peak area of standard.

2.6.5 cDNA synthesis

SuperScript III First-Strand Synthesis SuperMix for qRT-PCR (Invitrogen) was used in accordance to the manufacturer’s instructions to synthesis cDNA. The enzyme mix contains a Moloney Murine Leukemia Virus (M-MLV) reverse transcriptase (RT), which has been genetically modified for improved performance.

For each 20 µl reaction, 1000 ng of total RNA was added, and cDNA synthesis was conducted using a mix of oligo$(dT)_{20}$ and random hexamer primers.

2.6.6 Quantitative RT-PCR (qRT-PCR)

qRT-PCR was performed on an AriaMx machine (Agilent) using Brilliant III ultra-fast SYBR® Green low ROX qPCR master mix (Cat. 600892, Agilent). For each 20 µl reaction mixed according to the manufacturer's instructions, the final template concentration was 1 $ng.\mu l^{-1}$ and this was based on the concentration of RNA added to the cDNA synthesis reaction, which was then diluted 1:4 with nuclease-free water before use. The final primer concentration per reaction was 200-400 nM depending on primer efficiency. Three technical repeats were carried out for each sample on a plate. The cycling parameters were then as follows: a hot start cycle of 10 min at 95°C and then 40 amplification cycles with a denaturation step for 30 s at 95°C and annealing step for 1 min at 60°C. Due to the short length of the amplicon, elongation occurs as the temperature is ramped back up to the next denaturation. This was followed by a melt curve cycle of 30 s at 95°C, followed by 30 s at 65°C and finally ramped back to 95°C for 30 s at a ramp rate of 0.3°C every 2 s to produce a dissociation curve.

2.6.7 qRT-PCR data analysis

The fold-change in relative expression of each gene in *myb93* mutant and over-expressor plants was calculated relative to WT. Cq values were normalised against two housekeeping genes, actin 2 (ACT2) and ubiquitin-conjugating-enzyme 21 (UBC21). Fold-changes were calculated using the $2^{-\Delta\Delta C_T}$ method (Livak &

Schmittgen, 2001), with WT expression set to one. Error bars represent the upper and lower ranges of the fold-change expression levels and were calculated using plus and minus standard deviations (SD).

2.7 Gene cloning

2.7.1 Purification of PCR reactions

PCR reactions were either purified directly or first run on a gel so that specific bands could be excised. An ISOLATE II PCR and Gel Kit (Bioline) was used for purification in either case. The manufacturer's instructions were followed as written apart from during the elution step where, to increase DNA recovery rates, the elution incubation was performed at 70°C for 5 min before centrifuging as normal.

2.7.2 Extraction of plasmid DNA from bacteria

Plasmid DNA was isolated using the ISOLATE Plasmid Mini Kit (Bioline). This was carried out in accordance to the manufacturer's instructions. A combination of SDS and alkaline lysis is used before the sample is neutralised and adjustments for high-salt binding are made. The plasmid is then bound to a silica membrane and washed to remove impurities before elution with a buffer.

2.7.3 Restriction enzyme digest

Restriction enzymes (New England Biolabs) were used to cut double-stranded DNA for cloning and plasmid linearisation and were carried out using the recommended buffers supplied. A minimum of 10 units of each enzyme were used per 50 ml reaction, which were used to digest around 1-2 µg of DNA. Digests were

performed at 37°C for plasmid prep for either 30 min for a diagnostic digest or between 2-3 h for complete digestion. As much larger volumes were required for plasmid transformations into yeast, reaction volumes were increased to 200 µl to cut up to 10 µg of DNA.

2.7.4 Single colony PCR

Individual colonies of bacteria or yeast were picked with an empty pipette tip and mixed into 10 µl of dH_2O. A 1 µl sample of this mixture was then used as the DNA template in PCR reactions as outlined in 2.6.1. For bacteria the diluted colony could be used directly, however for yeast it was necessary to heat the mixture at 95°C for 10 min to help break down the cell membrane.

2.7.5 Plasmid DNA de-phosphorylation and precipitation

pINT1-HIS3NB is a vector with a low copy number and hence can be more difficult to use for cloning steps. Therefore, in order to maximise transformation efficiency, plasmid DNA was de-phosphorylated and then precipitated prior to transformation into yeast. For de-phosphorylation of digested DNA ends, the number of moles of DNA ends was calculated by using the online calculator 'Mass to Moles Converter', which can be found at New England Biolabs. It is inevitable that the amount of DNA will be reduced during the precipitation process and so an excess of plasmid DNA was required in order to obtain the amount that was needed for each transformation.

For a 50 µl reaction in dH_2O, 2.5 pmol of digested DNA was added to 2.5 units of CIP (10,000 units.ml^{-1}) (New England Biolabs) in a 1.5 ml Eppendorf tube. The

mixture was incubated at 37°C for 1 h. DNA precipitation was performed by adding 0.1 volume of 3 M sodium acetate, followed by 2 volumes of 100% (w/v) ethanol. The sample was mixed by briefly vortexing and incubated at -20°C for a minimum of 2 h, or preferably overnight. Samples were centrifuged for 20 min at 14,500 g and the supernatant was discarded. A washing step was carried out using enough 70% (w/v) ethanol to cover the pellet, followed by centrifugation for 10 min at 14,500 g and removal of the ethanol. The wash step was repeated, and pellets were left to dry for ~10 min in order to allow all remaining ethanol to evaporate. The dried pellet of DNA could then be resuspended in 30 µl of dH_2O.

2.7.6 Ligation of DNA into plasmids

For each 20 µl ligation reaction, 1 µl (400 units) of T4 DNA-Ligase (Cat. M0202, New England BioLabs) was used in conjunction with 2 µl of 10x ligase buffer to a final concentration of 1x ligase buffer. A molar ratio of 1:3 of DNA for the vector and insert respectively was most commonly used (although ratios of up to 1:10 were used for sticky end cloning and ratios of up to 1:20 were used for blunt end cloning) and the reactions were incubated at 16°C overnight.

2.7.7 Transformation of plasmid DNA into E. coli

Half of a 20 µl ligation reaction was heat-shock transformed into chemically competent DH5α *E. coli* that had been defrosting on ice for 30 min. Heat-shocking was carried out at 42°C for 45 s before chilling on ice for 5 min. A 1 ml volume of LB broth was then added and a recovery incubation was carried out at 37°C, 200 rpm for up to 2 h. Samples were centrifuged at 3000 g for 10 minutes and most of the

supernatant was removed. The bacterial pellet was resuspended in the remaining supernatant (~100 µl) and spread onto LB agar plates containing antibiotics for selection. Plates were incubated at 37°C overnight.

2.7.8 Transformation of plasmid DNA into S. cerevisiae

In order to be transformed into yeast, cloning steps were first carried out in *E. coli* as outlined above and then stored in bacterial glycerol stocks at -80°C. Resulting plasmids containing the construct of interest were then be extracted and – in the case of pINT-HIS3NB – linearised and phosphorylated before transformation into the Y187 strain of *S. cerevisiae*. Transformation was performed using an EZ-Yeast™ Transformation Kit (MPBio) which utilises a combination of lithium acetate, single-stranded carrier DNA and PEG. Furthermore, specially prepared competent cells are not required, a yeast colony is simply grown overnight in liquid culture and can then be used for transformation. The transformation incubation step was carried out at either 30°C or 37°C for 45 min, with 37°C generally producing a greater number of colonies. Plates were incubated inverted at 30°C for between ~48-72 h on appropriate selection.

2.8 Yeast one-hybrid assays

At*MYB93* was cloned into a pGADT7 vector, while the promoters of the candidate target genes of MYB93 were cloned into the pINT1-HIS3NB vector. Firstly, pINT1-HIS3NB was linearised and de-phosphorylated, before transformation into Y187 yeast for integration into the yeast genome and selection on YPD medium containing G418 antibiotic. MYB93::pGADT7 was then transformed into the yeast

containing the integrated vector and selected for on SD single drop-out medium lacking leucine. Finally, cultures of colonies containing both the integrated bait sequence and prey plasmid were plated onto SD double drop-out medium lacking leucine and histidine with 3mM 3-AT additionally added to reduce autoactivation (as discussed in section 3.3.2 of chapter III).

2.9 Floral dip and transgenic plant selection

2.9.1 Transformation of plasmid DNA into A. tumefaciens

Electrocompetent GV3101 *A. tumefaciens* was defrosted on ice for around 30 min before being transferred to a 0.2 cm electroporation cuvette (BioRad). Around 100 ng of pGreen0229 and pSoup DNA was added and the mixture was incubated on ice for ~20 min. Electroporation was performed using a Micropulser (BioRad) set to EC2 at 2.2 kV for ~5 ms. Then 1 ml of LB broth was immediately added and a recovery incubation was carried out at 28°C, 200 rpm for a minimum of 2 h but preferably up to 4 h. The bacterial pellet was resuspended in the remaining supernatant (~100 µl) and spread onto LB agar plates containing antibiotics for selection and were incubated at 28°C for 2-3 days.

2.9.2 A. tumefaciens-mediated transformation into Arabidopsis using floral dip

The method outlined by Clough and Bent (1998) was used to facilitate the insertion of the plasmid DNA into *Arabidopsis* via inoculation with *A. tumefaciens*. Cultures of *A. tumefaciens* were prepared as described in section 2.3.2 once *Arabidopsis Atmyb93-1* mutant plants had reached the growth stage where buds were present, but few flowers had opened. Primary bolts were removed by cutting in

order to encourage side bolting to increase the number of buds. Bacterial cultures were centrifuged at 3000 g for 15 min and the pellet was resuspended into a 250 ml solution of 5% (w/v) sucrose with 125 µl of Silwet L-77 (Cat. VIS_02, Lehle Seeds). Plants were then inverted and dipped into a beaker containing the dipping solution at a depth to cover all buds and were agitated for at least 10 seconds. The plants were wrapped in clingfilm and kept in the dark for ~24 h, at which point the clingfilm was removed and plants were returned to growing under normal conditions until their seeds could be harvested.

2.9.3 Selection of transgenic plant lines

T1 seeds collected from dipped plants were selected for with BASTA® in one of two ways. The first method was as follows: agar plates of plant growth media supplemented with BASTA® were made. Large numbers of sterilised and cold treated seeds, which were additionally treated with PPM (Sigma), could then be easily sown across the plates by adding them to 15 ml agarose (cooled to ~40°C) and pouring a layer over the growth medium plates. The second method involved sowing seeds directly onto soil watered with a 2% BASTA® solution. A small number of clearly separated WT seeds were also sown as a control to ensure the effectiveness of selection.

2.9.4 Selection of mutant plant lines

Seeds from heterozygous mutant lines obtained from NASC were sown onto plates containing selection growth media to kill WT seedlings. Kanamycin was used for selection with Salk Institute Genomic Analysis (SALK) lines, while BASTA® was

used for Syngenta *Arabidopsis* Insertion Library (SAIL) lines. Remaining seedlings were then planted onto soil after 14 days of selection and genotyped to obtain homozygous mutant seedlings that were transferred to a glasshouse and grown to collect seed.

2.10 Phenotyping assays

2.10.1 Root bioassays

Three plates containing a row of 20 *Arabidopsis* seedlings grown vertically as described in 2.3.1 were used for each biological repeat. Roots were analysed after seven, eight, nine and ten days of growth by counting the number of lateral roots (LRs) by eye. Photographs of each plate of seedlings were used to assess the length of primary root length using Fiji (Image J) (Schindelin *et al*., 2015) and the LR density was calculated for each genotype measured for each day as: primary root length / number of LRs.

2.10.2 Germination assays

Plates of *Arabidopsis* were grown horizontally as described in 2.3.1. For each biological repeat, 90 seedlings of each genotype were grown on medium with and without 150 mM salt. Seedlings were assessed using a low-powered light microscope after one, two, three and four days of growth, by scoring them as being in one of the following stages of germination. These were ruptured testa, radicle emergence and cotyledon emergence. Germination was considered to have occurred at the point of radicle emergence. The germination percentage was calculated as: (number of germinated seeds / total number of seeds) x 100.

2.10.3 Inductively Coupled Plasma Mass Spectrometry (ICP-MS)

2.10.3.1 Preparation and growth of Arabidopsis tissue

For work on plants under normal, non-stressed conditions, plant growth media was made using MS powder (Sigma) as described in Appendix 2.2, but with 0.7% (w/v) agar to encourage root penetration of the media. Around 100 ml of media was added per magenta pot and four evenly spaced, sterilised and cold treated seeds were sown on each pot. Plants were put in a growth room for 25 days (22°C, 16 h light: 8 h dark).

For experiments where plants were put under salt and sulfur stress, media was made using a recipe that allows precise control over its composition (see Appendix 2.2). All compounds were added for control conditions and 100 mM salt was added for salt stress. For sulfur stress, any sulfur-containing compounds were omitted.

2.10.3.2 Harvest and experimental analysis of tissue samples

Plants were harvested after growing for 25 days. They were removed from the agar growth medium and all root tissue was removed. Some plants had bolted in this time and so flower buds were also removed. Remaining shoot tissue was transferred into 50 ml falcon tubes and dried in an oven at 50°C for 48-72 h. Dried shoot tissue was sent to the University of Nottingham, where it was analysed against known tomato standards by Dr. Neil Graham's research group using ICP-MS (Thomas *et al.*, 2016).

2.10.3.3 Analysis of elemental data

Analysis was conducted following the method described in Ghaderiardakani *et al*., (2019). Results from ICP-MS were based on ~200 mg of dried Arabidopsis shoot tissue and data were supplied as $mg.L^{-1}$ for a range of 28 elements: Ag, Al, As, B, Ba, Ca, Cd, Cr, Co, Cs, Cu, Fe, K, Mg, Mn, Mo, Na, Ni, P, Pb, Rb, S, Se, Sr, Ti, U, V and Zn.

2.11 Proteins

All recipes for buffers, gels and staining solutions can be found in Appendix 2.2.

2.11.1 MG-132 treatment

Seven-day-old *Arabidopsis* and ten-day-old wheat samples were treated with 50 mM MG-132 100% ethanol solution by pipetting (Cat.ab147047, abcam), a proteasome inhibitor, prior to protein extraction. Care was taken to ensure the MG-132 solution was covering the entire root and samples were treated for 2h alongside an ethanol control.

2.11.2 Protein extraction from plants

A sample of tissue frozen in liquid nitrogen was homogenized with a pestle and mortar. A volume of ~0.5-1 ml was transferred to a 2 ml Eppendorf tube and left to defrost before 1 ml of protein extraction buffer was added. Samples were centrifuged at 13,000 g for 45 min and the resultant supernatant was collected for use.

2.11.3 Estimation of protein concentration

A Bradford (AppliChem) assay was used to estimate the concentration of extracted protein samples. Samples were diluted in dH_2O (1:5) and 4 µl was added to 250 µl of Bradford reagent (Cat. A6932, AppliChem). A spectrophotometer was used to measure absorbance at 595 nm and measurements were made against a BSA standard curve.

2.11.4 SDS-PAGE

Gels were cast and run using a Mini-PROTEAN® Tetra Handcast System (BioRad) and each gel was made with 10% (w/v) resolving gel and 4% (w/v) stacking gel. Protein samples were mixed with 5x protein loading buffer to a final concentration of 1x loading buffer before loading the gel. Gels were run vertically in a tank containing 1x ELFO buffer at 100 V for ~2 h.

2.12 Western Blotting

2.12.1 Protein transfer

Once the protein samples had been sufficiently separated, the stacking gel was detached and discarded while the resolving gel was carefully placed onto a piece of PVDF transfer membrane (Hybond P 0.45, Amersham). The gel and transfer membrane were placed between two sheets of electroblotting paper (BioRad) and two electroblotting sponges (BioRad) and were then submerged in protein transfer buffer. A plastic case was used to hold each component in place inside an electroblotting tank (BioRad) containing protein transfer buffer, which was run at 100 V for 1 h at room temperature, or alternatively at 20 V at 4°C overnight

2.12.2 Blocking and antibody probing

The transfer membrane was placed in blocking solution and put on a rocker either at room temperature for 1 h, or at 4°C overnight. The membrane was washed 3 x 5 min in TBST before it was incubated with primary antibody – 1:1000 dilution rabbit α-GFP polyclonal antibody (Cat. PABG1, Chromotek) in 10 ml of blocking solution. Incubation was carried out on a rocker for 3 h at room temperature. The membrane was washed 3x 5 min in TBST once more and was then incubated with secondary antibody – 1:2000 dilution goat α-rabbit immunoglobin (Cat. ab6721 abcam) in 10 ml of blocking solution. Incubation was carried out on a rocker for 2 h at room temperature. Finally, the membrane was washed again for 3x 5 min in TBST.

2.12.3 Protein detection and analysis

Before protein visualisation, the transfer membrane was briefly left to dry. Enhanced chemiluminescence (ECL) was used to detect presence of the secondary antibody using ECL Western Blotting Detection Reagents (Amersham). Detection reagents one (peroxide solution) and two (luminol enhancer solution) were mixed at a 1:1 ratio and 3 ml of this pipetted over the dry membrane for 3 min. To analyse the blot in the dark, the transfer membrane was pressed onto photographic film (Hyperfilm^{TM} ECL^{TM}, Amersham). The photographic film was then passed through a X-o-Graph machine (Fuji) where the blot was developed, fixed and then washed.

2.12.4 Coomassie staining

Nitrocellulose membranes were incubated in Coomassie stain after protein transfer and left on a rocker at room temperature for 2-4 h. Following this,

membranes were submerged in destaining solution and left on a rocker, also at room temperature overnight.

2.13 Statistical analysis

Statistical analyses were carried out either using GraphPad Prism version 8.0.2 or Microsoft Excel. Data were tested for normality using a Kolmogorov-Smith test. Then for instances of normally distributed data with two groups (e.g. qRT-PCR), a t-test was performed. For three or more groups of normally distributed data (e.g. qRT-PCR), an ANOVA was used with a Dunnet's post-hoc test to identify the significance between each sample relative to a control sample (i.e. the WT). Where data were non-parametric and made up of three or more groups (e.g. germination assays), a Kruskal Wallis test was used. Following Kruskal Wallis, a Bonferonni correction was applied and a post-hoc Dunn test was performed to identify the significance between pairs of groups. For elemental analysis, differences between genetic lines or environmental conditions were analysed using a Welch's t-test. Exact p-values were recorded to at least three decimal places.

2.14 Graphical software used for figure preparation

Diagrams were created using a combination of Microsoft PowerPoint and BioRender, a website offering a library of graphical icons that can be used to make figures (BioRender, 2019). The version of BioRender used was a free trial of the academic license. Tables and heat maps were produced with Microsoft Excel. Graphs were made using either Microsoft Excel or GraphPad Prism version 8.0.2, and Venn diagrams were made using R version 3.2.2 (R Core Team, 2015).

Phylogenetic trees were designed using iTOL version 4.4.2, an online tool that can be used to modify and annotate trees (Letunic & Bork, 2019). BOXSHADE version 3.21 (Hofmann & Baron, 1996) was used to shade sequence alignments and sequence logos were created with WebLogo (Crooks *et al.*, 2004).

Chapter III

IDENTIFYING UPSTREAM REGULATORS AND DOWNSTREAM TARGETS OF *At*MYB93

3.1 Introduction

Plant root development is well studied, yet much remains to be discovered about how we can manipulate root architecture to create more adaptable crop varieties. As previously discussed in chapter I, *At*MYB93 is a TF known to inhibit the development of new LRs (Gibbs *et al.*, 2014; Gibbs & Coates, 2014), though currently very little is known regarding the molecular network through which *At*MYB93 performs its role. Gibbs *et al.* (2014) previously identified a direct interaction between *At*MYB93 and *At*ARABIDILLO, a protein that promotes LR initiation. The mechanism of this interaction is unknown; however, it has been hypothesised that *At*MYB93 could be repressing a LR development gene or alternatively activating a LR development inhibition gene (Gibbs & Coates, 2014). Moreover, *AtMYB93* is upregulated by a variety of endogenous and exogenous factors including the hormones auxin and ABA and stresses like salt (Gibbs *et al.*, 2014; Winter *et al.*, 2007). Identifying the target genes of *At*MYB93 could help us understand more about its function.

A study of the transcriptome by microarray in *Arabidopsis* hypocotyl tissue provided evidence of *AtMYB93* upregulation in response to auxin (Chen *et al*., 2012), whereas *AtMYB93* upregulation was observed in response to ABA application in an RNA-seq experiment by Zhan *et al.* (2015). In addition, several other genome-wide studies have also previously highlighted the effects of auxin and ABA on *AtMYB93* expression (Kranz *et al.*, 1998; Vanneste *et al.*, 2005; Yanhui *et al.*, 2006) and data from these experiments are further reinforced by a qRT-PCR analysis by Gibbs *et al.* (2014) in which *AtMYB93* was found to be upregulated by auxin and ABA in a dose-dependent manner.

Nitrogen and sulfur are nutrients that are often linked and their deficiency results in *AtMYB93* upregulation (Bielecka *et al.*, 2015; Iyer-Pascuzzi *et al*., 2011; Scheible *et al.*, 2004). Data mined from the literature for phosphate starvation is slightly less convincing but suggests that *AtMYB93* may be upregulated in response to a lack of available phosphate. For example, a transcriptomic microarray analysis by Bustos *et al.* (2010) revealed a two-fold increase in expression of *AtMYB93* and a similar increase in expression was seen by Lin *et al.* (2011) in roots measured 24 hours after transfer to phosphate-starved conditions. Yet this is in contrast with an earlier transcriptome analysis that failed to show a significant difference in *AtMYB93* expression in response to phosphate stress (Morcuende *et al.*, 2007).

Finally, *AtMYB93* is also upregulated following the application of salt stress (Dinneny *et al.*, 2008). Unpublished data by the Coates lab has suggested that salt stress further accentuates the reduced LR phenotype of the *AtMYB93* over-expressor and preliminary data by (Wilkinson, 2018) identified that *Atmyb93-1* mutants accumulate more salt in their shoots than WT plants when grown in high-salt media. As for drought, if salt stress is severe enough plants can experience physiological drought and so this implies that insufficient water availability could produce a similar effect. Mustroph *et al*. (2009) conducted a translatome-wide study of the effects of gene expression in *Arabidopsis* in a hypoxic environment. This experiment found that *AtMYB93* is upregulated under such conditions.

Expression of *At*MYB93 is limited to just a few cells in the root endodermis (Gibbs *et al.*, 2014), a layer that lies directly above the pericycle (Dolan *et al.*, 1993; Malamy & Benfey, 1997a), from which new LRs begin their development. This expression profile implies that *AtMYB93* might be repressed in the rest of the

endodermis, suggesting an endodermal TF may be responsible for this restricted expression. *AtSCR* is a TF expressed in the endodermis that directly binds to the *pAtMYB93* promoter (Iyer-Pascuzzi *et al.*, 2011; Sparks *et al.*, 2016). In addition, *AtSCR* expression is downregulated shortly before AtMYB93 upregulation, as shown in an analysis of gene expression after the induction of a new LR (J.C. Coates, unpublished observation; Voß *et al.*, 2015). One hypothesis I will test in this chapter is that the upstream regulator of *At*MYB93 that modulates its specific tissue localisation is *At*SCR.

The endodermal cell wall contains a layer of suberin that acts as a protective coating against a range of stresses (Beisson *et al.*, 2012), though the trade-off of this barrier is that it must be remodelled in order to make way for the developing LR primordium (Péret *et al.*, 2009; Robbins *et al.*, 2014; Vermeer *at al.*, 2014). Several closely related MYBs are thought to have roles in regulating suberin, including *AtMYB107* and *AtMYB9* (Gou *et al.*, 2017; Lashbrooke *et al.*, 2016), *AtMYB41* (Kosma *et al.*, 2014), *AtMYB53*, *AtMYB92* and *AtMYB93* (Hu, 2018) and also *MdMYB93* (Legay *et al.*, 2016). It is therefore hypothesised that *At*MYB93 has a role in modifying the suberin layer in some way during the initial development of new LRs.

The primary aim of this project was to start uncovering the molecular network of MYB93 in *Arabidopsis* and presented in this chapter is an analysis of potential downstream target genes of *At*MYB93 as well as the identification of a possible upstream regulator of the TF.

3.2 Bioinformatic analysis to select putative downstream targets of AtMYB93

3.2.1 Analysis of microarray expression data after lateral root induction

The target genes of *At*MYB93 are so far unknown; therefore, the first section of work set out to select a number of genes likely to be promising targets on which to carry out molecular analyses. Figure 3.1 contains a flow diagram depicting this process. A study by Voß *et al.* (2015), looked at the expression of the *Arabidopsis* transcriptome during LR initiation. This was achieved through a time course microarray analysis of developing LRs induced in the hours after mechanical bending of *Arabidopsis* seedling primary roots to result in the formation of a new LR primordium (Malamy and Benfey, 1997b; Péret *et al.*, 2012).

It was predicted that any genes that are targets of *At*MYB93 are likely to have peak expression of that gene during the same time frame as the significant fold-change in upregulation of *At*MYB93 (15 h vs. 12 h), or in the time frame immediately after this (18 h vs. 15 h). Some genes that initially appeared to be targets based on their upregulation during these time periods could be eliminated after looking at the time period prior to peak *At*MYB93 expression (12 h vs. 9 h). This is highlighted in the case of *At*FAR4, *At*FAR5 and *At*GPAT5 (Figure 3.2), which represent genes potentially involved in suberin biosynthesis (Beisson *et al.*, 2007; Domergue *et al*., 2010) but unlikely targets of *At*MYB93 as their peak expression occurs during this earlier time period (Voß *et al*., 2015) and their expression is more widespread, as shown through their *promoter::GUS* reporter gene pattern (Beisson *et al.*, 2007; Domergue *et al*., 2010). Interestingly, however, a significant downregulation in expression was observed in the later timepoint in which *AtMYB93* is also significantly downregulated. This initial analysis provided time series data of fold-change gene

Figure 3.1. Flowchart depicting selection of candidate gene targets of AtMYB93 through bioinformatic analysis. *An initial longlist of 337 genes was ultimately reduced to 16 genes considered likely downstream targets of AtMYB93. White, green, grey, orange, magenta and blue boxes match the colour coding of genes in Appendix 3.1.*

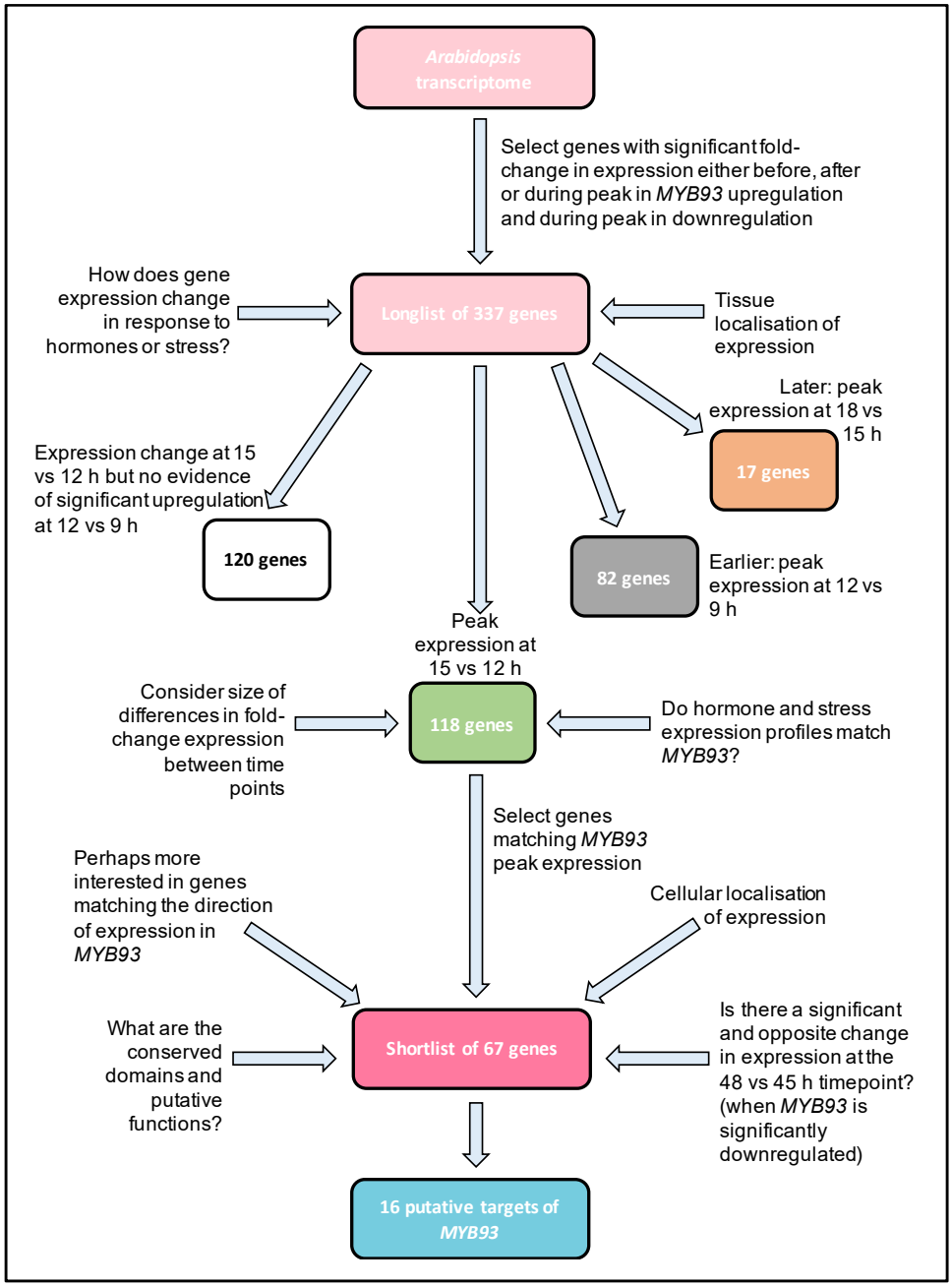
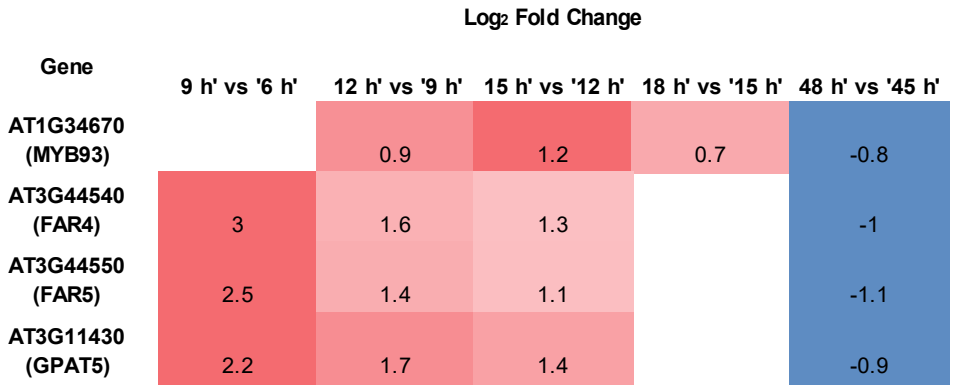

Arabidopsis transcriptome
Select genes with significant fold-change in expression either before, after or during peak in MYB93 upregulation and during peak in downregulation
How does gene expression change in response to hormones or stress?
Longlist of 337 genes
Tissue localisation of expression
Later: peak expression at 18 vs 15 h
17 genes
Expression change at 15 vs 12 h but no evidence of significant upregulation at 12 vs 9 h
120 genes
82 genes
Earlier: peak expression at 12 vs 9 h
Peak expression at 15 vs 12 h
Consider size of differences in fold-change expression between time points
118 genes
Do hormone and stress expression profiles match MYB93?
Select genes matching MYB93 peak expression
Perhaps more interested in genes matching the direction of expression in MYB93
Cellular localisation of expression
What are the conserved domains and putative functions?
Shortlist of 67 genes
Is there a significant and opposite change in expression at the 48 vs 45 h timepoint? (when MYB93 is significantly downregulated)
16 putative targets of MYB93

Figure 3.2. Heatmap of fold-change in expression of well-known suberin lateral root biosynthesis genes not thought to be AtMYB93 targets. *Time series data from Arabidopsis transcriptional expression profile has been used for fold-change data. Positive values = fold-change in upregulation relative to previous time point. Negative values = fold-change in downregulation relative to previous time point. Only four time points are shown as these are where differences in fold-change expression of AtMYB93 were significant ($p < 0.05$).*


Log2 Fold Change
Gene
9 h' vs '6 h'
12 h' vs '9 h'
15 h' vs '12 h'
18 h' vs '15 h'
48 h' vs '45 h'
AT1G34670 (MYB93)
0.9
1.2
0.7
-0.8
AT3G44540 (FAR4)
3
1.6
1.3
-1
AT3G44550 (FAR5)
2.5
1.4
1.1
-1.1
AT3G11430 (GPAT5)
2.2
1.7
1.4
-0.9

expression matching these criteria for 337 genes out of the 19,380 genes in the original dataset. The next step was to narrow this longlist down further.

3.2.2 Genes matching the expression profile of AtMYB93 in response to hormones and stresses

The EMBL-EBI expression atlas database (Petryszak *et al.*, 2016) contains information on the relative expression of genes in response to an array of factors that have been tested and published. The relative gene expression of *AtMYB93* and the 337 genes from the microarray dataset were recorded for several relevant conditions. These were: ABA, auxin, hypoxia, iron application, pathogen challenge, phosphate stress, nitrogen stress, sulfur stress, drought and salt stress.

Direct experimental data were not available for the response of *AtMYB93* to some factors, namely iron application, pathogen challenge and, as mentioned earlier in this chapter, drought. Nevertheless, the direction of change in expression of *AtMYB93* could be predicted in response to these conditions based on its known LR function. For example, a study by Reyt *et al.* (2015) found that an excess application of iron results in the inhibition of LR initiation in *Arabidopsis*. This ties in with the reduced LR phenotype seen in the *AtMYB93* over-expressor (Gibbs *et al*., 2014). Additionally, a transcriptome microarray of protoplasted *Arabidopsis* root cells identified *AtMYB93* downregulation in these cells under conditions of iron deprivation, inferring that *AtMYB93* could be upregulated by iron (Dinneny *et al*., 2008). Expressions looking at gene expression changes in response to pathogens tend to focus on shoot tissue, however salicylic acid has been shown to increase *AtMYB93* expression (Singh *et al*., 2015). Salicylic acid is a hormone that has roles in

orchestrating immune responses to pathogen challenge and wounding (Bari & Jones, 2009), thus implying a possible role of *AtMYB93* in response to infection. More likely, however, it signifies a response to the breaking of the different layers of cells in the primary root during LR emergence.

The assumption was made that changes in *At*MYB93 expression in response to changing conditions would also be observed in its targets and thus the genes most closely following this pattern were selected. Further to this, EMBL-EBI also displays expression data pertaining to confirmed gene homologues in other species such as crops. Therefore, the database was mined for experiments performed on crop roots for each gene (Petryszak *et al.*, 2016). Unfortunately, these data were relatively sparse and no existing crop data were available for any of the final selection of candidate targets. Finally, the data were assessed for any other noteworthy factors promoting significant changes in expression – to see if any common themes were revealed, but nothing of merit was uncovered and so these data have not been included.

Data from the *Arabidopsis* eFP browser (Winter *et al.*, 2007) provided evidence of the levels and locations of gene expression in different parts of the plant. These data have been collated from different experiments and were used to supplement the other bioinformatic analysis data. As *AtMYB93* expression has only been identified in the roots it was expected that genes most likely to be direct targets would also likely be restricted to the roots in their expression but some putative *At*MYB93 target genes are also expressed in other tissues (Winter *et al.*, 2007). Most notably, many of them are expressed in the seed coat; hence it is possible that not all target genes are root specific as there could be other TFs, including closely related

MYBs, switching on their expression in other parts of the plant. For example, the *AtMYB93* relatives *AtMYB41*, *AtMYB107* and *AtMYB9* are all expressed in the seed and *AtMYB93*'s closest relatives (*AtMYB53* and *AtMYB92*) are expressed more widely than just in the roots (Gibbs *et al.*, 2014, Winter *et al.*, 2007), therefore non root-specific genes were not excluded at this stage.

3.2.3 Potential AtMYB93 target genes identified

Together, analysis of patterns in the gene expression profiles and expression data from the eFP brower were used in conjunction with the microarray data to filter the 388 genes to 67 (Appendix 3.1). This list of 67 genes was then narrowed down by again assessing all of the elements of the bioinformatic analysis together, along with looking at conserved domains and putative gene function by searching on databases such as NCBI and TAIR. Genes with conserved elements or putative functions deemed relevant to LR development, including those involved in cell wall remodelling, lipid biosynthesis, and responses to factors known to be associated with *AtMYB93*, were considered more compelling candidate genes, particularly when this corroborated with data from the rest of the analysis.

After this the bioinformatic analysis has revealed 16 genes that look promising as putative targets of *At*MYB93 (Figure 3.3). *At*MYB93 was determined to be one of the least strongly upregulated genes with a significant fold-change in expression of 1.2 during the 15 h vs 12 h time point ($p < 0.05$). There were also three further time periods where *At*MYB93 was up- or downregulated, namely the time points adjacent to the 15 h time point (12 h vs 9 h and 18 h vs 12 h with a fold-change difference in expression of 0.9 and 0.7 respectively (both $p < 0.05$)) and the 48 h vs 45 h time

Figure 3.3 Heatmap of fold-change in expression of the 16 genes selected as possible transcriptional targets of AtMYB93 (also shown). *Time series data from Arabidopsis transcriptional expression profile has been used for fold-change data. Positive values = fold-change in upregulation relative to previous time point. Negative values = fold-change in downregulation relative to previous time point. Genes are ranked in order of expression at the 15 h vs 12 h time point from most to least strongly upregulated. Only four time points are shown as these are where differences in fold-change expression of AtMYB93 were significant ($p < 0.05$).*

	Log$_2$ Fold Change			
Gene	12 h' vs '9 h'	15 h' vs '12 h'	18 h' vs '15 h'	48 h' vs '45 h'
AT5G49350	1.3	3.7	1.3	-1.1
AT5G44550 (CASPL1B1)	2.7	3.6	1.5	-1.1
AT2G18370 (LTP8)	1.5	2.8	1.7	-0.1
AT2G43670	1.5	2.5	1	-0.8
AT3G04370 (PDLP4)	0.9	2.4	1.2	-1.2
AT3G47950 (AHA4)	0.5	2.2	1.3	-0.7
AT2G38110 (GPAT6)	1.7	2.1	1.1	-0.7
AT2G37360 (ABCG2)	1.3	1.7	0.4	-0.6
AT1G55940 (CYP708A1)	0.6	1.6	0.9	-0.5
AT5G55590 (QRT1)	0.5	1.5	1.2	-1.1
AT1G73410 (MYB54)	1.1	1.4	0.9	-0.6
AT5G47550 (CYS5)	0.5	1.4	1.2	-0.5
AT2G19150 (PME10)	0.9	1.3	0.3	-0.7
AT3G27270	0.2	1.3	0.5	-0.9
AT4G17480	0.4	1.3	1.1	-0.4
AT1G34670 (MYB93)	0.9	1.2	0.7	-0.8
AT5G16240 (S-ACP-DES1)	0.9	1	0.7	-0.6

point, where *At*MYB93 expression is significantly downregulated (a fold-change of 0.8) ($p < 0.05$).

Table 3.1 shows the expression profiles of *At*MYB93 and the final list of genes. Where data were available from the EMBL-EBI expression atlas database, the 16 chosen candidates mostly match the expression patterns of *At*MYB93 for the range of hormones and stresses that were assessed. Generally, they are upregulated in response to ABA, auxin, the application of iron, under phosphate deficiency, and finally, in drought conditions. On the other hand, hypoxic conditions and the presence of pathogens typically result in downregulation of the genes. Data for iron stress and hypoxia are scarcer and so the result is less robust than for the other treatments and additionally, it is *At*MYB93 translation that has been shown to be regulated by hypoxia (Mustroph *et al.*, 2009), rather than transcription.

In Table 3.2, the absolute expression of each candidate gene in root and hypocotyl tissue is shown, along with a heatmap representative of their expression in root, hypocotyl, seed and expression in any other tissue types to give an idea of tissues where gene expression is particularly enriched. The absolute values for *AtMYB93* expression were 80.77 in young roots and the hypocotyl, while this has reduced slightly to 64.43 in more mature root tissue. Three of the putative targets of *At*MYB93 are much more strongly expressed in both hypocotyl and root tissue than any of the other genes. They are: *AtCASPL1B1*, *AtLTP8* and *AtCYS5*. The fact that not all genes appear to show expression in the root or hypocotyl for relative expression data is likely due to particularly strong expression in other tissue types, as absolute values make clear that these genes are all in fact expressed in the root. The relative expression data implies that none of the genes are root specific, with all but

Table 3.1. Hormone and stress expression profiles for the 16 genes selected as possible transcriptional targets of AtMYB93 (also shown). *ABA, auxin and iron+, salt = application. Iron-, phosphate, nitrate, sulfur = deprivation. Red = gene is upregulated. Blue = gene is downregulated. White = no/inconclusive data.*

Gene	ABA	Auxin	Hypoxia	Iron +	Iron -	Pathogens	Phosphate	Nitrate	Salt	Drought	Sulfur
AT1G34670 (MYB93)											
AT5G49350											
AT5G44550 (CASPL1B1)											
AT2G18370 (LTP8)											
AT2G43670											
AT3G04370 (PDLP4)											
AT3G47950 (AHA4)											
AT2G38110 (GPAT6)											
AT2G37360 (ABCG2)											
AT1G55940 (CYP708A1)											
AT5G55590 (QRT1)											
AT1G73410 (MYB54)											
AT5G47550 (CYS5)											
AT2G19150 (PME10)											
AT3G27270											
AT4G17480											
AT5G16240 (S-ACP-DES1)											

Table 3.2. Expression of the 16 putative AtMYB93 (also shown) target genes in different tissues. *Values derived from the Arabidopsis EFP database. Absolute expression of expression in root and hypocotyl tissue. Relative tissue values are given as the tissue with the maximum overall expression relative to all other tissue types measured.*

	Absolute expression		Expression relative to highest tissue expression			
Gene	Root	Hypocotyl	Root	Hypocotyl	Seed	Other
AT1G34670 (MYB93)	64.43	80.77	64.43	80.76	14.13	12.01
AT5G49350	65.82	68.07	65.81	68.06	188.10	3.46
AT5G44550 (CASPL1B1)	270.90	225.82	270.90	225.81	744.35	23.41
AT2G18370 (LTP8)	259.32	244.58	279.15	244.58	1850.06	15.93
AT2G43670	56.80	45.58	56.79	45.58	223.33	10.86
AT3G04370 (PDLP4)	36.85	55.08	36.85	55.08	482.73	8.88
AT3G47950 (AHA4)	23.13	20.93	24.65	20.93	321.37	140.63
AT2G38110 (GPAT6)	47.42	50.53	47.41	50.53	83.73	2011.11
AT2G37360 (ABCG2)	34.23	31.97	34.23	31.96	237.13	19.46
AT1G55940 (CYP708A1)	46.83	33.58	46.83	33.58	67.26	27.41
AT5G55590 (QRT1)	10.13	17.02	10.13	17.01	16.06	42.88
AT1G73410 (MYB54)	19.82	19.25	19.81	19.25	47.93	29.53
AT5G47550 (CYS5)	116.72	329.47	116.71	329.46	4367.20	2881.20
AT2G19150 (PME10)	34.48	24.98	34.48	24.98	28.81	10.91
AT3G27270	59.00	47.53	59.00	47.53	52.70	64.73
AT4G17480	5.45	25.65	5.45	25.64	386.81	181.51
AT5G16240 (S-ACP-DES1)	17.18	54.65	17.18	54.65	558.60	491.91

AtGPAT6 also being expressed in the seed coat. This is perhaps expected as closely related MYBs such as *AtMYB92* and *AtMYB53* are expressed in this tissue and limited expression of *AtMYB93* begins in germinating seedlings (Gibbs *et al*., 2014). Ten of the genes are relatively strongly expressed in other tissue types (such as in the flower during its different stages of development), and this potentially implicates these genes in having wider roles in the plant. It is important to note that these data are often taken from large scale, whole genome experiments and therefore should be considered in conjunction with more robust evidence, rather than viewed alone.

To analyse gene function an NCBI BLAST® search (Altschul *et al.*, 1990) of these genes indicated that three have been linked to carbohydrate cell wall remodeling, seven are thought to be involved in lipid biosynthesis or metabolism and that another three genes are related to endodermal functions. Importantly, the characterised or putative roles of these genes are all factors associated with LR emergence (Table 3.3).

3.2.4 Candidate gene promoter analysis

An analysis of the promoter regions of the 16 candidate target genes of *At*MYB93 was carried out. This was achieved by searching the Plant Transcription Factor Database (PlantTFDB) (Jin *et al.*, 2017) using the estimated promoter sequences for the candidate genes. This is an online resource for plant TFs and their regulatory interactions with target genes. Predicted TF binding motifs are detected in input sequences by making use of data from a range of sources. These include regulatory elements – non-coding DNA that can bind TFs to regulate gene expression – which have been identified through high-throughput sequencing from

Table 3.3. Functional categorisation of putative AtMYB93 targets. *14 of the 16 putative AtMYB93 are grouped into three functional roles. Purple = an endodermal membrane protein, blue = fatty acid or suberin biosynthesis or metabolism gene, green = a cell wall remodeling gene. References are given for published characterisations of function.*

Gene	Description	Implicated Role
AT5G49350	Glycine-rich protein	Endodermal membrane protein
AT5G44550 (CASPL1B1)	Casparian strip membrane protein family	Endodermal membrane protein
AT2G18370 (LTP8)	Non-specific lipid-transfer protein 8	Putative fatty acid/suberin biosynthesis
AT2G43670	Carbohydrate-binding X8 domain superfamily protein	
AT3G04370 (PDLP4)	Plasmodesmata-located protein 4	Endodermal membrane protein
AT3G47950 (AHA4)	Plasma membrane ATPase 4	Endodermal membrane protein
AT2G38110 (GPAT6)	Glycerol-3-phosphate acyltransferase 6	Putative fatty acid/suberin biosynthesis
AT2G37360 (ABCG2)	ABC transporter G family member 2	Putative fatty acid/suberin biosynthesis
AT1G55940 (CYP708A1)	Cytochrom P450 family 708 polypeptide 1 sufamily A	Putative fatty acid/suberin biosynthesis
AT5G55590 (QRT1)	Pectin methylesterase	Cell wall remodelling
AT1G73410 (MYB54)	MYB domain protein 54	Cell wall remodelling
AT5G47550 (CYS5)	Cystatin proteinase inhibitor 5	
AT2G19150 (PME10)	Pectin lyase/pectin methylesterase 10	Cell wall remodelling
AT3G27270	TLC lipid sensing domain protein	Putative fatty acid/suberin biosynthesis
AT4G17480	Alpha-beta hydrolase superfamily protein	Putative fatty acid/suberin biosynthesis
AT5G16240 (S-ACP-DES1)	Stearoyl-acyl-carrier-protein 9-desaturase 1	Putative fatty acid/suberin biosynthesis

ChIP-seq and DNase-seq experiments (Jin *et al.*, 2017). The database identified predicted *At*MYB93 binding domains in 9 of the 16 genes (Table 3.4), with two motifs found in *At5G49350*, a mostly uncharacterised gene. Additional to these, *At*MYB93 itself and three genes initially not expected to be downstream targets were analysed as negative controls: *AtSCR*, *AtARABIDILLO-1* and *AtARABIDILLO-2*. A *At*MYB93 binding domain motif was identified in *At*MYB93, while no putative domains were indicated to be present in either *AtARABIDILLO-1* or *AtARABIDILLO-2*. In *AtSCR*, two motifs with a high degree of sequence overlap were found just upstream of the 5' untranslated region (UTR). As I later demonstrate that *At*SCR is an upstream regulator of *AtMYB93* (section 3.5), this suggests the possibility of a feedback loop between *At*MYB93 and *At*SCR.

The sequences of the putative *At*MYB93 binding motifs were aligned using MEGA version 7 (Figure 3.4) and this revealed one fully conserved thymine nucleotide across each of the motifs as well as a further thymine and a guanine that is fully conserved except for in the case of *AtGPAT6*. Several other nucleotides are still relatively highly conserved; five in the ClustalO alignment (Figure 3.4.A) and three in the MUSCLE alignment (Figure 3.4.B).

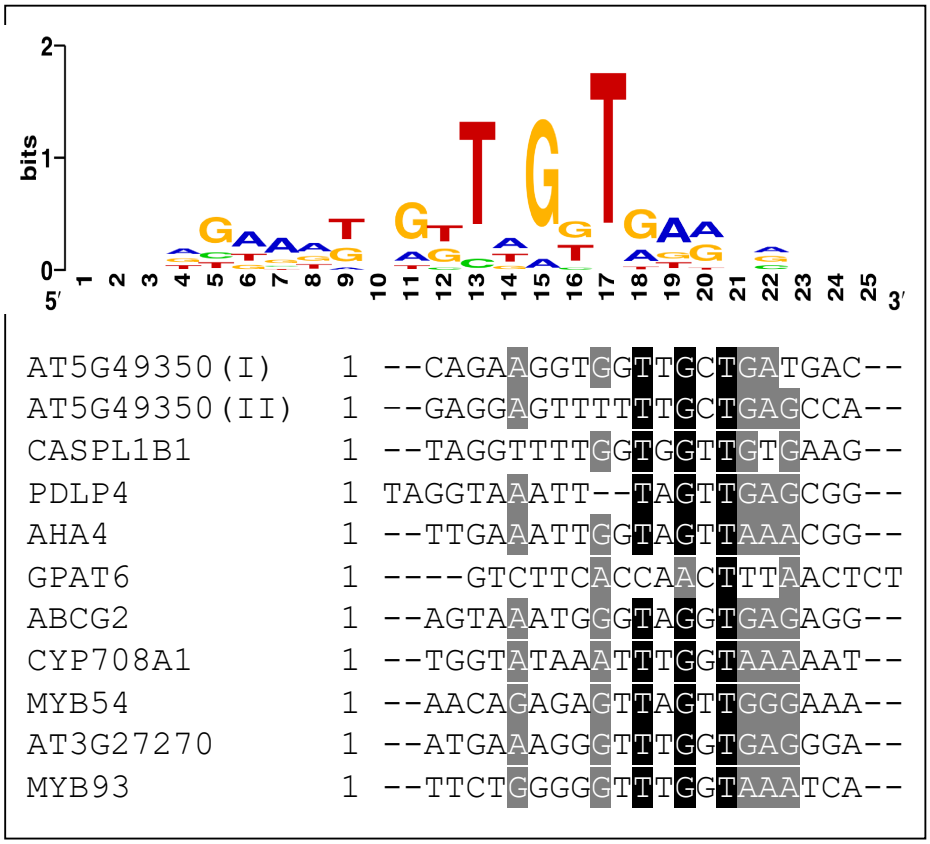
These findings, when taken together with the complementary bioinformatics analyses, provide additional confirmation that most of the candidate target genes may be direct transcriptional targets of the *At*MYB93 TF.

Table 3.4. Predicted AtMYB93 DNA binding domain motifs identified from a promoter analysis of the 16 putative AtMYB93 targets. *13 putative AtMYB93 binding motifs identified in 11 estimated promoter regions using the PlantTFDB. AtMYB93, AtARABIDILLO-1, AtARABIDILLO-2 and AtSCR analysed as controls. Upstream position of start codon of motifs, P-values, q-values and the direction of both the promoters and motifs are given.*

Gene	Motif Sequence	Nucleotide Position Upstream of Start Codon	P Value	Q Value	Promoter Direction	Motif Direction
AT5G49350 (I)	CAGAAGGTGGTTGGTGATGAC	144-165	5.64E-05	0.111	←	→
AT5G49350 (II)	GAGGAGTTTTTGGTTGGACCA	527-548	7.82E-05	0.111	←	→
CASPL1B1	TAGGTTTTGGTGGTTGTGAAG	134-155	6.32E-05	0.287	←	←
LTP8	-	-	-	-	-	-
AT2G43670	-	-	-	-	-	-
PDLP4	TAGGTAAATTTAGTTGAGCGG	499-520	8.63E-05	0.246	→	→
AHA4	TTGAAATTGGTAGTTAAACGG	1729-1750	3.34E-05	0.161	→	→
GPAT6	GTCTTCACCAACTTTAACTCT	465-486	2.44E-05	0.091	←	→
ABCG2	AGTAAATGGGTAGGTGAGAGG	238-259	1.05E-05	0.043	←	←
CYP708A1	TGGTATAAATTTGGTAAAAAT	950-971	7.29E-05	0.440	←	→
QRT1	-	-	-	-	-	-
MYB54	AACAGAGAGTTAGTTGGGAAA	1271-1292	7.59E-05	0.366	→	←
CYS5	-	-	-	-	-	-
PME10	-	-	-	-	-	-
AT3G27270	ATGAAAGGGTTTGGTGAGGGA	94-115	4.14E-06	0.017	→	←
AT4G17480	-	-	-	-	-	-
S-ACP-DES1	-	-	-	-	-	-
MYB93	TTCTGGGGGTTTGGTAAATCA	1365-1386	1.12E-05	0.047	→	←
ARABIDILLO1	-	-	-	-	-	-
ARABIDILLO2	-	-	-	-	-	-
SCARECROW	**GGGAAGAGATTAGGTGGGTGA**TCGG	451-472 & 447-468	1.03E-05 & 4.86E-05	0.051 & 0.079	→	←

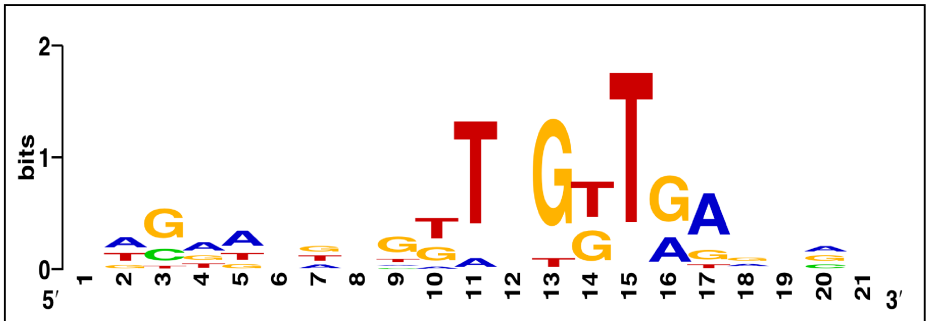
Figure 3.4. Conserved motifs in the putative AtMYB93 DNA binding domains of the 16 candidate target genes. *(A) ClustalO and (B) MUSCLE alignments of the putative AtMYB93 DNA binding domains. Sequence logos made using WebLogo and size of letter represents level of conservation between sequences. Sequence alignments coloured by BOXSHADE. Grey ≥80% nucleotide conservation, Black ≥ 90% nucleotide conservation.*

A




```
AT5G49350(I)     1 --CAGAAGGTGGTTGCTGATGAC--
AT5G49350(II)    1 --GAGGAGTTTTTTGCTGAGCCA--
CASPL1B1         1 --TAGGTTTTGGTGGTTGTGAAG--
PDLP4            1 TAGGTAAATT--TAGTTGAGCGG--
AHA4             1 --TTGAAATTGGTAGTTAAACGG--
GPAT6            1 ----GTCTTCACCAACTTTAACTCT
ABCG2            1 --AGTAAATGGGTAGGTGAGAGG--
CYP708A1         1 --TGGTATAAATTTGGTAAAAAT--
MYB54            1 --AACAGAGAGTTAGTTGGGAAA--
AT3G27270        1 --ATGAAAGGGTTTGGTGAGGGA--
MYB93            1 --TTCTGGGGGTTTGGTAAATCA--
```

B




```
AT5G49350(I)     1 --CAGAAGGTGGTTGCTGATGAC--
AT5G49350(II)    1 --GAGGAGTTTTTTGCTGAGCCA--
CASPL1B1         1 --TAGGTTTTGGTGGTTGTGAAG--
PDLP4            1 --TAGGTAAATTTAGTTGAGCGG--
AHA4             1 --TTGAAATTGGTAGTTAAACGG--
GPAT6            1 --GTCTTCACCAACTTTAACTCT--
ABCG2            1 --AGTAAATGGGTAGGTGAGAGG--
CYP708A1         1 --TGGTATAAATTTGGTAAAAAT--
MYB54            1 --AACAGAGAGTTAGTTGGGAAA--
AT3G27270        1 --ATGAAAGGGTTTGGTGAGGGA--
MYB93            1 --TTCTGGGGGTTTGGTAAATCA--
```

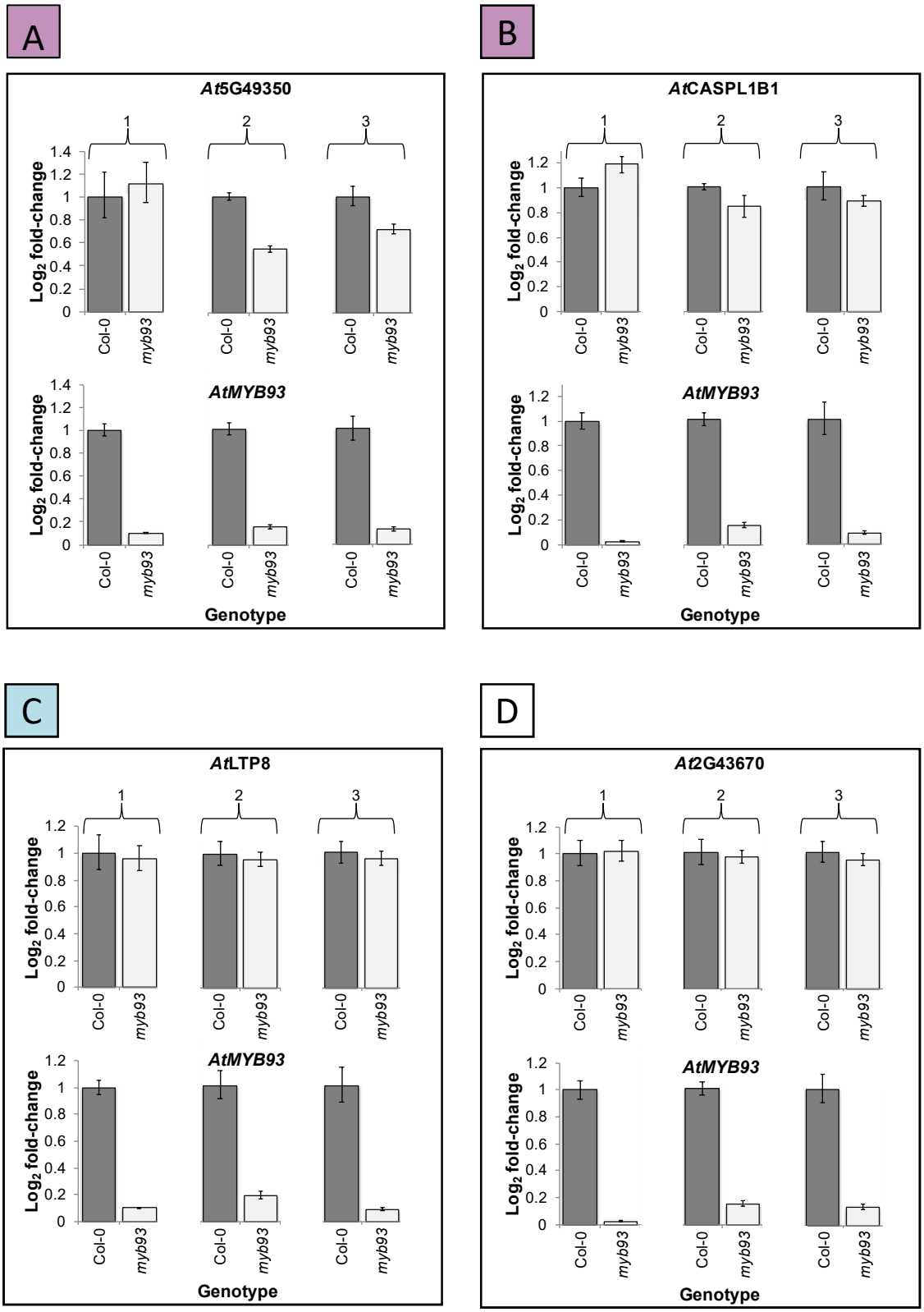
3.3 Strategies for confirming putative AtMYB93 targets

3.3.1 Quantitative reverse transcription PCR was used to analyse relative expression patterns

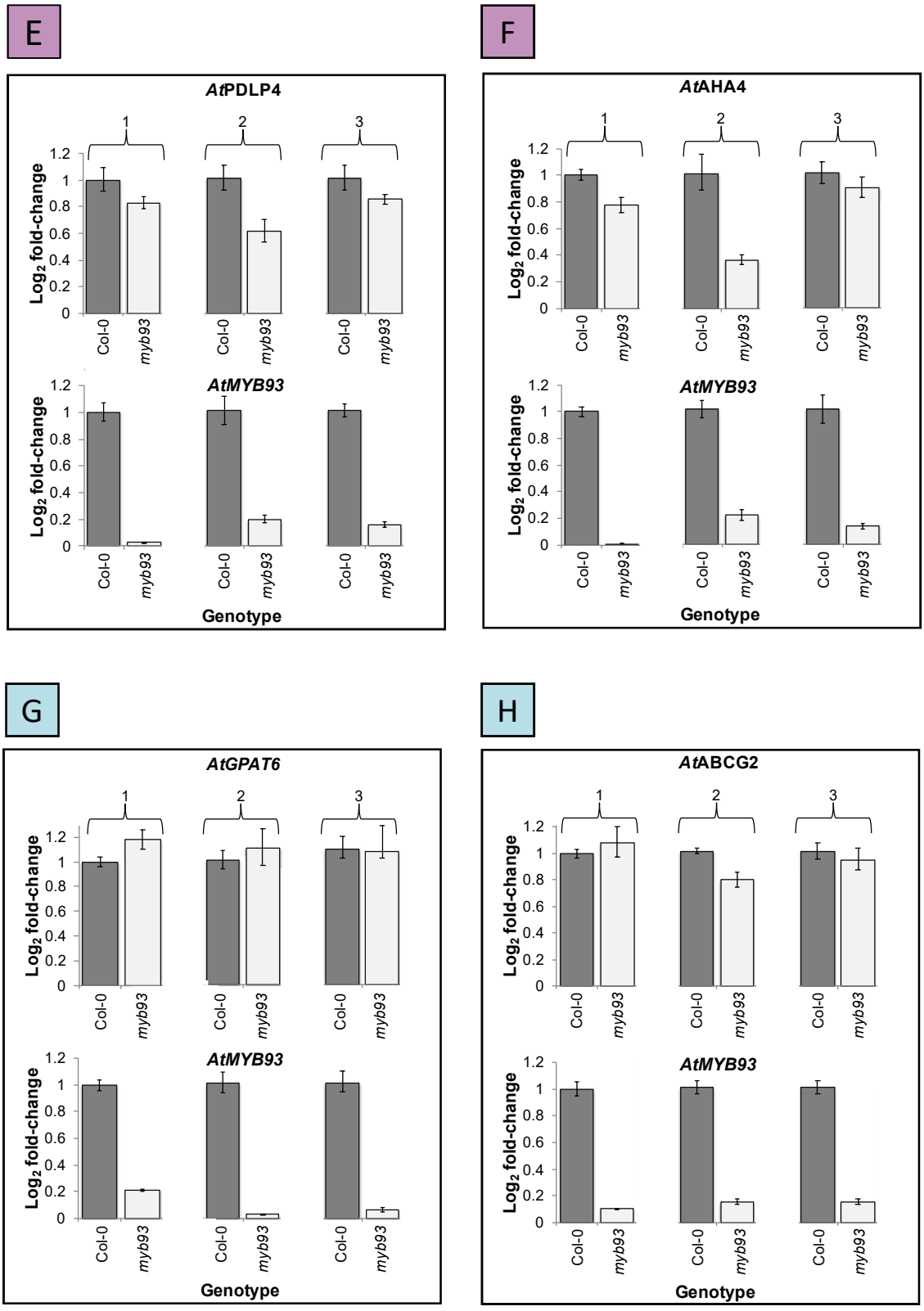
qRT-PCR was used to ask which of the 16 genes *At*MYB93 is potentially acting on. *AtMYB93* expression in wild type (WT), *Atmyb93-1* knock out mutant and *AtMYB93* over-expressor seedling tissue was quantified as a control and compared with the expression of the 16 candidate genes in the same genotype backgrounds. Two housekeeping genes (*AtACT2* and *AtUBC21*) were chosen to normalise the experimental Cq values against. Averaging the results of multiple housekeeping genes increases the accuracy of the experiment (Vandesompele *et al.*, 2002). *AtACT2* was selected due its widespread use in comparable studies, while *AtUBC21* shows very good expression stability (Czechowski, 2005).

To balance the limitations of this experiment it was considered better to use whole seedlings rather than roots from which to extract RNA, as this allowed a larger amount of high-quality RNA to be obtained more quickly and easily. However, as explained in the discussion, a variety of attempts were proposed to make it more feasible to obtain enough root tissue, or at least boost the proportion of root tissue in comparison with shoot tissue. Figure 3.5 represents qRT-PCR fold-change expression data for the 16 candidate genes coloured according to their putative function as in Table 3.3, with the relevant *AtMYB93* sample for each replicate as a control. The results highlight *AtPDLP4* (Figure 3.5.E), *AtAHA4* (Figure 3.5.F), *AtCYP708A1* (Figure 3.5.I), *AtMYB54* (Figure 3.5.K), *AtPME10* (Figure 3.5.M) and also *At4G17480* (Figure 3.5.O) as the six genes that consistently showed reduced expression in the *myb93-1* mutant, though only expression in *AtPME10* was found to

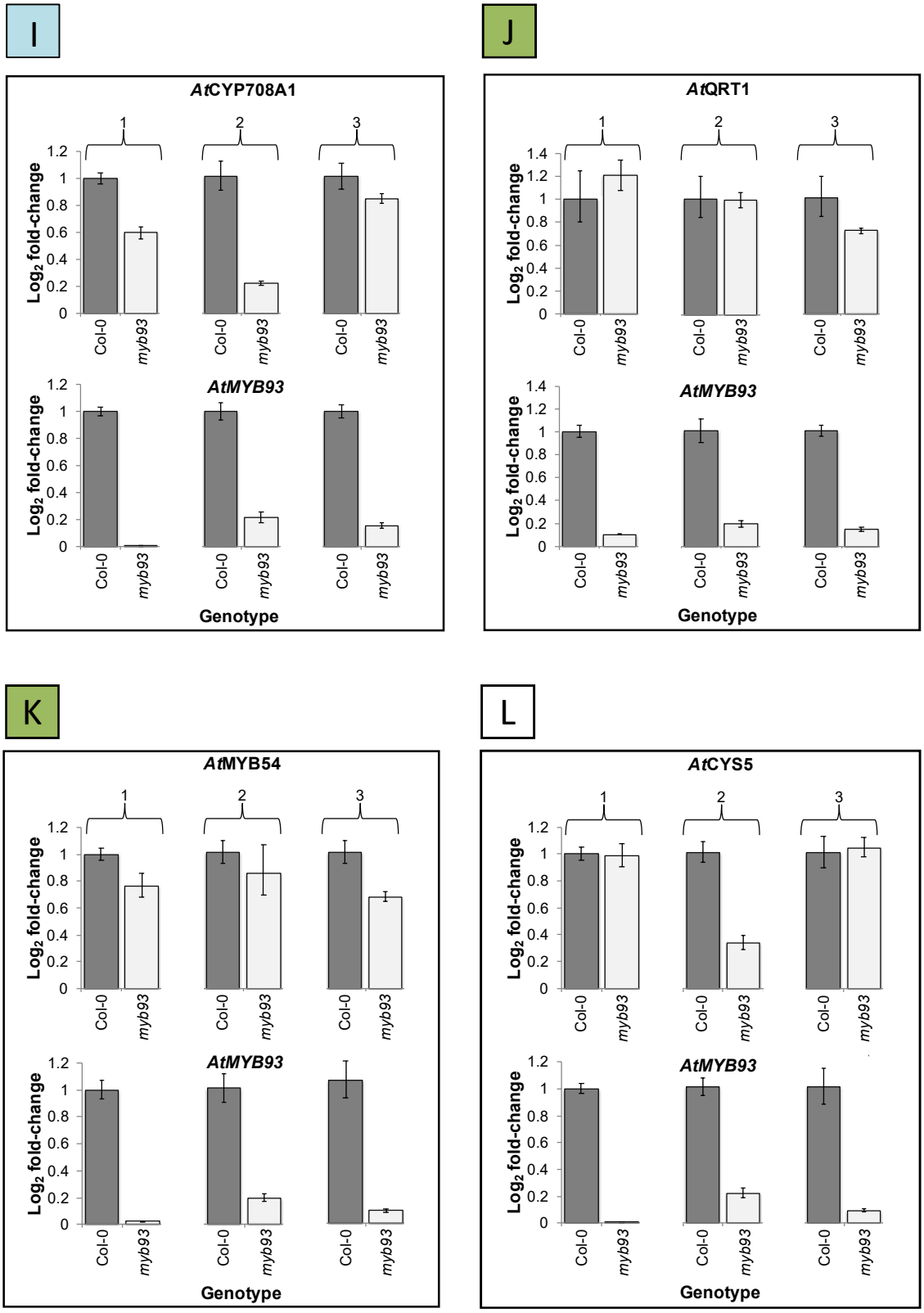
Figure 3.5. qRT-PCR analysis of relative expression of AtMYB93 putative targets in 7-day-old seedlings. *Fold-change relative expression of the 16 potential target genes of AtMYB93, in wild type and Atmyb93 mutant tissue. (A) At5G49350, (B) AtCASPL1B1, (C) AtLTP8, (D) At2G43670, (E) AtPDLP4, (F) AtAHA4, (G) AtGPAT6, (H) AtABCG2, (I) AtCYP708A1, (J) AtQRT1, (K) AtMYB54, (L) AtCYS5, (M) AtPME10, (N) At3G27270, (O) At4G17480 and (P) AtS-ACP-DES1. Cq values normalised to AtACTIN2 and AtUBIQUITIN10 housekeeping controls and fold changes calculated using ΔΔCt (Livak & Schmittgen, 2001). Data shown as three separate biological repeats with corresponding expression of AtMYB93 as a control. Error bars represent upper and lower ranges of fold-change calculated by incorporating the standard deviation of ΔΔCt into the fold-change. Statistics performed on ΔCt values of combined replicates.*



Gene expression p-values	*At5G49350*	*AtCASPL1B1*	*AtLTP8*	*At2G43670*
Two-tailed unpaired t-test on dCT values	0.136	0.810	0.931	0.412

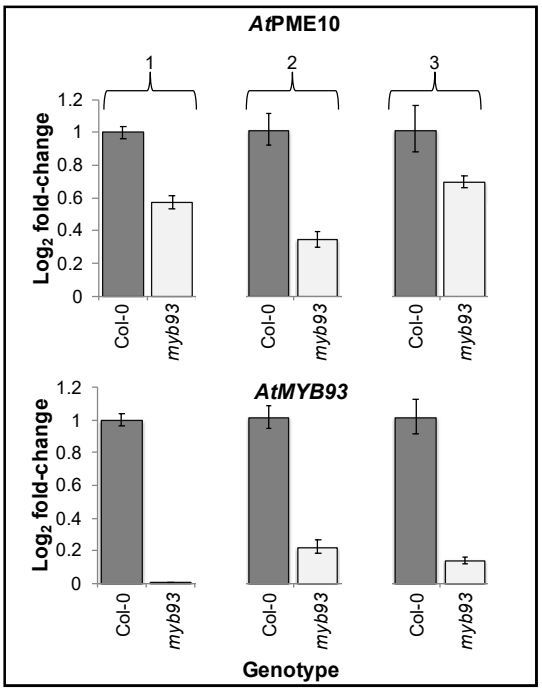


Gene expression p-values	AtPDLP4	*AtAHA4*	*AtGPAT6*	*AtABCG2*
Two-tailed unpaired t-test on dCT values	0.373	0.218	0.822	0.807

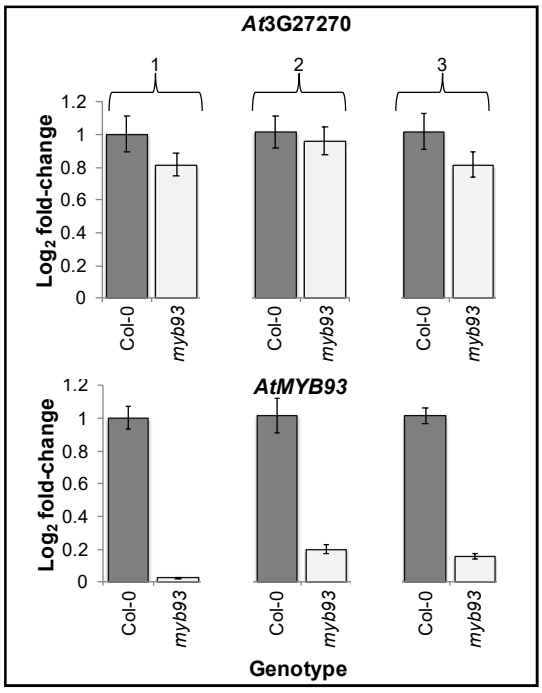


Gene expression p-values	*AtCYP708A1*	*AtQRT1*	*AtMYB54*	*AtCYS5*
Two-tailed unpaired t-test on dCT values	0.670	0.952	0.633	0.899

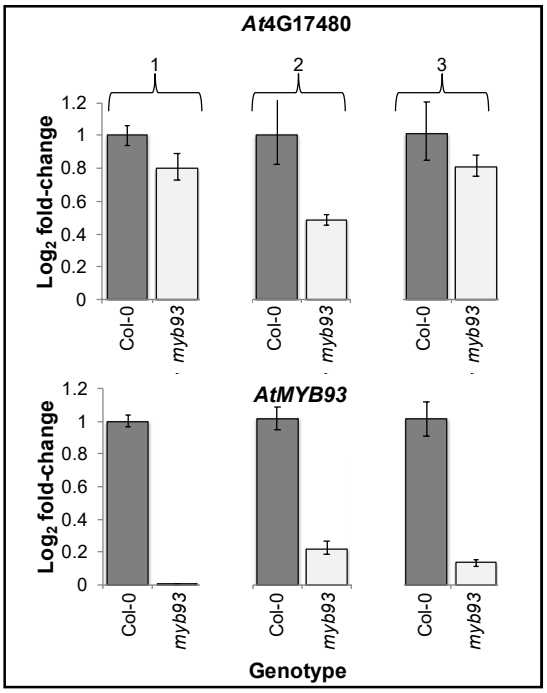
M



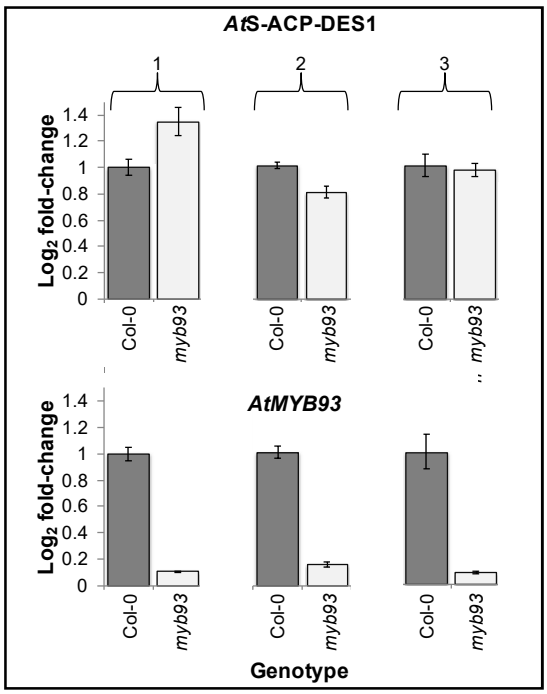
N



O



P



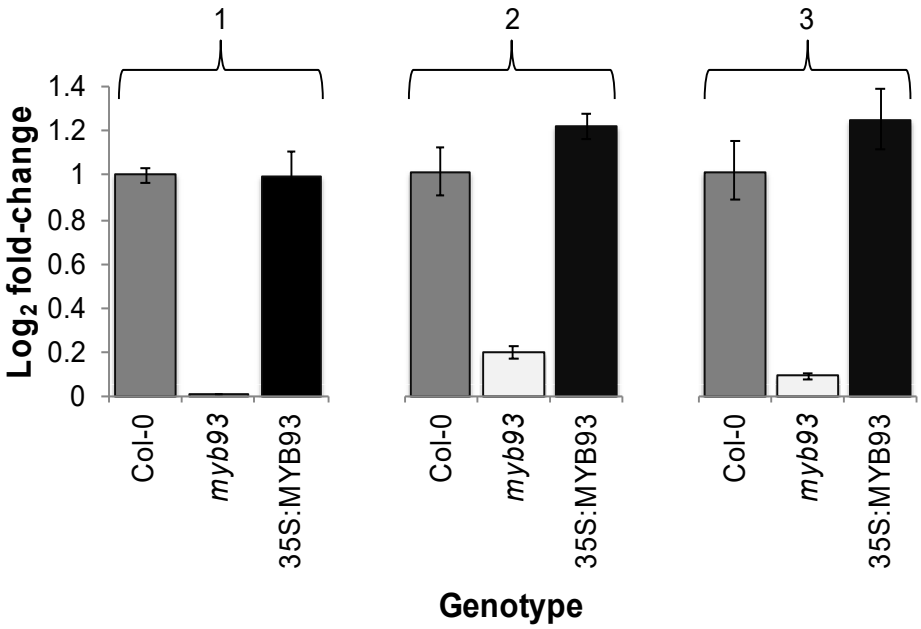
Gene expression p-values	*AtPME10*	*At3G27270*	*At4G17480*	*AtS-ACP-DES1*
Two-tailed unpaired t-test on dCT values	**0.036** *	0.602	0.823	0.953

be significantly different when compared to WT (p=0.036). The over-expressor data are less reliable, as indicated by the minimal upregulation of *AtMYB93* in the controls in Figure 3.6 and have largely been disregarded. Due to variations between biological repeats, each has been presented as a separate graph to show the shared trends but avoid producing error bars that may be misleading. These variations are likely due to the signal to noise ratio extending from the fact that the data represent changes in expression in just a few cells expressing *At*MYB93, whereas samples are from whole seedlings.

3.3.2 Yeast one-hybrid to detect AtMYB93 protein-candidate gene interactions

The next question was to ask if *At*MYB93 is actually binding to the promoters of any of the 16 candidate genes. Yeast one-hybrid (Y1H) was used to assess this, as it is a common technique used for examining at protein-DNA interactions (Ouwerkerk & Meijer (2001); Reece-Hoyes & Walhout, 2012; Uno *et al.*, 2000). The basic principle of Y1H involves the use of a 'prey' plasmid expressing a protein and 'bait' plasmid containing promoter DNA (Ji *et al.*, 2014; Ouwerkerk & Meijer, 2001). In addition to the promoter region for the gene of interest, the bait plasmid also contains a reporter gene, which is found immediately downstream of the promoter sequence and is regulated by a non-endogenous promoter. Additionally, the bait plasmid should have the ability to integrate into the yeast genome (Reece-Hoyes & Walhout, 2012). The prey plasmid contains a selection marker gene and, crucially, the protein of interest is translated into a fusion protein with a constitutively expressed trans-activation domain (Reece-Hoyes & Walhout, 2012). In the assay, if a protein-DNA interaction occurs the activation domain and non-endogenous promoter will come

Figure 3.6. qRT-PCR analysis of relative AtMYB93 expression between Col-0, Atmyb93 mutant and 35S:AtMYB93 over-expressor mutants. *Fold-change relative expression of AtMYB93 shows a lack of significant upregulation in the AtMYB93 over-expressor line. Cq values normalised to AtACTIN2 and AtUBIQUITIN10 housekeeping controls and fold changes calculated using ΔΔCt (Livak & Schmittgen, 2001). Data shown represents an example of three separate biological repeats. Error bars represent upper and lower ranges of fold-change calculated by incorporating the standard deviation of ΔΔCt into the fold-change. Statistics performed on ΔCt values of combined replicates.*



MYB93 expression p-values	*myb93* vs Col-0	35S:MYB93 vs Col-0
one-way ANOVA with Dunnett's comparison	**0.017** *	0.893

into close proximity and the reporter gene is expressed. A diagram representing the Y1H principle is shown in figure 3.7. Y1H has advantages over other cheap techniques such as luciferase reporter assays, since it can detect interactions where transcription is not necessarily being activated – for example if one of the candidate genes is being repressed by *At*MYB93.

Figure 3.8.A shows the cloning steps used to create the Y1H bait constructs. Initially, the estimated promoter regions of each putative target gene were amplified by PCR and then firstly cloned into pCR^{TM}-Blunt, a vector that can be used to clone blunt-ended DNA with high-efficiency. Using pCR^{TM}-Blunt also makes it easier to sequence constructs before transferring them to downstream vectors. The promoters were digested using combinations of *Not*I, *Spe*I and *Xba*I to release the sequenced construct from the pCR^{TM}-Blunt vector (section 2.7), so that it could be ligated into an appropriate vector for use in yeast, in this case, pINT1-HIS3NB. pINT1-HIS3NB was created by Ouwerkerk and Meijer (2001) as a hybrid between the pINT1 and pHIS3NB plasmids, to produce a yeast vector that is able to integrate itself, and therefore the construct, into the yeast genome. To create the construct required to express the prey protein full length *AtMYB93* was amplified by PCR, before being cloned into pGADT7 in *E. coli* (Figure 3.8.B).

Table 3.5 shows the stage of cloning that was reached for each of the 16 candidate gene promoters. Unfortunately, no binding assays with *At*MYB93 could be performed because I was unable to successfully transform the pGADT7 vector containing *AtMYB93* into the Y187 yeast strain, despite the empty vector being easily transformed into the same strain. Several transformation attempts were made, including incubating transformed yeast at different temperatures, subsequently

Figure 3.7. Diagram of the yeast one-hybrid interaction procedure. *(A) Positive interaction between 'prey' protein and 'bait' promoter sequence upstream of HIS3 reporter gene. Proximity of Gal4 activation domain (AD) to Gal4 promoter drives transcription of HIS3 reporter gene. Yeast nutritional requirements for growth are met by HIS3 expression. (B) No interaction so no activation of Gal4 promoter and no transcription of reporter gene. Yeast unable to grow without HIS3 expression.*

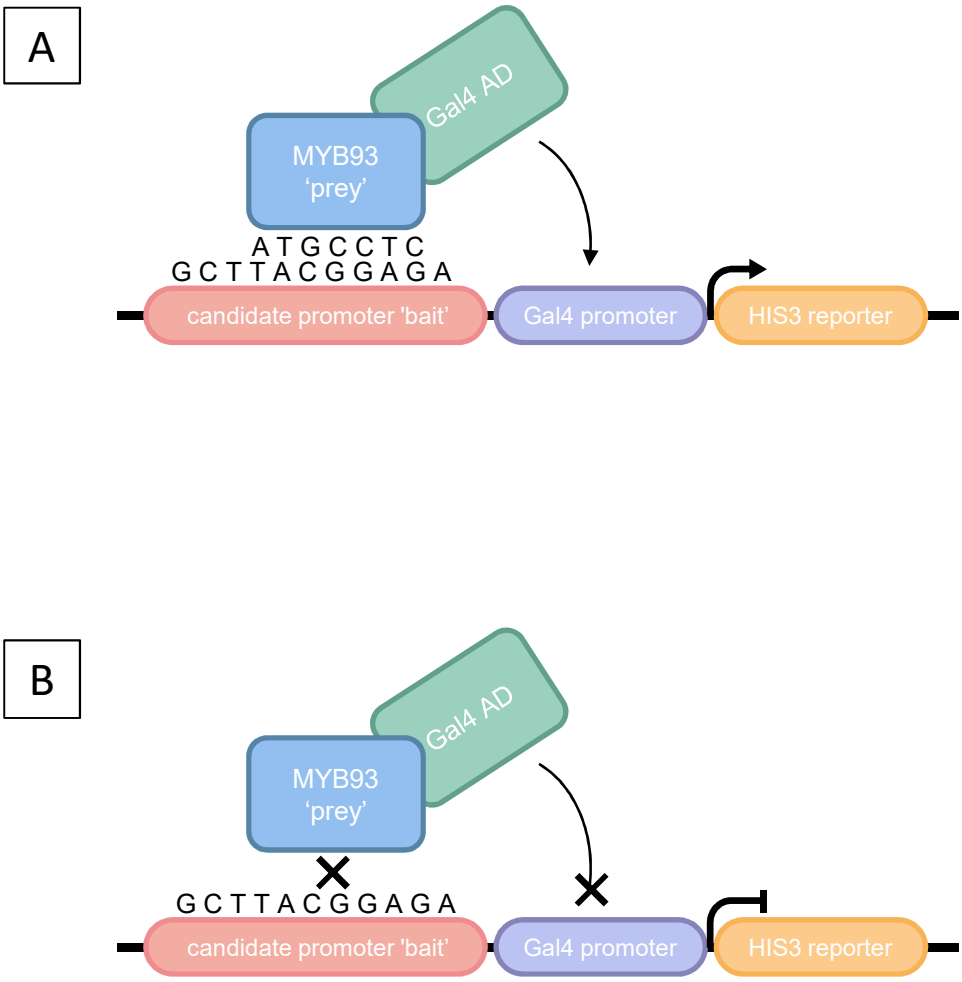
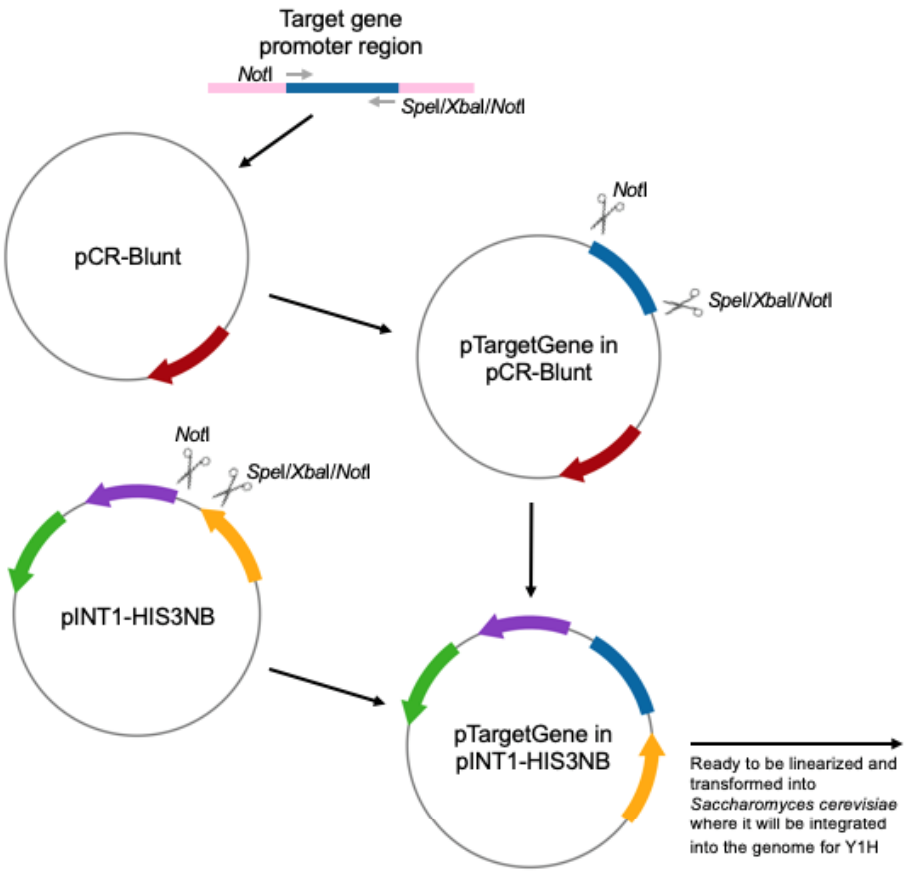

A
Gal4 AD
MYB93
'prey'
A T G C C T C
G C T T A C G G A G A
candidate promoter 'bait'
Gal4 promoter
HIS3 reporter
B
Gal4 AD
MYB93
'prey'
G C T T A C G G A G A
candidate promoter 'bait'
Gal4 promoter
HIS3 reporter

Figure 3.8. Schematic diagram of restriction enzyme cloning process used to create yeast one-hybrid prey and bait plasmids. *Creation of the (A) 'bait' and (B) 'prey' plasmids. Blue = target gene promoter DNA, dark red arrow = kanamycin resistance gene, green arrow = ampicillin resistance gene, purple arrow = histidine gene, yellow arrow = APT1 gene, grey arrow = leucine gene, maroon and bright blue = GAL4 activation domain and promoter, bright red = AtMYB93 gene. (Created using BioRender).*

A


Target gene
promoter region
NotI
SpeI/XbaI/NotI
pCR-Blunt
NotI
SpeI/XbaI/NotI
pTargetGene in
pCR-Blunt
NotI
SpeI/XbaI/NotI
pINT1-HIS3NB
pTargetGene in
pINT1-HIS3NB
Ready to be linearized and
transformed into
Saccharomyces cerevisiae
where it will be integrated
into the genome for Y1H

B

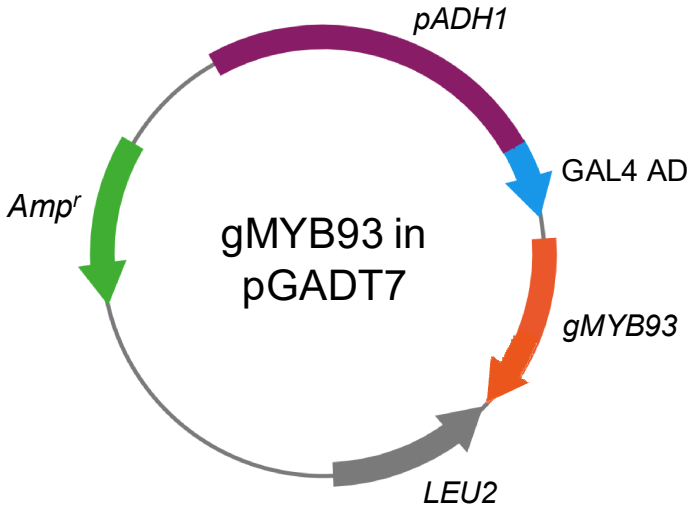

pADH1
GAL4 AD
gMYB93
LEU2
Ampr
gMYB93 in
pGADT7

Table 3.5. Yeast one-hybrid progress. *Stage reached for cloning yeast one-hybrid promoter constructs into yeast. Cloning was paused after AtMYB93 failed trainsformation into the Y187 strain of yeast.*

Gene	Amplified by PCR	Ligation into PCRblunt & transformation into Dh5α	Ligation into pINT & transformation into Dh5α	Linearisation and transformation into Y187
AT5G49350	✓	✓		
AT5G44550 (CASPL1B1)	✓	✓		
AT2G18370 (LTP8)	✓	✓	✓	
AT2G43670	✓	✓	✓	✓
AT3G04370 (PDLP4)	✓	✓	✓	
AT3G47950 (AHA4)	✓	✓	✓	
AT3G47950 (GPAT6)	✓	✓	✓	
AT2G37360 (ABCG2)	✓	✓	✓	✓
AT1G55940 (CYP708A1)	✓	✓	✓	✓
AT5G55590 (QRT1)	✓			
AT1G73410 (MYB54)	✓	✓	✓	✓
AT5G47550 (CYS5)	✓	✓	✓	
AT2G19150 (PME10)	✓	✓	✓	
AT3G27270	✓			
AT4G17480	✓	✓		
AT5G16240 (S-ACP-DES1)	✓			

culturing the plates at different temperatures and increasing the nutrient content of the growth medium. As a preliminary test a binding assay between the empty pGADT7 vector and two of the promoter regions (that of *AtABCG2* and *At2G43670*) was conducted to check for autoactivation of the *HIS3* reporter gene. Unfortunately, this was confirmed by the assay and so the use of 3-AT, a competitive inhibitor of the HIS3 protein will be required to reduce the activation threshold once the system is working.

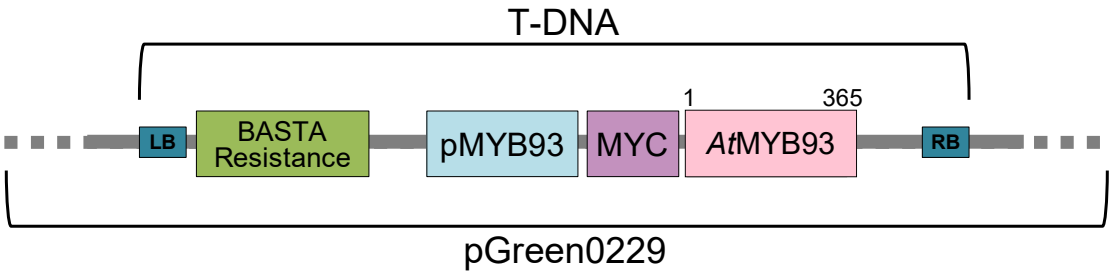
3.3.3 Targeted chromatin immunoprecipitation techniques

Use of chromatin immunoprecipitation–sequencing (ChIP-seq) and ChIP-qPCR was planned to validate the results of the qRT-PCR experiment and to support Y1H evidence that a direct interaction is taking place between *At*MYB93 and the candidate genes. Importantly, this type of experiment confirms such interactions *in planta* and so it is possible to determine the functional consequences of a protein-DNA interaction identified from Y1H. ChIP methods reveal protein-DNA interactions by first using chromatin immunoprecipitation to pull down DNA bound by a protein, followed by DNA sequencing to identify the genes (Park, 2009). Crucially, the difference between the two methods is that ChIP-seq identifies any DNA that the TF binds to in the plant, while ChIP-qPCR is a targeted approach, which looks at the binding of a TF to a group of preselected genes based on prior data, such as my 16 candidate genes.

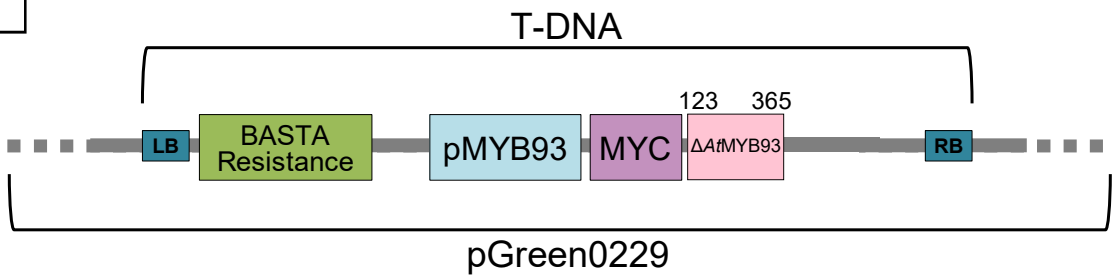
Figure 3.9.A depicts the construct that was made to transform into *Atmyb93-1* mutant background *Arabidopsis* plants. *Atmyb93* mutant plants were used as the background for T-DNA insertion to avoid overexpression of *AtMYB93* as a result of

Figure 3.9. Graphical Illustration of the constructs that have been created to generate transgenic Arabidopsis for ChIP-qPCR study. *The AtMYB93 promoter, myc tag and AtMYB93 were cloned into the T-DNA region of pGreen0229. A nos-bar cassette at the left border confers BASTA resistance to allow for plant screening of positive transformants. Constructs contain either (A) full length AtMYB93 gDNA or (B) truncated AtMYB93 gDNA lacking its predicted DNA binding domains identified in Gibbs et al. (2014).*

A


T-DNA
1
365
LB
BASTA Resistance
pMYB93
MYC
AtMYB93
RB
pGreen0229

B

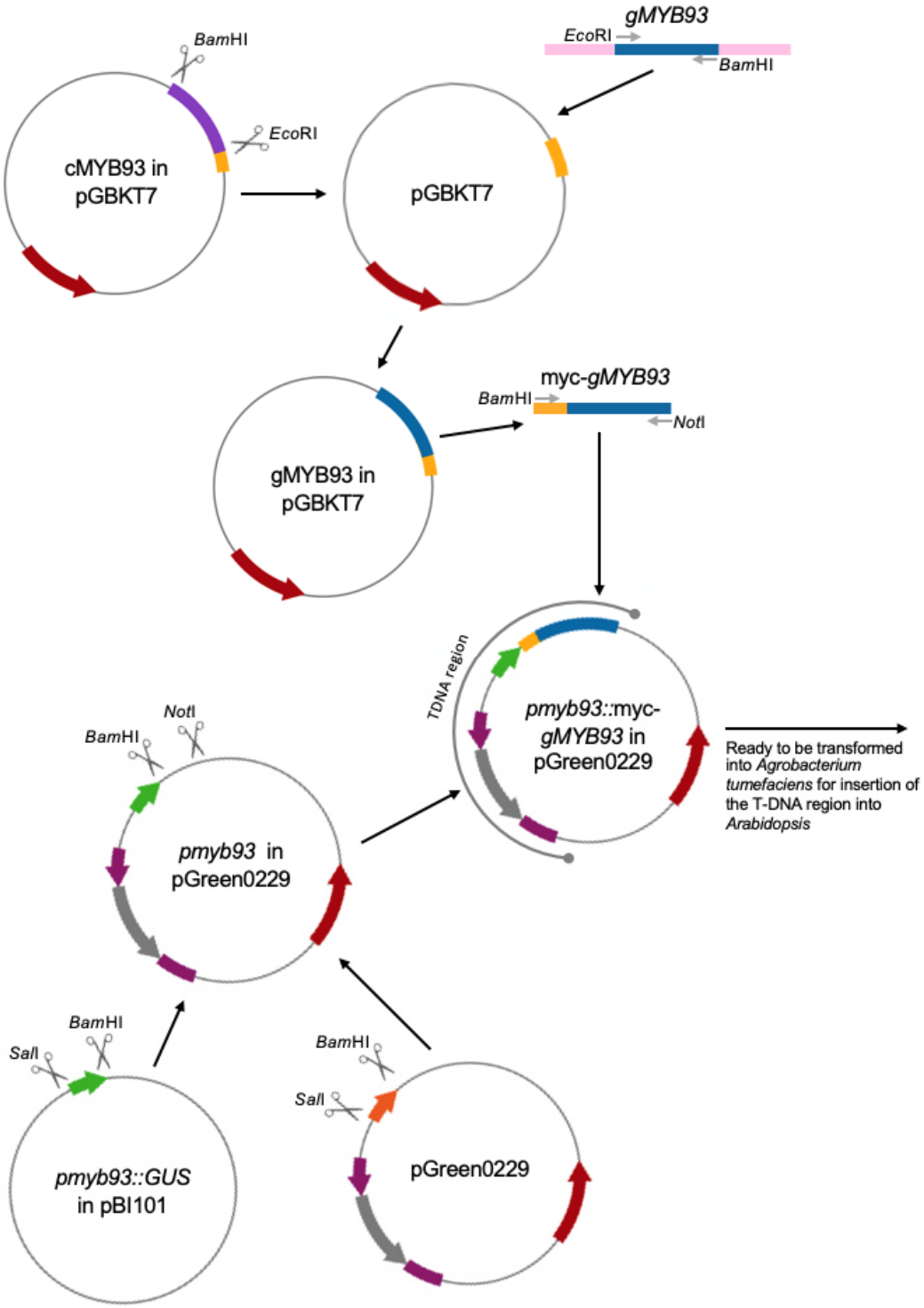

T-DNA
123
365
LB
BASTA Resistance
pMYB93
MYC
ΔAtMYB93
RB
pGreen0229

the plant containing more than one copy of the gene. Likewise, the construct contains the *AtMYB93* promoter in order to drive expression of near-endogenous levels of *AtMYB93* in *Arabidopsis*. Finally, a copy of *AtMYB93* consisting of gDNA was used to make the construct, rather than the transcriptional product, cDNA, to more closely mimic endogenous conditions. Together, these measures are designed to ensure the most accurate levels of DNA binding and help to reduce false positive results. A second construct (Figure 3.9.B) was also made to act as a control in order to determine the background levels of binding. This construct is the same as the first, except a truncated version of *At*MYB93 had been used that lacks the DNA binding domain (Gibbs *et al.*, 2014): this should not interact with the target genes of *At*MYB93.

The cloning steps followed to create these constructs is shown in Figure 3.10. Genomic *AtMYB93* was first amplified via PCR using primers containing *Eco*RI and *Bam*HI restriction sites. A c-myc epitope tag needed to be added to *AtMYB93* and the plasmid pGBKT7 contains a copy of this tag. The pGBKT7 vector already contained another gene downstream of c-myc that was removed. The *AtMYB93* PCR product was ligated into the multiple cloning site (MCS) of pGBKT7 and the initial construct was sequenced to confirm that *AtMYB93* had been inserted correctly.

The next step involved using primers containing *Bam*HI and *Not*I restriction sites to amplify myc-*At*MYB93 so that it could then be cloned into pGreen0229, along with its promoter. The pMYB93 promoter was obtained from a pBI101 derived vector (Gibbs *et al.*, 2014) and was used to replace the pre-existing 35S promoter in a pGreen0229-based construct 35S::myc-*At*MYB93 (cDNA) (Gibbs *et al.*, 2014) in order to generate transgenic plants with endogenous expression of *At*MYB93. Once

Figure 3.10. Schematic diagram of restriction enzyme cloning process used to create ChIP-qPCR constructs. *The AtMYB93 promoter, myc tag and AtMYB93 were cloned into the T-DNA region of pGreen0229. Purple = AtcMYB93, blue = AtgMYB93, yellow = myc tag, dark red arrow = kanamycin resistance gene, green arrow = 35S promoter, bright red arrow = MYB93 promoter, maroon and grey = nos-Bar cassette. (Created using BioRender).*


gMYB93
EcoRI
BamHI
BamHI
EcoRI
cMYB93 in pGBKT7
pGBKT7
gMYB93 in pGBKT7
myc-gMYB93
BamHI
NotI
TDNA region
pmyb93::myc-gMYB93 in pGreen0229
Ready to be transformed into Agrobacterium tumefaciens for insertion of the T-DNA region into Arabidopsis
NotI
BamHI
pmyb93 in pGreen0229
BamHI
SalI
pmyb93::GUS in pBI101
BamHI
SalI
pGreen0229

the complete construct was in pGreen0229 it was sequenced again. In particular, it was necessary to confirm the integrity of the boundaries between each component of the construct and the vector, for example, to check the protein fusion and correct reading frame.

Agrobacterium tumefaciens was used to provide a mechanism for the integration of the construct into the plant and the BASTA® resistance marker on the T-DNA sequence was used for plant screening. The presence of successful transformants was confirmed by amplifying full length and truncated *AtMYB93* from plasmids extracted from individual colonies.

Atmyb93-1 Arabidopsis plants were inoculated with *A. tumefaciens* containing the two constructs on five occasions, but all were unsuccessful in producing transgenic plants. Attempts were made to improve the dipping efficiency after the initial dipping event failed. One way was to use cultures from different positive colonies obtained after transforming the construct into *A. tumefaciens*. Although these had been stored in glycerol at -80°C some may have contained more viable bacteria than others as it was noticed that growth was not always even when growing up bacterial cultures for dipping. A commonly used method that was utilised in order to increase transformation efficiency was to cut the primary *Arabidopsis* bolts as they were forming (Clough & Bent, 1998). This has the effect of promoting side-bolting and so produces more buds, however it is possible that the siliques produced will be smaller and contain fewer seeds. Lastly, the selection of progeny from the first dipping event, which was carried out on agar plates containing BASTA®, selection of T1 plants was performed using a combination of both agar plates and soil as an alternative method to improve the selection process. Furthermore, non-dipped seeds

were used as a control to ensure that a sufficient amount of BASTA® was used. Unfortunately, none of these methods led to the creation of selectable transgenic plants and the effectiveness of dipping may have been reduced by the fact that the glasshouse was experiencing temperature control issues during most of the attempts.

The pGreen 0229 vector that was used to create the ChIP constructs had previously been used for other cloning experiments in the lab. As a result of this it was considered that it could be possible that a mutation had occurred in the Bar gene of the nos-Bar cassette, which confers resistance to BASTA®, as this region of the plasmid was not sequenced during the cloning of the ChIP constructs. Primers were obtained from Rao *et al.* (2017) that span the internal region and two additional primers were designed to amplify across the 5' and 3' boundaries (as in Appendix 2.2). Sequencing of the starting plasmid material for the creation of these constructs revealed that the BASTA® was intact (sequencing sample preparation performed by V. Gupta). The final constructs transformed into *A. tumefaciens* were, however, not tested and this would be necessary to confirm a mutation did not occur during the cloning process.

3.4 Analysis of RNA-seq data

In parallel to my experiments, an RNA-seq experiment performed on *Arabidopsis* whole root tissue (X. Cao, unpublished) looked at differences in gene expression between WT and *Atmyb93-1* mutant roots in seven-day-old seedlings. This experiment pulled out 256 genes with significant fold-changes in expression (Cao, unpublished). These genes were cross matched with the 16 proposed targets

of *At*MYB93 that we have selected bioinformatically. Though none of these 16 genes were found in the list of significant results, four were found when ranking the entire gene set by significance and narrowing this down to look at the top 1,000 hits. These matches were *AtMYB54*, *AtPDLP4*, *AtQRT1* and *At5G49350.*

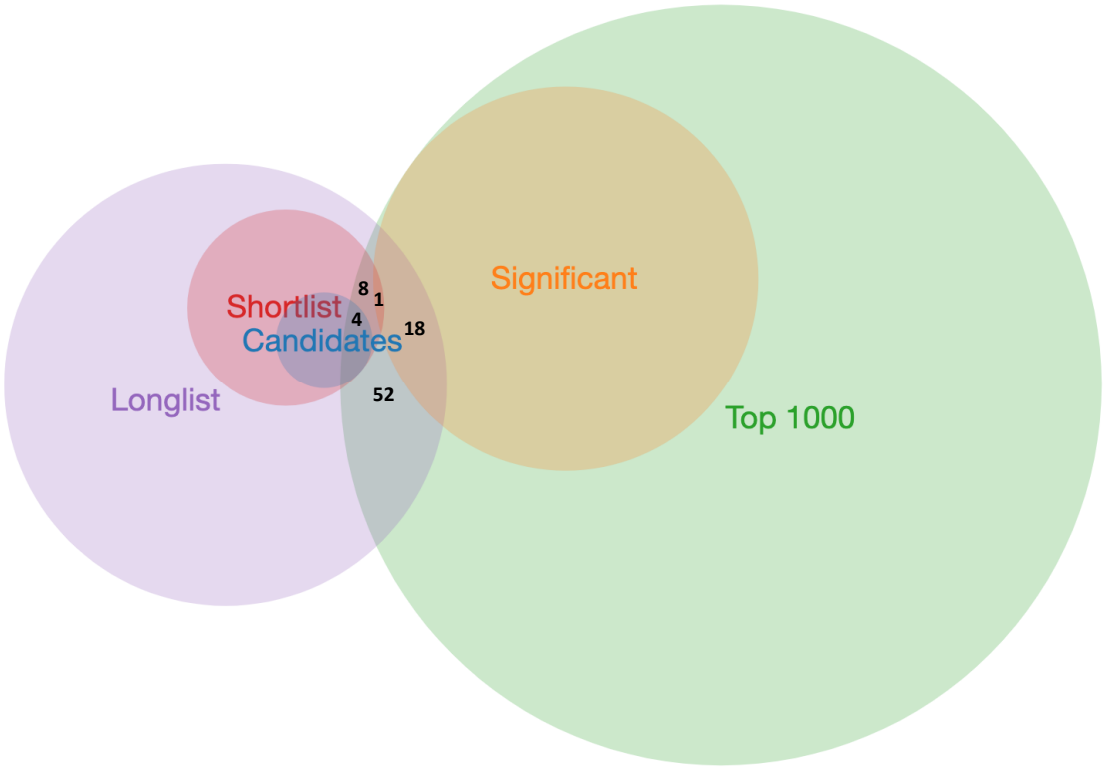
Following this, I decided to look against the initial longlist of 337 genes based on the transcriptome microarray experiment by Voß *et al.* (2015). The genes identified from the RNA-seq experiment were cross matched against the initial longlist genes (Figure 3.11). In total, 18 genes were also found in the 256 significant RNA-seq, and when the data were ranked by significance a further 34 longlist genes had been identified by RNA-seq in the top 1,000 hits. One gene, *AtGH9C3*, found to be significant in the RNA-seq experiment that was also in the 67 gene shortlist.

Table 3.6 provides the data obtained in the RNA-seq experiment for the list of putative *At*MYB93 targets (X. Cao, unpublished). The genes are ranked by p-value for the differences in their fold-changes in expression, though none the data for these genes was found to be significant. The direction of expression in *Atmyb93-1* mutant seedlings versus WT is also shown and compared to the qRT-PCR data measuring the same expression.

As described previously, the longlist was originally cut down to 67 genes before the final 16 were chosen for experimental analysis and some of these genes could prove to be promising *At*MYB93 targets. Of these 67 genes, one (*AtGH9C3*) was present in the list of significant genes from the RNA-seq experiment and three more of these genes were in the top 1,000 (*AtBHLH82*, *AtPOX2* and *At4G17215*).

Figure 3.11. Venn diagram comparing genes obtained through bioinformatic analysis vs RNA-seq during lateral root initiation. *Green = top 1000 RNA-seq genes ranked by significance, orange = 256 significant genes from RNA-seq, purple = longlist of 337 genes from bioinformatic analysis, red = shortlist of 67 genes from bioinformatic analysis, blue = 16 chosen genes from bioinformatic analysis as targets of AtMYB93.*

Table 3.6. Candidate AtMYB93 target gene expression from RNA-seq during lateral root initiation. *RNA-seq expression values of genes in Atmyb93 and Col-0 whole root tissue and respective fold-change of expression. Direction of expression in RNA-seq and qRT-PCR experiments also shown. Blue = downregulated in Atmyb93 mutant relative to WT, red = upregulated in Atmyb93 mutant relative to WT. Genes in bold are those candidates that overlap with the RNA-seq data.*



Gene	*myb93* value	Col-0 value	Log_2 fold-change	RNA-seq (expression in *myb93* relative to WT)	qRT-PCR (expression in *myb93* relative to WT)		
AT3G04370 (PDLP4)	**4.4499**	**9.0385**	**-1.0221**				
AT5G49350	**12.7558**	**21.5313**	**-0.7553**				
AT1G73410 (MYB54)	**4.1222**	**7.0206**	**-0.7680**				
AT5G55590 (QRT1)	**3.5957**	**6.0523**	**-0.7513**				
AT2G19150 (PME10)	5.1533	6.2848	-0.2864				
AT3G47950 (AHA4)	8.3005	7.2298	0.1994				
AT1G55940 (CYP708A1)	2.7583	2.7138	0.0235				
AT2G18370 (LTP8)	138.1387	142.3181	-0.0430				
AT2G38110 (GPAT6)	17.8147	16.5376	0.1073				
AT2G43670	24.2091	22.1465	0.1286				
AT3G27270	9.4100	9.2813	0.0197				
AT4G17480	2.7628	2.9232	-0.0816				
AT5G16240 (S-ACP-DES1)	5.4590	5.5512	-0.0241				
AT5G44550 (CASPL1B1)	120.9992	128.0215	-0.0813				
AT5G47550 (CYS5)	4.7748	5.3175	-0.1557				
AT2G37360 (ABCG2)	0.0181	0.0183	-0.0125				

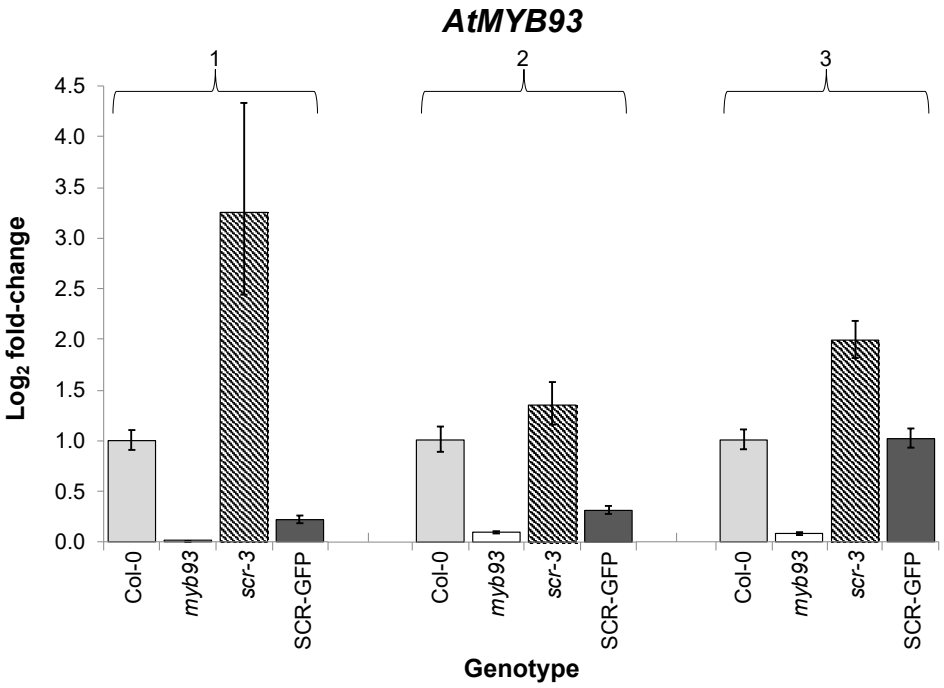
3.5 A potential upstream repressor of AtMYB93

As discussed in section 1.4.3 of chapter I, a review of the literature revealed *AtSCR* to be a possible upstream regulator of *AtMYB93* (Iyer-Pascuzzi *et al.*, 2011; Sparks *et al.*, 2016; Voß *et al.*, 2015). It was hypothesised that the TF *At*SCR is suppressing *AtMYB93* in most of the endodermis and qRT-PCR was used to test this hypothesis. In this experiment, *At*MYB93 expression was quantified in WT, *Atscr-3* mutant and *Atmyb93-1* mutant lines. The *Atscr-3* allele has a point mutation that produces a stop codon part way through the sequence and leads to the formation of a truncated protein (Fukaki *et al.*, 1996). The resultant mutant phenotype is less severe than other *Atscr* mutants, making it ideal for the experiments in this study as plants retain a single-celled endodermal-type layer (Benfey *et al.*, 1993). This is important as *AtMYB93* is expressed in the endodermis and in stronger *Atscr* mutant backgrounds it would not be possible to distinguish whether changes in *AtMYB93* expression were due to lack of *AtSCR* expression and thus reducing *At*SCR mediated repression of *AtMYB93,* or due a lack of endodermal tissue in the plant. Additionally, *Atscr-3* is the only *Atscr* mutant that has been generated in a Columbia (Col-0) background (Fukaki *et al.*, 1998) and this is the same background ecotype as the *Atmyb93-1* mutant, thus producing more easily comparable data.

The data in this experiment shows that *At*MYB93 was upregulated in the *Atscr-3* mutant by a magnitude of 1.34, 1.98 and 3.25-fold relative to the WT over three biological repeats (Figure 3.12.A). These data were not found to be significant, however a similar trend between the biological replicates can be observed. *Atmyb93-1* expression was used as a control and minimal amplification was detected (p=0.009). The *pSCR*::*SCR*-GFP;*Atscr-3* (Gallagher *et al.*, 2004; Goh *et al.*, 2016)

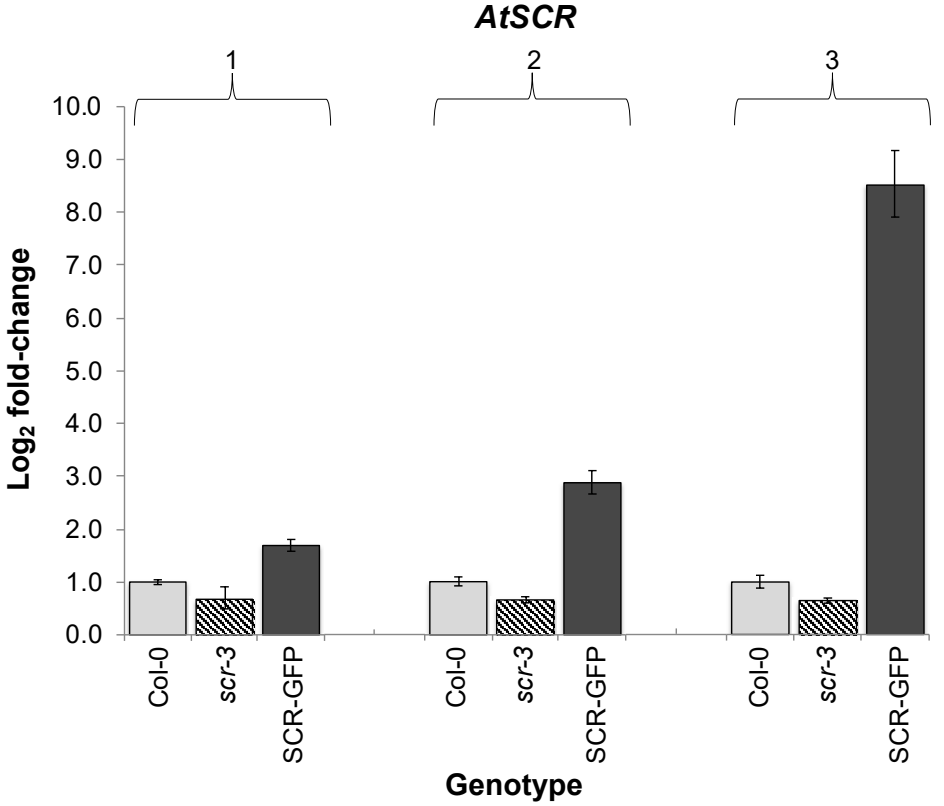
Figure 3.12. qRT-PCR analysis of AtSCR as a possible transcriptional target of AtMYB93. *(A) Fold-change relative expression of AtMYB93 in Col-0, Atscr-3 mutant, Atmyb93 mutant and pSCR::SCR-GFP;Atscr-3 using 7-day-old whole seedling tissue. (B) Control experiment showing fold-change relative expression of AtSCR in wild type, Atscr-3 mutant and pSCR::SCR-GFP;Atscr-3 using 7-day-old whole seedling tissue. In both cases, three biological repeats are shown. Cq values normalised to AtACT2 and AtUBC21 housekeeping controls and fold changes calculated using ΔΔCt (Livak & Schmittgen, 2001). Error bars represent upper and lower ranges of fold-change calculated by incorporating the standard deviation of ΔΔCt into the fold-change.* Statistics performed on *ΔCt* values of combined replicates.

A



MYB93 expression p-values	*myb93* vs Col-0	*scr-3* vs Col-0	SCR-GFP vs Col-0
one-way ANOVA with Dunnett's comparison	**0.009** **	0.741	0.612

B



SCR expression p-values	*scr-3* vs Col-0	SCR-GFP vs Col-0
one-way ANOVA with Dunnett's comparison	0.391	**0.016** *

line was created to complement expression in the *Atscr-3* mutant background, but as this experiment demonstrates, *Atmyb93-1* expression in the *SCR*::*SCR*-GFP;*Atscr-3* line was less than the WT in two out of three biological replicates, though this was not significant. Since the *Atscr-3* mutant does not show complete loss of *AtSCR* it is likely that *SCR*::*SCR*-GFP;*Atscr-3* is 'over-complementing' *AtSCR* expression in *Atscr-3*. This was confirmed by the results in Figure 3.12.B, as expression of *AtSCR* was revealed to be reduced in the *Atscr-3* mutant line, but not significantly so – in contrast to the near total loss of *AtMYB93* expression in the *Atmyb93-1* mutant line. Likewise, *AtSCR* expression in the *SCR*::*SCR*-GFP;*Atscr-3* line was increased by up to seven-fold compared to WT: statistical analysis did demonstrate this as significant ($p=0.016$).

3.6 Discussion

16 candidate *At*MYB93 target genes were identified based on co-expression data from a number of sources. Of these, nine have putative *AtMYB93* binding sites and six show changes in expression that correlate with *At*MYB93 in qRT-PCR analysis with *AtPME10* demonstrating a significant difference in expression in the *Atmyb93* mutant relative to WT. Constructs were made for the 16 candidate gene promoter regions to check for protein-DNA interactions in Y1H. Of these, 4 were transformed into Y187 yeast but the vector containing *AtMYB93* could not be transformed for binding assays to be performed. *AtSCR* was identified as a potential upstream regulator of *AtMYB93*. Two constructs were made to start ChIP-qPCR or ChIP-seq to test downstream target genes of *At*MYB93. Lastly, *AtSCR* was identified as a potential upstream regulator of *AtMYB93*.

3.6.1. Candidate target genes of AtMYB93 can be categorised by function

The candidate gene targets identified in this work can be grouped into three functional categories (Table 3.3): endodermal membrane proteins, cell wall remodelling genes and fatty acid or suberin biosynthesis or metabolism genes. These three elements can each be linked to LR development (section 1.2). The endodermis is the first cell layer that must be crossed by the LR primordium as it advances from the pericycle and begins to emerge, and the presence of complex lignin and suberin lipid biopolymers in addition to the cellulose cell wall poses a specific challenge. Furthermore, cell wall remodelling in the cortex and epidermis is central to the ability of the LR primordia to traverse these tissues. While endodermal cells must change their shape and volume dramatically to accommodate the primordia (Vermeer *et al.*, 2014), the simple loosening of cortical and endodermal cell walls by hydrolytic enzymes is less onerous and allows the cells to separate (Robbins *et al.*, 2014). *AtMYB93* could be regulating expression of genes involved in these functions at sites of LR initiation, and of the 16 candidate genes, 14 are predicted to have relevant roles.

3.6.2 Comparing the candidate gene mutants with RNA-seq data from the Atmyb93 mutant

It is interesting to see that *AtPDLP4* and *AtMYB54*, two of the genes identified by qRT-PCR (section 3.3.1), are represented in the top four of the 16 candidate genes that are in the top 1000 genes from the *AtMYB93* RNA-seq dataset, when ranked by significance. This further validates that these genes appear to be promising candidates as downstream targets of *At*MYB93. The presence of *At*GH9C3 as a

gene with a significant difference in expression from the RNA-seq and as a shortlist candidate gene merits further examination. *At*GH9C3 is a glycosyl hydrolase that is downregulated in the *Atmyb93-1* mutant (X. Cao, unpublished), while being upregulated in response to sulfur (Winter *et al.*, 2007). It is expressed in the root epidermis and has putative functions in cell wall breakdown, meaning it could have a role in enabling LR emergence. Below are details on the 16 candidate targets of *AtMYB93*, including any published work characterising their roles or that of closely related genes. Based on the evidence presented in this chapter, the genes to prioritise for further experimental work are *AtPME10, AtABCG2 and AtMYB54*. Additional next steps should also include the interrogation of published root-cell type data to assess the specific localisation of the putative target genes, though it may be difficult to draw conclusions for some genes since their expression is low and standard deviation can be high in the datasets.

3.6.3 Identities of candidate genes found to match the expression profile of AtMYB93

3.6.3.1 AtPME10: a putative cell wall remodelling enzyme

AtPME10 is predicted to encode a pectin methylesterase enzyme. Pectin is a polysaccharide important for binding plant cells and for cell wall expansion (Mohnen, 2008) and based on its sequence homology *AtPME10* is implicated in the first stage of pectin degradation. The role of cell loosening in LR emergence is well documented (section 1.2.4) (Péret *et al.*, 2009; Péret *et al.*, 2012) as this enables the cells of cortical and epidermal cells overlying the new LR primordia to separate, in turn allowing the developing LR to traverse these tissues and reach the soil (Swarup *et*

al., 2008). A study by Ociepa (2017) looked at the effect of PME expression in response to polyphenon-60 (PP60) treatment. PP60 acts in a similar way to PME inhibitors by binding the active site of the enzyme, and this inhibition can sometimes promote PME upregulation in response. This was revealed to be the case for *AtPME10,* and a similar result was observed in another candidate – *AtQRT1,* while the other genes tested were expressed similarly to untreated conditions (Ociepa, 2017). Growing *Arabidopsis* with PP60 leads to a lack of primary root elongation and reduced number of LR, thought to be the result of increased PME expression.

Pectin methylesterases are often found to be acting redundantly and four other pectinesterases were identified in the original longlist of candidates, though not in the top 1000 RNA-seq genes. These are: *AtQRT2*, *AtPME8*, *AtPME54* and *AtPME24,* and while *AtPME8*, *AtPME54* and *AtPME24* were not ultimately selected as candidates due to their significant fold-change in expression occurring in the wrong time frame, *AtQRT2* was close to being selected as a putative *At*MYB93 target. *AtQRT2* has enriched root expression (Kim & Patterson, 2006; Ogawa *et al.*, 2009) and encodes a polygalacturonase involved in pectin degradation during cell wall separation. Experimental characterisation of this gene has so far found a role in reproductive development (Kim & Patterson, 2006; Ogawa *et al.*, 2009). Nevertheless, Ogawa *et al.* (2009) did look at expression in the roots with a *pAt*QRT2::GUS reporter and observed localised expression in cells surrounding emerging LRs, however no obvious root phenotype was observed in mutants by the authors, but this is something that could be re-examined.

3.6.3.2 AtCYP708A1: a gene putatively involved in fatty acid or suberin metabolism

AtCYP708A1 is part of the cytochrome P450 superfamily of proteins. Plant P450s can encode enzymes that catalyse reactions involving structural polymers including lignin, cutin and suberin (Bak *et al.*, 2011). Though there is no published data on the role of *AtCYP708A1* itself, a number of other P450s have been implicated in suberin and complex polymer biosynthesis (Pinot & Beisson, 2011), though the P450 family comprises a very large and diverse family.

Three additional P450s that were originally picked out as potential *At*MYB93 targets in the initial longlist of candidate genes were also identified through RNA-seq (Cao, unpublished). *AtCYP86A4* demonstrated a significant change in expression after LR initiation, and both *AtCYP86A2* and *AtCYP705A1* were in the top 1,000 genes ranked by significance. *AtCYP86A4* was not selected as a candidate gene as its significant peak expression after LR initiation occurred in the earlier time period (12 h vs 9 h), though this peak was only a 0.1 fold-change difference compared to the later time period when *AtMYB93* expression peaks (15 h vs 12 h) (1.3 and 1.2 respectively). The situation was the same for *AtCYP86A2* and this gene did not change in hormone treatment and stress expression profiling. Very little experimental data were available for *AtCYP705A1* too, and where it existed it was inconsistent. Notably, both *AtCYP86A4* and *AtCYP86A2* have been found to have roles in cutin, suberin and wax biosynthesis (Kannangara *et al.*, 2007; Xiao *et al.*, 2004), though unfortunately, the phylogeny of these genes shows them to be fairly divergent (Nelson & Werk-Reichhart, 2011). Consequently, although it is interesting, given their roles, that *AtCYP86A4* and *AtCYP86A2* were initially considered as potential targets

of *AtMYB93*, the relevance of their function has only limited bearing on the uncharacterised role of *AtCYP708A1*.

3.6.3.3 AtMYB54: a transcription factor involved in cell wall biosynthesis

AtMYB54 encodes a TF involved in secondary cell wall biosynthesis (Zhong *et al.*, 2008; Zhong & Lee, 2012). The secondary cell wall is a suberin-containing structure immediately adjacent to the plasma membrane primary cell wall (Anderson *et al.*, 2015), though in this case *AtMYB54* has been implicated in cellulose biosynthesis (Zhong *et al.*, 2008). A recent paper has begun to elucidate a gene network containing *AtMYB54* and identified *AtMYB54* as among the direct targets of *At*GCN5 (Zheng *et al.*, 2019). *AtGCN5* encodes a histone acetyltransferase that is necessary for cell wall integrity and tolerance to salt stress. Bioinformatic analysis in this work found previous evidence of *AtMYB54* upregulation in response to salt stress, a result also confirmed in this study. Zheng *et al.* (2019) suggest that *At*GCN5-mediated activation of *AtMYB54* occurs under salt stress and furthermore, that this process is conserved in wheat.

TFs like MYBs can often be found working in hierarchical networks, that is, MYBs regulating other MYBs and interacting with each other (Wang & Dixon, 2012; Zhong & Ye, 2014) and the regulation of different biopolymers – including the cellulosic cell wall, must be coordinated in their regulation during LR emergence (section 1.2.4). Adding to this complexity, two further MYBs (*AtMYB52* and *AtMYB53*) were contemplated for selection as putative *At*MYB93 targets, though neither of these were detected by RNA-seq, in contrast to *AtMYB54,* which was identified (X. Cao, unpublished). Like *AtMYB54, AtMYB52* has been implicated in

secondary cell wall biosynthesis, but this time as a negative regulator of pectin demethylesterification in the seed coat mucilage (Shi *et al.*, 2018). *AtMYB53* is thought to have a role in suberin biosynthesis (Hu, 2018) and the experimental work completed on this gene is presented in the next chapter.

3.6.3.4 AtPDLP4: an endodermal membrane protein

AtPDLP4 encodes a plasmodesmata-located protein and is the second candidate gene to be both found to match the expression profile of *AtMYB93* by qRT-PCR and be present in the top 1000 genes from the RNA-seq analysis. However, there are no published studies available for this gene and no other *PDLP*'s were located in the longlist of candidate genes. Plasmodesmata link plant cells via cell wall channels that allow the passage of molecules between cells, and characterisation of other *AtPDLP* family members, of which eight have been discovered, suggest they may affect plasmodesmata permeability (Sager & Lee, 2018). Cell-to-cell communication and mobile signals are important for LR development and *AtPDLP5* regulates cell-to-cell movement throughout different tissues including roots by promoting callose deposition at the plasmodesmata (Lee *et al.*, 2011). Callose is a polysaccharide that controls the movement of molecules through the plasmodesmata and so is important for symplastic connectivity between cells. Callose deposition is associated with LR patterning (Benitez-Alfonso *et al*., 2013) and it is possible that other *PDLP* genes are similarly involved in its accumulation.

3.6.3.5 AtAHA4: a plasma membrane ATPase

AtAHA4 is part of a group of plasma membrane ATPases that contain a large number of sites for phosphorylation – a process often involved in regulation (Rudashevskaya *et al.*, 2012). Specific investigation of *AtAHA4* found only a subtle phenotype in *Ataha4* mutants of slightly reduced shoot and root growth compared to WT plants (Vitart *et al.*, 2001). Yet an obvious phenotype could be observed under salt stress as *Ataha4* mutant leaves acquired significantly more salt than WT (Vitart *et al.*, 2001). The link between candidate genes such as *AtAHA4* and *AtMYB54* with salt tolerance further makes the case to scrutinise the role of *AtMYB93* and its putative targets under salt stress.

3.6.3.6 At4G17480: a possible fatty acid or suberin metabolism gene

Some of the candidate gene mutants examined in this study are almost entirely uncharacterised and *At4G17480* was one such candidate also found to match the expression profile of *AtMYB93* by qRT-PCR. As very little is known about uncharacterised genes, the NCBI database was utilised to search for conserved domains and provide an indication of their function (Marchler-Bauer *et al.*, 2015). *At4G17480* contains an alpha/beta-hydrolases superfamily protein domain. Proteins containing this domain encompass hydrolytic enzymes involved in a wide range of processes, nevertheless this family is not considered well characterised in plants (Mindrebo *et al.*, 2016).

3.6.4 Identities of other putative target genes

3.6.4.1 AtCASPL1B1: a Casparian strip membrane protein

*At*CASPL1B1 is a Casparian strip membrane domain-like protein that is strongly homologous with CASP proteins linked to Casparian strip metabolism (Lashbrooke *et al.*, 2016). The gene was discovered to interact with the aquaporin, *At*PIP2;1 (Bellati *et al.*, 2016), which is important for LR emergence through the endodermis (1.2.4) and in a recent study expression of *At*CASPL1B1 has been revealed in cells overlying LR primordia and specifically, it is restricted to suberised endodermal cells (Champeyroux *et al.*, 2019). An earlier study also identified *At*CASPL1B1 in RNA-seq and qRT-PCR experiments demonstrating its downregulation in *Atmyb9* and *Atmyb107* mutants (Lashbrooke *et al.*, 2016).

3.6.4.2 AtLTP8: a gene putatively involved in fatty acid or suberin metabolism

*At*LTP8 is a lipid transfer protein and its similarity to others of its type link it to a possible role in wax or cutin deposition in expanding epidermal cell walls. Its homologue in *Camellia sinensis* (tea) has a function in conferring cold tolerance by potentially increasing fatty acid desaturation, thus maintaining the fluidity of cellular membranes (Li *et al.*, 2019). This recent evidence and a lack of other data perhaps makes *AtLTP8* a less attractive candidate gene as a target by *At*MYB93, since there is no evidence of *AtMYB93* being involved in cold tolerance. Nevertheless, the tea MYB93 homologue has not been characterised and so its role in response to extreme temperatures could be of interest to evaluate.

3.6.4.3 AtGPAT6: a cutin biosynthesis gene

AtGPAT6 regulates cutin biosynthesis during sepal and petal formation (Li-Beisson *et al.*, 2009) and is also important for tapetum and pollen development (Li *et al.*, 2012). While expression of *AtGPAT6* is highly enriched in inflorescence and has clearly been shown to have a primary function in fertility, further studies could still elucidate a more minor role in the roots. Interestingly, a more recent study in tomato unearthed a role for *SlGPAT6* fruit cutin biosynthesis (Petit *et al.*, 2016). *SlGPAT6* expression was strongest in the epidermis of young, developing fruit and *Slgpat6* mutant plants had a thinner and less robust cuticle (Petit *et al.*, 2016). Not only this, but the group also noted that plants mutant in *Slgpat6* alone impacted expression of a wide range of genes throughout the exocarp (the fruit 'skin'), including pathways associated with cell wall modelling. Moreover, *SlMYB93* is also expressed in the young fruit of tomato (X. Cao, unpublished).

3.6.4.4 AtQRT1: a putative cell wall remodelling enzyme

*At*QRT1 is another pectin methylesterase implicated with a similar role to *At*PME10 (section 3.6.3.1), though this has been only confirmed in pollen. Previous characterisations of *Atqrt1* mutant plants have uncovered that they appear to be phenotypically WT aside from the failure of tetraspore separation in pollen (Francis *et al.*, 2006). This is despite the fact that *AtQRT1* has been shown to be expressed in a variety of tissue types, including roots, hence indicating functional redundancy with other pectin methylesterases.

3.6.4.5 AtCYS5: a cystatin proteinase inhibitor

AtCYS5 is a phytocystatin and these genes encode proteinaceous inhibitors that are implicated in responses to a range of biotic and abiotic stresses and are considered to represent a promising avenue for crop improvement (Kunert *et al.*, 2015). A recent paper found that plants overexpressing *AtCYS5* had increased insensitivity to ABA (Song *et al.*, 2017) and as part of the study the group looked at the primary root length of plants exposed independently to a 50°C heat stress and ABA treatment, but did not look at LRs. They found that in both cases length of the primary root in *AtCYS5* over-expressor mutants was significantly less reduced under stress or hormone application than in WT plants, and that the opposite was true in *Atcys5* mutants (Song *et al.*, 2017). The bioinformatic analysis in this work identified the upregulation of *AtCYS5* after treatment with ABA, which is in line with the results of this study. However, previous analysis of the *Atmyb93* mutant failed to see a large change in response to ABA compared to WT and it is difficult to draw conclusions from this study since it looked at different aspects of the roots.

3.6.4.6 AtABCG2: a suberin biosynthesis gene

The ATP-binding cassette (ABC) transporter superfamily of proteins are membrane associated proteins that are responsible for transporting molecules, including lipids, across cellular membranes (Linton, 2007). *AtABCG2* encodes a protein belonging to the G subfamily of ABC transporters and has been implicated in having an important role in the synthesis of a suberin barrier layer in both roots and in the seed coat (Yadav *et al.*, 2014). In this role, *AtABCG2* has been shown to act redundantly alongside two other ABCG genes; *AtABCG6* and *AtABCG20*, though it

has been suggested that none of these genes are directly involved in synthesising suberin and instead could have a role in exporting suberin from the cell (Yadav *et al.*, 2014). *Atabcg2-1* single mutants did not have an obvious phenotype, but when characterising plants also deficient in *Atabcg6-1* and *Atabcg20-1*, these triple mutants were found to have greater amounts of suberin in the root and fewer LRs than WT. This LR phenotype is comparable to the phenotype anticipated from our hypothesis: that plants with less suberin would have increased numbers of LRs than WT plants due to the expected disruption of the suberin barrier layer and therefore a more 'leaky' or 'stretchy' endodermis.

Several other ABCG genes are discussed in the study, three of which were previously considered as potential target genes of *At*MYB93 after the initial longlist from data by Voß *et al.* (2015) was produced. *AtABCG6*, one of the genes shown to be functionally redundant with *AtABCG2* (Yadav *et al.*, 2014), was initially considered as a *At*MYB93 target. Ultimately, this gene was not selected because although its stress and hormone expression patterns were similar to that of *AtMYB93*, its peak expression occurs in the earlier time period of (12 h vs 9 h), rather than when *AtMYB93* peaks during the 15 h vs 12 h measurement.

AtABCG10 and *AtABCG23*, are two closely related genes that are co-expressed alongside suberin biosynthesis genes and could have a role in barrier function (Yadav *et al.*, 2014). *AtABCG23* was not chosen since its peak expression was before that of *AtMYB93*. *AtABCG10* on the other hand, was found to peak in its expression at the same time as *AtMYB93*, in addition to also being downregulated at the same time point (48 h vs 45 h), though this downregulation was not significant. As previously described, the Voß *et al.* (2015) microarray looked at gene expression

in the cells surrounding the site where the primary root was bent. The significant increase in fold-change expression in *AtABCG10* was 1.0 at the 15 h vs 12 h time point, up from 0.9 at 12 h vs 9 h. Given the specificity of the tissue examined, we might expect a somewhat larger difference in expression if the gene was a target of *AtMYB93* and so this gene was not chosen for the final shortlist.

3.6.4.7 Uncharacterised genes

AtS-ACP-DES1 a root-enriched, largely uncharacterised gene encoding a glycine-rich protein with a predicted plastid function. *At5G49350*, *At3G27270* and *At2G43670* are also entirely uncharacterised, though based on the homology of *At5G49350* with other genes it is predicted to be a glycine-rich endodermal membrane protein. *At*3G27270 is a TLC domain containing protein and animal studies have suggested they could have a role in lipid sensing, though this has not been explored in plants (Winter & Ponting, 2002). The sequence of *At2G43670* suggests that it may belong to the X8 protein superfamily. The X8 domain is thought to have a role in carbohydrate binding and proteins in this family may provide structural support to the plasmodesmata (Simpson *et al.*, 2009).

3.6.5 AtSCARECROW is an upstream repressor of AtMYB93

The trend observed from the qRT-PCR experiment of increased *AtMYB93* expression in the *Atscr-3* mutant relative to WT validates the hypothesis of *AtSCR* repressing *AtMYB93* expression and further corroborates previous evidence from Iyer-Pascuzzi *et al.* (2016) and Sparks *et al.* (2016). Nevertheless, it was important to use the less severe *Atscr-3* mutant which still differentiates a layer resembling the

endodermis. This is because any upregulation of *AtMYB93* expression in a mutant completely lacking an endodermis could be the result of ectopic expression since the plant lacks the proper tissue for *AtMYB93* to localise to, or alternatively, with no endodermis the plant may not express *AtMYB93* at all.

An important next step is to confirm the localisation of *AtSCR* expression *in vitro*. This was attempted using confocal microscopy with the *pSCR*::*SCR*-GFP;*Atscr-3* (B. Hutton & N. McMulkin, unpublished), however only limited evidence of the disappearance of GFP in cells overlying LR primordia was observed, necessitating further study. One option that was considered was to create a *pMYB93*::GFP or *pMYB93*::GUS line in the *Atscr*-3 mutant background, however as *AtSCR* has a role so early on in the development of the endodermal layer, resulting primary- and LR development could be affected and so the source of any impacts on *At*MYB93 would be unclear. Preliminary investigation of the LR density of the *pSCR*::*SCR*-GFP;*Atscr-3* line has also started where it was originally hypothesised that the *pSCR*::*SCR*-GFP;*Atscr-3* line would have a LR phenotype close to WT. Yet the initial data suggests a phenotype of increased LR density similar to the *Atmyb93* (N. McMulkin, unpublished). This data is line with the expression patterns seen from qRT-PCR, but more data, as suggested in the final discussion, is needed before any conclusions can be drawn.

3.6.6 Challenges and limitations

3.6.6.1 Limitations of assessing expression patterns

The Voß *et al*. (2015) time course dataset of the *Arabidopsis* transcriptome provided the best platform for beginning the selection of putative downstream targets

of *At*MYB93 because it tested expression in a very specific part of the root, where *AtMYB93* is expressed. A recent study by Trinh *et al.* (2019) outlines the use of a similar approach for the bioinformatical analysis of gene targets to the one presented in this thesis. Utilising the same dataset, the group used an algorithm to interrogate the data for genes with similar expression patterns to their gene of interest – *AtPUCHI*. A key difference in this method was in the group's preference for looking at genes with an expression pattern shifted to the later time zone than where the peak expression of *AtPUCHI* occurred. When applying the same time shift to *AtMYB93* it can be seen that *AtMYB93* is expressed at this time point, though this is no longer significantly different, whereas the change in expression during the 15 h vs 12 h period where *AtMYB93* peaks is significantly different. Nevertheless, it is possible that *AtMYB93* expression could be less detectable at earlier or later time points than those examined but could still be having an effect on genes expressed sooner after LR initiation. However, there were not many of these genes with peak expression in the later time period and most look less convincing after comparing their expression profiles to that of *AtMYB93*.

3.6.6.2 Limitations of promoter analysis

A notable limitation of computational promoter analysis is the rate of false positives that are produced (Zia & Moses, 2012). In many statistical analyses it is usual to use a correction factor to make adjustments to the p-value to account for the incidence of false positive results in terms of their significance. This is due to the fact that the probability of incorrectly obtaining at least one statistically significant result, where this outcome was actually down to chance, increases as the number of

hypotheses being tested increases (Zia & Moses, 2012). This is known as a type one error. The Bonferroni correction is one such method that is commonly to adjust the p-value in cases of multiple hypothesis testing. However, for large scale analyses such as promoter analysis, q-values are a more appropriate p-value adjustment than the Bonferroni correction (Noble, 2009). This is because the q-value controls the proportion of false positives for data deemed significant, whereas the Bonferroni correction controls for the probability of at least one false positive in a set of data and can actually lead to type two errors, where an important effect is incorrectly considered to be the result of chance (Noble, 2009). Thus, analyses based on very large datasets such as this should be treated with caution and could reduce the reliability of predicted binding motifs identified in the promoter regions being tested.

3.6.6.3 Experimental limitations of using whole seedling tissue

There are a number of limitations of using whole seedling tissue to detect gene expression that is occurring in small regions of root tissue. *At*MYB93 is expressed is restricted to just a few endodermal root cells and consequently, any changes in expression of both *At*MYB93 and its up- and downstream target genes will be more difficult to detect in RNA extracted from whole seedlings, where the signal to noise ratio is therefore low. Ultimately, however, the qRT-PCR experiments were carried out using this method because of the large number of roots required to get a mass of tissue that is comparable to that of whole seedlings – specifically, around eight times as many seven-day-old roots are needed as whole seedlings of the same age. The use of whole seedlings will therefore have increased the amount of background signal noise, reducing the ability to see differences in the data. A

future repeat of this experiment could utilise FACS, which would allow even more specificity with the use of just endodermal tissue.

The small increases recorded in *At*MYB93 expression in over-expressing plants relative to WT were considered to be a consequence of attempts made by the plant to shut down the transgene. This can become an issue after several generations are produced from over-expressing lines utilising a 35S promoter, particularly if the effects of the overexpressed gene are detrimental to the plant. This is likely to be the case with *At*MYB93 over-expressor mutants since in earlier generations some plants produced a white waxy substance that coats the leaves (E. Chapman, unpublished).

3.6.6.4 Root tissue mass yield optimisation

As growing *Arabidopsis* using the current protocol to harvest whole seedlings is not ideal for roots, two further methods were attempted to see if more root tissue could be more easily obtained. The first method was to grow fewer seedlings on each plate so that they could be grown for longer and therefore develop larger roots. *Arabidopsis* seedlings could be easily grown for 14 days, twice the usual length of time, without becoming crowded or reaching the bottom of the plate. Alternatively, larger plates could be obtained to grow seedlings for even longer. Growing seedlings for 14 days allows enough room for one row of seeds to be put on plates, whereas two rows of seedlings can grow for seven days. We postulated that 14-day-old seedlings may have a mass greater than that of double the number of seven-day-old seedlings, however this was not shown to be the case. Furthermore, previous data has been collected using seven-day-old plants (Coates lab, unpublished; Gibbs *et al.*,

2014) and so it would be better to use the same conditions for more comparable data.

The second technique involved germinating and growing entire seedlings in liquid Gamborg's medium kept on a rotator. Similar methods have been shown to enhance the amount of root tissue that grows (Benzle & Cornish, 2017; Hétu *et al.*, 2005) and it was easy to cultivate large numbers of seedlings. However, growing *Arabidopsis* in this way resulted in a tangled clump of seedlings that did not allow root tissue to be easily separated from the rest of the seedling. The time taken to do this could have had impacts on gene expression and so whole seedlings still needed to be harvested using this method and as a result no measurements confirming enhanced root growth could be made. Samples grown in this way in liquid culture were compared experimentally against samples grown on plates. qRT-PCR was used to test the expression of a selection of AtMYB93 target genes in *Atmyb93-1* and WT tissue (data not shown). Seedling growth in liquid culture appeared to have a different effect on gene expression than in seedlings grown on plates, potentially because the seedlings may have been under stress, and so ultimately this method was not taken any further.

A third option was proposed, which would have involved the use of a hydroponic growth system. This method would allow older plants to be grown without the issue of soil or agar sticking to the root tissue. Options for this included a system that was available for purchase (Araponics, Liège, Belgium), or a DIY setup. Cost was considered a prohibitive factor for using a purchased kit and in addition to this, the kit has been developed for the growth of mature plants and so large number of kits would be required to obtain enough tissue from seedlings. After looking into

creating a DIY hydroponic system this was also decided against because growing plants hydroponically has the potential to create more hypoxic conditions than when using plates and as *AtMYB93* is regulated by hypoxia (Mustroph *et al.*, 2009) this could be an issue.

3.6.6.5 Technical issues with qRT-PCR

Varying amounts of primer-dimer was observed in the amplification curves of some genes; however, this was largely unavoidable due to the expense and time required to fully optimise qRT-PCR conditions. These constraints mean that two different primer concentrations were tried (200 nM and 400 nM) with both the 5' and 3' primers at the same concentration as each other. Ideally a concentration matrix is carried out, where each concentration of the 5' primer is tested against each concentration of the 3' primer. The presence of primer-dimer can reduce the sensitivity and efficiency of the reaction by contributing to the fluorescence signal and taking away components from the desired reaction. Ultimately, limited optimisation was considered the best approach since a relatively large number of primers were being used in these experiments.

3.6.6.6 Yeast-one-hybrid considerations

Y1H relies on a robust interaction between protein and DNA and so to increase the chance of the protein binding successfully, tandem repeats of the upstream region spanning the promoter of the genes of interest would ideally be produced. However, there is a trade-off and, in this experiment, the resultant size of a construct with tandem repeats would have been much more difficult to clone.

Although *At*MYB93 binding motifs had been putatively identified bioinformatically, these were only predicted and so the actual location of any potential binding sites remained unknown. Hence, large regions upstream of the genes of up to 2 kb were used to try and capture any *At*MYB93 binding domains. Moreover, the experiment involved working with a large number of promoters, whereas it is more typical in Y1H to assess one promoter against various putative upstream proteins.

The critical issue faced during this experiment was that I was unable to transform *AtMYB93* into the Y187 strain of yeast in order for binding assays to be conducted. The choice of vector was not found the be the issue since this could be modified, while the yeast strain was also not the likely problem as a *At*MYB93 construct has previously been successfully transformed into the AH109 yeast strain (Gibbs *et al.*, 2014). For future work on this research line, a new Y1H system is being trialled (J.C. Coates, personal communication).

3.6.6.7 Glasshouse issues

Glasshouse conditions for light and temperature were not well maintained for a large part of the duration of this project. Though *Arabidopsis* has a wide growth range, controlled conditions are important for experimental reproducibility and any results obtained from plants grown under different conditions could be unreliable. Dipping plants to infect them with *A. tumefaciens* results in a minor stress response in *Arabidopsis* and so the addition of other stresses such as high temperatures would inevitably stress plants further. It is thought that this may have contributed to the poor transformation efficiency of the dipping events conducted during this project. Moreover, the glasshouse reached temperatures of up to 65°C, which can affect

Arabidopsis fertility (Rieu *et al.*, 2017) and consequently the facility was out of use for the remainder of the project, with very limited growing space available elsewhere.

3.7 Conclusion

The work in this chapter has identified *At*SCR as an upstream repressor of *AtMYB93* and a series of genes that *At*MYB93 itself could be regulating. The aim of the next chapter is to characterise these potential downstream targets of *AtMYB93* and to look more closely at the roles of close relatives *AtMYB53* and *AtMYB92*.

Chapter IV

CHARACTERISATION OF MUTANTS RELATED TO *At*MYB93 SIGNALLING

4.1 Introduction

Suberin is an insoluble lipid biopolymer that creates a highly impermeable barrier layer in the endodermal cell wall (Beisson *et al.*, 2012; Bernards, 2002). A cascade of genes that include a number of MYBs comprise the gene networks involved in the construction of biopolymers such as cutin and lignin (section 1.5.1) and though suberin biosynthesis is less well studied, a series of closely related MYBs have been implicated in regulating its synthesis in roots and seeds. These include the *AtMYB93*, *AtMYB92* and *AtMYB53* clade that have been suggested to have roles in regulating root suberin (Hu; 2018) and are the focus of this chapter. Additionally, *AtMYB41*, *AtMYB107* and *AtMYB9* – MYBs located in clades neighbouring *AtMYB93*, have been found to regulate suberin biosynthesis in the seed coat (Gou *et al.*, 2017; Lashbrooke *et al.*, 2016; Kosma *et al.*, 2014). *AtMYB53* and *AtMYB92* are thought to be *Brassicaceae*-specific innovations, whereas *AtMYB93* is established as the ancestral member of the S24 clade (Gibbs *et al.*, 2014). There are examples of *AtMYB93* in other species and this is something which is explored in chapter V. Furthermore, as examined in the previous chapter, a series of genes identified as potential downstream targets of *At*MYB93 includes several with confirmed or putative roles in fatty acid or suberin biosynthesis.

The *Atmyb93* mutant phenotype presents as having more rapid LR development and an increase in LR density in seedlings (Gibbs *et al*., 2014). The closest relative of *AtMYB93*, *AtMYB92*, is still predominately expressed in the root but a *Atmyb92* mutant does not have a LR phenotype. However, in a *Atmyb92-1/Atmyb93-1* double mutant, *Atmyb92* enhances the *Atmyb93* mutant phenotype. As more suberin-related MYBs are knocked out, such as in the *Atmyb53-1/Atmyb92-*

1/Atmyb93-1 and Atmyb53-2/Atmyb92-1/Atmyb93-1 triple mutants, a redundancy effect becomes more likely. Functional redundancy is where a gene is able to compensate for the loss of another closely related gene with an overlapping expression domain and function to ensure the maintenance of an essential step such as in signalling networks.

Sulfur deprivation and salinity pose critical challenges for agriculture and food security, and *AtMYB93* is upregulated in response to both stresses (Bielecka *et al.*, 2015; Dinneny *et al.*, 2008). Initial data collected from elemental analysis of *Arabidopsis* shoot tissue grown under sulfur deficient conditions suggests that under normal conditions *Atmyb93-1* (Gibbs) mutants can increase their sulfur uptake, though when sulfur starved, they may lose more sulfur than WT plants (Wilkinson, 2018). Likewise, under salt stress *Atmyb93-1* (Gibbs) mutants appear to take up more sodium than WT, but this difference is no longer apparent when unstressed (Wilkinson, 2018).

The first aim of this chapter was to determine if there were any mutants available for the putative *At*MYB93 targets and establish if any of these mutants had relevant LR phenotypes. The second aim was to characterise LR density and germination rates under varying physiological conditions in *Atmyb93-1* and also in mutants additionally deficient in the MYB family members *Atmyb53-1/-2* and *Atmyb92-1*. Presented in the following pages of this chapter is a phenotypic characterisation of a number of the candidate gene target mutants of *At*MYB93 and of the MYB single, double and triple mutants.

4.2 Genotyping of AtMYB93 putative target gene mutants

The T-DNA express database was interrogated for transfer DNA (T-DNA) insertions present in any of the candidate target genes. Overall, 13 of the 16 candidates had promising mutant lines available (Figure 1.A in Appendix IV), while no mutants considered suitable had yet been isolated for the remaining targets (*At2G18370* (*AtLTP8*), *At2G19150* (*AtPME10*) and *At5G49350*). Mutants were chosen from the database after considering the following features. The position of the T-DNA insert was the most important factor since if it is located towards the 3' end of the gene then the resultant mutant may not be a complete knock out. SAIL lines were considered more desirable than SALK because of the known issue of potential kanamycin resistance silencing in SALK lines. Other lines such as GABI-Kat were not selected due to the added complication of obtaining homozygous lines whilst working with a relatively large number of mutants.

Based on data from T-DNA express, nine of the chosen mutant lines were reported to be homozygous, with the remaining four annotated as heterozygous. Therefore initially, the mutants were genotyped to both confirm the putative homozygous mutants and to obtain homozygous mutant plants from the heterozygous lines. Two combinations of primers were needed to distinguish between WT, homozygous mutant, and heterozygous lines. The first set amplified the left flanking sequence of the gene using a "left primer" (LP) and the right border primer of the T-DNA insert with a "border primer" (BP). The second combination amplified the left and right flanking sequences of the uninterrupted gene, once again using a LP and also a "right primer" (RP) (Figure 1.B in Appendix IV).

Initial attempts to genotype the first of the mutants (grown from bulked seed), indicated that the genotypes of some lines may have been incorrectly characterised. The first purported homozygous mutant lines tested were *At1g73410* (*Atmyb54*), *At2g43670* and *At2g38110* (*Atgpat6*), and based on their genotyping PCR results *Atmyb54* initially appeared to be WT, *At2g43670* presented as heterozygous and only *Atgpat6* was correctly identified as a homozygous mutant. The four heterozygous lines (*At3G04370* (*AtPDLP4*), *At1G55940* (*AtCYP708A1*), *At2G37360* (*AtABCG2*) and *At5G16240* (*AtS-ACP-DES1*)) were also genotyped and while for *Atabcg2* and *Ats-acp-des1* homozygous mutant plants were successfully obtained, all 16 plants that were genotyped for each of *AtPDLP4* and *AtCYP708A1* were WT. It was therefore decided that the next step should be to germinate seeds from each line under selection, as it was now unknown how many of the remaining lines were homozygous mutants. Additionally, it would be easier to obtain homozygous mutants for *AtPDLP4* and *AtCYP708A1* by first screening out WT lines and seedlings.

Table 4.1 shows the expected genotype of each mutant line according to T-DNA express, the percentage that survived under selection and the genotype that this result suggests, and finally the result of genotyping varying numbers of plants. The four mutants denoted as heterozygous and the *At5g49350* homozygous mutant were from the SAIL collection and and so were selected for with BASTA®, while the remaining claimed homozygous mutants were SALK lines and so contain a kanamycin resistance gene to select for the presence of T-DNA inserts. Nearly all of the seedlings for the two heterozygous lines (*Atpdlp4* and *Atcyp708a1*) survived selection, thus implying that they were homozygous mutants. This contradicted both the T-DNA express database and the genotyping results and these lines were not

Table 4.1. Expected vs actual genotyping phenotypes of T-DNA mutants. *The genotype of each candidate gene mutant as indicated by the T-DNA express database, after selection and after PCR genotyping. N/A denotes line not tested.*

Mutant gene name	Expected genotype according to T-DNA express	Survival under selection	Inferred genotype from selection	Genotyping result
PDLP4	Heterozygous	~100% survived BASTA®	Homozygous	WT
ABCG2	Heterozygous	N/A	N/A	Heterozygous
CYP708A1	Heterozygous	~100% survived BASTA®	Homozygous	WT
S-ACP-DES1	Heterozygous	N/A	N/A	Heterozygous
CASPL1B1	Homozygous	~75% Survived BASTA®	Heterozygous	Heterozygous
AT2G43670	Homozygous	~75% survived kanamycin	Heterozygous	Heterozygous
AHA4	Homozygous	~50% survived kanamycin	Likely heterozygous	Homozygous?
QRT1	Homozygous	~75% survived kanamycin	Heterozygous	Heterozygous
MYB54	Homozygous	~50% survived kanamycin	Likely heterozygous	WT?
CYS5	Homozygous	None survived kanamycin	WT	Not done
AT3G27270	Homozygous	~75% survived kanamycin	Heterozygous	Homozygous?
AT4G17480	Homozygous	~50% survived kanamycin	Likely heterozygous	Homozygous?
GPAT6	Homozygous	N/A	N/A	Homozygous

used in further experiments. All but one of the supposed homozygous mutants actually appeared to be heterozygous when put onto selection medium, as around 25-50% of the seedlings were killed. The other homozygous mutant, *Atcys5*, appeared to be WT as none of the seedlings survived selection.

Of the seven mutants with healthy seedlings that were genotyped after selection, PCR revealed at least one homozygous plant for five of the SALK lines (*At2g43670*, *Ataha4*, *At3g27270*, *Atgpat6* and *At4g17480*) and the *Atcaspl1b1* SAIL line. Heterozygous plants were obtained for *Atqrt1*, while only WT plants were confirmed for *AtMYB54*, both also SALK lines. Figure 4.1 shows the PCR genotyping result of the candidate gene mutants and for those with homozygous mutants, while an RT-PCR confirmed that the transcript could not be formed (data not shown).

4.3 Phenotypic characterisation of AtMYB93 putative target gene mutants

4.3.1 Lateral root density phenotypes of AtMYB93 putative target gene mutants

Many of the candidate gene mutants are more widely expressed than *AtMYB93* (section 3.2.3) (Winter *et al.*, 2007), yet no obvious above ground phenotypes were observed when growing mature plants of the mutant lines. In order to examine any predicted root phenotypes, LR density assays were planned to be conducted on the candidate gene mutants. Initially the aim was to characterise any potential phenotypes of the six genes identified by qRT-PCR (section 3.3.1), for which pre-existing mutants were available: *AtPDLP4*, *AtAHA4*, *AtCYP708A1*, *At4G17480*, however after the issues with genotyping it was decided to begin characterising the mutants where enough homozygous mutant seed had already been obtained. These were the *Atabcg2* and *Ats-acp-des1* lines. Figures 4.2-4.4

Figure 4.1. Genotyping PCR results of candidate gene mutants. *Col-0 sample order = LP/RP positive control then LB/RP negative control. Primers are always designed to amplify from the T-DNA left border and the LB/RP product is up to 300 bp larger than LP/RP (as described on SIGnAL). White stars denote poorly amplified samples. (A) At least one homozygous or heterozygous mutant was obtained (lanes confirming homozygous mutant individuals marked by black stars). (B) Only WT plants were obtained. (Completed with Helen Wilkinson).*

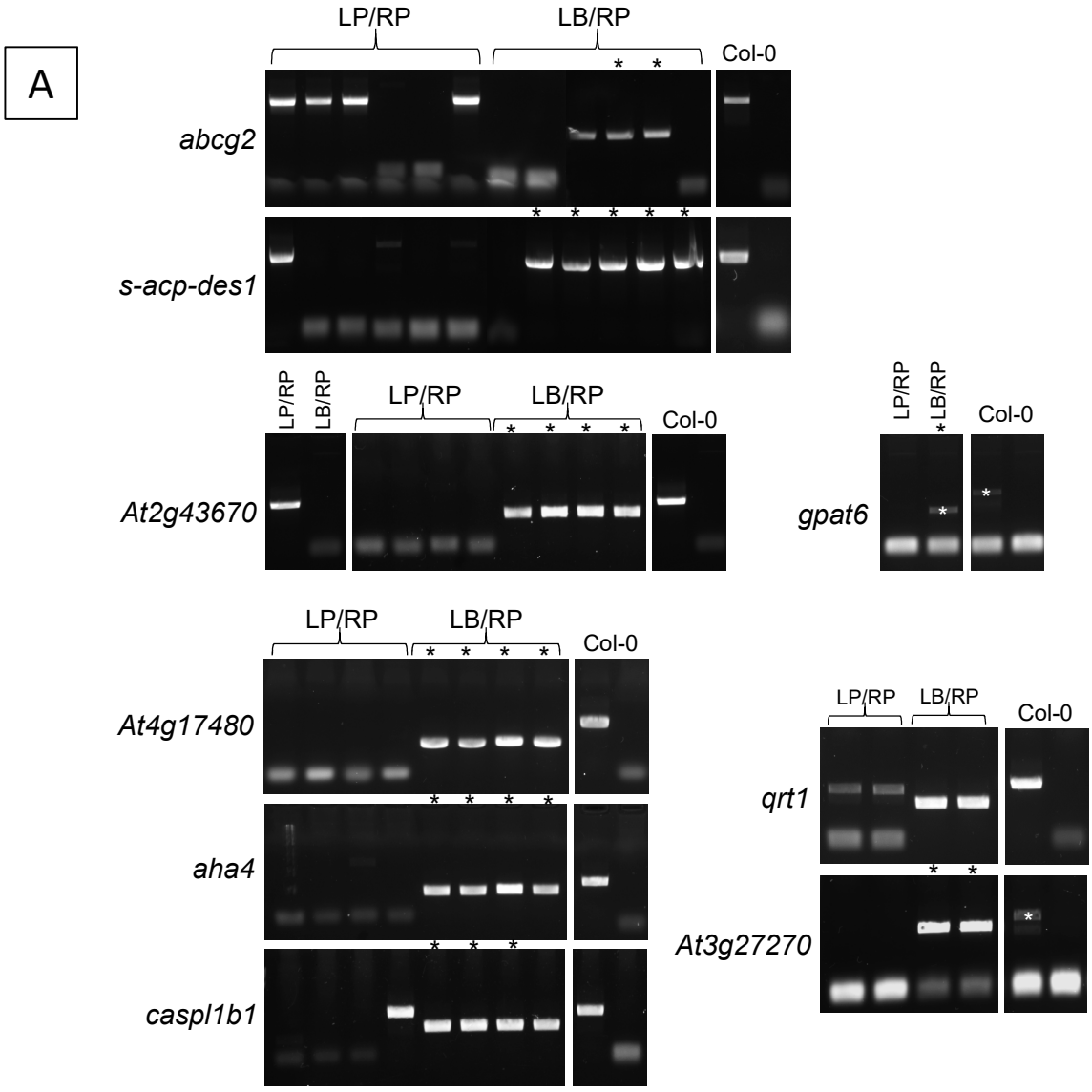
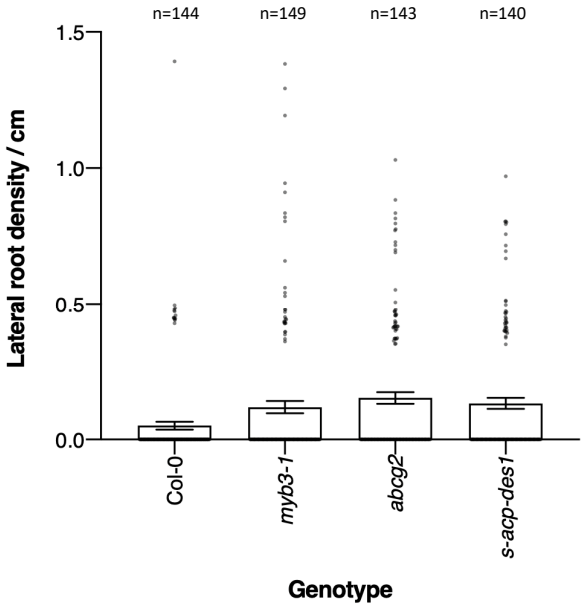

A
LP/RP
LB/RP
Col-0
abcg2
s-acp-des1
At2g43670
gpat6
At4g17480
aha4
caspl1b1
qrt1
At3g27270
B
cyp708a1
pdlp4
myb54

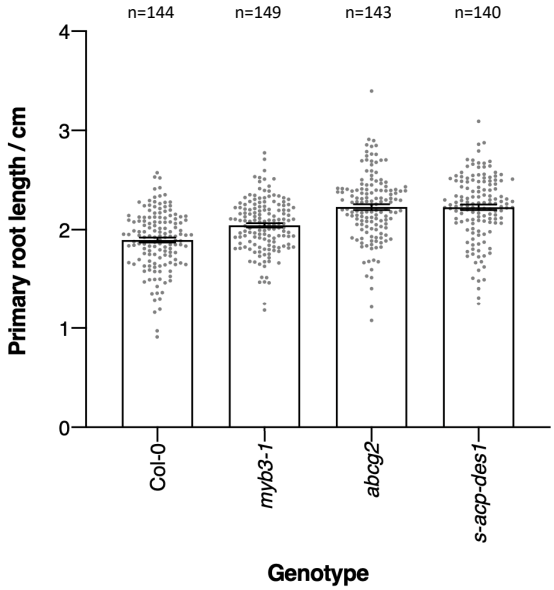
Figure 4.2. Primary and lateral root phenotypes of putative AtMYB93 targets of 7-day-old seedlings. *(A) Lateral root density, (B) primary root length. Lateral roots were counted by eye. Primary root length was calculated manually by measuring photographs in ImageJ. Lateral root density was calculated as follows: number of lateral roots / length of primary root. (Performed by Helen Wilkinson under supervision).*

A



Lateral root density p-values	*myb93* vs Col-0	*abcg2* vs Col-0	*s-acp-des1* vs Col-0
Kruskal Wallis with Dunn's test	0.113	**0.003 *****	**0.0048 ****

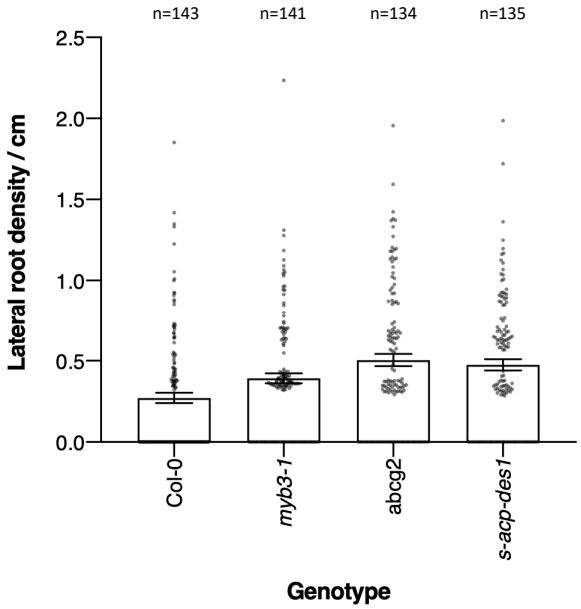
B



Primary root length p-values	*myb93* vs Col-0	*abcg2* vs Col-0	*s-acp-des1* vs Col-0
Kruskal Wallis with Dunn's test	**0.0011 ****	**<0.0001 *****	**<0.0001 *****

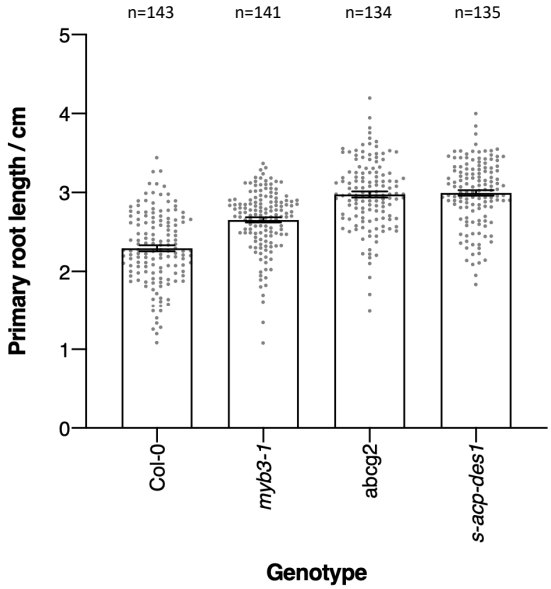
Figure 4.3. Primary and lateral root phenotypes of putative AtMYB93 targets of 8-day-old seedlings. *(A) Lateral root density, (B) primary root length. Lateral roots were counted by eye. Primary root length was calculated manually by measuring photographs in ImageJ. Lateral root density was calculated as follows: number of lateral roots / length of primary root. (Performed by Helen Wilkinson under supervision).*

A



Lateral root density p-values	*myb93* vs Col-0	*abcg2* vs Col-0	*s-acp-des1* vs Col-0
Kruskal Wallis with Dunn's test	**0.0081 ****	**<0.0001 *****	**<0.0001 *****

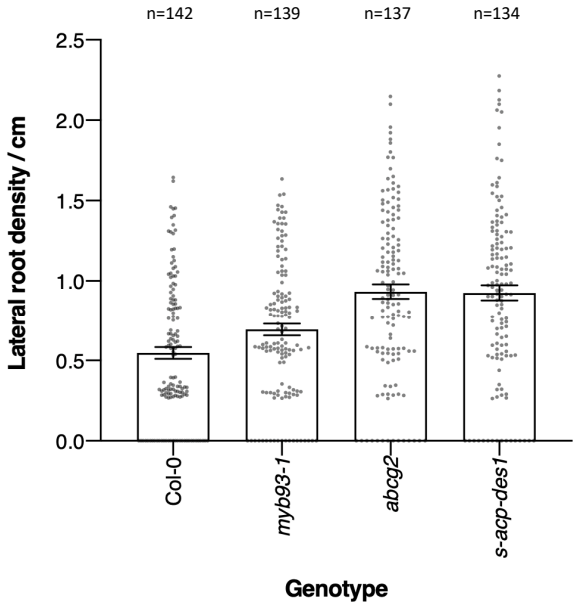
B



Primary root length p-values	*myb93* vs Col-0	*abcg2* vs Col-0	*s-acp-des1* vs Col-0
Kruskal Wallis with Dunn's test	**<0.0001 *****	**<0.0001 *****	**<0.0001 *****

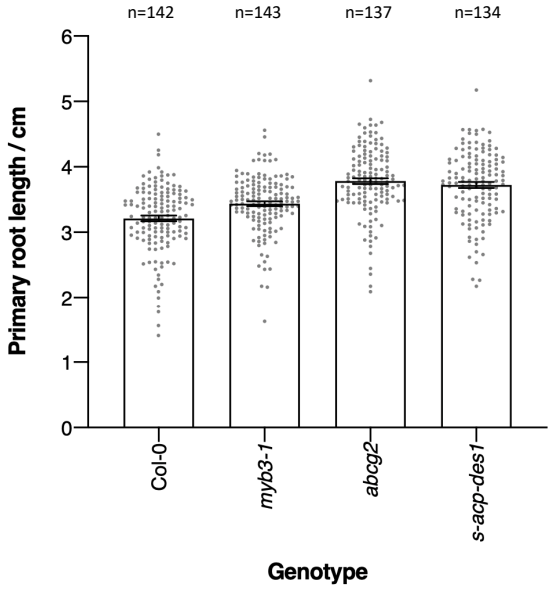
Figure 4.4. Primary and lateral root phenotypes of putative AtMYB93 targets of 9-day-old seedlings. *(A) Lateral root density, (B) primary root length. Lateral roots were counted by eye. Primary root length was calculated manually by measuring photographs in ImageJ. Lateral root density was calculated as follows: number of lateral roots / length of primary root. (Performed by Helen Wilkinson under supervision).*

A



Lateral root density p-values	*myb93* vs Col-0	*abcg2* vs Col-0	*s-acp-des1* vs Col-0
Kruskal Wallis with Dunn's test	**0.048 ***	**<0.0001 *****	**<0.0001 *****

B



Primary root length p-values	*myb93* vs Col-0	*abcg2* vs Col-0	*s-acp-des1* vs Col-0
Kruskal Wallis with Dunn's test	**<0.0001 *****	**<0.0001 *****	**<0.0001 *****

show the combined biological replicates of three LR density assays conducted on seven-, eight- and nine-day-old *Arabidopsis* seedlings (performed under supervision by Helen Wilkinson). The *Atmyb93-1* (Gibbs) mutant was used as a positive control and its phenotype of increased LR density was observed to be significant on both eight- and nine-day-old seedlings compared to WT ($p < 0.0001$ for both), though not on seven-day-old seedlings. Likewise, *Atabcg2* and *Ats-acp-des1* both had higher LR densities than WT plants and this result was highly significant across all three days ($p < 0.0001$ for all). Finally, the primary root lengths of the three mutants were significantly longer than WT ($p < 0.001$ for all comparisons with WT except for *Atmyb93-1* (Gibbs) on seven-day-old seedlings), with a trend of even longer primary roots in both candidate gene mutants when compared with *Atmyb93-1* (Gibbs).

4.3.2 Elemental analysis of AtMYB93 putative target gene mutants

ICP-MS is a type of mass spectrometry that can be used to accurately quantify the presence of a range of elemental ions in a sample (Thomas *et al.*, 2016). Root architecture, including LR density and endodermal barrier properties affect root nutrient and ion uptake, which ultimately impacts shoot ion content and so mature *Arabidopsis* shoot tissue was analysed to assess any differences in the shoot tissue of the candidate gene targets. Figure 4.5 shows heatmaps representing the relative differences in ion content in *Atmyb93* (Gibbs) and mutants for four of the putative *At*MYB93 target genes compared to WT. The candidate gene mutants tested were: *Atabcg2*, *Atgpat6*, *Ataha4* and *Ats-acp-des1*. Sulfur and sodium content were both measured to complement the potential phenotypes observed by Wilkinson (2018), which are outlined in the introduction of this chapter. While it is possible that there is

Figure 4.5. Comparison of the sulfur, sodium and magnesium ion content of Atmyb93 and mutant candidate gene targets of AtMYB93 relative to WT in 25-day-old Arabidopsis shoot tissue. *Data from two biological repeats. Values are presented as a ratio relative to the WT for each biological repeat in order to represent relative loss or gain in ion content.*

S

	RPT 1	RPT 2
Col-0	100.00	100.00
myb93	101.13	108.18
abcg2	107.62	104.77
gpat6	97.88	105.90
aha4	98.13	104.70
s-acp-des1	94.86	99.43

Na

	RPT 1	RPT 2
Col-0	100.00	100.00
myb93	111.69	99.44
abcg2	108.24	107.69
gpat6	91.32	88.84
aha4	102.81	92.27
s-acp-des1	95.40	77.71

Mg

	RPT 1	RPT 2
Col-0	100.00	100.00
myb93	96.17	100.87
abcg2	102.07	103.00
gpat6	98.26	95.19
aha4	98.89	103.05
s-acp-des1	99.86	94.21

a small increase of around 1.08-fold in the sulfur content of the *Atmyb93* mutant relative to the WT, which is in line with the result from Wilkinson (2018), this is only seen in one biological repeat. Likewise, with sodium one biological repeat matches the data from Wilkinson (2018) and shows no difference in ion content, however there is a roughly 1.1-fold increase in *Atmyb93* mutant compared to WT in the other biological repeat. Magnesium content was also assessed because there appeared to be an increase in this ion in *Atmyb93* mutant plants when examining the preliminary dataset from Wilkinson (2018), however this trend was not observed in the *Atmyb93* mutant data in this experiment.

The *Atabcg2* mutant appears to show an increase in the uptake of all three ions conserved across both biological repeats, and this is one of the mutants with a clear phenotype of increased LR density. The ion content of *Ataha4* relative to WT is, on the other hand mostly unchanged, except where roughly a 1.08-fold decrease in sodium is seen in the mutant, though only in one biological repeat. Overall, the *Ats-acp-des1* mutant appears to show a decrease in ion uptake relative to the WT. This is interesting since this is the other gene shown to have an increased LR density phenotype and is in contrast to the data for *Atabcg2.* The largest decrease for *Ats-acp-des1* is a 1.23-fold decrease in sodium content compared to WT, but once again this is only observed in one biological repeat. The trend for the uptake of the three ions varies more in *Atgpat6*. There appears to be a decrease of around 1.09 and 1.11-fold in sodium uptake in the mutant compared to in WT tissue. Yet there is little change in either the sulfur or magnesium ion content of the *Atgpat6* mutant relative to WT.

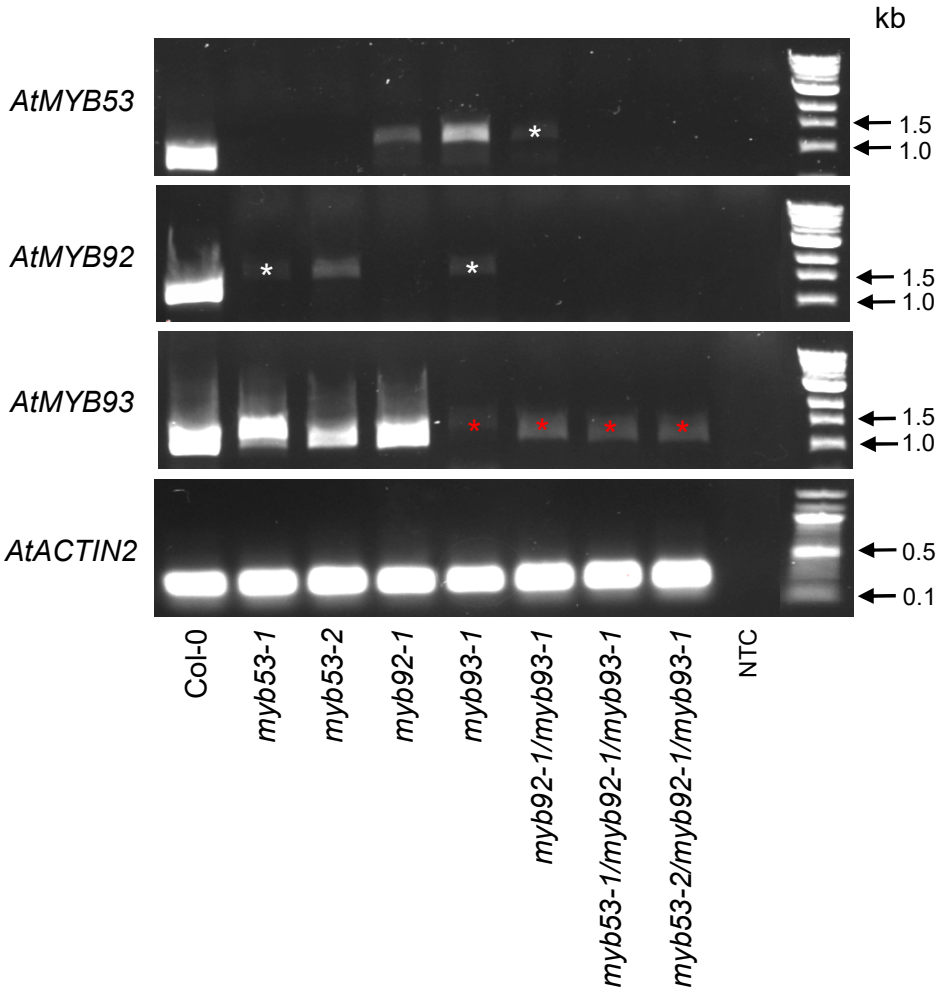
This experiment consists of two biological repeats and so additional repeats are required to allow statistical analysis and to more accurately draw conclusions about the ability of the *AtMYB93* target genes to uptake nutrients on normal plant growth medium. In addition, the remaining candidate gene mutants should be tested.

4.4 Genotyping of single, double and triple MYB mutants

In order to study *At*MYB93 and its close relatives *At*MYB92 and *At*MYB53, a series of KO lines mutant in these genes were isolated by Hu (2018). As two different allelic mutants have been created for *AtMYB53*, two triple mutants were also created to include both of these allelic variations (Hu, 2018). The first step was to ensure the relevant genes had been successfully knocked out in the seed stocks obtained from the Rowland lab, by checking for the presence of full-length transcripts by RT-PCR. Each of the single (*Atmyb53-1, Atmyb92-1* and *Atmyb93-1* (Hu)), double (*Atmyb92-1/Atmyb93-1*) and triple (*Atmyb53-1/Atmyb92-1/Atmyb93-1* and *Atmyb53-2/Atmyb92-1/Atmyb93-1*) mutants were shown to be lacking the predicted gene expression, as the full length transcript was not amplified in lines where the genes should have been knocked out (Figure 4.6). The mutant lines had also previously been genotyped by PCR (Hu, 2018).

It should be noted that amplification of the full-length transcripts of these genes proved difficult, despite redesigning some primers, and there was suspected gDNA contamination in some of the *Atmyb93-1* (Hu) samples, even after treatment with DNase.

Figure 4.6. RT-PCR results of MYB mutants. *DNA extracted from 20 pooled seedlings per line. White stars denote poorly amplified samples. Note that suspected gDNA contamination is present in Atmyb93 mutant lines, indicated by red stars.*


kb
AtMYB53
AtMYB92
AtMYB93
AtACTIN2
1.5
1.0
1.5
1.0
1.5
1.0
0.5
0.1
Col-0
myb53-1
myb53-2
myb92-1
myb93-1
myb92-1/myb93-1
myb53-1/myb92-1/myb93-1
myb53-2/myb92-1/myb93-1
NTC

4.5 Phenotypic characterisation of single, double and triple MYB mutants

4.5.1 Germination phenotypes of MYB mutants

Suberin is an important component of the seed coat (Molina *et al.*, 2006; Molina *et al.*, 2008) and both AtMYB53 and AtMYB92 are expressed in the germinating seed (Gibbs *et al.*, 2014). Additionally, *At*MYB93 expression in the root becomes apparent early on in the germinating seedlings development. Moreover, experiments have shown that mutant seedlings have reduced suberin levels with triple mutants containing just 40% of the levels seen in WT (Hu, 2018). This raised the question of whether any of these mutants exhibit a germination phenotype since it might be expected that germination would happen more readily with a reduced suberin barrier layer, as is seen in other mutants with less suberin (Beisson *et al.*, 2007; Fedi *et al.*, 2017; Yadav *et al.*, 2014). Germination assays were performed using both of the *Atmyb53* single mutants and both triple mutants against a WT control. The other mutants were not used since *Atmyb92* and *Atmyb93* have already been shown to have no germination phenotype in single mutants (Harding, Carlson & Coates, unpublished). The assumption was that the rate of germination would be affected, rather than the mutants being unable to fully complete the process.

AtMYB53, *AtMYB92* and *AtMYB93* have each been implicated in plant stress responses (Stracke *et al.*, 2001; Winter *et al.*, 2007) and so seedlings were additionally grown on medium containing salt in order to characterise their ability to germinate under a salt stress. It was hypothesised that any differences in germination between mutant and WT would be amplified in the presence of salt. Salt is known to have an inhibitory effect on germination in WT plants (Almansouri *et al.*, 2001; Jamil *et al.*, 2007), however it was hypothesised that a lack of suberin in the

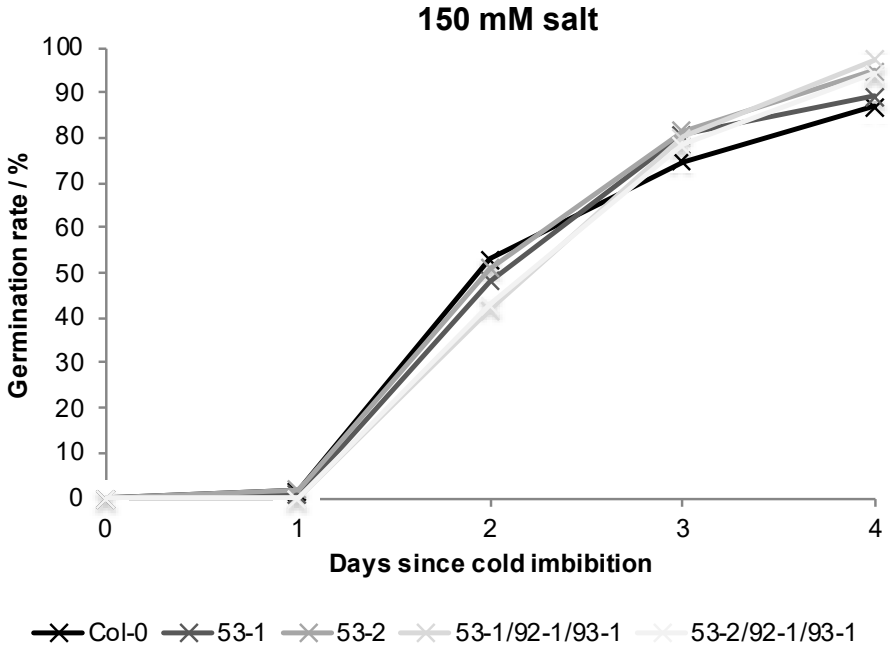
mutant lines means that they would fail to suppress their germination. A relatively high concentration of salt (150 mM) was chosen for this experiment as WT seedlings can only tolerate these conditions for a short time and it was predicted that any phenotypic differences would be more apparent under more highly stressed conditions.

As expected, no negative effect was observed on the ability of seeds to develop through each stage of germination, as all seeds that did germinate did so fully to reach the point where the cotyledon has fully emerged. When looking at germination under salt stress (Figure 4.7.A), a Kruskal Wallis test revealed evidence of a significant difference in germination between one or more of the lines. A post hoc Dunn's test with Bonferroni correction was subsequently performed, which determined that the difference between the WT and both triple mutants was significant. When comparing WT with *Atmyb53-1/myb92-1/93-1*, the calculated p-value was 0.005, while the p value for the difference between WT and *Atmyb53-2/myb92-1/93-1* was 0.025. No significant difference was indicated between any other pairs. Conversley, from Figure 4.7.B it is apparent that there seems to be little difference in germination rate between any of the lines when grown under normal conditions on plant growth medium.

In this experiment triple mutants germinated more quickly than WT, but the fact that the salt stressed seedlings did not reach 100% germination during the recorded time period means it is unclear if total capacity or speed of germination are more important factors. Therefore, in future studies seedlings should be left for longer to enable this question to be addressed. Furthermore, given that the triple mutants contain considerably less suberin than WT seedlings (Hu, 2018), these differences in

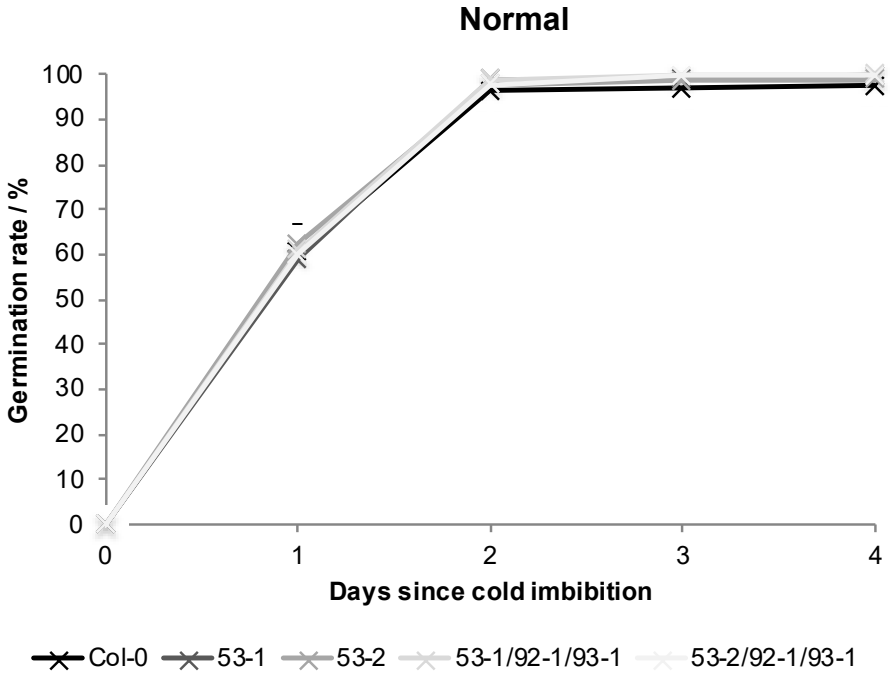
Figure 4.7. Germination assays of MYB mutants under (A) salt stress or (B) normal conditions. *Results from three biological replicates. Germination was considered to have occurred at the point of radicle emergence. The germination rate percentage was calculated as: (number of germinated seeds / total number of seeds) x 100. Error bars: SEM.*

A



Salt p-values	Day 1	Day 2	Day 3	Day 4
Kruskal Wallis	0.118	0.893	0.935	0.042
Dunn's 53-1/92/93				**0.005 ****
Dunn's 53-2/92/93				**0.025 ***

B



Control p-values	Day 1	Day 2	Day 3	Day 4
Kruskal Wallis	0.970	0.685	0.435	0.748

germination rate are unexpectedly subtle. It is also perhaps surprising that no effect on germination was seen under normal growth conditions since seed coat suberin deficiency is associated with an increase in germination rate (Beisson *et al.*, 2007; Fedi *et al.*, 2017; Yadav *et al.*, 2014), although Hu (2018) did only assess the suberin content of root tissue. Based on these data it was therefore hypothesied that further functional redundancy could be playing a role in this process and that another MYB from a neighbouring clade is potentially being upregulated to compensate for the loss of other MYBs in the triple mutants. Since any changes to the germination process could lead to a potentially catastrophic impact on the ability of a plant to reproduce, it would make sense that there were measures in place to mitigate this and it was reasoned that *AtMYB41* could be the gene acting to compensate for the loss of the three other MYBs.

4.5.2 AtMYB41 expression in MYB mutants

As described in the introduction, *AtMYB41* is expressed in the seed coat (Kosma *et al.*, 2014) and has been implicated in suberin biosynthesis (Kosma *et al.*, 2014). Moreover, it is upregulated in response to a variety of stresses, including salt (Lippold *et al.*, 2009). An experiment was designed to test the hypothesis of functional redundancy by *AtMYB41* in the *Atmyb53-1/Atmyb92-1/Atmyb93-1* and *Atmyb53-2/Atmyb92-1/Atmyb93-1* triple mutants. Expression of *AtMYB41* was measured by RT-PCR in all the single, double and triple mutants and compared to WT expression, with *AtACTIN2* as a housekeeping control. Analysis of the first two biological repeats of RT-PCR did not appear to show any evidence of upregulation or downregulation in either of the *Atmyb53* single mutants, *Atmyb92-1, Atmyb93-1* (Hu)

or the *Atmyb92-1/Atmyb93-1* double mutant relative to WT, however one of the two biological repeats indicated subtle evidence of upregulation in both of the triple mutants. A third biological repeat was performed with Col-0, *Atmyb93-1* (Hu) and both of the triple mutant lines, which supported the original result of a small increase in *AtMYB41* expression in *Atmyb53-1/Atmyb92-1/Atmyb93-1* and *Atmyb53-2/Atmyb92-1/Atmyb93-1* (Figure 4.8).

4.5.3 Lateral root density phenotypes of MYB mutants

The LR density phenotype of *AtMYB53* is so far unknown, as is the phenotype of the triple mutants. In order to characterise these MYB mutants for LR phenotypes, a series of LR density assays were conducted on *Atmyb53-1*, *Atmyb53-2* and two triple mutants. *Atmyb93-1* (Hu), *Atmyb92-1*, a double mutant and WT were also included as 'known' controls, however, the 'known' phenotypes were not always identifiable in these assays.

During the initial assay it was noticed that primary root lengths were smaller than expected for the age of the seedlings and it was thought this could either be distorting the data or delaying LR development. Therefore, the first modification to the protocol was to add 1% sucrose to the plant growth medium. Adding 1% sucrose has been confirmed to have no effect on LR phenotype but can augment seedling growth and development (J.C. Coates, personal communication). Root growth did appear to be slightly enhanced (not measured), however this did not resolve the marginal *Atmyb93-1* (Hu) phenotype.

Figure 4.9 is the combined biological replicates of two LR density assays on the MYB mutants in 7-day-old seedlings, while Figure 4.10 and Figure 4.11 show the

Figure 4.8. Semi-quantitative RT-PCR analysis of AtMYB41 expression in MYB mutants. *AtMYB41 expression at 35 cycles, AtACTIN2 loading control expression at 25 cycles. Second biological repeat (not shown) reveals similar result from semi-quantitative analyse of bands but actin loading has not been normalised for gel image.*

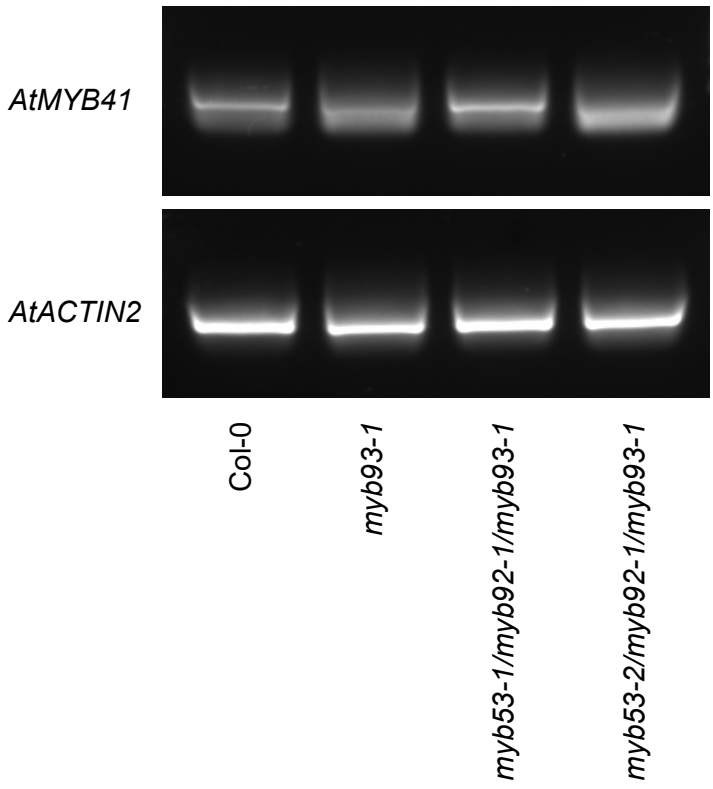
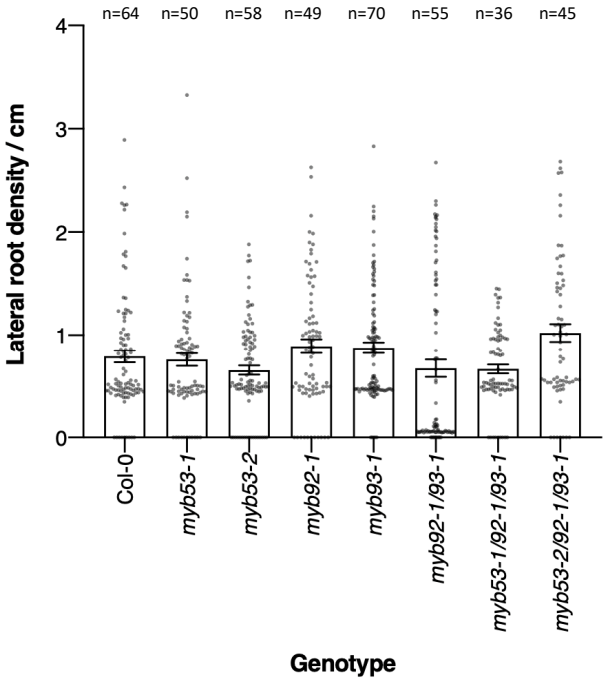

AtMYB41
AtACTIN2
Col-0
myb93-1
myb53-1/myb92-1/myb93-1
myb53-2/myb92-1/myb93-1

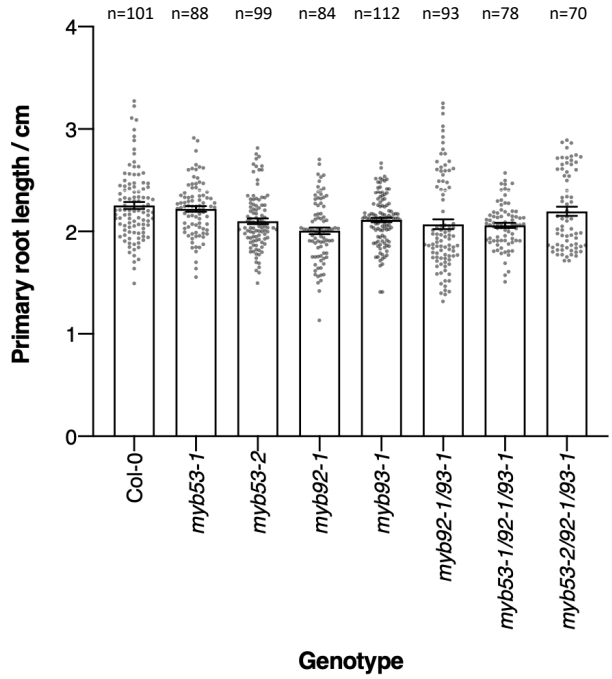
Figure 4.9. Primary and lateral root phenotypes of MYB mutants of 7-day-old seedlings. *(A) Lateral root density, (B) primary root length. Lateral roots were counted by eye. Primary root length was calculated manually by measuring photographs in ImageJ. Lateral root density was calculated as follows: number of lateral roots / length of primary root.*

A



Lateral root density p-values	*myb53-1* vs Col-0	*myb53-2* vs Col-0	*myb92* vs Col-0	*myb93* vs Col-0	*myb92/93* vs Col-0	*myb53-1/92/93* vs Col-0	*myb53-2/92/93* vs Col-0
Kruskal Wallis with Dunn's test	>0.999	>0.999	0.901	>0.999	0.454	>0.999	**0.040** *

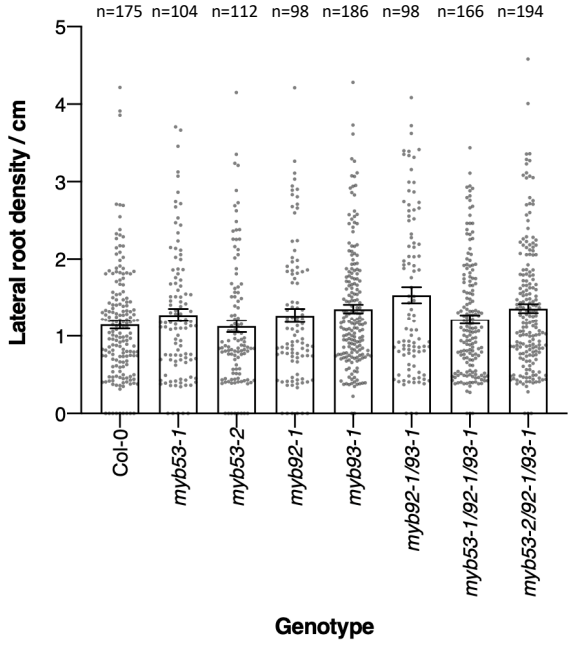
B



Primary root length p-values	*myb53-1* vs Col-0	*myb53-2* vs Col-0	*myb92* vs Col-0	*myb93* vs Col-0	*myb92/93* vs Col-0	*myb53-1/92/93* vs Col-0	*myb53-2/92/93* vs Col-0
Kruskal Wallis with Dunn's test	>0.999	**0.010** *	**<0.0001** ***	**0.082** *	**<0.0001** ***	**0.002** **	0.519

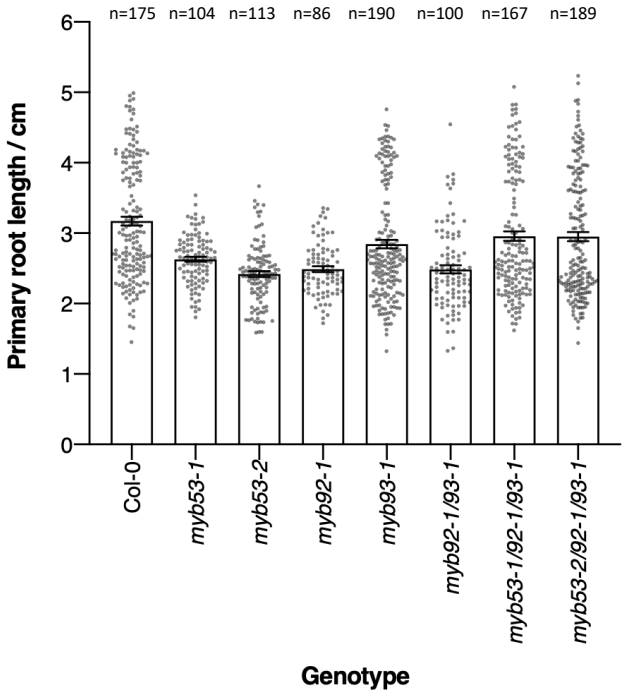
Figure 4.10. Primary and lateral root phenotypes of MYB mutants of 8-day-old seedlings. *(A) Lateral root density, (B) primary root length. Lateral roots were counted by eye. Primary root length was calculated manually by measuring photographs in ImageJ. Lateral root density was calculated as follows: number of lateral roots / length of primary root.*

A



Lateral root density p-values	*myb53-1* vs Col-0	*myb53-2* vs Col-0	*myb92* vs Col-0	*myb93* vs Col-0	*myb92/93* vs Col-0	*myb53-1/92/93* vs Col-0	*myb53-2/92/93* vs Col-0
Kruskal Wallis with Dunn's test	>0.999	>0.999	>0.999	0.157	**0.080** *	>0.999	0.159

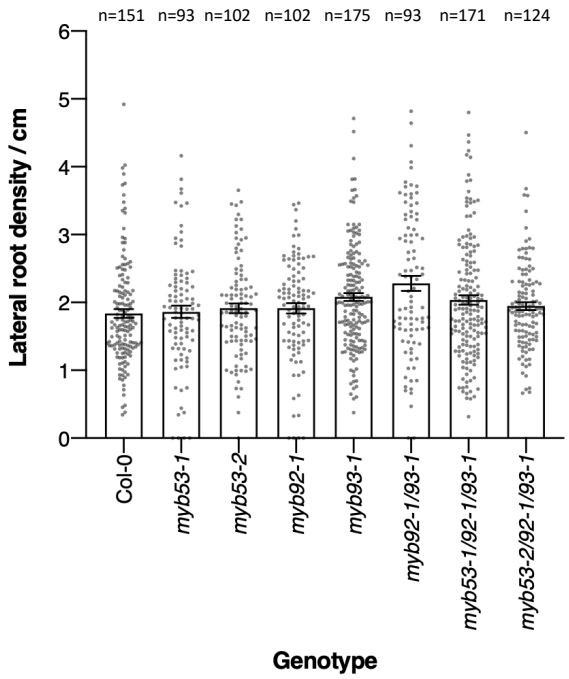
B



Primary root length p-values	*myb53-1* vs Col-0	*myb53-2* vs Col-0	*myb92* vs Col-0	*myb93* vs Col-0	*myb92/93* vs Col-0	*myb53-1/92/93* vs Col-0	*myb53-2/92/93* vs Col-0
Kruskal Wallis with Dunn's test	**0.0007** **	**<0.0001** ***	**<0.0001** ***	**0.0005** **	**<0.0001** ***	**0.017** *	**0.003** **

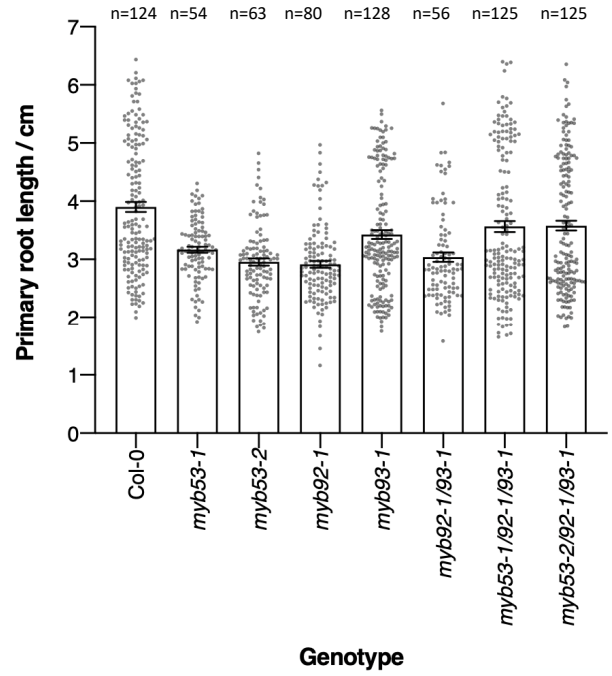
Figure 4.11. Primary and lateral root phenotypes of MYB mutants of 9-day-old seedlings. *(A) Lateral root density, (B) primary root length. Lateral roots were counted by eye. Primary root length was calculated manually by measuring photographs in ImageJ. Lateral root density was calculated as follows: number of lateral roots / length of primary root.*

A



Lateral root density p-values	*myb53-1* vs Col-0	*myb53-2* vs Col-0	*myb92* vs Col-0	*myb93* vs Col-0	*myb92/93* vs Col-0	*myb53-1/92/93* vs Col-0	*myb53-2/92/93* vs Col-0
Kruskal Wallis with Dunn's test	>0.999	>0.999	0.330	**0.0059 ****	**0.0007 *****	0.153	0.557

B



Primary root length p-values	*myb53-1* vs Col-0	*myb53-2* vs Col-0	*myb92* vs Col-0	*myb93* vs Col-0	*myb92/93* vs Col-0	*myb53-1/92/93* vs Col-0	*myb53-2/92/93* vs Col-0
Kruskal Wallis with Dunn's test	**0.0005 ****	**<0.0001 *****	**<0.0001 *****	**0.0013 ****	**<0.0001 *****	**0.0024 ****	**0.0063 ****

results of the combined biological repeats of four LR density assays carried out on 8-day-old and 9-day-old seedlings respectively. Here, the *Atmyb93-1* (Hu) and *Atmyb92-1/Atmyb93-1* phenotypes of increased LR density are revealed, though only on day nine for *Atmyb93-1* (Hu) (p= 0.0059) and on days eight and nine for the double mutant (p= 0.080 and 0.0007).

The other significant data were in the differences in length of the primary roots, with mutant roots for all lines being significantly shorter than Col-0 on days eight and nine. Differences on day seven were less pronounced. No significant difference in LR density was observed in either of the *Atmyb53* mutants. Furthermore, there appears to be very little effect on LR density as a result of increasing the number of genes being knocked out. The *Atmyb53-2/Atmyb92-1/Atmyb93-1* triple mutant does appear to have a higher LR density on day seven, and this is a significant difference (p = 0.040), but this is lost by day eight and nine. Based on these data, it is difficult to assess if the significant result is aberrant, though it is more likely that the two triple mutants would behave similarly. The day seven data for the *Atmyb53-2/Atmyb92-1/Atmyb93-1* triple mutant are also based on fewer seedling roots, as there were problems with some plates of this mutant. However, it is important to note that as the *Atmyb93* phenotype could not be reliably reproduced during these experiments, this brings into question the validity of the data collected for the other mutant lines. Importantly, however, a maternal effect can be ruled out since age-matched seed batches were used for each biological repeat, which were independently produced.

A different approach was used in an attempt to clarify the known *AtMYB93* phenotype under my assay conditions. It was thought possible that conditions in the

growth room were different than that of previous experiments and that this could be disrupting the phenotype. For instance, for an unknown period of time whilst conducting these experiments there had been issues with the photoperiod settings as lights had been staying on for much longer than the 16 h setting. Consequently, it was decided to try growing seedlings under different conditions available. Four different locations were chosen: A 21°C growth incubator, 24°C growth incubator, the glasshouse and the 22°C growth room as a control. Data were collected for seedlings grown under these conditions on two separate occasions, but the *Atmyb93* phenotype remained no more obvious under any of these growth conditions, whereas a noticeable difference was observed in the length of the primary root, most likely due to the differing temperatures (data not shown).

4.5.4 Elemental analysis of MYB mutants

Figure 4.12.A and 4.12.B show a comparison of the sulfur, sodium and magnesium content between the single, double and triple mutants of the S24 clade of MYBs, measured using ICP-MS. None of the data were found to be significant, although a subtle trend of increased uptake in sulfur, sodium and magnesium can be observed in the *Atmyb92-1*/*Atmyb93-1* double mutant and in the sodium content of both the *Atmyb53-1/Atmyb92-1/Atmyb93-1* and *Atmyb53-2/Atmyb92-1/Atmyb93-1* triple mutants. The largest variation is a 1.2-fold increase in sodium in the *Atmyb53-2/Atmyb92-1/Atmyb93-1* mutant. Both of the *Atmyb53* mutants show relatively little difference in the ion content of their shoot tissue compared to WT; however, strangely there is on average around a 1.07-fold decrease in sodium content in the

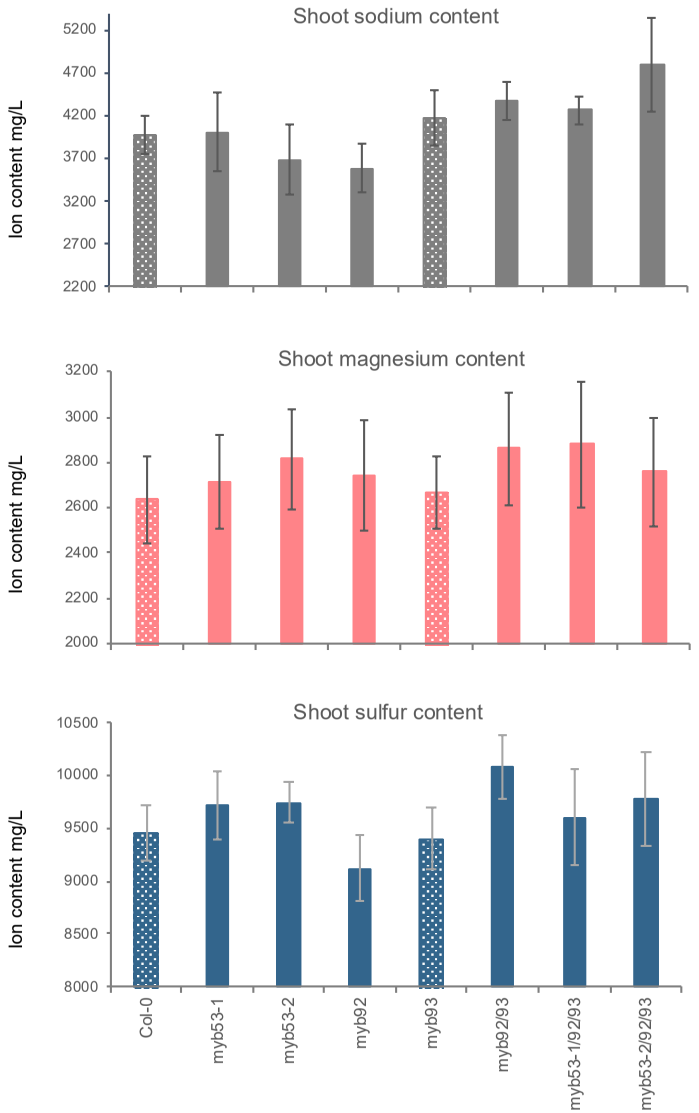
Figure 4.12. Comparison of the sulfur, sodium and magnesium ion content of MYB mutants relative to WT in 25-day-old Arabidopsis shoot tissue. *(A) Heatmap of the means from three biological repeats. Values are presented as a ratio relative to the WT to represent relative loss or gain in ion content. (B) Graphs showing actual values. Error bars: SEM.*

A

	S	Na	Mg
Col-0	100.00	100.00	100.00
myb53-1	102.66	100.89	102.91
myb53-2	103.04	92.75	106.78
myb92	96.44	90.27	104.11
myb93	99.40	105.18	101.23
myb92/93	106.66	109.83	108.63
myb53-1/92/93	101.55	107.44	109.28
myb53-2/92/93	103.35	120.65	104.73

Shoot ion content	S	Na	Mg
Kruskal Wallis	0.595	0.430	0.878

B



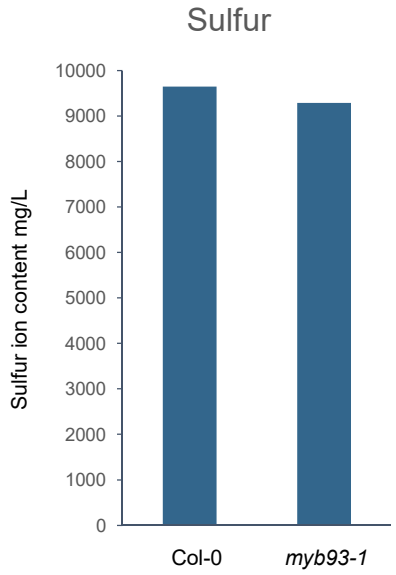
Atmyb53-2 mutant relative to WT. A decrease in sodium was also measured in the *Atmyb92* mutant and this was around a 1.1-fold change. Unlike in the preliminary data by Wilkinson, there is not an increase in sulfur or magnesium in *Atmyb93*, however there is relatively little change in sodium content.

4.5.5 Elemental analysis of the Atmyb93 mutant under sulfur (and salt) stress

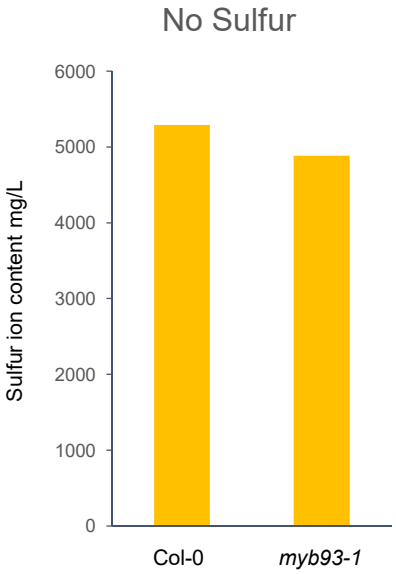
As outlined in the introduction, preliminary elemental analysis had already indicated that ion uptake in *Atmyb93-1* (Gibbs) mutants might be modified under sulfur or salt stress (Wilkinson, 2018) and so further experiments were conducted to gather additional data on the response of *AtMYB93* to these stresses. The plants did not grow well enough to collect enough tissue after two attempts at growing them under salt stress, which was first performed using 125 mM salt before being reduced to 100 mM. Nevertheless, data were obtained for two biological repeats from growing *Atmyb93* (Gibbs) mutant and WT plants on growth medium both with and without sulfur (Figure 4.13). A decrease of around 170 mg/L and 543 mg/L was measured in the sulfur content of *Atmyb93* grown on sulfur replete medium compared to WT. This trend is also seen in the *Atmyb93* vs WT data for the shoot ion analysis data presented in Figure 4.12. However, these data are in apparent contrast to previous preliminary data by Wilkinson (2018), which showed an increase in the sulfur content of *Atmyb93* mutant plants grown on sulfur containing medium, a trend that has been conserved in the equivalent data of figure 4.5. A small decrease in sulfur content in *Atmyb93* was also noted in plants grown on medium lacking sulfur. This was a decrease in uptake of around 470 mg/L and 403 mg/L of sulfur in the *Atmyb93*

Figure 4.13. Comparison of the sulfur ion content in Atmyb93 mutant tissue relative to WT in 25-day-old Arabidopsis shoot tissue grown on medium (A) with or (B) without sulfur. *Raw values provided. Data is from two biological repeats.*

A



B



	RPT1	RPT2
Col-0 normal sulfur	9882.25	9408.62
***myb93-1* normal sulfur**	9711.53	8865.15
Col-0 no sulfur	5293.42	5283.89
***myb93-1* no sulfur**	4823.31	4947.07

mutant compared to WT across the samples for the two biological repeats. For these data the result does match Wilkinson (2018).

4.6 Discussion

Mutants were obtained for 13 of the candidate gene mutants identified in the previous chapter and two of these mutants, *Atabcg2* and *Ats-acp-des1*, displayed LR phenotypes of increased LR density. The potential root phenotypes of the other candidate gene mutants and of *AtSCR*, a potential regulator of *At*MYB93, still need to be investigated. Elemental analysis was used to measure the differences in ion uptake in the shoot tissue of mature *Arabidopsis* plants. This was performed in both the S24 clade MYB mutants, and mutants for four of the candidate *At*MYB93 targets: *Atabcg2*, *Atgpat6*, *Ataha4* and *Ats-acp-des1*. This technique was also used to start looking at differences in the uptake of sulfur in *Atmyb93* mutant plants under sulfur stress. LR density and germination phenotypes were assessed for plants mutant in the S24 clade of MYBs, with subtle germination phenotypes detected in *Atmyb53-1/Atmyb92-1/Atmyb93-1* and *Atmyb53-2/Atmyb92-1/Atmyb93-1* under salt stress. Lastly, evidence of functional redundancy has been observed with *AtMYB41* putatively upregulated to counteract suberin reduction in MYB triple mutant seeds.

4.6.1 Lateral root phenotypes have been characterised in putative downstream targets of AtMYB93 but did not reveal a phenotype in Atmyb53 or MYB triple mutants

AtABCG2 and *AtS-ACP-DES1* were the two candidate mutants assessed for LR phenotypes, and both had LR phenotypes mimicking that of *Atmyb93*. Their

identities are discussed in section 3.6.4.6 and 3.6.4.7 in the previous chapter, though very little is known about *AtS-ACP-DES1*. *AtABCG2* on the other hand, has a characterised role as a regulator of suberin biosynthesis in the roots and seed coat and a *pAtABCG2:GUS* reporter highlighted expression as localised to the endodermis in roots (Yadav *et al.*, 2014). Together these data present *AtABCG2* as a particularly promising candidate for further study. Since *AtABCG2* has previously been reported as being redundant with two other ABCG genes (Yadav *et al.*, 2014), it would be interesting to investigate the LR phenotypes of these other mutants and their corresponding double and triple mutants. Moreover, this study revealed that the seed coat and roots of triple mutants are more permeable (Yadav *et al.*, 2014), which corroborates the LR phenotype found in this work. Neither of these genes were identified by the work in the previous chapter by either qRT-PCR of RNA-seq and the possible reasons behind this are explored in chapter VI.

The lack of phenotype in either of the *Atmyb53* mutants or triple mutants suggests a different role for *Atmyb53* in the roots, particularly as *AtMYB53* is expressed in a range of tissue types, whereas *AtMYB93* expression is specific to a very small area of the root (Gibbs *et al.*, 2014). Furthermore, *Atmyb92*, which is also more widely expressed than *AtMYB93*, has already been shown to have no LR phenotype (Gibbs *et al.*, 2014). Yet the fact that it has been found that a *Atmyb92-1*/*Atmyb93-1* double mutant enhances the *Atmyb93* phenotype suggested that the *Atmyb53* mutants could be expected to have an additive effect on the LR phenotype since this is shown to be the case in the suberin content of seedlings (Hu, 2018). However, the data presented here need to be treated with caution given the subtlety

of the *Atmyb93-1* (Hu) phenotype in these experiments and furthermore, the inconsistency of the primary root length data with the results of other experiments.

4.6.2 MYB triple mutants have a limited response to germination under salt stress but AtMYB41 could be functionally redundant

In two out of three biological repeats there were signs of *AtMYB41* upregulation in the *Atmyb53-1/Atmyb92-1/Atmyb93-1* and *Atmyb53-2/Atmyb92-1/Atmyb93-1* triple mutants. These lines are those with the largest loss of suberin, something typically associated with a germination phenotype (Beisson *et al.*, 2007; Fedi *et al.*, 2017; Yadav *et al.*, 2014), though the suberin content of seedlings was measured, rather than that of the seed coat itself. It makes sense that *AtMYB41* is showing evidence of enabling recovery or redundancy given its own role in suberin biosynthesis (Kosma *et al.*, 2014) and the otherwise limited germination phenotype in these triple mutant plants. However, the upregulation in expression appears to be subtle based on this semi-quantitative experiment and further experiments such as qRT-PCR need to be performed. Furthermore, there could be additional genes that are being upregulated in the triple mutants such as the other known seed- and suberin-based MYBs, *AtMYB107* and *AtMYB9* that have been shown to regulate suberin deposition in the seed coat (Gou *et al.*, 2017; Lashbrooke *et al.*, 2016).

4.6.3 Elemental analysis of the ion content of Arabidopsis shoot tissue

Based on the data presented in this thesis, only very limited conclusions can be drawn about the differences in shoot ion content between both the mutants and growth conditions tested. Further biological repeats are needed and the experimental

design may require additional development. Additionally, work on the MYB mutants was conducted using the Hu *Atmyb93* and matched Col-0 lines, whereas both of the other experiments and the study by Wilkinson (2018) were performed using the Gibbs *Atmyb93* and matched Col-0 lines. This could prove important once more data is obtained given the issues with the LR density phenotype and potential differences between these two lines.

A caveat of using *Arabidopsis* tissue in these experiments is that it is difficult to obtain a large enough mass of tissue for the analysis. Plants were harvested after 25 days so that they had grown enough for differences in their ion content to be meaningful, but at the point just prior to flower production. Around 100-200 mg tissue was provided for each sample and this consisted of around 40 plants. This mass of tissue is close to the minimum required for the technique, whereas in work on other species such as is ongoing for the *MYB93* homologue in tomato, obtaining 1000 mg is easily achievable and thus data from analysis on these plants could prove more valuable (X. Cao, unpublished). Ultimately, it means that this technique may be better suited to work on larger species and could be used to examine phenotypes in crops such as wheat, barley and rice. Should a significant phenotype be established, additional experiments will be required to assess whether differences in the ion content of plants are a direct or indirect effect of *AtMYB93*.

4.6.4 Challenges and limitations

4.6.4.1 Genotyping candidate gene mutant lines

The incorrect identification of the genotype from T-DNA express for at least seven of the 13 candidate gene mutant lines obtained from NASC highlights the

importance of checking all genetic material before any experimental analysis is conducted. The reasons for these inconsistencies were considered.

Firstly, two of the mutant lines all survived selection, despite a conflicting T-DNA prediction and genotyping result. One worry was that the amount of BASTA® used in the selection medium was not enough to kill seedlings lacking the resistance gene, particularly as most lines were selected for on kanamycin. However, one other line had been put under BASTA® selection and roughly 25% of these plants were killed and so the BASTA® strength that was used appeared to be sufficient. Therefore, it was thought probable that the seedlings appeared to be homozygous mutants under BASTA® selection because a second T-DNA insert was present in another part of the genome.

When *AtMYB54* was first genotyped using a few pooled seedlings, the results suggested that it was actually WT, though it was considered that the BP primer could also be causing an issue since there is not a positive control to test this. However, when put under selection, although some seedlings did survive, they appeared stressed and after the healthiest seedlings were transferred to soil, they were very slow to mature. It is feasible that these seedlings were being killed by the kanamycin but that transferring them to soil facilitated their rescue and subsequent survival.

Although *Atcys5* appeared to be WT under selection and indeed it is possible that the wrong seed was somehow received, it is also possible that silencing of the kanamycin resistance gene had occurred. This is all the more likely since the kanamycin resistance gene is prone to silencing after several generations (SIGnAL, undated). Genotyping could have been carried out on seedlings grown without selection in an attempt to find a homozygous mutant; however, ultimately it was

decided to focus on the other mutants since LR studies had already begun to reveal potentially interesting phenotypes in some of these mutants.

4.6.4.2 Replicating the AtMYB93 phenotype

Atmyb93 has a published mutant phenotype (Gibbs *et al.*, 2014) that has been documented by several members of the Coates lab. Unfortunately, its phenotype of increased LR density could not be replicated using *AtMYB93* seed obtained from Hu (2018), despite several attempts to address this problem described in section 4.5.3. Previous work on *AtMYB93* has been carried out on *Atmyb93* mutants sourced from NASC as heterozygotes and propagated as homozygotes by the Coates lab under their standard conditions for >7 years and the most likely cause for the lack of measurable phenotype was considered to be the different life history of the mutants. The mutants used in the experiments on the MYB mutants were obtained and produced by Hu (2018) in Canada and so the At*myb93-1* data has been generated by an independently created line, produced from the same mutant allele but with a different life history and growth conditions. The observation of significantly shorter primary roots in the *Atmyb93-1* mutant compared to WT is also at odds with the opposite phenotype seen in both the LR assay of the candidate gene mutants and data by Gibbs *et al.* (2014) and additionally, Gibbs *et al.* (2014) showed a trend of *AtMYB93* over-expressor seedlings with shorter primary roots. The foremost difference once again, is the origin of the mutants.

The first period of LR counting was done on seven-day-old seedlings and the last time point was on ten-day-old seedlings. This timescale was thought to be a suitable window in which to observe the LR phenotype. Younger seedlings have not

really begun to develop any LRs and it becomes difficult to count LRs beyond ten days using the current method. The spread of the data is large indicating a high level of natural variation in both primary root length and LR density.

The normal *Atmyb93* phenotype is sometimes not visible in seven-day-old seedlings, probably because LRs are only beginning to develop, however sometimes a difference in the number of seedlings with no LRs can be observed. By days eight and nine the phenotype should be visible, and though it is usually still present beyond this, at this point it can become difficult to count the developing LRs as the seedlings grow larger. The combination of a subtle phenotype on tissue that is difficult to analyse means that initial next steps should continue to attempt to develop a better method for analysis of the *AtMYB93* LR phenotype. This could include making changes to how the experiment is set up, such as assessing potential environmental factors that could cause variation in the phenotype, as well as looking into alternative software for analysis.

4.6.4.3 Failed growth of plants under salt stress for elemental analysis

Arabidopsis plants grown under salt stress for 25 days failed to grow large enough for elemental analysis. The difference between this experiment and Wilkinson (2018) was that Wilkinson used powdered ½ MS powder with additional Gamborg's vitamins, whereas I used a version of ½ MS made from a series of macro- and micronutrients put into solution. This was to enable the experiment to run alongside sulfur stress conditions meaning only one set of control plants were required per experiment due to space and material limitations. It is possible that even with the reduced salt concentration, the use of medium slightly poorer in nutrient

content was enough to overstress the seedlings, meaning that not enough tissue could be collected for analysis – as a minimum of 100 mg of dried shoot tissue was required.

4.7 Conclusion

This chapter reveals subtle germination phenotypes in MYB triple mutants under salt stress, but no LR density phenotype, while *AtMYB41* is presented as a having a putative role in functional redundancy in the MYB triple mutants, with further experiments required. Elemental analysis of shoot tissue also requires additional work, as potential phenotypes are currently unclear. The results of this chapter also provide further evidence substantiating the choice of some of the candidate gene targets of *At*MYB93. The development of MYB antibodies would be a useful next step to enable protein expression to be measured across different species, as well as in knock out or over-expressor lines. The next chapter aims to start looking at the evolution and possible functions of MYB93 in crops, with particular reference to important monocot species like cereals.

Chapter V

MYB93 HOMOLOGUES: EVOLUTIONARY CONSERVATION INCLUDING IN CROPS

5.1 Introduction

Climate change and an increasing population remain crucial challenges to plant biodiversity and food security (FAO, 2017; Müller & Robertson, 2014) and the manipulation of plant root systems has the potential to produce crop varieties with a range of phenotypes to best suit growth in different soil types or those able to withstand changing climatic conditions.

The goal for long term impact in many areas of plant research is to transfer knowledge from model organisms such as *Arabidopsis* to economically important crop species, ultimately enabling the improvement of such species (Hochholdinger *et al.*, 2004). Relatively little is known about the presence and function of *MYB93* beyond *Arabidopsis* and to date, the most notable research in a crop concerns the *MYB93* homologue in apple, *MdMYB93* (Legay *et al.*, 2016; Legay *et al.*, 2017). *Arabidopsis* has proven itself an exceptional candidate in studying many aspects of plant growth and development (Koornneef & Meinke, 2010), including expanding our knowledge of processes more typically found in species sharing little resemblance to the small, weed-like plant (Hertzberg *et al.*, 2001). Furthermore*, Arabidopsis* is a good model for developmental studies and has roots with a simple structure that are easily manipulated (Scheres & Wolkenfelt, 1998), so makes sense as a species in which to test hypotheses before moving onto more complex species such as wheat. For example, in depth studies of auxin signalling in *Arabidopsis* have formed the basis of translating this knowledge into cereals (Wang *et al.*, 2017).

The specificity of expression of *AtMYB93* in only a few root cells makes it an ideal target for manipulation in potential crop improvement studies, since this reduces the possibility of unwanted impacts on the morphology of other tissues. Moreover,

the expression of *AtMYB93* in the fruit of some species presents additional avenues for improving the storage or nutritional quality of crops.

Cereals are crucially important as both staple food crops across many areas of the globe and as the bulk of livestock feed. Gibbs *et al.* (2014) showed that MYB93s are present in monocots as well as dicots, although no MYB93 homologue exists in bryophytes (Bowman *et al.*, 2017; D. Gibbs, J.C. Coates, unpublished; Du *et al.*, 2015). The S24 clade contains *AtMYB93* and its close relatives investigated in chapter IV – *AtMYB92* and *AtMYB53* (Du *et al*., 2015). The clades neighbouring S24 include MYBs such as *AtMYB41*, *AtMYB9* and *AtMYB107* (section 1.4.4), though there is controversy over the fine details of these clades. Phylogenies by both Du *et al.* (2015) and Stracke *et al.* (2001) place *AtMYB107* and *AtMYB9* as in a separate, though closely related clade to the one containing *AtMYB93*. Lashbrooke *et al.* (2016), however, places them in the same clade with *AtMYB93* sitting at the base of these MYBs. The accuracy of these groupings is examined in this chapter through gene structure analysis.

5.2 MYB93 homologues

5.2.1 Phylogenetic analysis of MYB93 crop homologues

A phylogenetic analysis was performed in an attempt to identify putative homologues of *AtMYB93* in other species, with specific reference to cereal crops. Species included in the trees were selected to cover a range of monocots, including the important cereal crops *Triticum aestivum* (wheat), *Hordeum vulgare* (barley), *Oryza sativa* (rice), *Zea mays* (corn), *Sorghum bicolor* and *Setaria italica* (foxtail millet), as well as grasses such as *Brachypodium stacei*, *Brachypodium distachyon*,

Panicum virgatum and *Panicum hallii*. Some early diverging species such as *Ananus comosus* (pineapple) and *Musa acuminata* (banana) were incorporated, as well as the basal angiosperm *Amborella trichopoda*. Also included were *Brassica rapa* and *Brassica oleracea*, two dicot crop species that are closely related to *Arabidopsis*. Table 5.1 shows the putative homologues selected for each species. This analysis also complements the results of a previous phylogeny covering a more diverse range of species including dicot crops such as tomato, but also monocots like rice (Du *et al.*, 2015; X. Cao, unpublished). Cao identified two genes as putative MYB93 homologues in rice (LOC_Os06g11780 and LOC_Os02g51799), which were also previously isolated by Du *et al.* (2015), and provided monocot sequences by which to identify homologues in addition to the dicot *Arabidopsis AtMYB93* sequence. The protein sequences of both of these genes were also identified as MYB93 homologues in my phylogeny (*O. sativa* 1 and *O. sativa* 3, respectively), which improves the validity of these analyses.

Several known *Arabidopsis* MYBs were used in addition to *AtMYB93* for the phyogenetic analysis. *AtMYB36* and *AtMYB75* are two MYBs involved in different pathways of LR development and are not considered closely related to *AtMYB93* (Du *et al.*, 2015; Stracke *et al.*, 2001). These were chosen to represent an outgroup for the tree, as any putative MYB93 homologues that clustered with this outgroup could be considered to not be true MYB93 homologues. None of the putative MYB93 sequences that were used in creating phylogenetic trees clustered with either of these MYBs, which increases the validity of the methods used to identify putative MYB93 homologues. *AtMYB9*, *AtMYB107* and *AtMYB41* were included as MYBs that are closely related to *AtMYB93* (Du *et al.*, 2015; Stracke *et al.*,

Table 5.1. Putative AtMYB93 homologues identified in BLAST search. *Gene identifiers are designated as in Phytozome and table order relates to the numbered species names on phylogenetic trees. Putative homologues were numbered in order of similarity to the AtMYB93 peptide sequence. * Amborella_trichopoda_5 was originally identified on NCBI and its Phytozome equivalent could not be found. Green = identified in Phytozome against Arabidopsis, dark blue = identified in NCBI against Arabidopsis, red = identified in Phytozome against rice, light blue = identified in EMBL against Arabidopsis.*

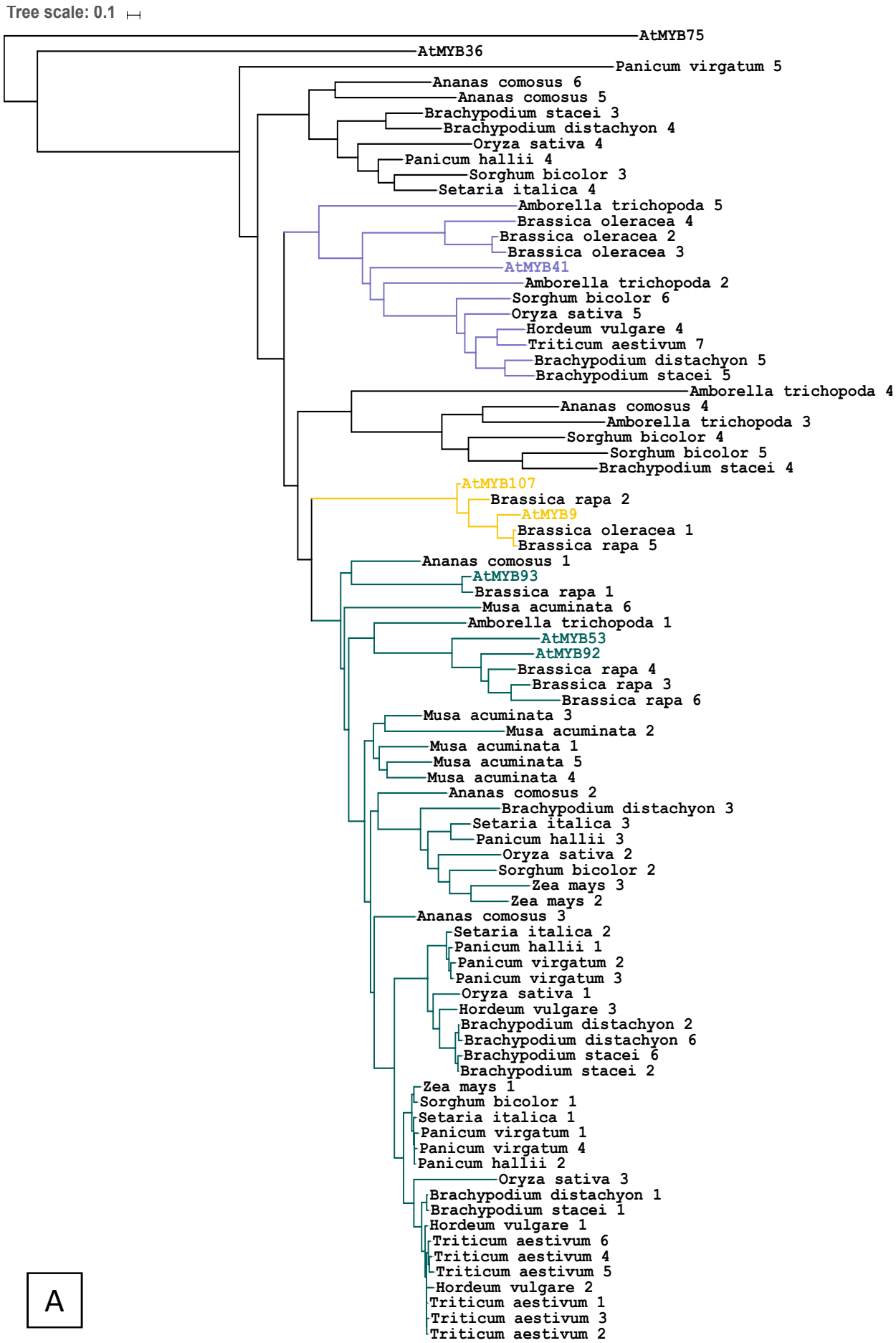
Species	1	2	3	4	5	6	7
***Amborella trichopoda* v1.0**	evm_27.model.AmTr_v1.0_scaffold00038.125	evm_27.model.AmTr_v1.0_scaffold00003.458	evm_27.model.AmTr_v1.0_scaffold00010.533	evm_27.model.AmTr_v1.0_scaffold00091.4	* XP_011624626.2		
***Ananas comosus* v3 (Pineapple)**	Aco004081.1	Aco001218.1	Aco002802.1	Aco001084.1	Aco013198.1	Aco020874.1	
***Brachypodium distachyon* v3.1**	Bradi3g58770.1	Bradi1g45510.1	Bradi3g38807.1	Bradi1g51747.2	Bradi1g24887.1	Bradi1g45510.2	
***Brachypodium stacei* v1.1**	Brast04G023500.1	Brast07G078400.1	Brast07G007700.1	Brast09G106500.1	Brast06G094700.1	Brast07G078400.2	
***Brassica oleracea* capitata v1.0**	Bol034418	Bol028417	Bol024303	Bol032639			
***Brassica rapa* FPsc v1.3**	Brara.I02373.1	Brara.E03547.1	Brara.C00413.1	Brara.B00350.1	Brara.C00739.1	Brara.C00413.1	
***Hordeum vulgare* v2 (Barley)**	HORVU6Hr1G078300.3	HORVU6Hr1G078270.2	HORVU7Hr1G037060.3	HORVU2Hr1G041300.2			
***Musa acuminata* v1 (Banana)**	GSMUA_Achr6T04130_001	GSMUA_Achr9T14850_001	GSMUA_Achr3T06470_001	GSMUA_AchrUn_randomT15560_001	GSMUA_Achr10T24560_001	GSMUA_Achr10T19130_001	
***Oryza sativa* v7_JGI (Rice)**	LOC_Os06g11780.1	LOC_Os08g37970.1	LOC_Os02g51799.1	LOC_Os06g02250.1	LOC_Os07g37210.1		
***Panicum hallii* v2.0**	Pahal.D02517.1	Pahal.A03821.1	Pahal.F01708.1	Pahal.J01057.1			
***Panicum virgatum* v1.1**	Pavir.J03585.1	Pavir.Db01850.1	Pavir.Da01312.1	Pavir.Aa00510.1	Pavir.J16675.1		
***Setaria italica* v2.2 (Foxtail millet)**	Seita.1G324800.1	Seita.4G091600.1	Seita.6G190300.1	Seita.4G008000.1			
***Sorghum bicolor* v3.1.1**	Sobic.004G231700.1	Sobic.007G224800.1	Sobic.010G008000.1	Sobic.006G107500.1	Sobic.004G191800.1	Sobic.002G337800.1	
***Triticum aestivum* v2.2 (Wheat)**	Traes_6BL_0D6CD031E.1	Traes_6AL_2BA02CAA9.1	Traes_6DL_8BF29D46C.1	Traes_6AL_52227091B.1	Traes_6BL_5A2DB8F78.1	Traes_6DL_499F64524.1	Traes_2AS_FA7059723.1
***Zea mays* Ensembl-18 (Corn)**	GRMZM2G159547_T01	GRMZM2G017268_T01	GRMZM2G017268_T02	GRMZM2G110135_T01			

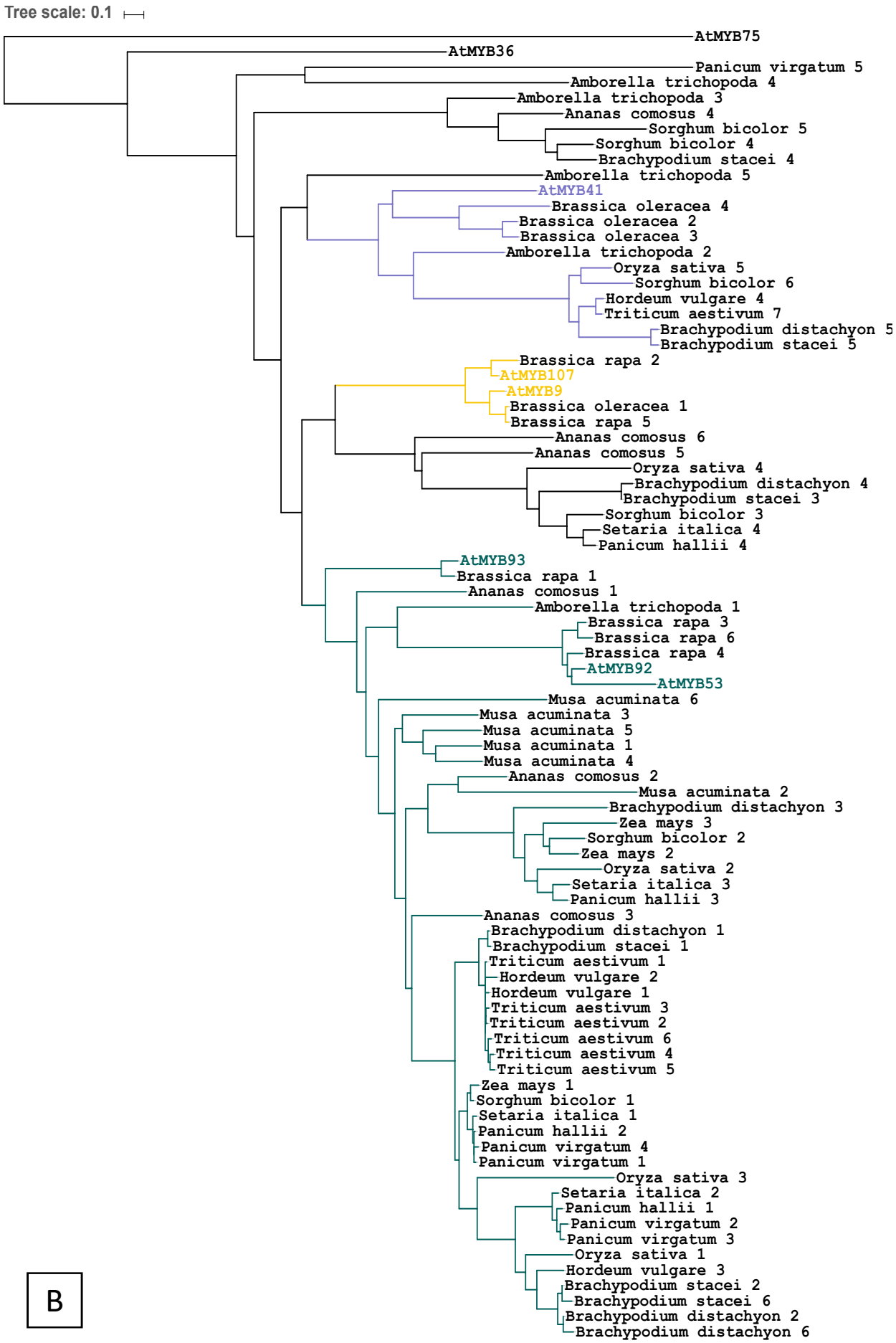
2001) but would ensure that MYB93-related proteins were placed confidently in the correct clade. *AtMYB53* and *AtMYB92* were included to complete the *Brassicaceae*-specific *AtMYB93* clade (S24, Du *et al.* 2015).

In total 77 of the protein-coding sequences identified in the pBLAST search with *At*MYB93 were used to create each phylogenetic tree and nearly two thirds of these proteins (48) clustered with *AtMYB93*, including the basal angiosperm evm_27.model.AmTr_v1.0_scaffold00038.125 (*A. trichopoda* 1), which sits near the base of the clade. Phylogenetic trees were created using two different alignment methods (MUSCLE and ClustalO) and two different algorithms (PhyML and BioNJ). Of the four trees created (Figure 5.1), all corroborate with each other and show similar groupings of protein sequences; in particular the grouping of sequences with the S24 clade is well conserved.

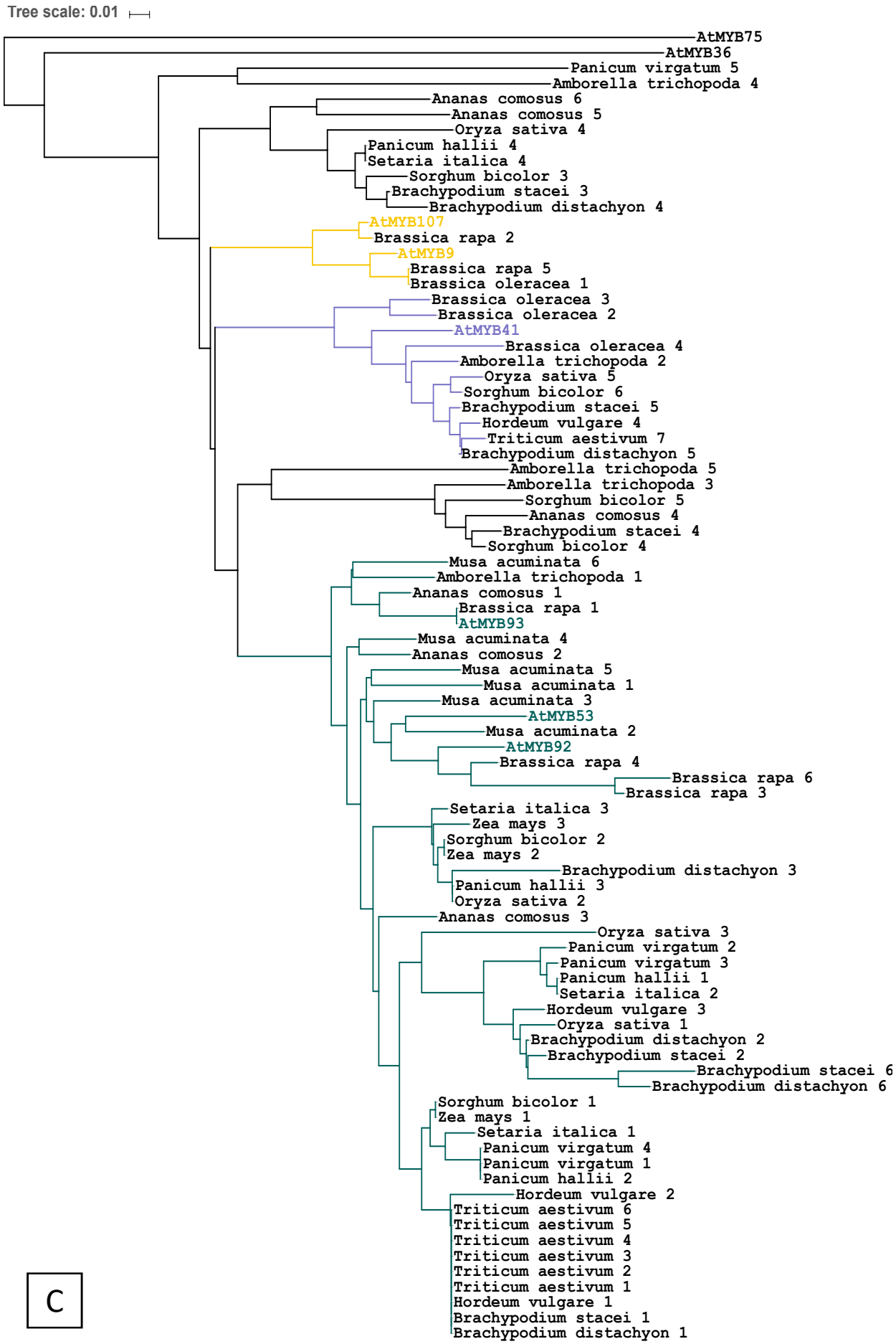
A number of of the putative homologues were found to cluster more in line with some of the other closely related MYBs of this analysis. A pBLAST search was performed using the *AtMYB41*, *AtMYB9* and *AtMYB107* sequences to identify a likely homologue for these MYBs in rice, barley and wheat. The results of the pBLAST suggested LOC_Os07g37210 (*O. sativa* 5), HORVU2Hr1G041300 (*H. vulgare* 4) and Traes_2AS_FA7059723 (*T. aestivum* 7) as possible MYB41 homologues. The phylogenetic tree was recreated with these protein sequences and they were confirmed to cluster with the MYB41 group (Figure 5.1). Moreover, no monocot sequences form a monophyletic clade with *AtMYB9* and *AtMYB107* in any of the four trees and the next closest monocot sequences are outside of this clade. The phylogeny by Du *et al.* (2015) is in agreement with this result, suggesting that these MYBs might be Brassica-specific or dicot-specific.

Figure 5.1. Phylogenetic trees identifying MYB homologues in other species. *(A) Using MUSCLE alignment with PhyML algorithm with 100 iterations. (B) Using Clustal O alignment with PhyML algorithm with 100 iterations. (C) Using MUSCLE alignment with BioNJ algorithm. (D) Using Clustal O alignment with BioNJ algorithm. Size of grey circle denotes Bootstrap values. Suggested MYB93 clade highlighted in green, MYB9 and 107 highlighted in yellow and MYB41 in purple. (Figure made with iTOL).*


Tree scale: 0.1
AtMYB75
AtMYB36
Panicum virgatum 5
Ananas comosus 6
Ananas comosus 5
Brachypodium stacei 3
Brachypodium distachyon 4
Oryza sativa 4
Panicum hallii 4
Sorghum bicolor 3
Setaria italica 4
Amborella trichopoda 5
Brassica oleracea 4
Brassica oleracea 2
Brassica oleracea 3
AtMYB41
Amborella trichopoda 2
Sorghum bicolor 6
Oryza sativa 5
Hordeum vulgare 4
Triticum aestivum 7
Brachypodium distachyon 5
Brachypodium stacei 5
Amborella trichopoda 4
Ananas comosus 4
Amborella trichopoda 3
Sorghum bicolor 4
Sorghum bicolor 5
Brachypodium stacei 4
AtMYB107
Brassica rapa 2
AtMYB9
Brassica oleracea 1
Brassica rapa 5
Ananas comosus 1
AtMYB93
Brassica rapa 1
Musa acuminata 6
Amborella trichopoda 1
AtMYB53
AtMYB92
Brassica rapa 4
Brassica rapa 3
Brassica rapa 6
Musa acuminata 3
Musa acuminata 2
Musa acuminata 1
Musa acuminata 5
Musa acuminata 4
Ananas comosus 2
Brachypodium distachyon 3
Setaria italica 3
Panicum hallii 3
Oryza sativa 2
Sorghum bicolor 2
Zea mays 3
Zea mays 2
Ananas comosus 3
Setaria italica 2
Panicum hallii 1
Panicum virgatum 2
Panicum virgatum 3
Oryza sativa 1
Hordeum vulgare 3
Brachypodium distachyon 2
Brachypodium distachyon 6
Brachypodium stacei 6
Brachypodium stacei 2
Zea mays 1
Sorghum bicolor 1
Setaria italica 1
Panicum virgatum 1
Panicum virgatum 4
Panicum hallii 2
Oryza sativa 3
Brachypodium distachyon 1
Brachypodium stacei 1
Hordeum vulgare 1
Triticum aestivum 6
Triticum aestivum 4
Triticum aestivum 5
Hordeum vulgare 2
Triticum aestivum 1
Triticum aestivum 3
Triticum aestivum 2
A


Tree scale: 0.1
AtMYB75
AtMYB36
Panicum virgatum 5
Amborella trichopoda 4
Amborella trichopoda 3
Ananas comosus 4
Sorghum bicolor 5
Sorghum bicolor 4
Brachypodium stacei 4
Amborella trichopoda 5
AtMYB41
Brassica oleracea 4
Brassica oleracea 2
Brassica oleracea 3
Amborella trichopoda 2
Oryza sativa 5
Sorghum bicolor 6
Hordeum vulgare 4
Triticum aestivum 7
Brachypodium distachyon 5
Brachypodium stacei 5
Brassica rapa 2
AtMYB107
AtMYB9
Brassica oleracea 1
Brassica rapa 5
Ananas comosus 6
Ananas comosus 5
Oryza sativa 4
Brachypodium distachyon 4
Brachypodium stacei 3
Sorghum bicolor 3
Setaria italica 4
Panicum hallii 4
AtMYB93
Brassica rapa 1
Ananas comosus 1
Amborella trichopoda 1
Brassica rapa 3
Brassica rapa 6
Brassica rapa 4
AtMYB92
AtMYB53
Musa acuminata 6
Musa acuminata 3
Musa acuminata 5
Musa acuminata 1
Musa acuminata 4
Ananas comosus 2
Musa acuminata 2
Brachypodium distachyon 3
Zea mays 3
Sorghum bicolor 2
Zea mays 2
Oryza sativa 2
Setaria italica 3
Panicum hallii 3
Ananas comosus 3
Brachypodium distachyon 1
Brachypodium stacei 1
Triticum aestivum 1
Hordeum vulgare 2
Hordeum vulgare 1
Triticum aestivum 3
Triticum aestivum 2
Triticum aestivum 6
Triticum aestivum 4
Triticum aestivum 5
Zea mays 1
Sorghum bicolor 1
Setaria italica 1
Panicum hallii 2
Panicum virgatum 4
Panicum virgatum 1
Oryza sativa 3
Setaria italica 2
Panicum hallii 1
Panicum virgatum 2
Panicum virgatum 3
Oryza sativa 1
Hordeum vulgare 3
Brachypodium stacei 2
Brachypodium stacei 6
Brachypodium distachyon 2
Brachypodium distachyon 6

B


Tree scale: 0.01
AtMYB75
AtMYB36
Panicum virgatum 5
Amborella trichopoda 4
Ananas comosus 6
Ananas comosus 5
Oryza sativa 4
Panicum hallii 4
Setaria italica 4
Sorghum bicolor 3
Brachypodium stacei 3
Brachypodium distachyon 4
AtMYB107
Brassica rapa 2
AtMYB9
Brassica rapa 5
Brassica oleracea 1
Brassica oleracea 3
Brassica oleracea 2
AtMYB41
Brassica oleracea 4
Amborella trichopoda 2
Oryza sativa 5
Sorghum bicolor 6
Brachypodium stacei 5
Hordeum vulgare 4
Triticum aestivum 7
Brachypodium distachyon 5
Amborella trichopoda 5
Amborella trichopoda 3
Sorghum bicolor 5
Ananas comosus 4
Brachypodium stacei 4
Sorghum bicolor 4
Musa acuminata 6
Amborella trichopoda 1
Ananas comosus 1
Brassica rapa 1
AtMYB93
Musa acuminata 4
Ananas comosus 2
Musa acuminata 5
Musa acuminata 1
Musa acuminata 3
AtMYB53
Musa acuminata 2
AtMYB92
Brassica rapa 4
Brassica rapa 6
Brassica rapa 3
Setaria italica 3
Zea mays 3
Sorghum bicolor 2
Zea mays 2
Brachypodium distachyon 3
Panicum hallii 3
Oryza sativa 2
Ananas comosus 3
Oryza sativa 3
Panicum virgatum 2
Panicum virgatum 3
Panicum hallii 1
Setaria italica 2
Hordeum vulgare 3
Oryza sativa 1
Brachypodium distachyon 2
Brachypodium stacei 2
Brachypodium stacei 6
Brachypodium distachyon 6
Sorghum bicolor 1
Zea mays 1
Setaria italica 1
Panicum virgatum 4
Panicum virgatum 1
Panicum hallii 2
Hordeum vulgare 2
Triticum aestivum 6
Triticum aestivum 5
Triticum aestivum 4
Triticum aestivum 3
Triticum aestivum 2
Triticum aestivum 1
Hordeum vulgare 1
Brachypodium stacei 1
Brachypodium distachyon 1

C


Tree scale: 0.01
AtMYB75
AtMYB36
Panicum virgatum 5
Amborella trichopoda 4
Amborella trichopoda 3
Ananas comosus 4
Sorghum bicolor 5
Brachypodium stacei 4
Sorghum bicolor 4
Ananas comosus 6
Ananas comosus 5
Oryza sativa 4
Panicum hallii 4
Setaria italica 4
Sorghum bicolor 3
Brachypodium stacei 3
Brachypodium distachyon 4
Amborella trichopoda 5
Brassica oleracea 3
Brassica oleracea 2
Amborella trichopoda 2
AtMYB41
Brassica oleracea 4
Brachypodium stacei 5
Brachypodium distachyon 5
Sorghum bicolor 6
Triticum aestivum 7
Hordeum vulgare 4
Oryza sativa 5
AtMYB107
Brassica rapa 2
AtMYB9
Brassica rapa 5
Brassica oleracea 1
Musa acuminata 6
Amborella trichopoda 1
Ananas comosus 1
Brassica rapa 1
AtMYB93
Musa acuminata 4
Ananas comosus 2
Brachypodium distachyon 3
Zea mays 3
Setaria italica 3
Panicum hallii 3
Oryza sativa 2
Sorghum bicolor 2
Zea mays 2
Musa acuminata 5
Musa acuminata 3
Musa acuminata 2
AtMYB53
Brassica rapa 4
AtMYB92
Brassica rapa 6
Brassica rapa 3
Musa acuminata 1
Ananas comosus 3
Setaria italica 2
Panicum virgatum 2
Panicum virgatum 3
Panicum hallii 1
Hordeum vulgare 3
Brachypodium stacei 2
Brachypodium distachyon 2
Oryza sativa 1
Brachypodium stacei 6
Brachypodium distachyon 6
Oryza sativa 3
Sorghum bicolor 1
Zea mays 1
Setaria italica 1
Panicum virgatum 4
Panicum virgatum 1
Panicum hallii 2
Brachypodium stacei 1
Brachypodium distachyon 1
Triticum aestivum 6
Triticum aestivum 5
Triticum aestivum 4
Triticum aestivum 3
Triticum aestivum 2
Triticum aestivum 1
Hordeum vulgare 2
Hordeum vulgare 1

D

Though the four trees are highly similar, the MUSCLE alignment with PhML algorithm (Figure 5.1.A) most closely matches the existing literature and was therefore considered to produce the most reliable phylogeny (Du *et al.*, 2015; Stracke *et al.*, 2001). The largest difference in the four trees is the positioning of the *AtMYB9* and *AtMYB107* clade. Both the MUSCLE/PhyML (Figure 5.1.A) and ClustalO/BioNJ (Figure 5.1.D) trees put this group as sharing a more recent common ancestor with the *AtMYB93* clade. This is similar to that seen in the ClustalO/PhyML (Figure 5.1.B) tree; however, an additional clade is closely linked with the *AtMYB9* and *AtMYB107* sequences. On the other hand, the MUSCLE/BioNJ (Figure 5.1.C) tree is the most dissimilar and positions *AtMYB9* and *AtMYB107* as more divergent than both *AtMYB93* and *AtMYB41*.

A number of the protein sequences identified from the pBLAST search were clustered within the region of the tree containing the more closely related MYBs but as a potentially separate group. A reciprocal best BLAST protein search of the two *A. trichopoda* proteins of this group evm_27.model.AmTr_v1.0_scaffold00010.533 (*A. trichopoda* 3) and the base of the group evm_27.model.AmTr_v1.0_scaffold00091.4 (*A. trichopoda* 4) was performed, which identified *AtMYB16* and *AtMYB43*, respectively, as putative homologues in *Arabidopsis*.

Another series of proteins were estimated to be more evolutionarily dissimilar than the closely related MYBs but were still positioned on the tree with a more recent common ancestor than the outgroup MYBs (*AtMYB36* and *AtMYB75*). This group also contained the fourth identified *O. sativa* protein (LOC_Os06g02250), which was therefore not considered to be closely related enough to be a rice *MYB93* homologue.

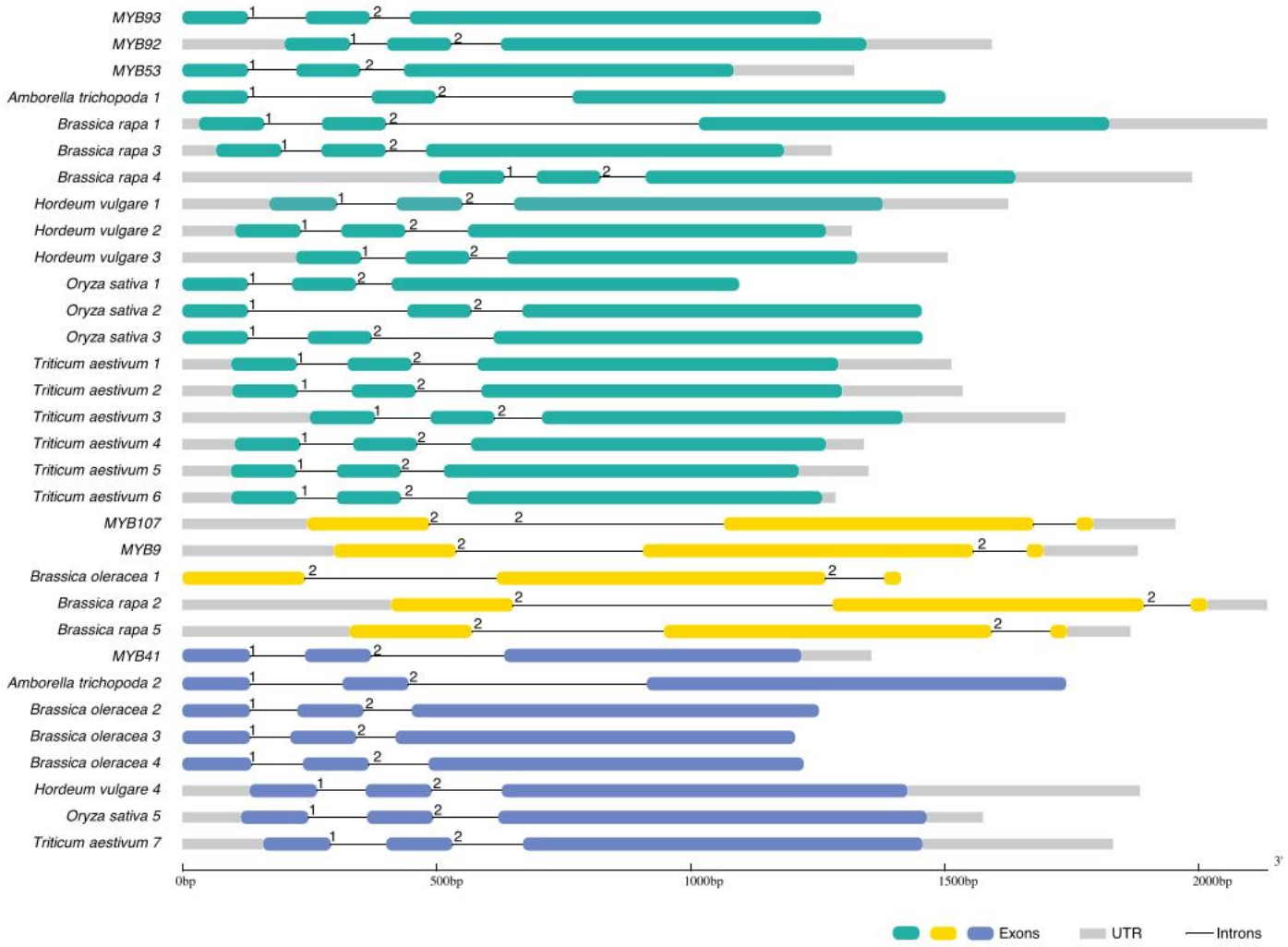
5.2.2 Intron analysis of MYB93 crop homologues

Non-coding DNA has historically been overlooked in favour of examining protein-coding regions, yet the advent of technologies better able to sequence and analyse whole genomes has enabled investigation into the functions and relevance of these regions of DNA (Palazzo & Gregory, 2014). Introns comprise the non-coding regions of genes and although transcribed, are removed from the final mRNA molecule by RNA splicing. The sequences of introns are much less highly conserved than exons, however the number, position and size of introns are more likely to be similar between closely related species or gene families (Carmel *et al.*, 2007). Thus, intron analysis presents another tool that can provide clues about the evolution of a gene. This may shed some light on the disparities between the previously published phylogenies (Du *et al.* (2015); Lashbrooke *et al.* (2016); Stracke *et al*. (2001)), and allow the selection of MYB homologues with greater confidence.

All of the potential MYB gene homologues analysed have three exons and two introns (Figure 5.2). *AtMYB9* and *AtMYB107* and their homologues are the most divergent in structure compared to the *AtMYB93* and *AtMYB41* groups. Namely, they are composed of a longer first and second exon and much shorter third exon, while the structure of the other genes is that of two shorter exons of similar length and a much longer third exon more similar in size to the second exon of the *AtMYB9* and *AtMYB107* homologues.

The conservation of intron positioning considers the phase of an intron. Phase zero means the intron does not interrupt a codon, whereas in phase one the intron is positioned between the first and second nucleotides of a codon, and in phase two the intron interrupts the second and third nucleotides. In all of the putative homologues of

Figure 5.2. Comparison of MYB homologue gene structure. *Box diagrams constructed from alignment of full length gDNA to CDS sequence using GSDS. Exons denoted by coloured boxes (green = MYB93 homologues, yellow = MYB9 and MYB107 homologues, purple = MYB41 homologues). UTR regions represented by grey boxes, introns denoted by black lines. Intron phase: 0, 1 or 2. All genes contain 3 exons and 2 introns.*


MYB93
MYB92
MYB53
Amborella trichopoda 1
Brassica rapa 1
Brassica rapa 3
Brassica rapa 4
Hordeum vulgare 1
Hordeum vulgare 2
Hordeum vulgare 3
Oryza sativa 1
Oryza sativa 2
Oryza sativa 3
Triticum aestivum 1
Triticum aestivum 2
Triticum aestivum 3
Triticum aestivum 4
Triticum aestivum 5
Triticum aestivum 6
MYB107
MYB9
Brassica oleracea 1
Brassica rapa 2
Brassica rapa 5
MYB41
Amborella trichopoda 2
Brassica oleracea 2
Brassica oleracea 3
Brassica oleracea 4
Hordeum vulgare 4
Oryza sativa 5
Triticum aestivum 7
0bp
500bp
1000bp
1500bp
2000bp
Exons
UTR
Introns

both *AtMYB93* and *AtMYB41* the first intron is in phase one, while the second is in phase two. In *AtMYB9* or *AtMYB107* homologues on the other hand, both introns are in phase two (Figure 5.2). The conservation of intron phase between *AtMYB93* and *AtMYB41* homologues, but not the *AtMYB9*/*AtMYB107* homologues indicates that *AtMYB93* and *AtMYB41* are likely to be more closely related than the other MYB relatives.

5.2.3 Predicted expression of putative MYB93 homologues in cereals

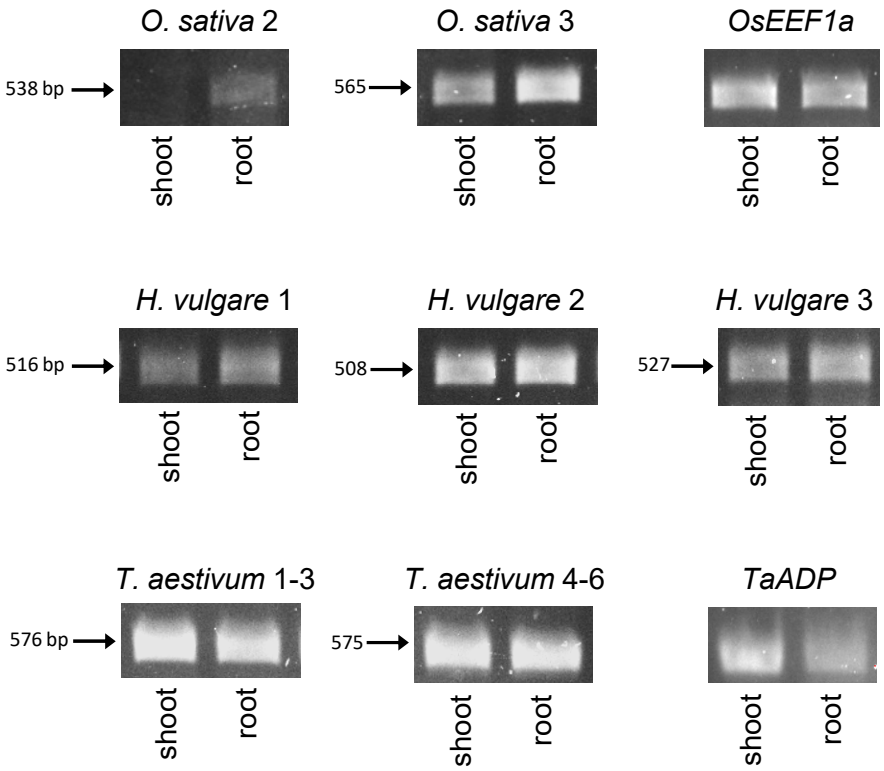
The next experiment aimed to look at the levels of gene expression in different tissue types for the putative homologues of wheat, barley and rice. Since *AtMYB93* is expressed solely in the roots it was hypothesised that any true homologues of *AtMYB93* in crops would also be root-specific, thus enabling the homologues identified through the phylogenetic analysis to be narrowed down further. The eFP browser contains limited expression data on species other than *Arabidopsis*, such as wheat, barley and rice, and so as a preliminary check these datasets were consulted for the putative MYB93 homologues (Winter *et al*., 2007). The barley and wheat databases proved unusable since the format of my gene identifiers is different to those used by eFP and there appears to be no way to convert these. Conversely, the rice data could be exploited and data were available for all five MYB homologues. Of the three putative MYB93 homologues, *O. sativa* 1 is highly expressed in the root and seed, with some expression in young leaves, LOC_Os08g37970 (*O. sativa* 2) expression is strong in roots and also in the panicle, but slightly weaker in seeds, and *O. sativa* 3 is enriched in the root and young leaf, with some expression also in the seed. *O. sativa* 4 was the rice protein that is expected to be the homologue of a

different closely related MYB, though not *AtMYB41*, *AtMYB9* or *AtMYB107.* While *O. sativa* 4 is expressed in the root, it is also more widely expressed than the putative MYB93 homologues and has the strongest expression in panicle tissue and the early developing seed. The likely MYB41 rice homologue, *O. sativa* 5, is also more widely expressed, though still with relatively strong root expression. The expression of the three putative *AtMYB93* homologues in the roots is as expected and although *AtMYB93* is not expressed in seeds, since it is considered less likely that rice possesses *AtMYB53* or *AtMYB92* homologues, it is possible that a rice *MYB93* could have a wider role and therefore be expressed in the seed. Evidence of expression in other tissues such as young leaves is more unexpected and could indicate a different function for these putative rice homologues.

5.2.4 RT-PCR expression of putative MYB93 homologues in cereals

Due to the time and space requirements of working on mature cereal crops, seedling tissue was used to compare expression in shoot and root tissue. Any homologues found to be root-specific could then be examined in a wider range of tissue types in mature plants. *TaADP* and *HvADP* – both ADP-ribosylation factors, and *OsEEF1a* – an elongation factor gene, were selected as housekeeping controls with stable expression across a range of tissues and conditions (Ferdous *et al.*, 2015; Jain *et al.*, 2006; Paolacci *et al.*, 2009). Of the three putative rice homologues, only *O. sativa* 2 was root-specific, as shown in Figure 5.3, providing initial evidence of this gene being a functional homologue. This figure also demonstrates that the three putative barley homologues are expressed in both root and shoot tissue and so are unlikely to be orthologous with *AtMYB93*. It should be noted that the *O. sativa* 1

Figure 5.3. RT-PCR analysis of putative MYB93 homologues in cereals. *RNA extracted from root and shoot tissue of four seedlings of wheat, barley or rice per biological replicate. Root specific expression is shown in rice homologue, Oryza sativa 2. OsEEF1a and TaADP were used as housekeeping controls for rice and wheat respectively. Note that expression of Oryza sativa 1 is not shown since results were inconclusive, and the barley housekeeping control (HvADP) is not shown because the primers did not successfully amplify a product.*


O. sativa 2
538 bp
shoot
root
O. sativa 3
565
shoot
root
OsEEF1a
shoot
root
H. vulgare 1
516 bp
shoot
root
H. vulgare 2
508
shoot
root
H. vulgare 3
527
shoot
root
T. aestivum 1-3
576 bp
shoot
root
T. aestivum 4-6
575
shoot
root
TaADP
shoot
root

primers failed to clearly amplify a product, however a faint band was present in the shoot sample (not shown). This is in line with the eFP data but optimisation of the RT-PCR conditions should be continued.

The situation is more complicated for wheat and further inspection of the six putative wheat homologues uncovered that they actually appear to be made up of two homologous genes, each with duplications on the A, B and D genomes that comprise cultivated wheat. Thus, Traes_6BL_0D6CD031E (*T. aestivum* 1) and Traes_6AL_2BA02CAA9 (*T. aestivum* 2) are revealed to be homeologs with Traes_6DL_8BF29D46C (*T. aestivum* 3), while Traes_6AL_52227091B (*T. aestivum* 4) and Traes_6BL_5A2DB8F78 (*T. aestivum* 5) are revealed to be homeologs with Traes_6DL_499F64524 (*T. aestivum* 6). Therefore, initially primers were designed to differentiate between these two sets of gene, but as seen from Figure 5.3 neither were found to be root-specific when amplified by RT-PCR. A recent study found that around 30% of homeolog triads (one each from the A, B and D genomes) exhibit differential expression across different tissue types (Ramírez-González *et al*., 2018) and so a new set of primers able to discriminate between each homeolog was designed. It was thought this could reveal whether any of the homeologs were actually root-specific, which would only be visible when amplified individually. RT-PCR analysis of these homeologs was inconclusive and requires more work. Additionally, a more in-depth analysis of these putative MYB93 homologues is required to clarify if the expression of any of these genes are enriched in the roots relative to other tissue types, if not root-specific.

5.3 Analysis of an AtARABIDILLO over-expressor in wheat

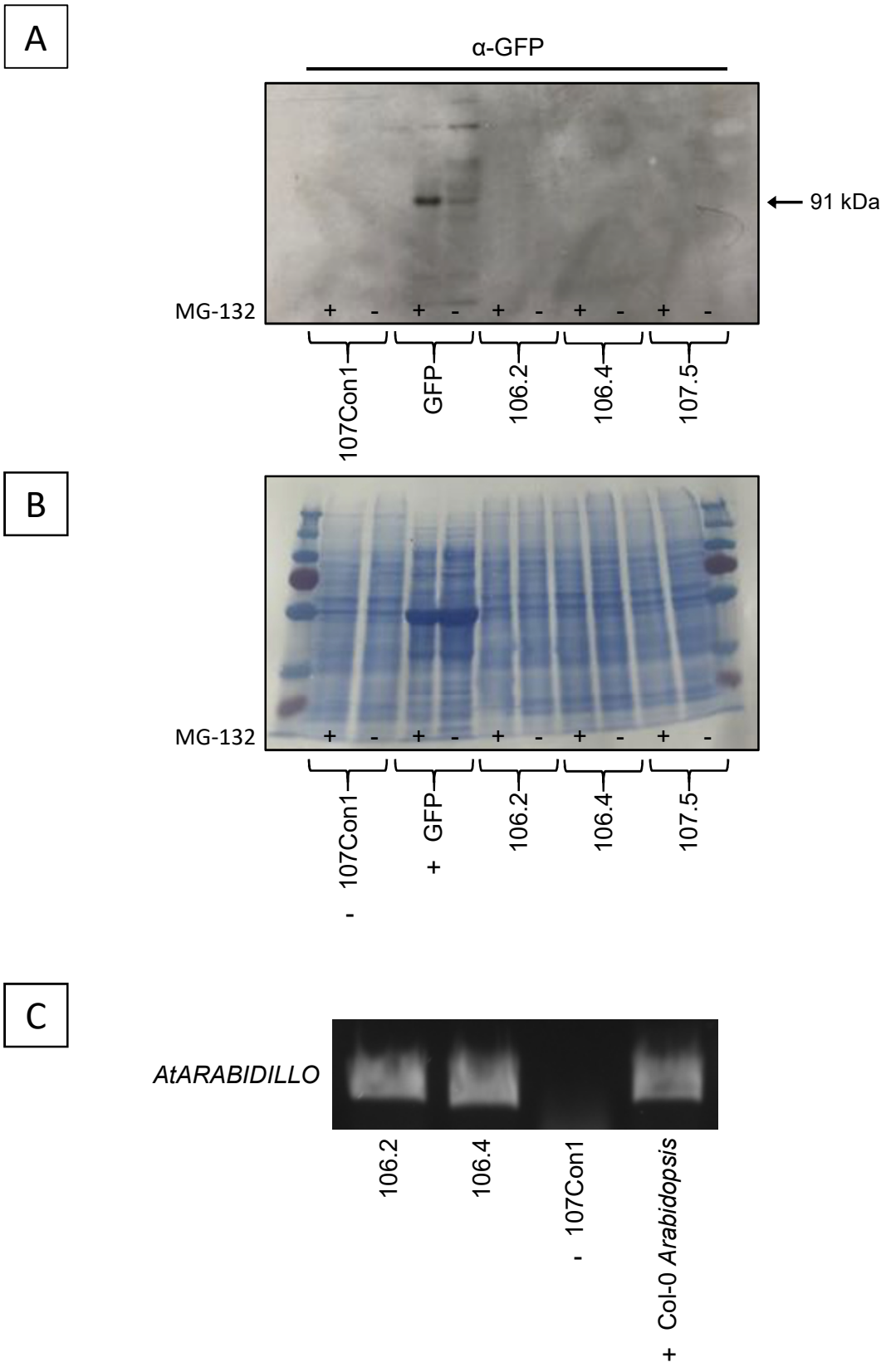
A series of transgenic wheat lines containing 35S::*AtARABIDILLO*-YFP were generated by the National Institute of Agricultural Botany (NIAB) through the Community Resource for Wheat Transformation project (NIAB, 2017). These constructs had been inserted into the wheat genome via the pEW324 vector which was partly constructed by J.C. Coates and M. Bailey and transformed by Emma Wallington and colleagues at NIAB. Two independent transformations produced a total of 31 lines alongside four control lines. It was anticipated that these *AtARABIDILLO* over-expressor lines would display a phenotype with increased LR density similar to that seen in *Atmyb93* mutant *Arabidopsis*. Complete knock out mutants are difficult to produce in wheat due to its hexaploid nature, but equally, the number of successfully inserted copies of the construct is likely to be highly variable between different lines of over-expressors. An estimate of the copy number in each line was provided, which was based on the number of copies of the *nptII* gene (conferring kanamycin resistance). This was able to approximate copy numbers between one and greater than four but could not differentiate higher copy numbers. The plan was to use Southern Blotting to accurately identify the number of inserted transgenes in each line and so DNA was extracted from mature leaf tissue samples of each line.

Initially, a selection of mutants containing an assortment of predicted copy numbers, alongside the relevant control lines, were selected for confocal microscopy. Unfortunately, no fluorescent signal for the YFP-tagged *At*ARABIDILLO protein could be detected in any of the lines. The next step was therefore to check the expression of the protein and this was performed by Western Blotting. Four lines were chosen:

106.2 to represent a line thought to contain greater than four transgene copies, 106.4 with 3 copies predicted, 107.5 with 2 copies estimated and 107.Con1 as a control line with no transgene inserted. Optimisation was required to see the appropriate band in the positive control – a GFP-tagged *Arabidopsis*, but ultimately, Western Blotting failed to positively identify YFP-tagged protein in any of the wheat lines. Treatment of seedlings with MG-132 (a proteasome inhibitor) was conducted in order to ascertain whether the proteins were being degraded by the proteasome, but Western Blotting suggested that this was not the case and that the protein was not being expressed at detectable levels (Figures 5.4.A and 5.4.B).

Following this, RT-PCR was carried out in order to determine if either the *AtARABIDILLO* gene or YFP tag were being transcribed. Successful amplification of the *AtARABIDILLO* gene is shown in Figure 5.4.C, while an RT-PCR spanning the YFP tag did not indicate any evidence of expression (data not shown). Amplification of the boundary across the YFP tag and *AtARABIDILLO* was also attempted, but once again failed to provide any confirmation of the presence of the YFP transcript (data not shown).

Figure 5.4. Western blot and RT-PCR analysis of 35S::AtARABIDILLO:YFP in wheat *(A) Western blot and (B) Coomassie stain of samples treated with MG-132 or an ethanol control. 107Con1 negative control and GFP positive control. (C) RT-PCR analysis of AtARABIDILLO in transgenic wheat lines. 107Con1 negative control and Col-0 Arabidopsis positive control. (Completed with Helen Wilkinson).*


A
α-GFP
91 kDa
MG-132
+ - + - + - + - + -
107Con1
GFP
106.2
106.4
107.5
B
MG-132
+ - + - + - + - + -
- 107Con1
+ GFP
106.2
106.4
107.5
C
AtARABIDILLO
106.2
106.4
- 107Con1
+ Col-0 Arabidopsis

5.4 Discussion

Several MYB homologues have been putatively identified through phylogenetic analysis, including three possible MYB93 homologues for the monocot cereals rice and barley, and six for wheat; later revealed to be two homeolog triplicates. A gene structure analysis of MYB homologues is in corroboration with previous phylogenies by Du *et al.* (2015) and Stracke *et al*. (2001), while dismissing the view of an alternative published phylogeny by Lashbrooke *et al*. (2016). Finally, RT-PCR analysis revealed root-specific expression in a putative MYB93 homologue in rice.

5.4.1 Phylogenetic analysis identified putative MYB homologues in monocots

Although it was necessary to identify gene homologues in this analysis using protein sequences, a possible caveat of this is that alternative splice variants might be misidentified as multiple, individual homologues in a species – i.e. gene duplications (Iñiguez & Hernández, 2017). In this analysis the protein sequences of some homologues identified for *Panicum virgatum* were almost 99% identical and so were pruned from the data as likely splice variants.

Sequence alignment quality and tree building algorithm both contribute towards the accuracy of a phylogenetic tree (Ogden & Rosenberg, 2006). Neighbour joining trees such as the BioNJ algorithm produce a single tree and make pairwise comparisons to establish relationships according to their genetic distance. This can overlook the finer details of an alignment. Conversely, maximum likelihood trees made using algorithms like PhyML produce many trees and aim to determine which is best according to their ‘likelihood’, but the trade-off is that they are much more

computationally taxing. The ClustalO alignment method has been shown to outperform MUSCLE, however it is slower and more computationally demanding (Pais *et al*., 2014). In the alignment produced for this phylogeny, it was found that the MUSCLE alignment method required fewer manual adjustments (moving a small number of amino acids to better align with the consensus) to be made to the output alignment than ClustalO. As these observations highlight, it is best to build trees using multiple methods as each has its own merits and can be compared in order help identify a consensus phylogeny. The corroboration between the different trees constructed during this project suggests that the monocot *MYB93* homologues have been identified correctly and helps validate the methods used and the resultant MYB homologues predicted.

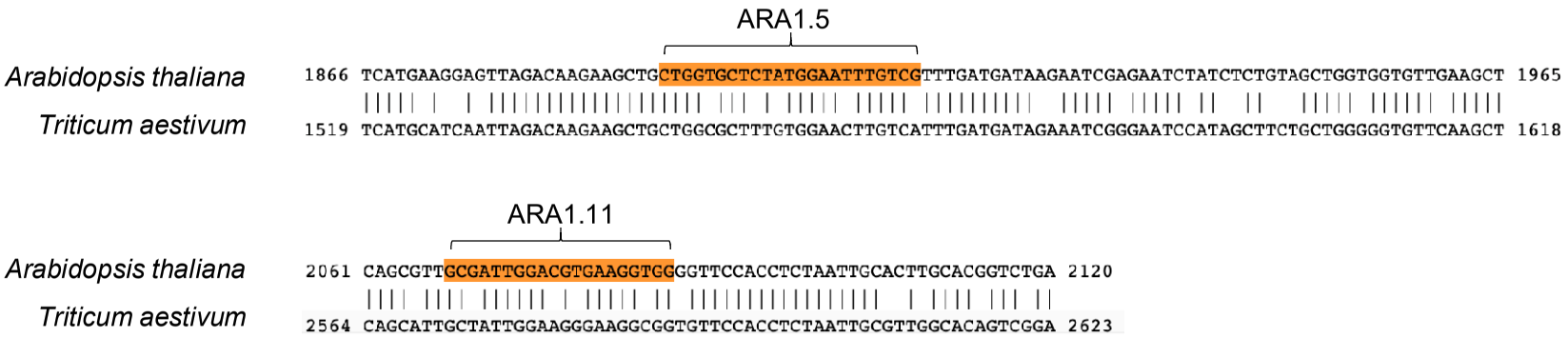
Gene structure analysis revealed information about the conservation of introns between MYB homologues. This included the intron phase, the point at which an intron intersects a codon, which is generally conserved (Long *et al.*, 1995). The shared intron phases between homologues of these MYBs suggest that *AtMYB9* and *AtMYB107* are indeed less closely related than *AtMYB93* and *AtMYB41*, and so therefore is in disagreement with the phylogeny by Lashbrooke *et al.* (2016). The *AtMYB93* and *AtMYB41* homologues have an asymmetrical exon, because the flanking introns do not match. This is important because if an exon translocates, is duplicated or deleted it will change the downstream reading frame of the DNA if its intron phase is not symmetrically aligned (Kolkman & Stemmer, 2001).

5.4.2 Transgenic wheat lines are not expressing YFP

The results of the confocal microscopy, Western Blot and RT-PCR are indicative of the successful insertion of a transgenic construct containing the *AtARABIDILLO* gene but suggest that the YFP tag cannot be successfully transcribed. Presence of the 35S promoter was not examined. The existence of ARABIDILLO but not YFP is puzzling as the coding sequence was confirmed by sequencing to be in frame in the construct (J.C. Coates & M. Bailey, personal communication). Crucially, a positive control was forgotten when amplifying YFP and it is essential that this is rectified.

Should a suitable positive control be successfully amplified there are other possible causes for the issues with the transgenic wheat. Both the wheat and *Arabidillo* protein sequences are very similar but it is feasible that there could still be a codon usage issue and that wheat is unable to read *Arabidopsis* DNA. It is also plausible that the primers designed to amplify *AtARABIDILLO* are not specific enough to avoid amplifying endogenous wheat sequence. Therefore, an alignment was made between wheat and the cDNA sequence of *AtARABIDILLO* and the primer sequences were sought for within the areas of highly conserved sequence. Three primers were in used to amplify *AtARABIDILLO*, and of these, two could be located (Figure 5.5) and although the sequence is partly conserved, it would be unlikely, but possible, that they could amplify the wheat genome. Raising the annealing temperature of the reaction could help increase the specificity of the reaction or designing new primers may be necessary. Alternatively, sequencing the amplification product would indisputably reveal which sequence is being amplified.

Figure 5.5. Alignment of AtARABIDILLO RT-PCR primers to wheat sequence. *BLAST alignment of AtARABIDILLO genomic sequence to wheat genomic sequence with ARA1.5 and ARA1.11 primer sequences highlighted in orange on the AtARABIDILLO sequence. Dashes between sequences indicate conserved nucleotides.*


ARA1.5
Arabidopsis thaliana
1866 TCATGAAGGAGTTAGACAAGAAGCTGCTGGTGCTCTATGGAATTTGTCGTTTGATGATAAGAATCGAGAATCTATCTCTGTAGCTGGTGGTGTTGAAGCT 1965
Triticum aestivum
1519 TCATGCATCAATTAGACAAGAAGCTGCTGGCGCTTTGTGGAACTTGTCATTTGATGATAGAAATCGGGAATCCATAGCTTCTGCTGGGGGTGTTCAAGCT 1618
ARA1.11
Arabidopsis thaliana
2061 CAGCGTTGCGATTGGACGTGAAGGTGGGGTTCCACCTCTAATTGCACTTGCACGGTCTGA 2120
Triticum aestivum
2564 CAGCATTGCTATTGGAAGGGAAGGCGGTGTTCCACCTCTAATTGCGTTGGCACAGTCGGA 2623

5.5 Conclusion

The data presented in this chapter has revealed a number of potential MYB93 homologues in monocots though these may not all be functional homologues. It has also corroborated the phylogenies placing the *AtMYB107* and *AtMYB9* in a separate clade to *AtMYB93*, *AtMYB92* and *AtMYB53* and maintains the separation of *AtMYB41* into a neighbouring clade.

Chapter VI

GENERAL DISCUSSION

6.1 Introduction

*At*MYB93 is a TF previously shown to be involved in a complex network for the regulation of LR development, though the mechanism of action of *At*MYB93 remained unknown. The principal aims of this work were to: (i) understand how *At*MYB93 functions at the molecular level, (ii) establish how modifying *At*MYB93 affects LR development and formation of the root barrier, and therefore ultimately plant stress resilience, and (iii) uncover ways in which the results of this work could help tackle the issue of global food security.

To address these aims, the project had the following objectives:

(i) To identify the downstream transcriptional targets and upstream regulators of *At*MYB93,

(ii) To characterise the physiological characteristics of *Atmyb93*-clade mutants and mutants in identified regulators/targets identified in objective (i).

(iii) To identify, characterise and ultimately manipulate *MYB93* homologues in crops to determine whether they could provide a promising opportunity for crop improvement.

In the sections below, the findings presented in the results chapters of this thesis are put into the context of the general research area and next steps are discussed.

6.2 Downstream targets of AtMYB93

Though no downstream targets have been conclusively identified, presented in this thesis is a list of promising candidates that merit further study. The data

provided here, when taken with the mounting evidence characterising a number of these genes or their closest relatives, further substantiates their suggested roles in LR development and the possibility of discovering one of them as a direct target of *At*MYB93.

The presence of many endodermal and cell wall specific genes upregulated at the same time as *AtMYB93* (Voß *et al.*, 2015) poses one of the biggest unanswered questions. Is the association because some are direct targets of *At*MYB93, or is the presence of these genes merely correlative because the expression measurements are taken at the point of LR initiation? For example, many Casparian strip related genes are downregulated at the time of significant upregulation of *AtMYB93*, which could be in response to the emerging LR itself. Thus, experiments crucial to furthering this work remain to be completed and involve demonstrating the direct binding of *At*MYB93 to its transcriptional targets. Ideally this would be via ChIP-seq as, unlike ChIP-qPCR this would pull down everything that *At*MYB93 binds to, and unlike Y1H, demonstrates binding *in planta*. Should problems persist in creating transgenic plants containing the *pMYB93*::myc-*AtgMYB93* constructs, transgenic plants already exist that contain a different *At*MYB93-based construct: *35S::*myc-*AtcMYB93* in a *Atmyb93* mutant background (Gibbs *et al.*, 2014). This construct is not ideal for the experiment, because ectopic expression is likely from its overexpression and the cDNA sequence is used in place of genomic DNA, however this could prove a starting point for confirming my candidate *At*MYB93 targets via ChIP-qPCR.

Another important aspect that remains to be considered is the reason behind the mismatch of genes identified by the different experiments and datasets explored

in this thesis. For example, *AtABCG2* and *AtS-ACP-DES1* both exhibited LR phenotypes but were not identified by either qRT-PCR or RNA-seq discovery methods. I propose that this could be down to their relatively low expression levels being masked by the background 'noise' created by using whole seedlings for qRT-PCR or whole roots for RNA-seq. This narrative also seems more likely when considering that more of the *At*MYB93 candidate genes were found in the top 1000 significant genes from the RNA-seq experiment than when these data were ranked by differences in fold-change expression. Future work mitigating this issue could include another RNA-seq experiment performed using FACS, since the specificity of experiments such as the microarray by Voß *et al.* (2015) are especially useful when considering genes with expression patterns as restricted as *AtMYB93*.

6.3 Upstream repression of AtMYB93 by AtSCARECROW

Experiments in this project built on previous research that revealed a direct interaction between *At*SCR and the *AtMYB93* promoter (Iyer-Pascuzzi *et al.*, 2011; Sparks *et al.*, 2016) and consequently demonstrate *At*SCR as a probable repressor of *AtMYB93* expression in the endodermis. The early developmental functions of *AtSCR* in differentiating cell types to create the endodermis make KO mutant studies more difficult. This is potentially problematic even with the use of a less severe mutant like *Atscr-3* (Fukaki *et al.*, 1996). Hence, future study would benefit from experiments allowing *AtSCR* expression in the endodermis to be knocked out or reduced at a later point during development, which would be possible with the use of an RNA interference (RNAi) system with an inducible promoter such as dex or oestradiol (Borghi, 2010). Alternatively, CRISPR/Cas9 provides greater precision in

gene editing and more control over gene expression levels (Belhaj *et al.*, 2013). Another avenue is to overexpress *AtSCR* in the endodermis, something which already appears to be happening to some extent in the *pSCR*::*SCR*-GFP;*Atscr-3* line, in turn reducing *AtMYB93* expression and so leading to an increased LR density.

6.4 Close relatives of AtMYB93

This section of work looked at characterising the root and germination phenotypes of a series of KO mutants made for *AtMYB93* and its close relatives *AtMYB92* and *AtMYB53*. These experiments provided perhaps one of the more surprising results of the project, given the lack of a discernible LR phenotype or more obvious germination phenotype for the *Atmyb53-1/Atmyb92-1/Atmyb93-1* and *Atmyb53-2/Atmyb92-1/Atmyb93-1* triple mutants. Previous research has demonstrated that reduced suberin in the seed coat leads to a distorted germination phenotype (Beisson *et al.*, 2007; Fedi *et al.*, 2017; Yadav *et al.*, 2014) and more permeable roots (Yadav *et al.*, 2014). Consequently, functional redundancy was insinuated to be at play. *AtMYB41* was originally suggested to be involved in this role because it is associated not only with suberin biosynthesis in seeds, but also in stress response. However, semi-quantitative RT-PCR hints that although upregulation is likely present in the triple MYB mutants, it is subtle. A quantitative experiment is required, such as qRT-PCR and there are other possible candidates that should be considered for future experiments, namely *AtMYB107* and *AtMYB9* – two other seed related MYBs with roles in suberin biosynthesis.

Developments of the work on sulfur and salt stress should also continue. These could include elemental analysis of shoot tissue in mutants grown under stress for the other S24 clade MYBs and also for some of the candidate targets of

*At*MYB93. For example, Yadav *et al.* (2014) revealed an increase in solute permeability to salt in their *Atabcg2-1/Atabcg6-1/Atabcg20-1* triple mutant compared to WT roots, which was measured using a pressure probe and in conditions with the addition of NaCl or KNO_3. Moreover, while the use of elemental analysis has provided a new way to phenotype the *Atmyb93* and associated mutant phenotypes, trialling different ways to assess their root architecture could prove valuable, particularly when considering that primary root lengths were sometimes more significantly different than differences in LR density. However, most automated software fails to provide satisfactory results for *Arabidopsis* or when using large numbers of seedlings. An in-depth assessment of the capabilities of currently available software should be made and additionally, photographs can be revisited for novel analysis in the event of new software developments.

6.5 MYB93 crop homologues

Putative MYB homologues for MYB93 and its close relatives were identified by phylogenetic analysis in a variety of species, with a focus on monocots. RT-PCR analysis of the potential MYB93 homologues in the cereals of wheat, barley and rice, showed most were expressed in both root and shoot seedling tissue. It seems unlikely that that a functional homologue of MYB93 would be expressed in seedling shoot tissue, given that MYB93 homologues in other species are expressed in the roots and absent from leaf tissue (X. Cao, unpublished; Du *et al.*, 2015; Legay *et al.*, 2016), though it is possible that these genes could be expressed more widely than the roots, such as in the seed. Furthermore, the expression profiles of MYB93 homologues in apple and tomato have revealed a broader expression pattern in the

fruit of these species (Legay *et al.*, 2016; X. Cao, unpublished). Although *AtMYB93* is not expressed in the 'fruit' of *Arabidopsis* – defined as the mature ovary of a plant, which in the case of *Arabidopsis* arises from the development of the fertilised gynoecium into a seed filled silique (Roeder & Yanofsky, 2006) – the presence of MYB93 homologue expression in other crops has potentially significant implications in future plant breeding projects. On the other hand, the *At*MYB93 interactor *At*ARABIDILLO is found in the fruit and so perhaps a similar, closely related MYB could be expressed in this tissue. For example, Gibbs *et al*. (2014) noted *AtMYB92* expression in *Arabidopsis* siliques.

According to the rice eFP browser, the putative MYB93 homologue encoded by LOC_Os08g37970 (*O. sativa* 2) was found to be expressed not only in the roots and seed, but also in late stage panicle development. RT-PCR also confirmed the expression of this potential *OsMYB93* in the roots of seedlings. Though panicle development is probably too early to be considered the "fruit" stage of rice, it would be interesting to see the effects of rice mutant in this gene, as this could have consequences for grain development. Furthermore, the potential role of this gene on LR density and germination would shed some light on exposing LOC_Os08g37970 as a MYB93 orthologue in an important monocot crop.

The initial next step could be to grow mature rice plants and confirm the expression of the putative MYB93 homologues in other tissues, particularly to check the panicle-based expression of LOC_Os08g37970. Further study on the potential MYB93 homologues of wheat and barley is also necessary and qRT-PCR could be used to show root enrichment of some of these genes. Subsequent to this, the CRISPR/Cas9 should be considered for the generation of mutants in these cereals in

order to assess their phenotypes. It can prove very difficult to create transgenic plants using traditional methods in complex species such as hexaploid wheat. Therefore, as a trade-off the *At*ARABIDILLO over-expressor wheat lines were created. More recently developed genome editing technologies such as the CRISPR/Cas9 system could instead prove less complicated for producing mutants in these species (Belhaj *et al.*, 2013; Ferrie *et al.*, 2019; Zhang *et al.*, 2019)

6.6 Final remarks

The work presented in this thesis has begun to answer questions about *AtMYB93* and its regulatory network by identifying *AtSCARECROW* as an upstream regulator of *AtMYB93,* as well as several putative *At*MYB93 downstream targets. It has also started to shape our understanding of the responses of *AtMYB93* and its S24 clade members, *AtMYB92* and *AtMYB53* to sulfur and salt stress, and hints that functional redundancy plays a role. Finally, it has initiated work on MYB93 homologues in globally important cereal crops. Continued work in this area will reveal whether MYB93 manipulation could be used for crop improvement.

The initial focus on future work should be to continue the ongoing experiments looking for direct interactions between *At*MYB93 and the putative downstream targets, with priority given to *AtPME10, AtABCG2 and AtMYB54*. Additionally, the rest of the candidate gene mutants should be assessed for LR phenotypes, as this would provide valuable data even for those genes that are not direct targets of *At*MYB93. Further study on *AtSCR* should also be prioritised, given its interesting potential role as an upstream *AtMYB93* repressor and in the first instance experiments should aim to confirm localisation of *AtSCR* expression *in vitro.*

Chapter VII

REFERENCES

Alman, M., Kinet, J-M., & Lutts, S. (2001) Effect of salt and osmotic stresses on germination in durum wheat (*Triticum durum* Desf.) *Plant Soil.* **231**(2):243-254

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) "Basic local alignment search tool" *J. Mol. Biol.*, **215**:403-410

Anderson, T., Barberon, M. & Geldner, N. (2015) Suberization – the second life of an endodermal cell. *Curr. Opin. Plant Biol.,* **28**:9-15

Anjum, N.A., Gill, R., Kaushik, M., Hasanuzzaman, M., Pereira, E. & Ahmad, I. *et al*., (2015) ATP-sulfurylase, sulfur-compounds, and plant stress tolerance. *Front. Plant Sci.*, **6**:1-9

Applied Biosystems (undated) *Tm Calculator*. [online] Available at: <http://www6.appliedbiosystems.com/support/techtools/calc/index.cfm>

Arif, M.R., Islam, T. & Robin, A.H.K. (2019) Salinity stress alters root morphology and root hair traits in *Brassica napus*. *Plants*, **8**(7):192

Arnell, N.W. (2004) Climate change and global water resources: SRES emissions and socio-economic scenarios. *Global Environ. Chang.*, **14**(1), 31–52

Bach, L., Gissot, L., Marion, J., Tellier, F., Moreau, P. & Satiat-Jeunemaître, B. *et al.* (2013) Very-long-chain fatty acids are required for cell plate formation during cytokinesis in *Arabidopsis thaliana*. *J. Cell Sci.*, **124**:3223-3234

Bak S., Beisson F., Bishop G., Hamberger B., Höfer R. & Paquette S. *et al.* (2011) Cytochromes P450. *Arabidopsis* Book, **21**(4):770-782

Bao, F., Shen, J., Brady, S.R., Muday, G.K., Asami, T. & Yang, Z. (2004) Brassinosteroids interact with auxin to promote lateral root development in *Arabidopsis*. *Plant Physiol.*, **134**(4):1624-1631

Barberon, M., Vermeer, J.E.M., Bellis, D., Wang, P., Naseer, S. & Anderson, T.G. *et al*. (2016) Adaptation of root function by nutrient-induced plasticity of endodermal differentiation. *Cell*, **164**:447-459

Bari, R. & Jones, J.D.G. (2009) Role of plant hormones in plant defence responses. *Plant Mol. Biol.*, **69**(4):473-488

Baxter, I., Hosmani, P.S., Rus, A., Lahner, B., Borevitz, J.O. & Muthukumar, B. *et al*. (2009) Root suberin forms an extracellular barrier that affects water relations and mineral nutrition in *Arabidopsis*. *PLoS Genet*., **5**(5):1-12

Beisson, F., Li, Y.H., Bonaventure, G., Pollard, M. & Ohlrogge, J.B. (2007). The acyltransferase GPAT5 is required for the synthesis of suberin in seed coat and root of *Arabidopsis*. *Plant Cell*, **19**:351–368

Beisson, F., Li-Biesson, Y. & Pollard, M. (2012) Solving the puzzles of cutin and suberin polymer biosynthesis. *Curr. Opin. Plant Biol.*, **15**(3):329-337

Belhaj, K., Chaparro-Garcia, A., Kamoun, S. & Nekrasov, V. (2013) Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. *Plant Methods*, **9**:39

Bellati, J., Champeyroux, C., Hem, S., Rofidal, V., Krouk, G. & Maurel, C. *et al.* (2016) Novel aquaporin regulatory mechanisms revealed by interactomics. *Mol. Cell. Proteom.*, **15**:3473-3487

Benitez-Alfonso, Y., Faulkner, C., Pendle, A., Miyashima, S., Helariutta, Y., & Maule, A. (2013) Symplastic intercellular connectivity regulates lateral root patterning. *Dev. Cell*, **26**:136-147

Benková, E. and Hejátko, J. (2009) Hormone interactions at the root apical meristem. *Plant Mol. Biol.*, **69**:383-396

Benková, E., Michniewicz, M., Sauer, M., Teichmann, T., Seifertová, D. Jürgens, G. & Friml, J. (2003) Local, efflux-dependent auxin gradients as a common module for plant organ formation. *Cell*, **115**:591-602

Bentsink, L. & Koornneef, M. (2008) Seed dormancy and germination. *Arabidopsis Book*, **6**

Benzle, K. & Cornish, K. (2017) Improved axenic hydroponic whole plant propagation for rapid production of roots as transformation target tissue. *Plant Methods*, **13**:37

Berardini, T.Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R. & Strait, E. *et al.* (2015) The Arabidopsis Information Resource: Making and mining the “gold standard” annotated reference plant genome. *genesis*, **53**:474-485

Bernards, M.A. (2002) Demystifying suberin. *Canad. J. Bot.*, **80**:227-240

Berry, P. *et al.* (2014) HGCA Oilseed Rape Guide (UK Agriculture and Horticulture Development Board). Retrieved from https://cereals.ahdb.org.uk/media/305093/g55-oilseed-rape-guide-jan-2014-update.pdf

Bielecka, M., Watanabe, M., Morcuende, R., Scheible, W.R., Hawkesford, M.J. & Hesse, H. *et al*. (2015) Transcriptome and metabolome analysis of plant sulfate starvation and resupply provides novel information on transcriptional regulation of metabolism associated with sulfur, nitrogen and phosphorus nutritional responses in *Arabidopsis*. *Front. Plant Sci.*, **5**:1-18

Bolingue, W., Rosnoblet, C., Leprince, O., Vu, B.L., Aubry, C. & Buitink, J. (2010) The *MtSNF4b* subunit of the sucrose non-fermenting-related kinase complex connects after-ripening and constitutive defense responses in seeds of *Medicago truncatula*. *Plant J.*, **61**:792-803

Bo, H., Jinpu, J., Guo, A., Zhang, H., Luo, J. & Gao, Ge. (2015). GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics*, **31**(8):1296-1297

Borghi, L. (2010) Inducible gene expression systems for plants. *Methods Mol. Biol.*, **655**:65-75.

Bustos, R., Castrillo, G., Linhares, F., Puga, M.I., Rubio, V. & Pérez-Pérez, J. *et al.* (2010) A central regulatory system largely controls transcriptional activation and repression responses to phosphate starvation in *Arabidopsis*. *PLoS Genet.*, **6**(9)

Cakmark, I. & Yazici, A.M. (2010) Magnesium: a forgotten element in crop production. *Better Crops*, **94**(2):23-25

Carmel, L., Rogozin, I.B., Wolf, Y.I. & Koonin, E.V. (2007) Patterns of intron gain and conservation in eukaryotic genes. *BMC Evol. Biol.*, **7**(192)

Casimiro, I., Marchant, A., Bhalerao, R.P., Beeckman, T., Dhooge, D. & Swarup, R. *et al.* (2001) Auxin transport promotes *Arabidopsis* lateral root initiation. *Plant Cell*, **13**:843-852

Champeyroux, C., Bellati, J., Baberon, M., Rofidal, V., Maurel, C. & Santoni, V. (2019) Regulation of a plant aquaporin by a Casparian strip membrane domain protein-like. *Plant Cell Environ.*, **42**(6):1788-1801

Chen, C.C., Fu, S.F., Lee, Y.I., Lin, C.Y., Lin, W.C. & Huang, H.J. Transcriptome analysis of age-related gain of callus-forming capacity in *Arabidopsis* hypocotyls. *Plant Cell Phys.*, **53**(8):1457-1469

Cheng, Y.E., Dong, M.Y., Fan, X.W., Nong, L.L. & Li, Y.Z. (2018) A study on cassava tolerance to and growth responses under salt stress. *Environ. Exp. Bot.*, **155**:429-440

Chen, R. Rosen, E. & Masson, P.H. (1999) Gravitropism in higher plants. *Plant Physiol.*, **120**:343-350

Christmann, A., Moes, D., Himmelbach, A., Yang, Y., Tang., Y. & Grill, E. (2006) Integration of abscisic acid signalling into plant responses. *Plant Biol.*, **8**(3):314-325

Clough, S.J. & Bent, A.F. (2008) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.*, **16**(6):735-743

Coates, J.C., Laplaze, L. & Haseloff, J. (2006) Armadillo-related proteins promote lateral root development in *Arabidopsis*. *PNAS*, **103**(5):1621-1626

Cominelli, E., Sala, T., Calvi, D., Gusmaroli, G. & Tonelli, C. (2008) Over-expression of the *Arabidopsis AtMYB41* gene alters cell expansion and leaf surface permeability. *Plant J.*, **53**:53-64

Crooks, G.E., Hon, G., Chandonia, J. & Brenner, S.E. (2004) WebLogo: A sequence logo generator. *Genome Res.*, **14**:1188-1190

Cui, H., Levesque, M.P., Vernoux, T., Jung, J.W., Paquette, A.J. & Gallagher, K.L. *et al.* (2007) An evolutionarily conserved mechanism delimiting SHR movement defines a single layer of endodermis in plants.

Czechowski, T. (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. *Plant Physiol.,* **139**(1):5-17

Dan, H., Yang, G. & Zheng, Z.L. (2007) A negative regulatory role for auxin in sulphate deficiency response in *Arabidopsis thaliana*. *Plant Mol. Biol.*, **63**(2)**:**221-235

de Dorlodot, S., Forster, B., Pagès, L., Price, A., Tuberosa, R. & Draye, X. (2007) Root system architecture: opportunities and constraints for genetic improvement of crops. *Trends Plant Sci.*, **12**(10):474-481

De Smet, I., Vanneste, S., Inze, D. & Beeckman T. (2006) Lateral root initiation or the birth of a new meristem. *Plant Mol. Biol.,* **60**:871-887

Dehaan, R.L. & Taylor, G.R. (2002) Field-derived spectra of salinized soils and vegetation as indicators of irrigation-induced soil salinization. *Remote Sens. Environ.*, **80**(3):406-417

Den Herder, G., Van Isterdael, G., Beeckman, T. & De Smet, I. (2010) The roots of a new green revolution. *Trends Plant Sci.*, **15**:600-607

D'hose, T., Cougnon, M., Vliegher, A.D., Vandercasteele, B., Viaene, N. & Cornelis, W. *et al.* (2014) The positive relationship between soil quality and crop production: A case study on the effect of farm compost application. *Appl. Soil Ecol.*, **75**:189-198

Dijkshoorn, W. & van Wijk, A.L. (1967) The sulphur requirements of plants as evidenced by the sulphur-nitrogen ratio in the organic matter a review of published data. *Plant Soil,* **26**(1):129-157

Di Laurenzio, L. Wysocka-Diller, J., Malamy, J.E., Pysh, L., Helariutta, Y. & Freshour, G. *et al.* (1996) The SCARECROW gene regulates an asymmetric cell division that is essential for generating the radial organization of the *Arabidopsis* root. *Cell*, **86**(3):423-433

Dinneny, J.R., Long, T.A., Wang, J.Y., Jung, J.W., Mace, D. & Pointer, S. *et al.* (2008) Cell identity mediates the response of *Arabidopsis* roots to abiotic stress. *Science*, **320**(5878):942-945

Dolan, L., Janmaat, K., Willemsen, V., Linstead, P., Poethig, S. & Roberts, K. *et al.* (1993) Cellular organisation of the *Arabidopsis thaliana* root. *Development*, **119**(1):71-84

Domergue, F., Vishwanath, S.J., Joubès, J., Ono, J., Lee, J.A. & Bourdon, M. *et al*. (2010) Three *Arabidopsis* fatty acyl-coenzyme A reductases, FAR1, FAR4, and FAR5, generate primary fatty alcohols associated with suberin deposition. *Plant Physiol.*, **153**(4):1539-1554

Duan, L., Dietrich, D., Ng, C.H. Chan, P.M.Y., Bhalerao, R. & Bennett, M.J. *et al.* (2013) Endodermal ABA signalling promotes lateral root quiescence during salt stress in *Arabidopsis* seedlings. *Plant Cell*, **25**:324-341

Du, H., Liang, Z., Zhao, S., Nan, M.G., Tran, L.S. & Lu, K. (2015) The evolutionary history of R2R3-MYB proteins across 50 eukaryotes: new insights into subfamily classification and expansion. *Sci. Rep.*, **5**:11037

Dubrovsky, J.G., Doerner, P.W., Colon-Carmona, A. & Rost, T.L. (2000) Pericycle cell proliferation and lateral root initiation in *Arabidopsis*. *Plant Physiol*., **124**:1648–1657

Dubrovsky, J.G., Rost, T.L., Colón-Carmona, A. & Doerner, P. (2001) Early primordium morphogenesis during lateral root initiation in *Arabidopsis thaliana*. *Planta*, **214**(1):30-36

Dufresne, J.L., Foujols, M.A., Denvil, S., Caubel, A., Marti, O. & Aumont, O. *et al.* (2013) Climate change projections using the IPSL-CM5 earth system model: From CMIP3 to CMIP5. *Climate Dyn*., **40**:2123–2165.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.*, **35**(5):1792-1797

Esau, K. (1940) Development anatomy of the fleshy storage organ of Daucus carota. *Hilgardia*, **13**:177–208

Evenson, R.E. & Gollin, D. (2003) Assessing the impact of the Green Revolution, 1960-2000. *Science*, **300**:758-762

Fan, M., Shen, J., Yuan, L., Jiang, R., Chen, X. & Davies, W.J. *et al.* (2012) Improving crop productivity and resource use efficiency to ensure food security and environmental quality in China. *J. Exp. Bot.*, **63**(1):13-24

FAOSTAT (2018) Food and agricultural commodities production: commodities by regions. Retrieved Aug. 18, 2014 from http://faostat3.fao.org/browse/rankings/commodities_by_regions/E

Fedi, F., O'Neill, C.M., Menard, G., Trick, M., Dechirico, S. & Corbineau, F. *et al.* (2017) Awake1, an ABC-type transporter, reveals an essential role for suberin in the control of seed dormancy. *Plant Physiol.*, **174**:276-283

Feng, C.P., Andreasson, E., Maslak, A., Mock, H.P., Mattsson, O. & Mundy, J. (2004) *Arabidopsis* MYB68 in development and responses to environmental cues. *Plant Sci.*, **167**(5):1099-1107

Ferdous, J., Li, Y., Reid, N., Langridge, P., Shi, B. & Tricker., P.J. (2015) Identification of Reference Genes for Quantitative Analysis of MicroRNAs and mRNAs in Barley under Various Stress Conditions. *PLoS One*, **10**(2)

Ferrie, A.M.R., Bhowmik, P., Rajagopalan, N. & Kagale (2019) CRISPR/Cas9-mediated targeted mutagenesis in wheat doubled haploids. *Methods Mol. Biol.*, **2072**:183-198

Food and Agriculture Organisation. (1996) Rome Declaration on World Food Security and World Food Summit Plan of Action. World Food Summit 13-17 November 1996. Rome, accessed from http://www.fao.org/3/w3613e/w3613e00.htm on 20/09/19

Food and Agriculture Organisation (2009) How to feed the world 2050, accessed from http://www.fao.org/fileadmin/templates/wsfs/docs/expert_paper/How_to_Feed_the_World_in_2050.pdf on 20/09/19

Food and Agriculture Organisation, International Fund for Agricultural Development & World Food Program. (2015) The State of Food Insecurity in the World 2015. Meeting the 2015 international hunger targets: taking stock of uneven progress. Rome, FAO

Food and Agriculture Organisation (2017) The future of food and agriculture – Trends and challenges: Rome: FAO, accessed from http://www.fao.org/3/a-i6583e.pdf on 20/09/19

Food and Agriculture Organisation & Intergovernmental Technical Panel on Soils (2015) Status of the World's Soil Resources (SWSR) – Main Report. Rome: Food and Agriculture Organization of the United Nations.

Francis, K.E., Lam, S.Y. & Copenhaver, G.P. (2006) Separation of *Arabidopsis* pollen tetrads is regulated by QUARTET1, a pectin methylesterase gene. *Plant Physiol.*, **142**:1004-1013

Friml, J. (2003) Auxin transport – shaping the plant. *Curr. Opin. Plant. Bio.*, **6**(1):7-12

Fukaki, H., Fujisawa, H. & Tasaka, M. (1996) SGR1, SGR2 and SGR3: novel genetic loci involved in shoot gravitropism in *Arabidopsis thaliana*. *Plant Physiol*., **110**:945–955

Fukaki, H., Wysocka-Diller, J., Kato, T., Fujisawa, H., Benfey, P.N. & Tasaka, M. (1998) Genetic evidence that the endodermis is essential for shoot gravitropism in *Arabidopsis thaliana*. *Plant J*., **14**:425–430

Fu, X. & Harberd, N.P. (2003) Auxin promotes *Arabidopsis* root growth by modulating gibberellin response. *Nature*, **421**:740-743

Gallagha, K.L., Paquette, A.J., Nakajima, K. & Benfet, P.N. (2004) Mechanisms regulating SHORT-ROOT intercellular movement. *Curr. Biol.*, **14**:1847-1851

Galloway, J.N., Aber, J.D., Erisman, J.W., Seitzinger, S.P., Howarth, R.W. & Cowling, E.B. *et al.* (2003) The nitrogen cascade. *Bioscience*, **53**(4):341-356

Galvan-Ampudia, C.S., Julkowska, M.M., Darwish, E., Gandullo, J., Korver, R.A. & Brunoud, G. *et al.* (2013) Halotropism is a response of plant roots to avoid a saline environment. *Curr. Biol.*, **23**(20):2044-2050.

Gascuel, O. (1997) BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Mol. Biol. Evol.*, **14**(7):685-695

Ghaderiardakani, F., Collas, E., Damiano, D.K., Tagg, K., Graham, N.S. & Coates, J.C. (2019) Effects of green seaweed extract on *Arabidopsis* early development suggest roles for hormone signalling in plant responses to algal fertilisers. *Sci. Rep.*, **9**

Gibbs, D.J. & Coates, J.C. (2014) *AtMYB93* is an endodermis-specific transcriptional regulator of lateral root development in *Arabidopsis*. *Plant Signal. Behav*., **9**(10):1-4

Gibbs, D.J., Lee, S.C., Isa, N.M., Gramuglia, S., Fukao, T. & Bassel, G.W. *et al.* (2011) Homeostatic response to hypoxia is regulated by the N-end rule pathway in plants. *Nature*, **479**:415-418

Gibbs, D.J., Voß, U., Harding, S.A., Fannon, J., Moody, L.A. & Yamada E. *et al.* (2014) *At*MYB93 is a novel negative regulator of lateral root development in *Arabidopsis*. *New Phytol.*, **203**:1194-1207

Goh, T., Toyokura, K., Wells, D.M. & Swarup, K. (2016) Quiescent center initiation in the *Arabidopsis* lateral root primordia is dependent on the SCARECROW transcription factor. *Development*., **143**(18)

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D. & Fazo, J. *et al*. (2012) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res*., **40**:1178-1186

Gou, M., Hou, G., Yang, H., Zhang, X., Cai, Y. & Kai, G. *et al.* (2017) The MYB107 transcription factor positively regulates suberin biosynthesis. *Plant Physiol.*, **173**(2):1045-1058

Gouy, M., Guindon, S. & Gascuel, O. (2010) SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.*, **27**(2):221-224

Grattan, S.R. & Grieve, C.M. (1999) Salinity mineral nutrient relations in horticultural crops. *Sci. Hort.*, **78**(1-4):127-157

Grierson, C. & Schiefelbein, J. (2002) Root hairs. *Arabidopsis Book*, **1**

Gruber, B.D., Giehl, R.F.H., Friedel, S. & von Wirén, N. (2013) Plasticity of the *Arabidopsis* root system under nutrient deficiencies. *Plant Physiol.*, **163**:161–179

Guindon, S., Dufayard, J.F., Leford, V., Anisimova, M., Hordijk, W. & Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.*, **59**(3):307-321

Hawkesford, M., Horst, W., Kichey, T., Lambers, H., Schjoerring, J. & Moller, I.S. *et al.* (2012) 'Functions on macronutrients' in Marschner, P. (ed.) *Marschner's mineral nutrition of higher plants, 3rd edition*. San Diego: Elsevier, pp. 135-189

Hellens, R.P., Edwards, A.E., Leyland, N.R., Bean, S. & Mullineaux, P.M. (2000) pGreen: a versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. *Plant Mol. Biol.*, **42**:819-832

Helliwell, R., Hartley, S., Pearce, W. & O'Neill, L. (2017) Why are NGOs skeptical of genome editing. *EMBO Rep*., **18**:2090-2093

Herdt R (2010) Handbook of Agricultural Economics, eds Pingali P, Evenson R (Elsevier, Amsterdam), pp 3253–3304

Hertzberg, M., Aspeborg, H., Schrader, J., Andersson, A., Erlandsson, R. & Blomqvist, K. *et al.* (2001) A transcriptional roadmap to wood formation. *PNAS* **98**(25):14732-14737

Hétu, M.F., Tremblay, L.J. & Lefebvre, D.D. (2005) High root biomass production in anchored arabidopsis plants grown in axenic sucrose supplemented liquid culture. *Biotechniques*, **39**(3):345-349

Hirota, A. ET AL (2007) The auxin-regulated AP2/EREBP gene PUCHI is required for morphogenesis in the early lateral root primordium of *Arabidopsis*. *Plant Cell*, **19**:2156-2168

Hofmann, K. & Baron, M.D. (1996) BOXSHADE 3.21, pretty printing of multiple-alignment files. [online] Available at: <https://embnet.vital-it.ch/software/BOX_form.html>

Hochholdinger, F., Park, W.J., Sauer, M. & Woll, K. (2004) From weeds to crops: genetic analysis of root development in cereals. *Trends Plant Sci.*, **9**(1):42-48

Hou, G., Hill, J.P. & Blancaflor, E.B. (2004) Developmetal anatomy and auxin response of LR formation in *Ceratopteris richardii*. *J. Exp. Bot.*, **55**(397):685-693

Hu, H. (2018) The role of transcription factor MYB53 from *Arabidopsis thaliana* in the regulated production of suberin, Master's Thesis, Carleton University, Ottowa.

Hu, Y. & Schmidhalter, U. (2005) Drought and salinity: a comparison of their effects on mineral nutrition of plants. *J. Plant Nutr. Soil Sci.* **168**:541-549

Iñiguez, L.P. & Hernández, G. (2017) The evolutionary relationship between alternative splicing and gene duplication. *Front. Genet.*, **8**:14

IPCC (2014) Summary for policymakers. In: IPCC. *Climate Change 2014: impacts, adaptation, and vulnerability*. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, pp. 1–32. Cambridge, UK and New York, USA, Cambridge University Press

Iyer-Pascuzzi, A.S., Jackson, T., Cui, H., Petricka, J.J., Busch, W. & Tsukagoshi, H. *et al.* (2011) Cell identity regulators link development and stress responses in the *Arabidopsis* root. *Dev. Cell*, **21**(4):770-782

Jain, M., Nijhawan, A., Tyagi, A.K. & Khurana, J.P. (2006) Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR. *Biochem. Biophys. Res. Com.*, **345**:646–651

Jamil, M., Jung, K.Y., Lee, D.B., Han, M.S. & Rha, E.S. (2007) Salt stress inhibits germination and early seedling growth in cabbage (*Brassica oleracea capitata* L.). *Pak, J. Biol. Sci.*, **10**(6):910-914

Jin, H., Cominelli, E., Bailey, P., Parr, A., Mehrtens, F. & Jones, J. *et al.* (2000) Transcriptional repression by *AtMYB54* controls production of UV-protecting sunscreens in *Arabidopsis*. *EMBO J.*, **19**:6150-6161

Ji, X., Wang, L., Nie, X., He, L., Zang, D. & Liu, Y. *et al.* (2014) A novel method to identify the DNA motifs recognized by a defined transcription factor. *Plant Mol. Biol.*, **86**:367-380

Jin, J.P., Tian, F., Yang, D.C., Meng, Y.Q., Kong, L. & Luo, J.C *et al.* (2017). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.*, **45**(1):1040-1045

Johnson, I.T. (2002) Glucosinolates in the human diet. Bioavailability and implications for health. *Phytochem. Rev.*, **1**:183-188

Jones, C.D., Hughes, J.K., Bellouin, N., Hardiman, S.C., Jones, G.S. & Knight, J. *et al*. (2011) The HadGEM2-ES implementation of CMIP5 centennial simulations. *Geosci. Model Dev*., **4**:543–570

Julkowska, M.M., Hoefsloot, H.C., Mol, S., Feron, R., de Boer, G. & Haring, M.A. *et al.* (2014) Capturing *Arabidopsis* root architecture dynamics with ROOT-FIT reveals diversity in responses to salinity. *Plant Physiol*., **166**:1387-1402

Kamiya, T., Borghi, M., Wang, P., Danku, J.M.C., Kalmbach, L. & Hosmani, P.S. *et al.*, (2015) The MYB36 transcription factor orchestrates Casparian strip formation. *PNAS*, **112**(33):10533-10538

Kannangara, R., Branigan, C., Liu, Y., Penfield, T., Rao, V. & Mouille, G. *et al.* (2007) The transcription factor WIN1/SHN1 regulates cutin biosynthesis in *Arabidopsis thaliana*. *Plant Cell*, **19**:1278-1294

Kanei-Ishii, C., Sarai, A., Sawazaki, T., Nakagoshi, H., He, D.N. & Ogata, K. *et al.* (1990) The tryptophan cluster: a hypothetical structure of the DNA-binding domain of the myb protooncogene product. *J. Biol. Chem.*, **265**:19990-19995

Kim, J. & Patterson, S.E. (2006) Expression divergence and functional redundancy of polygalacturonases in floral organ abscission. *Plant Signal. Behav.*, **1**:281-283

Kolkman, J.A. & Stemmer, W.P. (2001) Directed evolution of proteins by exon shuffling. Nat. Biotechnol., 19(5):423-428

Koornneef, M. & Meinke, D. (2010) The development of *Arabidopsis* as a model plant. *Plant J.*, **61**:909-921

Kosma, D.K., Murmu, J., Razea, F.M., Santos, P., Bourgault, R. & Molina, I. (2014) AtMYB41 activates ectopic suberin synthesis and assembly in multiple plant species and cell types. *Plant J.*, **80**(2):216-229

Kranz, H.D., Denekamp, M., Greco, R., Jin, H., Leyva, A. & Meissner, R.C., *et al.* (1998) Towards functional characterisation of the members of the *R2R3-MYB* gene family from *Arabidopsis thaliana*. *Plant J.*, **16**(2):263-276

Krasensky, J. & Jonak, C. (2012) Dought, salt and temperature stress-induced metabolic rearrangements of and regulatory networks. *J. Exp. Bot.*, **63**(4):1593-1608

Kruse, C., Jost, R., Lipschis, M., Kopp, B., Hartmann, M. & Hell, R. (2007) Sulfur-enhanced defence: effects of sulfur metabolism, nitrogen supply, and pathogen lifestyle. *Plant Biol.,* **9**:608-619

Kumar, S., Stecher, G. & Tamura, K. (2018) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.*, **35**:1547-1549

Kunert, K.J., van Wyk, S.G., Cullis, C.A., Vorster, B.J. & Foyer, C.H. (2015) Potential use of phytocystatins in crop improvement, with a particular focus on legumes. *J. Exp. Bot.*, **66**(12):3559-3570

Kutz, A., Muller, A., Hennig, P., Kaiser, W.M., Piotrowski, M. & Weiler, E.W. (2002) A role for nitrilase 3 in the regulation of root morphology in sulphur-starving *Arabidopsis thaliana*. *Plant J.*, **30**(1):95-106

Laplaze, L. Benkova, E., Casimiro, I., Maes, L., Vanneste, S. & Swarup, R. *et al.* (2007) Cytokinins act directly on lateral root founder cells to inhibit root initiation. *Plant Cell*, **19**(12):3889-3900

Lashbrooke, J., Cohen, H., Levy-Samocha, D., Tzfadia, O., Panizel, I. & Ziesler, V. *et al.* (2016) MYB107 and MYB9 homologs regulate suberin deposition in angiosperms. *Plant Cell*, **28**:2097-2116

Lavenus, J., Goh, T., Roberts, I., Guyomarc'h, S., Lucas, M. & Smet, I. *et al.,* (2013) Lateral root development in *Arabidopsis*: fifty shades of auxin. *Trends Plant Sci.*, **18**(8):450-458

Lee, J-Y., Wang, X., Cui, W., Sager, R., Modla, S. & Czymmek, K. *et al.* (2011) A plasmodesmata-localized protein mediates crosstalk between cell-to-cell communication and innate immunity in *Arabidopsis*. *Plant Cell*, **23**(9):3353-3373

Legay, S., Cocco, E., André, C.M., Guignard, C., Hausman, J.F. & Guerriero, G. (2017) Differential lipid composition and gene expression in the semi-russeted 'Cox Orange Pippin' apple variety. *Front. Plant Sci.*, **8**:1656

Legay, S., Guerriero, G., André, C., Guignard, C., Cocco, E. & Charton, S. *et al.*, (2016) MdMyb93 is a regulator of suberin deposition in russeted apple fruit skins. *New Phytologist*, **212**:977-991

Letunic, I. & Bork, P. (2019) Interactive Tree of Life (iTOL)v4 : recent updates and new developments. *Nucleic Acids Res.*, **47**(1):256-259

Lewanowska, M. & Sirko, A. (2008) Recent advances in understanding plant stress response to sulfur-deficiency stress. *Acta Biochim. Pol.*, **55**(3):457-471

Liang, Y., Mitchell, D.M. & Harris, J.M. (2007) Abscisic acid rescues the root meristem defects of the *Medicago trunculata ladt* mutant. *Dev. Biol.*, **304**:297-307

Li-Beisson, Y., Pollard, M., Sauveplane, V., Pinot, F., Ohlrogge, J. & Beisson, F. (2009) Nanoridges that characterize the surface morphology of flowers require the synthesis of cutin polyester. *PNAS*, **106**:22008-22013

Li, B., Kamiya, T., Kalmbach, L., Yamagami, M., Yamaguchi, K. and Shigenobu, S. *et al.* (2017) Role of LOTR1 in nutrient transport through organization of spatial distribution of root endodermal barriers. *Curr. Biol.*, **27**(5):758-765

Li, Y., Wang, X., Ban, Q., Zhu, X., Jiang, C. & Wei, *et al.* (2019) Comparative transcriptomic analysis reveals gene expression associated with cold adaptation in the tea plant *Camellia sinensis*. *BMC Genomics*, **20**:624

Li, X.C., Zhu, J., Yang, J., Zhang, G.R., Xing, W.F. & Zhang, S. *et al.* (2012) Glycerol-3-Phosphate Acyltransferase 6 (GPAT6) is important for tapetum development in *Arabidopsis* and plays multiple roles in plant fertility. *Mol. Plant*, **5**(1):131-142

Linton, K.J. (2006) Structure and function of ABC transporters. *Physiology*, **22**:122-130

Lin, W., Liao, Y., Yang, T.J.W., Pan, C., Buckhout, T.J. & Schmidt, W. (2011) Coexpression-based clustering of *Arabidopsis* root genes predicts functional modules in early phosphate deficiency signaling. *Plant Physiol.*, **155**:1383-1402

Lin, Z.F., Zhong, S.L. & Grierson, D. (2009) Recent advances in ethylene research. *J. Exp. Bot.*, **60**(12):3311-3336

Lippold, F., Sanchez, D.H., Musialak, M., Schlereth, A., Scheible, W.F. & Hincha, D.K. *et al.* (2009) AtMyb41 regulates transcriptional and metabolic responses to osmotic stress in *Arabidopsis*. *Plant Physiol*., **149**(4):1761-1772.

Livak, K.J. & Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C}{}_{T}$ method. *Methods*, **25**:402-408

Ljung, K. (2013) Auxin metabolism and homeostasis during plant development. *Development*, **140**(5):943-950

Long, M.Y., Rosenberg, C. & Gilbert, W. (1995) Intron phase correlations and the evolution of the intron exon structure of genes. PNAS, 92(26):12495-12499

Loomis, R.S. & Corner, D.J. (1992) Crop ecology: Productivity and management in agriculture systems. Cambridge University Press, Cambridge

Lu, C., Chen, M-X., Liu, R., Zhang, L., Hou, X. & Liu, S. *et al.* (2019) Abscisic acid regulates auxin distribution to mediate maize lateral root development under salt stress. *Front. Plant Sci.*, **10**:716

Maiti, R.K., Amaya, L.E.D., Cardona, S.I., Dimas, A.M.O., de la Rosa-Ibarra, M. & Castillo, H.L. (1996) Genotypic variability in maize cultivars (Zea mays L.) for resistance to drought and salinity at the seedling stage. *Plant Physiol.*, **148**:741-744

Malamy, J.E. & Benfey, P.N. (1997a) Down and out in *Arabidopsis*: The formation of lateral roots. *Trends Plant Sci.*, **2**(10):390–396

Malamy, J.E. & Benfey, P.N. (1997b) Organization and cell differentiation in lateral roots of *Arabidopsis thaliana*. *Development*, **124**(1):33-44

Marchler-Bauer, A. Derbyshire, M.K., Gonzales, N.R., Lu, S., Chitsaz, F. & Geer, L.Y. *et al.* (2015) CDD: NCBI's conserved domain database. *Nucleic Acids Res.*, **43**:222-226

Martinka, M, Dolan, L., Pernas, M., Abe, J. & Lux, A. (2012) Endodermal cell-cell contact is required for the spatial control of Casparian band development in *Arabidopsis thaliana*. *Ann. Bot.*, **110**(2):361-371

Masclaux-Daubresse, C., Daniel-Vedele, F., Dechorgnat, J., Chardon, F., Gaufichon, L. & Suzuki, A. (2010) Nitrogen uptake, assimilation and remobilization in plants: challenges for sustainable and productive agriculture. *Ann. Bot.*, **105**(7):1141-1157

Mazid, M., Khan, T.A. & Mohammad, F. (2011) Response of crop plants under sulphur stress tolerance: a holistic approach. *JSPB*, **7**(3):23-57

Meier, U. (2001). Growth stages of mono-and dicotyledonous plants. Berlin, Germany: Blackwell Wissenschafts-Verlag.

Mindrebo, J.T., Nartey, C.M., Seto, Y., Burkart, M.D. & Noel, J.P. (2017) Unveiling the functional diversity of the Alpha-Beta hydrolase fold in plants. *Curr. Opin. Struct. Biol.*, **41**:233-246

Mohnen, D. (2008) Pectin structure and biosynthesis. *Curr. Opin. Plant Biol.*, **11**(3):266-277

Molina, I., Bonaventure, G., Ohlrogge, J. & Pollard, M. (2006) The lipid polyester composition of *Arabidopsis thaliana* and *Brassica napus* seeds. *Phytochemistry*, **67**:2597-2610

Molina, I., Ohlrogge, J.B. & Pollard, M. (2008) Deposition and localization of lipid polyester in developing seeds of *Brassica napus* and *Arabidopsis thaliana*. *Plant J.*, **53**(3):437-449

Morcuende, R., Bari, R., Gibon, Y., Zheng, W., Pant, B.D. & Bläsing, O. *et al.* (2007) Genome-wide reprogramming of metabolism and regulatory networks of *Arabidopsis* in response to phosphorus. *Plant Cell Environ.* **30**:85-112

Mueller, N.D., Gerber, J.S., Johnston, M., Ray, D.K., Ramankutty N. & Foley, J.A. (2012) Closing yield gaps through nutrient and water management. *Nature*, **490**:254-257

Müller, C., & Robertson, R. (2014): Projecting future crop productivity for global economic modeling. *Agr. Econ.*, **45**:37-50

Mustroph, A., Zanetti, M.E., Jang, C.J., Holtan, H.E., Repetti, P.P. & Galbraith, D.W. *et al.* (2009) Profiling translatomes of discrete cell populations resolves altered cellular priorities during hypoxia in *Arabidopsis*. *Proc. Nati. Acad. Sci. USA.* **106**: 18843–18848.

Muttucumaru, N., Halford, N.G., Elmore, J.S., Dodson, A.T., Parry, M. & Shewry, P.R. *et al.* (2006) Formation of high levels of acrylamide during the processing of flour derived from sulfate-deprived wheat. *J. Agric. Food Chem.*, **54**:8951-8955

Naseer, S., Lee, Y., Lapierre, C., Franke, R., Nawrath, C. & Geldner, N. (2012) Casparian strip diffusion barrier in *Arabidopsis* is made of a lignin polymer without suberin. *PNAS*, **109**(25):10101-10106

Nelson, D. & Werk-Reichhart, D. (2011) A P450-centric view of plant evolution. *Plant J.*, **66**:194-211

New England Biolabs (undated) *NEBioCalculator v1.10.0 dsDNA: Mass to/from Moles Converter*. [online] Available at: <https://nebiocalculator.neb.com/#!/dsdnaamt>

Nakajima, K., Sena G., Nawy, T. & Benfey, P.N. (2001) Intercellular movement of the putative transcription factor SHR in root patterning. *Nature*, **413**:307-311
NIAB (2017) *Community resource for wheat transformation*. [online]. Available at: <http://www.niab.com/transgenic>

Nibau, C., Gibbs, D.J. & Coates, J.C. (2008) Branching out in new directions: The control of root architecture by lateral root formation. *New Phytol*., **179**:595-614

Nibau, C., Gibbs, D.J., Bunting, K.A., Moody, L.A., Smiles, E.J. & Tubby, J.A. *et al.* (2011) ARABIDILLO proteins have a novel and conserved domain structure important for the regulation of their stability. *Plant Mol. Biol*., **75**(1-2)77-92

Noble, W.S. (2009) How does multiple testing correction work. *Nat. Biotechnol.*, **27**(12):1135-1137

Noctor, G, Gomez, L., Vanacker, H. & Foyer, C.H. Interactions between biosynthesis, compartmentation and transport in the control of glutathione homeostasis and signaling. *J. Exp. Bot.*, **53**:1283-1304

Ociepa, P.M. (2017) Investigating the role of pectin methylesterases in regulating root development in *Arabidopsis thaliana*, PhD thesis, University of Southampton, Southampton.

Ogawa, M., Kay, P., Wilson, S. & Swain, S.M. (2009) ARABIDOPSIS DEHISCENCE ZONE POLYGALACTURONASE1 (ADPG1), ADPG2, and QUARTET2 are polygalacturonases required for cell separation during reproductive development in *Arabidopsis*. *Plant Cell*, **21**:216-233

Ogden, T.H. & Rosenberg, M.S. (2006) Multiple sequence alignment accuracy and phylogenetic inference. *Syst. Biol.*, **55**(2):314-328

Oshima, Y., Shikata, M., Koyama, T., Ohtsubo, N., Mitsuda, N. & Ohme-Takagi, M. (2013) MIXTA-like transcription factors and WAX INDUCER1/SHINE1 coordinately regulate cuticle development in *Arabidopsis* and *Torenia fournieri*. *Plant Cell*, **25**:1609-1624

Osmont, K.S., Sibout, R. & Hardtke, C.S. (2007) Hidden branches: Developments in root system architecture. *Annu. Rev. Plant Biol*., **58**:93-113

Ouwerkerk, P.B.F & Meijer, A.H. (2001) Yeast one-hybrid screening for DNA-protein interactions. *Curr. Protoc. Mol. Biol.*, **12**(12)

Overvoorde, P., Fukaki, H. & Beeckman, T. (2010) Auxin control of root development. *Cold Spring Harb. Perspect. Biol.*, **2**(6)

Pais, F.S-M., Ruy, P.C., Oliveira, G. & Coimbra, R.S. (2014) Assessing the efficiency of multiple sequence alignment programs. *Algorithms Mol. Biol.*, **9**:4

Palazzo, A.F. & Gregory, R.T. (2014) The case for junk DNA. *PloS Genet.* **10**(5)

Paolacci, A.R., Tanzarella, O.A., Porceddu, E. & Ciaffi, M. (2009) Identification and validation of reference genes for quantitative RTPCR normalization in wheat. *BMC Mol. Bio.*, **10**(11)

Pappas, K.M. & Winans, S.C. (2003) Plant transformation by coinoculation with a disarmed *Agrobacterium tumefaciens* strain and an *Escherichia coli* strain carrying mobilizable transgenes. *Appl. Environ. Microbiol.*, **69**(11):6731-6739

Park, P.J. (2009) ChIP-seq: advantages and challenges of a maturing technology. *Nat. Rev. Genet.*, **10**(10):669-680

Peachey, E.J. (2004) The Aral Sea basin crisis and sustainable water resource management in Central Asia. *JPIA*, **15**:1-20

Péret, B., De Rybel, B., Casimiro, I., Benková, E., Swarup, R. & Laplaze, L. *et al.* (2009) *Arabidopsis* lateral root development: an emerging story. *Trends Plant Sci.*, **14**:399–408

Péret, B., Zhao, J., Band, L.R., Voß, U., Postaire, O. & Luu, D.T. *et al.* (2012) Auxin regulates aquaporin function to facilitate lateral root emergence. *Nat. Cell. Biol.*, **14**(10):991-998

Péret, B., Middleton, A.M., French, A.P., Larrieu, A., Bishopp, A. & Njo, M. *et al.* (2013) Sequential induction of auxin efflux and influx carriers regulates lateral root emergence. *Mol. Syst. Biol.*, **9:**699

Petit, J. Bres, C., Mauxion, J-P., Tai, F.W.J., Martin, L.B. & Fich, E.A. *et al.* (2016) The Glycerol-3-Phosphate Acyltransferase GPAT6 from tomato plays a central role in fruit cutin biosynthesis. *Plant Physiol.*, **171**(2):894-913

Petryszak, R., Keays, M., Tang, A., Fonseca, N.A., Barrera, E. & Burdett, T. *et al.* (2016) Expression Atlas update – an integrated database of gene and protein expression in humans, animals and plants. *Nucleic Acids Res.*, **44**:746-752

Pingali, P.L (2012) Green Revolution: Impacts, limits and the path ahead. *PNAS*, **109**(31):12302-12308

Pinot, F. & Beisson, F. (2011) Cytochrome P450 metabolizing fatty acids in plants: characterization and physiological roles. *FEBS J.*, 195-205

Pollard, M., Beisson, F., Li, Y. & Ohlrogge, J.B. (2008) Building lipid barriers: biosynthesis of cutin and suberin. *Trends Plant Sci.*, **13**(5):1360-1385

Postma, J.A., Dathe, A. & Lynch, J.P. (2014) The optimal lateral root branching density for maize depends on nitrogen and phosphorous availability. Plant Physiol., 166:590-602

Preston, J., Wheeler, J., Heazlewood, J., Li, S.F., Parish, R.W. (2004) *AtMYB32* is required for normal pollen development in *Arabidopsis thaliana*. *Plant J., **40**:979-995*

Purushothaman, R. Krishnamurthy, L., Upadhyaya, H.D., Vadez, V. & Varshney, R.K. 2017) Root traits confer grain yield advantages under terminal drought in chickpea (*Cicer arietinum* L.) *Field Crops Res.*, **201**:146-161

Ramírez-González, R.H., Borrill, P., Lang, D., Harrington, S.A., Brinton, J. & Venturini, L. *et al.* (2018) The transcriptional landscape of polyploid wheat. *Science*. **361**

Rao, T.S.R.B., Naresh, J.V., Reddy, P.S., Reddy, M.K. & Malikarjuna, G. (2017) Expression of *Pennisetum glaucum* eukaryotic translational initiation factor 4A (*PgeIF4A*) confers improved drought, salinity, and oxidative stress tolerance in groundnut.

Ray, B. & Shaw, R. (2016) Urban Disasters and Resilience in Asia - Chapter 20: Water stress in the megacity of Kolkata, India and it's implications for urban resilience, Butterworth-Heinemann, 1st edition, UK.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL accessed on 20/09/19: https://www.R-project.org/

Reece-Hoyes, J.S. & Walhout, A.J.M. (2012) Gene-centred yeast one-hybrid assays. *Methods Mol. Biol.,* **812**:189-208

Rengasamy, P. (2006) World salinization with emphasis on Australia. *J. Exp. Bot.*, **57**(5):1017-1023

Reyt, G., Boudouf, S., Boucherez, J., Gaymard, F. & Briat, J.F. (2015) Iron- and ferritin-dependent reactive oxygen species distribution: impact on *Arabidopsis* root system architecture. *Molecular biology*, **8**:439-453

Rieu, I., Twell, D. & Firon, N. (2017) Pollen development at high temperature: from acclimation to collapse. *Plant Physiol.*, **173**(4):1967-1976
Robbins II, N.E., Trontin, C., Duan, L. & Dinneny, J.R. (2014) Beyond the barrier: Communication in the root through the endodermis. *Plant Physiol*., **166**:551-559

Rodriguez, M.V., Barrero, J.M., Corbineau, F., Gubbler, F. & Benech-Arnold, R.L. (2015) Dormancy in cereals (not too much, not so little): about the mechanisms behind this trait. *Seed Sci. Res.*, **25**:99-119

Roeder, A.H.K. & Yanofsky, M.F. (2006) Fruit Development in *Arabidopsis*. *TAB*, **4**

Rogers, E.D. & Benfey, P.N. (2015) Regulation of plant root system architecture: implications for crop advancement. *Curr. Opin. Plant. Bio.*, **32**:93-98

Romero, I., Fuertes, A., Benito, M.J., Malpica, J.M., Leyva, A., & Paz-Ares, J. (1998) More than 80 *R2R3-MYB* regulatory genes in the genome of *Arabidopsis thaliana*. *Plant J.*, **14**(3):273-284

Rosenzweig, C., Iglesias, A., Yang, X.B., Epstein, P.R. & Chivian, E. (2001) Climate change and extreme weather events – Implications for food production, plant diseases, and pests. *NASA Publications*, **24**:89-104

Rudashevskaya, E.L., Ye, J., Jenson, O.N., Fuglsang, A.T. & Palmgren, M.G. (2012) Phosphosite mapping of P-type plasma membrane H+-ATPase in homologous and heterologous environments. *J. Biol. Chem.*, **287**(7):4904-4913

Sager, R.E. & Lee, J-Y. (2018) Plasmodesmata at a glance. *J. Cell Sci.*, **131**

Sairam, R.K. & Tyagi, A. (2004) Physiology and molecular biology of salinity stress tolerance in plants. *Curr. Sci.*, **86**:407–421

Scheres, B. & Wolkenfelt, H. (1998) The *Arabidopsis* root as a model to study plant development. *Plant Physiol. Biochem.*, **36**(1-2):21-32

Schindelin, J., Rueden, C.T., Hiner, M.C. & Eliceiri, K.W. (2015) The ImageJ ecosystem: an open platform for biomedical image analysis. *Mol. Reprod. Dev.*, **82**(7-8):518-529

Serraj, R., Krishnamurthy, L., Kashiwagi, J., Kumar, J., Chandra, S., & Crouch, J.H. (2004) Variation in root traits of chickpea (*Cicer arietinum* L.) grown under terminal drought. *Field Crops Res.*, **88**:115-127

Sharma, D.A., Rishi, M.S. & Keesari, T. (2017) Evaluation of groundwater quality and suitability for irrigation and drinking purposes in southwest Punjab, India using hydrochemical approach. *Appl. Water Sci.*, **7**(6):3137-3150

Scheible, W.R., Morcuende, R., Czechowski, T., Fritz, C., Osuna, D. & Palacios-Rojas, N. *et al.* (2004) Genome-wide reprogramming of primary and secondary metabolism, protein synthesis, cellular growth processes, and the regulatory infrastructure of *Arabidopsis* in response to nitrogen. *Plant Physiol.*, **136**:2483-2499

Shi, D., Ren, A., Tang, X., Qi, G., Xu, Z. & Chai, G. *et al.* (2018) *MYB52* negatively regulates pectin demethylesterification in seed coat mucilage. *Plant Physiol.* **176**:2737-2749

Shin, R., Burch, A.Y., Huppert, K.A., Tiwari, S.B., Murphy, A.S. & Guilfoyle, T.J. (2007) The *Arabidopsis* Transcription Factor MYB77 Modulates Auxin Signal Transduction. *Plant Cell*, **19**:2440-2453

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K. & Li, W. *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.*, **7**:539

SIGnAL (undated) *T-DNA Primer Design*. [online] Available at: <http://signal.salk.edu/tdnaprimers.2.html>

Singh, M., Bag, S.K., Bhardwaj, A., Ranjan, A., Mantri, S. & Nigam, D. *et al.* (2015) Global nucleosome positioning regulates salicylic acid mediated transcription in *Arabidopsis thaliana. BMC Plant Biol.*, **15**(13)

Simpson, C., Thomas, C., Findlay, K., Bayer, E. & Maule, A.J. (2009) An *Arabidopsis* GPI-anchor plasmodesmal neck protein with callose binding activity and potential to regulate cell-to-cell trafficking. *Plant Cell.*, **21**:581-594

Smedema, L.K. & Shiati, K. (2002) Irrigation and salinity: a perspective review of the salinity hazards of irrigation development in the arid zone. *Irrig. Drain. Syst.* **16**:161-174

Solomon, S., Qin, D., Manning, M., Alley, M.B., Berntsen, T., & Bindoff, N.L. et al., in *Climate Change 2007: The Physical Science Basis: Contribution of Working Group*

I to the Fourth Assessment Report of the IPCC, Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., & Averyt, K.B. *et al.*, Eds. (Cambridge Univ. Press, New York, 2007), pp. 1–8

Song, C., Kim, T., Chung, W.S. & Lim, C.O. (2017) The *Arabidopsis* phytocystatin AtCYS5 enhances seed germination and seedling growth under heat stress conditions. *Mol. Cells,* **40**(8):577-586

Sparks, E.E., Drapek, C., Gaudinier, A., Li, S., Ansariola, M. & Shen, N. *et al.* (2016) Establishment of expression in the *SHORTROOT-SCARECROW* transcriptional cascade through opposing activities of both activators and repressors. *Dev. Cell.*, **39**:585-596

Stanislaus, A., Marafi, A. & Rana, M.S. (2010) Recent advances in the science and technology of ultra low sulfur diesel (ULSD) production. *Catal. Today*, **153**(1-2):1-68

Steffens, B. & Rasmussen, A. (2016) The physiology of adventitious roots. *Plant. Physiol.*, **170**:603-617

Stoeckle, D., Thellman, M. & Vermeer, E.J. (2018) Breakout – lateral root emergence in *Arabidopsis thaliana*. *Curr. Opin. Plant Biol.*, **41**:67-72

Stracke, R., Werber, M. & Weisshaar, B. (2001) The R2R3-MYB gene family in *Arabidopsis thaliana*. *Curr. Opin. Plant Bio.*, **4**(5):447-456

Swarup, K., Benková, E., Swarup, R., Casimiro, I., Péret, B. & Yang, Y. *et al.* (2008) The auxin influx carrier LAX3 promotes lateral root emergence. *Nat. Cell Biol.*, **10**:946-954

The Plant List (2013). Version 1.1. Accessed on 20/09/19 from http://www.theplantlist.org/

The Sulphur Institute (2009) Sulphur in Asia. *Fertilizer International* **433**:20-24

Thomas, C.L., Alcock, T.D., Graham, N.S., Hayden, R., Matterson, S. & Wilson, L *et al.* (2016) Root morphology and seed and leaf ionomic traits in a *Brassica napus* L. diversity panel show wide phenotypic variation and are characteristic of crop habit. *BMC Plant Biol.*, **16**(1):214

Traka, M.H. (2016) Chapter nine – health benefits of glucosinolates. *Adv. Bot. Res.*, **80**:247-279 . In Glucosinolates Edited by Stanislav Kopriva

Trinh, D.C., Lavenus, J., Goh, T., Boutté, Y., Drogue, Q. & Vaissayre, V. *et al.* (2019) PUCHI regulates very long chain fatty acid biosynthesis during lateral root and callus formation. *PNAS*, **116**(28):14325-14330

United Nations, Department of Economic and Social Affairs, Population Division (2017). World Population Prospects: The 2017 Revision, Key Findings and Advance Tables. Working Paper No. ESA/P/WP/248

Uno, Y., Furihata, T., Abe, H., Yoshida, R. Shinozaki, K. & Yamaguchi-Shinozaki, K. (2000) *Arabidopsis* basic leucine zipper transcription factors involved in an abscisic acid-dependent signal transduction pathway under drought and high-salinity conditions. *PNAS*, **97**(21):11632-11637

Vandesompele J., Preter B., Poppe B., Roy N.V., & Paepe A.D. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.,* **3**(7):1-12

Vanneste, S., De Rybel, B., Beemster, G.T., Ljung, K., De Smet, I. & Van Isterdael, G. *et al.* (2005) Cell cycle progression in the pericycle is not sufficient for SOLITARY ROOT/IAA14-mediated lateral root initiation in *Arabidopsis thaliana. Plant Cell*, **17**:3035-3050.

Vermeer, J.E.M., Wangenheim, D., Barberon, M., Lee, Y., Stelzer, E.H.K., & Maizel, A. *et al.* (2014) A spatial accommodation by neighboring cells is required for organ initiation in *Arabidopsis*. *Science*, **343**:178-183

Verstraeten, I. Schotte, S. & Geelen, D. (2014) Hypocotyl adventitious root organogenesis differs from lateral root development. *Front. Plant Sci.*, **5**.

Vitart, V., Baxter, I., Doerner, P. & Harper, J.F. (2001) Evidence for a role in growth and salt stress of a plasma membrane H-ATPase in the root endodermis. *Plant J.* **27**(3):191-201

Voß, U., Wilson, M.H., Kenobi, K., Gould, P.D., Robertson, F.C., & Peer, W.A. *et al.* (2015) The circadian clock rephrases during lateral root organ initiation in *Arabidopsis thaliana*. *Nat. Commun.*, **6**(7641):1-9

Wachsman, G., Sparks, E.E. & Benfey, P.N. (2015) Genes and networks regulating root anatomy and architecture. *New Phyto.*, **208**(1):26-38

Waines, J.G. & Ehdaie, B. (2007) Domestication and Crop Physiology: Roots of green-revolution wheat. *Ann. Bot.*, **100**:991-998

Wang, H.Z. & Dixon, R.A. (2012) On-off switched for secondary cell wall biosynthesis. *Mol. Plant*, **5**(2):297-303

Wang, Y., Li, K. & Li, X. (2009) Auxin redistribution modulates plastic development of root system architecture under salt stress in *Arabidopsis thaliana*. *J. Plant Physiol.*, **166**(15):1637-1645

Wang, Y., Zhang, T., Wang, R. & Zhao, Y. (2017) Recent advances in auxin research in rice and their implications for crop improvement. *J. Exp. Bot.*, **69**(2):255-263

Wilkinson, H.B. (2018) Investigating the MYB93 signalling pathway in *Arabidopsis thaliana* under nutrient stress and in wheat, Masters Thesis, University of Birmingham, Birmingham.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G.V., & Provart, N.J. (2007) An “Electronic Fluorescent Pictograph” browser for exploring and analyzing large-scale biological data sets. *PLoS One*, **2**(8)

Winter, E. & Ponting, C. (2002) TRAM, LAG1 and CLN8: members of a novel family of lipid-sensing domains? *Trends Biochem. Sci.*, **27**:381-383

Xiao, F., Goodwin, S.M., Xiao, Y., Sun, Z., Baker, D. & Tang, X. *et al.* (2004) *Arabidopsis CYP86A2* represses *Pseudomonas syringae* type III genes and is required for cuticle development. *EMBO J.*, **23**(14):2903-2913

Yadav, V., Molina, I., Ranathunge, K., Castillo, I.Q., Rothstein, S.J., Reed, J.W. (2014) ABCG transporters are required for suberin and pollen wall extracellular barriers in *Arabidopsis*. *Plant Cell*., **26**(9):3569-3588

Yanhui, C., Xiaoyuan, Y., Kun, H. Meihua, L., Jigang, L. & Zhaofeng, G. *et al.* (2006) The MYB transcription factor superfamily of *Arabidopsis*: expression analysis and phylogenetic comparison with the rice MYB family. *Plant Mol. Biol.*, **60**:107-124

Zhan, A., Schneider, H., & Lynch, J.P. (2015) Reduced lateral root branching density improves drought tolerance in maize. *Plant Physiol.*, **168**:1603-1615

Zhan, X., Quian, B., Cao, G., Wu, W., Yan, L. & Guan, Q. *et al.* (2015) An *Arabidopsis* PWI and RRM motif-containing protein is critical for pre-mRNA splicing and ABA responses. *Nat. Commun.*, **6**:8139

Zhang, S., Zhang, R., Gao, J., Gu, T., Song, G. & Li, W. *et al.* (2019) Highly efficient and heritable targeted mutagenesis in wheat via the *Agrobacterium tumefaciens*-mediated CRISPR/Cas9 system. *Int. J. Mol. Sci.*, **20**(17)

Zhao, F.J., Fortune, S., Barbosa, V.L., McGrath, S.P., Stobart, R. & Bilsborrow, P.E. *et al.* (2006) Effects of sulphur on yield and malting quality of barley. *J. Cereal Sci.*, **43**(3):369-377

Zhao, F.J., Hawkesford, M.J. & McGrath, S.P. (1999) Sulphur assimilation and effects on yield and quality of wheat. *J. Cereal Sci.*, **30**(1):1-17

Zheng, M., Liu, X., Lin, J., Liu, X., Wang, Z. & Xin, M. *et al.* (2019) Histone acetyletransferase GCN5 contributes to cell wall integrity and salt stress tolerance by altering the expression of cellulose synthesis genes. *Plant. J.*, **97**:587-602

Zhong, R., Lee, C., Zhou, J., McCarthy R.L., & Ye Z.H. (2008) A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in *Arabidopsis*. *Plant Cell*, **20**:2763-2782

Zhong, R. & Ye, Z.H. (2012) MYB46 and MYB83 bind to the SMRE sites and directly activate a suite of transcription factors and secondary wall biosynthetic genes. *Plant Cell Physiol.*, **53**:368-380

Zhong, R. & Ye, Z.H. (2014) Complexity of the transcriptional network controlling secondary wall biosynthesis. *Plant Sci.*, **229**:193-207

Zhou (2009) MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in *Arabidopsis*. *Plant Cell*, **21**:248-266

Zia, A. & Moses, A.M. (2012) Towards a theoretical understanding of false positives in DNA motif finding. *BMC Bioinformatics*, **13**:151

Zörb, C., Steinfurth, D., Seling, S., LangenkÄmper, G., Koehler, P. & Weiser, H. et al. (2009) Quantitative protein composition and baking quality of winter wheat as affected by late sulfur fertilization. *J. Agric. Food Chem.*, **59**(9):3877-3885

Appendix 2.1: Mutant plant lines

Gene	*AT Accession*	*SALK/SAIL ID*	*Reference*
MYB53	AT5G65230		Hu, H. (2018) 'The roll of transcription factor MYB53 from Arabidopsis thaliana in the regulated production of suberin', Master of Science
MYB92	AT5G10280		Hu, H. (2018) 'The roll of transcription factor MYB53 from Arabidopsis thaliana in the regulated production of suberin', Master of Science
MYB93	AT1G34670		Hu, H. (2018) 'The roll of transcription factor MYB53 from Arabidopsis thaliana in the regulated production of suberin', Master of Science
AtCYP708A1	AT1G55940	SAIL_361_C06	
AtMYB54	AT1G73410	SALK_043499	
AtABCG2	AT2G37360	SAIL_697_C08	
AtGPAT6	AT2G38110	SALK_056554	
At2G43670	AT2G43670	SALK_133229	
PDLP4	AT3G04370	SAIL_243_E02	
AT3G27270	AT3G27270	SALK_066426	
AHA4	AT3G47950	SALK_040519	
AT4G17480	AT4G17480	SALK_205932	
S-ACP-DES1	AT5G16240	SAIL_237_F10	
CASPL1B1	AT5G44550	SAIL_114_C02	
CYS5	AT5G47550	SALK_149928	
QRT1	AT5G55590	SALK_043429	
SCR	AT3G54220		Goh, T., Toyokura, K., Wells, D.M., Swarup, K., Yamamoto, M. & Mimura, T. *et al*. (2016)

Appendix 2.1: Recipes

Bacterial growth medium

LB broth:

Bacto-tryptone	10 g/L
Bacto-yeast extract	5 g/L
NaCl	10 g/L

For agar:

Bacto-agar	15 g/L

Low salt LB broth:

Bacto-tryptone	10 g/L
Bacto-yeast extract	5 g/L
NaCl	5 g/L

For agar:

Bacto-agar	15 g/L

Yeast growth medium

YPD broth:

Glucose	20 g/L
Bacto-peptone	20 g/L
Bacto-yeast extract	10 g/L

For YPDA:

Adenine (0.2%)	15 ml/L

For agar:

Bacto-agar	20 g/L

SD drop-out broth:

DO base	1.92 g/L
Sucrose	10 g/L
L-glutamate	1 g/L
YNB-antibiotic	1.7 g/L

Note: if using G418 the YNB should not contain ammonium sulfate

For agar:

Bacto-agar	20 g/L

Plant growth medium

1/2 Murashige and Skoog (MS) agar

MS basal medium with Gamborg's vitamins (Sigma M0404)	2.2 g/L
Sucrose	10 g/L

Adjust pH to 5.6-5.8

Agar	7-10 g/L

Extraction buffers

Quick DNA extraction buffer

KCl	0.25 M
EDTA	10 mM
Tris-HCL pH 9.5	100 mM

Quick DNA dilution buffer

BSA (3%)	3 g/L

Protein extraction buffer

Glycerol (5%)	5 g/L
Hepes	50 mM
NaCl	150 mM
Tween (0.1%)	1 ml/L
Complete Mini Protease Inhibitor Cocktail tablet	1

Agarose gel electrophoresis solutions

5x TBE:

EDTA	12.5 mM
Orthoboric acid	0.45 M
Tris	0.45 M

SDS-page and western blotting solutions

10% resolving gel:

dH2O	4 ml
Acrylamide	3.3 ml
Tris pH 8	2.5ml (1.5 M)
Sodium dodecyl sulfate (SDS)	100 µl 10% (w/v)
Ammonium per sulfate (APS)	100 µl 10% (w/v)
TEMED	4 µl

4% stacking gel:

dH2O	3.4 ml
Acrylamide	0.83 ml
Tris pH 6.6	0.63 ml (1 M)
SDS	50 µl 10% (w/v)
APS	50 µl 10% (w/v)
TEMED	5 µl

5x protein loading buffer:

Tris-HCl pH 6.8	67.5% (v/v) (2 M)
SDS	10% (w/v)
Glycerol	50% (w/v)
β-mercaptoethanol	5 ml/L
Bromphenol blue	0.005% (w/v)

Running buffer:

Tris	125 mM
Glycine	950 mM
SDS	0.1% (v/v)

Adjust pH to 8.3

Transfer buffer:

Tris	25 mM
Glycine	190 mM
Methanol (20%)	200 ml/L

Adjust pH to 8.0

Blocking solution:

TBS	10% (v/v)
Milk powder	5% (w/v)

Microbial selection

Selection Agent	Stock Concentration	Working Concentration
Kanamycin	50 mg/ml	50 µg/ml
Ampicillin	50 mg/ml	50 µg/ml
Rifampicin	50 mg/ml	50 µg/ml
Geneticin (G418)	50 mg/ml	150 µg/ml
Glufosinate ammonium (BASTA)	20 mg/ml	20 µg/ml

Macronutrient Stock for 10x 0.5 MS	**Mass/g in 1L dH_2O**
Ammonium nitrate	8.25
Calcium chloride anhydrous	1.65
Potassium nitrate	9.5
Potassium phosphate	0.85
Magnesium sulphate	0.9
Magnesium chloride	0.712

Micronutrient Stock for 100x 0.5 MS	**Mass/g in 1L dH_2O**
Na2-EDTA	1.863
Potassium iodide	0.042
Zinc sulphate - $7H_2O$	0.43
Ferrous sulphate - $7H_2O$	1.39
Manganese sulphate - H_2O	0.845
Zinc chloride	0.204
Ferrous Chloride - $4H_2O$	0.994
Manganese chloride - $4H_2O$	0.989
	Volume/ml in 1L dH_2O
Boric acid (0.6 M stock)	8.356
Cobalt chloride (hexahydrate 0.02 M stock)	0.263
Molybdic acid (0.03 M stock)	2.573
Cupric sulphate anhydrous (0.6 M stock)	**0.013**
Copper Chloride (0.02 M)	0.39

Appendix 2.3: Primers

qRT-PCR primers

Gene	*Oligonames*	*5′ primer sequence*	*3′ primer sequence*
AtCYP708A1	1G55940F & 1G55940R	ATGTTCAAGCACGGTGGTTC	GGAAGTCCTTGCCTTGCTTG
AtMYB54	1G73410F & 1G73410R	CTGGTCGCTCTGGTAAGAGTTG	GGTTCTTGACGGCGTTATCAG
AtLTP8	2G18370F & 2G18370R	TTCTGAATCCGCTATATCTTGCAG	AGCCACTGACTTGATGCATTG
AtPME10	2G19150F & 2G19150R	CCAATCCATTTTTGAGGGATGTAC	TGAATAAGATTTCCACGCTCGTC
AtABCG2	2G37360F & 2G37360R	GACCTACAGCGTGAAGATCCAG	ATCCGGTTGGCTAATGCATC
AtGPAT6	2G38110F & 2G38110R	TTTGGTGGCCTAGCGTCTG	ACAGAGCAACTAACGGCGTTG
At2G43670	2G43670F & 2G43670R	ACCAAGGCAGATCGCTCG	AGATTCATGGCAACAGAAGCAC
AtPDLP4	3G04370F & 3G04370R	CTGGCGCTCAAGTTAATAATCATC	GTAGCTATCCTCATAGAATCCGTGAC
At3G27270	3G27270F & 3G27270R	TGGATTCTCCTCCGTCAACAC	GCCGTGAGCAACCATGTAAC
AtAHA4	3G47950F & 3G47950R	GGTTGCGTTCACAAGGCA	GTTCACGCAACCTAGCAATTTC
At4G17480	4G17480F & 4G17480R	CACCGACTTCGTACAAGATCATATTG	CAAACATGACAAGGACCAAGTTG
AtS-ACP-DES1	5G16240F & 5G16240R	CATTGGTTCTGGAATGGATCC	GTTCATGGCGCCTCTCATC
AtCASPL1B1	5G44550F & 5G44550R	CCATTCTTGACATGCTAAATGTGAC	GGAGATGATGAGCATAAGGATTACG
AtCYS5	5G47550F & 5G47550R	CAAGTCGTAGAGATCGGTGAATTC	GTCCCATACGATTGCCAAGTAG
At5G49350	5G49350F & 5G49350R	TGAGAAACACATATACAGAGAGTGTGATG	TGTCGTGAATGGTGGCTCTG
AtQRT1	5G55590F & 5G55590R	TGGGATTTACAGGGAGAAGGTG	GCACAGAAGAAATCAGATTCAATGG
AtMYB93	MYB93 F & MYB93 R	AAGCTCGCAGATTTGAATAGGTG	ATCTGTACGACCTTGCAAATGC
AtACTIN2	ACTIN2 F & ACTIN2 R	TCGTACAACCGGTATTGTGCTG	TTACAATTTCCCGCTCTGCTG
AtUBC21	UBC21 F & UBC21 R	CGATTCTTGACCAAGATATTCCATC	TTAGAAGATTCCCTGAGTCGCAG
AtSCR	SCR_F & SCR_R	GCAGATAAGCTTGGCCTGCC	GGAGCTAATCTTTGGAGTAACCAG
AtMYB41	MYB41 qPCR F & MYB41 qPCR R	TGTTATGGGAAACAAGTGGTC	CATGTGTTGCAACTGCTG

Genotyping primers (for full length cDNA amplification)

Gene	*Oligonames*	*5′ primer sequence*	*3′ primer sequence*
AtMYB53	MYB53F & MYB53R	ATGGGAAGATCTCCTAGCTCAG	TTAAGATTGATAAGAAATGTCTGGAAAC
AtMYB92	MYB92F & MYB92R	ATGGGAAGATCTCCTATCTCTGATG	CTAAGGAATGTCGGAAAATATAGAATC
AtMYB93	MYB93ECOF & MYB93BAMR	AAAGAATTCGGGAGGTCGCCTTGTTGCG	AAAGGATCCCTAAGATATAACGTTCATGAG
AtCYP708A1	1G55940F_full & 1G55940R_full	ATGCCAGAGCATTGGGTC	AAAGATTTGCATTTTGATGAGATGAAC
AtMYB54	1G73410F_full & 1G73410R_full	ATGATCATGTGCAGCCGAG	CTAGGAGGCAGAGTTTCCAAC
AtABCG2	2G37360F_full & 2G37360R_full	ATGTCTGGCTTGCTCGG	TCACTTCCGTTTGTTCTTGCTAC
AtGPAT6	2G38110F_full & 2G38110R_full	TGGGAGCTCAGGAGAAACG	TTCACCGGAACCCTACCG
At2G43670	2G43670F_full & 2G43670R_full	ATGGCTAAAGCACAAATATGTCTTTG	CTTACGAAATTGATAAATGCAACTTCC
AtPDLP4	3G04370F_full & 3G04370R_full	TCCTCACACAAACCCTAGCC	AAGATAGCAACAAAGACCAATGC
At3G27270	3G27270F_full & 3G27270R_full	ATGGAACCAATTACTTATTCCAAAGATC	ACAGTACCCACAAATTCCAAATCC
AtAHA4	3G47950F_full & 3G47950R_full	ACGACGACTGTGGAGGAC	TCAGACAGTGTAAGCTTGTTGAATGG
At4G17480	4G17480F_full & 4G17480R_full	GGAGAAGAGTTTCCAGAAATCAGC	GATGCAGTATCTCCAAATCTTCACTC
AtS-ACP-DES1	5G16240F_full & 5G16240R_full	GGTTATGGCTATGGATCGGATC	TAAGCCCTGACTTCTCGACC
AtCASPL1B1	5G44550F_full & 5G44550R_full	TCAAAGCTTACATTGGCTGCTAC	GGACGACCGAGGGAGATG
AtCYS5	5G47550F_full & 5G47550R_full	TGACTAGTAAGGTCGTCTTCCTTC	TAAAGGAAGCGACCATTATTGGC
AtQRT1	5G55590F_full & 5G55590R_full	ATTCCCGCCGTTTTATTGCTC	TAGAGTCTCAGCCATTGATCTCC
AtACTIN2*	ACTIN2 F & ACTIN2 R	TCGTACAACCGGTATTGTGCTG	TTACAATTTCCCGCTCTGCTG

* Not full length but used as a positive control

Genotyping primers (for DNA and T-DNA amplification)

Gene	*Oligonames*	*LP primer sequence*	*RP primer sequence*
AtCYP708A1	SAIL361C06_LP & SAIL361C06_RP	GAAAAAGTCCGATGAGCACAC	TGAGACCATCCTCCAAAACAG
AtMYB54	SALK043499_LP & SALK043499_RP	AATCCTAGGAGGCAGAGTTTCC	TACGGTCCTCACAATTGGAAC
AtABCG2	SAIL697C08_LP & SAIL697C08_RP	TCGATTGATTATGGAGGACAAG	TGTTCACCGAGGAATCATTTC
AtGPAT6	SALK056554_LP & SALK056554_RP	TGCAAGGTTATTTCCGTCATC	ATATTCGTGTAGACGCGGATG
At2G43670	SALK133229_LP & SALK133229_RP	GCACCTCCTTCCGAGATAATC	TGAAAGGAAACCATATTTGTGC
AtPDLP4	SAIL243E02_LP & SAIL243E02_RP	CCGACTATTTCGCTCTCACAC	TCCTCGAAGATGAATACGAGC
At3G27270	SALK066426_LP & SALK066426_RP	TCCAAGTCCAGAGATAATGGG	CAACCATCTCCATCAACCATC
AtAHA4	SALK040519_LP & SALK040519_RP	AATTTCAGCTCTCCTCTTGGC	ATCAAAGGACAGGGTGAAACC
At4G17480	SALK205932_LP & SALK205932_RP	ATGACAAGGACCAAGTTGTGC	TCATTCCAGTTTCAATCTCGG
AtS-ACP-DES1	SAIL237F10_LP & SAIL237F10_RP	CATTGGCTTTTGATCATACGC	GTCTCATCTCTAACCCCGTCC
AtCASPL1B1	SAIL114C02_LP & SAIL114C02_RP	GCGGTTTCAAAGCTTACATTG	GTTGGACGAGACGAGAGATTG
AtCYS5	SALK149928_LP & SALK043429_RP	GCGTGTCTTCTTCTGTTTTCG	ATTATCCGGCATAATTACCGG
AtQRT1	SALK043429_LP & SALK043429_RP	TATCTCTTTGCGGTTGAGTGG	TGGGTAACCACAAACGCTAAC

Oligonames	*LB primer sequence*
LB1*	GCCTTTTCAGAAATGGATAAATAGCCTTGCTTCC
LBb1.3*	ATTTTGCCGATTTCGGAAC

*T-DNA left border primers, LB1 used for SAIL lines and LBb1.3 used for SALK lines

ChIP construct primers

Name of construct	*Oligonames*	*Primer sequence*
pGreen0229::pMYB93-myc-gAtMYB93	MYB93ECOF	AAA**GAATTC**GGGAGGTCGCCTTGTTGCG
	MYB93BAMR	AAA**GGATCC**CTAAGATATAACGTTCATGAG
	MYC-MYB93 5' BAMHI	AAA**GGATCC**ATGGAGGAGCAGAAGCTGAT
	MYC-MYB93 3' NOTI	AAA**GCGGCCGC**CCTAAGATATAACGTTCATG
pGreen0229::pMYB93-myc-ΔgAtMYB93	MYB93 C-term F	AAA**GAATTC**GTGACTCATCAGCCAAGAACC
	MYB93BAMR	AAA**GGATCC**CTAAGATATAACGTTCATGAG
	MYC-MYB93 3' NOTI	AAA**GCGGCCGC**CCTAAGATATAACGTTCATG

Restriction sites: GAATTC, *Eco*RI; GGATCC, *Bam*HI; GCGGCCGC, *Not*I; GATATC, EcoRV

Sequencing primers

Oligonames	*Primer sequence*
T7 promoter	TAATACGACTCACTATAGGG
M13 forward	GTAAAACGACGGCCAGT
M13 reverse	CAGGAAACAGCTATGAC
Nostersac3	AAAGAGCTCCCGATCTAGTAACATAGATGAC
35S_pGreen0229_F	AAGATGCCTCTGCCGACAG
MYB93-internal-R-(full)	GCTCTCCAGCTTCCATGACC
MYB93-internal-R-(trunc)	ACGTTGGAGATATTGTAGATTGGC
Nos-cassette_F	GAACCGCAACGTTGAAGGAG
Nos-cassette_R	GTAACGTTATCAGCTTGCATGCC
Nos-cassette_internal_F	AACCACTACATCCAGACA
Nos-cassette_internal_R	GAAGTCCAGCTGCCAGAAAC

MYB93 crop homologue primers for RTPCR

Gene	*Oligonames*	*5' primer sequence*	*3' primer sequence*
LOC_Os06g11780.1	Rice1_F & Rice1_R	CCTTGGCAACAAGTGGTCAG	CTCCATGCTGCACCACTC
LOC_Os08g37970.1	Rice2_F & Rice2_R	CCTCGGCAACAAGTGGTC	CGGTGTCGTCTGGGTCTG
LOC_Os02g51799.1	Rice3_F & Rice3_R	CTTGGCAACAAGTGGTCAGC	ATTGTCCAGCGAGGAATGAG
HORVU6Hr1G078300.1	Crop_F & Barley1_R	GCAACAAGTGGTCAGCCATC	GGTTGATACCATTGATGATAGGCTG
HORVU6Hr1G078270.6	Crop_F & Barley2_R	GCAACAAGTGGTCAGCCATC	CCACTGATGATAGGCTGTTTGTG
HORVU7Hr1G037060.3	Barley3_F & Barley3_R	CAACAAGTGGTCGGCTATCG	CACTGCAATGCCGCTCAC
Traes_6BL_0D6CD031E.1	Wheat1_F & Wheat1_R	CTGCAAGCCGATGCAGTG	TCTTGGCCATTGATACCCTCTAG
Traes_6AL_2BA02CAA9.1	Wheat2_F & Wheat2_R	CACCATCCCCACTGACCTAC	AACGGTTTCTGGGTGCCG
Traes_6DL_8BF29D46C.1	Wheat3_F & Wheat3_R	GGGATGACCAAAGTGCCAAC	GCGAAGGGCTCAGGAGAC
Traes_6AL_52227091B.1	Wheat4_F & Wheat4_R	AGCTCCAAGCCGATGCAG	CCAAGTCATGGCCATTGACG
Traes_6BL_5A2DB8F78.1	Wheat5_F & Wheat5_R	CTGCGACGACCAAAGCGG	GAGACGACATCGAAGACAAGGATG
Traes_6DL_499F64524.1	Wheat6_F & Wheat6_R	CCAGAACTCAGATTATAGTGCAAATG	GCAAACTCGCTCATGAACTCG
Traes_6BL_0D6CD031E.1 Traes_6AL_2BA02CAA9.1 Traes_6DL_8BF29D46C.1	Crop_F & Wheat1,2,3_R	GCAACAAGTGGTCAGCCATC	GTTTCTGGGTGCCATTCTCC
Traes_6AL_52227091B.1 Traes_6BL_5A2DB8F78.1 Traes_6DL_499F64524.1	Crop_F & Wheat4,5,6_R	GCAACAAGTGGTCAGCCATC	CAGGAGCAGCAGTTTCTCC
	Rice_eEF1a_F & Rice_eEF1a_R	TTTCACTCTTGGTGTGAAGCAGAT	GACTTCCTTCACGATTTCATCGTAA
	Wheat&Barley_ADP_F & 2Barley_ADP_R	GCTCTCCAACAACATTGCCAAC	GAGACATCCAGCATCATTCATTCC
	Wheat&Barley_ADP_F & Wheat_ADP_R	GCTCTCCAACAACATTGCCAAC	GCTTCTGCCTGTCACATACGC

AtARABIDILLO-YFP construct primers

Oligonames	*Primer sequence*
ARAB1.11	GCGATTGGACGTGAAGGTGG
ARAB1.5	CTGGTGCTCTATGGAATTTGTCG
ARAB1.6	CGTGAGGATCTATAAAACTTAGGAC
GFPSAC 3'	GGCGAGCTCTTATTTGTATAGTTCATCCATGCC
GFP Eco 3'	GAAAATTTGTGCCCATTAAC

Yeast one-hybrid primers

Gene	*Oligonames*	*5′ primer sequence*	*3′ primer sequence*
AtCYP708A1	1G55940 promoter F & 1G55940 promoter R	AAA**GCGGCCGC**GGTTCCGTTGTTGTAGCCAAG	AAA**TCTAGA**CAATGCTCTGGCATTGTCG
AtMYB54	1G73410 promoter F & 1G73410 promoter R	AAA**GCGGCCGC**GTTCTGGTAACAGGTTCAAGTCTCAG	AAA**ACTAGT**CTTCACGCCGATCTATTTAATTACG
AtLTP8	2G18370 promoter F & 2G18370 promoter R	AAA**GCGGCCGC**GATAAGTGACAACTTGCCAAATAGC	AAA**GCGGCCGC**AATACATTCATCGTTGAAGAAATTGAG
AtPME10	2G19150 promoter F & 2G19150 promoter R	AAA**GCGGCCGC**ACGAATCTTACTCTCTGGCCAAC	AAA**TCTAGA**TGACATCACAAGTAGACACACCTTG
AtABCG2	2G37360 promoter F & 2G37360 promoter R	AAA**GCGGCCGC**TTAGCTGCCAGAACTGGTTCTC	AAA**ACTAGT**ATTTTCCGAGCAAGCCAGAC
AtGPAT6	2G38110 promoter F & 2G38110 promoter R	AAA**GCGGCCGC**ACCGTAGTTTCAATGGGTCAGAC	AAA**ACTAGT**CATTAGTGAAGAAGTTGGAGAGAGGAG
At2G43670	2G43670 promoter F & 2G43670 promoter R	AAA**GCGGCCGC**GCCTTAGGTCACAGTCACACTAGC	AAA**ACTAGT**GACGTCTGAAAGCTCGTCTACAG
AtPDLP4	3G04370 promoter F & 3G04370 promoter R	AAA**GCGGCCGC**TAAGATGAGGAAGGAGCTTGTGG	AAA**TCTAGA**ATCAAGTGGACAACCATGTTTGAG
At3G27270	3G27270 promoter F & 3G27270 promoter R	AAA**GCGGCCGC**TTGATAATAATTCTCTTCCGTGCAG	AAA**TCTAGA**ACTTGAAGCTTTTGAAGTCGTCG
AtAHA4	3G47950 promoter F & 3G47950 promoter R	AAA**GCGGCCGC**GTTAAACGGTAAATCTCGTATTGCAG	AAA**GCGGCCGC**CGCACTGCACCGAAATAAAAC
At4G17480	4G17480 promoter F & 4G17480 promoter R	AAA**GCGGCCGC**TGGTTTAACAGTTAATCTTTCTTTGGC	AAA**ACTAGT**TCCATGGAACAACAACTTAGAAGAG
AtS-ACP-DES1	5G16240 promoter F & 5G16240 promoter R	AAA**GCGGCCGC**GATCAGTCTTGAAACAACAATGAGG	AAA**TCTAGA**CGATCCATAGCCATAACCATTG
AtCASPL1B1	5G44550 promoter F & 5G44550 promoter R	AAA**GCGGCCGC**GCCGGAAATTGAGGTAATTGAG	AAA**TCTAGA**GTTTGAAGATGAAAATGGTCAAATTG
AtCYS5	5G47550 promoter F & 5G47550 promoter R	AAA**GCGGCCGC**TCGTAGGTAGAAGTTGGCCATTG	AAA**GCGGCCGC**GAAGACGACCTTACTAGTCATTTTGC
At5G49350	5G49350 promoter F & 5G49350 promoter R	AAA**GCGGCCGC**GTAACTTGTGGCTCCCAACATAAG	AAA**ACTAGT**GAAAACGCCATTGCTAAGAGTTTAC
AtQRT1	5G55590 promoter F & 5G55590 promoter R	AAA**GCGGCCGC**ATTAATCATCTTCTGATTTGCCCTTC	AAA**ACTAGT**GAATGATGTCTCCTGCACAAGG

Restriction sites: GCGGCCGC, *Not*I; ACTAGT, *Spe*I; TCTAGA, *Xba*I

Appendix 3.1 Genes identified from bioinformatic analysis to select candidate downstream targets of AtMYB93. *Grey = peak expression of gene at 12 vs 9 h, green = peak expression of gene at 15 vs 12 h, orange = 18 vs 15 h, pink = shortlist candidate gene, blue = final selection of candidate genes. Time point data from Voß et al. 2015. Red = significant upregulation of gene relative to previous time point, blue = significant downregulation of gene relative to previous time point. Differential expression data from EMBL-EBI. Red = gene is upregulated, blue = gene is downregulated, white = no/inconclusive data. For eFP data 1 = high expression in this tissue, 0.5 = moderate expression in this tissue, 0 = no or very limited expression in this tissue.*

						Differential Expression Data							eFP Data		
AT	Gene name	12h vs 9h	15h vs 12h	18h vs 15h	48h vs 45h	ABA	Auxin	Hypoxia	Iron	Pathogen	Pi	Salt/ Drought	Root	Seed	Other
AT5G37690	AT5G37690	2.2	3.9	1.4		0	1	0	1	0	1	0	0.5	1	0
AT5G49350	AT5G49350	1.3	3.7	1.3	-1.1	1	1	NA	1	0	1	1	0.5	1	0
AT5G44550	AT5G44550	2.7	3.6	1.5	-1.1	1	1	0	1	0	1	0	0.5	1	0
AT1G78990	AT1G78990	1.9	3.4	1.1	-1.8	0	0	0	1	0	1	0	1	0	0
AT2G47200	AT2G47200	2.7	3.2	1.2		1	1	NA	1	1	1	0	0.5	1	0
AT4G24140	AT4G24140	1.7	3.1	1.2		0	1	NA	1	1/0	1	NA	0	0.5	1
AT2G19970	AT2G19970	1.5	2.9			0	1	0	NA	1	0	0	1	0.5	0.5
AT1G49960	NAT4	2	2.8	1.2		1	1	NA	NA	0	1	0	0	1	0
AT2G18370	LTP8	1.5	2.8	1.7	-0.1	1	1	0	NA	0	1	1	0	1	0
AT2G35380	PER20	3.3	2.8	1	-1.1	0	1	0	NA	0	1	0	1	0	0
AT2G01520	MLP328	1.3	2.7	2.1		NA	0	NA	NA	0	1	1/0	1	0	0
AT2G23540	AT2G23540	4.9	2.6	1	-1	1	1	0	NA	0	1	0	0	1	0
AT5G58860	CYP86A1	4.2	2.6	1.1		1	1	0	NA	0	1	0	1	1	0
AT2G43670	AT2G43670	1.5	2.5	1	-0.8	1	NA	NA	NA	0	1	NA	0	1	0
AT3G04370	PDLP4	0.9	2.4	1.2	-1.2	NA	1	NA	NA	NA	1	1	0	1	0
AT4G20390	AT4G20390	3.3	2.4			NA	1	0	NA	0	1	1/0	0.5	1	0
AT5G09480	AT5G09480	4.2	2.4	1.1		0	1	0	1	0	1	1/0	0	1	0
AT4G00360	CYP86A2	2.4	2.3			1	NA	NA	NA	0	NA	1/0	0	1	1
AT2G19200	AT2G19200	1.6	2.2			NA	1	NA	NA	NA	1	NA	1	1	0.5
AT2G43390	AT2G43390	2.5	2.2		-1.4	NA	1	NA	1	0	1	NA	0	1	0
AT3G47950	AHA4	0.5	2.2	1.3	-0.7	1	NA	NA	NA	0	1	NA	0	1	0.5
AT4G34510	KCS2		2.2			NA	NA	NA	1	0	NA	NA	0	0.5	1
AT5G01870	LTP10		2.2	3.1		0	1	0	NA	1/0	1	NA	1	1	0.5
AT1G18250	ATLP-1	1.4	2.1	1.1		NA	1/0	0	NA	0	0	0	0.5	0	1
AT2G22510	AT2G22510	3.9	2.1		-1.3	1	1	NA	NA	0	1	0	0	1	0
AT2G38110	GPAT6	1.7	2.1	1.1	-0.7	1	1	NA	NA	0	1	0	0	0	1
AT2G46830	CCA1	1.6	2.1		-1	NA	1/0	0	NA	1/0	0	1/0	0	0.5	1
AT4G30140	CDEF1		2.1	1.8		0	1	1/0	1	1	NA	0	0	0	0.5
AT4G38080	AT4G38080	2.7	2.1			1	1	NA	NA	0	1	0	0	1	0
AT3G07970	QRT2	0.7	2	1		1	1	NA	1	1/0	NA	1	0	1	0
AT4G23680	AT4G23680		2	2.1		1/0	0	0	NA	1/0	1	1	0	0	1
AT1G04360	ATL1	1.5	1.9			1	1/0	NA	NA	1/0	NA	1/0	0.5	1	0
AT3G48450	AT3G48450	2	1.9	1.1	-1.1	1	1	1	0	1/0	1/0	1	0	0	0
AT5G45670	AT5G45670		1.9	2.3		0	1	0	NA	0	0	0	0	1	1
AT1G49430	LACS2	2.4	1.8			1/0	NA	1	NA	0	1	1/0	0	0.5	1
AT1G75030	ATLP-3	2.1	1.8			1	1	0	1	0	NA	1/0	0	0	1
AT2G03200	AT2G03200	1.9	1.8			1	1	0	NA	0	1	1	0	0.5	1
AT3G16530	AT3G16530		1.8	1.5		1/0	0	1/0	1	1	1/0	1/0	0	0	1
AT3G19710	BCAT4	1.6	1.8	1.6		0	1/0	0	NA	1/0	1	1/0	0.5	0	1
AT5G19140	ATAILP1		1.8	1.1	-1.2	0	1/0	1	NA	1/0	0	1/0	0	0.5	1
AT5G20860	PME54	1.8	1.8			1	NA	NA	NA	0	NA	NA	0	1	0
AT1G68850	PER11	2.7	1.7			1	1	NA	1	0	1	0	0	1	0
AT1G73290	scpl5	1.8	1.7			NA	NA	NA	NA	NA	NA	NA	0	1	0
AT2G37360	ABCG2	1.3	1.7	0.4	-0.6	NA	1	NA	NA	0	1	NA	0	1	0
AT3G58350	RTM3	1.4	1.7	-1.3		NA	NA	NA	NA	0	NA	NA	1	0.5	1
AT4G17215	AT4G17215		1.7	0.5	-0.9	1	1	NA	NA	1	1	1/0	0	1	0
AT5G07130	LAC13	2.4	1.7			NA	1	NA	1	0	1	NA	1	1	1
AT5G09530	AT5G09530	3.1	1.7		-1.2	1	1	1	NA	1/0	1	1/0	0	1	0
AT5G57780	AT5G57780	1.1	1.7			NA	0	0	NA	0	0	0	0	0.5	1
AT1G55940	CYP708A1	0.6	1.6	0.9	-0.5	NA	1	NA	NA	0	1	NA	1	1	0.5
AT1G64000	WRKY56	1.6	1.6	0.2	-0.7	1	1	NA	NA	NA	1	NA	0	1	0
AT2G21100	DIR23	0.9	1.6	0.3	-1.4	1	1	0	1	0	1	1/0	0.5	1	0
AT2G27140	AT2G27140	2	1.6	-1.4		0	1	NA	NA	1	NA	1/0	1	0.5	0.5
AT2G44260	AT2G44260	-0.9	1.6	0.8		1	NA	0	NA	0	1	1	0	1	0
AT3G06390	AT3G06390	2	1.6			0	1/0	NA	NA	0	1/0	1/0	0	1	0
AT4G16640	AT4G16640	1.2	1.6	-0.5		1	1	NA	NA	0	NA	0	0	1	1
AT5G19410	ABCG23	2	1.6		-1	1	1	NA	NA	0	1	0	0	1	0
AT1G17950	ATMYB52	1	1.5	0.2	-0.5	1	1	NA	NA	0	NA	1	0	0.5	1

X

AT1G31490	AT1G31490	0.6	1.5	0.5		NA	NA	NA	1	NA	NA	NA	0	0.5	1
AT1G51640	ATEXO70G2	1.3	1.5			NA	1	NA	NA	1	NA	NA	0	1	0.5
AT1G63530	AT1G63530		1.5	1.2		NA	1	1	NA	1/0	1	1	0	0	1
AT1G68230	RTNLB14	1.3	1.5			0	0	NA	NA	1/0	NA	NA	1	1	1
AT2G28950	EXPA6		1.5			0	1	0	NA	1/0	0	0	0	0.5	1
AT3G07420	SYNC2	2.2	1.5	-1.3		NA	1	NA	NA	1	NA	NA	1	1	1
AT4G11650	OSM34		1.5			1/0	0	1	1/0	1	1	1	1	0.5	0
AT5G02540	AT5G02540		1.5	1.1		0	1	NA	1	0	0	0	0	1	1
AT5G04950	NAS1	1.8	1.5			0	0	0	NA	0	0	0	0	0	1
AT5G55180	AT5G55180	1.9	1.5			1	1	NA	1	0	NA	1	1	1	1
AT5G55590	QRT1	0.5	1.5	1.2	-1.1	1	NA	NA	NA	NA	NA	NA	0	0.5	1
AT1G20030	AT1G20030		1.4			1	0	0	NA	1/0	1	1	0	1	1
AT1G73370	SUS6	2	1.4			0	0	NA	NA	0	NA	1	1	1	1
AT1G73410	ATMYB54	1.1	1.4	0.9	-0.6	1	1	NA	NA	1	1	1	0.5	1	1
AT1G74460	AT1G74460	2	1.4		-1.3	1	1	0	1	1/0	1	0	1	1	0
AT2G39310	JAL22	1	1.4			1/0	0	0	1	1/0	0	1/0	0.5	0	0
AT3G11430	GPAT5	1.7	1.4			1	1	NA	NA	1/0	1	1/0	0	0	1
AT3G18400	anac058	1.4	1.4	0.5	-0.6	1	1	NA	NA	0	1	1	0	1	0
AT3G25855	AT3G25855	0.5	1.4	-0.3		NA	NA	NA	NA	0	NA	NA	0	1	0
AT4G14440	ECI3	1.7	1.4			1	1	0	NA	0	NA	1/0	0	1	1
AT4G35420	TKPR1	1.2	1.4		-1	NA	1/0	NA	NA	0	NA	0	0	0	1
AT5G04220	SYT3		1.4			1/0	0	1	NA	1	NA	0	0	0.5	1
AT5G08250	AT5G08250	2	1.4			1	1	NA	NA	0	1	NA	0	1	1
AT5G09520	AT5G09520	2	1.4		-1.7	1	1	0	1	0	1	0	0	1	0
AT5G23020	MAM3	1.5	1.4			0	0	0	NA	1/0	1	0	1	0	0.5
AT5G47550	CYS5	0.5	1.4	1.2	-0.5	1/0	1	0	1	1/0	1	1	0	1	1
AT1G05310	PME8	-1.2	1.3			NA	0	NA	1	0	0	1/0	0.5	0	1
AT1G05760	RTM1	1.9	1.3			0	0	NA	NA	0	NA	1/0	1	1	1
AT1G28290	AGP31	0.5	1.3	0.3		1/0	1	0	NA	0	NA	0	1	0	1
AT1G43790	TED6		1.3			0	1	0	NA	0	NA	1/0	0	0	1
AT1G43800	S-ACP-DES6		1.3			NA	1/0	1	NA	1	0	1	0	1	1
AT1G67870	AT1G67870		1.3			NA	0	0	NA	0	1	0	0	1	1
AT2G02850	ARPN		1.3	2.1		1	1	0	0	1/0	1	1/0	0	1	0.5
AT2G19150	PME10	0.9	1.3	0.3	-0.7	1	NA	NA	0	NA	NA	NA	1	1	0.5
AT2G24430	ANAC039	0.2	1.3	0.8	-0.2	NA	1	NA	NA	0	NA	NA	0	1	0
AT2G40370	LAC5	2.8	1.3			1	1	NA	1	0	1	1/0	0	1	0
AT2G43800	FH2	0.2	1.3		-0.2	NA	1	NA	NA	1/0	1/0	1/0	0	0.5	1
AT3G07320	AT3G07320	1	1.3			0	1	0	NA	0	NA	0	1	0.5	1
AT3G13760	AT3G13760		1.3			0	0	0	NA	0	0	0	1	0.5	0
AT3G27270	AT3G27270	0.2	1.3	0.5	-0.9	1	1	NA	NA	NA	1	1	1	1	1
AT3G44540	FAR4	1.6	1.3		-1	1	1	0	NA	0	1	1/0	0	1	0
AT3G47180	AT3G47180	1.4	1.3			NA	1	NA	NA	NA	NA	1/0	1	0.5	1
AT4G13770	CYP83A1	1.1	1.3	1.1	-0.3	1	1/0	0	NA	1/0	1	0	0	0.5	1
AT4G17480	AT4G17480	0.4	1.3	1.1	-0.4	1	NA	NA	NA	NA	NA	1	0	1	0.5
AT5G12970	AT5G12970	0.8	1.3	0.8		NA	1	NA	NA	1	NA	0	0	1	0
AT5G47635	AT5G47635	0.8	1.3	0.5	-0.7	1	1	NA	NA	1	1	NA	0	1	1
AT5G65230	AtMYB53	0.5	1.3	0.6	-0.6	1	NA	NA	NA	NA	NA	NA	1	0.5	1
AT1G01600	CYP86A4	1.3	1.2			1	1	0	NA	0	1	1	0	0	1
AT1G05450	AT1G05450	1.4	1.2			1	1	NA	NA	0	1	NA	0	1	0.5
AT1G06040	BBX24		1.2			0	0	1/0	NA	0	0	NA	0	1	1
AT1G34670	AtMYB93		1.2			1	1	1	NA	NA	1	NA	1	0	0
AT1G53100	AT1G53100		1.2			1	NA	NA	NA	1	1	1	0	1	1
AT1G56320	AT1G56320	0.9	1.2	0.4	-0.7	1	1	0	1	0	1	0	0	1	0
AT1G80520	AT1G80520		1.2			0	0	NA	NA	0	NA	1/0	1	1	1
AT2G16280	KCS9		1.2			NA	0	0	NA	0	NA	1/0	0	1	1
AT2G26910	ABCG32		1.2			0	0	0	NA	0	NA	0	0	0	1
AT2G28780	AT2G28780		1.2			0	0	NA	1	0	0	0	1	0.5	0
AT2G29340	AT2G29340	1	1.2			NA	0	1	NA	1/0	0	1/0	0	1	1
AT3G49070	AT3G49070		1.2			NA	NA	NA	NA	NA	NA	0	1	1	1
AT3G54500	AT3G54500		1.2			0	0	0	NA	1/0	0	0	0	1	1
AT4G19980	AT4G19980		1.2			0	1/0	1	NA	NA	NA	NA	1	1	1

AT4G26790	AT4G26790	1.4	1.2			0	1	0	NA	0	1	0	0	1	1
AT5G13580	ABCG6	2.2	1.2		-1.3	1	1	NA	NA	0	1	1/0	0	1	1
AT5G23010	MAM1	1.4	1.2	1.1		NA	0	NA	NA	1/0	NA	0	0	0	1
AT5G54040	AT5G54040		1.2			1/0	1/0	1	NA	1/0	1	1/0	1	0.5	0.5
AT5G61570	AT5G61570		1.2			NA	0	1	NA	0	NA	0	0.5	0.5	1
AT1G01610	GPAT4	1	1.1	-0.1		1	1	0	NA	0	NA	1	0	0.5	1
AT1G27950	LTPG1		1.1	3.1		1	1/0	0	NA	0	NA	1/0	0	1	0.5
AT1G73480	AT1G73480		1.1			1	1	1	NA	1/0	1	1	0	0	0
AT1G77590	LACS9	1	1.1			1	1	NA	NA	1	NA	0	0	0.5	1
AT2G05910	AT2G05910	0.2	1.1	1	-0.6	1	1/0	NA	NA	1/0	NA	NA	0	1	0.5
AT2G18980	PER16		1.1			0	1/0	0	NA	0	1	0	1	0	0
AT2G31380	BBX25		1.1		-1.2	0	1/0	0	NA	1/0	0	1/0	0	1	1
AT2G43140	BHLH129		1.1	1.3		0	1/0	0	0	1	1	1	0.5	1	0
AT2G46130	WRKY43	1.2	1.1			1	1	NA	1	0	1	NA	0	1	0
AT3G11210	CPRD49		1.1			0	1	NA	NA	1/0	NA	0	0	0	1
AT3G26380	AT3G26380	1.2	1.1			1	NA	NA	NA	1/0	NA	1/0	0	1	0.5
AT3G44550	FAR5	1.4	1.1		-1.1	1	1	NA	NA	1	1	1/0	1	1	1
AT3G50400	AT3G50400	1.4	1.1		-1	1	1	NA	1	1/0	1	1/0	1	1	0
AT3G60980	AT3G60980		1.1			1	0	0	NA	1	1	1	1	0.5	1
AT4G11050	AtGH9C3		1.1	0.3		NA	1	NA	NA	NA	1	0	1	0.5	0.5
AT4G17670	AT4G17670		1.1			NA	1	1	NA	1	NA	1/0	0	0.5	1
AT4G18910	NIP1-2		1.1			0	1	0	NA	1/0	NA	1	0	1	0
AT4G26760	MAP65-2		1.1			0	1	0	NA	1/0	0	0	0.5	0.5	1
AT4G28250	EXPB3		1.1			0	1/0	0	NA	0	1/0	1/0	0	0.5	1
AT4G37970	CAD6	0.7	1.1			1	1/0	NA	NA	0	1	1	0.5	0	0.5
AT4G38970	FBA2		1.1			0	0	NA	1	1/0	0	0	0	0.5	1
AT5G04890	RTM2	1.3	1.1			0	NA	NA	NA	0	NA	1	1	1	1
AT5G23190	CYP86B1	1.3	1.1		-1.2	1	1	NA	NA	0	1	0	0	1	0
AT5G52170	HDG7	0.2	1.1	0.8		NA	1	NA	NA	NA	NA	NA	1	0.5	1
AT1G01060	LHY	1.6	1			0	1	0	NA	1/0	0	1/0	0	0.5	1
AT1G03700	AT1G03700	1.7	1			0	1	NA	NA	NA	1	NA	0.5	0.5	1
AT1G28400	AT1G28400		1			0	1	0	NA	1/0	0	1/0	1	1	0
AT1G48480	RKL1		1	1.1		1/0	0	0	NA	0	0	0	0	0	1
AT1G53270	ABCG10	0.9	1		-0.6	1	1	NA	NA	0	1	NA	0	1	0
AT1G55330	AGP21		1			0	0	0	NA	0	NA	0	1	1	1
AT1G75520	SRS5	1	1			0	1	NA	NA	NA	NA	NA	0	0.5	1
AT1G78020	AT1G78020		1			0	0	1	0	0	NA	0	0	0	1
AT2G01900	AT2G01900		1			0	0	NA	NA	0	0	0	1	0	0
AT2G02450	ANAC034		1			0	0	1	NA	0	0	1/0	0	0.5	1
AT2G03280	AT2G03280		1			1	NA	NA	NA	1	NA	NA	0	1	0.5
AT2G46450	CNGC12		1			0	1	1	NA	1/0	NA	0	0	1	1
AT2G47260	WRKY23		1			1/0	1	NA	NA	1	0	1	1	1	0.5
AT3G07390	AIR12		1			0	1	0	NA	1	1	1	0	1	1
AT3G15790	MBD11	0.8	1			1	1	NA	NA	1/0	NA	1	0	1	1
AT3G45010	SCPL48		1			1	1	1	NA	0	NA	1	0	0.5	1
AT3G45160	AT3G45160		1			0	0	NA	NA	0	0	0	1	0	1
AT3G62590	AT3G62590		1			1	1	NA	NA	1	1	1	0	1	1
AT4G02850	AT4G02850		1	1.1		0	0	0	NA	0	0	0	0.5	0	0.5
AT4G08780	PER38		1	1.9		0	1	NA	NA	1	1/0	1	1	0	0
AT4G16260	AT4G16260		1			1	1	0	NA	1	1	1/0	1	0	1
AT4G24780	AT4G24780		1			0	1/0	0	NA	0	0	0	1	1	0.5
AT4G28940	AT4G28940		1	1		1	0	0	1	1/0	1	NA	1	0	1
AT4G39720	AT4G39720		1			0	1	1	NA	1	NA	1	1	1	1
AT5G14760	AO		1			0	1	1	NA	1/0	0	1/0	0	1	1
AT5G16240	S-ACP-DES1	0.9	1	0.7	-0.6	1/0	1	NA	NA	0	NA	1	0	1	1
AT5G56720	AT5G56720	1.8	1			0	1/0	NA	NA	1	NA	1	1	1	1
AT5G57660	COL5		1			0	0	1/0	NA	0	1/0	1/0	0	0.5	1
AT5G67070	RALFL34		1	0.6		1/0	1	0	NA	1/0	NA	0	0	0.5	1
AT1G08340	ROPGAP5	-1.2	-1			NA	1	NA	NA	1/0	NA	1	0	0	1
AT1G23720	AT1G23720		-1			0	NA	0	NA	0	1	1/0	1	0	0
AT1G60010	AT1G60010		-1			0	1	1	NA	1/0	NA	0	0	0.5	1

AT1G62570	FMOGS-OX4		-1		-0.2	1	1	1/0	NA	1/0	1	1	0.5	1	1
AT2G03980	AT2G03980	-0.8	-1			1	1/0	0	NA	0	NA	1/0	0	0	0
AT2G27740	AT2G27740		-1	-1		0	1	0	NA	0	NA	0	1	0	1
AT2G45750	AT2G45750		-1	-1		1/0	1/0	0	NA	0	1	1/0	1	0.5	1
AT3G16180	NPF1.1		-1			0	0	0	NA	1/0	0	1/0	0	1	1
AT3G16690	SWEET16		-1			1	0	NA	0	0	1	0	1	0	0.5
AT3G19390	AT3G19400		-1			0	1/0	1	NA	1/0	1	1	0	0	1
AT3G20820	AT3G20820		-1			1/0	0	0	NA	0	NA	0	0	0	1
AT3G28210	SAP12		-1			1	1	1	NA	1	1/0	1	0	0	0
AT3G59710	AT3G59710	-1	-1			NA	0	0	NA	0	1	1	0.5	0.5	1
AT4G15330	CYP705A1	-1	-1			NA	1/0	NA	NA	1/0	NA	1	1	0	0
AT4G16146	AT4G16146	-1	-1			1	1	1	NA	1/0	1	1	0	1	0.5
AT4G16750	ERF039		-1			1	1/0	0	NA	0	1	1	0.5	1	1
AT4G18510	CLE2	-1.3	-1			NA	0	0	1	0	1/0	1/0	1	0	0
AT4G26320	AGP13	-1.2	-1			0	0	0	NA	0	1	1/0	1	0	1
AT4G39675	AT4G39675		-1	-1.3		1	1/0	1	0	1/0	0	1/0	1	0	0
AT5G06570	CXE15		-1			0	0	1	NA	1/0	1	1	1	1	0.5
AT5G17340	AT5G17340	-1	-1	-1.8		NA	1	1	NA	NA	1	NA	0	0	0.5
AT5G24410	PGL4		-1			0	0	NA	NA	0	NA	0	1	0.5	0
AT5G26930	GATA23	-1.5	-1			NA	1	NA	NA	1	NA	NA	0.5	0.5	0.5
AT5G37180	SUS5		-1	-1.3		NA	1	NA	NA	0	NA	NA	1	1	0.5
AT5G44110	ABCI21		-1	-1.1		1/0	1	NA	NA	0	1	1	0	0	1
AT5G56540	AGP14	-1.4	-1			1	0	0	NA	0	1	0	0	1	0.5
AT5G60660	PIP2-4	-1.2	-1			1/0	0	0	NA	0	NA	0	1	0	0.5
AT1G01750	ADF10	-2.7	-1.1			0	1/0	0	NA	0	1	1/0	1	0	1
AT1G03870	FLA9	-1.6	-1.1		1	0	0	0	NA	1/0	0	0	1	0	1
AT1G05340	AT1G05340		-1.1			1	1/0	1	NA	1/0	1/0	1	0	0.5	1
AT1G27380	RIC2		-1.1			0	1	NA	NA	0	NA	1	0	0.5	1
AT1G30820	AT1G30820		-1.1			0	1/0	1/0	0	1/0	0	1	0.5	1	1
AT1G75720	AT1G75720		-1.1			NA	NA	NA	NA	0	NA	NA	1	1	1
AT1G78070	AT1G78070		-1.1			1	1	1	NA	0	1	1	0	1	0
AT2G20670	AT2G20670	-1	-1.1	-1		1	0	1/0	NA	0	1/0	0	0	0	1
AT2G21880	RABG2	-2	-1.1			1	0	0	NA	0	1	0	0	0	0
AT2G30930	AT2G30930	-1.5	-1.1			1	0	0	NA	1/0	0	0	0.5	0	1
AT3G30775	POX1		-1.1	-1		1	1/0	1	NA	1/0	0	1/0	0	0.5	0
AT3G48940	AT3G48940	-1.2	-1.1			NA	0	0	NA	NA	NA	0	1	1	0.5
AT4G04840	MSRB6		-1.1		-1.2	0	0	1	NA	0	1	1/0	0	1	0.5
AT4G09990	GXM2		-1.1	-1		0	0	0	NA	0	1	0	0	0	1
AT4G12580	AT4G12580		-1.1			1	1	NA	NA	1	NA	1	0	0.5	1
AT4G29180	RHS16	-1.5	-1.1			0	0	0	NA	NA	1	0	1	0.5	0.5
AT4G37220	AT4G37220		-1.1			1	0	1	NA	0	1	0	0	0.5	1
AT4G39660	AGT2		-1.1			0	0	1	NA	1	NA	0	1	0.5	0.5
AT5G06630	AT5G06630	-1.7	-1.1			0	0	0	1	0	1	1/0	1	0.5	1
AT5G06690	WCRKC1		-1.1			NA	0	1	NA	1/0	1/0	0	0	0	1
AT5G09440	EXL4		-1.1			1/0	1/0	1	NA	1/0	1	1/0	0	0	1
AT5G14090	AT5G14090		-1.1			0	0	NA	NA	NA	0	1/0	0	0	1
AT5G19800	AT5G19800	-2.4	-1.1			0	1/0	0	1	NA	1	0	1	1	1
AT5G22920	AT5G22920		-1.1	-1		0	1/0	1/0	NA	1/0	0	0	0.5	0.5	1
AT5G24570	AT5G24570	-1	-1.1			1/0	0	0	NA	0	1/0	1	0	0.5	0.5
AT5G42590	CYP71A16		-1.1			0	1/0	0	NA	0	1/0	0	1	0	0
AT5G45840	AT5G45840	-0.7	-1.1	-0.2		NA	1	NA	NA	1/0	NA	0	0	0	0
AT5G61440	ACHT5		-1.1			0	0	1	NA	1/0	0	0	0.5	0.5	1
AT5G62210	AT5G62210	-0.8	-1.1	-1	-0.5	1	1	NA	NA	1/0	1	1/0	0	0.5	1
AT1G62480	AT1G62480	-1	-1.2			0	0	0	NA	0	0	0	0.5	0.5	1
AT1G70720	AT1G70720	-0.1	-1.2	-0.3		NA	NA	NA	NA	NA	NA	1	0	0	1
AT2G19410	PUB34	-1.4	-1.2			NA	NA	NA	NA	NA	0	0	1	0.5	1
AT2G22470	AGP2		-1.2			1	1	0	NA	1	1	1	1	0.5	0.5
AT2G24610	CNGC14	-1.9	-1.2		1	NA	1	NA	NA	0	1	NA	1	0	0
AT2G32150	AT2G32150		-1.2			1/0	1/0	1	NA	1/0	1/0	1/0	1	1	1
AT2G43590	AT2G43590		-1.2			1	1	1/0	NA	1	1	1/0	1	1	0
AT2G47930	AGP26		-1.2			NA	1/0	0	NA	0	0	0	0	1	0.5

AT3G28550	AT3G28550	-1.4	-1.2	-1.5		0	1/0	NA	NA	0	1	1/0	NOT FOUND				
AT4G02090	AT4G02090	-1.6	-1.2			0	1	NA	1	0	1	0	1	0.5	0		
AT5G15290	CASP5	-3.8	-1.2			NA	0	0	0	NA	1	1	1	1	0.5		
AT5G20250	RFS6		-1.2	-1.2		0	0	1	NA	1/0	0	0	0.5	0.5	1		
AT5G53250	AGP22	-1.6	-1.2			0	0	0	NA	0	1	1/0	1	0	0		
AT1G17190	GSTU26		-1.3			0	0	0	NA	1/0	1	1	1	0.5	1		
AT2G34180	CIPK13		-1.3	-1.2		NA	1/0	1	NA	0	0	1	1	0	0		
AT3G14850	TBL41	-1	-1.3			1/0	1/0	0	NA	1	NA	1/0	0	1	0		
AT3G45700	NPF2.4	-1	-1.3			NA	0	0	NA	0	NA	1/0	1	0	0		
AT4G08040	ACS11	-0.5	-1.3	-0.5	-1	1	1	NA	0	0	1	1/0	0	0	1		
AT4G14130	XTH15	-1	-1.3			1	1/0	0	NA	0	0	1/0	1	1	0.5		
AT5G01190	LAC10	-0.4	-1.3	-0.4	-0.4	NA	1	NA	NA	NA	NA	1	0	0	1		
AT5G06640	AT5G06640	-1.1	-1.3			0	1/0	0	1	0	1	0	1	0.5	0		
AT5G38710	POX2	0.3	-1.3	0.1	0.1	1	1/0	NA	NA	1/0	1	1	0	0	1		
AT5G44610	PCAP2	-1	-1.3			0	0	NA	0	1/0	1	1/0	0	0	0.5		
AT1G05260	PER3		-1.4			0	1/0	0	NA	1/0	1	0	1	0.5	0		
AT1G12090	ELP	-2	-1.4			0	0	NA	NA	0	NA	0	0.5	1	1		
AT1G60030	NAT7	-0.2	-1.4	-0.2	-0.1	NA	1	NA	NA	NA	NA	1	0	1	0.5		
AT1G67070	PMI2		-1.4			1/0	0	1	NA	1	NA	1	0	0.5	1		
AT2G14900	GASA7	-0.9	-1.4	-1.4	0.8	1/0	1	0	NA	1/0	NA	0	0	1	0.5		
AT2G25240	AT2G25240	-1	-1.4			NA	1/0	NA	1	1	1	0	1	1	0.5		
AT2G34810	AT2G34810		-1.4	-1.5		1	1/0	1	NA	1	1	1	0	0	1		
AT2G35980	YLS9		-1.4			1/0	1/0	1	0	1	1	1	0	0	1		
AT3G57520	RFS2		-1.4			1	0	1	0	1/0	0	1	1	1	1		
AT4G08770	PER37	-2	-1.4			0	1	1	NA	1	1/0	1	1	0	0		
AT5G05410	DREB2A		-1.4			1	1/0	1	0	1	1	1	1	1	0.5		
AT5G17220	GSTF12		-1.4	-1.7		1	0	0	NA	1/0	1	1	0	0.5	1		
AT5G44130	FLA13		-1.4			0	0	0	NA	0	1	0	0	0	1		
AT1G34510	PER8	-2.1	-1.5			0	1	0	1	NA	1	1/0	1	0	0		
AT1G43160	RAP2-6		-1.5	-2.1		1	0	1	NA	1/0	1	1	0	1	0		
AT1G62510	AT1G62510	-1.3	-1.5	-1.2		1	0	1	NA	1/0	0	1/0	0	1	0		
AT2G19800	MIOX2		-1.5	-1.1		0	1/0	1	NA	1/0	0	1/0	0	0	0		
AT2G34910	AT2G34910		-1.5			0	1/0	0	NA	0	1	1	1	0	0		
AT3G01680	SEOB		-1.5	-1.2		0	1/0	NA	NA	0	0	1/0	0.5	0	0.5		
AT3G45730	AT3G45730		-1.5			1	1	1	NA	1/0	0	1/0	0	0.5	1		
AT3G60140	BGLU30		-1.5			1	1/0	1	0	1	NA	1	0	1	1		
AT3G62950	GRXC11		-1.5			NA	0	1/0	NA	1/0	0	0	0	1	1		
AT4G25830	AT4G25830	-1.2	-1.5	-1		0	0	NA	NA	0	NA	0	0	1	1		
AT4G30320	AT4G30320	-2.4	-1.5			0	1/0	0	1	NA	1	1/0	1	0	0		
AT5G24140	SQE4		-1.5			1	0	0	1	1	1	NA	1	0.5	0.5		
AT5G45280	PAE11		-1.5			1/0	1	1	NA	0	1	1	1	0.5	1		
AT5G47240	NUDT8		-1.5			1	0	1/0	NA	1/0	1	1	0.5	0	1		
AT1G02813	AT1G02813	-1.7	-1.6			NA	NA	NA	NA	NA	NA	NA	0	0	1		
AT1G22880	CEL5		-1.6	-1.4	-1.5	0	1/0	0	NA	1	NA	0	1	1	0.5		
AT1G43020	AT1G43020	-2	-1.6			0	1/0	NA	NA	NA	NA	0	1	0.5	0.5		
AT1G52240	ATROPGEF11		-1.6			0	1/0	1	NA	0	NA	1/0	1	0.5	1		
AT2G22980	SCPL13		-1.6			NA	0	0	NA	0	1/0	0	0	0	1		
AT3G10710	PME24	-1.7	-1.6			0	1/0	0	1	NA	1	0	1	0	0		
AT3G45970	EXLA1		-1.6			1	1/0	1	NA	1/0	1	1/0	0.5	1	1		
AT5G66390	PER72		-1.6		-1.1	0	1	1/0	NA	0	1	1/0	1	0	0		
AT1G15040	AT1G15040		-1.7	-1.1		NA	1/0	1	NA	1/0	1	1	1	0	0		
AT1G62980	EXPA18	-2.1	-1.7			0	1/0	0	1	NA	1	1/0	1	0.5	0		
AT2G24980	AT2G24980	-1.7	-1.7			0	0	0	1	0	1	0	1	0	0		
AT2G44110	MLO15	-1.1	-1.7			0	0	NA	NA	0	1	0	1	1	0		
AT3G18450	PCR5		-1.7			0	0	1	0	NA	1	0	1	0	0		
AT4G25790	AT4G25790	-2.4	-1.7			0	1/0	1	1	1	1	1/0	1	0.5	0		
AT5G02230	AT5G02230	0.2	-1.7	-0.6		1	1/0	0	NA	1/0	1	1	0.5	1	1		
AT5G06200	CASP4	-2.5	-1.7			0	1/0	0	NA	0	1	1/0	1	0.5	0		
AT5G35190	AT5G35190	-4.2	-1.7			0	1/0	0	1	0	1	0	1	0	0		
AT5G41080	GDPD2		-1.7			0	1/0	1	0	1	1	1/0	1	1	0.5		
AT5G52790	CBSDUF5	-1.2	-1.7			0	1/0	NA	NA	0	0	0	NOT FOUND				

AT2G25150	SCT		-1.8			NA	NA	0	NA	NA	1	0	1	0	0
AT3G17130	AT3G17130	-1.3	-1.8	-0.6		1	1/0	NA	NA	0	1	1	0	1	0
AT5G06760	LEA4-5		-1.8			1	1	1	NA	1	1	1	0	1	0
AT5G22870	AT5G22870	0.3	-1.8	-1.6	-0.3	NA	1	NA	NA	NA	NA	0	1	0.5	0.5
AT3G54590	ATHRGP1		-1.9			0	0	0	1	0	1	1/0	1	0.5	0
AT4G00680	ADF8	-2.2	-1.9			1/0	1/0	0	1	0	1	0	1	0	0
AT4G34580	SFH1	-2.4	-1.9			0	1/0	0	NA	0	1	1/0	1	0	1
AT5G40730	AGP24	-1	-1.9			1/0	0	0	NA	1/0	NA	1/0	0	0	0
AT4G33730	AT4G33730	-2.5	-2.1			0	1/0	0	1	NA	1	0	1	0.5	0.5
AT5G15180	PER56	-1.5	-2.1	-1.3		NA	1/0	0	NA	0	1	1/0	1	1	0
AT5G47450	TIP2-3	-1	-2.1	-1.6		0	0	0	NA	0	0	1/0	1	0	0
AT5G58010	BHLH82	-1.2	-2.1	-1.1		NA	1/0	NA	NA	NA	1	1	0.5	0.5	0
AT4G25220	RHS15	-1.4	-2.2			0	NA	NA	1	NA	NA	0	1	1	0.5
AT4G25820	XTH14	-1.5	-2.3	-1.2	-0.9	1/0	1/0	0	1	0	1	1/0	1	0	0
AT4G26010	PER44	-1.9	-2.3	-0.7	-0.7	1/0	1/0	0	NA	1/0	1	1/0	1	0	0
AT5G65530	AT5G65530	-1.8	-2.3	-0.5	0.2	1/0	1/0	0	NA	0	1	1/0	0	0	0
AT1G13420	SOT8	-1.4	-2.4			0	1/0	0	NA	1/0	1	0	1	0	0
AT1G30750	AT1G30750	-1.2	-2.4			0	1/0	0	NA	0	1	1/0	1	0	0
AT5G42180	PER64	-1.6	-2.4			0	1/0	0	NA	0	1	1/0	0	0	1
AT4G22460	AT4G22460	-2.3	-2.6			NA	0	NA	NA	0	1	0	1	0.5	1
AT4G36430	PER49		-2.6			0	1	NA	1	1	0	1	1	0	0
AT5G05500	AT5G05500	-1.9	-2.6			0	1/0	0	NA	0	1	0	1	0	0
AT1G71740	AT1G71740	-1.8	-2.7			0	1	0	1	0	1	0	1	0.5	0
AT2G28670	DIR10	-1.2	-2.7	-1.5		NA	1/0	0	NA	0	1	0	1	0	0
AT1G30870	PER7	-1.7	-2.8			0	1/0	0	NA	0	1	1/0	1	0	0
AT4G13580	DIR18	-1.5	-2.9	-1.6	-0.4	1/0	1/0	0	NA	0	1	0	1	0	0
AT4G02270	RHS13	-1.8	-3	-2.3		0	1/0	0	1	0	1	1/0	1	0	0
AT4G40090	AGP3	-2.1	-3			0	1/0	0	1	0	1	0	1	0.5	0
AT5G67400	PER73	-1.7	-3.1	-1.5		0	1/0	0	1	0	1	1/0	1	0	0
AT3G11550	CASP2	-2.1	-3.3			0	1/0	0	NA	0	1	0	1	0	0
AT3G62680	PRP3	-2.3	-3.4			0	1/0	0	1	0	1	1/0	1	0	0
AT2G27370	CASP3	-1.7	-3.5			0	1/0	0	NA	0	1	0	1	1	0
AT1G44970	PER9		-3.8		2.5	0	1/0	0	1	0	NA	1/0	0	0	1
AT2G32300	UCC1	-1.8	-3.8			0	0	0	NA	0	1	1/0	1	0	0
AT2G39430	DIR9	-2	-3.9			0	1	NA	0	0	1	0	1	0	0
AT2G36100	CASP1	-1.4	-4.4	-1.5		0	1/0	0	NA	0	1	1/0	1	0	0

Appendix Figure 4.1. LBb1.3 and LB1 border primers used for SALK and SAIL lines respectively. *(A) Table of genes and their respective mutants and primers. (B) Diagram showing the position of primers.*

A

Gene	T-DNA mutant	Reported genotype	LP	RP	BP
AT5G49350	n/a	n/a	-	-	-
CASPL1B1	SAIL_114_C02	homozygous	GCGGTTTCAAAGCTTACATTG	GTTGGACGAGACGAGAGATTG	GCCTTTTCAGAAATGGATAAATAGCCTTGCTTCC
LTP8	n/a	n/a	-	-	-
AT2G43670	SALK_133229	homozygous	GCACCTCCTTCCGAGATAATC	TGAAAGGAAACCATATTTGTGC	ATTTTGCCGATTTCGGAAC
PDLP4	SAIL_243_E02	heterozygous	CCGACTATTTCGCTCTCACAC	TCCTCGAAGATGAATACGAGC	GCCTTTTCAGAAATGGATAAATAGCCTTGCTTCC
AHA4	SALK_040519	homozygous	AATTTCAGCTCTCCTCTTGGC	ATCAAAGGACAGGGTGAAACC	ATTTTGCCGATTTCGGAAC
GPAT6	SALK_056554	homozygous	TGCAAGGTTATTTCCGTCATC	ATATTCGTGTAGACGCGGATG	ATTTTGCCGATTTCGGAAC
ABCG2	SAIL_697_C08	heterozygous	TCGATTGATTATGGAGGACAAG	TGTTCACCGAGGAATCATTTC	GCCTTTTCAGAAATGGATAAATAGCCTTGCTTCC
CYP708A1	SAIL_361_C06	heterozygous	GAAAAAGTCCGATGAGCACAC	TGAGACCATCCTCCAAAACAG	GCCTTTTCAGAAATGGATAAATAGCCTTGCTTCC
QRT1	SALK_043429	homozygous	TATCTCTTTGCGGTTGAGTGG	TGGGTAACCACAAACGCTAAC	ATTTTGCCGATTTCGGAAC
MYB54	SALK_043499	homozygous	AATCCTAGGAGGCAGAGTTTCC	TACGGTCCTCACAATTGGAAC	ATTTTGCCGATTTCGGAAC
CYS5	SALK_149928	homozygous	GCGTGTCTTCTTCTGTTTTCG	ATTATCCGGCATAATTACCGG	ATTTTGCCGATTTCGGAAC
PME10	n/a	n/a	-	-	-
AT3G27270	SALK_066426	homozygous	TCCAAGTCCAGAGATAATGGG	CAACCATCTCCATCAACCATC	ATTTTGCCGATTTCGGAAC
AT4G17480	SALK_205932	homozygous	ATGACAAGGACCAAGTTGTGC	TCATTCCAGTTTCAATCTCGG	ATTTTGCCGATTTCGGAAC
S-ACP-DES1	SAIL_237_F10	heterozygous	CATTGGCTTTTGATCATACGC	GTCTCATCTCTAACCCCGTCC	GCCTTTTCAGAAATGGATAAATAGCCTTGCTTCC

B