The Role of MicroRNAs in Arabidopsis Lateral Root Development

Tom Denyer, BSc (hons), MRes

Thesis submitted to the

University of Nottingham

for the degree of Doctor of Philosophy

September 2015

ABSTRACT

The root system of a plant serves to anchor a plant to its terrain, to take up water and nutrients and to provide resistance to various biotic and abiotic stresses. Understanding the mechanisms dictating root, and lateral root (LR) development is crucial for work towards global food security. In the last two decades, microRNAs have been found to play crucial roles in many plant developmental processes. Up to now, no global expression profiling has focused specifically on *Arabidopsis thaliana* LR development.

This study utilised mRNA and small RNA sequencing in order to build two large datasets. RNA was taken from developing LR primordia at selected stages and sequenced. The two datasets were validated and found to be of great depth and quality. They were complimented by an Affymetrix dataset (Voβ *et al.,* 2015). A large number of differentially expressed, known microRNAs were identified as were a large number of targets. In many cases, expression of miRNAs and their targets during LR development revealed seemingly reciprocal expression patterns including a number which have previously been linked to LR development.

In some cases, miRNA/target couples were highlighted as having a potential, unreported, role in LR development. This study focused particularly on the potential role of the miR159 family and one if its targets, the transcription factor MYB33. This study reports a dramatic LR defect in a *miR159ab* double mutant and a miR159 target mimic line and discusses a potential regulatory role of this miRNA and this target during LR initiation.

As well as this, ten novel microRNAs were identified and validated. Putative targets were identified for a number of these microRNAs. This study produced overexpression lines for two of these microRNAs though no phenotype was identified. No significant developmental defects were identified in the roots of mutants for a number of the putative targets.

This study presents the first comprehensive analysis of global expression profiling during LR development and has resulted in the production of two high quality datasets.

CONTENTS

	Page number
1. INTRODUCTION	1-31
1.1. PLANT ROOT ARCHITECTURE	2-8
1.1.2. THE ROLE OF AUXIN IN LATERAL ROOT DEVELOPMENT	4-6
1.1.3. AUXIN TRANSPORT	7-8
1.2. MICRORNAS	8-30
1.2.1. MIRNA BIOGENESIS AND FUNCTION	9-14
1.2.2. ANIMAL VS PLANT MICRORNAS	14-15
1.2.3. MIRNAS ASSOCIATED WITH LATERAL ROOT DEVELOPMENT	15-30
1.2.3.1. MIR160	16-18
1.2.3.2. MIR164	19-20
1.2.3.3. MIR393	20-22
1.2.3.4. MIR167	22-24
1.2.3.5. MIR390	24-26
1.2.3.6. MIR169	26-29
1.2.3.7. OTHER MIRNAS WITH LINKS TO LR DEVELOPMENT	29-30
1.3. AIMS	31

2. MATERIALS AND METHODS 32-48

2.1. PLANTS 32-36

2.1.1. PLANT MATERIALS 32

2.1.2. PLANT GROWTH MEDIA 32

2.1.3. PLANT SELECTION 32-33

2.1.4. PLANT GROWTH 34

2.1.5. HORMONE AND NUTRIENT ASSAY TREATMENT 35

2.1.5.1. GA, NAA, ABA & ACC TREATMENT OF MIRNA159 TARGET MIMIC (MIM159) 35

2.1.6. PLANT PHENOTYPING 35-36

2.1.6.1. ROOT CLEARNG 35

2.1.6.2. IMAGE ANALYSIS 36

2.2. BACTERIAL GROWTH 36

2.2.1. BACTERIAL GROWTH MEDIA 37

2.3. MOLECULAR STUDIES 37-43

2.3.1. RNA EXTRACTION & VALIDATION 37

2.3.2. DNASE TREATMENT 37

2.3.3. REVERSE TRANSCRIPTION 38

2.3.4. RT QPCR 38-39

2.3.5. DNA EXTRACTION 39-40

2.3.6. PLASMID ISOLATION 40

2.3.7. POLYMERASE CHAIN REACTION 40-41

2.3.7.1. PCR PROTOCOL 41

2.3.7.2. PCR PURITIFATION 41

2.3.8. DNA RESTRICTION 41-42

2.3.9. GEL ELECTROPHORESIS 42-43

2.3.9.1. AGAROSE GEL ELECTROPHORESIS 43

2.3.9.2. POLYACRYLAMIDE GEL ELECTROPHORESIS 43

2.3.10. MOLECULAR CLONING 43

2.3.10.1. RESTRICTION 43

2.3.10.2. LIGATION 43

2.4. BACTERIAL TRANSFORMATION 43-45

2.4.1. PREPERATION OF CHEICALLY-COMPETENT *E.COLI* CELLS 43-44

2.4.2. *ESCHERICHIA COLI* TRANSFORMATION 44

2.4.3. PREPERATION OF ELECTRICALLY-COMPETENT

AGROBACTERIAUM CELLS 44-45

2.4.4. *AGROBACTERIUM TUMEFACIENS* TRANSFORMATION 45

2.5. PLANT TRANSFORMATION 46-47

2.5.1. TRANSFORMATION INTO *ARABIDOPSIS* 46

2.5.2. TRANSFORMATION INTO *NICOTIANA BENTHAMIANA* LEAVES 46-47

2.6. BIOCHEMICAL ASSAYS 47

2.6.1. GUS STAINING 47

2.7. RNA AND SMALL RNA SEQUENCING SPECIFICATIONS 47

2.8. GENERAL CHEMICALS AND REAGENTS 48

3. SEQUENCING 49-74

3.1. SEQUENCING OVERVIEW 50

3.2. LATERAL ROOT DEVELOPMENT & GRAVITROPISM 50-52

3.3. SAMPLE ACQUISITION 52-57

3.4. SEQUENCING PROCESS 58

3.4.1. CDNA LIBRARY CONSTRUCTION AND SEQUENCING 58-59

3.4.2. PROCESSING OF SEQUENCE DATA 59-60

3.4.3. PROCESSING OF SMALL RNA SEQUENCE DATA 60-62

3.5. FASTQC ANALYSIS OF ALL SAMPLES 62-65

3.6. RNASEQ QC 66-72

3.6.1. RNASEQ QC: AFFYMETRIX 66

3.6.2. RNASEQ QC: REPORTER STUDIES 67-68

3.6.2.1. FOCL1 MUTANTS SHOW A REDUCTION IN LR NUMBERS 68-69

3.6.2.2. FOCL1 MUTANT PHENOTYPE CAN BE RESCUED WITH AUXIN TREAMTENT 69-70

3.6.2.3. FOCL1 PROMOTER:GUS LINES REVEAL A DELAY IN EXPRESSION DURING LR INITIATION 71-72

3.7. SMALL-RNASEQ QC 72-75

3.7.1. SMALL-RNASEQ QC: DISTRIBUTION OF READ-LENGTHS 72-74

3.7.2. SMALL-RNASEQ QC: DEPTH 73-74

3.7.3. SMALL-RNASEQ QC: CLUSTERING 74

3.7.4. SMALL-RNASEQ QC: KNOWN MICRORNAS 74

4. KNOWN MICRORNAS 75-144

4.1. INTRODUCTION 76

4.2. KNOWN MIRNAS INVOLVED IN LR DEVELOPMENT 77-86

4.2.1. MOST EXPECTED MICRORNAS WERE IDENTIFIED BY SMALL-RNASEQ 77-84

4.2.2. SEVERAL MIRNA* STRANDS WERE ALSO IDENTIFIED BY SMALL-RNASEQ 86

4.3. KNOWN MIRNA TARGETS 86-90

4.3.1. MANY KNOWN MIRNA TARGETS WITH DYNAMIC EXPRESSION PROFILES WERE IDENTIFIED BY MRNA-SEQUENCING 87-90

4.4. IDENTIFICATION OF POTENTIAL MICRORNA/TARGET COUPLES REGULATING LRD 90-110

4.4.1. THE MIR165/166 AND MIR157 FAMILIES ARE TWO OF THE MOST HIGHLY-EXPRESSED AND SHOW ANTI-CORRELATION WITH CERTAIN TARGETS 91-99

4.4.1.1. MIR165/166 91-95

4.4.1.2. MIR156/157 FAMILY 95-99

4.4.2. MIR158, MIR395, AND MIR162 TARGETS SHOW ANTI-CORRELATION WITH THEIR MICRORNA REGULATORS EXPRESSION PROFILES 99-106

4.4.3. THE LOWER-EXPRESSED MIR403, MIR399, MIR823, AND MIR840 FAMILIES, ALSO SHOW STRONG ANTI-CORRELATING EXPRESSION PROFILES WITH SELECTED TARGETS 107-110

4.5. USING TARGET MIMICRY TO STUDY IMPACT OF KNOWN MIRNAS ON ROOT DEVELOPMENT 110-139

4.5.1. TARGET MIMICRY OVERVIEW 111-112

4.5.2. MICRORNAS SELECTED FOR TARGET MIMICRY INVESTIGATION 112-114

4.5.3. SELECTED TARGET MIMIC LINES SHOW NO OBVIOUS LATERAL ROOT PHENOTYPE WITH THE EXCEPTION OF MIM159 114

4.5.3.1. MIR159 EXPRESSION SHOWS STRONG ANTI-CORRELATION WITH THAT OF A NUMBER OF TARGETS 115-120

4.5.3.2. EXPRESSION OF MYB33 IS DRAMATICALLY INCREASED IN MIM159 120-122

4.5.3.3. MYB33 EXPRESSION IS INCREASED IN THE ROOTS HUGELY IN A MIR159-RESISTANT GUS LINE 123-124

4.5.3.4. THE MIM159 TARGET MIMIC AND MIR159AB DOUBLE MUTANT BOTH REVEAL A ROOT PHENOTYPE 124-126

4.5.3.5. THE MIM159 ROOT PHENOTPYE CANNOT BE SIGNIFICANTLY RECOVERED BY A VARIETY OF ROOT-RELATED HORMONES 126-128

4.5.3.5.1. AUXIN 128-130

4.5.3.6. NO CHANGE IN LR PRIMORDIA MORPHOLOGY OR DEVELOPMENTAL TIMING CAN BE IDENTIFIED IN MIR159AB OR MIM159 131-134

4.5.4. MIR164 AND MIR160 HAVE BEEN PREVIOUSLY LINKED TO CENTERAL ROLES IN ROOT DEVELOPMENT 135-138

4.5.5. EXPRESSION STUDIES ON MIR393 & ITS TARGETS

REVEALS LITTLE TO SUGGEST A ROLE IN LR DEVELOPMENT 139

4.6. CHAPTER DISCUSSION 140-144

5. NOVEL MICRORNAS **145-201**

5.1. SEVERAL POTENTIAL NOVEL MIRNAS WERE IDENTIFIED BY

SMALL RNA SEQUENCING 145-152

5.1.1. NOVEL MIRNAS ARE LOCATED IN EXONS, INTRONS

AND INTERGENIC REGIONS 149-150

5.1.1.1. NOVEL MIRNA11 IS LOCATED WITHIN AN EXON 149-150

5.1.2. VALIDATION OF NOVEL MICRORNAS BY RT QPCR 150-152

5.2. NOVEL MICRORNAS REVEAL STRONG AND VARIED SMALL

RNASEQ EXPRESSION PROFILES 153-155

5.3. CONSERVATION OF NOVEL MICRORNAS 155-158

5.3.1. CONSERVATION IN ARABIDOPSIS 156-157

5.3.2. CONSERVATION ACROSS SPECIES 158

5.4. INVESTIGATION INTO THE ROLES OF NOVEL MICRORNAS

IN LR DEVELOPMENT 159-163

5.4.1. EXPRESSION PROFILING OF NOVEL MICRORNAS 159

5.4.1.1. LATERAL ROOT INDUCTION USING NPA/NAA TREATMENT 159-161

5.4.1.2. RT QCPR ANALYSIS OF EXPRESSION OF NOVEL

MICRORNAS DURING LR DEVELOPMENT 161-163

5.5. INVESTIGATING THE ROLE OF NOVEL MICRORNAS IN

PLANT DEVELOPMDENT 164-171

5.5.1. GENE CLONING 164

5.5.1.1. SELECTION OF TRANSGENIC PLANTS 164-166

5.5.2. RT qPCR REVEALS OVEREXPRESSION IN NEW MIRNA19 AND

NEW MIRNA39 LINES 166-168

5.5.3. PHENOTYPIC ANALYSIS OF NEW-MIRNA19 AND NEW

MIRNA39 OVEREXPRESSION LINES 168-171

5.6. TARGET PREDICTION FOR TEN NOVEL MICRORNAS 172-

5.6.1. BIOINFORMATIC PIPELINE TO PRIORITISE POTENTIAL TARGETS 176

5.6.1.1. POTENTIAL TARGETS: EXPRESSION IN OTHER

INDEPENDENT DATASETS 178-181

5.6.1.1.1. LATERAL ROOT INITIATION DATASET

(VANNESTE ET AL., 2005) 178

5.6.1.1.2. ASYMMETRIC CELL DIVISION DATASET

(DE SMET ET AL., 2008) 178

5.6.1.1.3. AUXIN RESPONSE DATASET (OKOSHIMA ET AL., 2005) 179

5.6.1.2. UNIGENE NOVEL MIRNA TARGETS 181-182

5.6.1.3. NOVEL MIRNA AND POTENTIAL TARGETS SHOW CONTRASTING EXPRESSION 182-185

5.6.1.4. NEW MICRORNA - TARGET COUPLES: MOST PROMISING CANDIDATES 185-189

5.6.2. PHENOTYPIC ANALYSIS OF SELECTED PROMISING NOVEL MIRNA TARGET CANDIDATES 189-192

5.6.3. EXPRESSION ANALYSIS OF PREDICTED TARGETS IN NEW MIRNA19 OXP LINES 193-195

5.7. CHAPTER DISCUSSION 196-201

6. DISCUSSION **202-214**

6.1. RNA SEQUENCING PIPELINE 203-210

6.2. KNOWN MICRORNA EXPRESSION DURING LR DEVELOPMENT 210-213

6.3. NOVEL MICRORNAS 213-214

7. REFERENCES 215-248

APPENDICES 249-272

A1. PRIMER DETAILS 250-256

A.1.1. PRIMERS FOR GENOTYPING MIRNA TARGET MIMIC LINES 250

A.1.2. PRIMERS FOR PRODUCING NOVEL MICRORNA

OVEREXPRESSION CONSTRUCTS 251

A.1.3. PRIMERS FOR RT QPCR 252

A1.4. PRIMERS FOR MIRNA (STEM-LOOP) RT AND QPCR

(FOLLOWING THE PRIMER DESIGN PRODOCOL DETAILED IN

VARKONYI-GASIC ET AL., 2011) 253-255

A.1.5. TDNA INSERTION MUTANT GENOTYPING PRIMERS 255-256

A2. RNASEQ RAW READ COUNTS 256

A.3. FASTQC 257-259

A.3.1. 15 TIME POINT RNASEQ FASTQC 257

A.3.2. SMALL-RNASEQ FASTQC 258

A.3.3. SMALL-RNASEQ FASTQC REPEATS 259

A.4. MIR159 TARGET QPCR IN MIM159 260

A.5. NOVEL MICRORNA PREDICTED HAIRPINS (UEA SMALL RNA

WORKBENCH) 261-262

A.6. NOVEL MICRORNA IDENTIFICATION & ANALYSIS 263

A.7. KNOWN MICRORNA CONSERVATION IN ARABIDOPSIS 264

A.8. MICRORNA OVEREXPRESSION LINES 265-267

A.9. NOVEL MICRORNA SEQUENCING EXPRESSION 268

A.10. GENOTYPING 269-272

COMMON ABBREVIATIONS

ABA – Abscisic acid

AGO - ARGONAUTE

APX – ASCORBATE PEROXDASE

ARF – AUXIN RESPONSE ELEMENT

ARF – AUXIN RESPONSE FACTOR

AUX – AUXIN TRANSPORTER PROTEIN

bp – Base pair

Col – Columbia

CPIB – Centre for Plant Integrative Biology

DCL1 –DICER LIKE

GA – Giberellic Acid

GFP – Green Fluorescent Protein

HEN – HUA ENHANCER

HYSL – HYPONASTIC LEAVES

IPS – INDUCED BY PHOSPHATE STARVATION

LAX – AUXIN TRANSPORTER PROTEIN LIKE

LBD – LATERAL ORGAN BOUDARIES-DOMAIN

lncRNA – long non-coding RNA

LR – Lateral Root

LRD – Lateral Root Development

MIM - Mimic

miRNA – microRNA

MS –Murashige and Skoog

NAC – NAM/ATAF/CUC Domain

NASC – National Arabidopsis Stock Centre

NFY – Nuclear Factor Y

NGS – Next Generation Sequencing

nt – Nucleotide

P – Phosphate

PCC – Pierson Correlation Coefficient

PHB – PHABULOUSA

PHV – PHAVLOUTA

Pol II – RNA Polymerase 2

PR – Primary Root

QC – Quality Control

REV – REVLOUTA

RISC – RNA INDUCED SILENCING COMPLEX

RNASeq – RNA Sequencing

RT – Reverse Transcription

SE – SERRATE

SHR – SHORT ROOT

SLR – SOLITARY ROOT

SPL – Squamosa-Promoter Binding Like

sRNA – Small RNA

TAIR –The Arabidopsis Information Resource

TCL – TRICHROMELESS

TF – Transcription Factor

TM – Target Mimic

TP – Time Point

TRY – TRIPTYCHON

UTR – Untranslated region

LIST OF TABLES

Table 4.1: *MIRBASE-recognised microRNAs identified in this study by small RNA sequencing of developing lateral root primordia*

Table 4.2: *microRNA family groups identified by small RNA sequencing with differentially-expressed profiles*

Table 4.3: Differentially expressed *microRNA*sequences identified in small RNA sequencing*

Table 4.4: *Targets and potential targets for strong miRNA candidate families*

Table 4.5: *miRNA selected for Target mimic production in this study*

Table 4.6: *Highly-expressed targets and predicted targets of miR159 family members showing anti-correlation in expression profile between target and microRNA*

Table 5.1: *Potential novel miRNAs identified by small-RNA sequencing*

Table 5.2: *Conservation of the 10 predicted novel microRNA locations across sequenced genomes in the SALK 1,001 Genomes database*

Table 5.3: *Targets for novel miRNAs predicted by psRNATarget*

Table 5.4: *Targets for novel miRNA* sequences predicted by psRNATarget*

Table 5.5: *psRNATarget predicted targets of novel miRNA (and equivalent miRNA* sequences) found in one or more selected datasets*

Table 5.6: *PCC between selected, predicted target RNASeq expression profiles and those of the miRNA which are predicted to target them*

Table 5.7: *Most promising potential novel miRNA targets, (based on expression profiles (between RNASeq and Affymetrix and between target & miRNA), prediction scores, presence in relevant, external datasets and nature of target itself) selected for further investigation*

Table 5.8: *Details of tDNA-insertion mutants selected and obtained for predicted targets of novel microRNAs*

Table 6.1: *MicroRNAs identified as being differentially expressed during LR development based on a variety of combinations of various analysis tools.*

LIST OF FIGURES

Figure 1.1: *Cross-section of the Arabidopsis primary root showing three expanded pericycle cells which have undergone divisions*

Figure 1.2: *Simplified diagram of the basic auxin-dependant gene response*

Figure 1.3: *Model for the auxin-responsive regulation of LR initiation by ARF7 and ARF19*

Figure 1.4: *Simplified diagram of microRNA biogenesis and gene silencing through mRNA transcript cleavage*

Figure 1.5: *Overview of some of the most comprehensively studied miRNAs associated with root development in Arabidopsis, and their targets*

Figure 1.6: *An arf10/arf16 mutant and miR160c overexpression line both show agravitropic root growth in young seedlings*

Figure 1.7: *MiR164 regulates LR development through its repression of NAC1. Comparison of lateral root numbers for wild-type (Col), 35S-Myc-NAC1 and miRNA-resistant 35S-Myc-NAC1m*

Figure 1.8: *Expression of pmiR393:GUS in the lateral roots of 10 day-old seedlings (adapted from Parry et al., 2009)*

Figure 1.9: *A model of the primary and secondary circuits of miR167 pathways in response to water stress or nitrogen*

Figure 1.10: *A model for the miR390-tasiRNA-ARF module which has an effect on lateral root development*

Figure 1.11: *Expression of pMIR390A:GUS-GFP in developing lateral root primordia of 10 day-old wild type (Col-0) and solitary root mutant (slr) plants treated with 10μM IAA*

Figure 1.12: *There is a dramatic effect on root system architecture when the miR169defg isoform is sequestered (Target mimic line - MIM169defg)*

Figure 3.1: *Eight numbered stages of lateral root primordia development after a gravitropic stimulus is applied*

Figure 3.2: *Various stages of lateral root development in young wild-type (Col-0) seedlings at selected time points after the application of a gravistimulus*

Figure 3.3: *LP2, ALDH3F1, and PATL2 all show dynamic expression profiles in an Affymetrix dataset of Lateral root development*

Figure 3.4: *RT qPCR QC expression profiles of selected genes over five time points for two replicates of RNA extracted for small RNASeq*

Figure 3.5: *RT qPCR QC expression profiles of selected genes from two replicates of RNA extracted for RNASeq*

Figure 3.6: *Basic outline of the initial RNASeq processing pipeline used in this study*

Figure 3.7: Bioinformatic p*ipeline used for small RNA identification*

Figure 3.8: *FASTQC analysis of individual nucleotide bases across all reads of a selected RNASeq samples across four flow-cell lanes.*

Figure 3.9: *Individual gene Expression profile correlation between 15 time-point RNASeq data and equivalent Affymetrix data*

Figure 3.10: *Range normalised expression profiles of FOCL1 from RNASeq and Affymetrix datasets*

Figure 3.11: *LR number, density and PR length of focl1 knockout lines all show a reduction after 10 days growth*

Figure 3.12: The l*ateral root phenotype of focl1-1 is rescued by treatment of auxin after 9 days. 3 day old seedlings grown on 1/2 MS transferred to 1/2 MS containing 100 nM NAA*

Figure 3.13: *FOCl1 is not expressed in very young stage (stage 1) lateral root primordia in FOCL1-PROMOTER: GUS line nor is it expressed in the primary root tip. However, it does express during the latter stages of LR development.*

Figure 3.14: *Distribution of small RNA Sequencing read lengths from all samples combined*

Figure 3.15: *Small RNA sequence expressions could be grouped into four distinct clusters between replicates*

Figure 4.1: *Expression profiles of the eight most highly expressed, conserved, microRNA families identified by small RNA sequencing based on log2 counts of reads*

Figure 4.2*: Expression profiles of conserved, highly-expressed microRNA families identified by small RNA sequencing based on log2 counts of reads*

Figure 4.3: *Expression profiles of differentially expressed, conserved microRNA* strands, identified in small RNA sequencing*

Figure 4.4*: Range normalised expression patterns for miR165/166 family members and selected targets during lateral root development*

Figure 4.5: *PHB:GUS expression in developing lateral root primordia in Arabidopsis*

Figure 4.6: *Range normalised expression patterns for miR157 family members and selected targets during lateral root development*

Figure 4.7: *A model for miR156/SPL regulation of LR development (adapted from Yu et al., 2015)*

Figure 4.8: *Range normalised expression patterns for miR158 family members and selected targets during lateral root development*

Figure 4.9: *A model of the negative feedback regulation of DCL1 by miR162 in Arabidopsis during lateral root development*

Figure 4.10: *Range normalised expression patterns for miR162 family members and their target, DCL1*

Figure 4.11: *The DCL1 mutant - dcl1-9- shows a dramatic lateral root phenotype in Arabidopsis thaliana when compared to wild-type*

Figure 4.12*: Range normalised expression patterns for miR395 family members and selected targets during lateral root development*

Figure 4.13: *MIR395d promoter-GFP shows high expression in the developing lateral root in Arabidopsis thaliana*

Figure 4.14: *Range normalised expression patterns for miR408, miR399c, miR823, and miR840 and selected targets during lateral root development*

Figure 4.15: *Basic outline of the action of Target mimics*

Figure 4.16: *Aerial phenotypes of young seedlings (3 weeks) of selected miRNA target mimic lines (MIM164, MIM397 and MIM164)*

Figure 4.17: *Phenotyping of Target mimic lines for miR159, miR160, miR164, and miR397: Lateral root numbers, density, and primary root length*

Figure 4.18: *Range-normalised expression profiles of miR159 family members identified in small RNA-sequencing of developing lateral root primordia*

Figure 4.19: *Range normalised RNASeq expression profiles of highly-expressed miR159 targets and predicted targets during lateral root primordia development alongside expression profiles of miR159 family members.*

Figure 4.20: *Range normalised RNASeq expression profiles of MYB33 during lateral root primordia development alongside expression profiles of miR159 family members.*

Figure 4.21: *miRNA RT qPCR of miR159 targets; MYB33, MYB65, and MYB101 on whole, 10 day-old MIM159 and Col-0 whole-seedlings.*

Figure 4.22: *relative RT qPCR expression of MYB33 in 10-day-old MIM159 and Col-0 root systems following NAA treatment*

Figure 4.23: *Expression of MYB33 in the roots of MYB33:GUS and miRNA159-resistant mMYB33:GUS lines*

Figure 4.24: *Expression of MYB33 in the early-leaves of MYB33:GUS and miRNA159-resistant mMYB33:GUS lines*

Figure 4.25: *Lateral root number, primary root length, and lateral root density of MIM159 and miR159ab after 14 and 12 days*

Figure 4.26: *MIR159 and miR159ab both show a dramatic reduced lateral root phenotype after 12 days growth when compared against wild type (Col-0)*

Figure 4.27: *Primary root length and lateral root density of MIM159 after 12 days' growth on 1/2MS media treated with various concentrations of selected hormones (GA, ABA, and ACC)*

Figure 4.28: *Primary root length (cm) and lateral root density (LRs per cm PR) of MIM159 after 12 days' growth on 1/2MS media treated with two concentrations of auxin (NAA)*

Figure 4.29*: ProAUX1:AUX-YFP (adapted from Swarup et al., 2012)*

Figure 4.30: *MIM159 primordia at various stages after initiation by a gravity stimulus compared to wild-type (Col-0)*

Figure 4.31: *Staging of LR primordia for MIM159 and miR159ab, compared against wild-type (Col-0)*

Figure 4.32: *Averaged expression profile of NAC100 from averaged RNA sequencing and Affymetrix expression profiles from developing root primordia*

Figure 4.33: *RT qPCR expression of NAC1, NAC80, and NAC100 in 10-day-old MIM164 (blue) and Col-0 (red) root systems following NAA treatment, relative to actin expression*

Figure 4.34: *Averaged, range normalised, expression profiles of ARF10 and ARF16 from RNA sequencing and Affymetrix*

Figure 4.35: *Averaged, range normalised, expression profiles of AFB2 and AFB3 from RNA sequencing and Affymetrix*

Figure 5.1: *Predicted, novel miRNA precursor hairpins, developed using RNAfold*

Figure 5.2: *RNASeq expression profiles of SDH5 and potential, novel microRNA 11 during lateral root development*

Figure 5.3: *RT qPCR detection of novel microRNA on samples used for initial small RNA sequencing*

Figure 5.4: *Averaged expression profiles of novel microRNA with a PCC between replicates of 0.6 or above*

Figure 5.5: *Expression profiles of new miR11 and new miR29 from RNA sequencing of developing lateral root primordia*

Figure 5.6: *Conservation of the 16 predicted novel microRNA locations across sequenced genomes in the SALK 1,001 Genomes database*

Figure 5.7: *Developing LR primordia in wild type Arabidopsis after treatment on NAA (following Germination on 1/2 MS (5 days) and growth on NPA (3 days)*

Figure 5.8: *Expression profiles of novel miRNA on NPA/NAA treated whole root samples*

Figure 5.9: *Fold change values of novel miRNA expression (6-24 hours) from small RNASeq data*

Figure 5.10: *Plasmid map of pH7WG2,0*

Figure 5.11: *RT qPCR of overexpression lines for new miRNA19 and new miRNA39. Expression of a region before the start of the microRNA precursor, behind the 35S promoter*

Figure 5.12: *miRNA RT qPCR of overexpression lines for new miRNA19 and new miRNA39 using the stem-loop method (Varkonyi-Gasic et al., 2007).*

Figure 5.13: *Phenotyping of new miRNA 39 and 19 Arabidopsis 35S overexpression lines. Lateral root numbers, density, and primary root length were measured after 12 days growth.*

Figure 5.14: *Developing primordia for novel microRNA overexpression lines*

Figure 5.15: *Developing root primordia of novel microRNA overexpression lines, 24 hours after gravistimulus*

Figure 5.16: *Comparison of range normalised expression profiles for novel microRNAs and selected, predicted targets during lateral root primordia development, from RNASeq*

Figure 5.17: *Phenotyping of tDNA insertion-knockout lines for selected, predicted targets for novel miRNA. Lateral root numbers, density, and primary root length were measured after 12 or 11 days growth.*

Figure 5.18: *RT qPCR on T3 new miRNA19 overexpression lines of ATG309640 (Ascorbate Peroxidase 1B)*

Figure 6.1: *Novel microRNAs detected using four different sets of parameters.*

Figure 6.2: *Differentially expressed, known microRNAs identified using an alternative bioinformatics pathway.*

Figure 6.3: *Range normalised RNASeq expression profiles of MYB33 during lateral root development alongside expression profiles of miR159 family members.*

Acknowledgements

I would like to express my sincerest gratitude to my supervisor Dr Ranjan Swarup who has been endlessly generous with his time, patience, and in sharing his expertise. From him I have learned so much during the course of this study, so much of which will aid me greatly in my future career.

I would also like to thank my co-supervisors: Dr Dylan Sweetman and Professor Jerry Roberts from the University of Nottingham for their time and advice. This study was done in collaboration with Dr Martin Crespi at the Institut des Sciences du Végétal, CNRS, France, and I would like to thank him and his lab, notably Dr Federico Ariel and Dr Thomas Blein, for all of their help with so many aspects of this study, and for their patience while I was over there. I would like also to offer my many thanks to Dr Michael Wilson for all of his help on the bioinformatic side of things.

Furthermore I would like to thank all of the following who have helped out over the course of this study: Dr Jonathan Atkinson, Dr Ruth Cornock, Dr Rupert Fray, Dr Julie Gray and her lab (Sheffield University), George Janes, Dr Anthony Millar and his lab (Australian National University), Dr Evan Murphy, Dr Rebecca Smith, Dr Stephanie Smith, Dr Céline Sorin, Dr Kamal Swarup, Dr Ute Voβ and anyone who I may have forgot to mention here.

CHAPTER 1

INTRODUCTION

1. INTRODUCTION

The study and understanding of roots is hugely important for work towards global-food security. Since small RNAs play a hugely significant regulatory role in both plants and animals, it is important that a strong understanding is gained of them and how they work within this particular system. This study focuses on one particular class of small RNAs; microRNAs (miRNAs). Most microRNAs identified in plants were first discovered in Arabidopsis and show remarkable conservation across species. A large number of miRNAs have been predicted to be involved in Arabidopsis root development and function (Breakfield *et al.,* 2012, Khan *et al.*, 2010). However, as yet, only a handful have been directly-linked with lateral root development (Guo *et al*., 2005, Wang *et al*., 2005, Mallory *et al*., 2005, Marin *et al*., 2010, for example). Whilst no miRNAs have yet been reported to be directly involved in root hair formation in Arabidopsis, differentially-expressed microRNAs have been identified in soybean root hairs during nodulation (Yan *et al*., 2015).

1.1. PLANT ROOT ARCHITECTURE

A plant's root system anchors the plant to its terrain as well as taking up water and nutrients from the ground and providing a response to various biotic and abiotic stresses. A root system typically consists of a primary root formed during embryogenesis, with lateral roots, developed post-embryonically, branching off in an alternating fashion from either-side of the primary root.

The volume, length and angles of the primary and lateral roots dictate the root system architecture. In the model plant, *Arabidopsis thaliana*, lateral roots grow in a generally acropetal manner and are covered in root hairs which serve to uptake water and nutrients from the soil. Arabidopsis has a very simple root morphology consisting of concentric layers of epidermal, columella, endodermic, pericyclic and vascular tissues.

Lateral roots initiate in the pericycle of the primary root, opposite the xylem poles (figure 1.1.). These xylem pole pericycle cells divide anti-clinally to form a single-cell layer of primordium before dividing peri-clinally producing the dome-shaped lateral root primordium (Malamy *et al*., 1997, Peret *et al*., 2009). As this primordium develops, it breaks through the surface of the primary root and extends at 90 degrees, cell-layer by cell-layer, forming the lateral root. Auxin has been found to play a key role in lateral root development (Fukaki *et al*., 2007, Peret *et al*., 2009, Swarup *et al*., 2008).

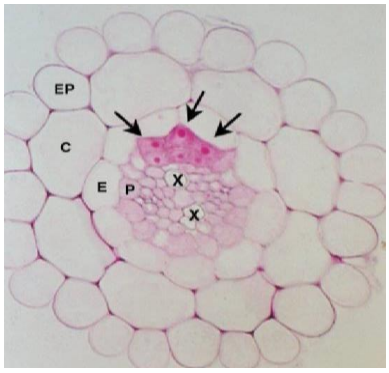


Figure 1.1: *Cross-section of the Arabidopsis primary root showing three expanded pericycle cells which have undergone divisions (arrows). C – cortex, E – endodermis, EP – epidermis, P – pericycle, X – xylem (adapted from Casimiro et al., 2003)*

1.1.2. THE ROLE OF AUXIN IN LATERAL ROOT DEVELOPMENT

The plant phytohormone auxin is a key regulator of lateral root development both directly and indirectly. Auxin is produced in the young roots and shoots of plants and is transported root-wards. Lateral root initiation is prompted by the concentration of auxin in particular regions of the pericycle; priming the cells there for the divisions necessary for the development of the new root. Other hormones including Gibberellic Acid (GA) and cytokinin also play roles in many plant developmental processes, both independently and in conjunction with auxin.

The major form of auxin in higher plants is IAA (indole-3-acetic acid) which plays a role through its interaction with the TRANSPORT INHIBITOR RESPONSE 1 and AUXIN F-BOX PROTEIN (TIR1/AFB) family of auxin-receptor proteins. Auxin stabilises the interaction of TIR1/AFB proteins with AUXIN/IAA (AUX/IAA) repressor proteins, which act as transcriptional repressors (Dharmasiri *et al*., 2005, Vanneste & Friml, 2012). TIR1 is part of an ubiquitin-ligase complex which earmarks AUX-IAAS for targeting by the 26S proteasome (Maraschin Fdos *et al*., 2009). The destruction of Aux/IAAs frees Auxin Response Factors (ARFs) from repression, allowing them to interact with genes containing Auxin Response Elements (AREs), which control cell division (Guilfoyle & Hagen, 2007). A simplified version of this whole process can be seen in figure 1.2. The whole process is described in more detail by Smith *et al*., 2012.

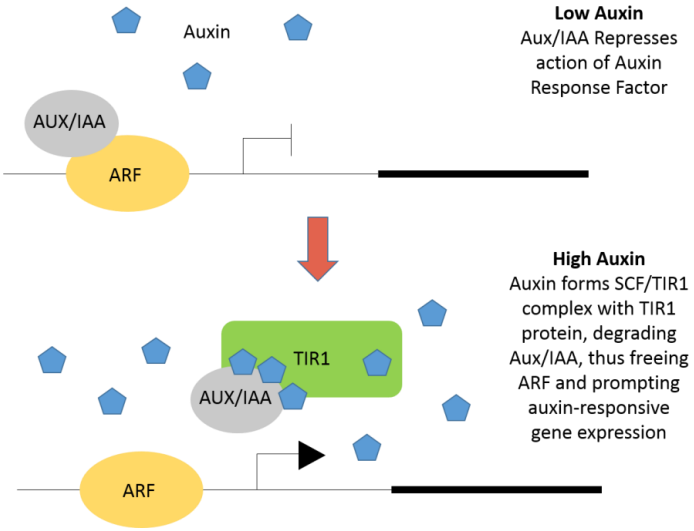


Figure 1.2: *Simplified diagram of the auxin-dependant gene response pathway. Under low auxin (top); Aux/IAA represses the action of Auxin Response Factors (ARFs). Under higher auxin concentrations; a complex incorporating TIR1 is formed which degrades AUX/IAA thus releasing ARFs from repression and allowing auxin-responsive gene expression.*

Various components of the auxin pathway have been studied in great depth. Auxin Response Factors are crucial components of auxin regulation and ARF7 and ARF19 have been identified as directly regulating the auxin-mediated transcription of LATERAL ORGAN BOUNDARIES-DOMAIN16 (LBD16) and/or LBD29 in roots (Okushima *et al*., 2007). ARF7 and ARF19 are themselves regulated by the SLR1/IAA14 module and the solitary root mutant (*slr1*) has a striking phenotype with no lateral roots whatsoever (Fukaki *et al*., 2002, Fukaki *et al*., 2005, Okushima *et al*., 2007).

Auxin prompts the degradation of IAA14/SLR thus de-repressing ARF7 and ARF19 and leading to the activation of LBD16 and LBD29, amongst others; which activate the downstream transcriptional network for LR initiation (figure 1.3.) (Okushima *et al*., 2007).

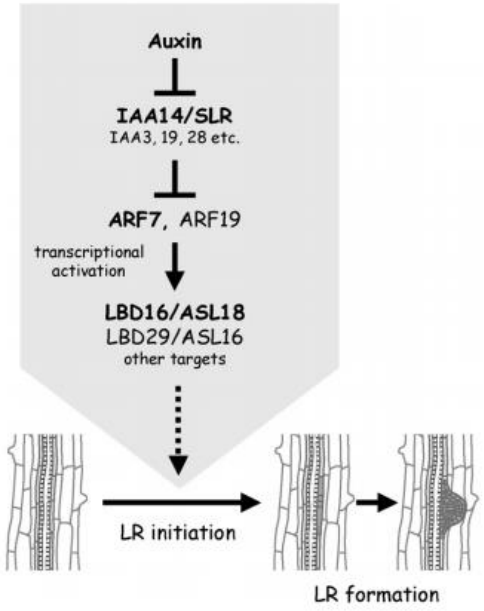


Figure 1.3: *Model for the auxin-responsive regulation of LR initiation by ARF7 and ARF19 (adapted from Okushima et al., 2007). Higher auxin concentrations lead to the de-repression of ARF7 & ARF19 (via the degradation of IAA14/ALR). ARF7 & AF19 activate LBD16 & LBD29 which act as transcriptional regulators to genes involved in lateral root initiation.*

Because of the importance of the auxin signalling pathway to lateral root initiation, many mutants of various modules of the pathway, similarly to *slr1,* show altered root phenotypes. For example: mutants of SHORT HYPOCOTYLE 2-2/IAA3 and IAA28 both also show reduced lateral root numbers while AXR2-1/IA7 mutants show an increase in lateral root numbers (Rogg *et al*., 2001, Tian *et al*., 2003, Nagpal *et al*., 2000).

1.1.3. AUXIN TRANSPORT

The auxin transport pathway is crucial for the proper growth and development of higher plants. Auxin transport is polar in nature and, following synthesis in young, apical tissues, polarity is provided through auxin influx and efflux carriers; most notably AUX1, LAX3 and the PIN Proteins (Bennett *et al.,* 1998, Swarup & Peret, 2012, Křeček *et al*., 2009, Friml *et al*., 2002). This polar transport is necessary for lateral root initiation and blocking the process has been shown to stop lateral root development (Cassimiro *et al.,* 2001).

Prominent amongst the auxin efflux carriers are the PIN proteins, comprising a family of eight (PIN1-8) (Paponov *et al*., 2005). These proteins are involved in auxin distribution throughout the plant and maintaining the auxin gradients necessary for organ formation (Benkova *et al*., 2003). PIN4 has been specifically linked to patterning in the root (Friml *et al*., 2002).

There are four members of the auxin influx carrier family: AUX1 (AUXIN TRANSPORTER PROTEIN 1) and LAX1-3 (AUXIN TRANSPORTER LIKE PROTEIN 1-3) (Parry *et al*., 2001). LAX3 has been found to promote lateral root emergence through the establishing of an auxin 'maximum' at the lateral root tip (pre-emergence) (Swarup *et al*., 2008). LAX3, induced by auxin itself, prompts the accumulation of auxin and the subsequent up-regulation of cell wall remodelling enzymes which initiate the cell separation necessary for emergence (Swarup *et al*., 2008). AUX1 is also required for LR development and the *aux1* mutant shows reduced lateral roots (Marchant *et al.,* 2002). This is a phenotype shared by LAX3, at least with regards emerged roots (Swarup *et al*., 2008). As lateral roots develop they begin to synthesise their

own auxin and become less-dependant on polar transport (Celenza *et al*., 1995, Laskowski *et al*., 1995).

1.2. MICRORNAS

Several types of small non-coding RNAs have been identified including microRNAs (miRNAs), which are typically around 22 nucleotides long. MicroRNAs are post-transcriptional regulators, particularity associated with gene silencing. While non-protein coding transcripts had been suggested to regulate gene activity long before (Britten and Davidson, 1969), the first microRNA to be discovered was found to be a key molecule in the negative regulation of larval developmental timing in the nematode *Caenorhabditis elegans* (Lee *et al.,* 1993). This miRNA was found during the study of the lin-4 gene and was found to repress the lin-14 gene (Lee *et al*., 1993). A second microRNA was later identified in the same organism, targeting the same family of genes (in this case lin-41) (Reinhart *et al*., 2000). These, and other early examples, were not initially described as microRNAs but were instead referred to simply as 'small temporal RNAs' due to their transient expression pattern during certain developmental processes (Pasquinelli and Ruvkun, 2002). By 2001 many more of this same class of small RNA had been identified across a variety of organisms including humans; subsequently the term 'microRNA' was established (Lagos-Quintana *et al*., 2001, Lau *et al*., 2001, Lee and Ambros, 2001).

While some of these initial microRNAs were being identified, another group described how certain double-stranded RNAs silence gene expression through RNA interference (RNAi) (Fire *et al.,* 1998). This work would win its authors the Nobel Prize for physiology or medicine in

2006. Following this, gene silencing through RNAi in plants, triggered by a small RNA with a matching nucleotide sequence to its target, was defined in 1999 (reviewed in Tomari and Zamore, 2005). In a recent review, Huntzinger and Izaurralde (2011) highlight the extensive span of influence that these small RNAs have and summarise how microRNAs have now been associated with numerous, widespread roles in plant development, cell proliferation and differentiation, apoptosis and metabolism.

In plants, miRNAs are known to regulate almost all aspects of plant growth and development (Grennan, 2008, Rhoades *et al*., 2002, Jones-Rhoades and Bartel, 2004). They are also involved in the regulation of plant responses to abiotic and biotic stresses (Dugas *et al.,* 2008, Jones-Rhoades and Bartel, 2004, Piriyapongsa *et al*., 2007, Piriyapongsa and Jordan 2008).

1.2.1. MIRNA BIOGENESIS AND FUNCTION

There are two methods by which microRNAs affect expression of mRNA transcripts following binding to complementary sequences of mRNA: by either degrading it through transcript cleavage, or by attenuating its translation. However, it has been described that these two processes can work in conjunction in-order to silence gene-expression (Djuranovic *et al*., 2012). mRNA cleavage is by-far the more common form of gene silencing directed by plant miRNAs (Saumet and Lecellier, 2006).

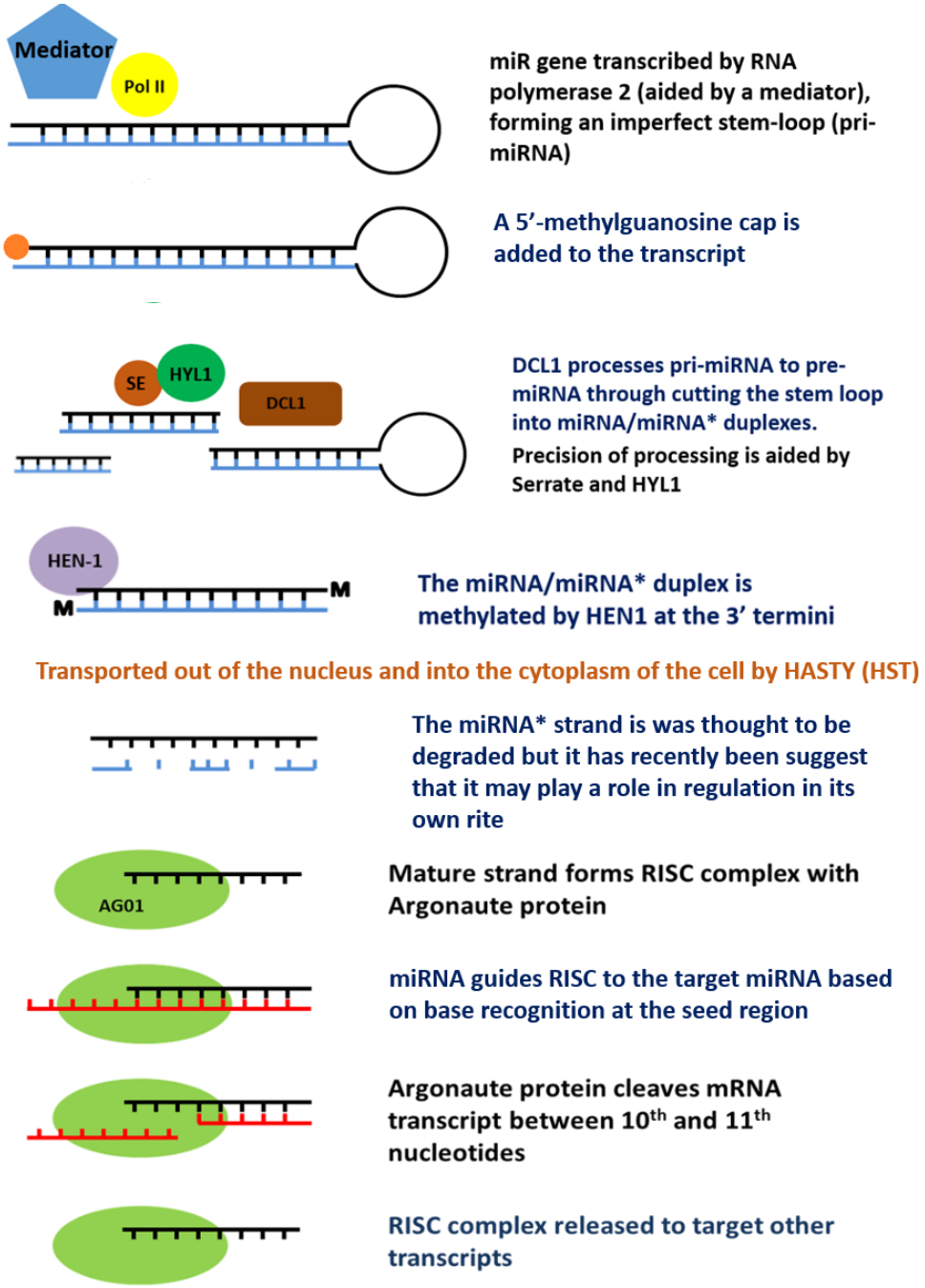


Figure 1.4: *Simplified diagram of microRNA biogenesis and gene silencing through mRNA transcript cleavage. The miR gene is transcribed by RNA polymerase 2 forming a stem-loop to which a 5' methylguanosie cap is attached. The processed miRNA/miRNA* duplex is processed by DCL1 (aided by SE and HYL1) before 3' methylation by HEN1 and transport out of the nucleus by HST. The mature strand of the duplex form RISC with AGO1 which acts to cleave a target mRNA transcript.*

Generally, miRNAs are contained within a non-coding sequence in the genome by RNA Polymerase II, primarily from regions between protein-coding genes (Xie *et al.,* 2005); although some have been found to be transcribed from introns, known as mirtrons (Ruby *et al.*, 2007). The miRNA transcript folds back on itself with intra molecular base pairings forming a stem-loop double stranded RNA (figure 1.4.). This is referred to as the primary miRNA (pri-miRNA) and is modified by the addition of a 5' 7-methylguanosine cap (Xie *et al*., 2005) and a 3' poly-adenylated tail (Jones-Rhoades and Bartel, 2004) In plants, these stem-loop structures vary wildly in length, from around 100 to 1,000 nucleotides (Jones-Rhoades *et al*., 2006).

Pri-miRNAs are cleaved by a RNase III-like enzyme called DICER-LIKE 1 (DCL1) to a stem-loop intermediate called miRNA precursor (pre-miRNA) before being converted into a mature miR/miR* duplex with 5’ phosphates and two-nucleotide 3’ overhangs (Schauer *et al*., 2002, Kurihara and Watanabe, 2004). This is also performed by DCL1. The duplex nomenclature typically refers to the strand which will become the mature, active miRNA (miR), and the complimentary strand (miR*). Due to the size of the pri-miRNA stem loop structure and the fact that it may contain multiple RNA duplexes, the cutting of the stem loop by DCL1 is repeated multiple times.

Various proteins and complexes assist in the processing of pri-RNA and accumulation of miRNAs. These include: HYL1 (HYPONASTIC LEAVES 1), a dsRNA binding protein which forms a complex with DCL1 called ‘microprocessor’ (Kurihara *et al*., 2006, Vazquez *et al*., 2004), and SERRATE (SE), a C2H2-type zinc finger (Lobbes *et al*., 2006). These two proteins assist in promoting accurate miRNA processing (Kurihara *et al*., 2006).

Each 3' end of the miR/miR* duplex is methylated by a methyl-transferase known as HUA ENHANCER 1 (HEN1) to protect it from degradation and further modifications such as 3' uridylation (Li *et al*., 2005) before it is transported out of the nucleus into the cytoplasm of the cell by an Exportin 5 homologue, HASTY (HST) (Bollman *et al*., 2003) (though it has been suggested that transport out of the nucleus can take on a HST-independent mechanism - Park *et al*., 2005). It is generally thought that it is in the cells' cytoplasm that miRNA sorting and RISC (RNA-induced silencing complex) assembly takes place (Hammond *et al*., 2000), transforming the miR/miR* duplex into mature, functional miRNA.

In the cytoplasm, the miRNA strand of the miRNA/miRNA* complex (also known as the guide strand) is incorporated into an Argonaute protein (AGO) in order to mediate gene silencing (Meister and Tuschl, 2004, Tomari and Zamore, 2005) whilst the miRNA* strand, which initially also joins the protein, was generally thought to be degraded. However, there are now strong suggestions that these passenger strands may themselves function as miRNAs (Chang *et al.,* 2013). Indeed, this has led to a change in nomenclature for the two strands by many institutions including the miRNA reference database, MIRBASE, who now refer simply to which arm of the precursor the sequence has originated from (5' or 3').

AGOs are large proteins (90-100 kDa) of the Argonaute protein family. They provide numerous possibilities for RNA-protein interactions and are involved in many types of small-RNA directed regulatory pathways. Cenik and Zamore (2011) consider endonucleic cleavage of target RNA as the simplest, and likely ancestral, Argonaute function.

Argonaute proteins contain four conserved regions. These are: the PAZ (Piwi Argonaute Zwille), and MID (Middle) domains, which have both been associated with RNA binding (Yan *et al*., 2003, Vaucheret, 2008), the poorly conserved N (N-terminal) domain which has been associated with multiple biochemical functions (Djuranovic *et al*., 2011), and the PIWI domain - originally P-element induced wimpy testis in Drosophila (Saito *et al.*, 2006), an endonuclease resembling ribonuclease H which catalyses the cleavage of RNA (Song *et al*., 2004).

Arabidopsis contains 10 different Argonaute proteins (AGO1 – AGO10). Of these, one, AGO1, is known to be necessary for microRNA function in plants. With the aid of ATP, this protein binds to the miRNA/miRNA* duplex, producing pre-RISC before the eviction of the miRNA* strand and the introduction of further proteins (known as the RISC loading machinery) produces mature RISC (Yoda *et al*., 2010). Within this complex, in the cytoplasm of the cell, following target recognition by Watson-Crick base pairing, the protein ultimately catalyses target cleavage once the RISC binds to the target mRNA (Qi *et al*., 2005). Recognition is determined by the seed region of the miRNA and the 3'UTR of its target.

Argonaute proteins comprise two lobes, with one lobe consisting of the PAZ domain and the N domain, and the other containing the other two domains. The PAZ domain binds the 3' end of the guide strand. The other end of the guide strand engages with a 5'-phosphate binding pocket in the MID domain and the strand tracks along a positively charged surface (Carthew and Sontheimer, 2005). The PAZ domain, in conjunction with the other three domains, forms a groove and it is along this groove that the miRNA strand tracks.

The binding of RISC to the mRNA prevents translation either by physically blocking translation (translational inhibition), or - more commonly in plants - by cleaving the mRNA before it is translated. The phenomenon of double-stranded RNA directing post translational silencing of genes is known as RNA interference (Fire *et al*., 1998).

Directing target cleavage is the best understood and most common mode of action used by small RNAs in plants. Cenik and Zamore (2011) highlight how the structure of the AGO protein ensures that the bond cleaved always lies between the target nucleotides paired to the tenth and eleventh nucleotides of the guide RNA. The action of cleaving the near complimentary target is performed by the AGO protein in the RISC, directed by the miRNA. Following cleavage, fragments are released into the cell's cytoplasm, freeing the RISC to continue on to another transcript.

1.2.2. ANIMAL VS PLANT MICRORNAS

There are many differences between plant and animal microRNAs; from the nature and processing of their precursors, to their mode of action on their targets. The specific evolutionary history of microRNAs is still unclear and questions such as whether plant and animal microRNAs arose separately, and to what extent, are still being debated (Jones-Rhodes *et al.,* 2006).

In spite of microRNAs having been discovered much later in plants than they were in animals, it has been suggested that there is a broader understanding of them, their regulatory roles,

and biochemical actions, in plants than there is in animals (Bartel and Bartel, 2003). This could be due to what might be considered the more straightforward nature of plant miRNAs and their target interactions.

A major reason for this comparative simplicity, is the near-perfect nucleotide complementarity between plant miRNAs and their mRNA targets, as well as a tendency to have fewer targets than their animal counterparts (Rubio-Somoza *et al.,* 2011). The nucleotide-pairing conservation proves a great benefit in simplifying the discovery and identification of plant miRNAs and their targets. There is far less complementarity between animal miRNAs and their targets rendering miRNA/target prediction potentially more troublesome. There is also high conservation of plant microRNAs between species and studies (such as those of Axtell and Bartel, 2005) have found that the majority of plant miRNAs are conserved across plant families. This conservation is another great benefit when discovering novel miRNA/target couples.

1.2.3. MIRNAS ASSOCIATED WITH LATERAL ROOT DEVELOPMENT

Whilst a large number of miRNAs have been identified in Arabidopsis, and a large proportion of these have been found to be active in the roots, only a very small number have been directly linked to the regulation of lateral root and primary root development (figure 1.5.). miR160, miR164, miR393, miR167, miR169 and miR390 are some of the most prominent amongst this small set and are briefly discussed in this section (Wang *et al.,* 2005, Guo *et al.*, 2005, Meng *et al*., 2010, Gifford *et al*., 2008, Sorin *et al*., 2014, and Marin *et al.,* 2010).

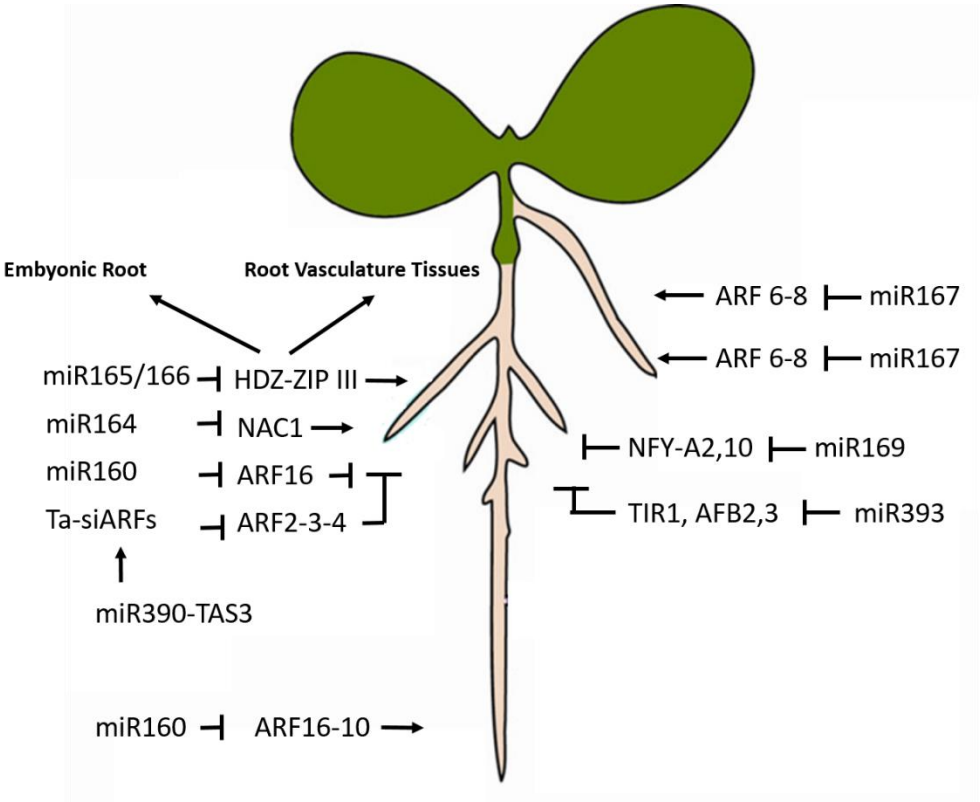


Figure 1.5: *Overview of some of the most comprehensively studied miRNAs associated with root development in Arabidopsis, and their targets (adapted from Khan et al., 2010).*

While many of the phenotypic effects of some of these microRNAs have been studied in some depth, there is often a question as to whether some of the root phenotypes are independent, direct consequences of the regulation of target genes by their microRNAs, or whether it is simply an indirect consequence of the other, perhaps more dramatic and global phenotypic changes.

1.2.3.1. MIR160

MiR160 is encoded by three genes in Arabidopsis: MIR160a, MIR160b, and MIR160c. Overexpression of any of these has been shown to result in agravitropic primary root growth (Wang *et al.,* 2005). MiR160 is found in a range of plants and has been shown to target the same three ARFs in *Medicago truncatula*, where an effect on root growth and nodule numbers has also been described (Bustos-Sanmamed *et al*., 2013).

MiR160 regulates three ARFs (ARF10, ARF16 and ARF17) which have been linked to lateral root development. Wang *et al.* (2005) describe how miR160 affects lateral root density by repressing ARF16 and ARF10. These two functionally redundant ARFs are principal components in primary root growth and the control of root-cap formation (Wang *et al*., 2005). Double *arf10/arf16* mutants show abnormal, agravitropic, primary root growth (figure 1.6.). These mutants developed more, and longer, lateral roots and overexpression of miR160 mirrors the *arf10/arf16* phenotype, suggesting its role in regulating ARF10 and ARF16 (Wang *et al*., 2005).

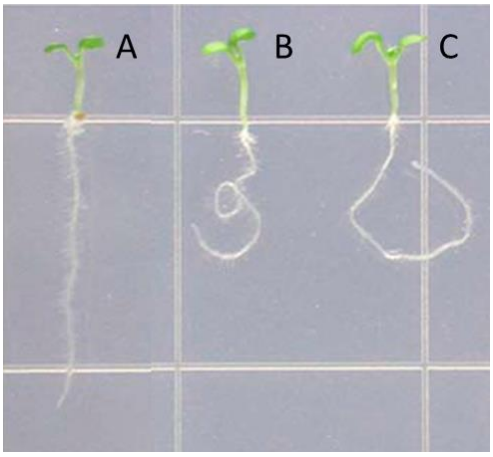


Figure 1.6: *An arf10/arf16 mutant and miR160c overexpression line both show agravitropic root growth in young seedlings (5 day-old) (A: wild type, B: aft10-arf16 double mutant, C: Pro35s:MIR160c overexpression line) (adapted from Wang et al., 2005)*

Recently, Turner *et al.* (2013) showed auxin hypersensitivity in Arabidopsis roots in plants ectopically expressing miR160. The repression of ARF10 by miR160 also results in poor seed germination (Liu *et al*., 2007).

MiR160 has also been liked with the repression of ARF17, although the effects of ARF17 repression are less prominent than those of ARF10 and ARF16. A miRNA-resistant version of AFR17 has been found to affect expression of two GH3-like mRNAs which encode auxin conjugating proteins, effecting auxin homeostasis (Mallory *et al.,* 2005). While this has a strong effect on embryo and emerging leaf symmetry (amongst other strong phenotypic traits), there are also some root-growth defects such as a reduced primary root length and reduced lateral root numbers.

1.2.3.2. MIR164

There are three miR164 family members found in *Arabidopsis thaliana*: miR164a, miR164b and miR164c; all of which have the same mature miRNA sequence. Guo *et al.* (2005) show that miR164 targets five members of the NAM/ATAF/CUC (NAC) domain transcription factor family, one of which (NAC1) regulates auxin signalling necessary for lateral root development (Xie *et al*., 2000). Plants expressing miR164-resistant versions of NAC1 have an increased number of lateral roots (figure 1.7.). However, there are still gaps in the understanding of the miR164/NAC1 module's relationship with other components of auxin signalling.

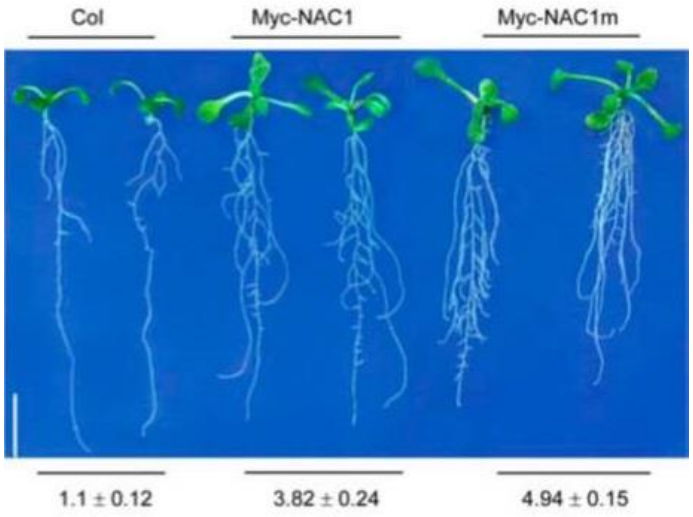


Figure 1.7: *MiR164 regulates LR development through its repression of NAC1. Comparison of lateral root numbers for wild-type (Col), 35S-Myc-NAC1 and miRNA-resistant 35S-Myc-NAC1m, 12 days after germination. Numbers of lateral roots (average ± SE) shown below. Bar = 1cm (adapted from Guo et al., 2005)*

In contrast to NAC1, the role of miR164 in regulating other targets such as CUC1 & CUC2 and NAC80 & NAC100 is less clear (Guo *et al*., 2005). CUC1 & CUC2 are involved in meristem development and the separation of aerial organs, and mutants of these genes show a strong aerial phenotype (Raman *et al.,* 2008). However, a T-DNA insertion mutant for miR164b showed no-aerial phenotype despite there being 15x less miR164 transcript produced in the line revealing a level of functional redundancy and hinting at spatial expression of the various family members having an effect on the targets (Guo *et al*., 2005).

Knockdown mutants for both miR164a & miR164b result in an increase in NAC1 expression resulting in increased numbers of lateral roots. In contrast, overexpression of miR164 can reduce NAC1 expression and, in turn, lateral root emergence (Guo *et al*., 2005). Furthermore, miR164 has been found to be induced by auxin leading to an increase of cleaved NAC1 mRNA transcripts which are not found in auxin-insensitive mutants. A miR164-resistant NAC1 line was also found to be unaffected by auxin treatment (Guo *et al*., 2005). Therefore, there seems to be an auto-regulatory loop at work, whereby the induction of the microRNA by auxin provides a homeostatic mechanism, down-regulating expression of NAC1 mRNA, and thus down-regulating auxin signals (Guo *et al.,* 2005).

1.2.3.3. MIR393

Many (if not most) miRNAs that have been identified in plants have been linked to stress responses. Due to the importance of the plasticity of the root-system to deal with stresses such as drought and nutrient deficiencies, many of these responses involve the adaptation of the plants' root-system architecture, and as such, the control of lateral root initiation and

development. Amongst the few miRNA which have been linked to root adaptation in response to stress is miR393.

MiR393 targets transcripts of TIR1 and its homologues, the F-BOX auxin receptors, AFB2 and AFB3 (Vidal *et al*., 2010). While it has also been identified as a target, AFB1 is partially resistant to miR393 cleavage due to a single nucleotide mismatch (Navarro *et al*., 2006). Repression of these targets leads to a repression of auxin signalling which has numerous consequences for the plant including development, stress responses and antibacterial resistance (Chen *et al*., 2012, Navarro *et al.,* 2006). Both *tir-1* mutants and miR393 overexpression lines have been found to be auxin-resistant (Parry *et al.,* 2009).

MiR393 is induced by Abscisic Acid (ABA) - a stress response hormone associated with drought response – and targets TIR1 and AFB2, leading to a reduction in lateral roots (amongst other things) under drought stress (Chang *et al.,* 2012). As well as this, Navarro *et al.* (2006) connect miR393 to *Pseudomonas syringae* resistance in Arabidopsis through auxin-signalling repression.

Vidal *et al*., (2010) point to how miR393 plays a crucial part in the adaptation of the Arabidopsis root-system architecture in response to nitrate stress through its subsequent targeting of AFB3. Vidal *et al.* point to a feed-forward mechanism induced by nitrate and repressed by nitrogen metabolites generated by nitrate reduction and assimilation.

In their review of 2010, Meng *et al.* explain how miR393 has a direct role in lateral root formation since NAC1 (itself a target for miR164) acts downstream of the miR393 target, TIR1, to transduce auxin signals promoting LR development. Furthermore, Parry *et al.* (2009) identified, using miR393-promoter:GUS lines, that miR393 expression is restricted to the later stages of lateral root development (figure 1.8.).

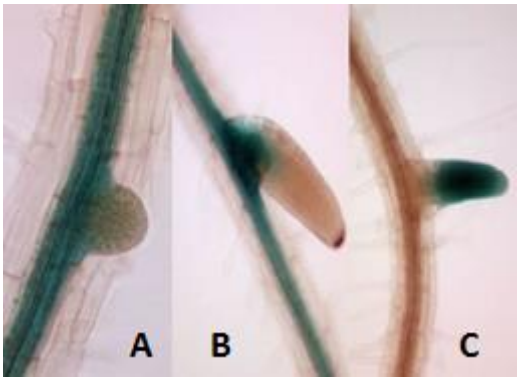


Figure 1.8: *Expression of pmiR393:GUS in the lateral roots of 10 day-old seedlings (adapted from Parry et al., 2009). Little expression is seen in the earlier stages of LR development (A & B) but becomes stronger in the later stages (C)*

1.2.3.4. MIR167

Many microRNAs have been identified as having roles regulating various nutrient responses including responses to phosphate and sulphate starvation (Chiou, 2007, Bari *et al.,* 2006). These responses typically have an effect on the root architecture of the plant. MiR167 has

been identified as having a role in Arabidopsis' nitrogen response including a lateral root phenotypic response (Gifford *et al*., 2008).

MiR167 is encoded by four genes in Arabidopsis (miR167a, b, c & d) and has been reported as targeting both ARF6 and ARF8 (Wu *et al*., 2006). MiR167a and miR167b have been found to be expressed in the pericycle and lateral root cap while ARF8 was repressed in these areas in response to nitrogen, suggesting an antagonistic effect on the target (Gifford *et al*., 2008). In their paper, Gifford *et al.* suggest that nitrogen represses the microRNA, allowing ARF8 transcript to accumulate in the pericycle under nitrogen treatment. This leads to the strong response which includes a change in the ratio of initiating/emerging lateral roots.

Expression-gradients of the microRNA in the Arabidopsis embryo have also hinted at a role in embryo development through the regulation of both ARF6 and ARF8 (Ágyi and Havelda, 2013).

In addition to the two ARF targets, more recently, IAA-Ala-Resistant 3 (IAR3) has also been identified as a target of miR167 and is highly expressed in the pericycle (Kinoshita *et al.,* 2012). IAR3 hydrolyses an inactive form of auxin (IAA-alanie), releasing the active form of auxin, IAA. Kinoshita *et al.* (2012) links this process to osmotic stress, whereby the free auxin released by IAR3 stimulates lateral root development through PIN3, an auxin efflux carrier which forms the auxin gradient (Blilou *et al.,* 2005).

There is no response by ARF6 or ARF8 to osmotic stress but ARF6 and ARF8 both promote Jasmonic acid (JA), which is a strong inducer of IAR3 and therefore these targets may interact through secondary responses (Kinoshita *et al.,* 2012) (figure 1.9.). This kind of complex miRNA/multiple target model is common in plants.

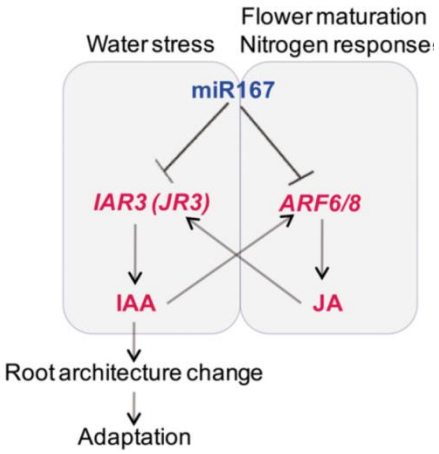


Figure 1.9: *A model of the primary and secondary circuits of miR167 pathways in response to water stress or nitrogen (adapted from Kinoshita et al., 2012). In brief: Jasmonic acid is produced via the ARF6/ARF8 pathway and in-turn prompts the accumulation of IAR3. Accumulation of IAR3 releases auxin which stimulates ARF6 & ARF8*

1.2.3.5. MIR390

MiR390 is encoded by two genes in Arabidopsis leading to the production of miR390a and miR390b isotypes. It has been linked to lateral root development indirectly through its action on another type of small, regulatory RNA: a tasiRNA, which itself targets particular Auxin Response Factors (Marin *et al.,* 2010). TasiRNAs also repress gene expression through RNA

interference; however, tasiRNAs achieve this through a different method to miRNA and there are large differences in their biogenesis and function.

TasiRNAs originate from a non-protein coding transcript called TAS. Four tasiRNAs have been identified so-far in *Arabidopsis thaliana* and one particular transcript (TAS3) requires cleaving in order to produce the active small RNA (ta-siARF). This is cleaved by the miR390/AGO7 complex (Montgomery *et al.,* 2008) and the cleaved product is then polymerised and diced into ta-siARF by RNA Polymerase 6 (RDR6) and Dicer-like 4 (DCL4), respectively (Yoon *et al*., 2009).

Cleavage of TAS3 and the subsequent tasiRNA produced leads to the regulation of leaf-patterning and developmental timing, directed by the three ARF targets of the tasiRNA (ARF3, ARF4, and ARF5) (Yoon *et al*., 2009, Fahlgren *et al*., 2006). These ARFs repress lateral root growth and also affect auxin-induced miR390 accumulation, forming an auto-regulatory loop (Marin *et al.,* 2010) (figure 1.10.).



Figure 1.10: *A model for the miR390-tasiRNA-ARF module which has an effect on lateral root development (adapted from Marin et al., 2010). MiR390 and ARF expressions are strongly regulated; the mutual repression of the microRNA and ARF4 maintains the miR390. Dashed arrows indicate indirect relationships.*

MiR390 is specifically expressed at the site of LR initiation (figure 1.11.) and overexpression lines for miR390 have been found to have longer lateral roots (Marin *et al.,* 2010). Marin *et al.* describe how the positive and negative regulation of miR390 by these ARFs ensures the 'proper definition of the miR390 expression pattern' thus forming a network which can specify the timing of LR growth, in the same manner that these ARFs have a role in leaf patterning and developmental timing (Marin *et al.,* 2010).

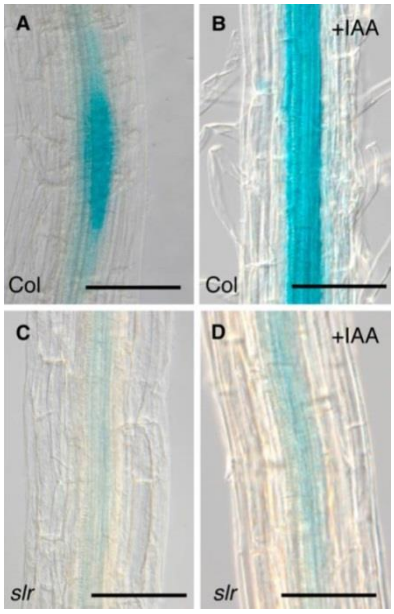


Figure 1.11: *Expression of pMIR390A:GUS-GFP in developing lateral root primordia of 10 day-old wild type (Col-0) and solitary root mutant (slr) plants treated with 10µM IAA (B & D), and untreated (A & C) (adapted from Marin et al., 2010)*

1.2.3.6. MIR169

MiR169 is encoded by 14 genes which together form four different isoforms of the mature microRNA, differing by just one or two nucleotides at the ends of the sequence (Li *et al*., 2010). Amongst the predicted targets of miR169 are some of the ten genes encoding a subunit (A) of the Transcription Factor 'Nuclear Factor Y' (NFY); involved in the regulation of CCAAT box-comprising promoters (Rhodes *et al*., 2002). MiR169 is highly conserved microRNA found in a wide range of plants, often comprising a very large family. It has been linked to a range of functions including lateral root development.

Sorin *et al.* (2014) looked into the role of the specific miR169 isoforms in root architecture development and describe how a particular isoform (comprising miR169d, e, f and g) targets NFY-A2 and NFY-A10 in order to control primary root length and lateral root density. Disturbance of the regulation of NFY-A2 by this particular isoform was found to indirectly affect LR initiation (Sorin *et al*., 2014) (figure 1.12.).

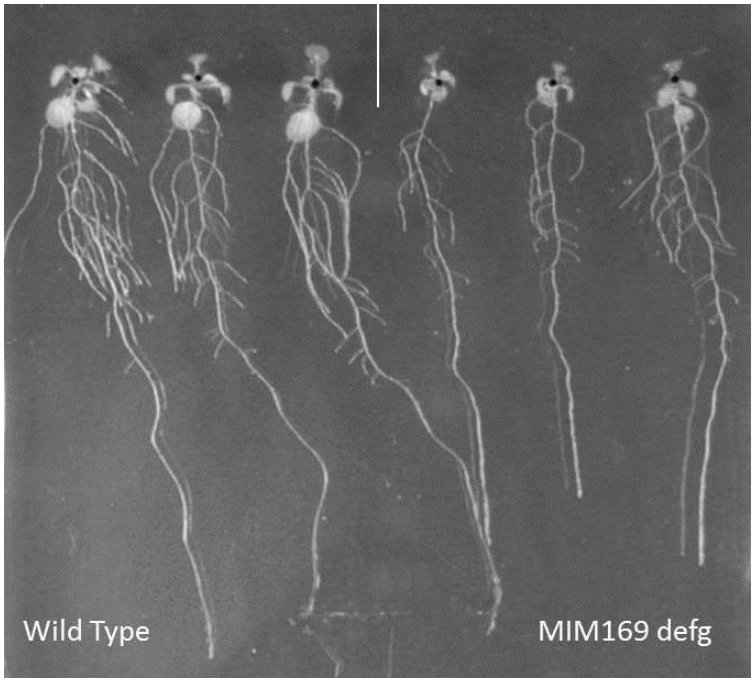


Figure 1.12: *There is a dramatic effect on root system architecture when the miR169defg isoform is sequestered (Target mimic line - MIM169defg) (adapted from Sorin et al., 2014). Seedlings are 8 days after germination; wild type (left) and MIM169defg (right). The mimic line shows shorter primary and lateral roots though LR numbers remain the same, LR density is affected*

As well as controlling root architecture, a link between miR169 and some of these targets during stress-responses has been noted. Accumulation of miR169 transcripts in seedlings correlates with a reduction in target (NFY-A) transcripts under cold-stress (Zhou *et al.,* 2008, Lee *et al*., 2010). On the other hand, reduction in miR169 expression in phloem sap correlates with the induction of NFY-A targets during nitrogen (N) and phosphorus (P) limitation (Pant *et al*., 2009, Buhtz *et al*., 2010).

This miRNA/target couple respond to a variety of stresses and play a fundamental role in root development. Conservation of miR169 between species and strong expression throughout various organs of the plant highlights the importance of miR169 in plant growth and development.

1.2.3.7. OTHER MIRNAS WITH LINKS TO LR DEVELOPMENT

A few other microRNAs have also been linked to lateral root development though in a more indirect way than those described above. In some cases they have been linked to targets involved in root development which, as a consequence, may have an impact on the lateral roots. Alternately, they may have a direct role in stress responses of which an adaptation in root system architecture might be a consequence.

MiR159, for example, targets various MYB transcription factors and when disrupted shows a very strong, negative, phenotypic effect on the plant globally; one element of which is a reduction in lateral root numbers (Alonso-Peral *et al*., 2010). This may simply be a

consequence of an overall-reduction in plant biomass (fewer and/or shorter roots are needed to supply nutrients) or alternately, there may well be a more direct link between microRNA and its target(s).

MiR171 targets the Scarecrow-like genes which encode a group of Transcription factors, some of which have multiple roles including organisation of the radial patterning in roots (Llave *et al*., 2002).

The miR165/166 family are also involved in the organisation of root cell patterning; prompted by SHORT ROOT (SHR), miR165/166 targets PHAVLOUTA (PHV) which controls proper xylem patterning (Zhong & Ye, 2007, Carlsbecker *et al.,* 2011).

Considering the number of microRNAs which have been identified, globally, in *Arabidopsis thaliana*; there are only a relatively small number which have been particularly well-defined. This study will utilise, primarily, next-generation sequencing technologies; and will look to identify any miRNA which may have a role, however indirectly, to lateral root development.

The microRNAs described already here will be considered in this study and further insights into their roles will be looked into. In some cases, microRNAs which have been particularly well defined (along with their targets) will serve as controls when studying less-well-understood microRNA/target couples and when identifying novel microRNA/target couples.

1.3. AIMS

This study aims to build a comprehensive view of the microRNAs involved in lateral root development in *Arabidopsis thaliana*. Using a variety of techniques, but primarily next generation sequencing, known microRNAs and targets will be identified and investigated in order to establish if any of these miRNA/target couples may be involved. As well as this, the study will aim to identify novel microRNAs and targets with particular roles in lateral root initiation and developmental processes.

Once potential candidates are identified and validated, the microRNAs and their targets will be studied further in order to elucidate their roles and networks of which they may belong.

A comprehensive understanding of microRNAs which dictate root-development could prove beneficial in future crop improvement programmes due to the high conservation of microRNAs and targets between Arabidopsis and crops such as rice and wheat.

CHAPTER 2

MATERIALS AND METHODS

2. MATERIALS AND METHODS

2.1. PLANTS

2.1.1. PLANT MATERIALS

Wild-type *Arabidopsis thaliana* Ecotype Colombia-0 was used routinely. T-DNA insertion mutants were obtained from the Nottingham Arabidopsis Stock Centre (NASC).

2.1.2. PLANT GROWTH MEDIA

For routine plant growth: half-strength Murashige and Skoog (MS) media (Sigma-Aldrich) (2.15g/L basal salts; 1.0% (w/v) Bacto™ agar (Becton Dickinson & Co.), pH 5.8 with 1.0M KOH). This was used with 1.0% (w/v) Bacto™ agar. The media was autoclaved and cooled to about 50°C before plating.

2.1.3. PLANT SELECTION

If required, antibiotics or plant hormones were added to media prior to plating. Where appropriate, antibiotics were used in plates and broth, following autoclaving, at the following

concentrations: hygromycin 50µg/ml, kanamycin monosulphate 50µg/ml, and spectinomycin 100µg/ml.

2.1.4. PLANT GROWTH

Arabidopsis thaliana seeds were surface sterilised in 70% EtOH for one minute followed by 25% (v/v) bleach (Fisher Scientific) for 7 minutes. They were then washed four times in sterile distilled water before being sown directly onto the agar with the use of a Gilson pipette.

Seeds were vernalised in the dark at 4^{o}C for 48 hours and then germinated vertically at 20 ±2°C under constant light conditions (122 µmol/m^2/s) until a time by which they would be used. For plants intended for full growth, seedlings were grown on plates (Greiner bio-line) for two weeks before being transferred to soil (unless otherwise stated).

For growth in soil, seedlings were transferred to individual 9cm (diameter) pots containing Levington M3 compost (Everits ltd). This compost was supplemented with MET52 fungicide (fargo). Disposable sleeves (Zwapak, NE) were placed over the plants to prevent cross-pollination during growth. Plants were grown in a glasshouse at a constant temperature between 21-23^{o}C under a 16h light and 8h dark cycle. Light intensity was 150µmols m-2s-1.

2.1.5. HORMONE AND NUTRIENT ASSAY TREATMENT

2.1.5.1. GA, NAA, ABA & ACC TREATMENT FOR ROOT PHENOTYPING

Seedlings were grown on MS plates containing the following concentrations of hormones (or as appropriate): Gibberellic acid (GA) (1µM & 1.5 µM), Auxin (NAA) (100 nM & 175 nM), (Abscisic acid) ABA (10 µM & 15 µM), & Ethylene (ACC) (100nM & 200 nM) prior to root phenotype analysis using a Leica light microscope.

2.1.6. PLANT PHENOTYPING

2.1.6.1. ROOT CLEARING

For slide production, roots were cleared as per the protocol of Malamy and Benry (1997). Seedlings were submerged in Acidified Methanol (4% conc. HCl + 20% MeOH) for 20 minutes at room temperature followed by 20 minutes in neutralisation solution (7% NaOH in 60% EtOH). Seedlings were then rehydrated in EtOH series (40%, 20%, and 10% for 5 minutes each) and water for 5 minutes. Seedlings were then mounted on slides in 50% Glycerol.

2.1.6.2. IMAGE ANALYSIS

Images were captured with a Leica light microscope fitted with a camera unless detailed otherwise in image legends. Primary root length was measured using ImageJ (Collins, 2007).

2.2. BACTERIAL GROWTH

2.2.1. BACTERIAL GROWTH MEDIA

Bacteria were grown in Luria-Bertani (LB) broth (0.5% (w/v) yeast extract, 1.0% (w/v) bacto-trytone, 1.0% (w/v) NaCl, pH 7.0). For plates, 1.0% (w/v) bacto agar was added. Where appropriate, antibiotics were used in plates and broth, following autoclaving, at the following concentrations: hygromycin 50µg/ml, kanamycin monosulphate 50µg/ml, rifampicin 25µg/ml, and spectinomycin 100µg/ml.

Glycerol stocks were made of liquid bacterial cultures by mixing 2 volumes of bacterial culture with 1 volume of 50% glycerol and stored at -80^{o}C.

2.3. MOLECULAR STUDIES

2.3.1. RNA EXTRACTION & VALIDATION

Plant tissues for RNA extraction were collected in 800µl Trizol (Invitrogen) and homogenised using beads in an MP Biomedical bench-top homogenizer (3x 30 second bursts at top speed). RNA was then extracted following the Trizol manufacturer's instructions: 160µl Chloroform was added to samples followed by separation of fractions and precipitation with 400µl Isopropanol and 2µl glycogen (Invitrogen). Samples were washed with 75% EtOH and re-suspended in 50µl sterile deionised water and stored at -80°C. RNA was validated using an Agilent Bioanalyzer RNA nano chip. RNA Integrity Number (RIN) was used as a measure of RNA quality and typically (though not exclusively), any sample with a RIN less than 7 was rejected. Further Validation was done using RT qPCR.

In certain cases RNA concentration and quality was validated using nanodrop rather than using a Bioanalyzer chip.

2.3.2. DNASE TREATMENT

RNAse-free DNAse I (Thermo Scientific) was used routinely for DNAse treatment of RNA preps prior to downstream applications following manufacturer's instructions

2.3.3. REVERSE TRANSCRIPTION

Typically, 500ng of total RNA (up to a volume of 10µl in sterile deionised water) was used for cDNA synthesis. RNA, 10mM dNTPs (1µl), and 100pmol oligo-dT primer (1µl,) were incubated at 65^{o}C for five minutes and then left to anneal on ice (In cases of miRNA RT procedures, selected miRNA RT primers (necessary for stem loop PCR/qPCR) were added alongside oligo-dT primers - 0.2µl, 100pmol – water volume reduced accordingly). Reverse transcription mix (5x First strand buffer 4µl, DTT 2µl 0.1M, Invitrogen) was then added to this reaction and incubated at 42^{o}C for 2 minutes. SuperScript II™ (1µl, Invitrogen) was then added and left to incubate at 42^{o}C for 90 minutes. The reaction was terminated with heating at 72^{o}C for 10 minutes, and the cDNA stored at –20^{o}C prior to use.

In some cases SuperScript III was used instead of SuperScript II. The protocol remains the same apart from the 42 ^{o}C incubations, which were raised to 50 ^{o}C.

2.3.4. RT QPCR

For standard RT qPCR (not stem-loop miRNA RT qPCR), 1µl of prepared, undiluted cDNA was used per reaction. The 10µl reaction also contained 5µl 2x SYBR Green Super Mix, ROX (Quanta Biosciences), 0.2µl each Primer (100pmol/L), 3.6µl Sterile, deionised water.

When possible, biological replicates were used, when unavailable, a minimum of 2x technical replicates were used for each reaction. For miRNA RT qPCR, the U6 primer was used (details in appendix A1) for normalisation and miR390 as a control. In the case of regular PCR, actin was used as a control.

Primer details in Appendix A1.

qPCR protocol (standard):

95°C	5 min	
95°C	15sec	45x
60°C	15 sec	
72°C	4.4 sec	
95°C	5 sec	
95°C	hold	

qPCRs were run on a Roche 'Light Cycler 480 II'

2.3.5. DNA EXTRACTION

Plant material was ground to a fine powder using liquid nitrogen. 400µl extraction buffer (200mM Tris-HCL pH 7.5, 250mM NaCl, 25mM EDTA, 0.5% SDS) was added to the tissue in an eppindorf and mixed by vortex. Samples were centrifuged at 15,000xg for 5 minutes. Equal volume of isopropanol was added to the supernatant and mixed by inversion. DNA was pelleted at 15,000xg for 5 minutes and then air dried (room temperature) for no more than

20 minutes. DNA was re-suspended in 35µl sterile water and stored at –20°C (adapted from the protocol of Edwards *et al.,* 1991).

2.3.6. PLASMID ISOLATION

Single colonies were grown overnight in 5ml of LB containing appropriate antibiotics at 37°C. The culture was harvested by centrifugation at full-speed for 1 minute, and plasmid preparations were performed using the SIGMA mini-prep kit, following manufacturers' instructions.

2.3.7. POLYMERASE CHAIN REACTION

2.3.7.1. PCR PROTOCOL

A typical PCR protocol contained: 2µl 5x Crimson PCR Buffer (New England Biolabs), 0.4µl $MgCl_2$ (50mM) (Bioline), 0.4µl dNTPs (10mM), 0.5µl each primer (10µM), 0.2µl Taq Polymerase, 4.5µl sterile, distilled water, 1.5µl DNA.

Lid - 110°C	
Denaturation - 94°C - 2 min	
94°C – 30 sec	30x Cycles
Annealing (typical) - 57°C – 30secs	
Extension - 72°C - 1 min	
Extension - 72°C 10m	
Hold - 10°C	

PCRs run on an Applied Biosystems 'Gene Amp PCR System 9700'

2.3.7.2. PCR PURIFICATION

PCR purification was done using a PCR clean up (GenElute™ PCR Purification - Sigma) following manufacturers' protocol. In the cases of a PCR product needing to be extracted from an agarose gel, a Gel extraction kit (MiniElute, (Qiagen)) was used.

2.3.8. DNA DIGESTION

DNA was restricted with appropriate restriction enzymes (New England Biolabs) following manufacturers' instructions regarding optimising buffers and conditions based on multiple, simultaneous restrictions on the same product. When multiple restrictions could not be done simultaneously, sequential restrictions were performed.

Reactions were typically undertaken in a volume of 20µl; 2-12µl (1µg) DNA in water (concentration as appropriate), 2µl 10x Buffer, 1µl Restriction enzyme (each), up to volume with sterile, distilled H_20. In most cases, digestions were performed at 37°C for 2 hours.

2.3.9. GEL ELECTROPHORESIS

2.3.9.1 AGAROSE GEL ELECTROPHORESIS

Typically, 1% (w/v) agarose (Fisher Scientific) in TBE buffer (90mM Tris-HCl, pH 8.0; 90mM boric acid and 2mM EDTA), 0.5µg/ml ethidium bromide. DNA was electrophoresed in a TBE buffer at 100V for 30 minutes or as necessary for clear band separation.

2.3.9.2. POLYACRYLAMIDE GEL ELECTROPHORESIS

Small RNAs were separated on 10% (v/v) Acrylamide gel (Biorad) in TBE Buffer (as above) with TEMED (Sigma) and Ammonium persulphate (10% w/v) was used.

2.3.10. MOLECULAR CLONING

2.3.10.1. RESTRICTION

Plasmid and insert were restricted with appropriate restriction enzymes (NEB) (section 2.3.8.).

2.3.10.2. LIGATION

Ligation was typically done in a 10µl reaction; a 3:1 molar ratio of insert to vector was used in each case with 100ng of vector and the equivalent amount of insert. Ligation reactions contain: 1µl of 10x Ligase Buffer and 1µ T4 DNA Ligase (3 U/µl) (NEB) and were undertaken at room temperature overnight. A negative control was performed in each case with no insert.

2.4 BACTERIAL TRANSFORMATION

2.4.1. PREPARATION OF CHEMICALLY-COMPETENT *E.COLI* CELLS

An overnight culture of *E.coli* was produced from a single colony in SOB media (2% Bacto tryptone (2g), 0.5% Bacto yeast extract, 10mM NaCl, 2.5mM KCl - Autoclaved prior to adding

10mM $MgCl_2$ and 10mM $MgSO_4$ - filter sterilised). 100ml of Pre-warmed SOB media was inoculated with 5ml of the overnight culture and grown at 37°C to an Optical Density (OD) 600 between 0.37 and 0.4). Cells were kept on ice for 15 minutes prior to centrifugation at 3,000xg for 15 minutes at 4°C before re-suspension in 50ml RF1 Solution (RbCl (100mM), MnCl2•4H2O (50mM), Potassium acetate (30mM), CaCl2•2H2O (10mM), and Glycerol (15% w/v); final pH 5.8 with acetic acid) and kept on ice for 15 minutes prior to centrifugation at 3,000xg for 15 minutes at 4°C. Pellet was re-suspended in 8ml RF2 Solution (RbCl (10mM), MOPS (10mM), CaCl2•2H2O (75mM), Glycerol (15% w/v); Final pH 6.8 with NaOH) and kept on ice for 15 minutes. Aliquots of 50µl were prepared in pre-chilled, pre-labelled Eppendorfs and stored at -80°C.

2.4.2 *ESCHERICHIA COLI* TRANSFORMATION

1-5ng of plasmid DNA was mixed with 100µl competent *E.coli* cells and incubated on ice for 20 minutes. Cells were heat-shocked for 90 seconds at 42°C and were then incubated on ice again for another 20 minutes. 900µl LB was added to the cells and incubated for one hour in a horizontal position at 37°C. Cells were then plated (100µl) on LB plates containing appropriate antibiotics and grown overnight at 37°C.

2.4.3. PREPARATION OF ELECTRICALLY-COMPETENT AGROBACTERIUM CELLS

Agrobacterium tumefaciens was grown on LB agar plates with selective antibiotics for 1-3 days at 28°C. 100ml of LB media with selective antibiotics was inoculated with a single colony

and grown overnight at 28°C, shaking (200rpm). 50ml of the overnight culture was transferred into 500ml LB media and grow at 28°C until OD600 = 1 - 1.5 (~4hrs). The culture was incubated on ice for 20 minutes. Cells were centrifuged at 4°C for 15 minutes (4000xg). The pellet was re-suspended in 250mls of cold, sterile water (total volume = 500ml) and centrifuged at 4°C for 15 minutes (4000xg). This was then were spun and the pellet re-suspended in water in this manner a further 4 times.

The pellet was re-suspended in 25ml of 10% sterile, cold glycerol (total volume 50ml). This was then centrifuged and he pellet re-suspended in10% glycerol. Samples were stored in 40µl aliquots in pre-chilled (-80°C) 0.5ml eppendorfs in liquid nitrogen, and stored at -80°C.

2.4.4. *AGROBACTERIUM TUMEFACIENS* TRANSFORMATION

An appropriate molar concentration of plasmid was electroporated into 40µl of electro-competent Agrobacterium cells (typically strain C58) (2.5 kV, 200 Ohms, room temperature - Bio-Rad electroporator). 450µl LB was then added to the cells and incubated for three hours at 28°C. Cells were then plates (100µl) on LB agar plates containing appropriate antibiotics and grown over 48 hours at 28°C. Glycerol stocks were made of each culture for storage at -80°C (as detailed in section 2.2.1.).

2.5 PLANT TRANSFORMATION

2.5.1. TRANSFORMATION INTO *ARABIDOPSIS*

Transformation of *Arabidopsis thaliana* was undertaken using the floral dip method (Clough & Bent, 1998). Cultures were grown up overnight (28°C) in 5ml LB broth containing the appropriate antibiotics. This was then transferred to 100ml LB, also containing appropriate antibiotics and grown at 28°C until 0D 600 = 1 (±0.1). At this point, the culture was centrifuged for 3 minutes at 3,000xg and re-suspended in LB containing Sucrose (5% w/v) (Sigma) in order to remove antibiotics from the culture. Pror to dipping, Silwet-L77 (0.5% v/v) (Lehle Seeds) was added to the mixture and shook well. The Aerial parts of flowering *Arabidopsis thaliana* (Col-0) were dipped in the Agrobacterium culture for ~20 seconds.

Plants were covered with folded plastic covers overnight to produce a humid environment before the sleeves were opened slowly over a series of days before being fully open and the plants left to grow entirely.

2.5.2. TRANSFORMATION INTO *NICOTIANA BENTHAMIANA* LEAVES

20ml Agrobacterium cultures in LB containing appropriate antibiotics were grown up overnight (30°C). Samples were centrifuged at 400xg for 20 minutes (20°C) and re-suspended in 5ml MMA buffer (10mM MES, 10mM MgCl2, 150µM Acetosyringone, adjusted to pH 5.8

with 1M KOH) with the addition of addition of further MMA until OD600 = 1 (±0.1). Cultures were then left to grow at RT for 3 hours. Using a blunt syringe, cultures were injected onto the underside of 3-week old tobacco leaves which were left to grow for 72 hours. Sections of infected leaves were collected in liquid nitrogen and stored at -80°C.

2.6. BIOCHEMICAL ASSAYS

2.6.1. GUS STAINING

Seedlings were stained for GUS activity when necessary, before being cleared (section 2.1.6.1.), mounted on slides, and observed under a Leica light microscope.

GUS Staining Protocol	4.25 ml	50 mM phosphate buffer (pH 7.2)
	0.25 ml	33 mg/ml K3Fe(CN)6 in PO4 buffer
	0.25 ml	44 mg/ml K4Fe(CN)6 IN P04 buffer
	0.25 ml	10 mg/ml X-Gluc in Dimethylformamide
	5 µl	Triton X-100

2.7. RNA AND SMALL RNA SEQUENCING SPECIFICATIONS

Detailed in detail in Chapter 3.

2.8. General Chemicals and Reagents

All chemicals used were of analytical grade and were supplied by the following companies: Amersham Biosciences Inc (Little Chalfont, U.K), Bioline Ltd (London, U.K), Biorad Laboratories Ltd (Hemel Hempstead, U.K), Fisher Scientific (Loughborough, U.K), Illumina (San Diego, U.S.A), Invitrogen Ltd (Paisley, U.K), New England Biolabs Ltd (Hitchin, U.K), Promega UK Ltd (Hampshire, U.K), Roche Diagnostics (West Sussex, U.K) Sarstedt (Leicester, UK), Scientific Laboratory Supply (Nottingham, U.K), Sigma-Aldrich Co. Ltd (Dorset, UK), Stratagene Ltd (Cambridge, U.K), and Thermo Scientific (London, U.K),

CHAPTER 3

SEQUENCING

3. SEQUENCING

3.1. SEQUENCING OVERVIEW

To identify microRNAs and their targets which are involved in lateral root development, next generation sequencing (NGS) was used. Sequencing involved two approaches: small RNA Sequencing and mRNA sequencing of RNA extracted from lateral root primordia at various developmental stages. This approach aimed to produce a comprehensive and deep dataset. NGS reaches far beyond the capabilities of the next-best technologies in allowing the identification of novel genes and molecules as well as allowing the study of less-well-known phenomena such as alternative splicing events. This study aimed to produce and mine a dataset which includes vast amounts of data about LR development.

3.2. LATERAL ROOT DEVELOPMENT & GRAVITROPISM

As discussed fully in chapter 1; Arabidopsis lateral roots initiate in the pericycle of the primary root, opposite the xylem poles. The cells divide anticlinally to form a single-cell layer of primordium before dividing periclinally producing the dome-shaped lateral root primordium (Malamy *et al*., 1997, Peret *et al*., 2009). As this develops, it breaks through the surface of the primary root as an emerged primordium.

In Arabidopsis, lateral roots develop in an acropetal manner with newly-formed LRs found close to the root tip and mature and more advanced LRs away from the root tip. Due to the acropetal nature of LR development, there are mixed populations of LRs at different stages of development in any given plant. Two different methods have been developed to obtain relatively homozygous populations of LRs at particular stages: NPA-NAA treatment (discussed more-thoroughly later on in this study), and the use of a gravi-stimulus. Using these methods, global expression profiling during LR development has been studied using microarray based approaches (Himanen *et al*., 2004, Voβ *et al*., 2015). These studies have been very successful in increasing our understanding of lateral root development. In recent years, the advancement of next generation sequencing has added new dimensions to global gene expression profiling and has led to the identification of novel transcripts and an understanding of alternative splicing.

Higher plants show gravitropism in their roots and other organs. When an Arabidopsis root bends in response to gravity, formation of a lateral root is initiated on the outside of the bend and the root grows outwards, in effect, anchoring the plant (Peret *et al*., 2009). When a gravi-stimulus is applied to an Arabidopsis seedling, the lateral root begins to develop in a regulated manner which Peret *et al.* (2009) staged against the time after stimulation (figure 3.1.). Specific, numbered stages were applied to the various steps in development up to emergence at around 48 hours after a stimulus is applied.

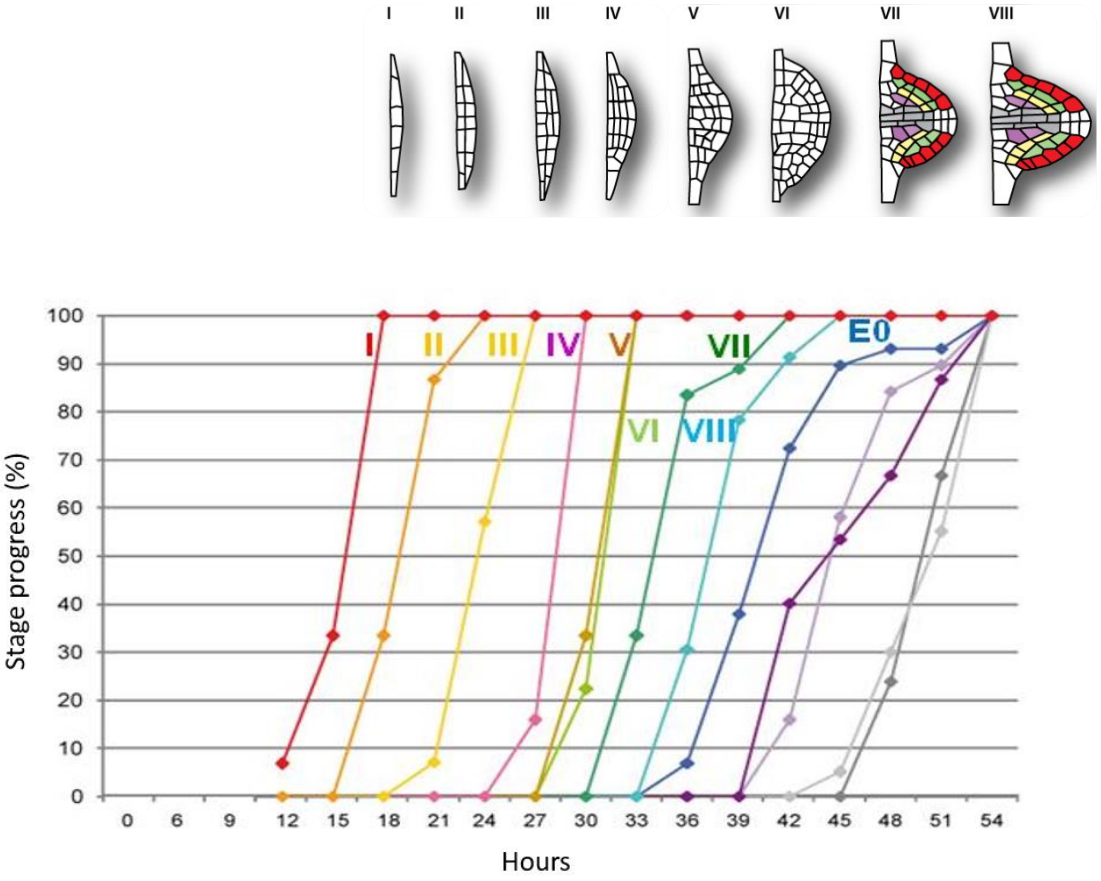


Figure 3.1: Eight numbered stages of lateral root primordia development after a gravitropic stimulus is applied *(x-axis – time after gravistimulus (hours), y-axis – primordia stage progress (%)) (Adapted from Peret et al., 2009)*

3.3. SAMPLE ACQUISITION

Five time points (following gravi-stimulus) were selected for small RNA sequencing, covering pre-initiation, initiation, pre-emergence, and post-emergence stages of the lateral root primordia development. (6, 12, 24, 36, and 48 hours - figures 3.1. & 3.2.). The six hour time-

point was used as a control since, though there is a gravitropic response by this point (and therefore a comparable bend which can be cut); there is no visible lateral root primordium.

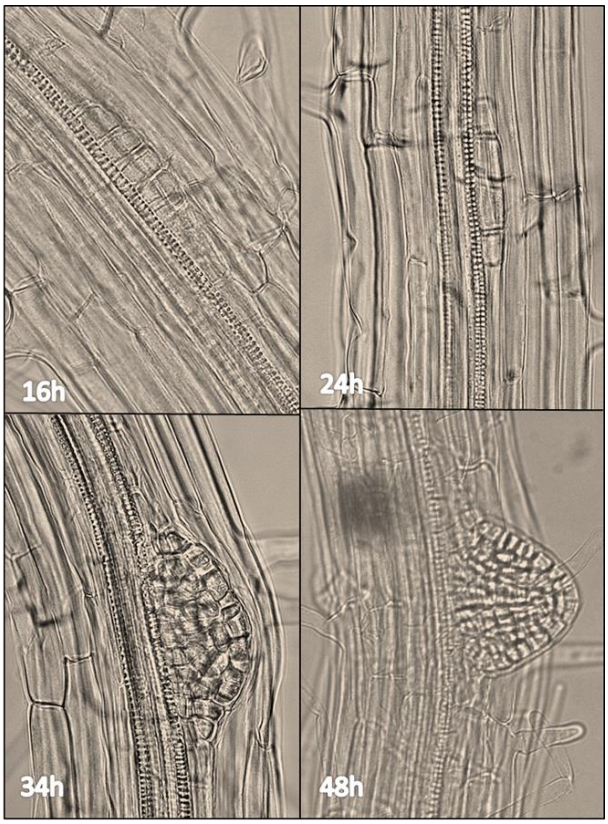


Figure 3.2: *Various, later stages of lateral root development in young wild-type (Col-0) seedlings at selected time points after the application of a gravistimulus (16h, 24h, 34h, & 48h)*

For mRNA sequencing, fifteen time-points were used covering the same time-range (6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36, 39, 42, 45, and 48 hours) to obtain a more in-depth dataset. To obtain samples; three-day old seedlings, grown vertically in constant light, were turned 90

degrees, prompting the gravitropic response that stimulates the development of a lateral root primordia on the outer side of the root bend.

Around 400 root bends containing the developing primordia were cut (around 2mm either side of the bend)/per replicate/per time-point. Excised root bends were collected in Trizol® for RNA extraction (following the Trizol protocol) and stored at -80°C. Two biological replicates were collected for each time-point.

RNA yield and quality was validated using Agilent Bioanalyzer RNA nano chips and RT qPCR was performed using selected genes with particular expression profiles based on comparable Affymetrix data from a lateral root development dataset (CPIB, Nottingham - Voß *et al*., 2015. These genes were PATL2, LP2, ALDHF3F1 & ARF7. While ARF7 is a well-known component of the lateral root initiation mechanism, the other three genes were selected based on their distinctive expression profiles in the Affymetrix data (figure 3.3.). All five time points for the small RNASeq replicates were tested (figure 3.4.) as well as four selected time points (6h, 21h, 33h, and 45h) for the fifteen time-point mRNASeq replicates (figure 3.5.).

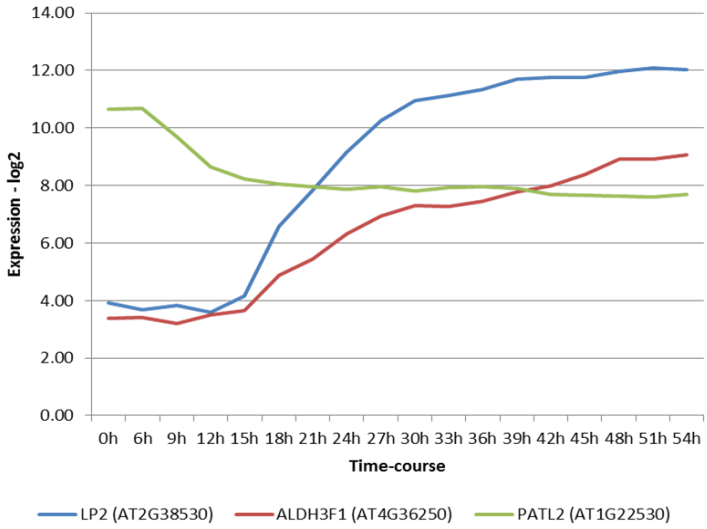


Figure 3.3: *LP2, ALDH3F1, and PATL2 all show dynamic expression profiles in an Affymetrix dataset of Lateral root development (CPIB, Voß et al., 2015)*

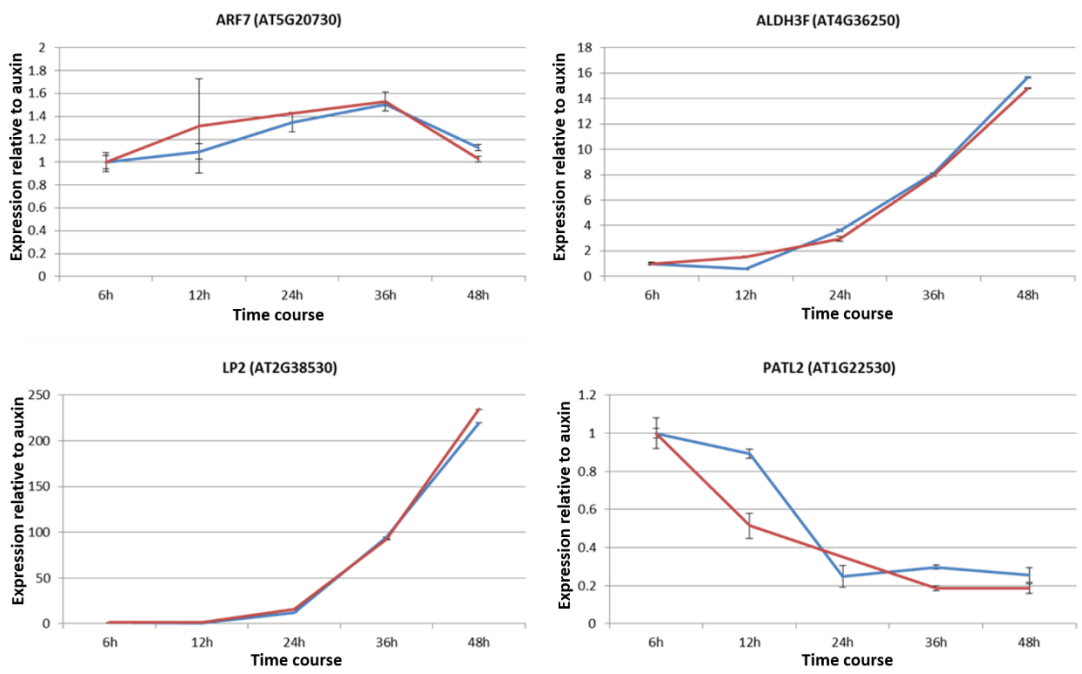


Figure 3.4: *RT qPCR QC expression profiles of selected genes over five time points for two replicates of RNA extracted for small RNASeq. RT qPCR was undertaken on RNA samples, prior to sequencing, to detect four genes; three of which have been found to have distinctive expression profiles (as seen in an Affymetrix dataset), as well as ARF7. Blue bars represent replicate A, red bars represent replicate B. Expression is relative to actin expression at 6 hours. Error bars represent standard error*

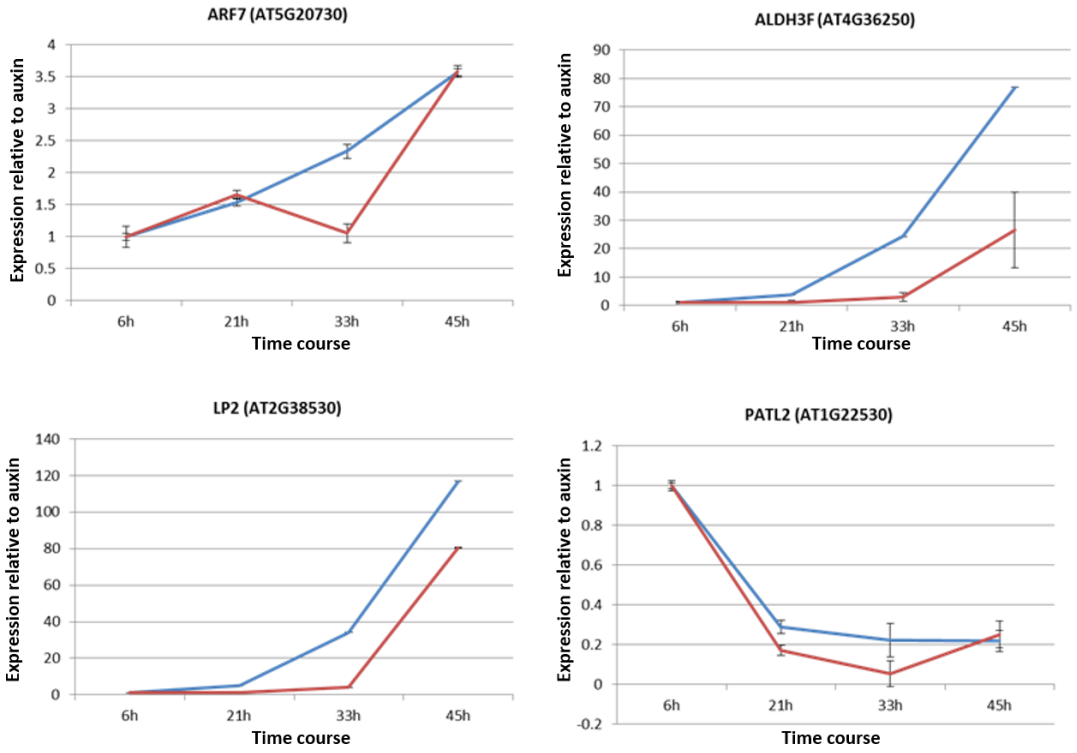


Figure 3.5: RT qPCR QC expression profiles of selected genes from two replicates of RNA extracted for RNASeq (from two selected time points of two replicates - 6h, 21h, 33h, and 45h). *qPCR was undertaken on RNA samples, prior to sequencing, to detect four genes; three of which have been found to have distinctive expression profiles (as seen in an Affymetrix dataset), as well as ARF7. Blue bars represent replicate A, red bars represent replicate B. Error bars represent standard error*

3.4. SEQUENCING PROCESS

Samples were sequenced using Illumina® (San Diego, California) HiSeq technology. The sequencing itself was undertaken by Eurofins mwg operon (Ebersberg, Germany) on an Illumina HiSeq 2000 machine (chemistry v3.0).

For both mRNA and small RNA sequencing, quality controlled total RNA extracted using the Trizol® precipitation method was used for sequencing. Following sequencing, data produced was quality controlled by Eurofins and returned in the FASTQ file format which includes information on sequence quality.

3.4.1. CDNA LIBRARY CONSTRUCTION AND SEQUENCING

For both RNA and small RNA sequencing, cDNA libraries were produced for sequencing from total RNA using the Illumina® RNA TruSeq™ and Small RNA TruSeq™ library production kits, respectively, following manufacturer's instructions. Adaptors were ligated to each end of the RNA with the 3’ adaptor containing a unique index sequence to identify samples. Following adapter ligation, Reverse Transcription was undertaken, followed by PCR amplification with random primers, resulting in the production of the cDNA library. These libraries were sequenced.

RNA replicates were multiplexed across 4 flow-cell channels with each replicate being run twice using identifying index adapters. In the case of mRNASeq, paired-end sequencing was then undertaken, for small RNASeq, single-end sequencing was used. Like the mRNA libraries, the small RNA sample replicates were also run twice each, across two flow-cell channels.

3.4.2. PROCESSING OF SEQUENCE DATA

Eurofins provided sequences in the FASTQ format. These sequences were further processed using a variety of well-defined steps.

In brief, reads were identified by their indexed adapters and separated into the original samples. All adapters were then trimmed from the FASQ files using FASTQ Clipper (from the FASTX toolkit, Hannon lab; http://hannonlab.cshl.edu/fastx_toolkit/index.html). All reads were then mapped against the reference genome (the Arabidopsis information resource reference genome version 10 (TAIR10)) using the well-established 'Bowtie' programme (Johns Hopkins University, Langmead *et al*., 2009) which is used in conjunction with TopHat (also John Hopkins University, Trapnell *et al*., 2009) to identify splice junctions. Output .BAM files (binary versions of .SAM files) were used for visualisation and quantification of the sequence data using various programmes (fgure 5.6). In all cases default settings were used unless specified otherwise.

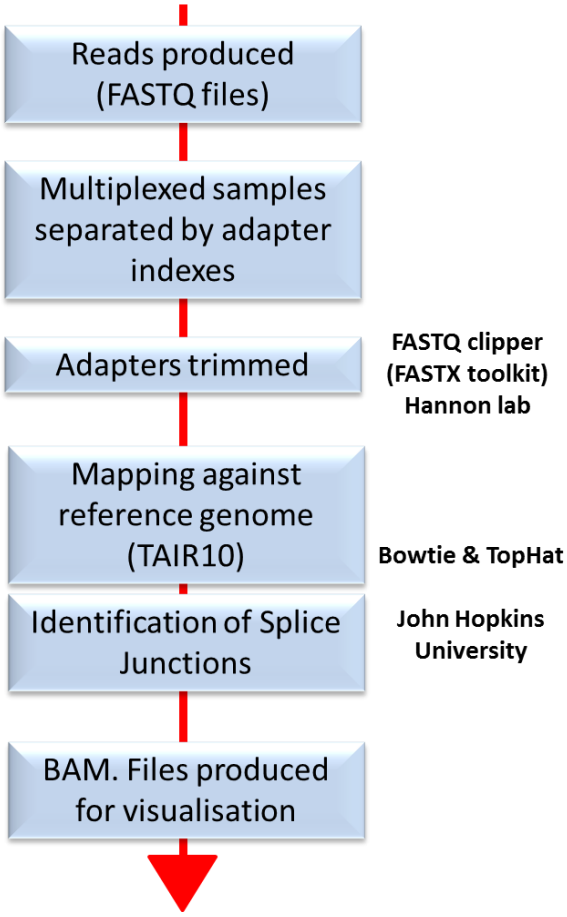


Figure 3.6: *Basic outline of the initial RNASeq processing pipeline used in this study. FASTQ data files were processed through de-multiplexing, adapter trimming, and mapping against TAIR10 prior to quality control and visualisation. Programmes used for certain processes are detailed on the right*

3.4.3. PROCESSING OF SMALL RNA SEQUENCE DATA

Various tools are available for the identification of novel miRNA and for the quantification of known miRNA, the identification of hairpins, and identification of ~23nt miRNA/miRNA* pairs on those hairpins. A number of miRNA prediction algorithms are available (for example: miRCat, miRanalyser, miRTrap, and miRDeep), and these will typically identify ~100nt

windows with many reads where the flanking region can be folded to produce a hairpin within a certain threshold. Further considerations include the identification of consistent 5' ends of reads, and miRNA/miRNA* pairs with a 3' overhang. For small RNA identification and initial QC, a pipeline was produced which would identify small RNAs based on read-lengths and secondary structure motifs (figure 3.7., described below).

In this study, small RNA sequences were processed in a similar manner as the RNASeq set though they were also size-selected. Reads with a length of 21-22 nucleotides were identified in each sample. Redundant reads were then removed and all reads of all samples, of both replicates, were combined and again reads of low incidence were removed (collapsing). Sequences were then aligned against the genome (TAIR 10, again using Bowtie) and extended to include precursor sequences. This was then crossed with MIRBASE in order to identify the conserved microRNAs and novel microRNAs were predicted using miRDeep (Friedländer *et al*., 2008). Quantification of microRNAs was done using CuffDiff (Trapnell *et al*., 2012) and differential expression identified using DESeq (Anders *et al*., 2010). In all cases, unless otherwise noted, default settings were used.

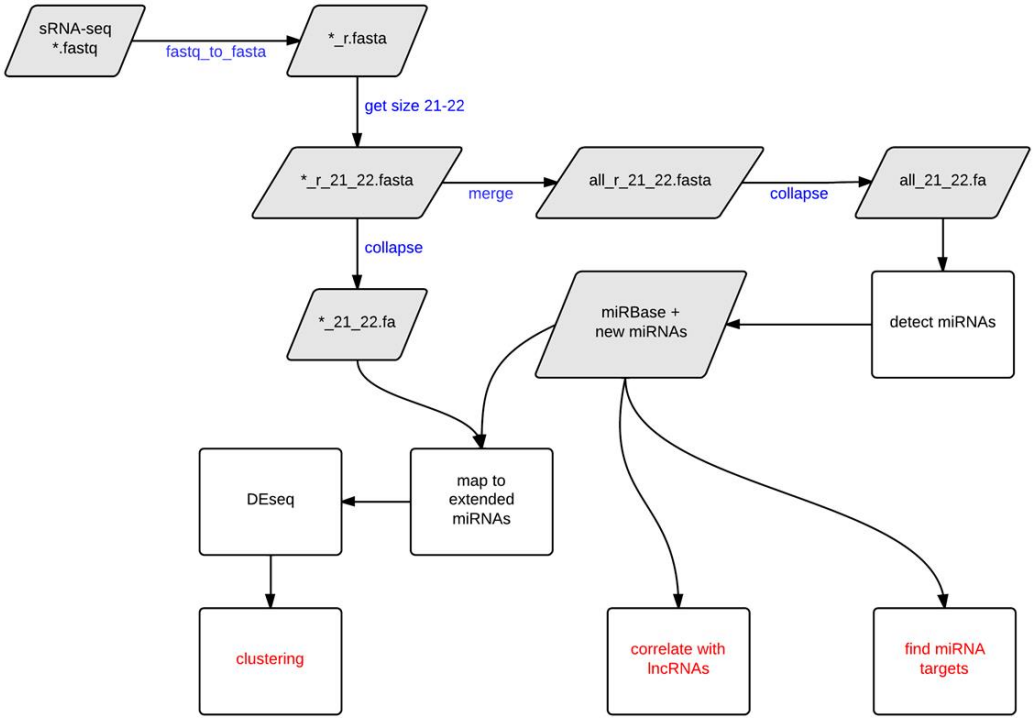


Figure 3.7: Bioinformatic p*ipeline used for small RNA identification. Size-selection (21-22nt); redundant reads removed (collapsing); all reads combined; reads of low incidence removed; alignment (TAIR10); extension (including precursors); crossed with MIRBASE - I.D. conserved microRNAs; novel microRNAs predicted (miRDeep); quantification (CuffDiff); differential expression (DESeq)*

3.5. FASTQC ANALYSIS OF ALL SAMPLES

The quality of sequenced reads and assessment of any bias in the data (such as unexpected sequence content) was done using FASTQC (Babraham Bioinformatics, Cambridge).

FASTQC was undertaken on all separate samples in all flow-cell lanes. This programme assigns a quality score to individual nucleotides of each read. Each nucleotide position is given a Phred quality score (Ewing *et al.*, 1998) between 0 and 40. A sample of some of this analysis for the RNASeq data can be seen in figure 3.8. (a more comprehensive figure can be found in appendix A.3.1.). Results for all small RNA samples in their primary two flowcell channels can be seen in figure 3.9.

The quality of reads, on the whole, drops towards the end of the nucleotide sequences, though this is very common. QC of the sequence-reads revealed that in general, sequences were of high-quality. However, in the case of three small RNASeq samples (24A, 48A, and 48B) the quality of sequences was poor and hence the sequencing was repeated. The new reads proved to be of a sufficient quality and depth. These few repeated samples can be found in appendix A.3.2.

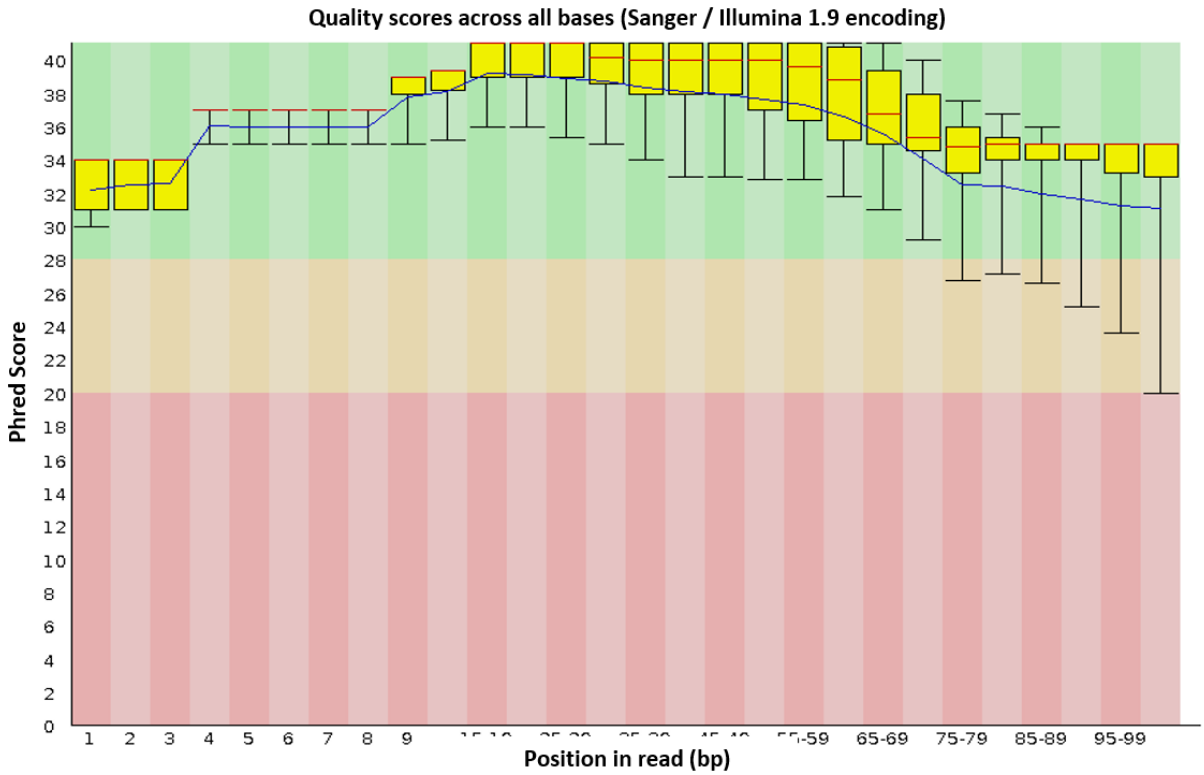


Figure 3.8: Example of *FASTQC analysis of individual nucleotide bases across all reads of a selected RNASeq sample (9h) across four flow-cell lanes. Y axis represents Phred quality score; X axis represents bp position in read. Yellow boxes represent 95% of reads with black lines indicating the range. Red lines represent the read mean. Equivalent FASTQC data for all other samples can be found in appendix A.3.1.*

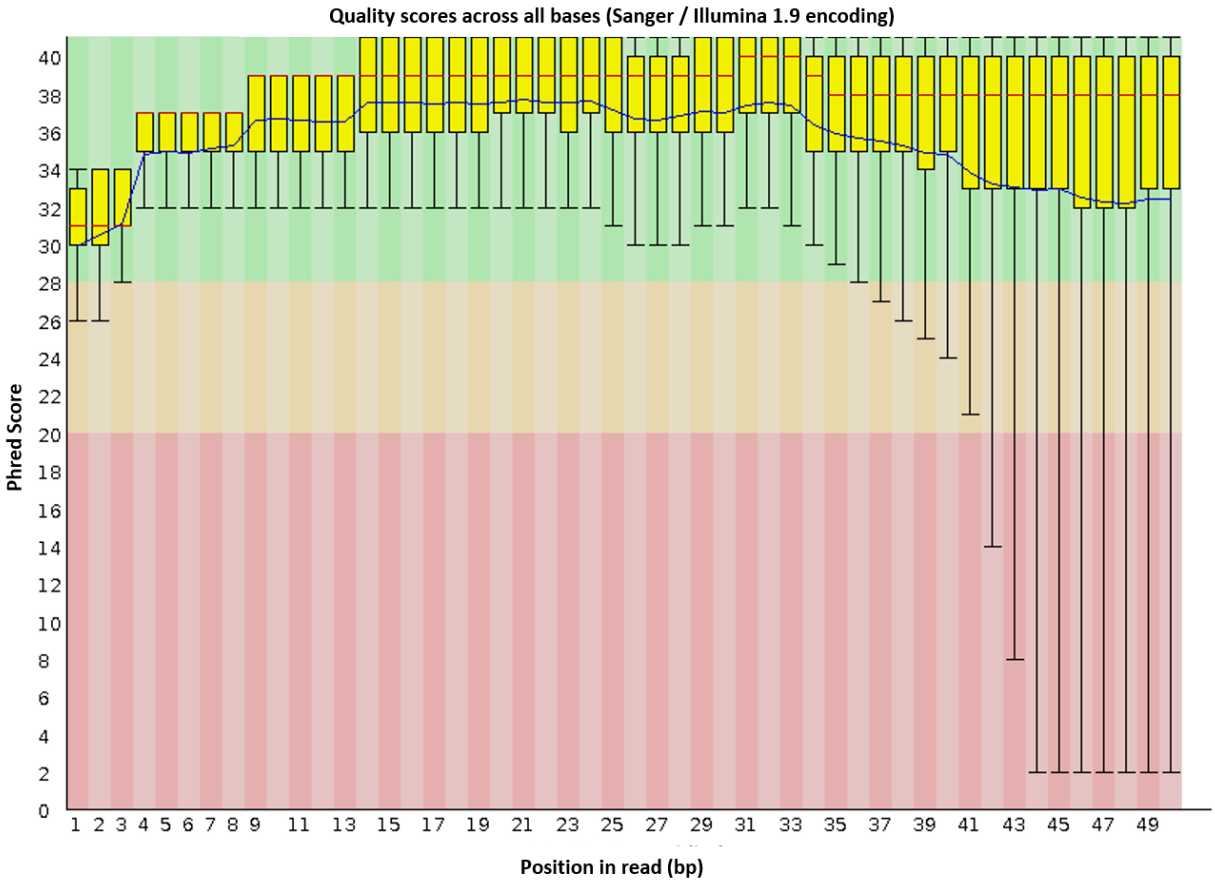


Figure 3.9: Example of *FASTQC analysis of individual nucleotide bases across all reads of a selected small RNASeq sample (36A) across two flow-cell lanes. Y axis represents Phred quality score; X axis represents bp position in read. Yellow boxes represent 95% of reads with black lines indicating the range. Red lines represent the read mean. Equivalent FASTQC data for all other samples including a third flow-cell channel can be found in appendix A.3.2.*

3.6. RNASEQ QC

3.6.1. RNASEQ QC: AFFYMETRIX

In order to check if RNASeq data is comparable to previously generated Affymetrix data, the expression profiles of all differentially-expressed genes identified by RNASeq were compared with the lateral root Affymetrix dataset (CPIB, University of Nottingham, Voβ *et al.*, 2015) using the Pierson correlation coefficient (PCC) (Pearson, 1895)[1]. Results show a very high correlation (figure 3.9.).

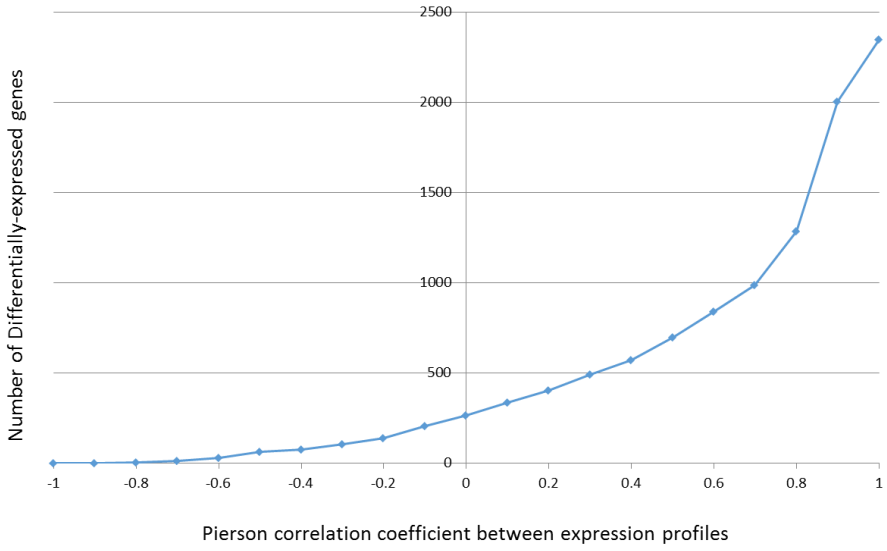


Figure 3.9*: Individual gene expression profile correlation between genes found in both the 15 time-point RNASeq data and the equivalent Affymetrix data*

[1] Using the Pearson Correlation Coefficient (PCC) - a score of 1 represents perfect correlation, 0 represents no-correlation, and -1 represents a perfect anti-correlation (Pearson, 1895).

3.6.2. RNASEQ QC: REPORTER STUDIES

In order to further validate the quality of RNASeq data, GUS lines were obtained for one of the genes that showed a very interesting expression pattern in both RNASeq and Affymetrix datasets (with a PCC of 0.99) (figure 3.10.). This gene encodes a protein which plays a role in stomatal shape and function (Gray Lab, University of Sheffield, in review) and has putatively been named 'Fused Outer Cuticular Ledge 1' (FOCL1)

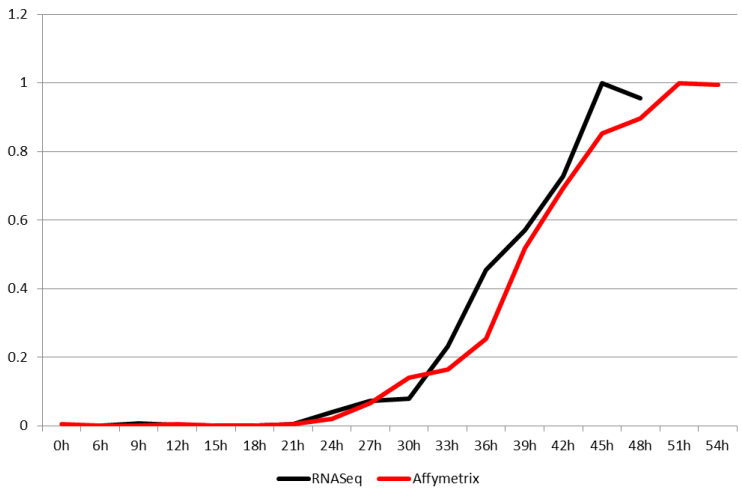


Figure 3.10: *Range normalised expression profiles of FOCL1 from RNASeq and Affymetrix datasets*[2]

These expression profiles strongly suggest that while this particular gene is not involved in lateral root initiation, it may-well be involved in the later stages of lateral root development.

[2] Range normalisation assigns values between 1 (the top value) and 0 (the lowest value) of a series. This approach has been used throughout this study as a way to best visualise relationships between different expression profiles, most significantly, the complimentary relationships between miRNA and their targets.

In order to test this, *focl1* knockout mutants and promoter:GUS lines were obtained (Gray lab).

3.6.2.1. FOCL1 MUTANTS SHOW A REDUCTION IN LR NUMBERS

Two-separate *focl1* mutant lines (*focl1-1* and *foc1-2*) were grown on half MS and emerged LR number, PR length and LR density were scored after 6, 7, 8 , 9, and 10 days. In the case of both of the mutants, all three levels were significantly reduced when compared to wild-type (Col-0). After ten days, lateral root numbers dropped from over 20 to less than 15, Primary root length dropping by around a centimetre and subsequently, lateral root density dropped by around 0.5 LRs per centimetre of primary root (figure 3.11.)

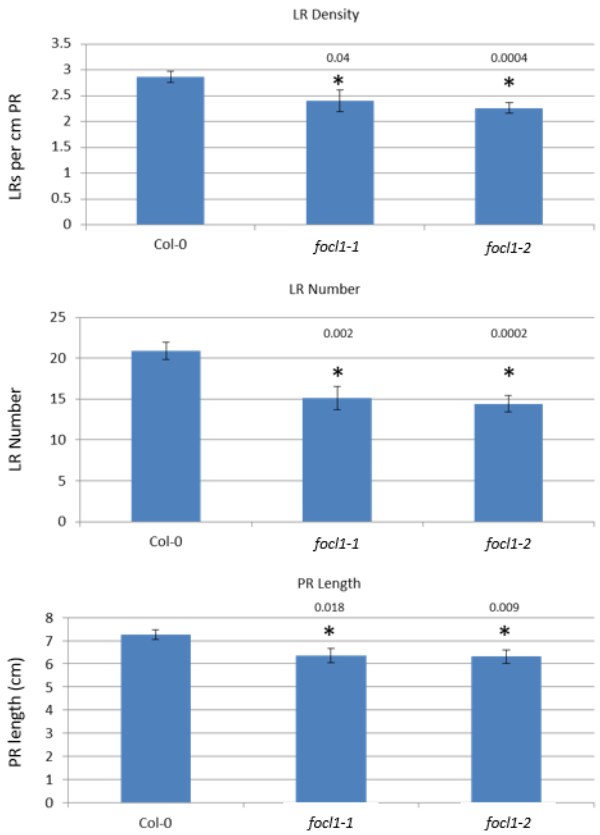


Figure 3.11: *LR number, density, and PR length of focl1 knockout lines (focl1-1 and focl1-2) all show a reduction after 10 days growth, compared against wild-type (Col-0). Asterisk indicates significant difference between mutant and wild-type (student's t-test p<0.05). Error bars indicate standard error*

3.6.2.2. FOCL1 MUTANT PHENOTYPE CAN BE RESCUED WITH AUXIN TREAMTENT

The lateral root defect found in *focl1-1* was subsequently found to be rescued under treatment with auxin (100nM NAA) after 9 days. This indicates that any role that FOCL1 may play in lateral root development occurs downstream of auxin's involvement in the process,

and again, indicates a role further on in the process than in the initiation of LR primordium development (figure 3.12.).

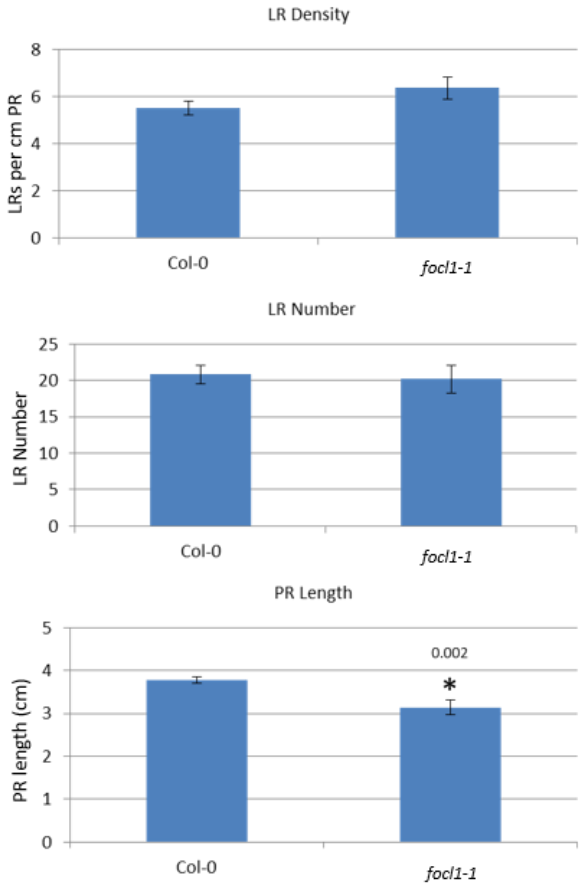


Figure 3.12: *The lateral root defect of focl1-1 is rescued by treatment of auxin after 9 days. 3 day old seedlings grown on half MS transferred to half MS containing 100nM NAA. Asterisk indicates significant difference between mutant and wild-type (Col-0) (student's t-test p<0.05). Error bars indicate standard error*

3.6.2.3. FOCL1 PROMOTER:GUS LINES REVEAL A DELAY IN EXPRESSION DURING LR INITIATION

Two independent FOCL1-Promoter:GUS lines both show FOCL1 expression during lateral root primordium development but not during the early initiation stages (figure 3.13.). This correlates very strongly with the expression profiles for the gene in both the RNASeq and Affymetrix datasets and suggests further that FOLC1 may play a role in the later stages of LR development.

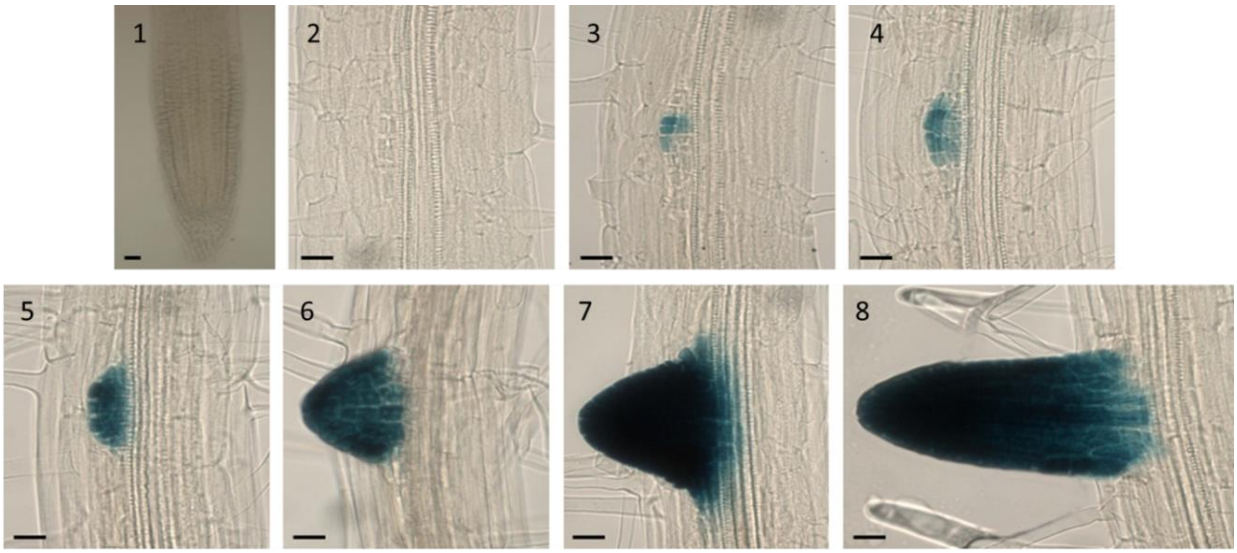


Figure 3.13: *FOCl1 is not expressed in very young stage (stage 1 – picture 2) lateral root primordia in FOCL1-PROMOTER:GUS line nor is it expressed in the primary root tip. However, it does express during the latter stages of LR development.*

FOCL1 seems to play a role in the later stages of LR development. This gene further validates the quality of the RNASeq and Affymetrix datasets. This gene was particularly-highly expressed and showed a very dynamic expression profile. This is supported by the promoter:GUS profile and the dramatic lateral root phenotype strongly suggests that this gene may be a significant player in lateral root development.

No microRNA has yet been found or predicted to target FOCL1.

3.7. SMALL-RNASEQ QC

3.7.1. SMALL-RNASEQ QC: DISTRIBUTION OF READ-LENGTHS

A strong indicator of small RNA sequence quality is the distribution of read lengths. An excess of reads of lengths at which certain small RNA are known to exist would be expected, such as in the case of microRNAs for which you should see a large number of 21 nucleotide reads, and the also-active 22 and 24 nucleotide small RNAs (Melnyk *et al*., 2011). All three of these types of small RNA are represented strongly in our data (figure 3.14.).

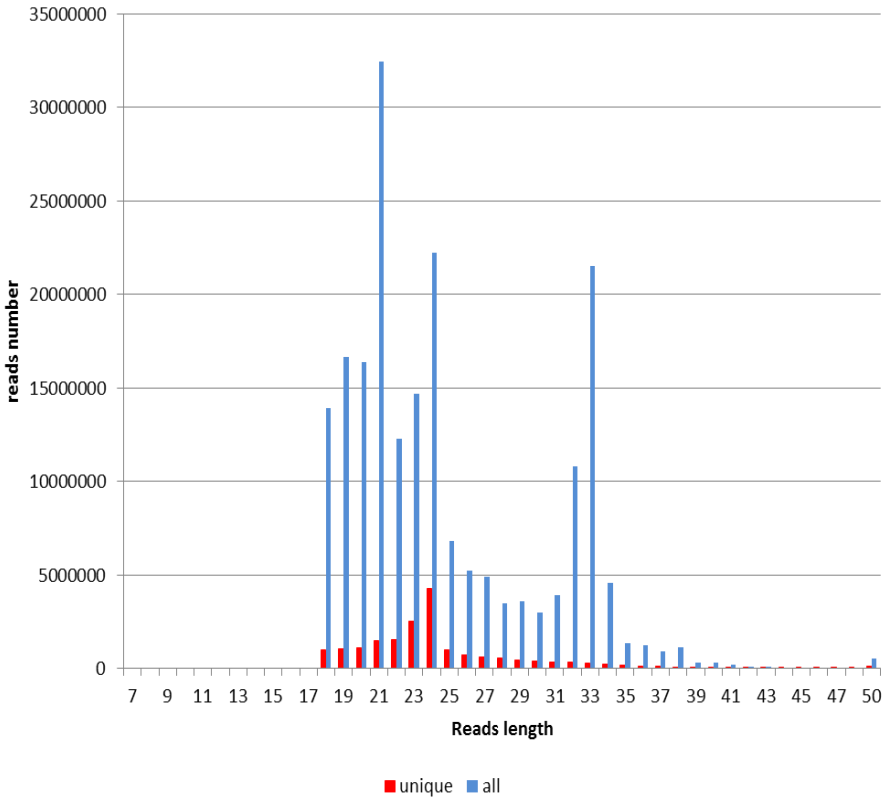


Figure 3.14: *Distribution of small RNA Sequencing read lengths from all samples combined. Red bars represent unique reads, blue bars represent all reads. Prominent small RNA family sizes could be detected strongly.*

3.7.2. SMALL-RNASEQ QC: DEPTH

MicroRNAs which were found in only one replicate were disregarded as being unreliable. In total, 86 different differentially-expressed microRNAs were identified in both replicates. Confidence in a sequence was provided by only accepting those with at least 1 read being found at every time point and 10 reads found in at least one of the time-points. Of these 86 identified miRNA, 74 were previously classified miRNA (identified in MIRBASE), and were

found in both replicates. As well as the 74, two more known miRNA were identified in each replicate: miR171a* and miR834* in replicate A, and miR5023 and miR838 in replicate B.

3.7.3. SMALL-RNASEQ QC: CLUSTERING

Small RNA reads could be classified into four distinct clusters based on expression profiles (figure 3.15.)

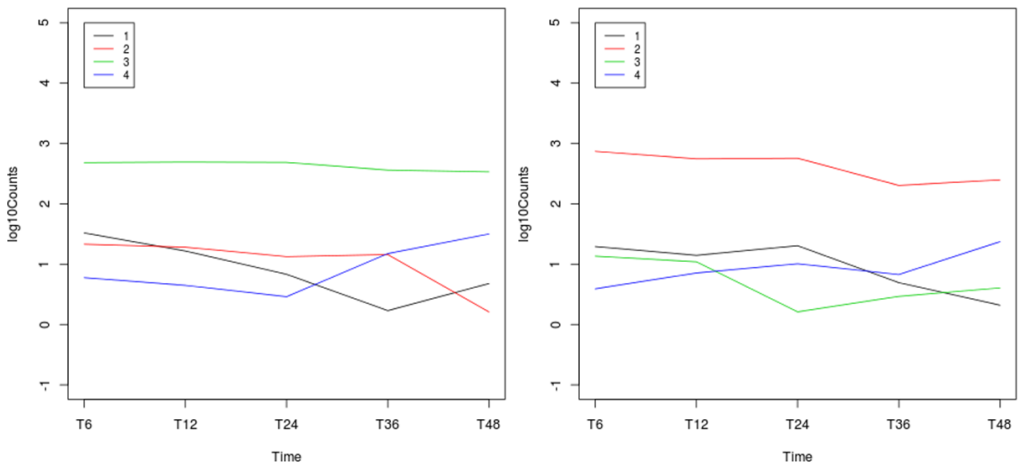


Figure 3.15: *Small RNA sequence expressions could be grouped into four distinct clusters (based on log10 counts) between replicates (A - left, B - right)*

3.7.4. SMALL-RNASEQ QC: KNOWN MICRORNAS

Small RNASeq data was further validated by the identification of known and expected microRNAs. This is discussed fully in the following chapter.

CHAPTER 4

KNOWN MICRORNAS

4. KNOWN MICRORNAS

4.1. INTRODUCTION

Though a large number of microRNAs are known to regulate plant development in *Arabidopsis thaliana*, very few of them have been implicated in lateral root development (LRD) (notably miR160, miR164, miR393, miR169 and miR390 - described in chapter 1). In order to get a comprehensive understanding of the role of microRNAs in LRD, a RNA-sequencing-based approach has been used to identify miRNAs that are expressed during this process.

Most known microRNAs have some well-established and well-defined targets, but many also have other predicted targets for which no-particular role has yet been described. Comparing the expression strengths and profiles delved from small-RNASeq, mRNASeq, and Affymetrix datasets; known miRNAs, and predicted targets that are known to play a part in lateral root development should be identified. One of the key aims of this study is to compare the expression patterns of known miRNAs and their predicted targets during LRD and to identify any miRNA/target couples that may play a role in LRD.[3]

[3] **A note on miRNA nomenclature (as outlined by miRBASE)**
The letters prefixing a particular microRNAs' name indicates the species in which the miRNA has been identified (in the case of all of those described in this study, it is 'ath' – *Arabidopsis thaliana*). Letters suffixing the miRNA name (e.g. miR160a) indicate that this particular miRNA, from a family of multiple miRNAs, is found in a particular genomic location, producing an individual isoform. Finally, an asterisk following the miRNA name (e.g. ath-miR393b*) indicates that this particular sequence is from the opposite end of the precursor to the mature sequence. These 'passenger strands' have more-recently been suggested to play a role in gene regulation themselves (as discussed in the introduction). In some cases, where the miRNA/miRNA* dynamic is less-clear, the miRNA name may be followed by the suffix -5p or -3p which indicates that the read is from the 5' or 3' arm of the precursor, respectfully (Ambros

4.2. KNOWN MIRNAS INVOLVED IN LR DEVELOPMENT

4.2.1. MOST EXPECTED MICRORNAS WERE IDENTIFIED BY SMALL-RNASEQ

Small RNA sequencing during LRD revealed a total of 57 known miRNA profiles. In all, 90 different microRNAs (as listed on miRBASE) are represented; 17 of which are represented by their * strand. Breaking down the numbers; 44 different microRNA families are represented in the data and 40 of these are represented by the traditional, mature, miRNA strand of at least one family-member. A further four families are represented solely by the * strand of at least one family member. In the case of seven families, both mature and passenger strands for some family member groups were identified. [4]

Expression patterns between replicates were compared using the Pearson Correlation Co-efficient (PCC) (Pearson, 1895). On the whole, these coefficient scores give good confidence in the two replicates, with the vast majority showing a strong correlation score (around 0.5 or above) (table 4.1.).

et al., 2003). However, this method has more recently become standard in miRBASE nomenclature due to the perceived ambiguity of the miRNA/miRNA* dynamic.

[4] When conserved miRNAs were being identified, families of microRNAs which have multiple members with identical mature sequences (based on their miRBASE records) were grouped together as they cannot be distinguished between from the small RNASeq data. For example, while miR160 has three family members (a, b, and c); miR160b and miR160c have identical mature sequences but are different to miR160a. As such, miR160 comprises 2 groups of family members.

Considering the 57 separate mature microRNA groups, ten reveal anti-correlation between replicate expression patterns; including members of the significant and well-studied miR160, miR164, miR390, and miR408 families. This difference between replicates makes what can be gained from the expression-profiles of these particular microRNAs limited. However, in cases where the microRNA expression level remains fairly consistent, or particularly low, then discrepancy between these values could be misleading, highlighting differences which, when taking into consideration the relatively flat, or low-level of expression, are not necessarily significant. Bearing this in mind, when it comes to analysing the data, every miRNA group and its expression profile must be considered on its own merits, independently.

Of the other 47 miRNA groups discovered, 34 showed a particularly high correlation between the two replicates (0.6 or more) with an average read count of 256,000. There is a huge difference (~25,500x) between the most highly expressed (5.4 million reads - miR165b/miR166cdfg) and lowest expressed (207 reads - miR2111ab) miRNA family-groups.

Known miRNA	Pierson correlation coefficient	Known miRNA	Pierson correlation coefficient
miR393b*	0.997	miR167ab	0.75
miR827	0.997	miR159a	0.741
miR156e*	0.994	miR841a*	0.726
miR5012	0.984	miR2111ab	0.667
miR394ab	0.983	miR164c*	0.648
miR399c	0.982	miR172e*	0.644
miR399f*	0.981	miR319abc	0.629
miR169d*	0.98	miR777	0.613
miR171bc	0.977	miR157abc	0.589
miR156ac*	0.964	miR829*	0.583
miR167ab*	0.96	miR5651	0.543
miR395abcdef	0.946	miR397ab	0.467
miR158a	0.944	miR156h	0.448
miR842	0.936	miR4245*	0.299
miR159c	0.932	miR399a	0.297
miR168ab	0.927	miR859*	0.277
miR840	0.92	miR160a	0.153
miR169g*	0.912	miR5654	0.112
miR169hijklmn	0.906	miR396ab	0.104
miR162ab	0.89	miR781ab	0.08
miR169e*	0.889	miR408	-0.029
miR156b*	0.87	miR160bc	-0.057
miR172b	0.848	miR398bc	-0.068
miR165b/miR166cdfg	0.832	miR861	-0.123
miR5650	0.822	miR393ab	-0.142
miR158b	0.819	miR860	-0.205
miR823	0.808	miR390ab	-0.237
miR166ab	0.804	miR8305p	-0.241
miR780.2	0.794	miR862	-0.486
miR403	0.776	miR164ab	-0.592

Table 4.1: *MIRBASE-recognised microRNAs identified in this study by small RNA sequencing of developing lateral root primordia; grouped by family members with identical sequences. Pearson correlation coefficient (PCC) is between two small RNASeq replicates*

Members of the miR165/166 family were by far the most highly expressed in the dataset. The group comprising miR165b and miR166cdfg (sharing an identical mature sequence) was by a huge margin the most highly-expressed group of microRNAs, followed by miR166ab. These two groups are the only two which have combined reads which are in the millions (5.4 million and 1.8 million, respectively). After this family, the next highest cumulative read number is for the miR319 family group, miR319ab; with 429,000 reads followed by miR159c and miR157ab with 419,000 and 312,000, respectively. There a very few microRNA families that are particularly highly expressed during lateral root development, with a read count of over 100,000 reads. Most of the miRNA families are below 20,000 reads (table 4.2.).

Taking into account differential expression, and both read-count numbers and a high-correlation between replicate expression profiles, 29 microRNA profiles were identified as attractive candidates with regards regulating LRD (table 4.2.).

miRNA	Sum of reads	PCC		miRNA	Sum of reads	PCC
miR165b/miR166cdfg	5494202	0.832		miR827	1608	0.997
miR166ab	1771960	0.804		miR171b	1091	0.977
miR319abc	429057	0.629		miR395abcdef	1067	0.946
miR159a	419566	0.741		miR171c	1063	0.914
miR157abc	312883	0.589		miR840	814	0.920
miR403	286845	0.776		miR5650	776	0.822
miR168ab	108923	0.927		miR780.2	600	0.794
miR158a	67763	0.944		miR394ab	500	0.983
miR162ab	30912	0.890		miR167ab	484	0.750
miR5651	16211	0.543		miR5012	346	0.984
miR399c	14824	0.982		miR159c	269	0.932
miR158b	14553	0.819		miR777	260	0.613
miR842	4665	0.936		miR172b	214	0.848
miR169hijklmn	4237	0.906		miR2111ab	207	0.667
miR823	2034	0.808				

Table 4.2: *MicroRNA family groups identified by small RNA sequencing with differentially-expressed profiles (expression-strength based on sum of reads of both replicates). Only previously-described, mature miRNAs, found in both replicates with a PCC between replicates of 0.5 or above are included*

In cases where there are multiple groups of family members with different mature microRNA sequences, there seems to be a fairly high correlation between the expression patterns of these different groups (such as in the case of the miR165/166 and miR156 family members, figure 4.1.); albeit the expression strengths can vary quite considerably.

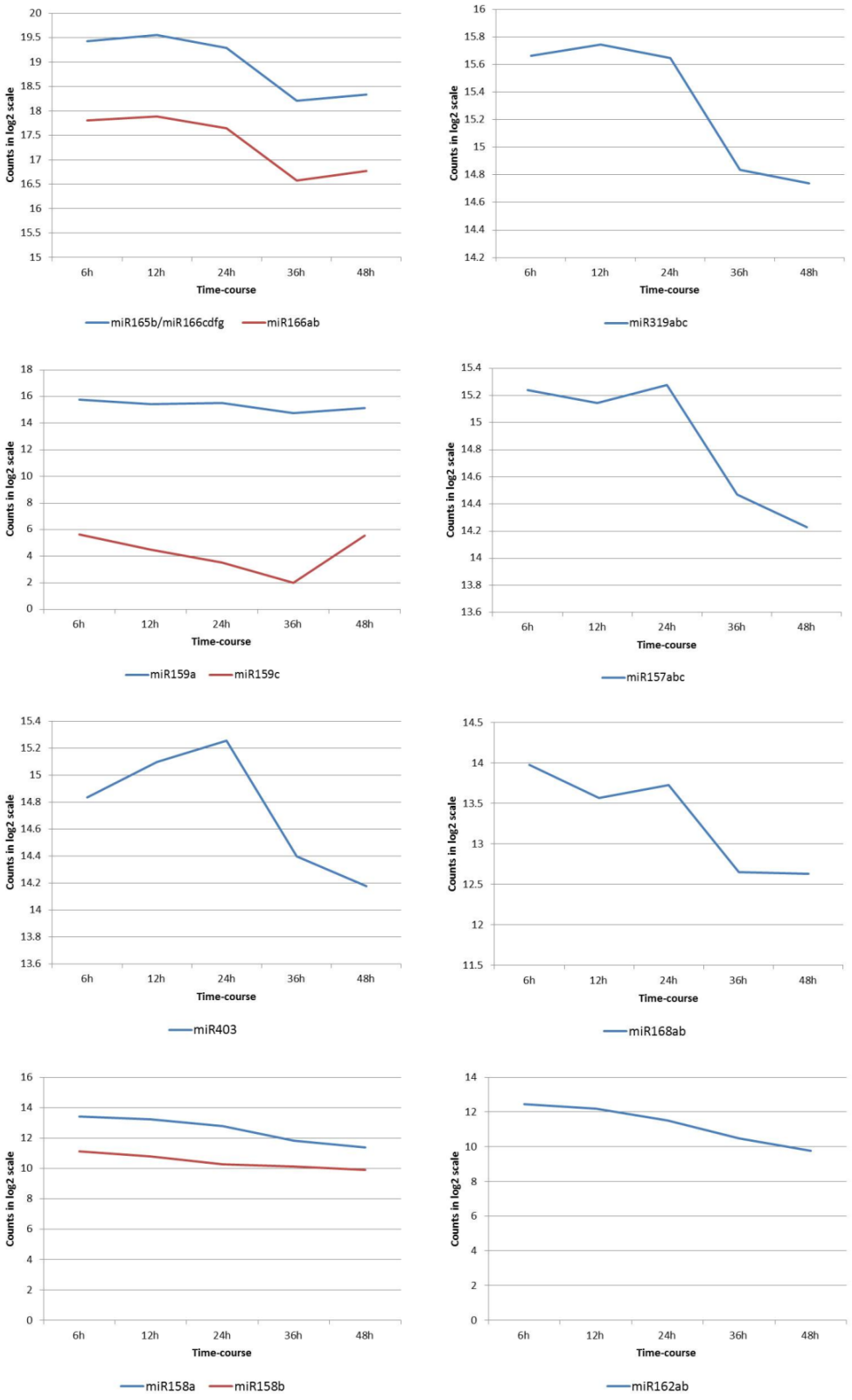


Figure 4.1: *Expression profiles of the eight most highly expressed, conserved, microRNA families identified by small RNA sequencing based on log2 counts of reads*

Taking into consideration the most highly-expressed microRNA families, there are clearly some very dynamic expression patterns and in many cases these show a downward trend, often dropping most dramatically from around the 24 hour mark (figure 4.1.).

Broadening further to the next 15 most-highly expressed microRNA family groups, similarly strong differential expression patterns can be seen across nearly all (figure 4.2.).

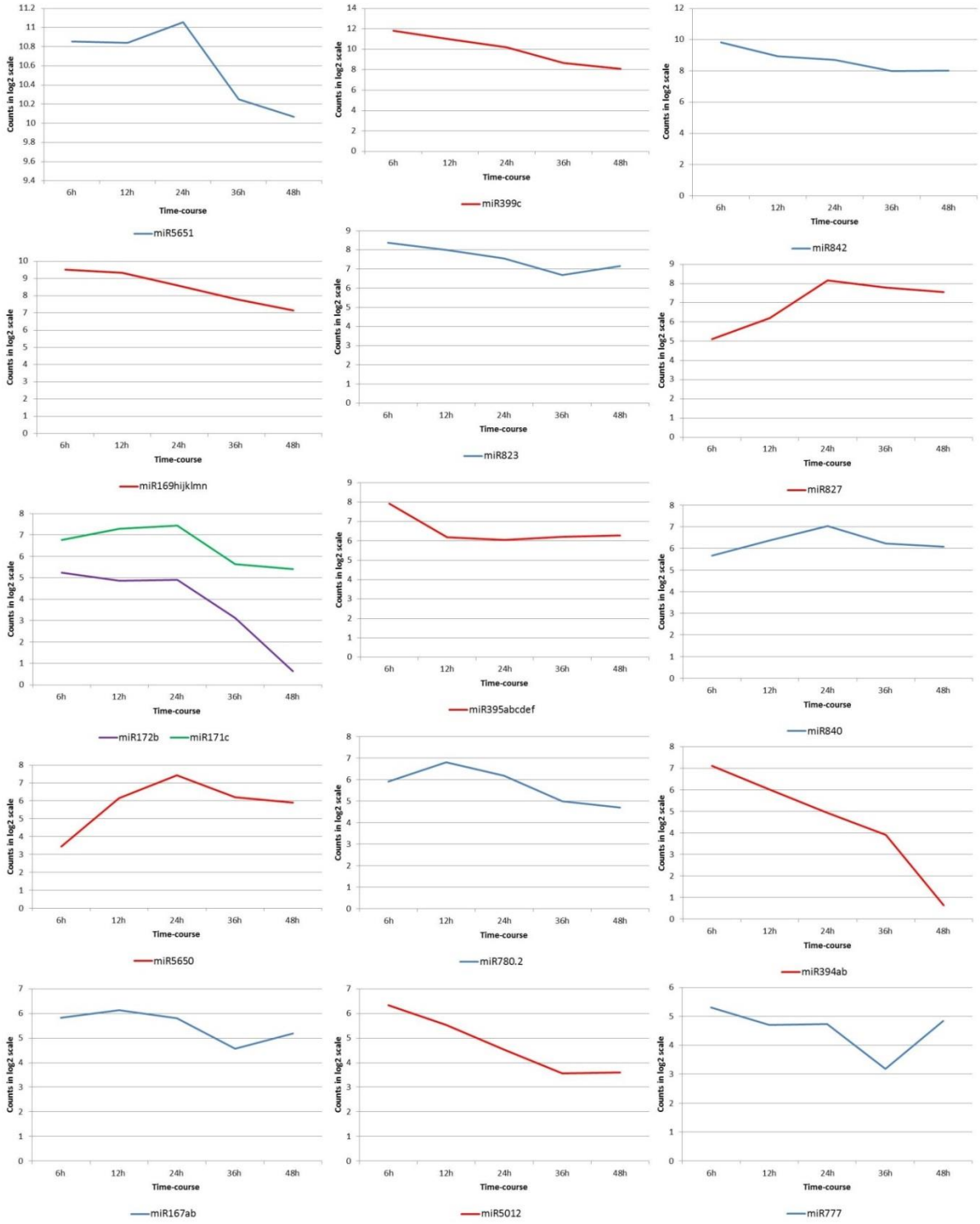


Figure 4.2*: Expression profiles of conserved, highly-expressed microRNA families identified in small RNA sequencing based on log2 counts of reads*

4.2.2. SEVERAL MIRNA* STRANDS WERE ALSO IDENTIFIED BY SMALL-RNASEQ

As well as the mature microRNA sequences mentioned, If considering also miRNA-passenger (*) strands, then thirteen miRNA family group candidates might also be considered (table 4.3.). However, in most cases, these * strand read-counts are on the lower-side so only a few might be considered as being potentially active. However, some are in-fact more-highly expressed than some of the mature microRNAs discussed earlier.

miRNA	Sum of reads	PCC
miR169d*	34,273	0.980
miR169g*	15,478	0.912
miR829*	5,189	0.583
miR167ab*	4,707	0.960
miR156b*	4,520	0.870
miR156ac*	2,268	0.964
miR156e*	1,849	0.994
miR164c*	1,604	0.648
miR393b*	1,121	0.997
miR841a*	727	0.904
miR169e*	484	0.889
miR399f*	163	0.981
miR172e*	118	0.644

Table 4.3: *Differentially expressed microRNA*sequences identified in small RNA sequencing (expression-strength based on sum of reads of both replicates, normalised). Only previously-described, miRNA '*' strands, found in both replicates with a PCC between replicates of 0.5 or above are included*

On the whole, similarly to many of the mature microRNAs identified, the miR* strands all show a downward expression trend, particularly between the 24 and 36 hour time-points

(figure 4.3.) before, generally, plateauing or rising between 36 and 48 hours. However, the very low expression of most of these * strands would lead to them being disregarded.

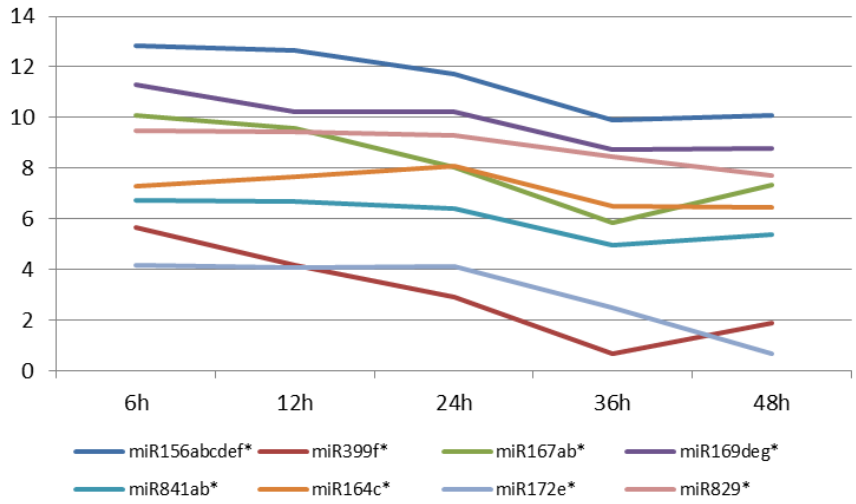


Figure 4.3: *Expression profiles (y-axis - counts in log2 scale, x-axis - time-course) of differentially expressed, conserved microRNA* strands, identified by small RNA sequencing with a PCC of 0.5 or above*

4.3. KNOWN MIRNA TARGETS

This study has identified several differentially expressed mature miRNAs during LR development. To further understand their importance to LRD, expression levels of their potential targets were investigated. It was rationalised that the true targets being regulated during this time course would also be highly differentially expressed and were likely to show a contrasting expression pattern compared to the miRNA regulating them.

To investigate the expression of potential target genes, two different datasets were used: one being the RNASeq dataset, and the other an Affymetrix dataset, produced by the Centre for Plant Integrative Biology (CPIB), the University of Nottingham (UK – Voβ *et al*., 2015). Both datasets were derived from samples collected in the same manner as those used for small-RNASeq and included the same time points used in miRNA sequencing, amongst many others.

4.3.1. MANY KNOWN MIRNA TARGETS WITH DYNAMIC EXPRESSION PROFILES WERE IDENTIFIED BY MRNA SEQUENCING

In order to look into the relationships between the identified microRNAs and their targets, expression of known and putative targets was investigated in both expression datasets. Table 4.4 includes both well-defined targets as well as other potential targets predicted in the literature or predicted using target prediction software (psRNATarget (Dai *et al*., 2011) – using a cut-off score of 2.5 unless there is only one predicted target, in which case this was raised to 3). In total, 96 targets or potential targets were identified.

Whereas all 96 targets were found in the RNASeq dataset, only 88 targets were found in the Affymetrix dataset. Five of the seven genes missing in the Affymetrix dataset, were very lowly expressed genes which might explain why they were not detected by Affymetrix; since RNASeq is so much more sensitive.

Of the 88 genes, 72 showed a positive correlation between the datasets, with 45 genes showing a PCC of over 0.5. This is a very strong correlation particularly considering that the PCC is based over such a range of time-points.

miRNA	Target	Name	PCC
miR156/157	AT2G47070	SPL1	-0.36
	AT5G43270	SPL2	0.33
	AT2G33810	SPL3	-0.28
	AT1G53160	SPL4	-0.20
	AT3G15270	SPL5	0.01
	AT1G69170	SPL6	0.12
	AT5G18830	SPL7	0.92
	AT1G02065	SPL8	A
	AT2G42200	SPL9	0.18
	AT1G27370	SPL10	-0.51
miR158	AT1G09040	Unknown	A
	AT1G49910	BUB3.2	0.71
	AT1G64100	PPR repeat-containing protein	-0.21
	AT3G03580	TPR-like superfamily protein	0.86
miR159	AT5G06100	MYB33	-0.18
	AT3G11440	MYB65	0.41
	AT2G32460	MYB101	-0.16
	AT2G26950	MYB104	0.18
	AT4G26930	MYB97	0.20
	AT5G55020	MYB120	-0.23
	AT2G26960	MYB81	-0.15
	AT3G60460	DUO1	A
	AT4G18390	TCP2	0.92
	AT4G37770	ACS8	0.13
	AT5G55930	OPT1	0.04
	AT5G18100	CSD3	0.64
	AT2G34010	MRG1	A
	AT3G61740	PHD(ATX3)	0.80
	AT3G06450	HCO3- transporter family	0.61
	AT5G17580	NPH3 family protein	0.22
	AT5G56080	NAS3	0.41
	AT4G27330	SPL	-0.01
	AT4G15530	PPDK	0.46
miR162	AT1G01040	DCL1	-0.22
miR164	AT5G07680	NAC80	0.23
	AT5G61430	NAC100	0.59
	AT3G15170	CUC1	0.18
	AT5G53950	CUC2	0.12
	AT1G56010	NAC1	-0.08
miR165/166	AT5G60690	REV	0.51
	AT1G30490	PHV	0.85
	AT2G34710	PHB	0.94
	AT4G32880	ATHB-8	0.64
	AT1G52150	ATHB-15	0.16
miR167	AT1G30330	ARF6	0.95
	AT5G37020	ARF8	0.94
miR168	AT1G48410	AGO1	0.68
mIR169	AT5G12840	NF-YA1	0.38
	AT3G05690	NF-YA2	0.26

miRNA	Target	Name	PCC
miR171	AT2G45160	HAM1	0.31
	AT3G60630	HAM2	0.4
	AT4G00150	HAM3	0.64
	AT3G47170	HXXXD-type acyl-transferase family protein	-0.22
	AT5G23250	Succinyl-CoA ligase, alpha subunit	0.95
miR172	AT5G60120	TOE2	0.82
	AT3G54990	SMZ	0.62
	AT4G36940	AP2	0.69
	AT2G39250	SNZ	0.22
	AT5G67180	TOE3	0.37
	AT2G28550	TOE1	0.69
miR319	AT2G26950	MYB104	0.18
	AT3G11440	MYB65	0.41
	AT5G06100	MYB33	-0.18
	AT1G30210	TCP24	0.96
	AT2G31070	TCP10	0.92
miR395	AT5G10180	SULTR2;1	0.83
	AT1G34355	ATPS1	0.82
	AT1G73010	ATPS2	-0.37
	AT3G47420	ATPS3	0.58
miR399	AT2G33770	PHO2	0.84
miR403	AT1G31280	AGO2	0.43
miR408	*AT2G30210*	LAC3	0.91
	AT5G05390	LAC12	0.57
	AT5G07130	LAC13	0.88
	AT2G47020	Peptide chain release factor 1	0.87
	AT3G02200	Proteasome component	0.97
	At2G02850	Plantacyanin	1.00
miR5012	AT1G53700	WAG1	0.23
	AT2G37678	FHY1	0.29
miR5650	AT2G48140	EDA4	0.97
	AT4G16144	AMSH3	A
	AT1G66880	Protein kinase superfamily protein	0.71
	AT1G80620	S15/NS1	0.98
	AT4G22200	AKT2	-0.57
	AT1G10300	NOG1-2	0.79
	AT5G62910	RING/U-box superfamily protein	0.34
	AT4G07517	CACTA-like transposase family	A
	AT4G19700	BOI	0.97
miR5651	AT1G37035	Mutator-like transposase family	A
miR777	AT1G30060	COP1-interacting protein-related	A
miR823	AT1G69770	CMT3	0.82
miR827	AT1G02860	NLA	0.52
miR829	AT5G18560	PUCHI	0.75
miR840	AT2G02740	WHIRLY 3	0.93
miR841	AT2G38810	HTA8	0.96
miR842	AT1G60130	Mannose-binding lectin superfamily protein	0.54

table legend next page

Table 4.4: *Targets and potential targets for strong miRNA candidate families. PCC indicates expression profile correlation between Affymetrix and RNASeq datasets (in cases of no PCC value, the gene was missing from the Affymetrix dataset – indicated by the letter A. Colours indicate the strength of correlation between datasets (green – PCC of 0.5 or above, orange positive correlation (PCC between 0 - 0.5), red – negative correlation or absence in Affymetrix dataset)*

4.4. IDENTIFICATION OF POTENTIAL MICRORNA/TARGET COUPLES REGULATING LRD

Based on negatively correlated expression profiles of the potential target genes, a number of strong miRNA/target candidate couples could be identified. On the whole there seem to be more strong candidates for the more conserved and typically well-described microRNA families (such as in the case for the miR165/166 family and the miR158 family) than for the less-well-described microRNA families.

4.4.1. THE MIR165/166 AND MIR156/MIR157 FAMILIES ARE TWO OF THE MOST HIGHLY-EXPRESSED AND SHOW ANTI-CORRELATION WITH CERTAIN TARGETS

4.4.1.1. MIR165/166 FAMILY

By some distance the most highly-expressed miRNA family in the sequencing data, the miR165/166 family, comprises 9 members (miR165a & b and miR166a, b, c, d, e, f, & g). These can be grouped into a number of sets based on their various shared mature miRNA sequences. As discussed in section 1.2.3.7., the miR165/166 family have been linked to the organisation of root patterning amongst other things. MiR165/166 regulates xylem differentiation by regulating the expression of the key HD-Zip III transcription factor PHAVOLUTA (PHV) in a non-cell autonomous and concentration dependent fashion (Miyashima *et al.,* 2011). Expression of miR165/166 is induced by SHR/SCR transcription factors in the endodermis. From there miR165/166 moves to the stele and regulates expression of PHV (Zhong & Ye, 2007, Sozzani *et al*., 2010, Carlsbecker *et al*., 2010) in the root. Besides this, the miR165/166 family also regulates other related HDZip III transcription factors: PHABULOSA (PHB) and REVOLUTA (REV), as well as the Homeobox transcription factors: ATHB-8 and ATHB-15, which have also been linked to cell-differentiation in the vascular meristem (Baima *et al*., 2001).

In the LR RNASeq datasets, PHB, PHV and ATHB-8 genes all show relatively strong expression and their expression patterns are strongly negatively correlated with miR165/166 expression (figure 4.4.).

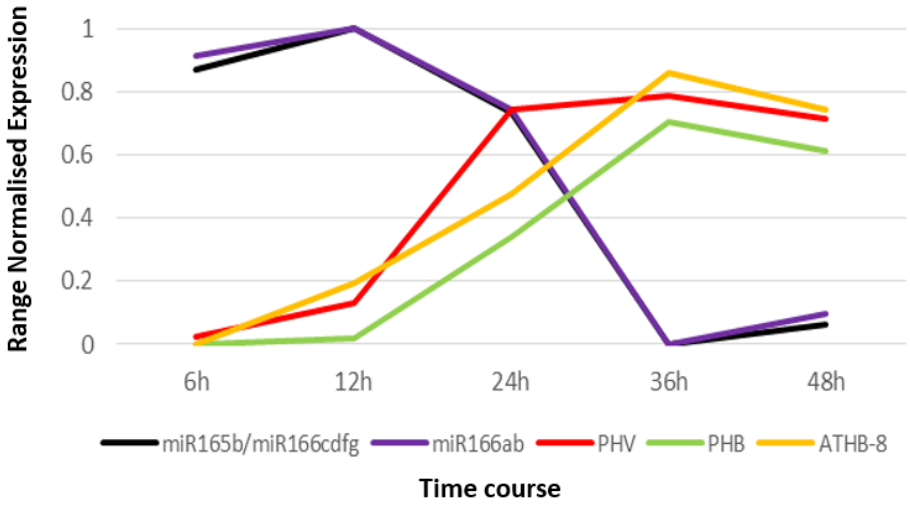


Figure 4.4*: Range normalised expression patterns (from normalised counts) for miR165/166 family members and selected targets during lateral root development (RNASeq)*

The rise in expression of these targets would hint at a possible role for the genes as the lateral root develops, regulated by the microRNAs. Hawker and Bowman (2004) have identified strong PHB and REV expression in developing lateral root primordia which, as the lateral root reaches emergence, becomes more limited to the centre of the primordia with expression localising to the vasculature as the root develops (Hawker and Bowman, 2004) (figure 4.5). The same expression pattern was seen in REV and ATHB-8. These results mirror the expression data seen here very nicely with an increase of expression over time as the primordia grows, before steadying as expression becomes more localised (figure 4.5.).

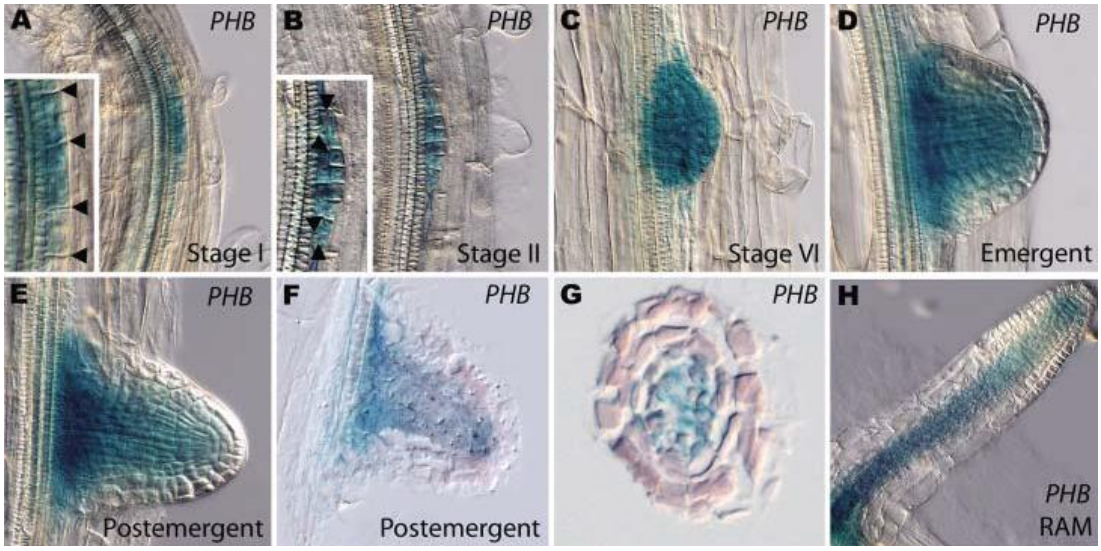


Figure 4.5: *PHB:GUS expression in developing lateral root primordia in Arabidopsis. A- F: GUS expression during LR primordia development (anticlinal divisions are represented by black arrows), G: GUS expression in a cross-section of the root apex, H: GUS expression in the root apical meristem (RAM). Expression of REV and ATHB8 mirrored this precisely - adapted from Hawker and Bowman, 2004.*

Further to these expression studies; a PHV, PHB, REV triple mutant (*phb-6,phv-5,rev-9*) has been found to show fewer lateral roots and these genes have already been suggested to promote LR formation in a similar way to the role they play in aerial lateral organ formation (Hawker and Bowman, 2004).

Grigg *et al*. (2009) have subsequently described how the restriction of PHV and PHB by miR165/166 is essential for proper embryonic root development and root meristem establishment. A failure to do this can lead to embryonic lethality. Furthermore, Grigg *et al.*

describe this pathway in relation to the importance of Serrate (SE – a mutant of which displays a short root phenotype) in regulating miR165/166 itself.

These studies all point towards the importance of these genes and their strict control by miR165/166 for many, crucial, plant developmental processes. Clearly HDZIP-III transcription factors (and their regulation) are crucial to proper root development and as such, also seem to be central to proper lateral root development. Their presence and striking expression profiles (showing striking anti-correlation with those of the miR165/166 family groups) in our LR development data set point towards regulation of these targets by the miR165/166 family during this particular process quite strongly.

Given the nature of miR165/166, its targets, and their varied roles including vasculature development and cell differentiation, it is to be expected that they are likely to be central to proper LR development in some capacity, not least as the vasculature of the LR begins to form. This potential role is supported strongly by the expression profiling of PHB and REV by Hawker and Bowman (2004), and the expression profiles revealed in this study, rising dramatically over the course primordial development. The drop seen after 36 hours would also point to a part in the early stages of vasculature development rather than a continuous, on-going role.

This microRNA/target family is one of the best described and the miR165/166 family are one of the highest-expressed in Arabidopsis. As such they were always one of the first families to identify with a view to validating the dataset. The expected, strong anti-correlation between the miR165/166-family microRNAs and their well-defined and strongly-expressed targets

clearly endorses the quality of both small RNA and mRNA datasets. The datasets can thus be comfortably considered for use in the investigation of other, negatively correlated miRNA/target couples for which a role in in lateral root development may be less-well defined or not known.

4.4.1.2. MIR156/157 FAMILY

Similarly to the miR165/166 family, miR156 and miR157 have also been classified as a single family due to their sequence similarity and conserved targets (Naya *et al*., 2010). In total, there are 14 family members (miR156a-j and miR157a, b, c, and d) and they regulate the expression of 11 of the 17 Squamosa-promoter binding-like proteins (SPLs) (Wu *et al*., 2006).

SPL genes play an important role in plant development, plant maintenance and defence. SPL3, 4, and 5 have all been linked to flowering timing (Yamaguchi *et al*., 2009), while SPL8 has been linked to the proper functioning of both male (Xing *et al*., 2010) and female (Xing *et al*., 2013) fertility. SPL2 has previously been found to be involved in shoot maturation in association with SPL10 and SPL11 (Shikata *et al*., 2009) and SPL6 has been linked to a regulatory role in pathogen defence, transcriptionally, in association with an immune receptor (Padmanabhan *et al.,* 2013).

With regards Arabidopsis roots; SPL9 has been found to directly regulate TRICHOMELESS1 (TCL1) and TRIPTYCHON (TRY), key genes in trichrome and root hair development (Yu *et al*., 2010, Wang *et al*., 2007). A miR156-resistant form of SPL9 has also shown very early

flowering (Wang *et al*., 2009, Wu *et al*., 2009). As such, this particular target should be expected to be highly-expressed in the root samples sequenced here.

With regards our small RNASeq data, the group of miR157a, b and c were the only mature microRNAs identified and were amongst the highest expressed (beneath only the miR165/166, and miR319a family members, and miR159a; the last of which is discussed in more detail later on in this chapter). Interestingly, although no mature miR156 family miRNAs were identified, four miR156* isoforms were found (miR156* a, b, c and e), though at a much lower level (in terms of expression) than the miR157 microRNAs.

Ten of the SPL targets which have been associated with miR156/miR157 were identified in the RNASeq and Affymetrix datasets and all are reasonably highly expressed with the exception of SPL8. Taking into consideration expression profiles/strength and anti-correlation to miRNA expression patterns, three of these targets stand out as being potentially regulated by miR157a and miR157b over the course of Arabidopsis LR development (SPL2, 6 and 9) (figure 4.6.). SPL2 and SPL6 are particularly highly expressed in the RNASeq data.

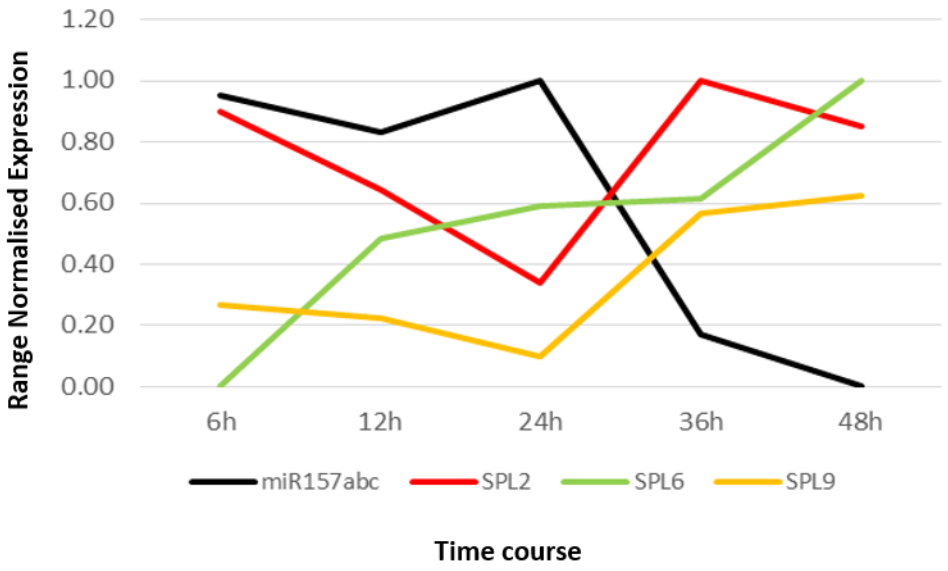


Figure 4.6: *Range normalised expression patterns (from normalised counts) for miR157 family members and selected targets during lateral root development (RNASeq)*

That SPL8 has been specifically linked to fertility, it is probably to be expected that its expression would be low in the roots (as it proves to be). Other targets such as the pathogen defence-involved SPL6 would be expected to be expressed more globally and may vary depending on necessity as prompted by outside conditions.

The small RNASeq and RNASeq results produced in this study; like those of miR165/166 and that family's targets; do hint at a role for some of these SPL targets in Arabidopsis lateral root development. SPL2 and SPL9 in particular are intriguing. The strong, dynamic expression of miR157 and these targets strongly suggests that repression by the microRNA is lessened over the time-course allowing far higher expression of these genes as the primordia develops and reaches root emergence stages.

Until recently, no lateral root phenotype had been described for any of these targets though one would expect, due to the huge change in expression in this dataset that there may-well be developmental defects in a mutant of SPL9 or SPL2.

However, very recently, a study has been published which does indeed describe a role for the miR156/SPL module in Arabidopsis lateral root development (Yu *et al*., 2015). Yu *et al*. describe how the overexpression of miR156 leads to the growth of more lateral roots and outline a module whereby SPL3, SPL9, and SPL10 are involved in the repression of lateral root development in a process which seems to be auxin responsive (figure 4.7.). This supports strongly our data whereby SPL9 expression rises as the LR develops suggesting its repression may be necessary for LR initiation and less-so as the root develops.

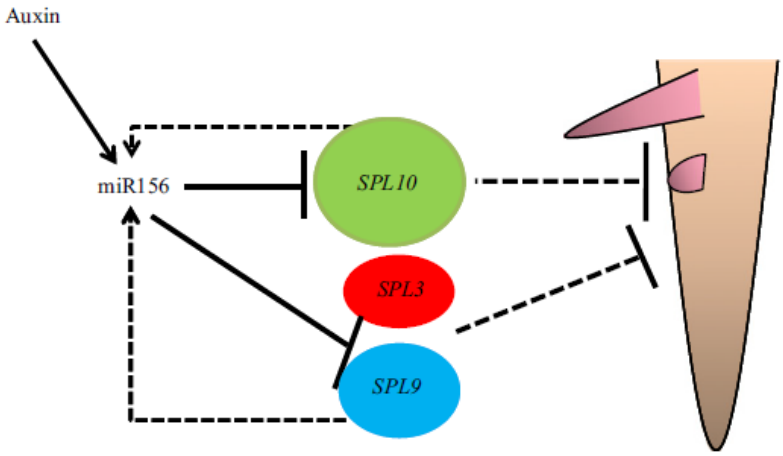


Figure 4.7: *A model for miR156/SPL regulation of LR development. SPLs repress lateral roots in response to auxin via repression by miR156. A positive feedback mechanism is also suggested whereby SPL10 & SPL9 repression can indirectly prompt miR156 activation (adapted from Yu et al., 2015)*

Yu *et al*. (2015) also found SPL2 to be particularly highly expressed in roots and the strong similarity between the expression patterns of SPL2 and SPL9 identified in this study does suggest that this too could be similarly involved though Yu *et al*. (2015) did not link this particular gene to LR development.

Ultimately, coupled with those of the similarly highly and dynamically expressed miR165/166 family, these results further emphasise the quality of the RNASeq and small RNASeq datasets created in this study.

4.4.2. MIR158, MIR395, AND MIR162 TARGETS SHOW ANTI-CORRELATION WITH THEIR MICRORNA REGULATORS EXPRESSION PROFILES

Following on from the miR165/166 and miR156/157 families, two more highly-expressed microRNA families (miR158 and miR162) also show strong anti-correlation with one or more of their defined targets. At the same time, the lower-expressed miR395 shows anti-correlation with a target which has been previously identified in Arabidopsis root vasculature, and linked to a response to sulphur starvation.

Both members (a & b) of the miR158 family were identified in the small RNA sequencing data; both showing high, and differential, expression. These microRNAs have been linked to four diverse targets: BUDDING UNINHIBITED BY BENZYMIDAZOL 3 (BUB3.2), a PPR repeat-containing protein, a TPR-like superfamily protein, and an unknown protein (from TAIR10 descriptions). Two of these show dynamic expression patterns in both RNASeq and

Affymetrix datasets which correlate negatively with the two miR158 family members (figure 4.8.).

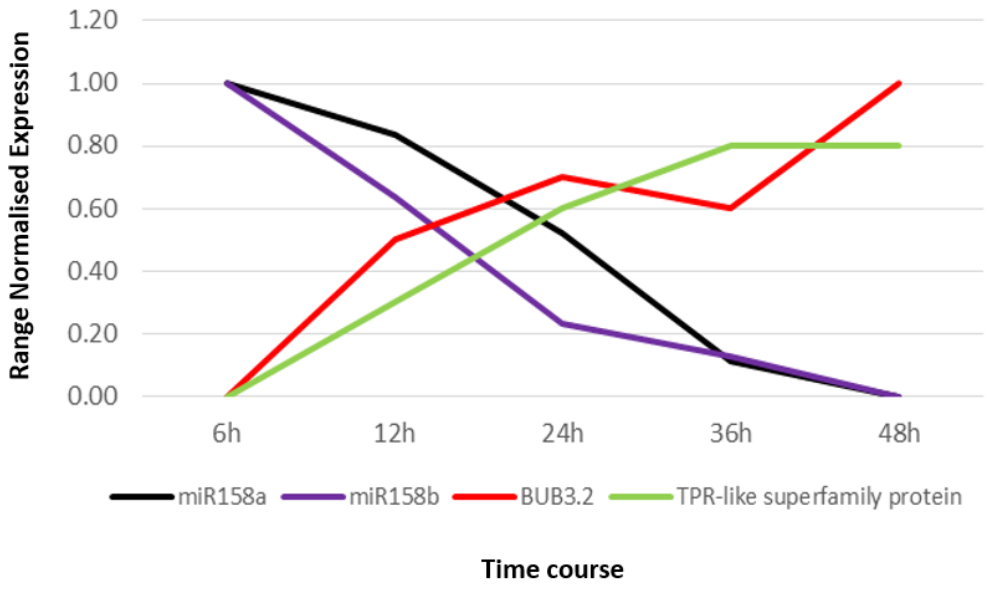


Figure 4.8: *Range normalised expression patterns (from normalised counts) of miR158 family members and selected targets during lateral root development (RNASeq)*

Whilst nothing substantial has been described about the TPR-like superfamily protein, BUB3.2 has been described as a mitotic checkpoint protein with a role in controlling proper spindle-assembly. As such this might be expected to be found to be expressed ubiquitously throughout the plant and more highly expressed during localised developmental processes such as LR development.

The rise in expression seen in our dataset may well reflect this due to the localised cell proliferation over the course of LR development. The very strong anti-correlation between miR158 and this target would hint at a necessary release of BUB3.2 repression during this process.

Similarly to the miR158 family, the miR162 family was found to be amongst the highest-expressed in the small RNASeq dataset. MiR162 has one clear, predicted target, DICER-LIKE ONE (DCL1), which is central to proper microRNA processing from precursor to mature sequence. As described in section 1.2.1., DCL1 cleaves the pri-miRNA hairpin, producing the miRNA/miRNA* duplex with 5' phosphates and two-nucleotide 3' overhangs (Schauer *et al*., 2002, Kurihara and Watanabe, 2004). Xie *et al*. (2003) describe the negative feedback regulation that DCL1 mRNA is subjected to by the activity of miR162 (figure 4.9.).

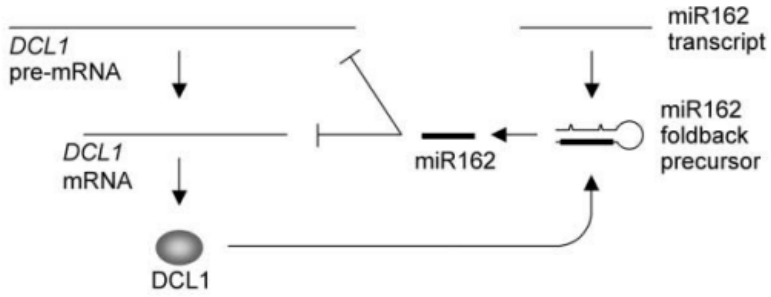


Figure 4.9: A m*odel of the negative feedback regulation of DCL1 by miR162 in Arabidopsis. DCL1 mRNA is maintained at a low level due to miR162-directed cleavage and subsequent degradation (adapted from Xie et al., 2003)*

The miR162 family comprises two members and the identical mature sequence produced by both was identified in our small RNA sequencing dataset here. The miR162 family was amongst the higher expressed of all the miRNAs identified. DCL1 was also fairly highly expressed and showed a dynamic expression pattern. However, unlike many of the other miRNA/target couples described already, this pattern did not correlate negatively with that of the miR162 family (PCCs of 0.89) (figure 4.10.).

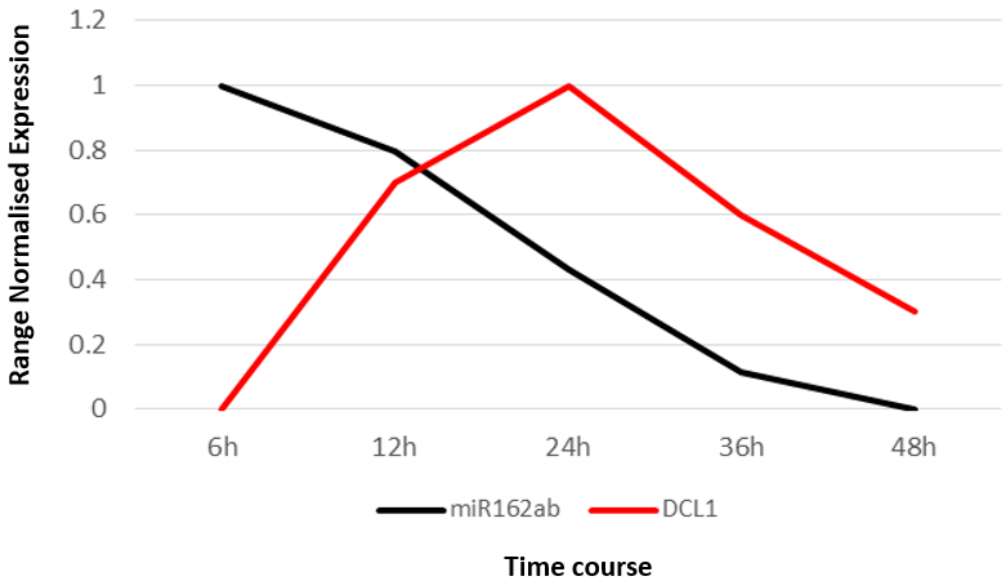


Figure 4.10: *Range normalised expression patterns (from normalised counts) for miR162 family members and their target, DCL1, during lateral root development (RNASeq)*

Because of the negative-feedback system, it is probably to be expected that microRNA and target levels would follow similar expression patterns, with one regulating the other and vice-versa; and in the case of the RNASeq data produced here, this does seem to be the case.

As to be expected, a mutant of DCL1 reveals reduced miRNA accumulation (Park *et al*., 2002). The mutant has a severe, compromised global phenotype which includes an increased lateral root phenotype (figure 4.11.) (Guo *et al*., 2005). This root phenotype has been specifically linked to the increase of NAC1 due to a reduction in miR164 activity (Guo *et al*., 2005), though it is fair to assume that due to the breadth of influence that DCL1 would have, this could be the culmination of the miss-processing of a number of microRNAs.



Figure 4.11: *The DCL1 mutant - dcl1-9 shows a dramatic lateral root phenotype in Arabidopsis thaliana when compared to wild-type (Landsberg erecta). White bar indicates 1cm. Numbers below images indicate the number of LRs per centimetre of primary root (average ± SE). Adapted from Guo et al. 2005.*

Although considerably lower-expressed than the miR158 and miR162 family members, all six members of the miR395 family share the same mature sequence and this sequence was identified as being differentially expressed by small RNA sequencing. MiR395 shows a very distinct expression pattern, dropping after the first time point and remaining very low from 12 to 48 hours (figure 4.12.). This profile would strongly suggest that if this miRNA family is indeed repressing a target, then higher-expression of this target would likely be necessary for the later stages of LR development, post-initiation.

Despite the lower-expression levels, the miR395 profile is still well-conserved between replicates and shows strong anti-correlation with three of their four predicted targets (the other one proved inconclusive due to poor correlation between its-own two replicates) (figure 4.12.).

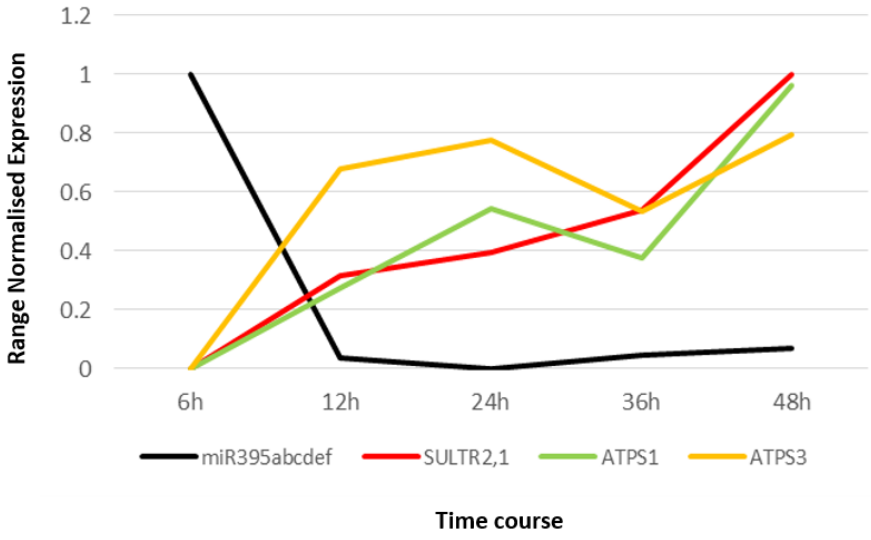


Figure 4.12*: Range normalised expression patterns for miR395 family members and selected targets during lateral root development (from RNASeq and small RNASeq)*

The miR395 family has previously been found to be highly expressed in the vasculature of both the primary and lateral roots in Arabidopsis (Kawashima *et al*., 2009) (figure 4.13.). It has been found to be induced by sulphur starvation, controlled by the transcription factor SULPHUR LIMITATION1 (SLIM1) (Kawashima *et al.,* 2009). Interestingly, Kawashima *et al*. describe how one of miR395's targets increases in expression during miR395 induction in roots. This particular target, SULTR2,1, encodes for one of a family of sulphate transporters involved in sulphur uptake from the soil by roots (Yoshimoto *et al*., 2002). Both miR395 and its target, SULTR2,1 are induced by sulphur starvation. Ultimately, this serves to limit expression of the target to the xylem parenchyma which aids in the transport of sulphate to the shoots while limiting shoot-to-root transport (Kawashima *et al*., 2011).

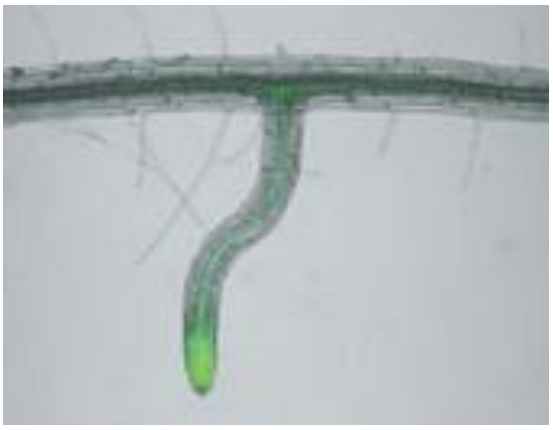


Figure 4.13: *MIR395d promoter-GFP shows high expression in the developed lateral root in Arabidopsis thaliana (Adapted from Kawashima et al., 2009)*

Interestingly, Kawashima *et al.* (2011) describe quite different responses to sulphur starvation from two more of mi395s targets; while SULTR2,1 expression rises, ATPS4 drops

and ATPS1 remains constant. Kawashima *et al.* point towards very different models of action for each of these targets in relation to the microRNA targeting them.

While ATPS4 expression was inconclusive in our dataset, the other two targets: APS1 and APS3, both show dynamic expression profiles, similar to SULTR2,1. APS1 and APS3 are ATP sulphurylases which are also involved in the sulphate metabolism pathway following uptake. Both have been predicted to localise in plastids (Hatzfield *et al*., 2000). These ATP sulphurylases activate sulphates and this ultimately leads to the production of sulphur-containing metabolites (Leustek and Saito, 1999, Saito, 2004). Liang *et al.* (2010) confirm that miR395 is involved in the regulation of sulphate accumulation under stress through the targeting of these genes.

According to our data, during LR primordia development, this miRNA family is not especially-highly expressed. No role has yet been described for any of these targets or this microRNA family with regards Arabidopsis LR development. Considering the relative low-expression of miR395 in our dataset it might be unwise to speculate too wildly on a role for it here, in spite of the strong anti-correlation between expression profiles for the microRNA and the dynamic targets. Nonetheless, the dramatic drop in expression might suggest a potential role, specific to the initiation of lateral root primordia, although this does seem contradictory to the high expression in the root described by Kawashima *et al*. (2009). Furthermore, no phenotype has yet been described.

4.4.3. THE LOWER-EXPRESSED MIR403, MIR399, MIR823, AND MIR840 FAMILIES, ALSO SHOW STRONG ANTI-CORRELATING EXPRESSION PROFILES WITH SELECTED TARGETS

Looking at the less-highly expressed microRNAs with clearly-defined targets, a few more anti-correlated miRNA/target couples can be identified. In some of these cases, less is known about the miRNA dynamic and the precise role that they may play.

MiR403 has been found to target the ARGONAUT protein, AGO2 (Allen *et al*., 2005). This particular target has subsequently been found to be induced by viral silencing suppressors (Lewsey *et al*., 2010, Endres *et al*., 2010) and an antiviral defence role has been confirmed by Harvey *et al.* (2011). A similar role has been identified for AGO1, the same protein which forms RISC with microRNAs which leads to target transcript cleavage (Qu *et al.,* 2008)

AGO2 was found to be very highly expressed in both RNASeq and Affymetrix screens with a very dynamic expression profile. MiR403 showed very similar expression patterns between replicates. In both cases, miR403 seems to drop in expression over the time-course (figure 4.14.). This results in a PCC of -0.3 in-between target and microRNA which would certainly hint at possible regulation at this time which may suggest the possibility of a specific role in root development between this miRNA/target couple though no such role has yet been described and no root phenotype identified.

There are six members of the miR399 family in *Arabidopsis thaliana* (miR399a-f). However, of these only one (miR399c) was identified by small RNA sequencing. MiR399c is a relatively highly-expressed microRNA and shows a dynamic expression profile which drops suddenly at around 24 hours (figure 4.14.). Of the six family members, it is also the one with by-far the most distinctive mature sequence. The other five share relatively similar sequences.

MiR399 has been found to target PHO2 (encoding an ubiquitin-conjugating E2 enzyme) (Aung *et al*., 2006; Bari *et al*., 2006). Regulation of this target by the miR399 family helps control inorganic phosphate (Pi) homeostasis through a complex regulatory network (Lin *et al*., 2008). Lin *et al*. explain a mechanism by which miR399 moves from roots to shoots in order to enhance Pi uptake and translocation during Pi deficiency (2008).

PHO2 is very highly expressed in the RNASeq data and shows particularly strong anti-correlation of expression with miR399c (figure 4.14.).

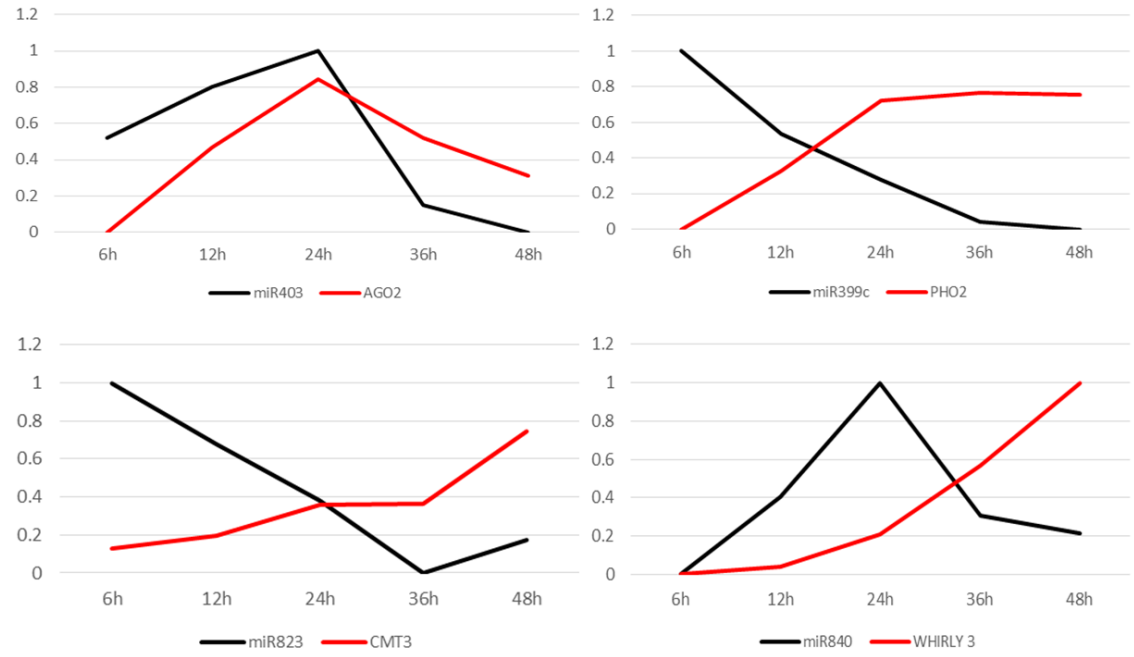


Figure 4.14: *Range normalised expression patterns (from normalised counts) for miR403, miR399c, miR823, and miR840 and selected targets, during lateral root development (RNASeq)*

Finally, miR823 and its target CHROMOMETHYLASE3 (CMT3), and miR840 and its target, WHIRLY3, also show a very strong anti-correlation with regards their RNASeq profiles (figure 4.16.). The miRNA823/CMT3 module is involved in DNA methylation and has been linked to pathogen defence with the observation that the microRNA is itself regulated by fungal elicitors (Baldrich *et al*., 2014). The WHIRLY family of transcription factors have, similarly to many miRNA targets described in this study, largely been linked to pathogen defence mechanisms in plants (Deseraux *et al*., 2005).

All of these miRNA/target couples which range wildly in their documented roles, whether they are induced by external stimuli, and with wildly differing expression strengths; serve mostly to reassure the quality of the dataset produced in this study. Many microRNAs which have been identified in Arabidopsis (many of which have not been discussed in any real detail here) were identified along with almost-all expected targets. Furthermore, there was often high anti-correlation between these expression patterns. While it may be inaccurate to speculate too much on possible roles in LR development for some of the lower-expressed miRNA/target couples, it is still to be expected that expression patterns would often show a certain degree of anti-correlation between regulator and target at any given time.

In this study we have identified several miRNA/target couples that may potentially play an active role in lateral root development. For the sake of this study, however, only a small number of microRNAs and their targets, were selected for further investigation (section 4.6. and beyond).

4.5. USING TARGET MIMICRY TO STUDY IMPACT OF KNOWN MIRNAS ON ROOT DEVELOPMENT

A number of microRNAs which were uncovered in this study (as well as a few more of initial potential interest) were selected for the production of target mimics. These target mimics would be used for various phenotypic and expression-based studies of the microRNAs themselves and their targets.

4.5.1. TARGET MIMICRY OVERVIEW

Target mimics act as 'RNA sponges' and sequester microRNAs, effectively producing miRNA knockdown lines. They were first considered as a potential tool for miRNA repression when the non-protein coding gene IPS1 (INDUCED BY PHOSPHATE STARVATION 1) was discovered to contain a motif with sequence complementarity to miR399 (Franco-Zorilla *et al.,* 2007). Curiously, it also contained a mismatched loop at the expected cleavage site. This meant that the miRNA would bind to the sequence but would be unable to cleave the transcript. As such, Franco-Zorilla *et al.* were able to develop a protocol for synthesising target mimics for any miRNA for which the target sequence was known. These mimic sequences, based around the IPS1 gene and driven by the strong 35S cauliflower mosaic virus (CaMV)-promoter, would bind to miRNAs in the organism and release the targets of that particular microRNA from repression.

Using this approach, target mimics for all known Arabidopsis miRNAs (as of 2010) have been produced (Todesco *et al*., 2010). Analysis of these mimic lines was deemed to be an extremely powerful approach with regards investigating some of the known miRNA which have a potential role in Arabidopsis lateral root development.

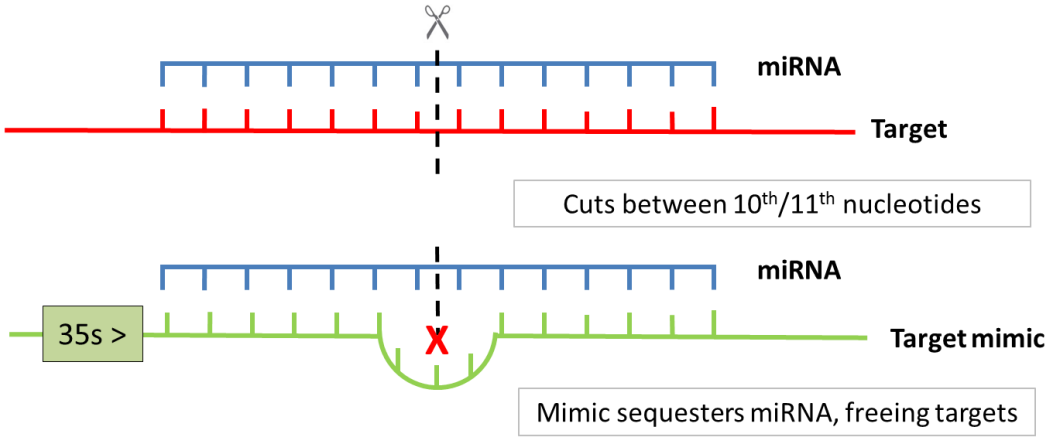


Figure 4.15: *Basic outline of the action of Target mimics. Target mismatches at the site at which the miRNA would ordinarily cut the transcript prevent cleavage from happening. An excess of these transcripts, driven by 35S, sequesters miRNAs, thus freeing the original targets from repression*

While Todesco *et al*. (2010) studied the aerial phenotypes of many of the target mimic lines to various degrees; the root phenotypes (and more-specifically, any potential lateral root phenotypes) were largely overlooked.

4.5.2. MICRORNAS SELECTED FOR TARGET MIMICRY INVESTIGATION

Fifteen microRNAs were selected for target mimic production and investigation (table 4.5.). These were selected on the basis of either strong expression profiles in the small RNASeq data or of their targets in the RNASeq data. Alternately, they were considered as potentially interesting based on the other roles of that particular miRNA, links of their targets to

developmental processes (such as in those which target ARFs), or previously noted roles or phenotypes (such as the strong aerial phenotype of miR159) (Todesco *et al*., 2010).

MiR156	MiR163	MiR172	MiR408
MiR157	MiR164	MiR393	MiR775
MiR159	MiR397	MiR396	MIR823
MiR160	MiR168	MiR165/166	

Table 4.5: *miRNAs were selected for Target mimic production in this study (originally designed by Todesco et al., 2010)*

MiRNA target mimic plasmids for the 14 miRNA, produced by Todesco *et al*., were obtained from the National Arabidopsis Stock Centre (NASC, Nottingham), transformed into plants through Agrobacterium strain LBA4404, and selected with BASTA. Seeds for one independent line of the target mimic for miR159 were obtained directly from Detlef Weigel (Max Planck Institute for Developmental Biology, Tübingen, GE). Due to various growth issues, of the fifteen, only a small number succeeded in producing plants.

Mimic lines were selected using BASTA, genotyped, and homozygous lines were identified. These lines were still used for subsequent studies. Ultimately, four miRNA target mimic lines were produced and considered for subsequent study (miR159, miR160, miR164, and miR397 (henceforth known and MIM159, MIM160, MIM164, and MIM397, respectively). However, in each case only one independent, homozygous line could be produced and/or validated. The strength of the mimics' action varies between different lines (Todesco et al., 2010).

Brief aerial phenotyping of young seedlings revealed a mild, serrated-leaf phenotype in MIM164 and a far more extreme, compromised leaf-phenotype in MIM159 (as described by Todesco *et al*., 2010). MIM397 did not seem to have an early-stage aerial phenotype (figure 4.16.)

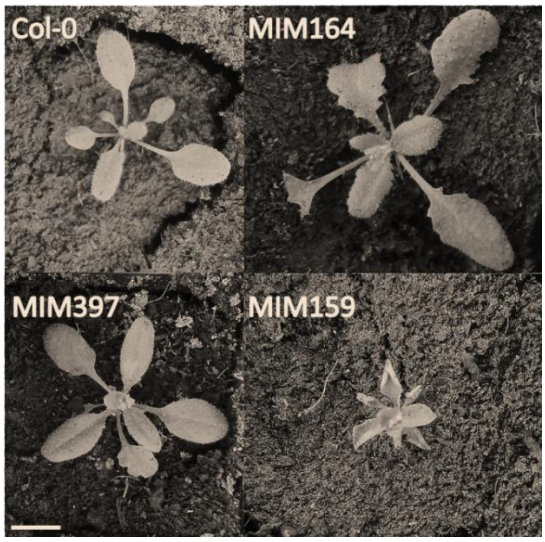


Figure 4.16: *Aerial phenotypes of young seedlings (3 weeks) of selected miRNA target mimic lines (MIM164, MIM397 and MIM164). White bar represents 1cm*

4.5.3. SELECTED TARGET MIMIC LINES SHOW NO OBVIOUS LATERAL ROOT PHENOTYPE WITH THE EXCEPTION OF MIM159

The miRNA target mimic lines which were produced were screened for lateral root numbers, primary root length and lateral root density after 14 days' growth. While one might expect a phenotype in the target mimic lines for miR160 and miR164 due to the nature of their

targets, no statistically-significant difference from wild-type could be detected. The same could be said of MIM397. However, MIM159 showed a dramatic and significant drop in both lateral root number and subsequently lateral root density (PR length remained comparable to wild-type) (figure 4.17.).

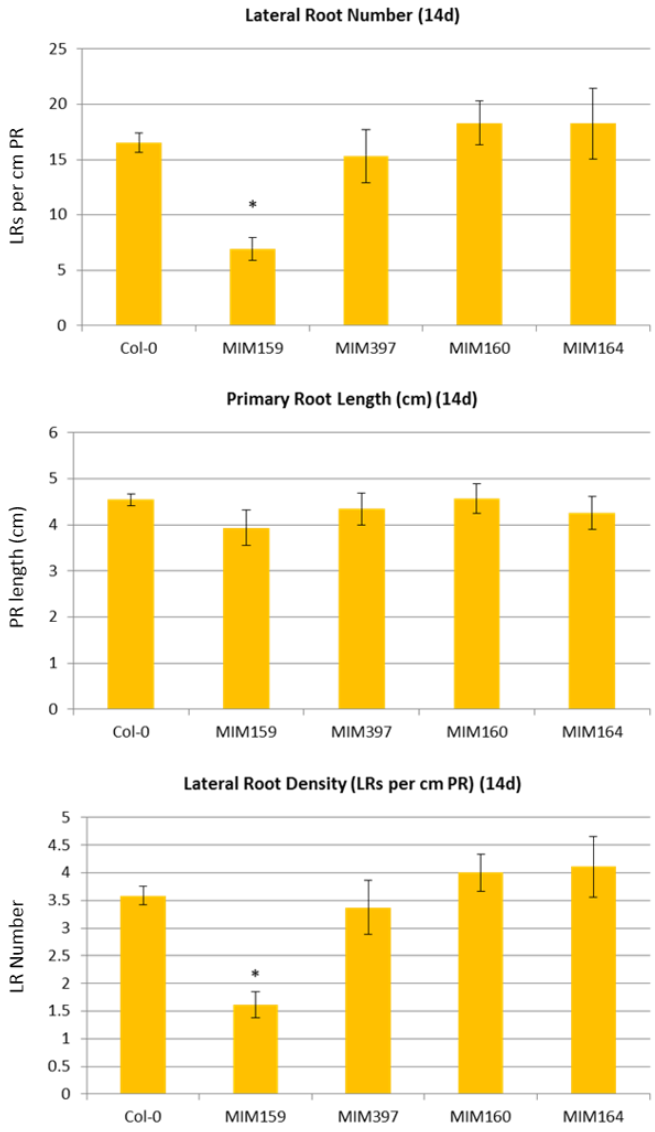


Figure 4.17: *Phenotyping of target mimic lines for miR159, miR160, miR164, and miR397. Lateral root numbers, density, and primary root length were measured after 14 days growth;*

compared against wild-type (Col-0) Arabidopsis. Statistically significant difference is identified with an asterisk (student's t-test <0.05). Error bars represent standard error.

MiR159 was immediately identified as a microRNA family of particular interest and the study turned to its targets and their action during lateral root development.

4.5.3.1. MIR159 EXPRESSION SHOWS STRONG ANTI-CORRELATION WITH THAT OF A NUMBER OF TARGETS

MiR159 is one of the most abundant miRNA families in Arabidopsis with around 20 predicted targets (Allen *et al*., 2010). These targets include a number of MYB transcription factors including MYB33 and MYB65, which are the two well-established targets of miR159a and miR159b. MYB101 has previously been highlighted as a target of miR159a through 35S:miR159a-based overexpression studies (Schwab *et al*., 2005) and the identification of cleavage products (Palatnik *et al*., 2007, Reyes *et al*., 2007, Alves-Junior *et al*., 2009). However, Allen *et al*. (2010) subsequently suggested that MYB101 expression is independent of miR159 regulation altogether. Besides these targets, functional specificity of miR159a and b for other predicted targets remains doubtful as regulation of these targets reveals negligible phenotypic effects (Allen *et al*. 2010).

Of the three miR159 family members (all of which have different mature sequences), only miR159a and miR159c were identified in the small RNASeq screen of developing lateral root

primordia. Both reveal dynamic expression patterns which fell towards the 36 hour time point of LR primordia development before rising again at 48 hours (figure 4.18.).

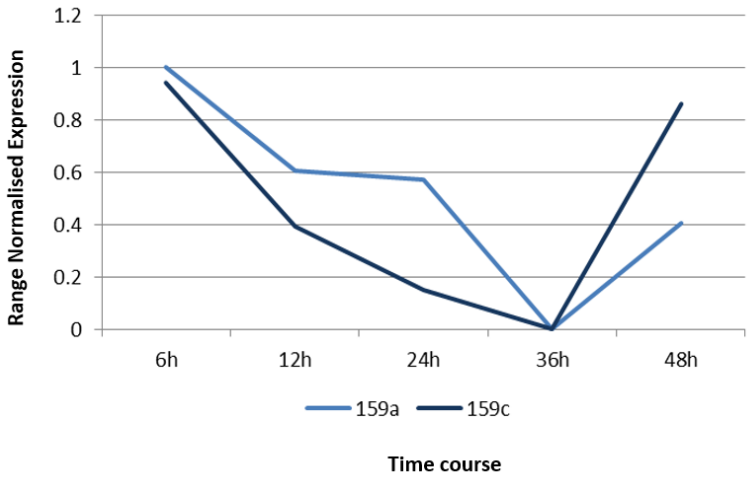


Figure 4.18: *Range-normalised expression profiles (from normalised counts) of miR159 family members identified in small RNA-sequencing of developing lateral root primordia (0-48 hours following gravi-stimulus). Profiles are an average of two replicates (PCC of 0.77 & 0.97 for miR159a and miR159c, respectively).*

Both MYB33 and MYB65 are expressed highly during LR development while MYB101 is weakly expressed. As well as these, several of the other predicted miR159 targets are also differentially expressed during LR development but only a small number showed significantly high expression with NICOTIANAMINE SYNTHASE 3 (NAS3) by far the highest-expressed of these others.

Encouragingly, expression profiles of five of the targets (MYB33, CSD3, PHD, NAS3 and PPDK) showed strong anti-correlation to both of the miR159 family members identified (figure 4.19., figure 4.20.), while one other (NPH3), showed a weak anti-correlation with miR159a (PCC -0.27) (table 4.6.). MYB65 showed the opposite with a strong positive correlation with both miR159 families.

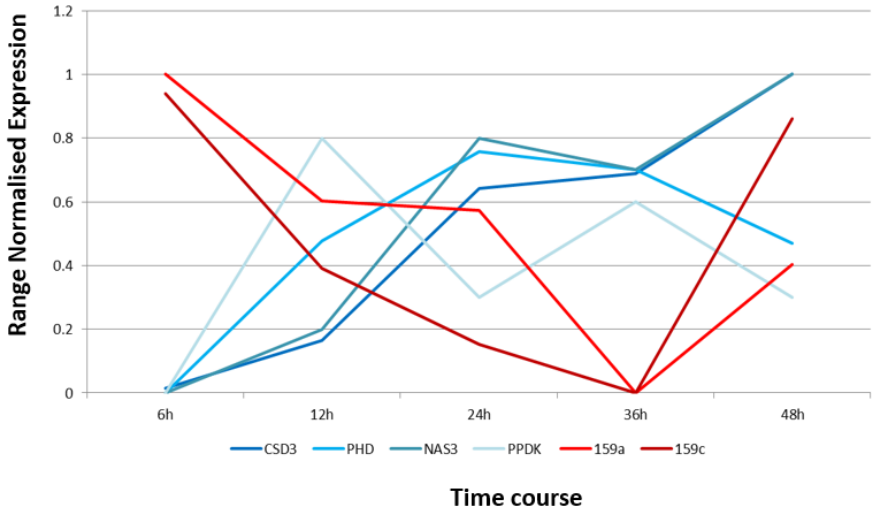


Figure 4.19: *Range normalised RNASeq expression profiles (from normalised counts) of highly-expressed miR159 targets and predicted targets during lateral root primordia development alongside expression profiles of miR159 family members.*

Target/Potential target		PCC of expression profile	
		against miR159a	against MiR159c
MYB33	AT5G06100	-0.810	-0.838
CSD3	AT5G18100	-0.697	-0.366
PHD(ATX3)	AT3G61740	-0.623	-0.718
NPH3	AT5G17580	-0.275	-
NAS3	AT5G56080	-0.684	-0.461
PPDK	AT4G15530	-0.543	-0.548

Table 4.6: *Highly-expressed targets and predicted targets of miR159 family members which show anti-correlation between target and microRNA RNASeq expression profiles*

MYB33 is possibly the most intriguing target. It was particularly highly expressed and its expression profile was the most negatively-correlated to the two miR159 family members (although it has not specifically been described as a target of miR159c) (figure 4.20.). This target has already been well validated and a double mutant of this target and MYB65 has revealed a globally compromised phenotype (Allen *et al*., 2010).

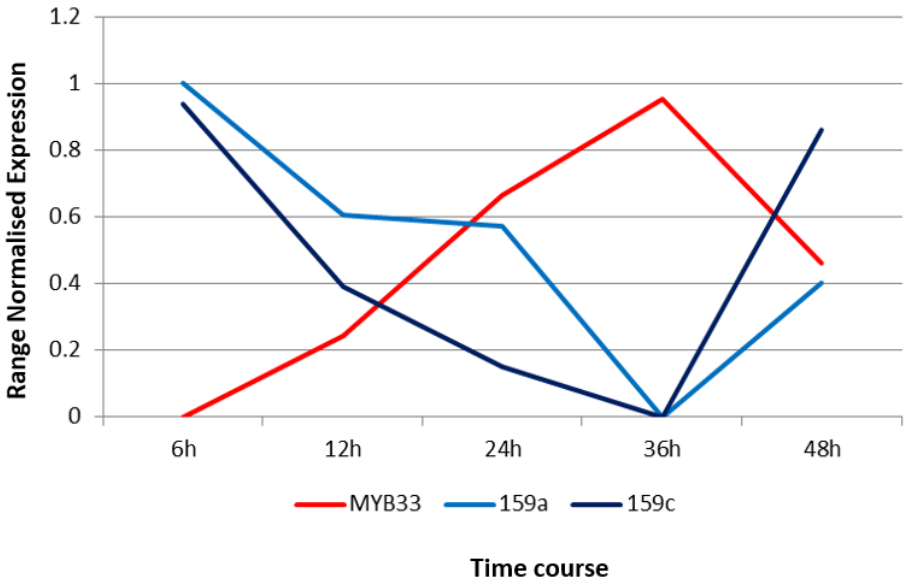


Figure 4.20: *Range normalised RNASeq expression profiles (from normalised counts) of MYB33 during lateral root primordia development alongside expression profiles of miR159 family members.*

4.5.3.2. MYB33 SHOWS A DRAMATIC INCREASE IN EXPRESSION BETWEEN MIM159 AND WILD-TYPE

Subsequent expression studies further highlighted the repression of MYB33 by miR159. RT qPCR was run on cDNA from MIM159 seedlings for the three primary miR159 targets of interest: MYB33, MYB65, and MYB101. In 10-day-old seedlings; MYB33 was expressed highly when compared to wild-type (Col-0) while the other two targets showed a far more comparable expression levels to Col-0 (figure 4.21.). This suggests that repression of MYB33 by miR159 is far stronger throughout the plant than the other two targets.

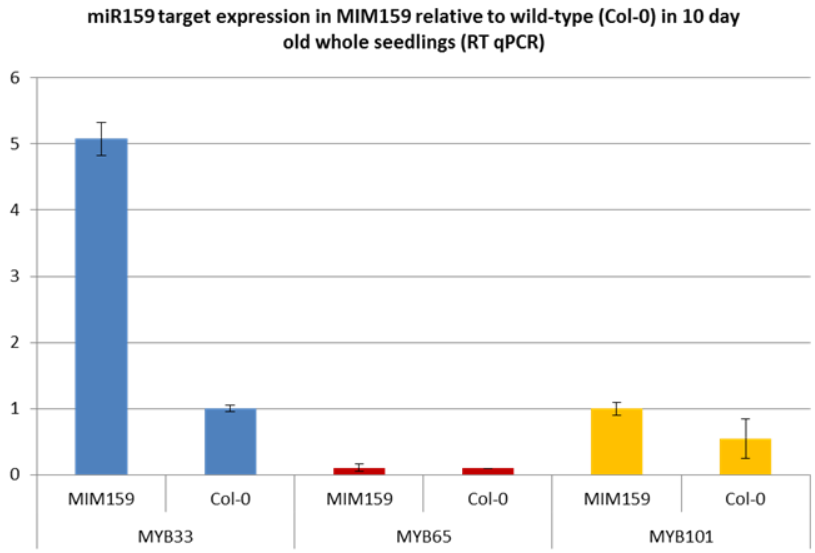


Figure 4.21: *miRNA RT qPCR of miR159 targets; MYB33, MYB65, and MYB101 on whole, 10 day-old MIM159 and Col-0 whole-seedlings. Error bars represent standard error.*

Further study looked to validate the expression change of miR159 targets during lateral root development. RT qPCR was used to identify MYB33 expression in NPA-NAA treated root samples from MIM159 and wild-type (Col-0). Seedlings were grown on NPA (10µM), blocking LR development before being transferred to NAA (10µM) which prompted excessive primordia development. RT qPCR was conducted on RNA from whole root systems of these two lines after 0, 12, and 24 hours NAA exposure (so the primary root had numerous, relatively-synchronised developing primordia).

Similarly to the study on whole seedlings, MYB33 expression was high but was found to rise significantly over time in the target mimic while staying relatively low in wild type (figure 4.22.). This was contrary to MYB65 and MYB101 which were run in the same circumstances

(see appendix A.4.). While MYB33 expression does rise over time in the wild type by about 1.5 to 2 times over 24 hours, in the mimic line there is an almost 10-fold increase in MYB33 expression upon auxin treatment. This may suggest that miR159/MYB33 module may play an important role during LR development and repression of MYB33 is crucial for normal LR development

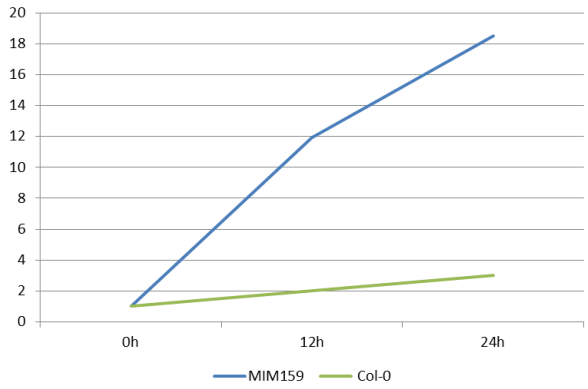


Figure 4.22: R*elative RT qPCR expression of MYB33 in 10-day-old MIM159 and Col-0 root systems following NAA treatment (1 – 0h, 2 – 12h, 3 – 24h). Expression was normalised against actin. Expression of MYB33 rose strongly over the time-course in the mimic line while it rose at a slower rate in wild-type.*

While MYB33 seems to be the obvious candidate for a potential role in LR development (as regulated by miR159), it is not clear whether the strong lateral root phenotype seen in MIM159 is simply a consequence of the compromised aerial parts. Nonetheless, the strong, dynamic expression and complimentary miRNA/target expression patterns would certainly suggest that miR159 may directly regulate LR development through repression of this particular target.

4.5.3.3. MYB33 EXPRESSION IS INCREASED IN THE ROOTS HUGELY IN A MIR159-RESISTANT GUS LINE

In order to look at spatial expression of MYB33 in the roots, a MYB33:GUS line and a miRNA159-resistant mMYB33:YFP line were obtained (from the Tony Millar lab, AUS). These revealed high expression of MYB33 throughout the roots (particularly in the vasculature, root tip, and throughout the developing primordia) while it was repressed, almost entirely, in the miR159-resistant line (figure 4.23.). This is consistent with mMYB33:GUS expression in the aerial parts of the plant varying hugely between the two lines (figure 4.24).

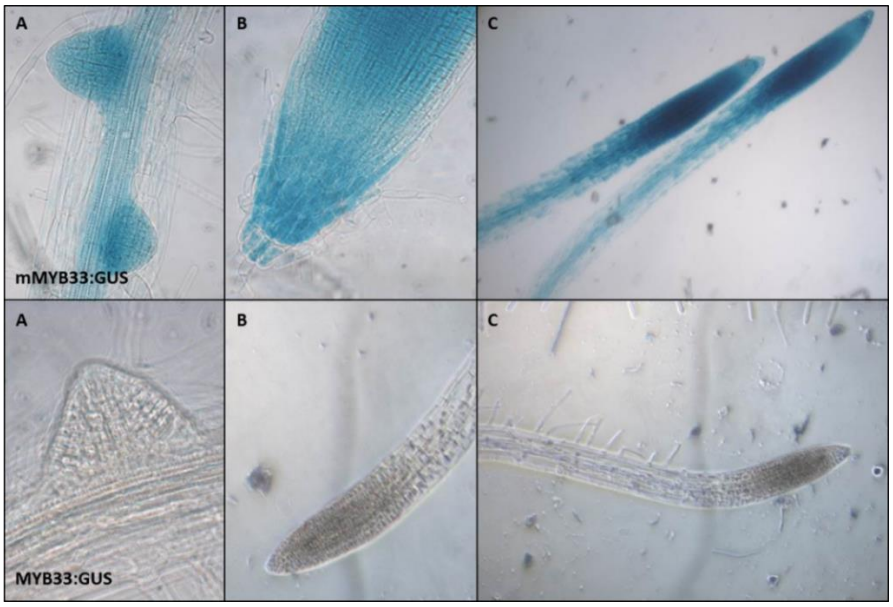


Figure 4.23: *Expression of MYB33 in the roots of MYB33:GUS and miRNA159-resistant mMYB33:GUS lines. A – emerging lateral root primordia, B – primary root tip, C – developed lateral roots. Expression can clearly be seen throughout the root in the miR159-resistant line (mMYB33:GUS)*

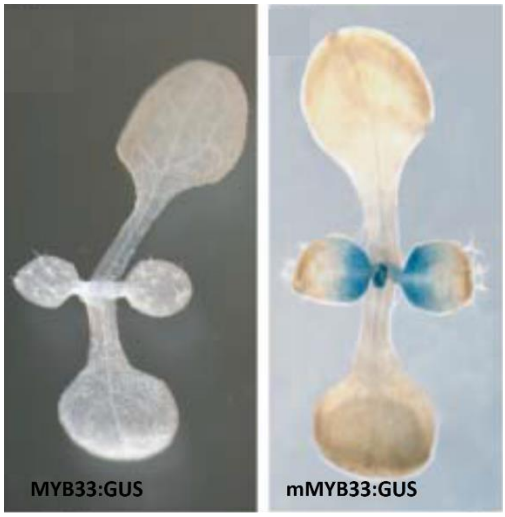


Figure 4.24: Expression of MYB33 in the early-leaves of *MYB33:GUS and miRNA159-resistant mMYB33:GUS lines. Adapted From Millar et al., 2005*

4.5.3.4. THE MIM159 TARGET MIMIC AND MIR159AB DOUBLE MUTANT BOTH REVEAL A STRONG, COMPROMISED ROOT PHENOTYPE

The miR159 family members and their two principal, validated targets (MYB33 and MYB65) have all been studied to some degree and a strong phenotype has been identified in many lines where either the microRNA or the target has been disturbed or affected (such as in MIM159). Because of strong differential expression of miR159 family members in the roots and the strong anti-correlation of some of the well-established and predicted targets' expressions during LR development, the lateral root phenotypes of a *miR159ab* double mutant (obtained from Dr Tony Millar, Australian National University (Allen et al., 2010)) and MIM159 were studied in more depth with LR numbers, PR length and LR density being measured across large numbers of seedlings after 12 and 14 days growth. Detailed morphology studies were also undertaken.

Consistent with the initial phenotyping of MIM159, these results show that while the primary root length was not significantly altered (compared to wild-type - Col-0), lateral root numbers and lateral root density were dramatically reduced in both MIM159 and *miR159ab* (figure 4.25. and figure 4.26).

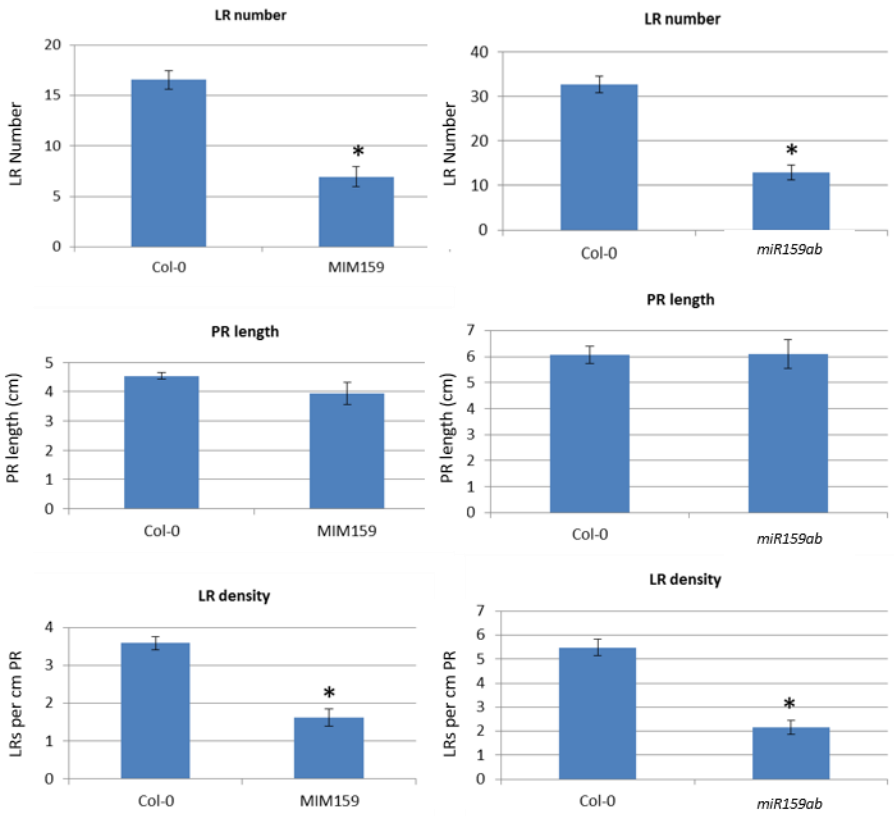


Figure 4.25: *Lateral root number (LRN), primary root length (PR length – cm), and lateral root density (LRD – lateral roots per cm PR) of MIM159 (left) and miR159ab (right), compared against wild type (Col-0) Arabidopsis. Observations were made at 12 (MIM150) or 14 day (mir159ab) old seedlings. Asterisks indicate statistically significant difference (student's t-test <0.05). Error bars represent standard error*

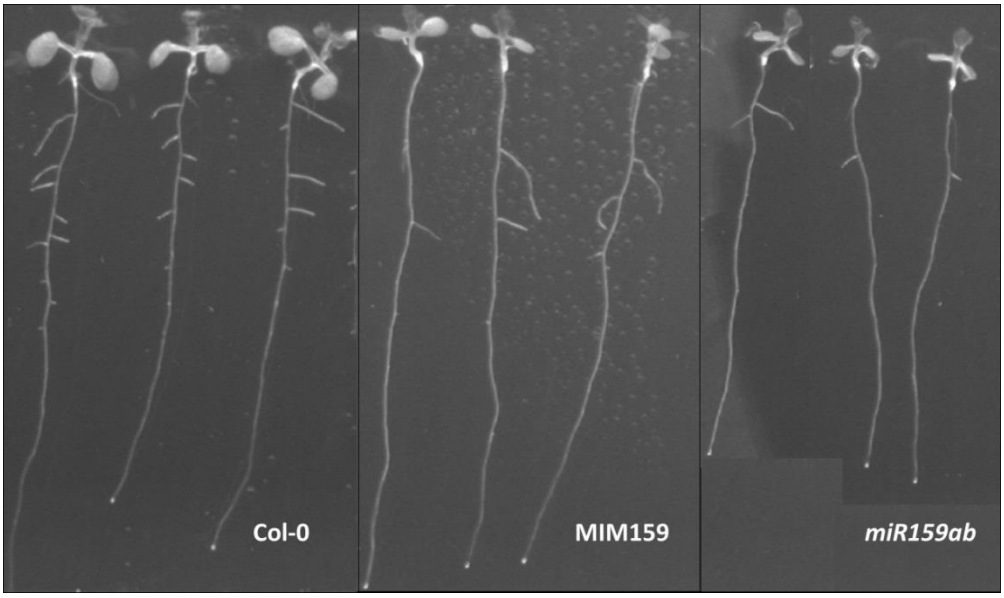


Figure 4.26: MIM159 and the m*iR159ab double mutant show a dramatic reduced lateral root phenotype after 10 days growth when compared against wild type (Col-0)*

4.5.3.5. THE MIM159 ROOT PHENOTPYE CANNOT BE SIGNIFICANTLY RECOVERED BY A VARIETY OF ROOT-RELATED HORMONES

Plant hormones are key regulators of root development. To investigate whether the LR phenotype of miR159 compromised plants was due to a defect in hormone signalling and/or transport pathways; MIM159 seedlings were grown in the presence of ethylene (ACC), gibberellic acid (GA), abscisic acid (ABA) and auxin (NAA). Each hormone was used at two different concentrations (ACC at 100nm & 200nm, GA at 1µM & 1.5µM, ABA at 10µM & 15µM, and NAA at 100nM & 175nM); and tested for the restoration of the lateral root defects of MIM159. With regards GA, ABA, and ACC, Investigation of primary root growth in the presence of these hormones suggests that hormone perception does not seem to be

affected in MIM159 and that it responds to these hormones in a more-or-less wild type fashion (figure 4.27).

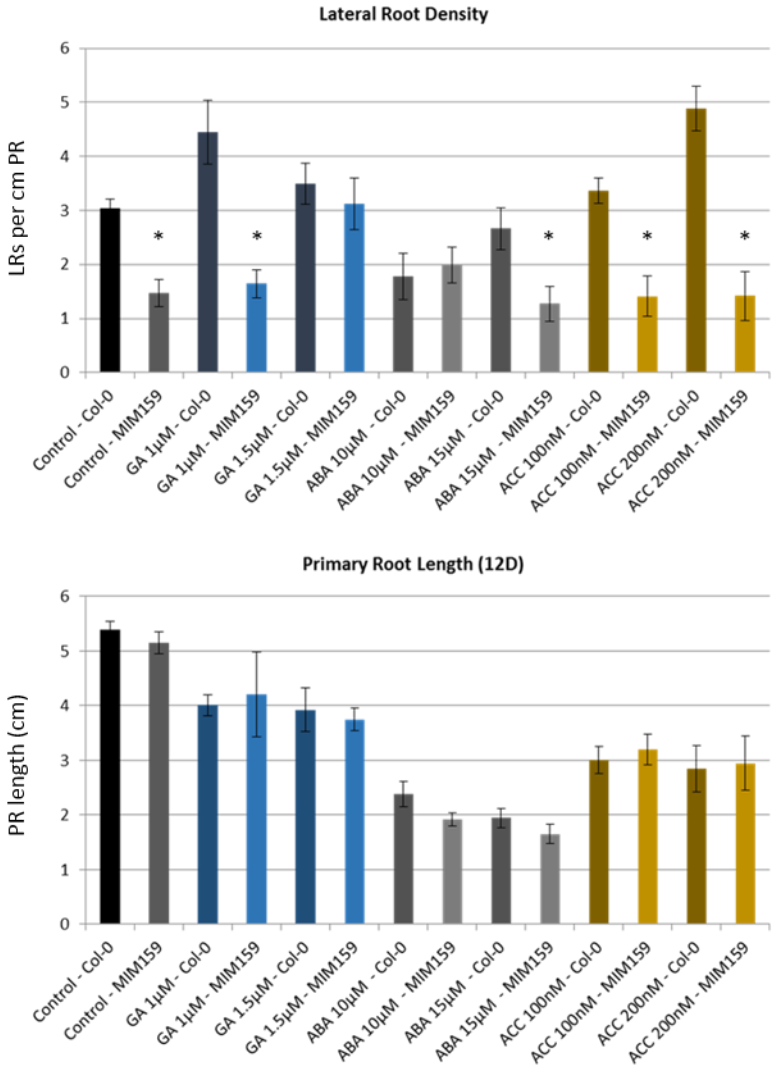


Figure 4.29: *Primary root length (cm) and lateral root density (LRs per cm PR) of MIM159 after 12 days' growth on 1/2MS media treated with various concentrations of selected hormones (GA, ABA, and ACC). Error bars indicate Standard error. Significant differences (student's t-test score of less than 0.05) between MIM159 and Col-0 are indicated with an asterisk*

GA at the higher concentration seems to completely restore the root phenotype and would suggest that miR159 may be involved in regulation of GA signalling. This may be linked to the fact that the GYMYB-like MYB gene-targets of miR159 (MYB33, MYB65 etc…) have been linked to GA-mediated processes (Gocal *et al*., 2001). MYB33, for-instance, has been found to be induced by GA, leading to the initiation of flowering processes (Gocal *et al.,* 2001) and these results may suggest that there is also a role for GA-regulation of some of these MYB genes in root development in conjunction with miR159. None of these other hormones seem to have an effect on restoration of MIM159 lateral root defect.

4.5.3.5.1. AUXIN

Auxin plays a very important role in LR development (as described in chapter 1) and of all of the hormones studied here; it is perhaps the most important in this regard. Auxin promotes LR development and as expected there is a significant increase in LR primordia in the wild-type control under both auxin concentrations. Interestingly, for MIM159 under the lower auxin concentration, there is also an increase in LR numbers and this increase is, relative to the numbers of LRs seen in this line without auxin, roughly the same, proportionality, than the increase seen in wild-type. The higher concentration of 1-NAA (175nM) seems to push LR density in MIM159 further towards the wild type comparative, effectively closing the gap. While even at this concentration, there is no significant difference in LR numbers between Col-0 and MIM159 lines (figure 4.28)

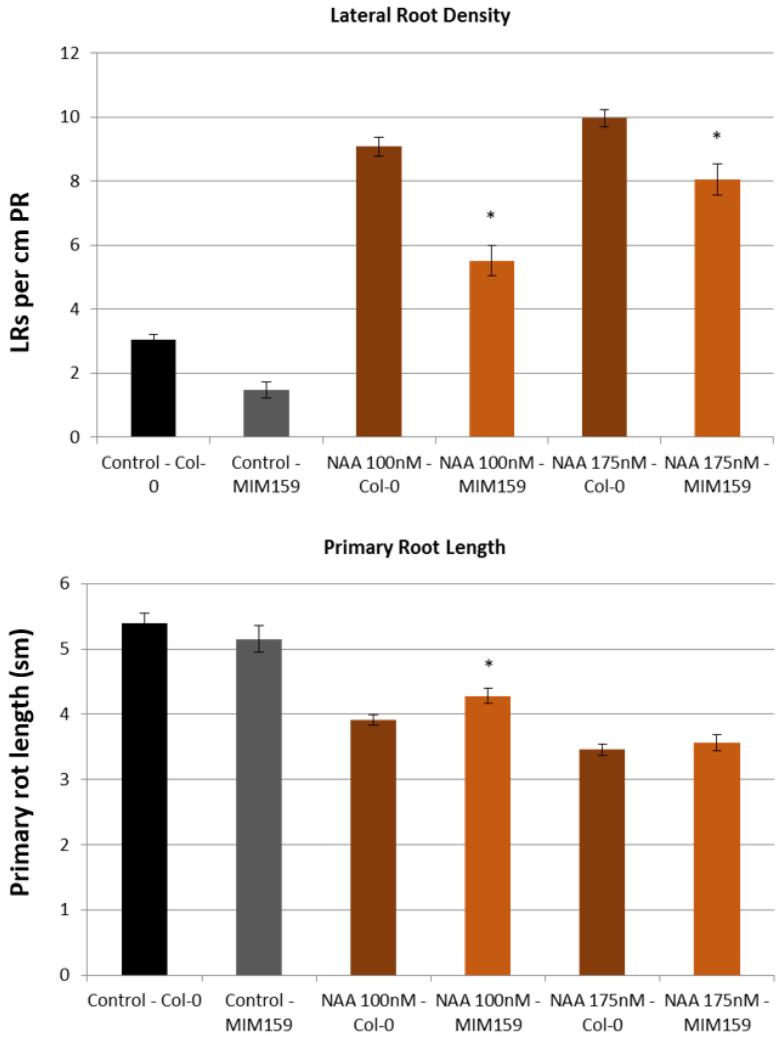


Figure 4.28: *Primary root length (cm) and lateral root density (LRs per cm PR) of MIM159 after 12 days' growth on 1/2MS media treated with two concentrations of auxin (NAA). Error bars indicate Standard error. Significant differences (student's t-test score of less than 0.05) between MIM159 and Col-0 are indicated with an asterisk*

Ultimately, this demonstrates that auxin at a particularly high concentration can produce a similar response in MIM159 to wild-type. This suggests that the defect in MIM159 roots must be upstream of the auxin pathway.

A cross between MIM159 and a AUX1:YFP line was made in order to better visualise any effect on LR primordia morphology during development. The AUX1 gene (similarly to AUX1-LIKE (LAX) genes, Swarup *et al*., 2008), is an auxin influx carrier responsible for the promotion of lateral roots (Marchant *et al.,* 2002). No clear difference could be identified in the expression of AUX1 between the MIM1959xAUX1:YFP line and the AUX1:YFP line (figure 4.29.). AUX1 was found to be clearly expressed in the rot epidermis and stele as well as the developing LR primordia. No change in cell morphology could be identified.

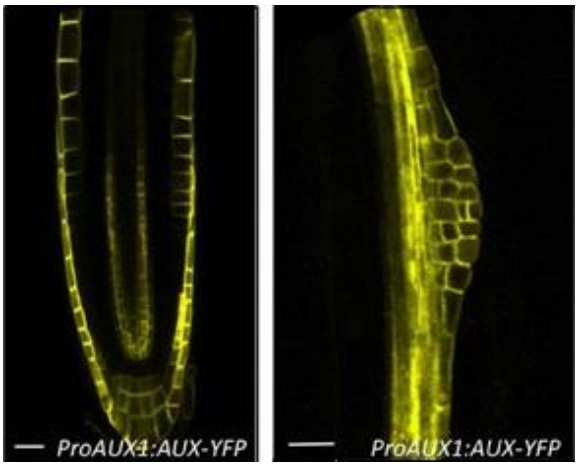


Figure 4.29: *ProAUX1:AUX-YFP (adapted from Swarup et al., 2012). MIM159xAUX1:YFP shows an identical expression pattern.*

4.5.3.6. NO CHANGE IN LR PRIMORDIA MORPHOLOGY OR DEVELOPMENTAL TIMING CAN BE IDENTIFIED IN *MIR159AB* OR MIM159

In order to further investigate lateral root defects in miR159-related lines; the morphology of developing LR primordia was studied more closely in MIR159 and *miR159ab* at specific stages of lateral root development following gravistimulus.

However, upon investigating developing primordia prompted by a gravistimulus in young seedlings, no significant difference in either root patterning nor developmental timing could be identified in between MIM159 or *miR59ab* and wild type (Col-0) (figure 4.30.). This would suggest that any defect in LR development in these lines would more likely be linked to early LR priming or initiation, prior to the development of the primordia itself.

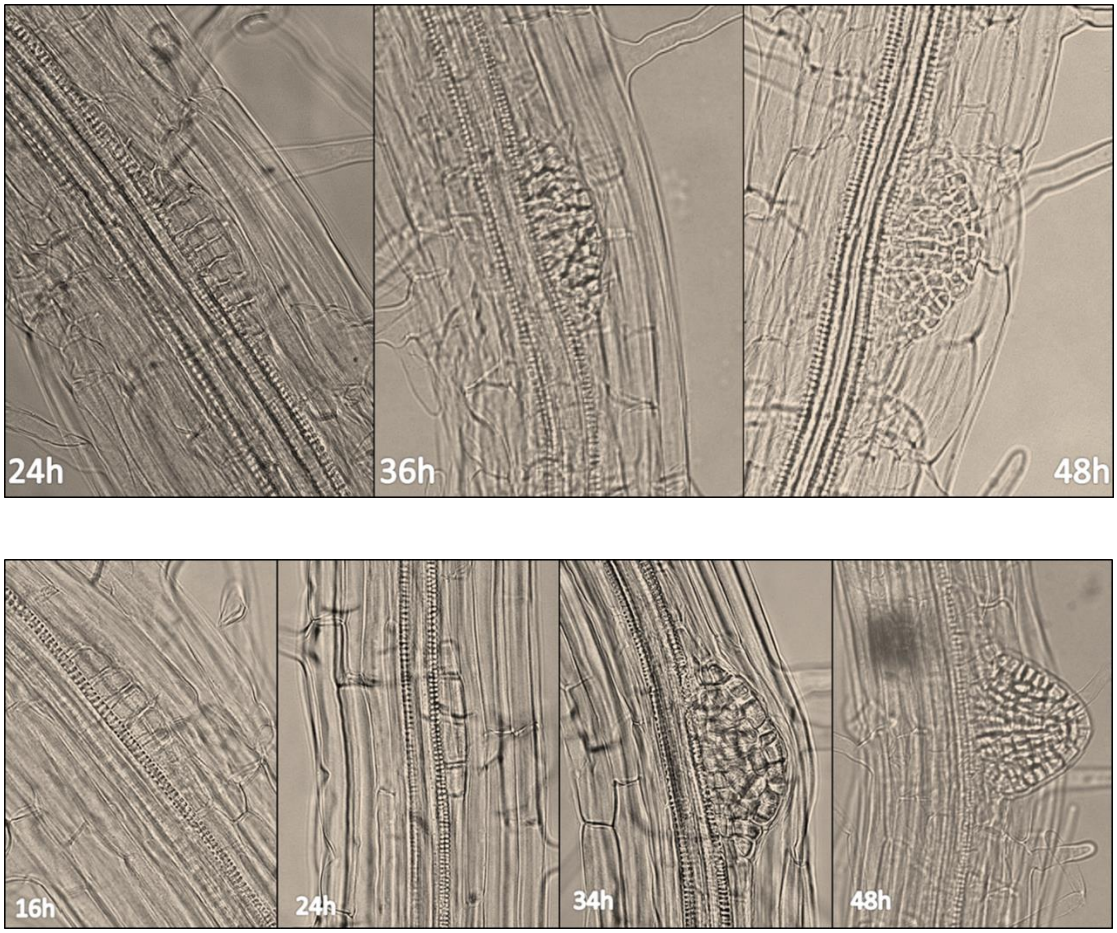


Figure 4.30: *MIM159 primordia (top) at various stages after initiation by a gravity stimulus compared to wild-type (Col-0), below. MiR159ab showed similar results*

This was further confirmed by a screen of the stages of developing LR primordia in 10 day old seedlings (figure 4.31.). Stages 1 to 5 were identified and any primordia at stage 6 up to emergence were identified and counted collectively. Counting the stages of developing primordia across a number of seedlings revealed that while there were significantly fewer in MIM159 and *miR159ab*, the spread of the various stages was roughly proportionate to that of wild-type (Col-0). This demonstrates that there seems to be no delay in LR progression in MIM159 or *miR159ab* but simply a lower number of primordia initiated.

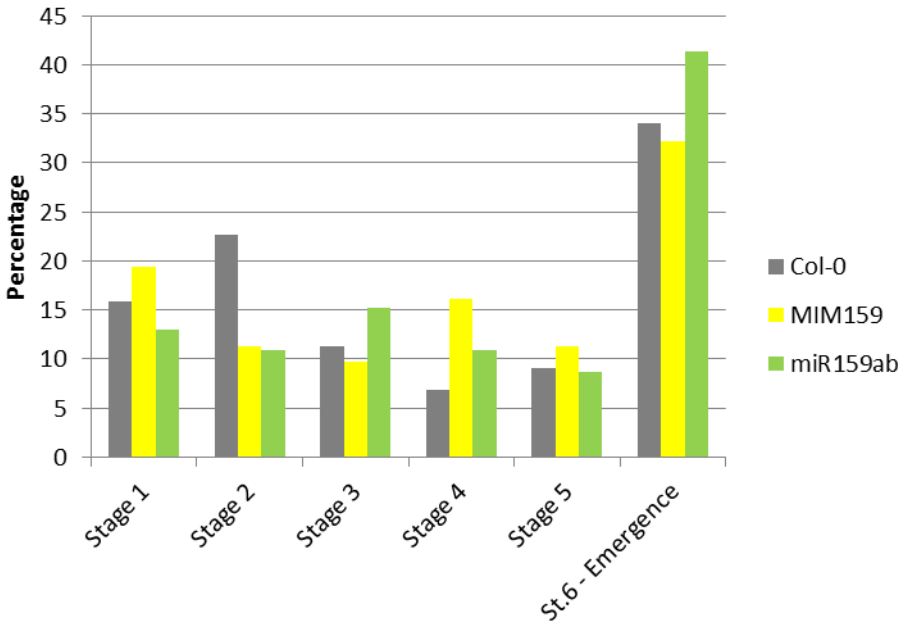


Figure 4.31: *Staging of LR primordia for MIM159 and miR159ab, compared against wild-type (Col-0). LR primordia were observed in 10 day-old seedlings (14 seedlings each line) and the primordia stages recorded (stages 1, 2, 3, 4, 5 and stage 6 to emerged primordia). Stages assigned are those described by Peret et al., 2009 (figure3.1.)*

MYB33 is clearly very strongly expressed throughout the plant but can also be very strongly repressed by miR159. While these results may prove limited so far that only suggestions can be made with regard whether this miRNA/target couple play an active role in LR development specifically; the dynamic and highly anti-correlated expression patterns do strongly hint at a possible role.

It appears that primordia morphology itself is not affected and that the strong phenotype seen in both MIM159 and *miR159ab* is a result of deficiencies upstream of the auxin pathway. This is supported by MIM159xAUX1:YFP expression which shows no change in

AUX1 expression. The sequencing and RT qPCR results show a dramatic rise in expression of MYB33 from 6 hours to 24 hours of lateral root development. This is complemented by miR159 expression profiles which drop over that time-course. This would suggest that repression of MYB33 could be necessary for proper LR initiation.

However, these results must be considered in the context of the other, broad roles of MYB33. Due to the ubiquitous expression seen throughout the plant of this gene and the many roles it has been linked with, the sheer strength of expression could lead one to jump to conclusions. While it is differentially expressed, this may-well be accounted for by something other than a role in LR development (in spite of the relative specificity of the datasets sequenced).

Further experimentation into whether the root phenotype seen when normal miR159 expression is disturbed is an independent phenotype or simply a consequence of a globally-compromised plant might consider a root grafting experiment whereby the aerial tissues of a wild type seedling are transferred onto the compromised root system of a MIR159 or *miR159ab* line. Recovery of the root phenotype may suggest the latter. Further MYB33 expression studies may also be considered.

4.5.4. MIR164 AND MIR160 HAVE PREVIOUSLY BEEN LINKED TO CENTERAL ROLES IN ROOT DEVELOPMENT

MiR164 is one of the microRNA families which has been most strongly linked to LR development through its repression of various members of the NAM/ATAF/CUC (NAC) domain transcription factor family (Guo *et al*., 2005) (section 1.2.3.1.). The five principal miR164 targets (NAC1, NAC80, NAC100, CUC1 and CUC2) were all found to be differentially-expressed in the RNASeq data. While the three NAC targets were found to be highly-expressed in the dataset, CUC1 and CUC2 were particularly-lowly expressed. Of the three NAC targets, only NAC100 showed strong expression profile correlation between RNASeq and Affymetrix datasets (PCC of 0.59) and a particularly dynamic expression pattern (figure 4.32.). Of the three miR164 family members (a, b and c), only miR164a and miR164b were identified in our sequencing data (both share a mature sequence whilst miR164c has a single nucleotide difference). Interestingly, unlike its targets, miR164ab does not seem to change much during lateral root development over the five time points.

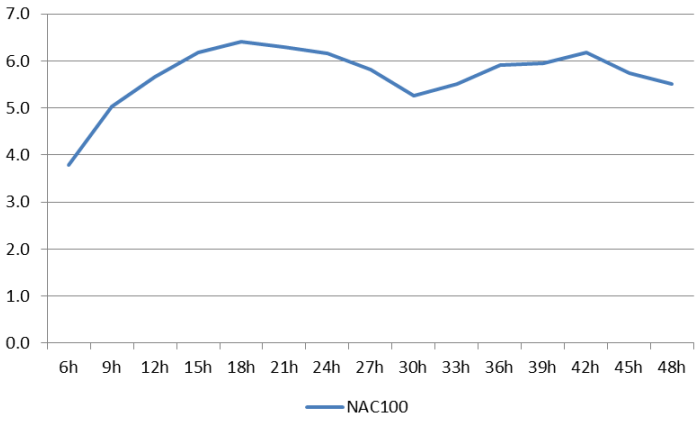


Figure 4.32: *Expression profile of NAC100 from averaged RNA sequencing and Affymetrix expression profiles from developing root primordia. Time-course is on the x-axis, y-axis represents normalised sequencing reads (log2 scale)*

For further validation of the expression profiles of three targets, RT qPCR was run for NAC1, NAC80, and NAC100 using NPA-NAA treated root systems on both the miR164 target mimic (MIM164) and wild-type (Col-0). A significant rise in expression in all three target-genes in MIM164 shows that the mimicry is working well and that miR164 is indeed repressing these three targets in Arabidopsis roots to a significant degree (figure 4.33.). RT qPCR also seems to show a rise in expression of all three targets as lateral roots develop.

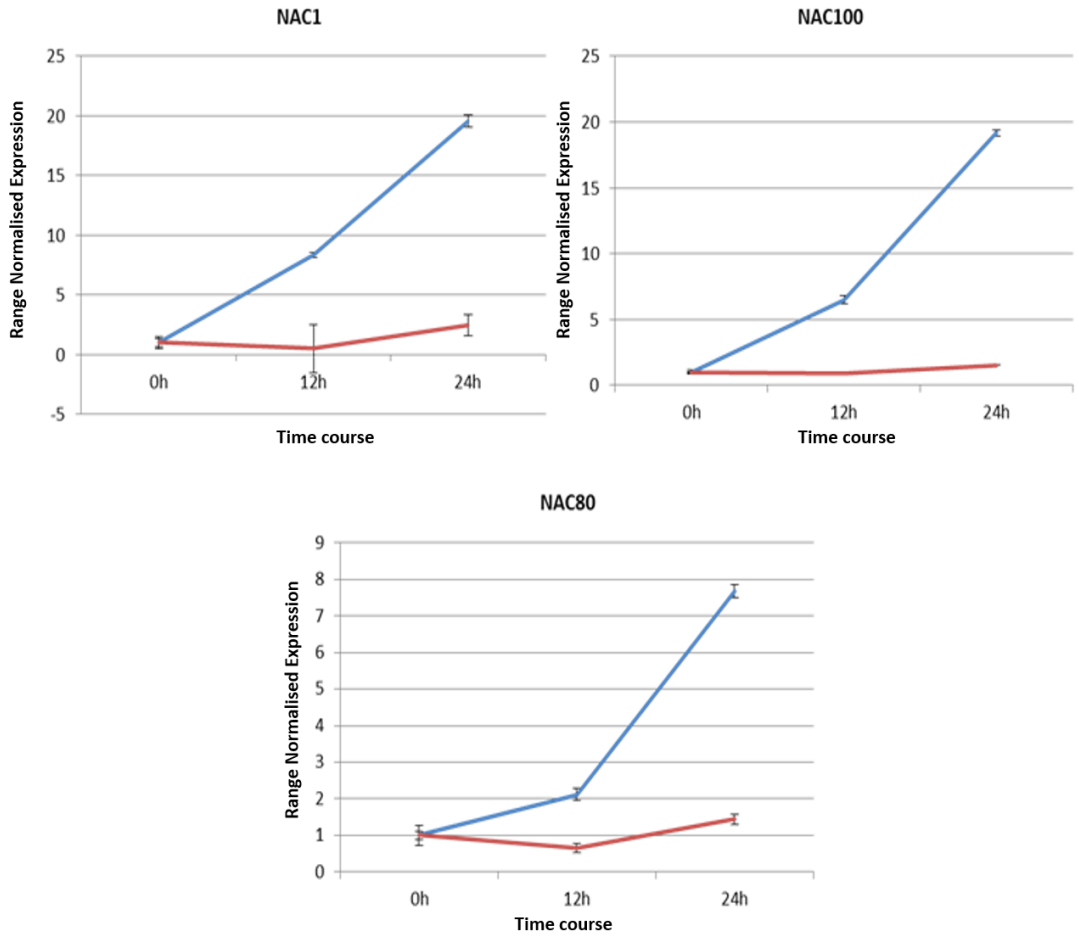


Figure 4.33: *RT qPCR expression of NAC1, NAC80, and NAC100 in 10-day-old MIM164 (blue) and Col-0 (red) root systems following NAA treatment (1 – 0h, 2 – 12h, 3 – 24h). Expression was normalised against actin. Error bars represent standard error*

Like miR164, miR160 has also been closely linked to lateral root development since it targets three auxin response factors (ARFS): ARF10, ARF16 and ARF17, which have all been found to play a role in Arabidopsis lateral root development (section 1.2.3.1.). The miR160 family also comprises three family members (a, b and c), all of which share identical mature sequences. Similarly to miR164, while this sequence was identified in the small RNA sequencing produced in this study, when comparing expression profiles between the two replicates

there is anti-correlation between the two. This is perhaps due to the relatively stable expression patterns of the microRNA which can lead to an over-representation of subtle differences.

Expression profiles for ARF10 and ARF16 correlated well between RNASeq and Affymetrix datasets and in both cases there seems to be a steady rise in expression over the time-course (figure 4.34.)

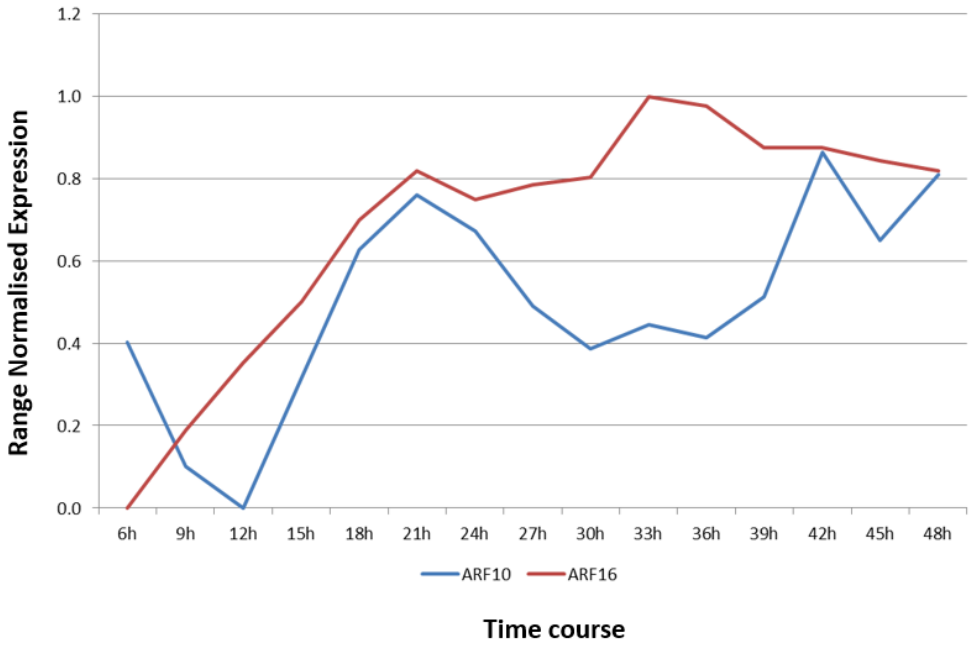


Figure 4.34: *Averaged, range normalised, expression profiles (from normalised counts) of ARF10 and ARF16 from RNA sequencing and Affymetrix datasets of developing root primordia*

4.5.5. EXPRESSION STUDIES ON MIR393 & ITS TARGETS REVEALS LITTLE TO SUGGEST A ROLE IN LR DEVELOPMENT

The miR393 family members, a and b, share an identical mature sequence and this sequence was found to be differentially expressed in our small RNA sequencing data although there was little correlation between replicates (with a PCC of -0.23). However, again, this might be due to the fact that, while the microRNA is quite highly expressed, its expression level remains very constant (and thus susceptible to a miss-representative PCC).

Some of the miR393 family's targets have revealed much more dynamic expression profiles; for example the well-defined targets; AFB2 and AFB3 (section 1.2.3.3.) (figure 4.35.).

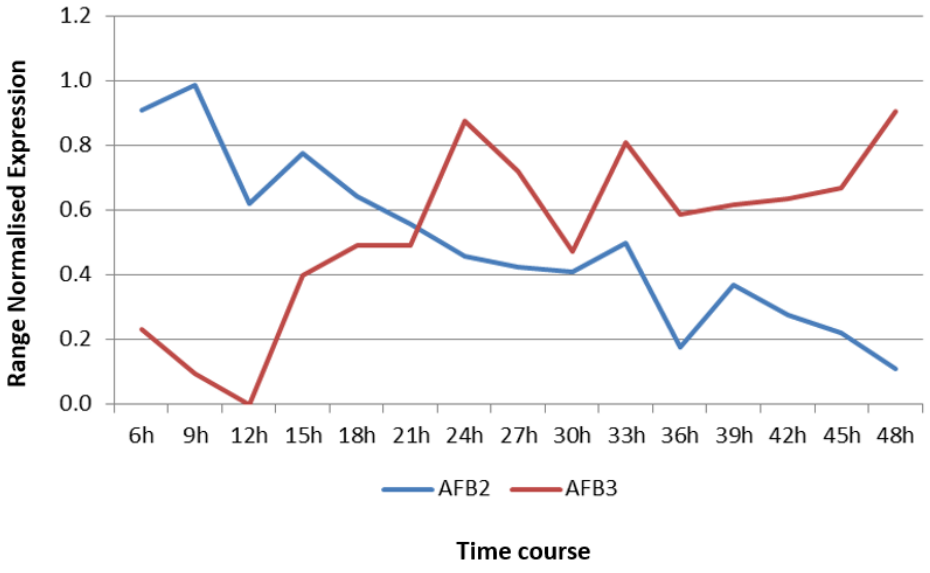


Figure 4.35: *Averaged, range normalised, expression profiles (from normalised counts) of AFB2 and AFB3 from RNA sequencing and Affymetrix of developing root primordia*

4.6. CHAPTER DISCUSSION

Only a small number of microRNAs have previously been linked to lateral root development in Arabidopsis, while a larger number have been linked to various developmental roles in the root system. This can perhaps be attributed to the difficulty in isolating a homogenous population of relatively pure lateral roots of any given age. This is due to the acropetal nature of LR development; there are mixed populations of LRs at different stages of development in any given plant. In recent years, methods have become available to obtain relatively homozygous populations of LRs of a particular stage (Peret *et al*., 2009). Using the gravi-stimulus method, RNA was extracted from relatively homogenous populations of LRs at different developmental stages and was then used for global gene expression profiling. This is the first report of a very detailed gene expression profiling including small RNA sequencing during LR development.

This study shows that several microRNA gene families are expressed during LR development. More well-known microRNA families such as miR164, miR160, and miR165/166 were found to be differentially expressed and in many cases, their root-related mRNA targets were found to be co-expressed during LR development. Encouragingly, very often, there was a negative correlation between microRNA expression and their targets suggesting a possible heightening or lowering of regulation of the target expression by the microRNA over the course of early LR development.

The quality of the data was established by a range of approaches including reciprocal relationships in expression between some known microRNAs and their targets over five time

points coupled with the correlation of expression profiles with the Affymetrix data (Voβ *et al*., 2015), the FOCL1 example (figure 3.11.), and the other validation outlined in chapter 3. These approaches show the data to be of very good quality and of great depth.

This study has identified several known microRNAs but their role in LR development is not clear.

The miR165/166 and miR156/157 microRNA families were found to be by far the highest-expressed of all of the families identified. One contributing reason for this might be due to the sheer number of family members within those groups and other studies have found the miR165/166 family in particular to often be amongst the highest-expressed in various parts of the plant. Both of these families had targets whose expression patterns were strongly anti-correlated with those of their microRNA regulators during LR development.

Encouragingly some well-defined targets show dramatic reciprocal relationships which hint at their possible role in lateral root development. This is also supported by the fact that some of these targets are often very highly expressed ubiquitously throughout the plant and are known to be involved in many, diverse functions. For example, PHV and PHB (targets of miR165/166, along with a number of others) are involved in root pattering, cell differentiation and the development of lateral root vasculature. It will be interesting to check the continuing expression of these genes and their corresponding microRNAs during post-emergence stages of LR development.

Similarly, the SPL targets of the miR156/157 family have been well defined. As described earlier in this chapter, one target which was identified in this study as a target of interest based on its expression profile (SPL9), was very recently found to be involved in LR initiation in an independent study published earlier this year (Yu *et al.,* 2015). This is really significant as it further validates the approaches taken in this study to identify potential targets.

Other targets of interest, showing high and strongly anti-correlated expression patterns with those of their microRNA regulators, include the mitosis checkpoint protein BUB3.2 (a target of miR158), the miRNA processing DCL1 (a target of miR162), and SULTR2,1 (which forms a feedback network with its regulator miR395). All three of these were found to be highly, and differentially, expressed and considering their expression profiles, they are highly likely to be regulated during LR development by their respective microRNAs. All three are also highly significant targets which are likely to play roles in various developmental processes and as such are strong candidates for a role in LR development.

Determining their precise role in LR development will be very interesting and would require a detailed phenotypic study using a variety of approaches including the use of materials such as microRNA target mimics. Investigation of the spatial expression patterns of these miRNAs and their targets during LR development using reporter based approaches would also be very valuable in determining their precise role in LR development.

Besides the targets discussed above, gene expression profiling revealed several genes showing strong differential expression during LR development such as the antiviral AGO2, the phosphate homeostasis protein PHO2, the chromomethylase CMT3, and the pathogen

defence-linked transcription factor WHIRLY3. However, their corresponding microRNAs (miR403, miR399, miR823, and miR840, respectively) were less-highly expressed. While these targets were intriguing due to their reported nature and their strong expression profiles, their relationships with their microRNA regulators during LR development were harder to ascertain.

Of all of the known microRNAs identified in this study, miR159 was the most intriguing with regards lateral root development. With two of its three members (miR159a and miR159c) differentially expressed in the data and one very highly expressed target (MYB33), attention rounded on this relationship and expression studies were continued. The *miR159ab* mutant and miR159 target mimic (MIM159) show dramatic root phenotypes and expression profiles of MYB33 suggest that it is strongly regulated by this microRNA family in the root. However, while root numbers are reduced, there does not seem to be an effect on the morphology of the primordium itself, nor a delay at a particular stage of development. This would suggest that miR159 acts to repress MYB33 in order to initiate lateral root development, and should the microRNA be perturbed, then lateral root numbers drop.

However, it can be argued that the LR phenotype of miR*159ab* mutants and MIM159 is a consequence of stunted aerial tissue growth due to a lack of resources available for root development. There are many experiments which can be done to test this including tissue grafting experiments. Further studies into the effects of exogenous auxin or gibberellic acid on expression of miR159 and its targets may also provide more details into any potential role in LR development.

One of the biggest difficulties with this strand of the study was with the production of target mimics. A large number of microRNAs were selected for target mimic production but ultimately only a small few could be produced due to a range of severe plant growth issues. Furthermore, only one line was produced for each successful mimic, rendering any insights gained limited to one biological replicate. It is not surprising considering the diverse role of microRNAs in plant growth and development and the potentially high number of targets that they can regulate. While other options are available, perhaps using other techniques, such as short tandem target mimics (STTMs) (Tang *et al*., 2012), might end up with similar issues. MicroRNA mutants may be a good option but there are only a very limited number of microRNA mutants available to date. Considering several microRNAs are part of gene families, it can also be technically very challenging to create multiple mutant combinations. Genome editing technologies are now well developed and may prove extremely useful but their use in Arabidopsis is currently very limited and has met with mixed success.

The use of miRNA biogenesis pathway mutants (such as those for DCL1 and HYL1) could also be used to study further the specifics of repression of specific targets by miRNAs during LR development but that will be an indirect approach and will not pinpoint the role of individual genes on lateral root development.

This chapter has investigated the role of known microRNAs in lateral root development using a very robust and logical pipeline and has identified several microRNA/target couples that require further investigation for their role in LR development to be clarified. However, as with any global gene expression techniques, it cannot be ruled out that some potentially interesting microRNA/target couples may have been overlooked.

CHAPTER 5
NOVEL MICRORNAS

5.1. SEVERAL POTENTIAL NOVEL MIRNAS WERE IDENTIFIED BY SMALL RNA SEQUENCING

Small RNA sequencing during different stages of lateral root development identified several novel reads. Based on the presence of these reads in both small RNASeq replicates, a predicted hairpin structure (miRDeep, Friedlander *et al*., 2008), and differential expression profiles; we report the identification of ten novel microRNAs. For these novel miRNAs; hairpins were predicted and visualised using both RNAfold (Lorenz *et al*., 2011) (figure 5.1.) and the UEA small RNA workbench (Stocks *et al*., 2012) (appendix A.5.). Besides these, six additional novel miRNA can also be identified based on their presence in only one of the small RNASeq replicate (*table 5.1*.).

A number of other shorter sequences were also identified as potential microRNAs but were deemed too short to provide confidence in their prediction. This chapter focuses on the ten microRNAs confidently identified in both replicates while also considering the other six strong candidates which were identified in just one. [5]

[5] A note on novel miRNA names: Novel microRNAs are referred to here by arbitrary numbers assigned to them during sequence analysis (table 5.1). These numbers are used throughout this study but would be changed should the miRNAs be submitted to MIRBASE.

Novel miRNA	Replicate	Mature sequence	Strand	Chromosome	Precursor co-ordinates
New-miR11	A & B	AACCGATGCGCTGCGTACTGCT	-	1	17396144-17396291
New- miR19	A & B	CCCGAGGTGTGTGGATCAGAA	-	1	30227829-30227895
New- miR29	A & B	TAGTGGAAGCAGCAACGAGAA	-	1	23355103-23355188
New- miR35	A & B	GGGGATTGACATGAATGACACT	+	2	3013112 - 3013262
New- miR36	A & B	CGATCCCCGGCAACGGCGCCA	+	2	3740828-3740958
New- miR39	A & B	ATATAATGGACTGAGACCGCCT	+	2	1111029-1111130
New- miR78	A & B	TTGTGTTGCGTTTCTGTTGATT	+	3	22678167-22678268
New- miR117	A & B	TCGGACGCTGACTGGCTTTTCA	-	5	7146675-7146751
New- miR120	A & B	AATCGCACATCTGTTAGGAGG	-	5	5590304-5590444
New- miR141	A & B	GAGTGGAGCGCGAATAGCACT	+	5	1913361-1913466
New- miR97	A	CGACTCTGGTCCCTTTGGCAC	-	4	18017274-18017339
New- miR 110	A	CGCTACAAGAGTTTTAGTCGAG	-	4	6539722-6539791
New- miR144	A	CCGAAGCTGGTAGTACTTGGCT	+	5	18397816-18397903
New- miR61	B	ATGTGAGGAAGACTGTTGAAG	-	3	10960249-10960315
New- miR62	B	ATAAGTCAAGTAGTAGTCCTTG	-	1	10445474-10445556
New- miR88	B	ATGTCACGGCTGAGTTATATGC	+	4	3817468-3817529

Table 5.1: *Potential novel miRNAs identified by small-RNA sequencing. Full precursor sequences can be found in appendix A.5.*

.

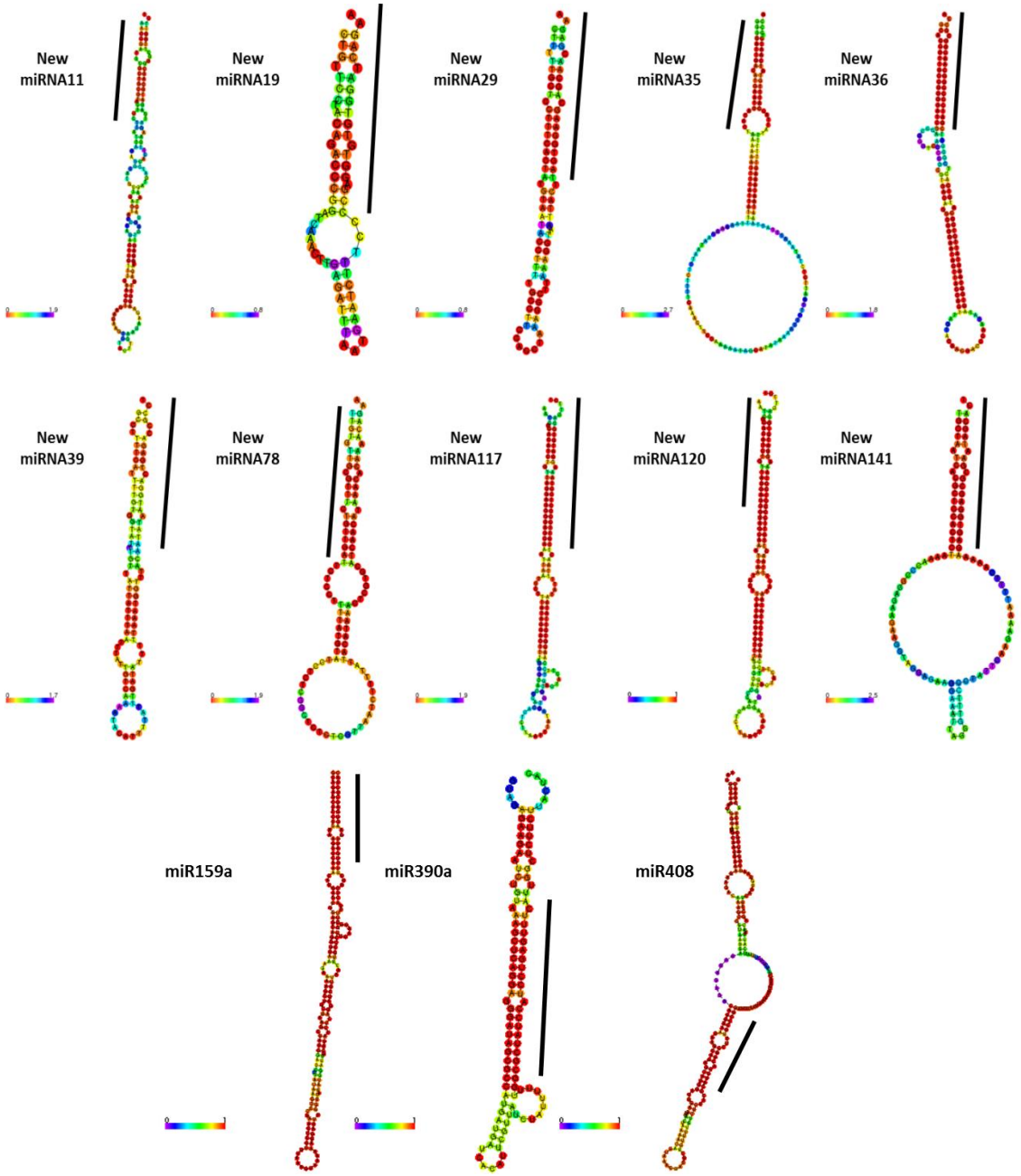


Figure 5.1: *Predicted, novel miRNA precursor hairpins, developed using RNAfold (Lorenz et al., 2011). Colours indicate the entropy of the sequence from red (low) to purple (high). Black bars indicate the mature microRNA location. The bottom three precursors show well-known, conserved microRNAs, for the purposes of comparison.*

5.1.1. NOVEL MIRNAS ARE LOCATED IN EXONS, INTRONS AND INTERGENIC REGIONS

Three of the ten novel miRNAs identified in both replicates (35, 36 and 39) were predicted to be located in intergenic regions of the genome while three more were found in the 3' UTR of a gene (19, 29 and 141 were found in the genes AT1G80410, AT1G63020 and AT5G06270, respectively). Two more were found in the 5' UTR (78 and 117 found in the genes of AT3G61270 and AT5G21040 respectively). On top of this, one new miRNA (120) was found within the intron of a protein-coding gene (AT5G17010) and another (11) was found in the exon of another (AT1G47420).

5.1.1.1. NOVEL MIRNA11 IS LOCATED WITHIN AN EXON

The new miRNA (new miRNA11) is located within the exon of a protein-coding gene AT1G47420 and could be potentially very interesting as, for microRNAs which have been defined; this is a relatively rare occurrence. Due to the nature of this microRNA, it is likely that microRNA111 regulates the expression of AT1G47420.

The gene AT1G47420 itself codes for the protein SUCCINATE DEHYDROGENASE 5 (SDH5), a subunit of a mitochondrial complex with a role in the respiratory chain. The expression profile of this gene from mRNASeq correlates relatively strongly with that from the Affymetrix dataset, with a PCC of 0.46. Both show a very stable, un-dynamic profile over the time-course. With regards the potential, novel microRNA, the two replicates for new miRNA11 are less well correlated together (PCC of -0.2). It is possible that the lack of high

correlation may be due to very low expression of this miRNA and subtle changes may be over-represented by the Pearson correlation scoring system. However, unlike SDH5, both miRNA replicates do seem to show quite dynamic expression patterns (figure 5.2.).

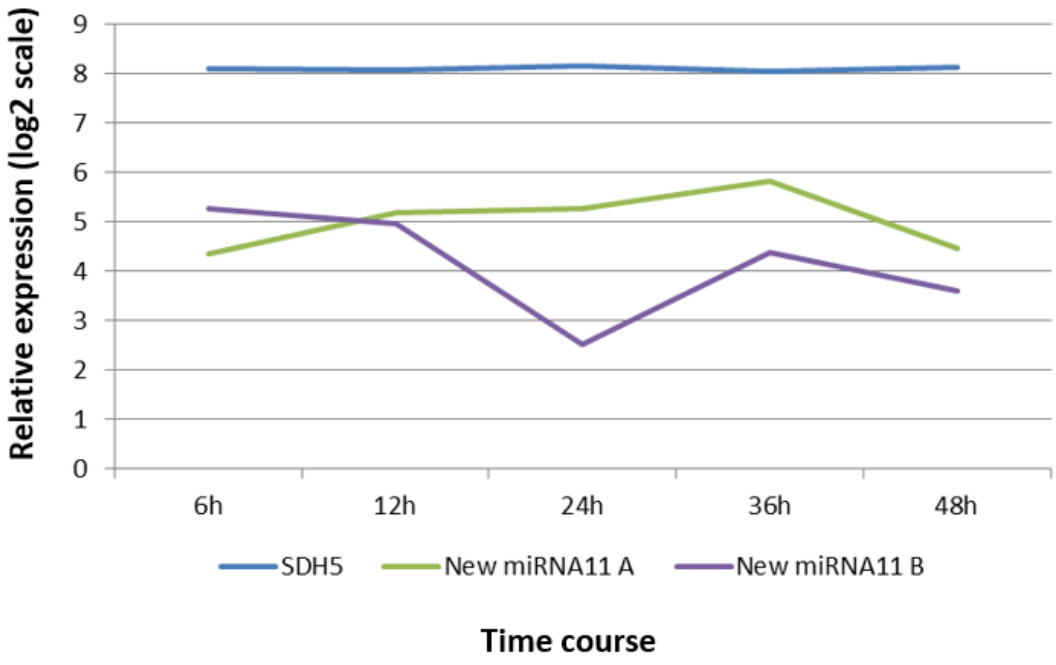


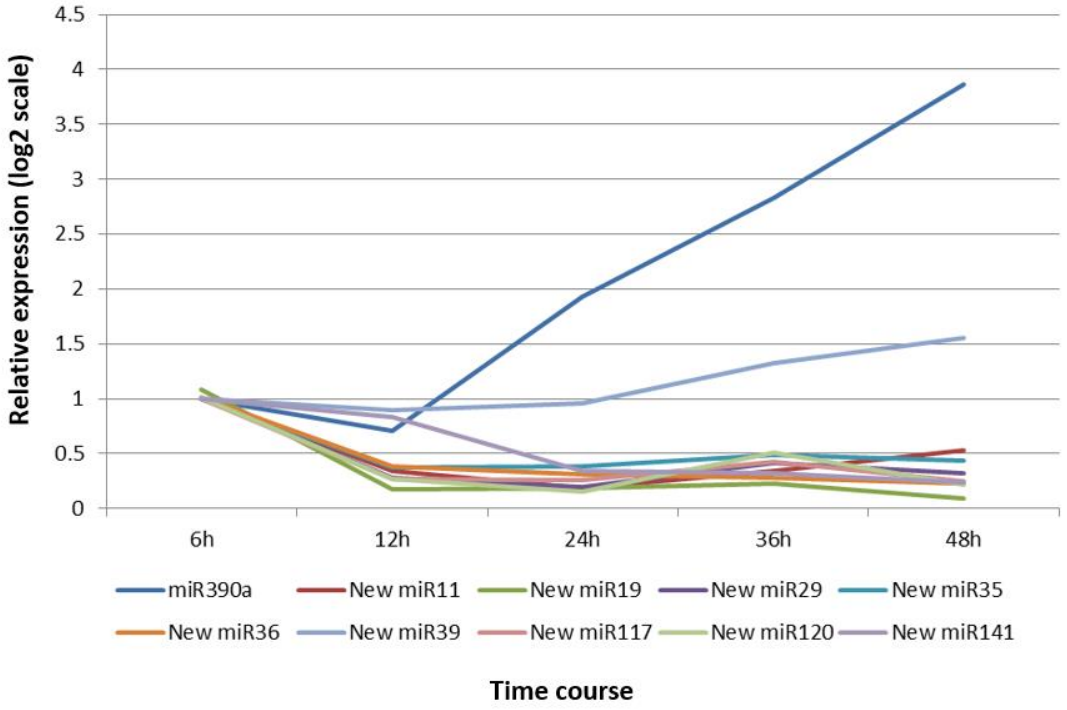
Figure 5.2: Relative *RNASeq expression profiles of SDH5 and potential, novel microRNA 11 (both replicates) during lateral root development (log2 scale from read counts)*

5.1.2. VALIDATION OF NOVEL MICRORNAS BY RT QPCR

Stem-loop RT qPCR (Varkonyi-Gasic *et al.,* 2007) was used for validation of novel microRNAs. This protocol is well established and which implements reverse transcription and PCR

primers which form a stem-loop structure and can be used to detect and quantify even lowly-expressed microRNAs (Varkonyi-Gasic *et al.,* 2007).

Primers were designed for all ten novel microRNAs and RT qPCR undertaken on the original replicates sent for sequencing following the sample preparation protocol of Varkonyi-Gasic *et al*. (2007). Samples were run using the Light Cycler protocol detailed in Chapter 2 alongside miR390a as a control. All ten novel microRNAs were detected with the exception of new miRNA78 (figure 5.3.). However, their expression levels appears to be very low compared to miR390a.



Sample	6h	12h	24h	36h	48h		6h	12h	24h	36h	48h
miR390a	1.003	0.710	1.933	2.832	3.866	**New miR36**	1.007	0.380	0.309	0.279	0.221
New miR11	1.002	0.344	0.180	0.338	0.524	**New miR39**	1.000	0.893	0.955	1.324	1.555
New miR19	1.079	0.174	0.184	0.228	0.087	**New miR117**	1.000	0.273	0.260	0.427	0.248
New miR29	1.012	0.276	0.198	0.415	0.324	**New miR120**	1.010	0.272	0.149	0.505	0.214
New miR35	1.005	0.371	0.382	0.488	0.432	**New miR141**	1.003	0.832	0.338	0.316	0.232

Figure 5.3: *RT qPCR detection of novel microRNA on samples used for initial small RNA sequencing (miR390a control). None of the ten novel microRNAs identified in both replicates could be detected. Values are relative to the 6h time-point of miR390a. Expression was normalised against the U6 small RNA.*

5.2. NOVEL MICRORNAS REVEAL STRONG AND VARIED SMALL RNASEQ EXPRESSION PROFILES

Focusing on the ten novel microRNAs identified in both replicates; there is great variation between the expression strengths and patterns of the different microRNAs. It is notable that on the whole, the expression of the novel miRNAs is much less than that of the known, and particularly the well-defined, miRNA families. This is not unexpected if they were not identified earlier. While the number of reads of some of these novel miRNAs is quite low, as they were identified in both replicates and their expression could be verified independently by RT qPCR, they are likely to be genuine miRNAs.

As mentioned previously, the lower the expression strength, the less significant the correlation coefficient value becomes so in the case of many of these novel miRNA, there is little correlation between the two replicates. In three cases (19, 36 and 78), there is positive correlation with a value of more than 0.6. Five more of the novel miRNA also show a positive correlation between replicates, albeit to a lower degree (29, 35, 117, 120 and 141). The other two novel miRNAs found in both replicates (11 and 39) show anti-correlation between replicates A and B.

Of these novel microRNAs, new-miRNA36 is by far the most highly expressed and has the highest correlation coefficient between replicates (PCC = 0.94). This is followed by new-miRNA39 though this is the novel miRNA with the highest anti-correlation between replicates. These two miRNAs, particularly new miRNA36, are quite strikingly more-highly expressed than the other eight.

New miRNA36 shows a steep drop in expression over the duration of the primordium development. The other two novel miRNA showing a positive correlation between replicates show this same pattern (figure 5.4.). Other notable expression patterns include new miRNA117 and new miRNA29 (PCC of 0.33 & 0.4, respectfully) both of which show a distinctive rise in expression towards the end of the time course, hinting at a possible role in lateral root emergence (figure 5.5.) (see appendix A.8. for full expression patterns of novel miRNA).

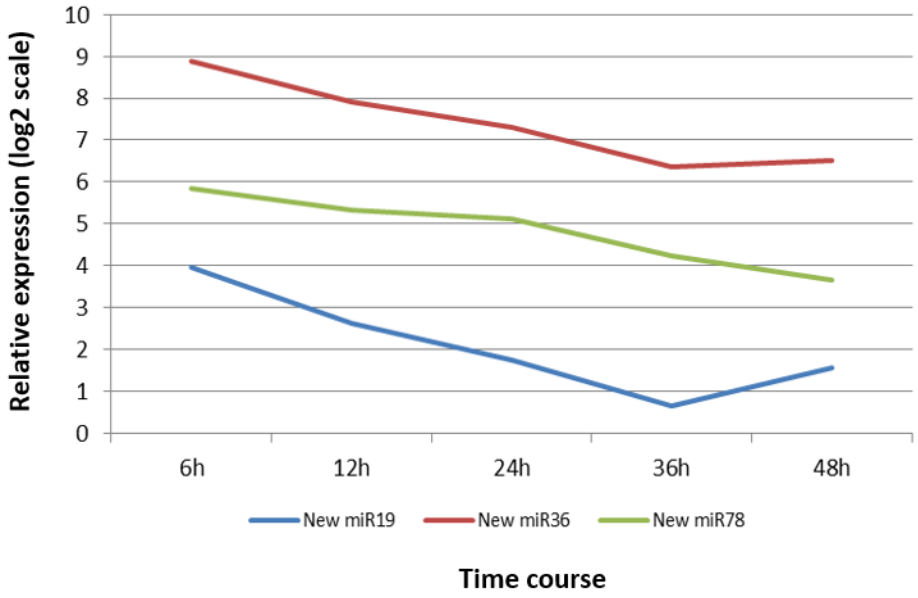


Figure 5.4: *Averaged expression profiles of novel microRNA with a PCC between replicates of 0.6 or above (log2 scale). All expression drops up to 36 hours before rising again in two cases.*

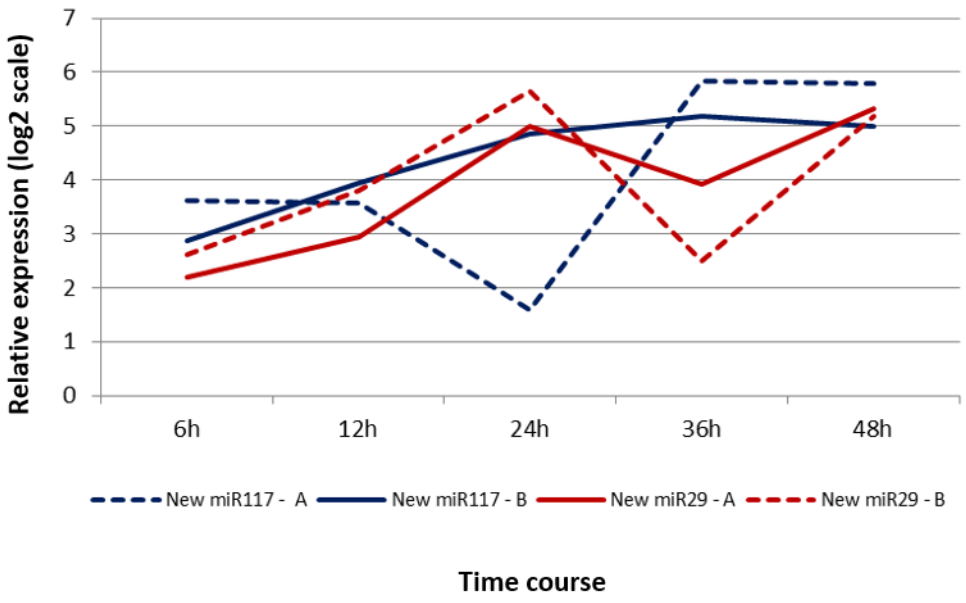


Figure 5.5: *Expression profiles of new miR11 and new miR29 from RNA sequencing of developing lateral root primordia (both replicates) (log2 scale)*

5.3. CONSERVATION OF NOVEL MICRORNAS

Globally, mature microRNAs are generally well conserved, often with just one or two nucleotide differences in the ~ 21 nucleotide mature sequences; highlighting their functional importance. However, the stem-loop precursor appears to show less conservation presumably due to less selection pressure.

5.3.1. CONSERVATION IN ARABIDOPSIS

The mature sequences and predicted precursor sequences of the 16 novel microRNAs were compared with all of the sequenced Arabidopsis accessions (Salk Arabidopsis 1,001 Genomes database, Gan *et al*., 2011). Extended precursor sequences (1.5kb upstream and 1 kb downstream) were also compared.

This analysis revealed that the ten novel microRNAs found in both replicates were, on the whole, very well conserved throughout all of accessions in the database. Interestingly, precursor sequences were also very well conserved (figure 5.6.). Most notable is new miRNA19 whose entire precursor was perfectly conserved across all accessions (table 5.2.). Also, new miRNA141, new miRNA120, new miRNA78, and new miRNA29 are all very well conserved with the equivalent of fewer than 7 SNP locations per hundred across all accessions (table 5.2.). By some distance, the least well conserved microRNA (with regards both its mature sequence and precursor sequence), is also the most highly expressed. New miR36 has the equivalent of 28.48 SNPs per 100nt on its precursor sequence, a number which is far higher than most conserved microRNAs (table 5.2.).

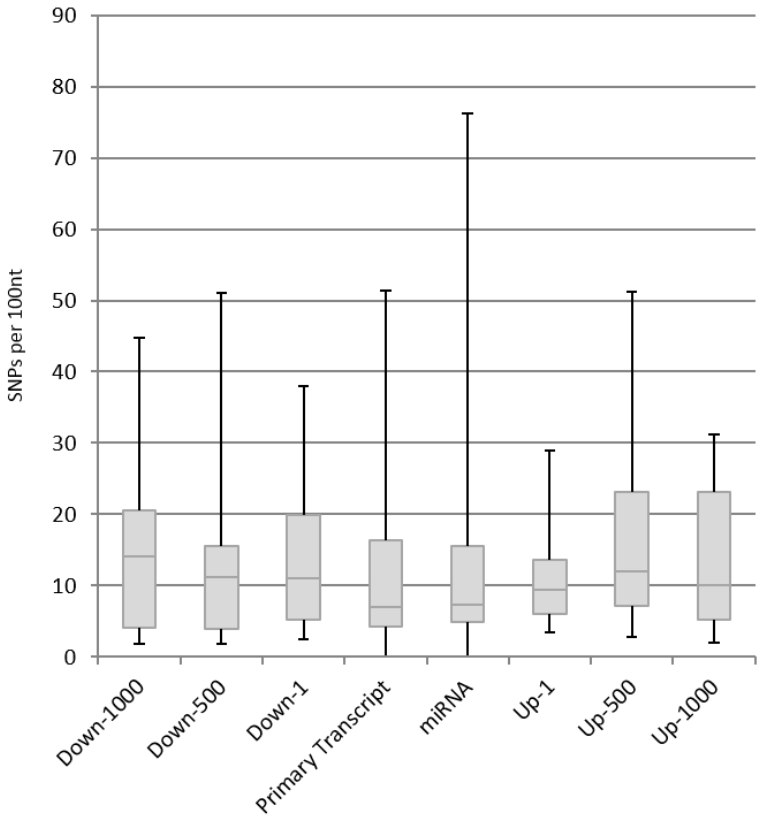


Figure 5.6: *Conservation of the 16 predicted novel microRNA locations across sequenced genomes in the SALK 1,001 Genomes database (Gan et al., 2011- Accessed 2nd April 2015). Box plots for the number of SNPs (y-axis: SNPs/100nt) at predicted miRNA precursor (primary transcript) and mature miRNA loci. Up-1000 refers to the upstream region from 1,500 to 1,000, up-500 refers to the upstream region from -999 to 500, and up-1, refers to the upstream region from -499 to -1. The same nomenclature applies to the downstream (down) regions. Grey boxes represent 50% of the data and the strike through them represents the median confidence interval. Whiskers indicate the upper and lower quartiles. A comparison graph of all of Arabidopsis microRNAs registered on MIRBASE can be found in appendix A.6.*

Novel miRNA	Mature SNPs	SNPs per 100 nt (Mature)	Precursor SNPs	SNPs per 100 nt (Precursor)	Location in Gene
New miR11	3	14.28	15	10.2	Exon
New miR19	0	0	0	0	3’ UTR
New miR29	1	5	5	5.88	3’UTR
New miR35	16	76.19	77	51.33	Intergenic
New miR36	8	40	37	28.46	Intergenic
New miR39	4	19.04	28	27.72	Intergenic
New miR78	1	5	4	3.96	5’ UTR
New miR117	1	4.76	9	11.84	5’ UTR
New miR120	1	5	9	6.42	Intron
New miR141	1	5	2	1.9	3’ UTR

Table 5.2: *Conservation of the 10 predicted novel microRNAs (found in both replicates) locations across sequenced genomes in the SALK 1,001 Genomes database (Gan et al., 2011 - Accessed 2nd April 2015).*

From this analysis, with regards conservation across accessions, most of the novel microRNAs identified are entirely comparable to those registered on MIRBASE; with the possible exception of new miR35, whose mature sequence in-particular is not well conserved at-all. Nonetheless, despite this being an outlier amongst the novel microRNAs, there are still comparably conserved miRNAs in the MIRBASE database.

5.3.2. CONSERVATION ACROSS SPECIES

Only one of the ten microRNAs identified in both replicates has a sequence which was identified in any other species. A routine nucleotide BLAST search (NCBI) revealed that the predicted precursor sequence of new miR36 (interestingly, the least conserved across Arabidopsis accessions, as well as the most highly expressed) is also found in *Medicago truncatula*, tomato, and species of melon, amongst other plant species; all with a high degree of conservation.

5.4. INVESTIGATION INTO THE ROLES OF NOVEL MICRORNAS IN LR DEVELOPMENT

5.4.1. EXPRESSION PROFILING OF NOVEL MICRORNAS

In order to investigate further whether some of these novel microRNAs may have a role in lateral root development, RT qPCR was carried out. However, this time a different set of samples was used.

5.4.1.1. LATERAL ROOT INDUCTION USING NPA/NAA TREATMENT

Producing a suitable amount of RNA for RT qPCR using the gravitropism-based method applied to obtain sequencing samples requires a lot of time and resources as fairly little RNA

can be obtained simply from developing root primordia. Therefore, in order to obtain representative samples (albeit slightly less refined), NPA/NAA treatment was used.

Auxin transport inhibitor NPA (1-N-Naphthylphthalamic acid) arrests lateral root initiation (Himanen *et al*., 2002) whereas auxin analog NAA (1-Naphthaleneacetic acid) promotes lateral root initiation (Casimiro *et al.,* 2001). They can be used in conjunction to synchronise lateral root development. The rationale is that seedlings are initially grown on NPA that blocks LR formation and then when the seedlings are transferred to NAA, they can be primed almost synchronously to promote LR development. The process appears to be more rapid than the gravistimulus method, initiating a large number of lateral roots and the primary root will contain multiple LR primordia at roughly the same stage (figure 5.7., figure 3.1.)

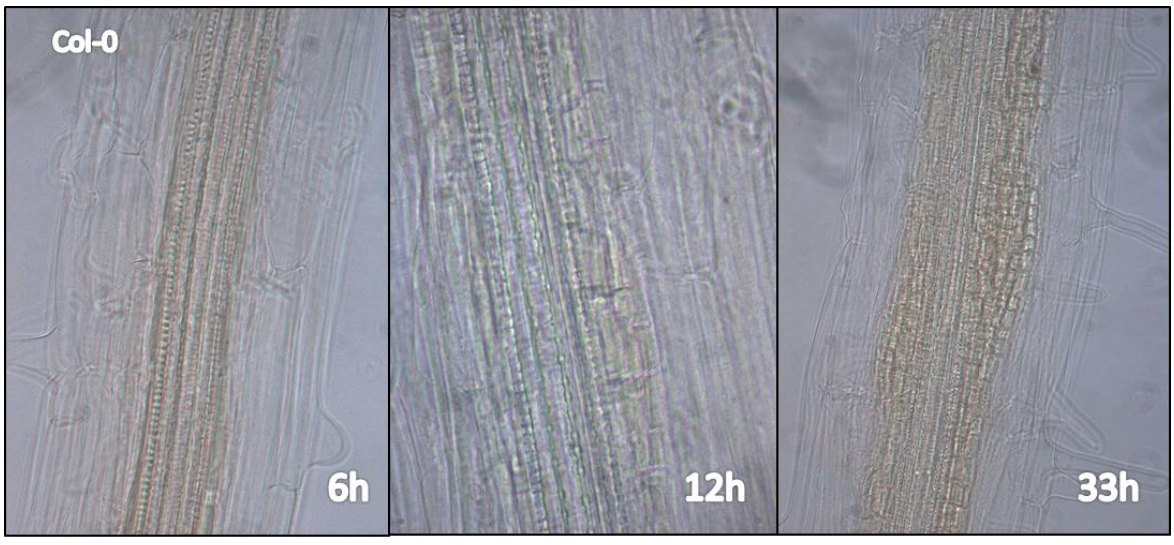


Figure 5.7: *Developing wild-type (Col-0) LR primordia at certain time-points after treatment on NAA (following Germination on 1/2 MS (5 days) and growth on NPA (3 days)*

Expression profiles of novel microRNAs during lateral root development were also further validated using this technique.

Seedlings were germinated and grown on half MS media for 5 days before being transferred to plates containing 10µM NPA and grown for a further 3 days in the same conditions. Seedlings were then transferred to plates containing 10µM NAA). Whole root systems were then excised from seedlings at various time-points (0 (time of transfer), 3, 6, 9, 12, 15, 18, 21, and 24 hours) and stored at -80°C prior to RNA extraction, DNAse treatment, and subsequent RT qPCR.

5.4.1.2. RT QCPR ANALYSIS OF EXPRESSION OF NOVEL MICRORNAS DURING LR DEVELOPMENT

RT qPCR was carried out (as in section 5.1.2.) and this time all 10 novel microRNAs were identified. New miRNA36 and new miNRA141 were found to be particularly abundant while new miRNAs 19, 39 and 78 (which were not identified in the previous round of RT qPCR) are particularly weak.

All ten novel microRNAs show differential expression and there is a wide-range of profiles between them (figure 5.8.). However, there is little obvious correlation between these profiles and those produced from the small RNASeq (figure 5.9.)

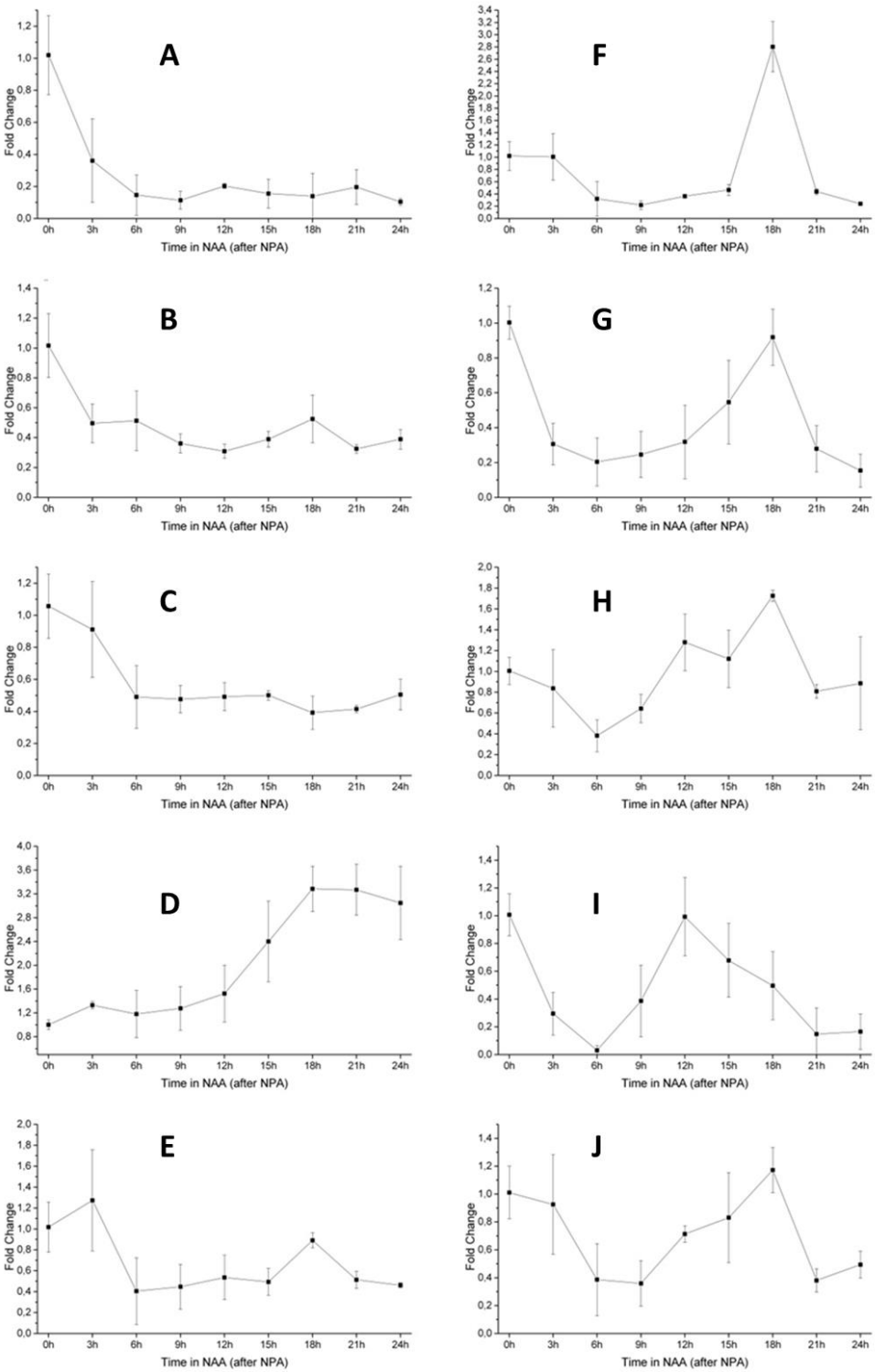


Figure 5.8: *Expression profiles of novel miRNA on NPA/NAA treated whole root samples. Bars represent standard error. Time is represented on the x-axis, fold change on the y-axis (A-J are New miRNA 11, 19, 29, 35, 36, 39, 78, 117, 120, and 141, respectively.*

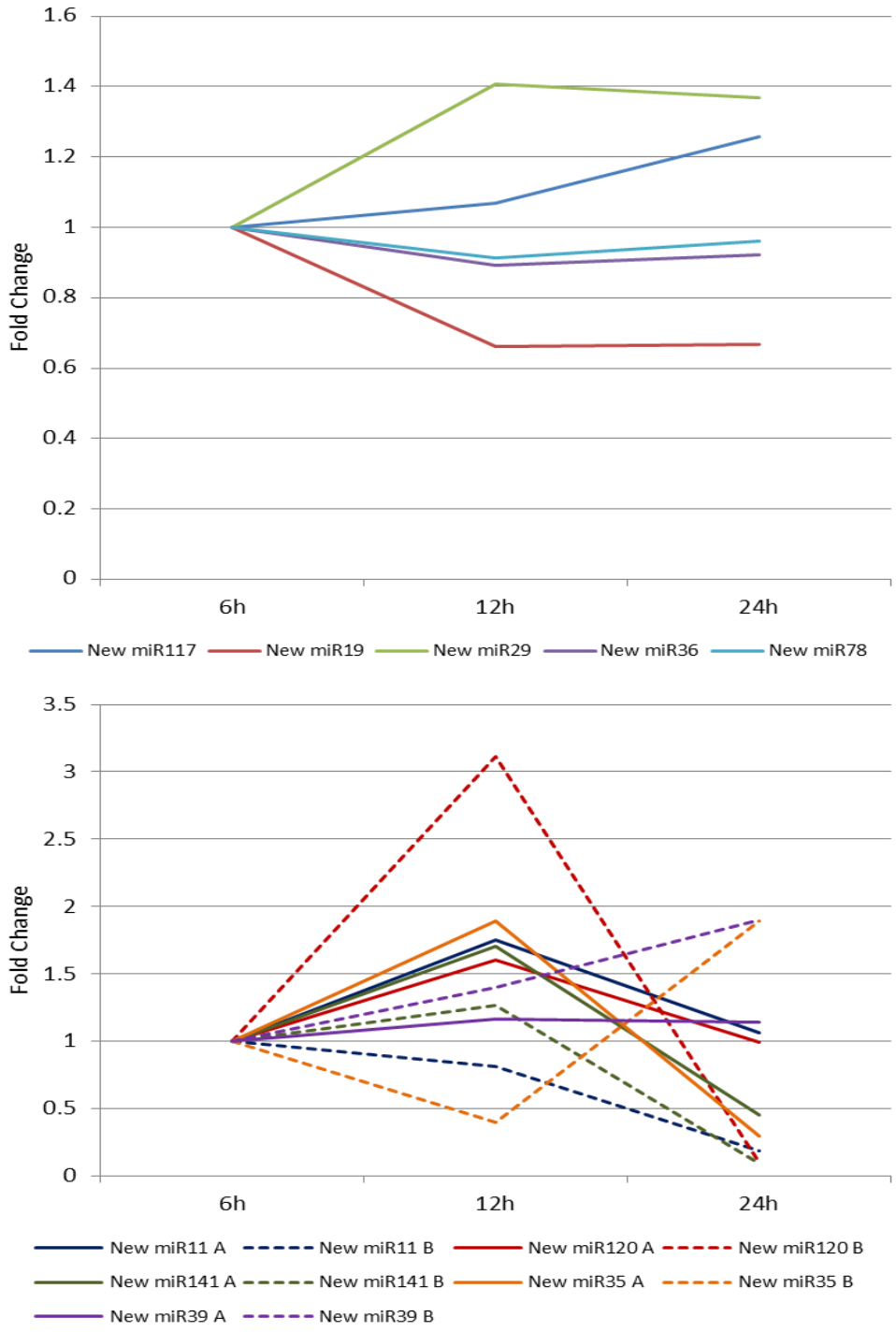


Figure 5.9: *Fold change values of novel miRNA expression (6-24 hours) from small RNASeq data for miRNA with PCC scores of less than 0.4 (left - two replicates shown) and 0.4 or above (right, averaged profiles shown)*

5.5. INVESTIGATING THE ROLE OF NOVEL MICRORNAS IN PLANT DEVELOPMDENT

To test the function of the novel microRNAs in plant development in general and LR development in particular, overexpression lines were generated for some of the novel microRNAs.

5.5.1. GENE CLONING

The whole microRNA precursor sequences (including a few hundred-base-pair of 3' downstream sequence to include any termination sequence - details in appendix A.7.) were PCR amplified and cloned into a Gateway entry vector; PCR8® (Invitrogen). The entry vectors were then checked by restriction and subsequently by sequencing. Good entry vectors were then recombined into a Gateway destination vector, pGWB8 (Nakagawa *et al*., 2007), containing the Cauliflower Mosaic virus 35S promoter. Cloning was validated by PCR. Eight of the ten microRNA constructs were successful and were then electroporated into Agrobacterium strain C58 for plant transformation.

5.5.1.1. SELECTION OF TRANSGENIC PLANTS

T1 seeds were grown on Kanamycin plates and several primary transformants were selected, genotyped and transferred to a green house. Successful transformants were also checked for expression using RT-qPCR. Surprisingly, neither overexpression of either mature microRNA or precursor could be detected in these plants.

There are many possibilities as to why these lines did not overexpress properly. There may be problems with processing the precursor; for example: it was thought that perhaps not enough, or too much, sequence was taken either side of the precursor which included signals which altered processing. However, this was subsequently disproved and since all eight were shown to be defective, it seems far more likely that the problem was not with the microRNAs themselves, but with the vector used.

As a result, a different cloning vector, pH7WG2,0 (Karimi *et al*., 2002), was used this time and as a proof of concept, three microRNAs were selected for the production of new overexpression lines (New microRNA11, 19 and 39). New miR11 and new miR39 were two of the most highly-expressed microRNAs whereas new miR19 is the best-conserved within Arabidopsis accessions.

The same region of precursor as described above was PCR amplified and was cloned into Gateway entry vector pENTR/D_TOPO® (Invitrogen). The cloning was verified as before and then correct entry vectors were used for LR reaction using Gateway destination vector pH7WG2,0 (Karimi *et al*., 2002, figure 5.10.).

Vectors were transformed into *Arabidopsis thaliana* and selected on the basis of Hygromycin resistance. Several primary transformants were selected for new miRNA 39 and 19 but none could be selected for new miRNA11. The transgenic seedlings were taken to T2 generation where they were genotyped and overexpression was checked using RT qPCR.

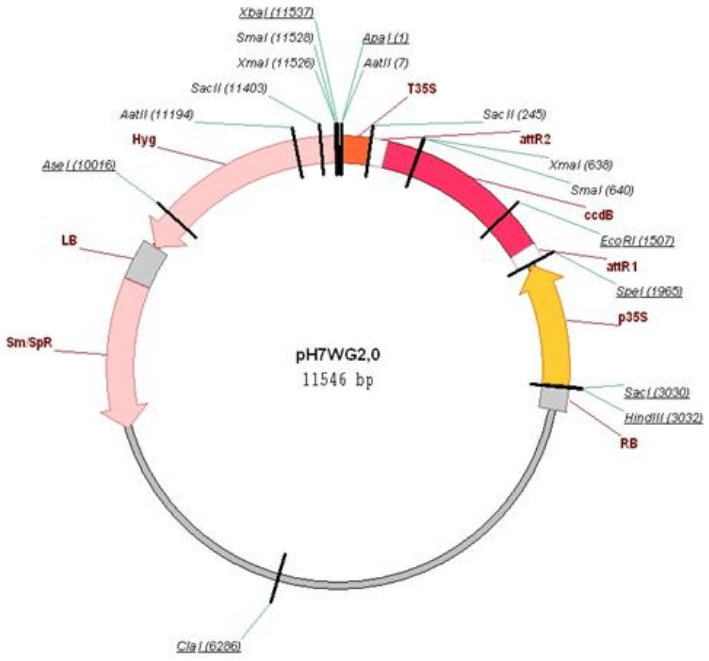


Figure 5.10: *Plasmid map of pH7WG2,0 (Adapted from Karimi et al., 2002)*

5.5.2. RT qPCR REVEALS OVEREXPRESSION IN NEW MIRNA19 AND NEW MIRNA39 LINES

Of ten independent T2 plants selected for the overexpressing lines for new miRNA19 and new miRNA39 (henceforth: OXP19 and OXP39, respectively), eight were genotyped by PCR each and in each case, seven were shown to be positive for the 35S promoter preceding the novel miRNA precursor. Of these, four separate plants each were selected for RT qPCR.

RT qPCR was performed on a specific region of the precursor between the 35S promoter and the mature sequence of the microRNA in each case. All four new miRNA19 overexpression

lines were found to be very strongly overexpressing this region when compared to wild-type while three of the new miRNA36 lines were found to be expressing stronger than wild-type, although this was not found to be a significant difference (figure 5.11.)

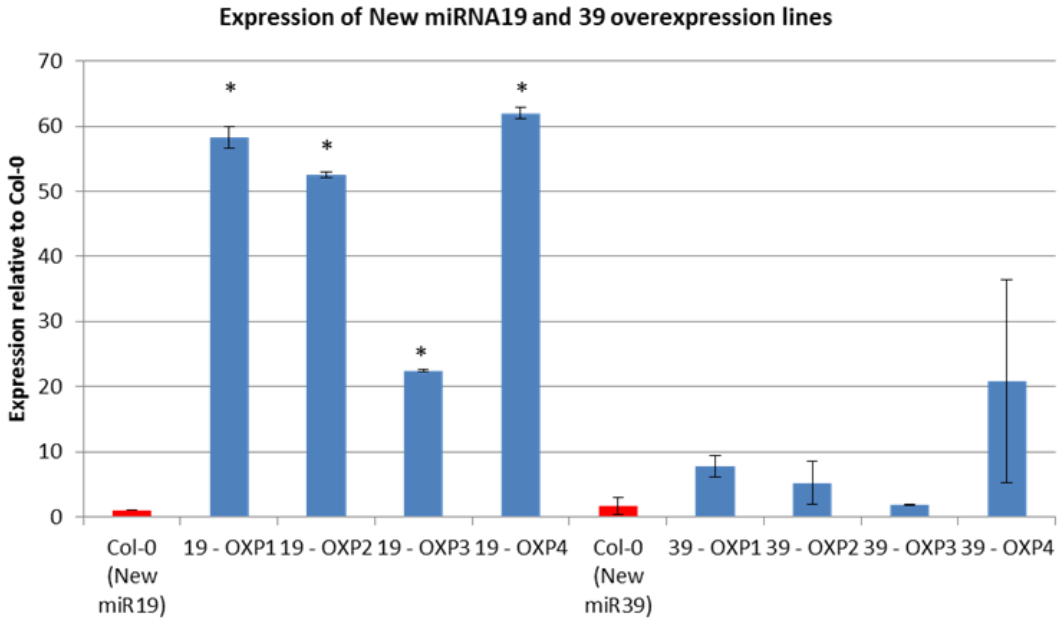


Figure 5.11: *RT qPCR of overexpression lines for new miRNA19 and new miRNA39. Expression of a region before the start of the microRNA precursor, behind the 35S promoter is compared against wild-type (Col-0). Statistically significant difference is identified with an asterisk (student's t-test <0.05). Error bars represent standard error.*

Three of the new miRNA19 overexpression line plants revealed to be significantly overexpressing were grown up to the next generation (T3) as well as three of the new miRNA39 lines which were most promising. Again, RT qPCR was used to detect the miRNAs themselves, using the stem-loop procedure (Varkonyi-Gasic *et al.,* 2007). This time,

significant overexpression could be detected in a number of independent lines for both novel microRNA19 and microRNA39 overexpression lines (figure 5.12.)

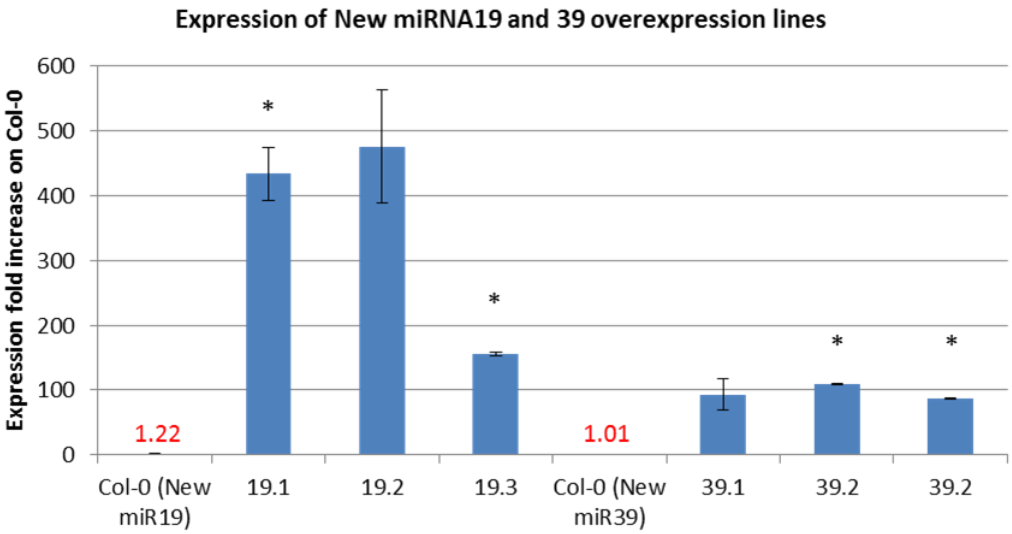


Figure 5.12: *miRNA RT qPCR of overexpression lines for new miRNA19 and new miRNA39 using the stem-loop method (Varkonyi-Gasic et al., 2007). Overexpression line expression is compared against wild-type (Col-0). Statistically significant difference is identified with an asterisk (student's t-test <0.05). Error bars represent standard error. Red numbers indicate values of Col-0 samples*

5.5.3. PHENOTYPIC ANALYSIS OF NEW-MIRNA19 AND NEW-MIRNA39 OVEREXPRESSION LINES

Three independent overexpression lines were then used for detailed lateral root phenotyping. Initially, lateral root numbers, primary root length and lateral root density were measured after 12 or 11 days' growth (figure 5.13.).

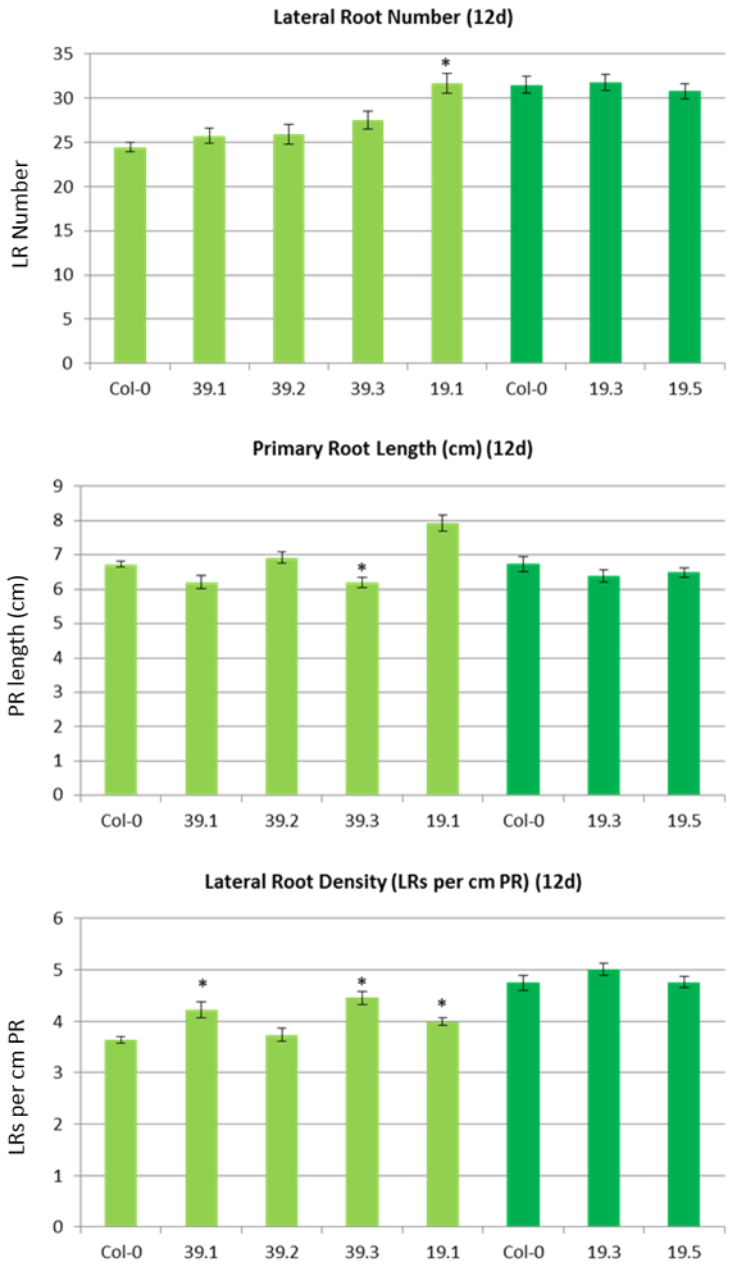


Figure 5.13: *Phenotyping of new miRNA 39 and 19 Arabidopsis 35S overexpression lines (3 x independent lines each). Lateral root numbers, density, and primary root length were measured after 12 days growth (min.24 seedlings/line). Light green bars are comparable to light green wild-type (Col-0) bar; dark green bars are comparable to dark green Col-0 bar. Statistically significant difference is identified with an asterisk (student's t-test <0.05). Error bars represent standard error.*

The results show that two of the lines overexpressing new microRNA39 and one of the three lines over expressing new microRNA 19 show a significant difference in lateral root density and primary root length . Though these results are significant it must be said that the differences are not sizable. While considerable numbers of seedlings were analysed, these results would need to be further validated and confirmed in order to give credence to this significance. Primordia morphology was also investigated in these same overexpression lines using the gravistimulus method used to obtain root primordia initially (figure 5.14. 5.15.)

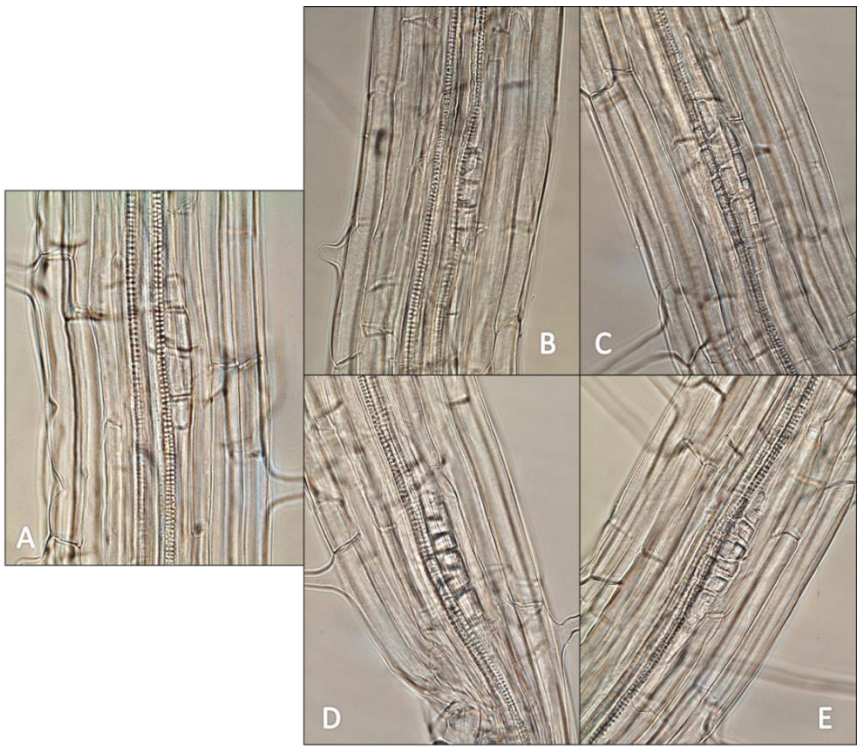


Figure 5.14: *Developing root primordia of novel microRNA overexpression lines 16 hours after gravistimulus, compared against wild-type (A – Col-0, B & C – OXP19, D & 5 – OXP39). 40X magnification*

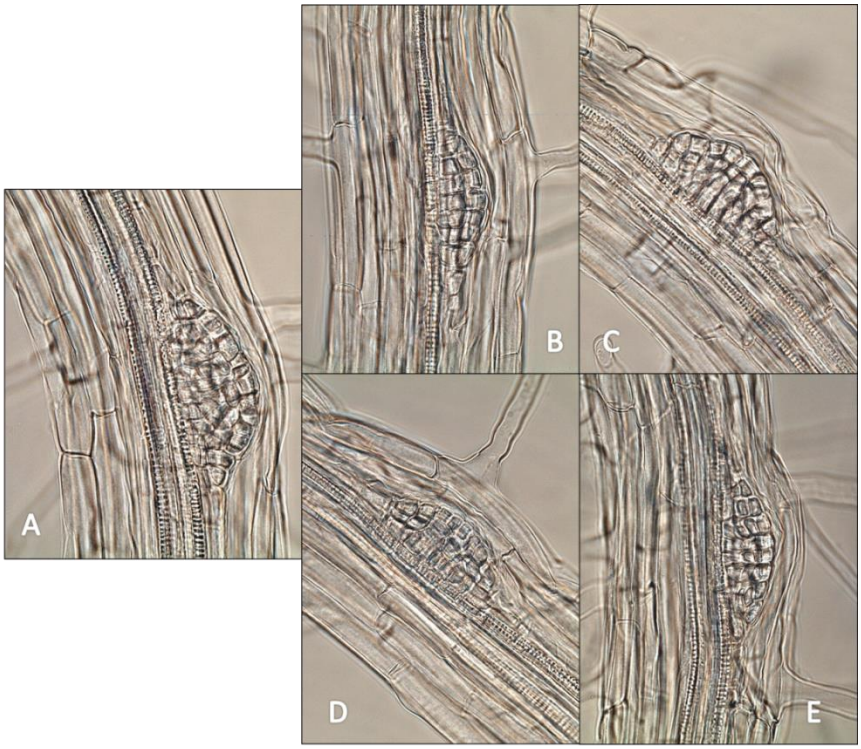


Figure 5.15: *Developing root primordia of novel microRNA overexpression lines, 24 hours after gravistimulus, compared against wild-type (A – Col-0, B & C – OXP19, D & 5 – OXP39). 40X magnification*

No change in morphology or the timing of LR primordia initiation or development could be detected in several independent lines of the new miRNA19 and new miRNA39 overexpression lines.

The study next turned to identification of potential targets for some of the microRNAs identified and validated.

5.6. TARGET PREDICTION FOR TEN NOVEL MICRORNAS

Potential targets for the ten novel miRNAs were predicted using psRNATarget (The Samuel Roberts Nobel Foundation, Dai and Zhao, 2011). As well as the mature miRNA sequences, a prediction was also run on what would be the equivalent miRNA* sequences in accordance with the widespread belief now that these sequences themselves may also have a role in regulation (for example, Yang *et al*., 2013). Default psRNTarget settings were used for the prediction and sequences were searched against the *Arabidopsis thaliana* genome from TAIR (Lamesch *et al*., 2011) (volume 10).

Targets were predicted for nearly every miRNA sequence and miRNA* sequence and these were narrowed down further by selecting only targets which were predicted to be cleaved by the potential miRNA sequence. Those targets predicted to be involved in transcriptional interruption of the target genes were overlooked at this point (this might be predicted when there is a mismatch at the point at which the miRNA would ordinarily cut the target sequence). psRNATarget provides an 'expectation' score for each prediction based on a range of considerations including number of base-pair mismatches and free energy scores, a system originally devised by Zhang (2005). The lower this number the more confidence in the prediction and anything with a score over 3 is discarded. Many miRNAs or miRNA* sequences revealed targets with a score of ≤2 which indicates a high confidence in the potential target. For the purposes of this study, targets with a score of 3 were included only if there were no, or few, other predicted potential targets for a particular miRNA or miRNA* sequence.

There were 40 predicted targets that met these criteria (18 for miRNA sequences and 22 for miRNA* sequences, tables 5.3. and 5.4.). Targets were identified for all miRNA or miRNA* with the exception of new miRNA36, new miRNA39*, and both the mature and passenger strands of new miRNA120 and new141 for which there were no predicted cleaved targets at all.

miRNA or miRNA*	Expectation Score	Target	TAIR description of potential target
11	3	AT4G37290	Unknown Protein
11	3	AT1G28290	AGP31, ARABINOGALACTAN PROTEIN 31
19	3	AT4G39300	Unknown Protein
19	3	AT3G55590	GLUCOSE-1-PHOSPHATE ADENYLYLTRANSFERASE FAMILY PROTEIN
19	3	AT3G09640	APX1B, ASCORBATE PEROXIDASE 1B
29	1.5	AT5G08400	Protein Of Unknown Function
29	2.5	AT3G62000	S-ADENOSYL-L-METHIONINE-DEPENDENT METHYLTRANSFERASES SUPERFAMILY PROTEIN
29	2.5	AT1G22020	SERINE HYDROXYMETHYLTRANSFERASE 6, SMH6
29	2.5	AT4G32300	S-DOMAIN-2 5, SD2-5
35	1.5	AT3G27500	CYSTEINE/HISTIDINE-RICH C1 DOMAIN FAMILY PROTEIN
35	2	AT3G27473	CYSTEINE/HISTIDINE-RICH C1 DOMAIN FAMILY PROTEIN
35	1.5	AT3G27480	CYSTEINE/HISTIDINE-RICH C1 DOMAIN FAMILY PROTEIN
35	1.5	AT3G27490	CYSTEINE/HISTIDINE-RICH C1 DOMAIN FAMILY PROTEIN
78	3	AT4G39690	Molecular Function Unknown
78	3	AT3G19720	ACCUMULATION AND REPLICATION OF CHLOROPLAST 5, ARC5
78	3	AT5G27030	TOPLESS-RELATED 3, TPR3
117	2.5	AT3G09850	D111/G-PATCH DOMAIN-CONTAINING PROTEIN
117	3	AT3G17890	Unknown Protein

Table 5.3: *Targets for novel miRNAs predicted by psRNATarget. A maximum expectation threshold of 3 was used - only targets predicted to be cleaved by the miRNA sequence were selected.*

miRNA or miRNA*	Expectation Score	Target	TAIR description of potential target
11*	2	AT4G22590	TPPG, TREHALOSE-6-PHOSPHATE PHOSPHATASE G
11*	2	AT4G22592	CONSERVED PEPTIDE UPSTREAM OPEN READING FRAME 27, CPUORF27
11*	2.5	AT4G31850	PGR3, PROTON GRADIENT REGULATION 3
19*	0	AT1G80410	EMB2753, EMBRYO DEFECTIVE 2753
19*	2.5	AT1G79790	CHLOROPLAST FMN HYDROLASE 1
19*	2.5	AT1G27750	nucleic acid binding
29*	2	AT2G28180	CATION/H+ EXCHANGER 8, CHX8
29*	2	AT4G17620	glycine-rich protein
29*	2	AT3G14810	MECHANOSENSITIVE CHANNEL OF SMALL CONDUCTANCE-LIKE 5, MSL5
29*	2	AT4G30993	Calcineurin-like metallo-phosphoesterase superfamily protein
29*	2.5	AT4G11160	Translation initiation factor 2
35*	3	AT2G44680	CASEIN KINASE II BETA SUBUNIT 4, CKB4
36*	2.5	AT2G25340	VAMP712, VESICLE-ASSOCIATED MEMBRANE PROTEIN 712
36*		AT1G53885	Protein of unknown function
36*		AT1G53903	Protein of unknown function
39*	2	AT1G44120	CELLULOSE SYNTHASE INTERACTIVE 2, CSI2
78*	3	AT3G43860	GH9A4, GLYCOSYL HYDROLASE 9A4
78*	3	AT1G14790	RDR1, RNA-DEPENDENT RNA POLYMERASE 1
78*	0	AT3G61270	Lipopolysaccharide-modifying protein
117*	1.5	AT5G21040	F-BOX PROTEIN 2, FBX2
117*	3	AT3G08020	PHD finger family protein

Table 5.4: *Targets for novel miRNA* sequences predicted by psRNATarget. A maximum expectation threshold of 3 was used - only targets predicted to be cleaved by the miRNA sequence were selected.*

5.6.1. BIOINFORMATIC PIPELINE TO PRIORITISE POTENTIAL TARGETS

A bioinformatics pipeline was used to narrow down the list of potential targets for further work.

Expression strength, dynamic expression profiles and a strong correlation in expression between RNASeq and Affymetrix datasets were used as key criteria to identify a subset of targets.

Firstly, expression data for each target from the two datasets was examined. All-but one of the predicted targets (for both miRNA and miRNA* sequences) were identified in the 15 time-point RNASeq; the one which was missing being CPUORF27 (CONSERVED PEPTIDE UPSTREAM OPEN READING FRAME 27, AT4G22590), a potential target for New miRNA11*. With regards the Affymetrix dataset, thirty of the forty targets were identified with three of the unidentified coming from those predicted for miRNA sequences and the other seven (including CPUORF27) predicted for miRNA* sequences.

In the case of five of the targets identified in only the one RNASeq dataset, reads form RNASeq were particularly low (AT3G27480, AT3G27490, AT3G14810, AT1G53903, and AT1G53903) which might account for them not being found in the Affymetrix dataset. However, in the case of the other four found only in RNASeq, expression levels were reasonably high (AT5G27030 (a target for New miRNA78), and the miRNA* targets: AT4G22592, AT2G28180, and AT3G61270).

With regards expression-profile correlation between datasets; the thirty predicted targets found in both show, on the whole, very strong correlation with all but six showing a positive PCC between RNASeq and Affymetrix profiles (although one more, VAMP712 (VESICLE-ASSOCIATED MEMBRANE PROTEIN 712, AT2G25340) could not be scored due to such low RNASeq expression). Furthermore, eighteen of the 24 remaining have strong PCC scores of above 0.5., twelve of which are targets for the mature miRNA sequences (as-opposed to the miRNA* sequences).

Targets which showed anti-correlation between datasets were disregarded at this point while the more lowly-expressed ones, only found in the RNASeq, were also disregarded. This left a total of 28 potential targets, 14 for miRNA sequences, and 14 for miRNA* Sequences. Targets were then further refined by expression-strength. In cases where expression was particularly low in both RNASeq and Affymetrix, targets were disregarded at this point. This removed a further five targets (AT3G55590, AT3G27500, AT3G27473, AT2G28180, and AT1G14790).

This bioinformatics pipeline resulted in a small number of 'most likely' targets which spans five of the predicted miRNA (11, 19, 29, 78, and 117), and six of the miRNA* sequences (the miRNA* sequences of the same five plus now miRNA35). The list was further narrowed down through consideration of the psRNATarget scores given to the targets (targets with a score of more than 2.5 only remained if there were no-other predicted targets or if the nature of the target as described on TAIR, was in itself, inherently interesting) as well as their presence in alternative data-sets.

5.6.1.1. POTENTIAL TARGETS: EXPRESSION IN OTHER INDEPENDENT DATASETS

To test the importance of putative targets in LR development, their expression was further investigated in three independent lateral root related datasets.

5.6.1.1.1. LATERAL ROOT INITIATION DATASET (VANNESTE *ET AL.*, 2005)

Vanneste *et al*., (2005) used microarrays to study lateral root initiation. In their study, they compared expression profiles between the roots of wild-type Arabidopsis and the auxin signalling mutant solitary root1 (*slr1*). As its name suggests, this mutant does not produce any lateral roots so the genes acting downstream of the auxin/indole-3-acetic acid (AUX/IAA) signalling pathway will be interrupted (see sections 1.1.2. and 1.1.3. for a further description).

5.6.1.1.2. ASYMMETRIC CELL DIVISION DATASET (DE SMET *ET AL.*, 2008)

De Smet *et al.* (2008) created a dataset concerned with lateral root initiation. This is a microarray based expression dataset for pericycle cells associated with the xylem poles (where lateral roots initiate) and contains expression data from plants treated with various concentrations of NAA, which promotes lateral root growth.

5.6.1.1.3. AUXIN RESPONSE DATASET (OKOSHIMA *ET AL*., 2005)

Finally, the potential targets were identified with an auxin response dataset (Okoshima *et al.,* 2005). This dataset is from microarray analysis of *arf7* and *arf19* mutants as well as an *arf7/arf19* double mutant. These two ARFs have been heavily linked with lateral root development (detailed in Chapter 1) and as such, any differential expression of the predicted target genes in this dataset is of immediate interest.

All 40 predicted target genes were checked against these databases. In the LR Initiation dataset, five were shown to have significant expression differences between wild type and *slr1* suggesting a potential role in lateral root development. Three of these genes have unknown functions (according to TAIR) while the others are 'Translation initiation factor 2' and a gene which encodes ACCUMULATION AND REPLICATION OF CHLOROPLAST 5 (ARC5).

With regards the Asymmetric cell division dataset; again, 8 genes were found to be significantly differentially expressed between different concentrations of NAA, one of which was also identified in the LR Initiation dataset (TRANSLATION INITIATION FACTOR 2). Others include an embryo defective phenotype and Ascorbate peroxidase 2 which has been linked to a drought and heat response (interestingly, this response is to both stresses together, but not independently) (Koussevitzky *et al*., 2008).

With regards the Auxin response dataset, only one target was highlighted as having a striking expression pattern, and the protein it encodes for is unknown to TAIR.

A list of the potential targets identified in these three datasets can be found in table 5.5. Reassuringly, of the twelve potential targets identified in at least one of these datasets (from the original 40 predicted by psRNATarget), ten correspond with those narrowed down by the criteria outlined in section 5.6.1.1. Furthermore, one of the other two (an unknown protein, AT5G08400) was only rejected due to a lower RNASeq expression strength, even though it still correlated well (PCC of 0.86) with the more strong Affymetrix expression profile.

The other rejected candidate was the potential new miRNA19 target AT3G09640, another unknown protein which was expressed lowly in both Affymetrix and RNASeq datasets with a very poor correlation (PCC = -0.33) but was present in the asymmetric cell division data set (De Smit *et al.* 2008).

miRNA or miRNA*	Expectation	Target	Vanneste (2008)	De Smet (2008)	Okushima (2005)	Tair Description
11	3	AT4G37290	*		*	unknown protein
19	3	AT4G39300	*			unknown protein
19	3	AT3G09640		*		APX1B, ASCORBATE PEROXIDASE 1B
19*	0	AT1G80410		*		EMB2753, EMBRYO DEFECTIVE 2753
19*	2.5	AT1G79790		*		CHLOROPLAST FMN HYDROLASE 1
29	1.5	AT5G08400	*			Protein of unknown function
29	2.5	AT4G32300		*		S-DOMAIN-2 5, SD2-5
29*	2	AT4G30993		*		Calcineurin-like metallo-phosphoesterase superfamily protein
29*	2.5	AT4G11160	*	*		Translation initiation factor 2
35*	3	AT2G44680		*		CASEIN KINASE II BETA SUBUNIT 4, CKB4
78	3	AT3G19720	*			ACCUMULATION AND REPLICATION OF CHLOROPLAST 5, ARC5
117	2.5	AT3G09850		*		D111/G-patch domain-containing protein

Table 5.5: *psRNATarget predicted targets of novel miRNA (and equivalent miRNA* sequences) found in one or more selected datasets (Vanneste et al., 2008, De Smet et al., 2008 and Okoshima et al., 2005). Asterisk indicates a significant (t-test) difference in expression under the various conditions of the experimental datasets.*

5.6.1.2. UNIGENE NOVEL MIRNA TARGETS

Among the 12 targets identified above, unigene targets, or those belonging to small gene families, are much more likely to result in phenotype due to less genetic redundancy. These were therefore prioritised for genetic studies. The number of related family members for each target was identified. The analysis revealed that except for four targets, all other targets belonged to small gene families (1-3 members). Among the other four, two were part of

multi-gene families of less than 11 family members. The last two targets were part of the same gene family of 124 members.

5.6.1.3. NOVEL MIRNA AND POTENTIAL TARGETS SHOW CONTRASTING EXPRESSION

If these potential target candidates are indeed regulated by these novel microRNAs, their expression patterns would be expected to contrast those of their miRNA regulators during lateral root development. Analysis revealed that, in general, there is a high anti-correlation between miRNA and potential target gene expression over five time points during LR development for many of these miRNA/target couples. Though this would strongly suggest that the identified targets for these novel microRNAs may indeed be the real targets, the significance of this is not yet clear particularly for lowly-expressed microRNAs, where the changes in expression may not have a significant effect on the expression of the target gene. In cases like this expression correlation may well simply be coincidental. Unfortunately, many of the novel microRNAs identified in this study are particularly lowly expressed.

The expression patterns of eleven potential targets spanning five of the novel microRNAs were compared to those of the microRNAs they are predicted to be targeted by (table 5.6. and figure 5.17.). In the cases of the targets predicted to be targeted by new miR11, since the two replicates of that particular microRNA are so different, expression profiles were compared against both the small RNASeq replicates (table 5.6. and figure 5.17.).

Target	miRNA	PCC (Average of reps)	PCC - Rep A	PCC Rep B
AT4G37290	11		-0.194	0.939
AT1G28290	11		-0.498	-0.093
AT4G39300	19	-0.590		
AT5G08400	29	0.836		
AT3G62000	29	0.867		
AT1G22020	29	-0.092		
AT4G32300	29	0.653		
AT4G39690	78	0.088		
AT3G19720	78	-0.710		
AT5G27030	78	-0.788		
AT3G17890	117	0.823		

Table 5.6: *PCC between selected, predicted target RNASeq expression profiles and those of the miRNA which are predicted to target them (average of two replicates where PCC between reps = 0.4 or above)*

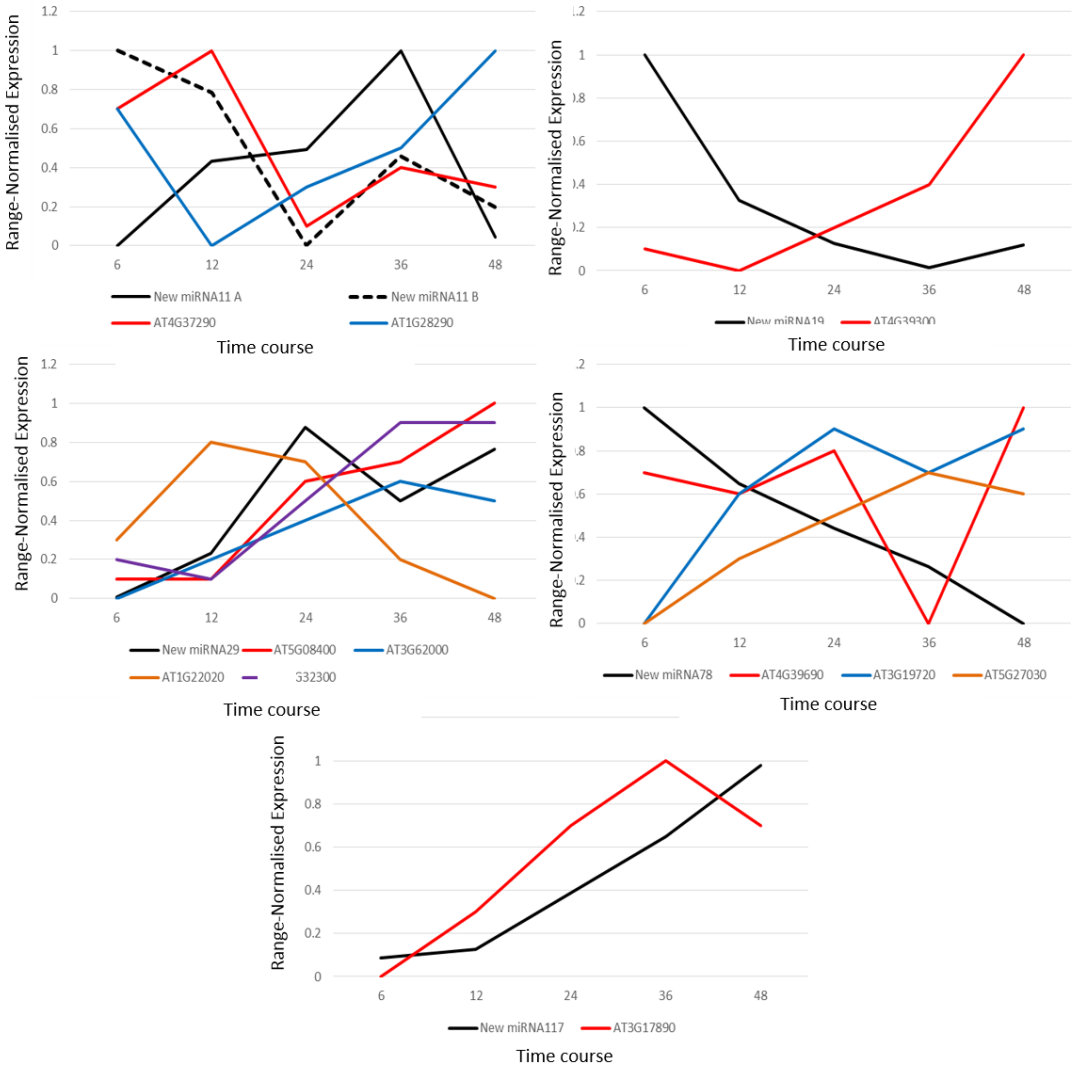


Figure 5.16: *Comparison of range normalised expression profiles for novel microRNAs and selected, predicted targets during lateral root primordia development, from small RNASeq and mRNASeq*

The expression profile comparisons between the microRNAs and predicted targets are varied with some showing strong anti-correlation (the new miRNA19 target AT4G39300 and new miRNA78 target AT5G27030 in particular) while others are more erratic or, in the case of new miR11, due to the lack of correlation between miRNA replicates, impossible to ascertain.

Potential miRNA* targets cannot be analysed in this manner since these are only potentially active sequences which did not come out of the RNASeq data as mature sequences. As such, these potential targets might only be considered should they look particularly enticing based on the other criteria outlined in this chapter.

5.6.1.4. NEW MICRORNA - TARGET COUPLES: MOST PROMISING CANDIDATES

Based on the narrowing down of potential targets described in this chapter, a small number of targets were selected as being the strongest candidates for novel miRNA regulation during LR development (table 5.7.). In total, eight potential targets were predicted spanning five novel microRNAs as well as a further three targets predicted for miRNA* sequences. These were predicted primarily on their strong expression in the RNASeq and Affymetrix datasets and on the very reassuring scores of '0' from psRNATarget (table 5.7.).

Target	New miRNA	PsRNATarget Score	PCC Affymetrix/RNASeq	PCC Target/miRNA (RNASeq)	Presence in other datasets	Details
AT1G28290	11	3	0.51	-0.498 (rep A) -0.093 (rep B)	Vanneste *et al*., 2005, Okushima *et al*., 2005	ARABINOGALACTAN PROTEIN 31 (AGP31)
AT4G39300	19	3	0.96	-0.590	Vanneste *et al*., 2005	Unknown protein
AT3G09640	19	3	-0.33		De Smet *et al*., 2008	ASCORBATE PEROXIDASE 1B (APX2)
AT5G08400	29	1.5	0.86	0.836	Vanneste *et al*., 2005	Unknown protein
AT1G22020	29	2.5	0.66	-0.092		SERINE HYDROXYMETHYLTRANSFERASE 6 (SMH6)
AT4G39690	78	3	0.52	0.088		Unknown protein
AT3G19720	78	3	0.74	-0.710	Vanneste *et al*., 2005	ACCUMULATION AND REPLICATION OF CHLOROPLAST 5 (ARC5)
AT5G27030	78	3	n/a	-0.788		TOPLESS-RELATED 3 (TPR3)
AT3G09850	117	2.5	-0.04		De Smet *et al*., 2008	D111/G-patch domain-containing protein
AT1G80410	19*	0	0.82	n/a	D De Smet *et al*., 2008	EMBRYO DEFECTIVE 2753 (EMB2753)
AT3G61270	78*	0		n/a		Unknown protein

Table 5.7: *Most promising potential novel miRNA targets, (based on expression profiles (between RNASeq and Affymetrix and between target & miRNA), prediction scores, presence in relevant, external datasets and nature of target itself) selected for further investigation*

While some of the predicted targets encode unknown proteins or encode proteins for which no clear role has yet been described, amongst this short-list of targets are a number which have been studied previously in some depth, and some which have been assigned quite specific roles in plant developmental processes and/or stress responses.

One of the most intriguing targets is one which was predicted to be regulated by miRNA19. Ascorbate Peroxidase 1B (also known as Ascorbate peroxidase 2 - AXP2) has been liked to plant responses to heat, salinity, light, and oxidative stresses with a knockdown mutant

found to be more susceptible to these pressures (Suzuki *et al*., 2013). The ascorbate peroxidase proteins act to detoxify peroxidases such as hydrogen peroxide and this particular one has been found to be up-regulated in the roots under any of these stresses, as well as in response to wounding (Zimmermann *et al*., 2004, Frank *et al*., 2009). Interestingly, the highly related Ascorbate peroxidase 1 (AXP1) has been linked to a response to combined stress pressures (Koussevitzky *et al.,* 2008). While no response was found to heat or drought separately, a combination to the both revealed a strong response in AXP1 expression (Koussevitzky *et al.,* 2008).

Arabinogalactan protein 31 (AGP31), a predicted target of new miRNA11, has been found to be highly expressed in Arabidopsis vasculature, regulated by jasmonic acid and localised in cell walls (Hijazi *et al*., 2012, Liu and Mehdy, 2007). As such, AGP31 has been linked to roles in vascular tissue function during developmental processes and defence responses (Liu and Mehdy, 2007).

Due to their nature and connection to global stress and defence responses, both of these potential targets might be expected to be found expressed quite ubiquitously throughout the plant. However, AXP2 was not identified as being differentially expressed in our RNASeq dataset. AGP31 did not reveal strong correlation with new miRNA11 though there was discrepancy between the two replicates of this microRNA anyway so this relationship cannot be clarified by expression profiling using these RNASeq datasets.

Accumulation and replication of chloroplast 5 (ARC5) (a potential target of new miR78) encodes a protein which is involved in chloroplast replication (Robertson *et al.,* 1996). As

such, this potential target probably has less relevance to root development though its expression during LR development did show very strong anti-correlation with the expression profile of its predicted regulator.

Perhaps most striking of all of the targets predicted is TOPLESS-RELATED 3 (TRP3 – a predicted target of new miR78). TOPLESS is a transcriptional co-repressor and along with AUX1/IAA proteins, it is a negative regulator of the auxin pathway. Perception of auxin by the auxin receptor TIR1 relieves the transcriptional repression induced by TOPLESS enabling the transcription of early-auxin responsive genes by ARFs and has been linked to proper embryonic development (Szemeynei *et al.,* 2008). This gene is hugely important because of the crucial role of auxin in primary and lateral root development (as described in Chapter 1). TOPLESS loss-of-function mutants show a dramatically compromised phenotype whereby they fail to correctly develop the top-parts of the plant including developing apical roots rather than shoots in the embryo (Long *et al.,* 2002).

As such, the TOPLESS RELATED 3 gene is of particular interest in this study and its highly dynamic expression pattern which rises and falls over time certainly might suggest a possible role in correct root development. Furthermore, expression of TRP3 also shows a strong anti-correlation with that of its potential miRNA regulator which drops over the time course which would hint at a relaxation of repression over the course of primordia development.

Of these narrowed-down targets, the expression profiles of three show particularly strong anti-correlation with the expression profiles of their potential microRNA regulators. As well as the aforementioned ARC5 and TRP3, an unknown protein predicted to be targeted by new

miRNA19 has a PCC of -0.59. However, these values must be tempered by the relatively low-expression of some of these microRNAs.

All of these predicted targets, along with some of the unknown ones for which no function has yet been assigned were considered for phenotypic analysis.

5.6.2. PHENOTYPIC ANALYSIS OF SELECTED PROMISING NOVEL MIRNA TARGET CANDIDATES

To investigate if the predicted targets are indeed the real targets for the new microRNAs, a genetic approach was used. Following identification of the most promising candidates based on expression profiles, expression strength and scores awarded to them by psRNATarget, T-DNA insertion mutants for the potential targets were identified and obtained (from the National Arabidopsis Stock Centre (NASC) – Nottingham), and their root and lateral root phenotypes were investigated. Where possible, homozygous exonic insertions were selected. Table 5.8. outlines mutants which were identified. All lines obtained were genotyped and selected in order to identify lines which were homozygous for the mutation.

miRNA	Target	Mutant line	tDNA insertion Location
11	AT1G28290	No mutant available	
19	AT4G39300	SALK_053004	Unavailable
19	AT3G09640	SALK_057686	Exon
29	AT5G08400	SALK_068624	Exon
29	AT1G22020	SALK_042297	Promoter
78	AT4G39690	SALK_087650	Exon
78	AT3G19720	No mutant available	
78	AT5G27030	SALK_029936	Intron
117	AT3G09850	SALK_055098	Promoter

Table 5.8: *Details of tDNA-insertion mutants selected and obtained for predicted targets of novel microRNAs.*

The root phenotypes for the predicted target mutants (homozygous) were screened to see if any significant difference in lateral root numbers, primary root length or lateral root density (compared to wild-type - Col-0) could be detected after either 12 or 11 days growth (figure 5.17.). Seeds used were bulked up alongside the same batch of Col-0 used in this experiment.

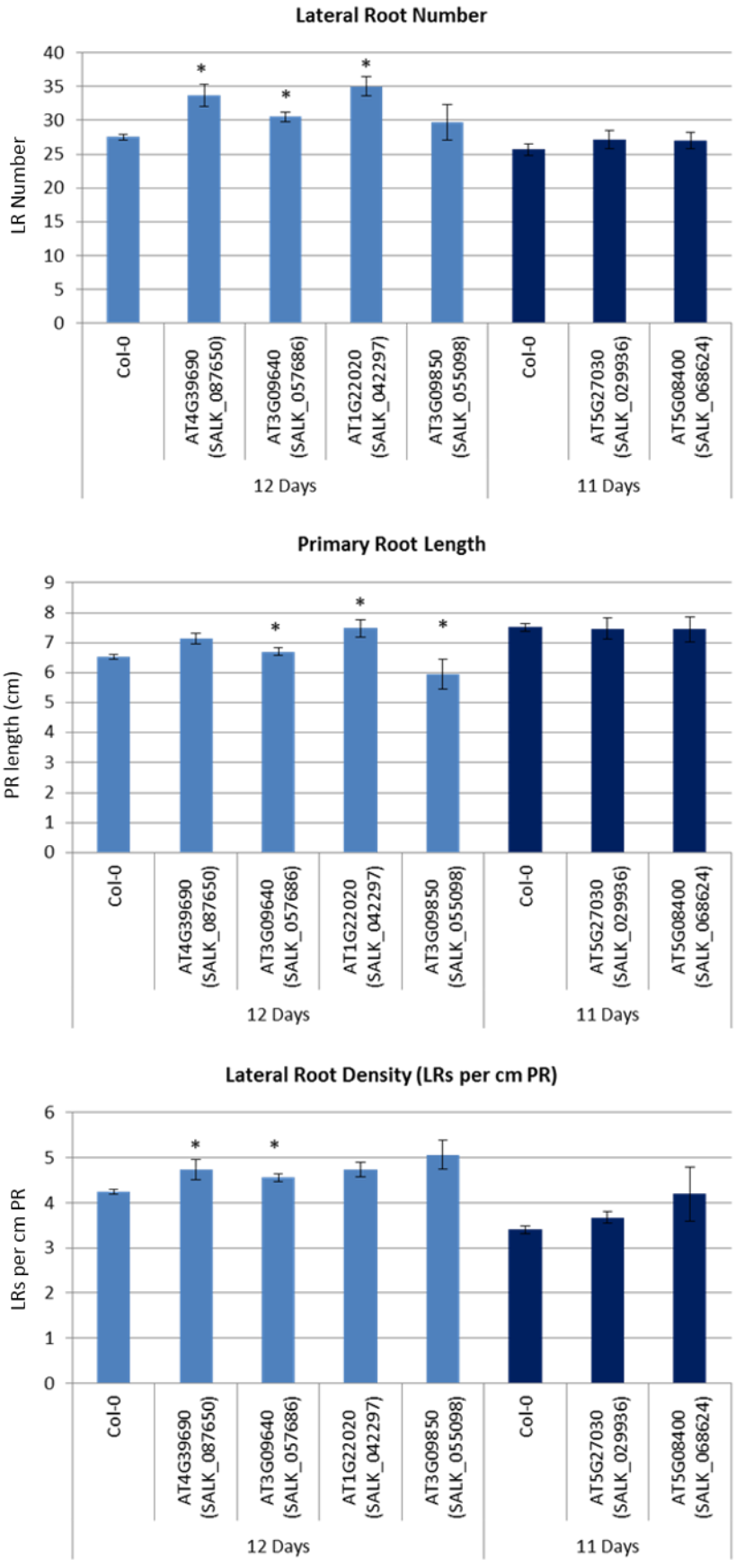


Figure 5.17. *Legend next page*

Figure 5.17: *Phenotyping of tDNA insertion-knockout lines for selected, predicted targets for novel miRNA. Lateral root numbers, density, and primary root length were measured after 12 or 11 days growth. Light blue bars are comparable to light blue wild-type (Col-0) bar; dark blue bars are comparable to dark blue Col-0 bar. Statistically significant difference is identified with an asterisk (student's t-test <0.05). Error bars represent standard error. Table describes genomic location of tDNA insert for each line and for which novel microRNA that particular gene is a predicted target*

Of the seven lines (spanning seven potential targets) which were selected for detailed root phenotyping analysis, four showed a statistically significant difference in lateral root number, primary root length and/or lateral root density (figure 5.18.). However, while significant, none of these phenotypes could be considered particularly dramatic with lateral root density only showing a difference of around 0.5 more lateral roots per cm of primary root for the most dramatically different mutant line.

Unfortunately, and perhaps surprisingly, no phenotype was noted in this analysis or the potential miRNA target TOPLESS RELATED PROTEIN 3 (TRP 3). Considering the TOPLESS proteins have been related to embryonic miss-development, early stages of development were observed closely and no obvious developmental defects were witnessed in the germinating seed or young seedling. This is possibly die to the fact that it is not a particularly strong mutation (with an intronic t-DNA insertion)

5.6.3. EXPRESSION ANALYSIS OF PREDICTED TARGETS IN NEW MIRNA19 OXP LINES

In order to confirm some of these targets, RT qPCR was undertaken on one of the novel miRNA overexpression lines produced. Three targets were predicted for novel miRNA19: Ascorbate Peroxidase 1B (APX2), a Glucose-1-phosphate adenylyl-transferase family protein, and an unknown protein (from TAIR10 descriptions). Two of these targets (APX2 and the unknown protein) were amongst the most promising candidates.

Unfortunately, no targets were predicted for the other overexpression line which was produced (new miRNA39).

cDNA was obtained from the RNA of young seedlings of T3 generation new miRNA19 overexpression lines. Two of the targets could not be amplified by RT qPCR while the other one (AT3G09640 - Ascorbate peroxidase 1B) was found to be particularly highly expressed in two of three overexpression-line plants studied. One of these seedlings (19.2) was found to be significantly highly expressed (student's t-test) at around three-times the level seen in wild-type. The lack of statistical significance in the other is largely due to the high standard deviation found in expression in the other seedling showing high expression (figure 5.18.).

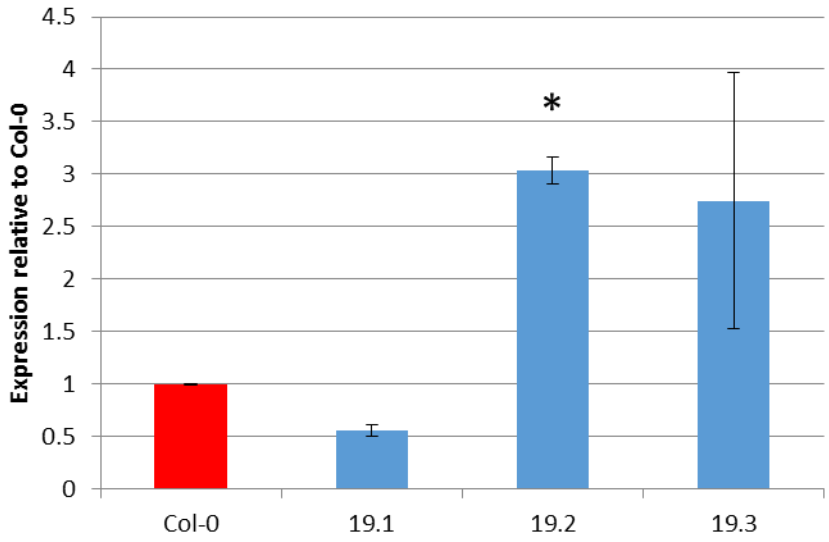


Figure 5.18: *RT qPCR on T3 new miRNA19 overexpression lines found ~3x overexpression of AT3G09640 (Ascorbate peroxidase 1B) in two of three seedlings compared to wild-type (Col-0). Asterisk indicated significant difference in expression (student's t-test). Error bars represent standard error. Expression was normalised against actin*

Unfortunately, these results provide no clear information about these three targets. The two targets which could not be identified in the new miRNA19 overexpression line could not be amplified in the wild-type Col-0 line either. While the other line (Ascorbate Peroxidase 1B) is potentially the most interesting with links to lateral root development, these results are inconclusive and reveal no clear repression of the target by this particular microRNA. With regards the unexpected higher-expression of this gene in two of the seedlings sampled, while in one case expression was found to be statistically significant, a three times difference in expression may not be considered particularly high.

A more thorough screen of these targets is crucial and one might suggest that due to the low-level expression of many of these novel microRNAs, perhaps of the forty potential targets which were predicted, only a handful might actually be targeted actively by their miRNA regulator. This study aimed to narrow down the original number to a more workable number in a logical manner. However, in reality it may be hard to identify the best possible candidates without using expression studies in lines such as the miRNA overexpression lines produced here.

Identifying suitable mutant lines for many of the targets proved challenging and expanding this arm of investigation to more of the targets predicted might be worthwhile in the hope of identifying a phenotype of interest.

In this study, two miRNA overexpression lines were produced (for new miRNA19 and new miRNA39). These were proved to be successful so far as overexpression could be detected in a number of seedlings. However, for a more comprehensive view of the novel microRNAs identified in this study and their relationships with potential targets, further overexpression lines must be produced for the other eight predicted miRNAs.

5.7. CHAPTER DISCUSSION

It is fair to assume that microRNA/target couples directly related to lateral root development which have not yet been identified are likely to be relatively lowly-expressed. One would expect that more highly expressed microRNAs (wherever they may be active in the plant) are more likely to have been identified earlier and that roles would have been validated, clarified, and studied previously. It was therefore to be expected that any novel microRNAs identified in this study were likely to be lower-expressed than the majority of the conserved microRNAs discussed in chapter 4. This did indeed turn out to be the case.

Of the many potential microRNAs identified, there were ten which were predicted with a high degree of confidence based on a robust pipeline that includes their differential expression profiles, predicted hairpin structures, and presence in both replicates. These ten were mostly among the lowest expressed of all microRNAs identified.

The ten microRNAs predicted in this study were further validated by RT qPCR. This is one of the more common methods by which to validate microRNAs and has been used to validate many of the miRNAs registered on MIRBASE. Other methods such as northern hybridisation are available and may be considered more definitive; however, in this case, RT qPCR coupled with the miRNAs' identification in two biological and technical RNASeq replicates was deemed sufficient for validation. This was further reinforced by the confidence of the bioinformatic programmes used to predict miRNA hairpins, and the conservation of the novel miRNA sequences between Arabidopsis accessions being comparable to that of well-defined microRNAs.

Encouragingly, shortly after the prediction of the microRNAs in this study, one of the miRNA sequences (that of new miRNA36) was identified independently, in a separate study (Vidal *et al*., 2013). Subsequently registered on MIRBASE as ath-miR8175, this microRNA was identified by Vidal *et al.* by RNA sequencing of the root transcriptome in a study aimed at looking into nitrate responsive genes. This independent discovery serves to reassure that the pipeline used to predict these novel microRNAs is sound. Curiously, this particular microRNA sequence was one of the less-well conserved across Arabidopsis accessions although it was the only one to be identified in other species.

The low-expression of these novel microRNAs proved challenging with regards linking them to potential roles in specific stages of LR development. The differential expression identified in all ten was coupled with low-level expression and as such could easily be misinterpreted. More troubling however, is the difficulty that this low-expression causes with regards identifying and validating microRNA/target relationships.

Several promising targets were predicted for most of the ten novel microRNAs identified in this study. High-confidence scores from the psRNATarget prediction software as well as dynamic expression profiles in both this and in other lateral root-related datasets provided a strong subset of possible targets to focus on. However, identifying anti-correlation in expression profiles between microRNAs and potential targets was far less clear-cut than in the case of many of the known microRNAs and targets described in chapter 4. While this method proved limited in what could be ascertained from the sequence data, through a pipeline narrowing them down, a number of potential targets were identified which seemed

most likely to play a role in LR development, whether or not they are regulated by these novel microRNAs. However, by far the most clear-cut way of identifying whether these potential targets are indeed regulated by the novel microRNAs would be with the use of overexpression or knockdown-based expression analysis.

To investigate the roles of these novel microRNAs in plant development, all ten novel microRNAs were overexpressed in Arabidopsis. Curiously, despite successfully generating the constructs and obtaining transgenic lines, none of the microRNAs were found to be overexpressed. It is unclear why these overexpression lines were not properly functional though, ultimately, the most likely reasons may be a defect in miRNA hairpin processing and/or a defect or incompatibility issue with the binary vector (pGWB8) used for this study (although the plasmid has previously been used successfully in Arabidopsis – Myouga *et al*., 2008 – for example).

To address the plasmid issue, another set of overexpression lines were generated for a smaller number of microRNAs using another binary vector pH7WG2,0 and two out of three of these lines proved to be successful. Both new miRNA19 and new miRNA39 could be overexpressed in Arabidopsis.

In both successful miRNA overexpression lines, no lateral root-related phenotypic defects could be identified (LR numbers, LR density, PR length or any defects in primordia morphology or developmental timing) and in the case of the new miRNA19 construct, no repression of predicted targets could be validated by RT qPCR. New miRNA19 is the lowest-expressed of all of the microRNAs identified the in lateral root primordia RNASeq dataset,

though it was also the best-conserved. Unfortunately, the other successful overexpression line, that of new miRNA36, was one of the few novel microRNAs for which no targets were predicted. It can be argued that these microRNAs may be inactive. They could even be evolutionary relics which no-longer serve a functional purpose.

With regard to the other seven novel microRNAs, due to the success of the new miRNA19 and new miRNA39 overexpression lines, in terms of over expression; it would be recommended to produce similar lines for them too. For low expressed genes, an overexpression approach may be very useful.

Use of overexpression lines is a routine method in gene investigation studies and has been used successfully in several studies across several plant species and remains one of the preferred ways to investigate gene function. However, in the absence of a phenotype in overexpression lines, another approach might be to produce miRNA knock-out lines; perhaps target mimics (such as those used when looking at known microRNAs), or by utilising the recently developed CRISPR technique of targeted silencing (Jinek *et al.,* 2012). A system has already been described employing this technique for miRNA knockdown (Zhao *et al.,* 2014).

To further pinpoint their role in LR development, investigating tissue specific spatial expression may prove useful. However, for very lowly expressed genes, this approach can be very challenging.

Ultimately, the identification and validation of ten novel miRNAs is encouraging. The prediction of potential targets for many of them is also encouraging. However, this study failed to find a strong candidate miRNA/target couple which could also be validated.

While these microRNAs were found to be differentially expressed in our lateral root-specific dataset; beyond that fact it is difficult to link any of them to lateral root development itself without further resources. Neither of the two miRNA overexpression lines revealed any phenotypic defect of interest, nor did mutants for the selected targets screened for root phenotypes. Neither physical LR development nor developmental timing seemed to be convincingly affected in any.

This is unlikely to be the case for all of the miRNAs or targets and further study would surely identify strong miRNA/target couples revealing strong phenotypic changes. As outlined in Chapters 1 and 4; there are several microRNAs which are involved in root-related processes which are responsive to external influences such as nutrient deficiencies, pathogen attack, and hormone gradients. Indeed, a number of targets predicted here have been linked to stress responses (such as ascorbate peroxidase 1B). Further study might investigate further any miRNA/target expression changes in response to some of these stimuli.

Narrowing down potential miRNA/target couples through logical means was undertaken in this study in order to identify a manageable number for further investigation. However, in reality, perhaps a wider screen would have been a better approach as ultimately it seems that miRNA/target couples which may be involved in LR development may well be more few-

and-far between than what might have been hoped, and identification of the best-possible candidates may be less straightforward.

CHAPTER 6
DISCUSSION

6. DISCUSSION

This study aimed to gain a comprehensive view of microRNAs and their corresponding targets involved in Arabidopsis lateral root development from the initiation of the root, through to emergence. Following a gravistimulus, lateral roots are initiated and using this technique, large numbers of synchronised primordia at selected stages (as described by Peret *et al*., 2009) could be collected. RNA was extracted from these primordia and mRNA sequencing and small RNA sequencing was undertaken. Two very large datasets were produced and this study is the first comprehensive analysis of global gene expression profiling during Arabidopsis LR development.

The sequence data produced was validated in a variety of ways and was found to be of great depth and of very good quality (as discussed in chapter 3). From here, attention turned to known microRNAs of which many root-related miRNA families were identified and relationships with known targets were clarified. This further served to validate the quality of the datasets. Furthermore, several novel microRNAs could be identified and putative targets identified.

6.1. RNA SEQUENCING PIPELINE

There are several different platforms available for the analysis of sequence data. When doing the data analysis, stringency and robustness are two crucial considerations. The key

question is which parameters are best for this particular set of data?; so-as not to be too lax that bad data gets through, but not too stringent so-as to miss out on good data.

Following the identification of potential microRNA reads; both known microRNAs and novel microRNAs are aligned against the genome. Alignment of known and predicted microRNAs against the genome was repeated using four different sets of parameters: alignment with no-more than one mismatch and multiple counts on that position, alignment with no-more than one mismatch and counts on that position extended by 2nt on both sides, alignment without any mismatches and with multi-hit[6], counts on the position extended by 2nt on each side, and alignment without any mismatches and with multi-hit counts on the position extended by 2nt on both side using a positioning quality of alignment set to 0 (instead of 10)[7].

Using these four sets of parameters, of the sixteen novel microRNAs identified originally in this study (ten of which were identified in both sequencing replicates); counts were identified for thirteen of them using all four sets of parameters. The other three (new microRNA36, new microRNA61, and new microRNA88 could only be identified when the quality of alignment was set to 0 (figure 6.1.). This is because the sequences are found in

[6] **Multi-hit:** when you have a read which can be mapped to different positions on the genome this is known as a 'multi-hit scenario'. Though the read can only have come from one position, we cannot know which one it is. This leaves three options to solve this problem: Firstly you can attribute the read to one of the positions, randomly (this was what this study did originally), secondly you can multiply the read so it registers on all possible locations. Thirdly, you can disregard it altogether

[7] **Quality of alignment:** The quality score is designed around the possibility that the read may have come from other position in the genome: a quality score of 10 means that there 1 chance over 10 to come from otherwise (Bowtie2)

multiple places on the genome. In all cases, expression was on the whole found to be very low

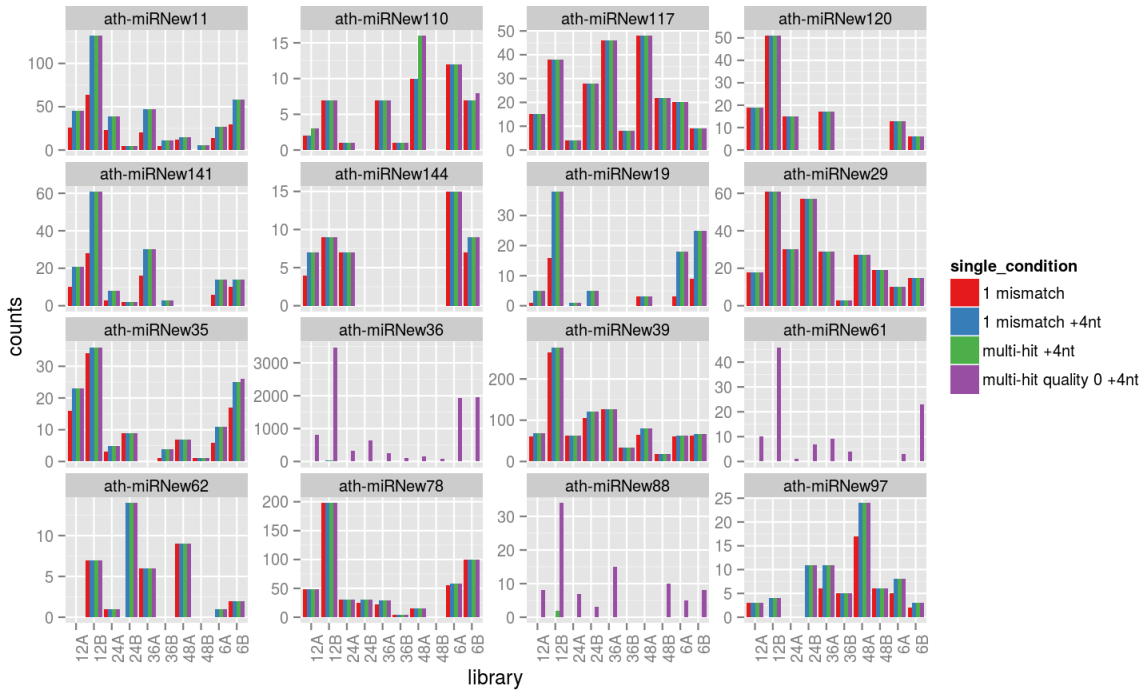


Figure 6.1: *Novel microRNAs detectable using four different sets of parameters. Bar graphs represent non-normalised counts at the various time points (y-axis). Individual sample time-points on x-axis*

In order to check the robustness of our pipeline, more rigorous parameters were used. The new methods of analysis alter the perception of the novel microRNAs' various expression patterns. However, the novel microRNAs were still identified and found to be differentially expressed to some degree in nearly all cases. All means of analysis revealed these microRNAs to be particularly lowly-expressed.

Encouragingly, both methods of analysis showed that all the novel microRNAs are differentially expressed during LR development. This was further validated by RT qPCR. It may well be the case that the variation in perceived differential expression between the various analytical methods may to a degree be a consequence of low expression.

After the preliminary less-robust analysis to ensure that nothing is missed out, a more stringent analysis was done to narrow down the numbers by alterations to parameters as well as some alternate normalisation techniques. With regards the different sets of parameters attempted, the following criteria were suggested as those best suited to the best possible alternative analysis for this data:

- In the case of a multi-hit situation, reads would be multiplied across all positions instead of being assigned randomly to the positions as was the case in the original analysis.
- Alignment quality was set to 0 as adding 2nt either side seemed to make little difference to the differentially expressed microRNAs identified.
- Counts of the sequence were limited to only the precise mature sequence as recognised on MIRBASE.
- Normalisation was performed according to the number of mapped reads on MIRBASE (which was not done in the initial analysis), followed by normalisation using DESeq2 (DESeq was used in the original analysis).

As expected, the new analysis does show changes to expression profile to the original analysis (figure 6.2).

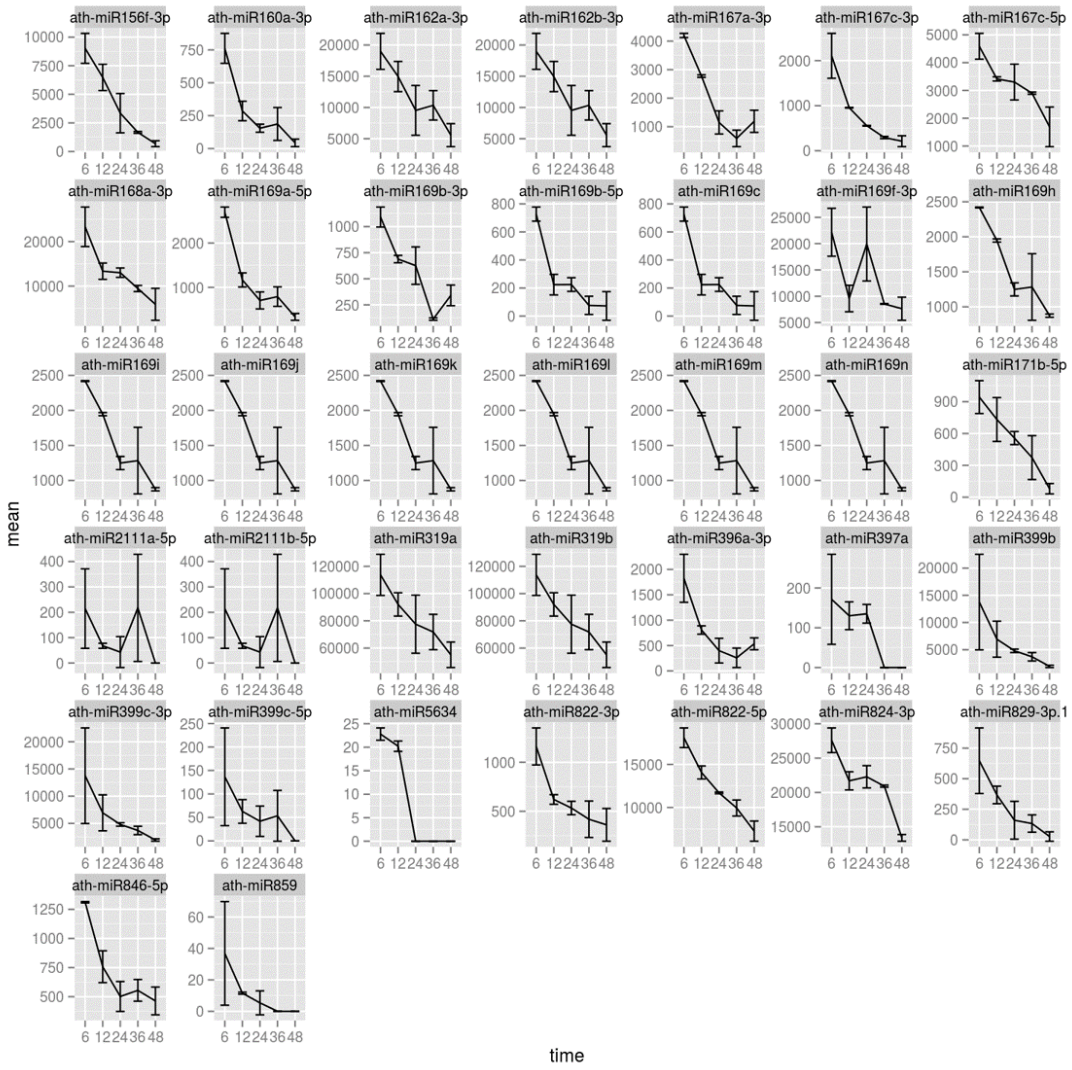


Figure 6.2: *Differentially-expressed, known microRNAs identified using an alternative bioinformatics pathway: multiplications of the reads in case of multi-hit, alignment quality set to 0, normalisation according to the number of mapped reads on MIRBASE prior to normalisation by DESeq2. Bars represent standard error. Time course is represented on the x-axis while an average of normalised reads is on the y-axis.*

What is quite striking about this data is the relatively small-number of microRNAs which could be found to be differentially expressed when compared to the original analysis. As

expected, these criteria seem to be considerably more stringent with regards narrowing down those microRNAs which are substantially (and perhaps more statistically significantly), differentially expressed. Members of common miRNA families including miR156, miR160, miR169, miR171, and miR399 were identified using both methods though certain families, including that which was found to be most strongly expressed in our initial analysis (miR165/166), and the interesting miR159 family, are curious by their absence.

This finalised method of analysis utilised DESeq2. Initially, DESeq was used for normalisation but there was high discrepancy between the results when using these two different programmes. However, there was still a high degree of differentially expressed miRNAs identified by both of these programmes (table 6.1.). Ultimately, microRNAs of interest (as highlighted through this round of analysis) could be grouped into 'best possible' candidates by way of appearing under different, varied, methods of analysis. This could also be applied to microRNAs which would be identified when either allowing, or not allowing, a nucleotide mismatch. Table 6.1. shows these groups of microRNAs. Regarding using either DESeq, or DESeq2, and allowing or not allowing nucleotide mismatches; the column on the left indicates microRNAs which were identified in all cases (the strongest candidates) while the central column includes microRNAs which were identified by more than one method. Finally, the right hand column are microRNAs identified only by DESeq2 (both with particularly low read counts).

Identified through all or most means of analysis (DESeq/DESeq2/with or without mismatch)	Identified in some types of analysis	Only identified using DSeq2 analysis (particularly low number of reads)
miR156f-3p	miR169b-3p	miR5634
miR160a-3p	miR169h	miR859
miR162a-3p	miR169i	
miR162b-3p	miR169j	
miR167a-3p	miR169k	
miR167c-3p	miR169l	
miR167c-5p	miR169m	
miR168a-3p	miR169n	
miR169a-5p	miR319a	
miR169b-5p	miR319b	
miR169c	miR396a-3p	
miR169f-3p	miR397a	
miR171b-5p	miR399c-5p	
miR2111a-5p	miR822-5p	
miR2111b-5p	miR824-3p	
miR2111b-5p		
miR829-3p		

Table 6.1: *MicroRNAs identified as being differentially expressed during LR development based on a variety of combinations of various analysis tools (DESeq, DESeq2, with or without mismatches allowed)*

That the two different analysis methods produced slightly different results clearly highlights the importance of carefully selecting the parameters. However, the robustness of our initial approach is highlighted from the fact that the microRNAs which were expected to be found, were indeed found, and that many of these families showed strong anti-correlation with their known targets. Furthermore, differential expression could also be detected independently using RT qPCR.

In any study using next generation sequencing data, alternative methods for data analysis might be employed. While it may be debated as to what would have been the best approach for this study, the method of data processing and analysis used primarily in this study was deemed to be sufficient for the purposes of studying miRNA/target couple relationships which may play a role in LR development.

6.2. KNOWN MICRORNA EXPRESSION DURING LR DEVELOPMENT

The global gene expression profiling clearly established known microRNAs with a role in Arabidopsis LR development.

This data further supported the role in LR development of certain targets such as PHV and PHB and had highlighted other targets to which roles have subsequently (and independently) been described, such as in the case of SPL9 (Yu *et al.*, 2015). Other targets such as BUB3.2 and SULTR2,1 have been found in this study and their expression patterns do hint at possible roles in LR development, and whose biological function supports this.

There are several other microRNA/target couples of interest and future work would reveal their roles in LR development. While these miRNA/target couples are of great interest to LR development, this study focused on the in depth analysis of one known microRNA and ten novel microRNAs.

The one known miRNA chosen for further investigation; miR159, has previously been shown to regulate plant development (Allen *et al*., 2010, Alonso-Peral *et al*., 2010). Not much is known about its role in root development.

This study showed that miR159 plays a central role in root and lateral root development as the *miR159ab* double mutant and miR159 target mimic (MIM159) both have a serious defect in LR development.

This study strongly suggests that if miR159 has a role in LR development, the most likely candidate target to be involved is MYB33. This highly-expressed microRNA is also strongly repressed by miR159 and expression studies undertaken here suggest that that this repression is dynamic during LR development with target expression rising dramatically from LR initiation through to pre-emergence at 36 hours; expression then drops again after 36 hours. This expression changes compliments very strongly those of miR159a and miR159c which strongly suggests at dynamic repression of MYB33 during LR primordium development (figure 6.3.). RT qPCR reflected this rise in expression of MYB33 over the course of LR development albeit in samples of a different type.

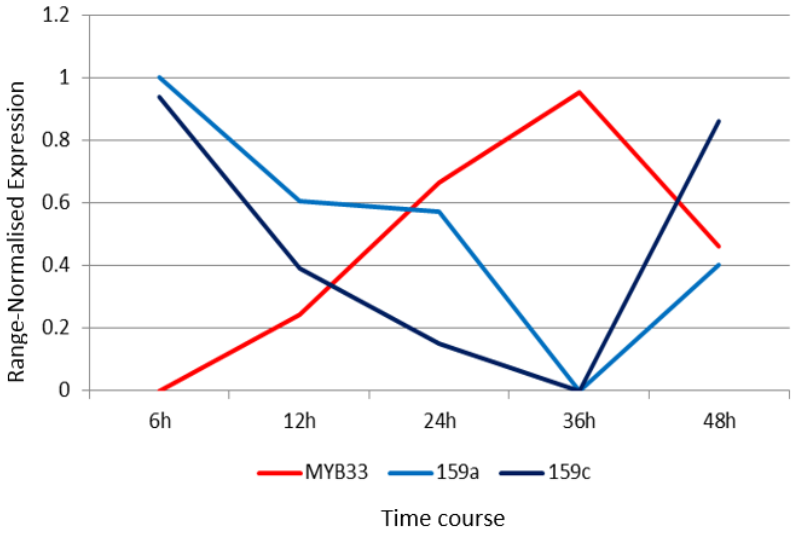


Figure 6.3: *Range normalised RNASeq expression profiles (from normalised counts) of MYB33 during lateral root primordia development alongside expression profiles of miR159 family members.*

MYB33:GUS lines and miRNA159-resistant MYB33:GUS lines revealed how the target is repressed strongly, and ubiquitously by miR159 in the plant including in the developing LR primordia. Hormone treatment also demonstrated that the defect seen in *miR159ab* and MIM159 is upstream of the auxin pathway. However, it does seem that the loss in lateral root numbers can be recovered by gibberellic acid (GA), suggesting a possible involvement of miR159 in GA signalling in the root, presumably via the GA-related MYB targets.

No defect could be found in LR primordia development itself with the primordia which do develop, seemingly developing properly and with the same timing as wild-type (Col-0). This was confirmed with the comparable staging of LR primordia between *miR159ab*, MIM159 and wild-type.

MYB33 is ubiquitously expressed and strongly regulated by miR159 throughout the plant. It is tempting to speculate on the basis of contrary expression of miRNA159 family members and MYB33 during lateral root development that regulation of MYB33 is crucial for LR development. Future work using grafting experiments will clarify the role of MYB33 in LR development.

6.3. NOVEL MICRORNAS

This study has identified and validated ten novel microRNAs expressed during LR development. Unfortunately, overexpression of some of these microRNAs did not result in any LR related developmental defects. However, this is a very standard approach and has been used successfully in many studies (Wu *et al*., 2006, Merchan *et al*., 2009).

As has been mentioned many times now, these microRNAs were, on the whole, amongst the lowest expressed of all miRNAs identified in this study. This is to be expected in novel microRNA identification considering the volume of studies undertaken.

Targets could be predicted, with confidence, for a number of these novel microRNAs. However, no significant lateral root phenotype could be identified amongst those targets for which knockout mutants could be obtained. Study of double, triple, or higher mutant combinations may be required to understand further their role in LR development.

STTM lines or miRNA knockdown lines may also be considered for looking at miRNA/target relationships (perhaps using the CRISPR technique) though the very low expression of these microRNAs in the datasets might suggest that overexpression would be the more effective method by which to identify potential roles.

In the end, this has been the first study of its kind which has investigates the expression profiles of mRNA and small RNAs during LR development and has produced two very strong, well validated datasets of mRNA and small RNA expression during LR development.

CHAPTER 7
REFERENCES

7. REFERENCES

Ágyi A, and Havelda Z. 2013. Analysis of Gradient-like Expression of miR167 in *Arabidopsis thaliana* Embryonic Tissue. Journal of Plant Biology **56:** 336-344.

Allen E, Xie Z, Gustafson AM, and Carrington JC. 2005. MicroRNA-directed phasing during trans-acting siRNA biogenesis in plants. Cell **121:** 207-221.

Allen RS, Li J, Alonso-Peral MM, White RG, Gubler F, and Millar AA. 2010. MicroR159 regulation of most conserved targets in Arabidopsis has negligible phenotypic effects. Silence **1:** 18.

Alonso-Peral MM, Li J, Li Y, Allen RS, Schnippenkoetter W, Ohms S, White RG, and Millar AA. 2010. The miR159 regulated GAMYB-like genes inhibit growth and promote Programmed Cell Death in Arabidopsis. Plant Physiology: DOI:10.1104/pp.110.160630.

Alves-Junior L, Niemeier S, Hauenschild A, Rehmsmeier M, and Merkle T. 2009. Comprehensive prediction of novel microRNA targets in *Arabidopsis thaliana*. Nucleic Acids Research **37 (12):** 4010-4021.

Ambros V, Bartel B, Bartel D, Burge CB, Carrington JC, Chen X, Dreyfuss G, Eddy SR, Griffiths-Jones S, Marshall M, Matzke M, Ruvkun G, and Tuschl T. 2003. A uniform system for microRNA annotation. RNA **9 (3):** 277-279.

Anders S, and Huber W. 2010. Differential expression analysis for sequence count data. Genome Biology **11:** R106.

Aung K, Lin S-I, Wu C-C-, Huang Y-T, Su C-L, and Chiou T-J. 2006. pho2, a Phosphate Overaccumulator, Is Caused by a Nonsense Mutation in a MicroRNA399 Target Gene. Plant Physiology 141: 1000-1011.

Axtell MJ, and Bartel DP. 2005. Antiquity of microRNAs and their targets in land plants. Plant Cell **17:** 1658-1673.

Baima S, Possenti M, Matteucci A, Wisman E, Altamura MM, Ruberti I, and Morelli G. 2001. The Arabidopsis ATHB-8 HD-zip protein acts as a differentiation-promoting transcription factor of the vascular meristems. Plant Physiology **126 (2)**: 643-55.

Baldrich P, Kakar K, Siré C, Moreno AB, Berger A, García-Chapa M, López-Moya JJ, Riechmann JL, and San Segundo B. 2014. Small RNA profiling reveals regulation of Arabidopsis miR168 and heterochromatic siRNA415 in response to fungal elicitors. BMC Genomics **10 (15):** 1083.

Bari R, Pant BD, Stitt M, and Scheible W-R. 2006. PHO2, MicroRNA399, and PHR1 Define a Phosphate-Signalling Pathway in Plants. Plant Physiology **141 (3):** 988-999.

Bartel B, and Bartel DP. 2003. MicroRNAs: at the root of plant development? Plant Physiology **132:** 709-717.

Benkova E, Michniewicz M, Sauer M, Teichmann T, Seifertova D , Jurgens G, and Friml J. 2003. Local, Efflux-Dependent Auxin Gradients as a Common Module for Plant Organ Formation. Cell **115:** 591-602.

Bennett MJ, Marchant A, May ST, and Swarup R. 1998. Going the distance with auxin: unravelling the molecular basis of auxin transport. Philosophical Transactions of the Royal Society B **370 (1664):** 1511-1515.

Blilou I, Xu J, Wildwater M, Willemsen V, Paponov I, Friml J, Heidstra R, Aida M, Palme K, and Scheres B. 2005. The PIN auxin efflux facilitator network controls growth and patterning in Arabidopsis roots. Nature **433 (7021):** 39-44.

Bollman KM, Aukerman MJ, Park MY, Hunter C, Berardini TZ, and Poethig RS. 2003. HASTY, the Arabidopsis ortholog of exportin 5/MSN5, regulates phase change and morphogenesis. Development **130 (8):** 1493-1504.

Breakfield NW, Cocoran DL, Peticka JJ, Shen J, Sae-Seaw J, Rubio-Somoza I, Weigel D, Ohler U, and Benfey PN. 2012. High-Resolution experimental and computational profiling of tissue-specific known and novel miRNAs in *Arabidopsis*. Genome Research **22:** 163-176.

Bustos-Sanmamed P, Mao G, Deng Y, Elouet M, Khan GA, Bazin J, Turner M, Subramanian S, Yu O, Crespi M, and Lelandais-Brière C. 2013. Overexpression of miR160 affects root growth and nitrogen-fixing nodule number in *Medicago truncatula*. Functional Plant Biology **40 (12):** 1208-1220.

Buhtz A, Pieritz J, Springer F, and Kehr J. 2010. Phloem small RNAs, nutrient stress responses, and systemic mobility. BMC Plant Biology **10:** 64.

Carlsbecker A, Lee J-Y, Roberts CJ, Dettmer J, Lehesranta S, Zhou J, LLindgren O, Moreno-Risueno MA, Vaten A, Thitamadee S, Campilho A, Sebastian J, Bowman JL, Helariutte Y, and Benfey PN. 2011. Cell signalling by microRNA165/6 directs gene dose-dependent root cell fate. Nature **465 (7296):** 316-321.

Carthew RW, and Sontheimer EJ .2009. Origins and mechanisms of miRNA and siRNAs. Cell **136:** 642-655.

Casimiro I, Marchant A, Bhalerao RP, Beeckman T, Dhooge D, Swarup R, Graham N, Inzé D, Sandberg G, Pedro J. Casero PJ, and Malcolm Bennett M. 2001. Auxin Transport Promotes Arabidopsis Lateral Root Initiation. The Plant Cell **13:** 843-852.

Casimiro I, Beeckman T, Graham N, Bhalerao R, Zhang H, Casero P, Sandberg G, and Bennett M. 2003. Dissecting *Arabidopsis* lateral root development. Trends in Plant Science **8 (4):** 165-171.

Celenza JL, Grisafi PL, and Fink GR. 1995. A pathway for lateral root formation in *Arabidopsis thaliana*. Genes and Development **9:** 2131-2142.

Cenik ES, and Zamore PD. 2011. Argonaute proteins. Current Biology **21 (12):** 446-449.

Chang K-W, Kao S-Y, Wua Y-H, Tsai M-M, Tu H-F, Liu C-J, Lui M-T, and Lin S-C. 2013. Passenger strand miRNA miR-31⁄ regulates the phenotypes of oral cancer cells by targeting RhoA. Oral Oncology **49:** 27-33.

Chen H, Zhuofu L, and Xiong L. 2012. A plant microRNA regulates the adaptation of roots to drought stress. FEBS Letters **586**: 1742-1747.

Chiou, TJ. 2007. The role of microRNAs in sensing nutrient stress. Plant Cell Environment **30 (3):** 323-332.

Clough SJ, Bent AF. 1998. Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. The Plant Journal **16 (6):** 735-743.

Collins TJ. 2007. ImageJ for microscopy. Bio Techniques **43:** 25-27.

Dai X, and Zhao PX. 2011. psRNATarget: a plant small RNA target analysis server. 2011. Nucleic Acids Research **39:** 155-159.

De Smet I, Vassileva V, Rybel BD, Levesque MP, Grunewald W, Van Damme D, Van Noorden G, Naudts M, Van Isterdael G, De Clercq R, Wang JY, Meuli N, Vanneste S, Friml J, Hilson P, Jürgens G, Ingram GC, Inzé D, Benfey PN, and Beeckman T. 2008. Receptor-Like Kinase ACR4 Restricts Formative Cell Divisions in the Arabidopsis Root. Science **322:** 594-597.

Dharmasiri N, Dharmasiri S, and Estelle M. 2005. The F-box protein TIR1 is an auxin receptor. Nature **435:** 441-445.

Djuranovic S, Nahvi A, and Green R. 2011. A parsimonious model for gene regulation by miRNAs. Science **331:** 550-553.

Dugas DV, and Bartel B. 2008. Sucrose induction of Arabidopsis miR398 represses two Cu/Zn superoxide dismutases. Plant Molecular Biology **67:** 403-417.

K Edwards K, Johnstone C, and Thompson C. 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucleic Acids Research **19 (6):** 1349.

Endres MW, Gregory BD, Gao Z, Foreman AW, and Mlotshwa S. 2010. Two plant viral suppressors of silencing require the ethylene-inducible host transcription factor RAV2 to block RNA silencing. PLoS Pathogy **6:** e1000729.

Fahlgren N, Montgomery TA, Howell MD, Allen E, Dvorak SK, Alexander AL,1 and James C. Carrington JC. 2006. Regulation of AUXIN RESPONSE FACTOR3 by TAS3 ta-siRNA Affects Developmental Timing and Patterning in Arabidopsis. Current Biology **16:** 939-944.

Fire A, Xu SQ, Montgomery MK, Kostas SA, Driver SE, and Mello CC. 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans.* Nature **39 (6669):** 806-811.

Franco-Zorrilla JM, Valli A, Todesco M, Mateos I, Puga MI, Rubio-Somoza I, Leyva A, Weigel D, Garcı JA and Paz-Ares J. 2007. Target mimicry provides a new mechanism for regulation of microRNA activity. Nature Genetics **39:** 1033-1037.

Frank G, Pressman E, Ophir R, Althan L, Shaked R, Freedman M, Shen S, and Firon N. 2009. Transcriptional profiling of maturing tomato (*Solanum lycopersicum L.)* microspores reveals the involvement of heat shock proteins, ROS scavengers, hormones, and sugars in the heat stress response. Journal of Experimental Botany **60:** 3891–3908.

Friedlander MR, Chen W, Adamidi C, Maaskola J, Einspanier R, Knespel S, and Rajewsky N. 2008. Discovering microRNAs from deep sequencing data using miRDeep.Nat. Biotechnology **26:** 407-415

Friml, J., Wiśniewska, J., Benková, E., Mendgen, K., and Palme, K. 2002. Lateral relocation of auxin efflux regulator PIN3 mediates tropism in *Arabidopsis*. Nature **415:** 806-9.

Fukaki H, Tameda S, Masuda H, and Tasaka M. 2002. Lateral root formation is blocked by a gain-of-function mutation in the SOLITARY-ROOT/IAA14 gene of Arabidopsis. The Plant Journal **29 (2):** 153-168.

Fukaki H, Okushima Y, and Tasaka M. 2007. Auxin-Mediated Lateral Root Formation in Higher Plants. International Review of Cytology **256:** 111-137.

Gan X, Stegle O, Behr J, Steffen JG, Drewe P, Hildebrand L, Lyngsoe R, Schultheiss SJ, Osborne EJ, Sreedharan VT, Kahles A, Bohnert R, Jean G, Derwent P, Kersey P, Belfield EJ, Harberd NP, Kemen E, Toomajian C, Kover PX, Clark RM, Rätsch G, and Mott R. 2011.

Multiple reference genomes and transcriptomes for *Arabidopsis thaliana*. Nature **477:** 419-423.

Gifford ML, Dean A, Gutierrez RA, Coruzzi GM, and Birnbaum KD. 2008. Cell-specific nitrogen responses mediate developmental plasticity. Proceedings of the National Academy of Science **105 (2):** 803-808.

Gocal GF, Sheldon CC, Gubler F, Moritz T, Bagnall DJ, MacMillan CP, Li SF, Parish RW, Dennis ES, Weigel D, and King RW. 2001. GAMYB-like genes, flowering, and gibberellin signalling in Arabidopsis. Plant Physiology **127:** 1682–1693.

Grennan AK. 2008. *Arabidopsis* MicroRNAs. Plant Physiology 146 **(1):** 3-4.

Grigg SP, Galinha C, Kornet N, Canales C, Scheres, and Tsiantis M. 2009. Repression of Apical Homeobox Genes Is Required for Embryonic Root Development in Arabidopsis. Current Biology **19:** 1485-1490.

Guilfoyle TJ, and Hagen G. 2007. Auxin response factors. Current Opinion Plant Biology **10 (5):** 453-460.

Guo H-S, Xie Q, Fei J-F, and Chua N-H. 2005. MicroRNA directs mRNA cleavage of the transcription factor NAC1 to down-regulate Auxin signals for *Arabidopsis* lateral root development. The Plant Cell **17:** 1376-1386.

Hammond SM, Bernstein E, Beach D, and Hannon GJ. 2000. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. Nature **404:** 293-296.

Harvey JW, Lewsey MG, Patel K, Westwood J, Heimstadt S, Carr JP, and Baulcombe DC. 2011. An Antiviral Defense Role of AGO2 in Plants. PLOS One **6 (1):** e14639.

Hatzfeld Y, Lee S, Lee M, Leustek T, and Saito K. 2000. Functional characterization of a gene encoding a fourth ATP sulfurylase isoform from *Arabidopsis thaliana*. Gene **248:** 51–58.

Hawker NP, and Bowman JL. 2004. Roles for Class III HD-Zip and KANADI Genes in Arabidopsis Root Development. Plant Physiology **135:** 2261–2270.

Hijazi M, Durand J, Pichereaux C, Pont F, Jamet E, and Albenne C. 2012. Characterization of the arabinogalactan protein 31 (AGP31) of *Arabidopsis thaliana*: new advances on the Hyp-O-glycosylation of the Pro-rich domain. Journal of Biological Chemistry **287 (12):** 9623-32.

Himanen K, Boucheron E, Vanneste S, Engler JA, Inzé D, and Beeckman T. 2002. Auxin-Mediated Cell Cycle Activation during Early Lateral Root Initiation. The Plant Cell **14:** 2339-2351.

Himanen K, Vuylsteke M, Vanneste S, Vercruysse S, Boucheron E, Alard P, Chriqui D, Van Montagu M, Inze´ D, and Beeckman T. 2004. Transcript profiling of early lateral root initiation. PNAS **101 (14):** 5146-5151.

Huntzinger E, and Izaurralde E. 2011. Gene silencing by microRNAs: contributions of translational repression and mRNA decay. Nature Reviews Genetics **12:** 99-110.

Jia F, Rock CD. 2013. MIR846 and MIR842 comprise a cistronic MIRNA pair that is

regulated by abscisic acid by alternative splicing in roots of Arabidopsis. Plant Molecular Biology 81 **(4-5):** 447-460.

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, and Charpentier E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science **337 (6096):** 816-821.

Jones-Rhoades MW, and Bartel DP. 2004. Computational identification of plant microRNAs and their targets including a stress-induced miRNA. Molecular Cell **14:** 787-799.

Jones-Rhoades MW, Bartel DP, and Bartel B. 2006. MicroRNAs and their regulatory roles in plants. Annual Review of Plant Biology **57**: 19-53.

Kant S, Peng M, and Rothstein SJ. 2011. Genetic Regulation by NLA and MicroRNA827 for Maintaining Nitrate-Dependent Phosphate Homeostasis in Arabidopsis. PLOS Genetics **7 (3):** e1002021.

Karimi M, Inze D, and Depicker A. 2002. GATEWAY vectors for Agrobacterium mediated plant transformation. Trends in Plant Sciences **7:** 193–195.

Kawashima CG, Yoshimoto N, Maruyama-Nakashita A, Tsuchiya YN, Saito K, Takahashi H, and Dalmay T. 2009. Sulphur starvation induces the expression of microRNA-395and one of its target genes but in different cell types. The Plant Journal **57:** 313–321.

Khan GA, Declerck M, Sorin C, Hartmann C, Crespi M, and Lelandais-Briere C. 2011. MicroRNAs as regulators of root development and architecture. Plant Molecular Biology **77:** 47-58.

Kinoshita N, Wang H, Kasahara H, Liu J, MacPherson C, Machida Y, Kamiya Y, Hannah MA, and Chuaa N-H. 2012. AA-Ala Resistant3, an Evolutionarily Conserved Target of miR167, Mediates Arabidopsis Root Architecture Changes during High Osmotic Stress. The Plant Cell.

Koussevitzky S, Suzuki N, Huntington S, Armijo L, Sha W, Cortes D, Shulaev V, and Mittler R. 2008. Ascorbate Peroxidase 1 Plays a Key Role in the Response of *Arabidopsis thaliana* to Stress Combination. Journal of Biological Chemistry **283:** 34197-34203.

Křeček P, Skůpa P, Libus J, Naramoto S, Tejos R, Friml J, and Zažímalová E. 2009. The PIN-FORMED (PIN) protein family of auxin transporters. Genome Biology **10:** 249.

Kurihara Y, Takashi Y, and Watanabe Y. 2006. The interaction between DCL1 and HCY1 is important for efficient and precise processing of pri-miRNA in plant microRNA biogenesis. RNA **12:** 206-212.

Kurihara Y, Takashi Y and Watanabe Y. 2006. The interaction between DCL1 and HYL1 is important for efficient and precise processing of pri-miRNA in plant microRNA biogenesis. RNA **12:** 206-212.

Lagos-Quintana M, Rauhut R, Lendeckel W, and Tuschl T. 2001. Identification of novel genes coding for small expressed RNAs. Science **294:** 853-858.

Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, Muller R, Dreher K, Alexander DV, Garcia-Hernandez M, Karthikeyan AS, Lee CH, Nelson WD, Ploetz L, Singh S,

Wensel A, and Huala E. 2012. The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic Acids Research **40:** 1202-1210.

Langmead B, Trapnell C, Pop M, and Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biology **10:** R25

Laskowski MJ, Williams ME, Nusbaum HC, and Sussex IM. 1995. Formation of lateral root meristems is a two-stage process. Development **121 (10):** 3303-3310

Lau NC, Lim LP, Weinstein EG, and Bartel DP. 2001. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. Science **294 (5543):** 858-862.

Lee RC, Feinbaum RL, and Ambros V. 1993. The *C. elegans* Heterochronic Gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell **75:** 843-854.

Lee RC, and Ambros V. 2001. An extensive class of small RNAs in *Caenorhabditis elegans*. Science **294:** 862-869.

Lee H, Yoo SJ, Lee JH, Kim W, Yoo SK, Fitzgerald H, Carrington JC, and Ahn JH. 2010. Genetic framework for flowering-time regulation by ambient temperature-responsive miRNAs in Arabidopsis. Nucleic Acids research **38:** 3081–3093.

Leustek T, and Saito K. 1999. Sulfate transport and assimilation in plant. Plant Physiology **116:** 1307–1313

Lewsey MG, Murphy AM, MacLean D, Dalchau N, Westwood JH, Macauley K, Bennett MH, Moulin M, Hanke DE, Powell G, Smith AG, and Carr JP. 2010. Disruption of Two Defensive Signaling Pathways by a Viral RNA Silencing Suppressor. MPMI **23 (7):** 835-845.

Li J, Yang Z, Yu B, Liu J, and Chen X. 2005. Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in *Arabidopsis*. Current Biology **15:** 1501-1507.

Li Y, Fu Y, Ji L, Wu C, and Zheng C. 2010. Characterization and expression analysis of the Arabidopsis mir169 family. Plant Science **178:** 271–280.

Liang G, Ynag F, and Yu D. 2010. MicroRNA395 mediates regulation of sulfate accumulation and allocation in *Arabidopsis thaliana.* The Plant Journal **62:** 1046–1057.

Lin S-I, Chiang S-F, Lin W-Y, Chen J-W, Tseng C-Y, Wu P-C, and Chiou T-J. 2008. Regulatory Network of MicroRNA399 and PHO2 by Systemic Signaling. Plant Physiology **147:** 732-746.

Liu C, and Mehdy MC. 2007. A Nonclassical Arabinogalactan Protein Gene Highly Expressed in Vascular Tissues, AGP31, Is Transcriptionally Repressed by Methyl Jasmonic Acid in Arabidopsis. Plant Physiology **145 (3):** 863-874.

Liu PP, Montgomery TA, Falgrehn N, Kasschau KD, Nonogaki H, and Carrington JC. 2007. Repression of AUXIN RESPONSE FACTOR 10 by microRNA160 is crucial for seed germination and post-germination stages. Plant Journal **52 (1):** 133-146.

Llave C, Xie Z, Kasschau KD, and Carrington J. 2002. Cleavage of SCARECROW-like mRNA Targets Directed by a Class of *Arabidopsis* miRNA. Science **297 (20):** 2053-2056.

Lobbes D, Rallapalli G, Schmidt DD, Martin C, and Clarke J. 2006. SERRATE: a new player on the planet microRNA scene. European Molecular Biology Organization Reports **7:** 1052-1058.

Long JA, Woody S, Poethig S, Meyerowitz EM, and Barton MK. 2002. Transformation of shoots into roots in Arabidopsis embryos mutant at the TOPLESS locus. Development **129 (12):** 2797-806.

Lorenz R, Bernhart SH, Siederdissen CHz, Tafer H, Flamm C, Stadler PF, and Hofacker IL. 2011. ViennaRNA Package 2.0. Algorithms for Molecular Biology **6:** 26.

Lucas M, Godin C, Jay-Allemand C, and Laplaze L. 2008. Auxin fluxes in the root apex co-regulate gravitropism and lateral root initiation. Journal of Experimental Biology **59:** 55-66.

Malamy JE, and Benfey PN. 1997. Organization and cell differentiation in lateral roots of *Arabidopsis thaliana.* Development **124:** 33-44.

Mallory AC, Bartel DP, and Bartel B. 2005. MicroRNA-directed regulation of *Arabidopsis* AUXIN RESPONSE FACTOR 17 is essential for proper development and modulates expression of early auxin response genes. Plant Cell **17:** 1360-1375.

Maraschin Fdos S, Memelink J, and Offringa R. 2009. Auxin-induced, SCF(TIR1)-mediated poly-ubiquitination marks AUX/IAA proteins for degradation. Plant Journal **59 (1):** 100-109.

Marchant A, Bhalerao R, Casimiro I, Eklöf J, Casero PJ, Bennett M, and Sandbergb G. 2002. AUX1 Promotes Lateral Root Formation by Facilitating Indole-3-Acetic Acid Distribution between Sink and Source Tissues in the Arabidopsis Seedling. The Plant Cell **14:** 589-597.

Marin E, Jouannet V, Herz A, Lokerse AS, Weijers D, Vaucheret H, Nussaume L, Crespi MD, and Maizel A. 2010. miR390, Arabidopsis TAS3 tasiRNAs, and their AUXIN RESPONSE FACTOR targets define an autoregulatory network quantitatively regulating lateral root growth. The Plant Cell **22:** 1104-1117.

Meister G, and Tuschl T. 2004. Mechanisms of gene silencing by double-stranded RNA. Nature **431:** 343-349.

Melnyk CW, Molnar A, Bassett A, and Baulcome DC. 2011. Mobile 24 nt small RNAs direct transcriptional gene silencing in the root meristems of *Arabidopsis thaliana.* Current Biology **21 (19):** 1678-1683.

Meng Y, Shao C, and Chen M. 2011. Toward microRNA-mediated gene regulatory networks in plants. Briefings in Bioinformatics **12 (6):** 645-659.

Merchan F, Boualem A, Crespi M, and Frugier F. 2009. Plant polycistronic precursors containing non-homologous microRNAs target transcripts encoding functionally related proteins. Genome Biology **10 (12):** R136.

Millar AA, & Gublera. 2005. The Arabidopsis GAMYB-Like Genes, MYB33 and MYB65, Are MicroRNA-Regulated Genes That Redundantly Facilitate Anther Development. The Plant Cell **17:** 705-721.

Miyashima S, Koi S, Hashimoto T, and Nakajima K. 2011. Non-cell-autonomous microRNA165 acts in a dose dependent manner to regulate multiple differentiation status in the Arabidopsis root. Development **138:** 2303-2313.

Montgomery TA, Howell MD, Cuperus JT, Li D, Hansen JE, Alexander AL, Chapman EJ, Fahlgren N, Allen E, and Carrington JC. 2008. Specificity of ARGONAUTE7-miR390 Interaction and Dual Functionality in TAS3 Trans-Acting siRNA Formation. Cell **133:** 128-141.

Myouga F, Hosoda C, Umezawa T, Iizumi H, Kuromori T, Motohashi R, Shono Y, Nagata N, Ikeuchi M, and Shinozakia K. 2008. A Heterocomplex of Iron Superoxide Dismutases

Defends Chloroplast Nucleoids against Oxidative Stress and Is Essential for Chloroplast Development in Arabidopsis. Plant Cell **20 (11):** 3148–3162.

Nagpal P, Walker LM, Young JC, Sonawala A, Timpte C, Estelle M, and Reed JW. 2000. AXR2 encodes a member of the Aux/IAA protein family. Plant Physiology **123:** 563–574.

Nakagawa T, Kurosea T, Hinoa T, Katsunori T, Kawamukaib M, Niwac Y, Toyookad K, Matsuokad K, Jinbof T, and Kimuraf T. 2007. Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. Journal of Bioscience and Bioengineering **104 (1):** 34-41

Navarro L, Dunoyer P, Jay F, Arnold B, Dharmasiri M, Estelle O, Voinnet JD, and Jones A. 2006. A plant miRNA contributes to antibacterial resistance by repressing auxin signalling. Science **312:** 436-439.

Naya L, Khan GA, Sorin C, Hartmann C, Crespi M, and Lelandais-Brière C. 2010. Cleavage of a non-conserved target by a specific miR156 isoform in root apexes of *Medicago truncatula*. Plant Signal Behaviour **5 (3)**: 328-331.

Okushima Y, Overvoorde PJ, Arima K, Alonso JM, Chan A, Chang C, Ecker JR, Hughes B, Lui L, Nguyen D, Onodera C, Quach H, Smith A, Yu G, and Theologisa A. 2005. Functional

Genomic Analysis of the AUXIN RESPONSE FACTOR Gene Family Members in *Arabidopsis thaliana*: Unique and Overlapping Functions of ARF7 and ARF19. The Plant Cell **17:** 444-463.

Okushima Y, Fukaki H, Onoda M, Theologis A, and Tasaka M. 2007. ARF7 and ARF19 regulate lateral root formation via direct activation of LBD/ASL genes in *Arabidopsis*. The Plant Cell **19:** 118-130 (2007).

Padmanabhan MS, Ma S, Burch-Smith TM, Czymmek K, Huijser P, and Dinesh-Kumar SP. 2013. Novel Positive Regulatory Role for the SPL6 Transcription Factor in the N TIR-NB-LRR Receptor-Mediated Plant Innate Immunity. PLOS Pathogens **9 (3):** e1003235.

Palatnik JF, Wollmann H, Schommer C, Schwab R, Boisbouvier J, Rodriguez R, Warthmann N, Allen E, Dezulian T, Huson D, Carrington JC, and Weigel D. 2007. Sequence and Expression Differences Underlie Functional Specialization of Arabidopsis MicroRNAs miR159 and miR319. Developmental Cell **13:** 115-125.

Pant BD, Musialak-Lange M, Nuc P, May P, Buhtz A, Kehr J, Walther D, and Scheible WR. 2009. Identification of nutrient-responsive Arabidopsis and rapeseed microRNAs by comprehensive real-time polymerase chain reaction profiling and small RNA sequencing. Plant physiology **150:** 1541–1555.

Paponov IA, William D. Teale WD, Trebar M, Blilou I, and Palme K. 2005. The PIN auxin efflux facilitators: evolutionary and functional perspectives, Trends in Plant Sciences **10 (4):** 170-177.

Park W, Li J, Song R, Messing J, and Chen X. 2002. CARPEL FACTORY, a Dicer Homolog, and HEN1, a Novel Protein, Act in microRNA Metabolism in *Arabidopsis thaliana*. Current Biology **12:** 1484-1495.

Park MY, Gonzales-Sulser A, Vaucheret H, and Poethig RS. 2005. Nuclear processing and export of microRNAs in *Arabidopsis*. Proceedings of the National Academy of Science 102 **(10):** 3691-3696.

Parry G, Delbarre D, Marchant A, Swarup R, Napier R, Perrot-Rechenmann C, and Bennett NJ. 2001. Novel auxin transport inhibitors phenocopy the auxin influx carrier mutation *aux1*. The Plant Journal **25 (4):** 399-406.

Parry G, Calderon-Villalobo L-I, Prigge M, Peret B, Dharmasiri S, Itoh H, Lechner E, Gray WM, Bennett M, and Estelle M. 2009. Complex regulation of the TIR1/AFB family of auxin receptors. PNAS **106 (52):** 22540-22545.

Pasquinelli A E, & Ruvkun G. 2002. Control of developmental timing by microRNAs and their targets. Annual Review of Cell Developmental Biology **18:** 495-513.

Pearson K. 1895. Notes on regression and inheritance in the case of two parents. Proceedings of the Royal Society of London **58:** 240–242.

Peret B, De Rybel B, Casimiro I, Benkova E, Swarup R, Laplaxe L, Beeckman T, and Bennett MJ. 2009. *Arabidopsis* lateral root development: an emerging story. Cell Press **14 (7):** 399-408.

Piriyapongsa J, Marino-Ramirez L, and Jordan IK. 2007. Origin and evolution of human microRNAs from transposable elements. Genetics **176:** 1323-1337.

Piriyapongsa J, and Jordan IK. 2008. Dual coding of siRNAs and miRNAs by plant transposable elements. RNA **14:** 814-821.

Popov VN, Eprintsev AT, Fedorin DN, and Igamberdiev AU. 2010. Succinate dehydrogenase in *Arabidopsis thaliana* is regulated by light via phytochrome A. FEBS Letters **584 (1):** 199-202.

Qi Y, Denli AM, and Hannon GJ. 2005. Biochemical specialization within Arabidopsis RNA silencing pathways. Molecular Cell Biology **19:** 421-428.

Qu F, Ye X, and Morris TJ. 2008. Arabidopsis DRB4, AGO1, AGO7, and RDR6 participate in a DCL4-initiated antiviral RNA silencing pathway negatively regulated by DCL1. Proceedings of the National Acadademy of Science; 105 **(38):** 14732–7.

Raman S, Greb T, Peaucelle A, Blein T, Laufs P, and Theres K. 2008. Interplay of miR164, CUP-SHAPED COTYLEDON genes and LATERAL SUPPRESSOR controls axillary meristem formation in *Arabidopsis thaliana*. The Plant Journal **55:** 65-76.

Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, Rougvie AE, Horvitz R, and Ruvkun. 2000. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. Nature **403:** 901-906.

Reyes JL, and Chua N-H. 2007. ABA induction of miR159 controls transcript levels of two MYB factors during Arabidopsis seed germination. The Plant Journal **49:** 592-606.

Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, and Bartel DP. 2002. Prediction of plant microRNA targets. Cell **110:** 513-520.

Robertson EJ, Rutherford SM, and Leech RM. 1996. Characterization of chloroplast division using the Arabidopsis mutant arc5. Plant Physiology **112 (1):** 149-159.

Rogg LE, Lasswell J, and Bartel B. 2001. A gain-of function mutation in IAA28 suppresses lateral root development. Plant Cell **13:** 465–480.

Rubio-Somoza I, and Weigel D. 2011. MicroRNA networks and developmental plasticity in plants. Trends in Plant Science **16 (5):** 258-264.

Ruby G, Jan CH, and Bartel DP. 2007. Intronic microRNA precursors that bypass Drosha processing. Nature **448 (7149):** 83-86.

Saito K. 2004. Sulfur assimilatory metabolism. The long and smelling road. Plant Physiology **136:** 2443–2450.

Saito K, Kazumichi M, Tomoko M, Kawamura Y, Miyoshi K, Nagami T, Siomi H, and Siomi MC. 2006. Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. Genes and Development **20:** 2214-2222.

Saumet A and Lecellier C-H. 2006. Anti-viral RNA silencing: do we look like plants? Retrovirology **3:** 3.

Schauer SE, Jacobsen SE, Meinket DW, and Ray A. 2002. DICER-LIKE1: blind men and elephants in Arabidopsis development. Trends in Plant Science 7 **(11):** 487-491.

Schwab R, Palatnik JF, Reister M, Schommer C, Schmid M, and Weigel D. 2005. Specific Effects of MicroRNAs on the Plant Transcriptome. Developmental Cell **8:** 517-527.

Shikata M, Koyama T, Mitsuda N, and Ohme-Takagi M. 2009. Arabidopsis SBP-box genes SPL10, SPL11 and SPL2 control morphological change in association with shoot maturation in the reproductive phase. Plant Cell Physiology **50 (12):** 2133-2145.

.

Smith S, and De Smet I. 2012. Root system architecture: insights from Arabidopsis and cereal crops. Philosophical Transactions of the Royal Society **367:** 1441-1452.

Song JJ, Smith SK, Hannon GJ, and Joshua-Tor L. 2004. Crystal structure of Argonaute and its implications for RISC slicer activity. Science **305:** 1434-1437.

Sorin C, Declerck M, Christ A, Blein T, Ma L, Lelandais-Briere C, Njo MF, Beeckman T, Crespi M, and Hartmann C. 2014. A miR169 isoform regulates specific NF-YA targets and root architecture in Arabidopsis. **New Phytologist 202:** 1197-1211.

Sozzani R, Cui H, Moreno-Risueno MA, Busch W, Van Norman JM, Vernoux T, Brady SM, Dewitte W, Murray JA, and Benfey PN. 2010. Spatiotemporal regulation of cell-cycle genes by SHORTROOT links patterning and growth. **Nature 466 (7302):**128-32.

Stocks, MB, Moxon S, Mapleson D, Woolfenden HC, Mohorianu I, Folkes L, Schwach F, Dalmay T and Moulton V. 2012. The UEA sRNA workbench: a suite of tools for analysing and visualizing next generation sequencing microRNA and small RNA datasets. Bioinformatics **28:** 2059-2061.

Suzuki N, Miller G, Sejima H, Harper J, and Mittler R. 2013. Enhanced seed production under prolonged heat stress conditions in *Arabidopsis thaliana* plants deficient in cytosolic ascorbate peroxidase 2. Journal of Experimental Botany **64 (1):** 253-263.

Swarup K, Benkova E, Swarup R, Casimiro I, Peret B, Yang Y, Parry G, Nielson E, De Smet I, Vanneste S, Levesque MP, Carrier D, James N, Calvo V, Ljung K, Kramer E, Roberts R, Graham N, Marillonnet S, Patel K, Jones JDG, Taylor CG, Schachtman DP, May S, Sandberg G, Benfey P, Friml J, Kerr I, Beeckman T, Laplaze L, and Bennet MJ. 2008. The Auxin influx carrier LAX3 promotes lateral root emergence. Nature Cell Biology **10 (8):** 946-954.

Swarup R, and Peret B. 2012. AUX/LAX family of auxin influx carriers – an overview. Frontiers in Plant Sciences **3:** Article 225.

Szemenyei H, Hannon M, and Long JA. 2008. TOPLESS mediates auxin-dependent transcriptional repression during Arabidopsis embryogenesis. Science **319 (5868):** 1384-1386.

Tang G, Yan J, Gu Y, Qiao M, Fan R, Mao Y, and Tang X. 2012. Construction of short tandem target mimic (STTM) to block the functions of plant and animal microRNAs. Methods **58 (2):** 118-125.

Tian Q, Nagpal P, and Reed JW. 2003. Regulation of Arabidopsis SHY2/IAA3 protein turnover. Plant Journal **36:** 643-651.

Todesco M, Rubio-Somoza I, Paz-Ares, and Weigel D. 2010. A collection of target mimics for comprehensive analysis on microRNA function in *Arabidopsis thaliana*. PLoS Genetics **6 (7):** e1001031.

Tomari Y, and Zamore PD. 2005. Perspective: machines for RNAi. Genes Development **19:** 517-529.

Trapnell C, Patcher L, and Salzberg S. 2009. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics **25 (9):** 1105-1111.

Trapnell C, Robert A, Goff L, Pertea G, Kim D, Kelley DR, Pimtel H, Salzberg SL, Rinn JL, and Patcher L. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature Protocols **7 (3):** 562-578.

Turner M, Nizampatnam NR, Baron M, Coppin S, Damodaran S, Adhikari S, Arunacalam SP, Yu O, and Subramanian S. 2013. Ectopic Expression of miR160 Results in Auxin Hypersensitivity, Cytokinin Hyposensitivity, and Inhibition of Symbiotic Nodule Development in Soybean. Plant Physiology **162:** 2042-2055.

Vanneste S, Rybel BD, Beemster GTS, Ljung K, Smet ID, Isterdael GV, Naudts M, Iida R, Gruissem W, Tasaka M, Inze´ D , Fukaki H, and Beeckmanan T. 2005. Cell Cycle Progression in the Pericycle Is Not Sufficient for SOLITARY ROOT/IAA14-Mediated Lateral Root Initiation in *Arabidopsis thaliana*. The Plant Cell **17:** 3035-3050.

Vanneste S, and Friml J. 2012. Plant Signalling: Deconstructing auxin sensing. Nature Chemical Biology **8:** 415-416.

Vaucheret H. 2008. Plant ARGONAUTES. Trends in Plant Science **13 (7):** 350-358.

Varkonyi-Gasic E, Wu R, Wood M, Walton EF, and Hellens RP. 2007. Protocol: a highly sensitive RT-PCR method for detection and quantification of microRNAs. Plant Methods **3:** 12.

Vazquez F, Gasciolli V, Crete P, and Vaucheret H. 2004. The nuclear dsRNA binding protein HYL1 is required for microRNA accumulation and plant development, but not posttranscriptional transgene silencing. Current Biology **14:** 346-351.

Vidal EA, Araus V, Lu C, Parry G, Green PJ, Coruzzi GM, and Gutierrez RA. 2010. Nitrate-responsive miR393/AFB3 regulatory module controls root system architecture in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences U.S.A **107:** 4477-4482.

Vidal EA, Moyano TC, Riveras E, Contreras-López O, and Gutiérrez RA. 2013. Systems approaches map regulatory networks downstream of the auxin receptor AFB3 in the nitrate response of *Arabidopsis thaliana* roots. Proceedings of the National Academy of Science **110 (31):** 12840-12845.

Voß U, Wilson MH, Kenobi K, Gould PD, Robertson FC, Lucas M, Swarup K, Casimiro I, Holman TJ, Wells DM, Péret B, Goh T, Fukaki H, Hodgman C, Laplaze L, Halliday KJ, Ljung K, Hall AJ, Webb AAR, and Bennett MJ. 2015. Lateral root organ initiation re-phases the circadian clock in *Arabidopsis thaliana*. Nature Communications **in review**

Wang J-W, Wang L-J, Mao Y-B, Cai W-J, Xue H-W, and Chen X-Y. 2005. Control of root cap formation by MicroRNA-targeted Auxin response factors in *Arabidopsis*. Plant Cell **17 (8):** 2204-2216.

Wang S, Kwak S-H, Zeng Q, Ellia BE, Chen X-Y, Schiefelbein J, and Chen J-G. 2007. TRICHOMELESS1 regulates trichome patterning by suppressing GLABRA1 in Arabidopsis. Development **134:** 2873-3882.

Wang JW, Czech B, and Weigel D. 2009. MiR156-regulated SPL transcription factors define an endogenous flowering pathway in *Arabidopsis thaliana*. Cell **138 (4):** 738-749.

Wu MF, Tian Q, and Reed JW. 2006. *Arabidopsis* microRNA167 controls patterns of ARF6 and ARF8 expression, and regulates both female and male reproduction. Development **133:** 4211-4218.

Wu G, and Poethig RS. 2006a. Temporal regulation of shoot development in *Arabidopsis thaliana* by miR156 and its target SPL3. Development **133**: 3539–3547.

Wu G, Park MY, Conway SR, Wang J-W, Weigel D, and Poethig RS. 2009. The sequential action of miR156 and miR172 regulates developmental timing in Arabidopsis. Cell **138 (4):** 750-759.

Xie Q, Frugis G, Colgan D, and Chua NH. 2000. *Arabidopsis* NAC1 transduces auxin signal downstream of TIR1 to promote lateral root development. Genes Development **14:** 3024-3036.

Xie Z, Allen E, Fahlgren N, Calmar A, Givan SA, and Carrington JC. 2005. Expression of *Arabidopsis* MIRNA genes. Plant Physiology **138:** 2145-2154.

Xie Z, Kasschau KD, and Carrington JC. 2003. Negative Feedback Regulation of Dicer-Like1

in Arabidopsis by microRNA-Guided mRNA Degradation. Current Biology **13:** 784-789.

Xing S, Salinas M, Höhmann S, Berndtgen R, AND Huijser P. 2010. MiR156-targeted and non-targeted SBP-box transcription factors act in concert to secure male fertility in Arabidopsis. Plant Cell **22 (12):** 3935-3950.

Xing S, Salinas, M, Garcia-Molina A, Hohmann S, Berndtgen R and Huijser P. 2013. SPL8 and miR156-targeted SPL genes redundantly regulate Arabidopsis gynoecium differential patterning. The Plant Journal: doi: 10.1111/tpj.12221.

Yamaguchi A, Wu MF, Yang L, Wu G, Poethig RS, and Wagner D. 2009. The microRNA-regulated SBP-Box transcription factor SPL3 is a direct upstream activator of LEAFY, FRUITFULL, and APETALA1. Developmental Cell **17:** 268-278.

Yan KS, Yan S, Farooq A, Han A, Zeng L, and Zhou M-M. 2003. Structure and conserved RNA binding of the PAZ domain. Nature **426**: 469-474.

Yan Z, Hossain MS, Valdés-López O, Hoang NT, Zhai J, Wang J, Libault M, Brechenmacher L, Findley S, Joshi T, Qui L, Sherrier J, Ji T, Meyers BC, Xu D, and Stacey G. 2015. Identification and functional characterization of soybean root hair microRNAs expressed in response to *Bradyrhizobium japonicum* infection. Plant Biotechnology Journal. **doi: 10.1111/pbi.12387.**

Yang GD, Yan K, Wu BJ, Wang YH, Cao YX, and Zheng CC. 2012. Genome-wide analysis of intronic microRNAs in rice and Arabidopsis. Journal of Genetics **91 (3):** 313-324.

Yang D, Sun Y, Hu L, Zheng H, Ji P, Pecot CV, Zhao Y, Reynolds S, Cheng H, Rupaimoole R, Cogdell D, Nykter M, Broaddus R, Rodriguez-Aguayo C, Lopez-Berestein G, Liu J, Shmulevich I, Sood AK, Chen K, and Zhang W. 2013. Integrated Analyses Identify a Master MicroRNA Regulatory Network for the Mesenchymal Subtype in Serous Ovarian Cancer. Cancer Cell **23:** 186-199.

Yoda M, Kawamata T, Paroo Z, Ye X, Iwasaki S, Liu Q, and Tomari Y. 2010. ATP-dependant human RISC assembly pathways. Nature Structures Molecular Biology **17**: 17-23.

Yoon EK, Yang JH, Lim J , Kim SH, Kim S-K, and Lee WS. 2010. Auxin regulation of the microRNA390-dependent transacting small interfering RNA pathway in Arabidopsis lateral root development. Nucleic Acids Research **38 (4):** 1382-1391.

Yoshimoto N, Takahashi H, Smith FW, Yamaya T, and Saito K. 2002. Two distinct high-affinity sulfate transporters with different inducibilities mediate uptake of sulfate in Arabidopsis roots. Plant Journal **29:** 465–473.

Yu N, Cai W-J, Wang S, Shan C-M, Wang L-J, and Chen X-Y. 2010. Temporal Control of Trichome Distribution by MicroRNA156-Targeted SPL Genes in *Arabidopsis thaliana*. The Plant Cell **22:** 2322-2335.

Yu N, Nui Q-W, Ng K-H, and Chua N-H. 2015. The role of miR156/SPLs modules in Arabidopsis lateral root development. The Plant Journal **83:** 673-685.

Zhang Y. 2005. miRU: an automated plant miRNA target prediction server. Nucleic Acids Research **33:** 701–704.

Zhao Y, Dai Z, Liang Y, Yin M, Ma K, He M, Ouyang H, and Tenga C-B. 2014. Sequence-specific inhibition of microRNA via CRISPR/CRISPRi system. Science Reports **4:** 3943.

Zhong R, and Ye Z-H. 2007. Regulation of HD-ZIP III Genes by MicroRNA 165. Plant Signalling Behaviour **2 (5):** 351-353.

Zhou X, Wang G, Sutoh K, Zhu JK, and Zhang W. 2008. Identification of cold-inducible microRNAs in plants by transcriptome analysis. Biochimica et Biophysica Acta **1779:** 780–788.

Zimmermann P, Hirsch-Hoffmann M, Hennig L, and Gruissem W. 2004. GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. Plant Physiology **136 (1):** 2621-2632.

APPENDICES

A1. PRIMER DETAILS

Details of all Primers used in this study:

A.1.1. PRIMERS FOR GENOTYPING MIRNA TARGET MIMIC LINES

Reverse:

MIM408	ATGCACTGCCAAGCCTTCCCTGGC
MIM163	TTGAAGAGGATTGATTGGAACTTCGAT
MIM775	TTCGATGTCTTCGTGCAGTGCCAA
MIM157	TTGACAGAAGTCGATAGAGAGCAC
MIM159	TTTGGATTGATTGATTGAGGGAGCTCTT
MIM160	GGCATACAGTCTAGAGCCAGGCA
MIM164	TGGAGAAGCATAGGGGCACGTGCA
MIM172	AGAATCTTGAACTCGATGCTGCAT
MIM397	TCATTGAGTGTAAAATCGTTGATG
MIM823	TGGGTGGTGAGGGTCATATAAGAT
MIM393	TCCAAAGGGAGGGCCGCATTGATCC

Forward (35S):

35S_F	ACTATCCTTCGCAAGAC

A.1.2. PRIMERS FOR PRODUCING NOVEL MICRORNA OVEREXPRESSION CONSTRUCTS

New19_L	GACTGAATTGAGCGACGGAC
New19_R	ACTCAAAGGTAGGCTCGTCA
New117_L	TCGTTGTTCCTCTGTTTCGT
New117_R	GAGTCCCCATCTCTCACGAT
New29_L	CCCGTGAACAGCCAAGTAAG
New29_R	AACAGAGCATAACCAAGAGAACT
New120_L	TATGGGGTCTCAGTTGGACT
New120_R	TCAGTTCCCGTAGCAAGAGT
New11_L	AGGGTGTGAAGCCTTTACCA
New11_R	ACATAACAAGCTTCGCACGT
New78_L	GACTTCTTAGACCGTGCCATG
New78_R	TGGCCGTAGATCTTCATGGA
New39_L	GGTTGTTTTGGTGTCACGTTG
New39_R	CGAATGAGTACTCGTGGCAA
New141_L	TGCTTCCTCTGGTCGCAA
New141_R	AAGCAATCTCTCTGCTACGC
New36_L	GTGATCCGAGAGCTTTGTGG
New36_R	GAACGGTGCCATAACAATCCT
New35_L	CGCGGTTTATCCTTCTTCGA
New35_R	AAGGACATCATGAAAAGGGCA

Genotyping was undertaken using 35S_F (see A.1.1.) and the specific new miRNA_R primer

A.1.3. PRIMERS FOR RT QPCR

ACT_F	CCGCTCTTTCTTTCCAAG		LAC13_F	AATGGTCAGCCTGGTGATCT
ACT_R	CCGGTACCATTGTCACACAC		LAC13_R	TCTCCTGACCCCACTAGAAATC
ALDH3_F	TCTGTCTAGCTGGAATTTTCCTATATC		LAC17_F	CAACGAGGCACTCTGTGGTA
ALDH3_R	AGATGACTTGAGAAGCACGGTA		LAC17_R	CAACGAGGCACTCTGTGGTA
ARF16_F	ATTCCGTGGTGCTCAGGTAT		LAC3_F	GACCAGGAAGCCCGTTTATC
ARF16_R	TGAGGTAGCCGCATCTTTTT		LAC3_R	CGTGACATTCCAATTCTTCTTG
ARF19_F	CACCGATCACGAAAACGATA		LBD16_F	CACTAGTTTGTTGCCACGTCATAAAC
ARF19_R	TGTTCTGCACGCAGTTCAC		LBD16_R	AGCCGCCGGAGATCTAAGACA
ARF7_F	GCTCATATGCATGCTCCACA		LBD29_F	GGTTTCGTTGTGGTAACTACTCTCTG
ARF7_R	GCAATGCATCTCTGTCATATTTGT		LBD29_R	GAAGTTCTGGGACGGTTCAACAC
AT1G53290_F	GACGTCGCCGGAAATAGTAA		LP2_F	CCAGCACCAACTGCAACA
AT1G53290_R	CGAAGTCATCATACTCTGCGA		LP2_R	TCGAAACATCCGCTAGTAGCTT
CMT3_F	GAAAATACTAAAGAGACTATCCCTG		MYB65_F	GGCTAATCATCTGAGGCCAA
CMT3_R	TGTCGACATAGCTCCACAGC		MYB65_R	TCTTCGCTCCAATGAAGGTC
LAC1_F	CGAAATTCCCATTATACTTGGTG		NAC1_F	TTTCTATCAAGGTCGAGCTCCT
LAC1_R	GGGTGTAAGCGTCGGAAAC		NAC1_R	GATCCATCAGTGGAGGAAGC
LAC10_F	CGATCAGATTCCGAGCAGAT		PATL2_F	CCAGAGTTTGTCGCTAAAGAGC
LAC10_R	GAATTGATTGGTTAGGGCCTTT		PATL2_R	TGTCCTTGGCGATGTAATGA
LAC11_F	CCCACAGCCATACCAAGAAT		SMZ_F	TAAGGTGGAGTTCGTGCAAA
LAC11_R	AACCGCAGTCTCAACGTCTT		SMZ_R	ATATGGGCACCAAACTTCATG

A.1.4. PRIMERS FOR MIRNA (STEM-LOOP) RT AND QPCR (FOLLOWING THE PRIMER DESIGN PRODOCOL DETAILED IN VARKONYI-GASIC ET AL., 2011)

Novel miRNAs:

11_F_QPCR	GCGGCGAACCGATGCGCTGCGT
11_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACAGCAGT
29_F_QPCR	GCGGCGTAGTGGAAGCAGCAA
29_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTTCTCG
19_F_QPCR	GCGGCGCCCGAGGTGTGTGGA
19_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTTCTGA
35_F_QPCR	GCGGCGGGGGATTGACATGAAT
35_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC AGTGTC
36_F_QPCR	GCGGCGCGATCCCCGGCAACG
36_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTGGCGC
39_F_QPCR	GCGGCGATATAATGGACTGAGA
39_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACAGGCGG
78_F_QPCR	GCGGCGTTGTGTTGCGTTTCTG
78_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACAATCAA
117_F_QPCR	GCGGCGTCGGACGCTGACTGGC
117_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTGAAAA
120_F_QPCR	GCGGCGAATCGCACATCTGTT
120_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACCCTCCT
141_F_QPCR	GCGGCGGAGTGGAGCGCGAAT
141_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACAGTGCT

Conserved miRNAs:

MIR156_F_Q	GCGGCGTGACAGAAGAGAGT
MIR156_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACGTGCTC
MIR157_F_Q	GCGGCGTTGACAGAAGATAGA
MIR157_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACGTGCTC
MIR159_F_Q	GCGGCGTTTGGATTGAAGGGA
MIR159_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTAGAGC
MIR160_F_Q	GCGGCGTGCCTGGCTCCCTGT
MIR160_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTGGCAT
MIR163_F_Q	GCGGCGTTGAAGAGGACTTGGAAC
MIR163_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACATCGAA
MIR164_F_Q	GCGGCGTGGAGAAGCAGGGCA
MIR164_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTGCACG
MIR165_F_Q	GCGGCGTCGGACCAGGCTTCA
MIR165_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACGGGGGA
MIR168_F_Q	GCGGCGTCGCTTGGTGCAGGT
MIR168_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTTCCCG
MIR172_F_Q	GCGGCGAGAATCTTGATGATG
MIR172_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACATGCAG
MIR390_F_Q	GCGGCGAAGCTCAGGAGGGAT
MIR390_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACGGCGCT
MIR393_F_Q	GCGGCGTCCAAAGGGATCGCAT
MIR393_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACGGATCA
MIR396_F_Q	GCGGCGTTCCACAGCTTTCTT
MIR396_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACCAGTTC
MIR397_F_Q	GCGGCGTCATTGAGTGCATCG
MIR397_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACCATCAA
MIR408_F_Q	GCGGCGATGCACTGCCTCTTC
MIR408_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACGCCAGG
MIR775_F_Q	GCGGCGTTCGATGTCTAGCA
MIR775_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTGGCAC
MIR823_F_Q	GCGGCGTGGGTGGTGATCATA
MIR823_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACATCTTA
MIR857_F_Q	GCGGCGTTTTGTATGTTGAAG
MIR857_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACATACAC

Reverse qPCR primer for all microRNAs:

MIRNA_R_Q	GTGCAGGGTCCGAGGT

U6 housekeeping primers:

U6_F_Q	CGCACAAATCGAGAAATGGTCC
U6_R_Q	CGATTTGTGCGTGTCATCCTTG

A.1.5. TDNA INSERTION MUTANT GENOTYPING PRIMERS:

SALK_041500_L	TGAATCTGCAGGGAAAATGAG
SALK_041500_R	TTCCCACAATGTGATTTCTCC
SALK_074816_L	CATGCGAATGAATAATAAAATCATC
SALK_074816_R	CATCAGGCAATTCCAAAGAAG
SALK_089888_L	CTTTTTCTGGGGATCAGGATC
SALK_089888_R	AAGGCCTAAGCTCAAGCAATC
SALK_048394_L	CAACAAAATGGAATAACAGTTAAAGC
SALK_048394_R	CAGGTCTTCTTCTTGCGTCTG
SALK_057686_L	TGGTCCTCCAGTGATCTCAAC
SALK_057686_R	CACGTGGTGTGTATCTGTTGG
SALK_046612_L	TGTTTGTCACGCAAAAGACAG
SALK_046612_R	AGAGTCTTTCGCCCTCTTCAG
SALK_087191_L	GCGTACATGGGAACATTAAGC
SALK_087191_R	CAACATAACATGCAACCGTTG
SALK_129128_L	CATCTAGCCAGCTAACGAACG
SALK_129128_R	CTTAAGAAATGGGGCAAAAGC
SALK_010447_L	GTTCCACGACCTCGTAGACAG
SALK_010447_R	AACACACCCAATCAGAGATCG
SALK_061031_L	CAAGCTCAAGCATTGGAAGAC
SALK_061031_R	CTACGGCGACAATCAAGCTAG
SAIL_124_B10_L	TGATAAATTCCACGTGGCTTC
SAIL_124_B10 _R	ACCTGTCATCACCCATGATTC
SAIL_384_G07_L	CTTCACTTAAGCATTTCCCCC
SAIL_384_G07_R	TGGTTTGACGAAGCTGAACTC
SALK_053004_L	AATTGGGAAACAACCTGATCC
SALK_053004_R	TGGGTTGATTCTGTTTTCTGC
SALK_068624_L	TGATGGCATGAGAGAGTGTTG
SALK_068624_R	CATGTCCAAACTTCACAATCG
SALK_029936_L	GTTCTCTTGCAGCCTCAATTG
SALK_029936_R	TTCCCACAATGTGATTTCTCC
SALK_014478_L	TCCACAAGCCCAAACTACTTG
SALK_014478_R	TAAACTACGAATCCCTGAGCG
SALK_012938_L	TAAACTACGAATCCCTGAGCG
SALK_012938_R	TTGCTTGTGGCTATGACAGTG

SALK_056910_L	TTTAAAAGGATCAATGAGATGAGC
SALK_056910_R	GCGTACGCTTATTGGAGAGG

Insert genotyping primers:

SALK_LBb1.3	ATTTTGCCGATTTCGGAAC
SAIL_LB1	GCCTTTTCAGAAATGGATAAATAGCCTTGCTTCC

A2. RNASEQ RAW READ COUNTS

Time Point	Total Reads	Read alignment rate	Aligned pairs	Multiple Alignments
6h	46809921	62.1%	27314503	1262055
9h	50790836	54.0%	25608066	1090788
12h	49018471	59.9%	27597100	1168475
15h	52210894	57.8%	28051433	1029934
18h	48433125	55.9%	25033979	998524
21h	47277079	52.0%	22743415	886793
24h	42349463	51.8%	20236625	797782
27h	32046203	54.3%	15922255	601137
30h	40842937	43.7%	16394546	646733
33h	34776136	42.7%	13648660	506813
36h	36472843	58.5%	19886867	782424
39h	42948470	55.9%	22419529	851693
42h	34669938	82.4%	27353404	902091
45h	32005163	78.3%	23888266	797328
48h	27061159	79.8%	20565125	769474

Table A1: *Read count data for 15 time-point RNASeq of developing lateral root primordia. 'Read alignment rate' is a percentage of reads which mapped from both left and right-directional reads. 'Aligned pairs' refers to the numerical value of this percentage.*

A.3. FASTQC

A.3.1. 15 TIME POINT RNASEQ FASTQC

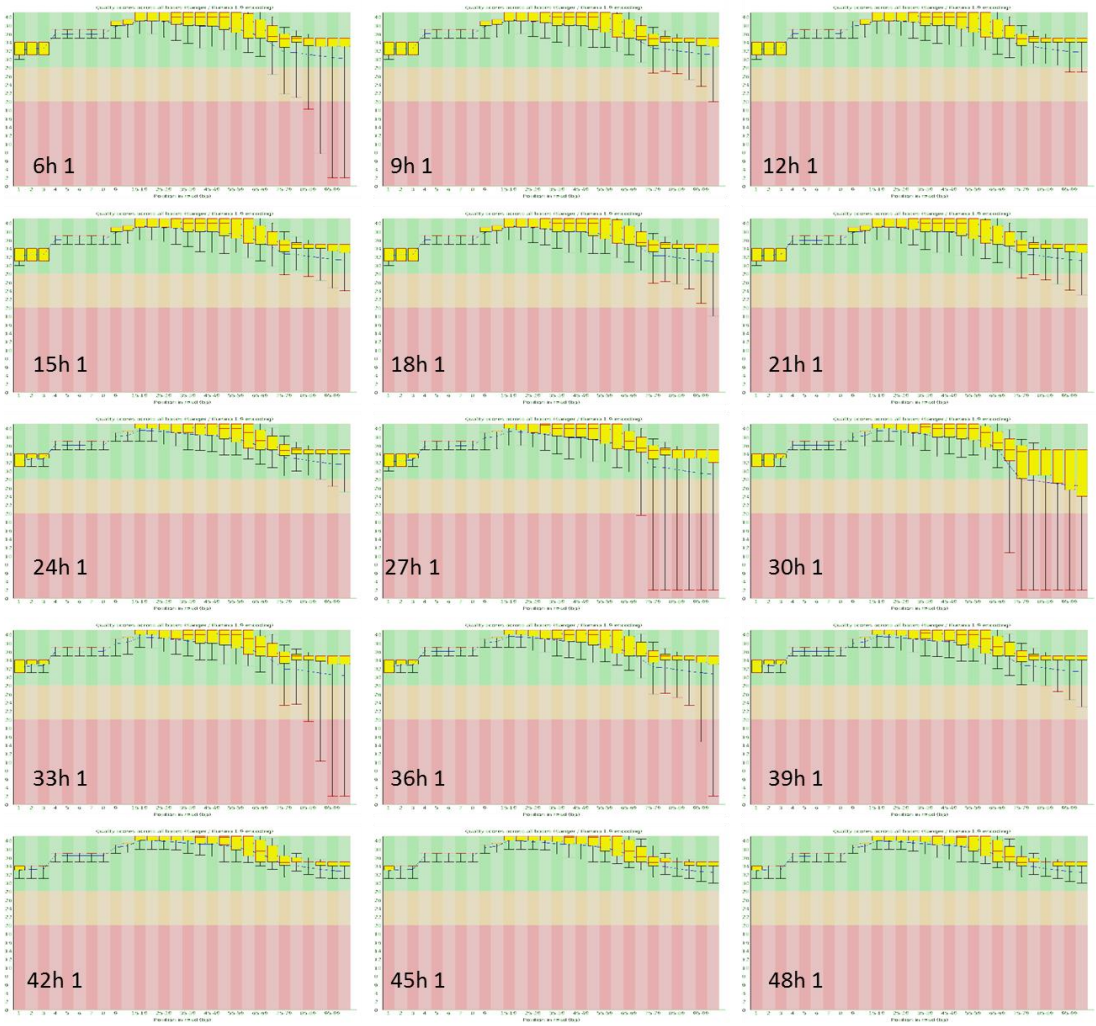


Figure A1: *FASTQC analysis of individual nucleotide bases across all reads of all RNASeq samples across four flow-cell lanes. Y axis represents Phred quality score. Yellow boxes represent 95% of reads with black lines indicating the range. Red lines represent the read mean.*

A.3.2. SMALL-RNASEQ FASTQC

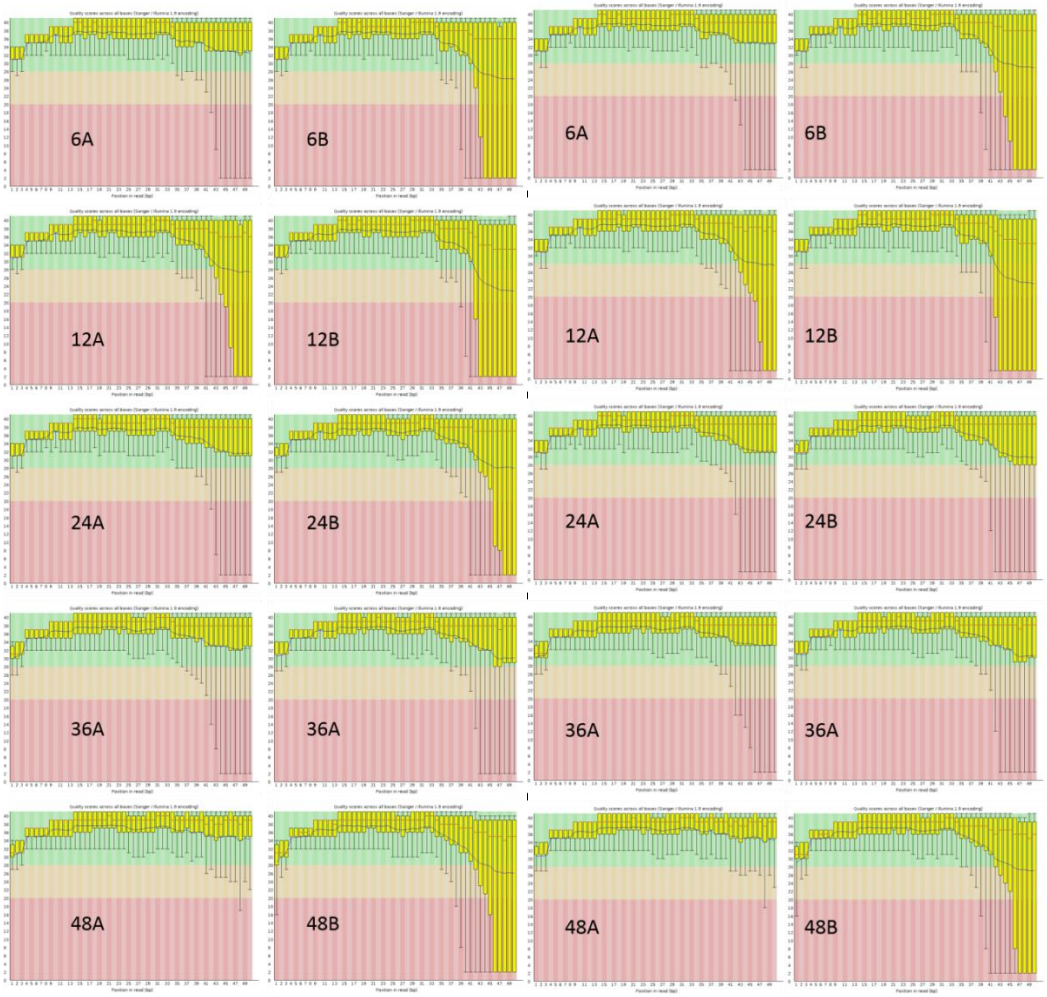


Figure A2: Example of *FASTQC analysis of individual nucleotide bases across all reads of small RNASeq samples across two flow-cell lanes. Y axis represents Phred quality score. Yellow boxes represent 95% of reads with black lines indicating the range. Red lines represent the read mean.*

A.3.2.1. SMALL-RNASEQ FASTQC REPEATS

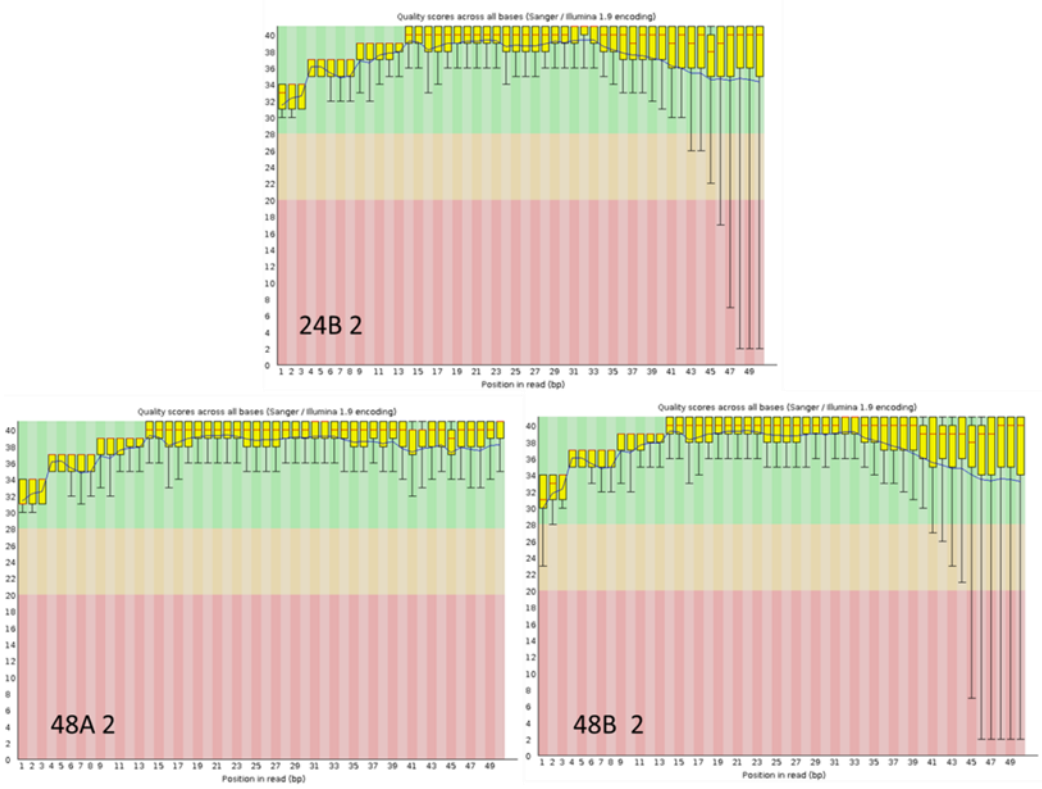


Figure A.3: *FASTQC analysis of individual nucleotide bases across all reads of three repeated small-RNASeq samples (24B, 48A, and 48B). Y axis represents Phred quality score. Yellow boxes represent 95% of reads with black lines indicating the range. Red lines represent the read mean.*

A.4. MIR159 TARGET QPCR IN MIM159

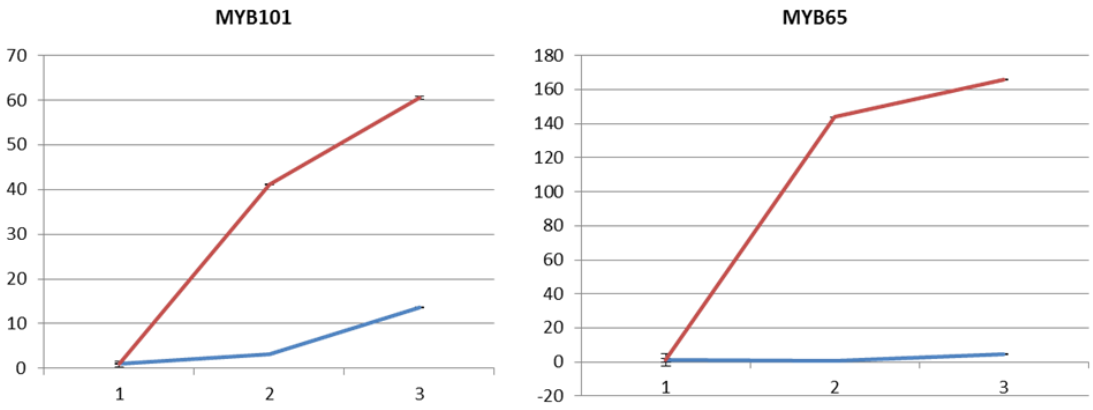


Figure A4: *Relative RT qPCR expression of MYB101 and MYB65 in 10-day-old MIM172 (blue) and Col-0 (red) root systems following NAA treatment (1 – 0h, 2 – 12h, 3 – 24h) Error bars represent standard error of technical repeat*

A.5. NOVEL MICRORNA PREDICTED HAIRPINS (UEA SMALL RNA WORKBENCH)

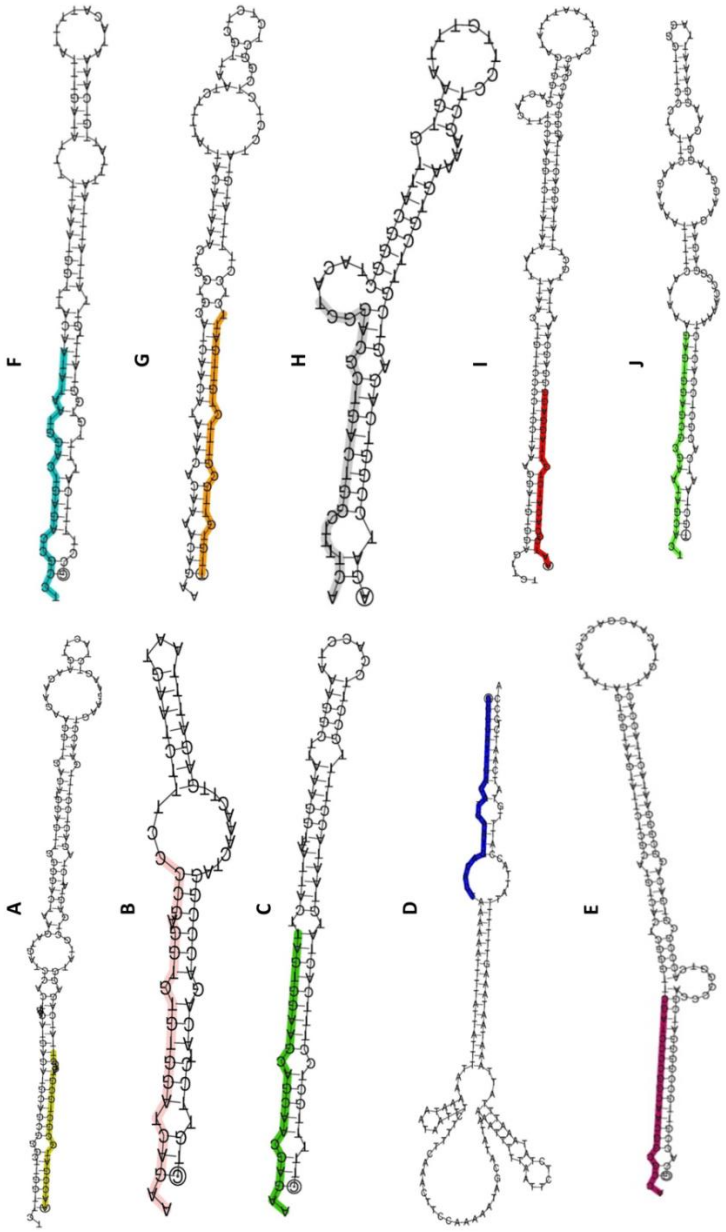


Figure A5: *Predicted, novel miRNA precursor hairpins, developed using The UEA small RNA workbench (Stocks et al., 2012). Coloured regions indicate the mature microRNA location. The letters A to J refer to new microRNA11, 19, 29, 35, 36, 39, 78, 117, 120, and 141, respectfully.*

Novel miRNA	Precursor Sequence	Precursor co-ordinates
miR11	AACCGATGCGCTGCGTACTGCTTATCAGAGGTATGCTGAGTATCTAGACTCCTTTG AACCTGAAGAAGTCTACTTGAAGAAGAAGGTTGAGATGGAGTTGGGGACTAAGA TGATCCACTTGAAGATGAGATGCAGCGGGCTTGGTTCT	17396144 - 17396291
miR19	CTGTTCCTACAGACCCGGATCAAACTTGAGATTTAATGAATCTTTCCCCGAGGTGT GTGGATCAGAA	30227829 - 30227895
miR29	CTTTTTGCTCCTTTCACTATGTAATACCTTTTGCCTTCCACCTAAAGGCTTAAAGGTT AATTACTTAGTGGAAGCAGCAACGAGAA	23355103 - 23355188
miR35	GGGGATTGACATGAATGACACTAAAAAATTTTATTATTTTAAAAATAATAATTTCT TTTCAAACTTCCAAAAATAGCATATAAAATTTTTAATTCTCTATAAATTTATTAAAT AATAAAGTTTTTTTATTACCATTTGTATCAATCTCCCA	3013112 - 3013262
miR36	GCACCGTTGCCGGGGATCGACCCGGGTCACCCGCGTGACAGGCGGGAATACTTA CCACTTAGTACAACGACCCAATATAGTGGTAAGTATTCTCGCATGTCACTCGGGTT CGATCCCCGGCAACGGCGCCA	3740828 - 3740958
miR39	GCCTTTTCATTTTGTGGTATTTGTTATTATTTAATTTATTGTCAAAATACATTTTATT TGATATTTTTAAATGGTTTACAATATAATGGACTGAGACCGCCT	1111029 - 1111130
miR78	TTGTGTTGCGTTTCTGTTGATTCTCCTTTTATGTATCCTCTCGGCTCTCTCGTTAATC TTTTATTACATAAACTCGTGCATCAACATAAACACAAAACAGAA	22678167 - 22678268
miR117	AGATCCGGTCAGAGTCGTTCGTGAAAAGCTCCTTGTTTAAGTGTTACGGGCTACA TCGGACGCTGACTGGCTTTTCA	7146675 - 7146751
miR120	AATCGCACATCTGTTAGGAGGGGAGCAAATTAATGCTTTATAGGACTTTAGGGCA CCCACACTCTTAATTTATAGTGGTGACTACTTTCCAAGGTCTTATAATATTTTAACT TGTTCCCCTCCTAAAGGATGTGGAGTTCT	5590304 - 5590444
miR141	TGCTAATCACGCTCCACTCTAAAACCCGGAGAAGAAGGTAGGAGAAGGAAATTA GGGTTTCCCTATTCAAGAAAATTTTCAAAAAGAGTGGAGCGCGAATAGCACT	1913361 - 1913466
miR97	GCTAGAGGACCAGAGCCAGCTCGAGTCTGGCTCTCTTAGCTCCCACGACTCTGGT CCCTTTGGCAC	18017274 - 18017339
miR 110	CGCTACAAGAGTTTTAGTCGAGTGGAATGAGGGTGAGCATGAAAGATATGATTC CAGCTACTTGAGTGAT	6539722 - 6539791
miR144	CTAAGACTAGCAGTTTCGAGGAAAAGCTCAATTTCTTCTGCGGTTACTGGTATTGG CTTTGCGAAACCGAAGCTGGTAGTACTTGGCT	18397816 - 18397903
miR61	ATGTGAGGAAGACTGTTGAAGTTGTTTTGACTGTAAGTTACTACTTCAATATTCTC TAATCACAATA	10960249 - 10960315
miR62	ATAAGTCAAGTAGTAGTCCTTGTTTGTCATCTAAGTTAAGAAGTCCTATATGATGA CAAACATGGCTTCTACTTGGCTTATGC	10445474 - 10445556
miR88	TTATAACTCAGCCATAACATGTAAGGTTCCGTAAATTTACATGTCACGGCTGAGTT ATATGC	3817468 - 3817529

Table A2: *Predicted precursor sequences (and locations) of predicted novel microRNAs*

A.6. NOVEL MICRORNA IDENTIFICATION & ANALYSIS

Novel miRNA	PCC between replicates
New miR11	-0.191
New miR19	0.588
New miR29	0.447
New miR35	0.558
New miR36	0.866
New miR39	-0.236
New miR78	0.833
New miR117	0.101
New miR120	0.234
New miR141	0.495

Table A3: *Pierson correlation coefficient (PCC) values between two replicates for the ten microRNAs identified in both*

Novel miRNA name	Chromosome	Co-ordinates	Strand	Found in Replicate
New miRNA33	Chr2	3565396-3565443	-	A & B
New miRNA13	Chr1	2460559-2460605	+	A & B
New miRNA60	Chr3	11174293-11174338	-	A & B
New miRNA34	Chr2	3416012-3416056	-	A & B
New miRNA5	Chr1	17717725-17717767	-	A & B
New miRNA56	Chr2	7757392-7757448	-	B
New miRNA72	Chr3	17919381-17919429	+	B
New miRNA44	Chr2	16217760-16217806	+	B
New miRNA7	Chr1	1203315-1203359	+	B
New miRNA100	Chr4	16429491-16429534	-	B
New miRNA47	Chr2	17802968-17803010	-	B

Table A4: *Potential novel miRNAs identified by small RNASeq deemed too-short for initial consideration or further investigation though still filling many if the criteria for being novel microRNAs*

A.7. KNOWN MICRORNA CONSERVATION IN ARABIDOPSIS

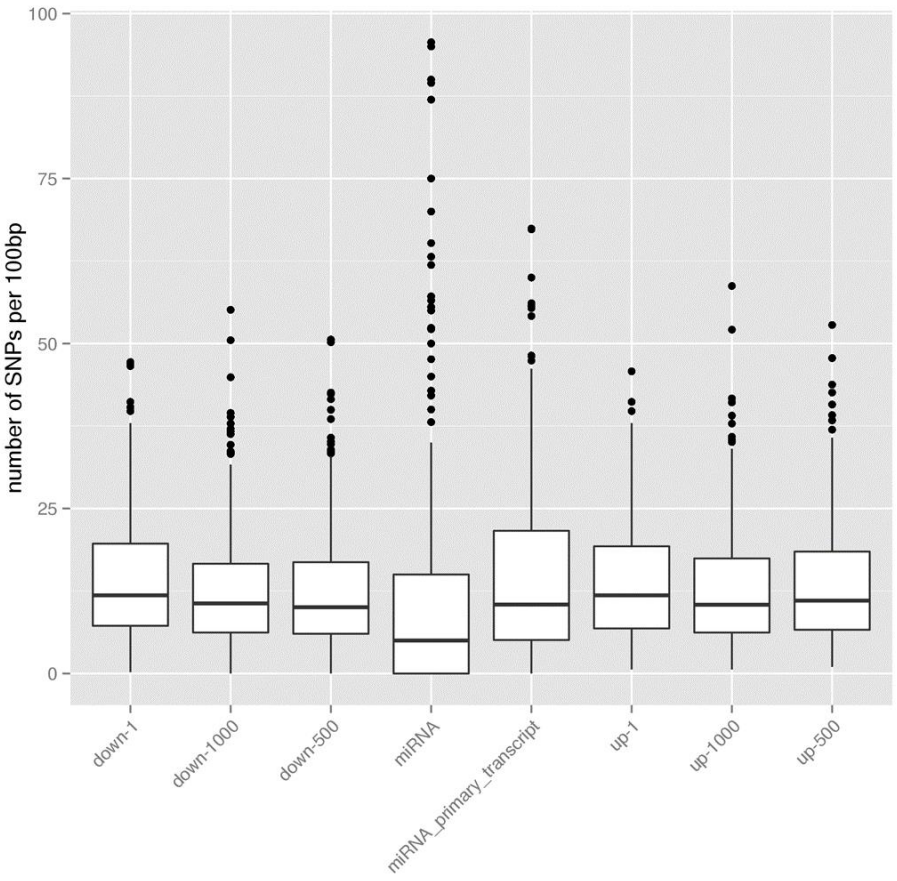


Figure A6: *Conservation of the all Arabidopsis microRNAs identified in MIRBASE (as of April 2015) based on various locations across sequenced genomes in the SALK 1,001 Genomes database (Gan et al., 2011- Accessed 2nd April 2015). Box plots for the number of SNPs (y-axis: SNPs/100nt) at predicted miRNA precursor (primary transcript) and mature miRNA loci. Up-1000 refers to the upstream region from 1,500 to 1,000, up-500 refers to the upstream region from -999 to 500, and up-1, refers to the upstream region from -499 to -1. The same nomenclature applies to the downstream (down) regions. Grey boxes represent 50% of the data and the strike through them represents the median confidence interval. Whiskers indicate the upper and lower quartiles.*

A.8. MICRORNA OVEREXPRESSION LINES

miRNA	Sequence taken for construction of novel miRNA overexpression line.
New miRNA 11	AGGGTGTGAAGCCTTTACCAAATGATATAACCGATGCGCTGCGTACTGCTTATCAGAGGTATGCT GAGTATCTAGACTCCTTTGAACCTGAAGAAGTCTACTTGAAGAAGAAGGTTGAGATGGAGTTGG GGACTAAGATGATCCACTTGAAGATGAGATGCAGCGGGCTTGGTTCTGAATGGGGAAAGGTAAT CAAATCTGATCTATACTCTCAAAACACGTACTCTGTTGATTTGCATCTTTGACAGCATCATATAAGC TAGAATAAGCCTGATTAAAGTATGAGTTTGCATAATTATTTACGATTTCAACTTGTCCTCCAAAGT CTCACTCGTACTATATGTGATCTTTTTTTCTCAACCTTATGCATCTTTTAGGCTAGCTATCCGAATGA AATCATATAGAATTTGGCAGAAACTAGTGTTTTCATTTAGACGAAACCACTGAACCAGTAGTTGT CTTGCTAACGCTTTGTTTTAATATTTTTGTAGGTTACTGTTCTTGGAACCTCAGGACTCTCGGGATC TTACGTGGAGCAAAGGGCCTAAAACAAAAAGAGAAGATCTAGAATTACGTGCGAAGCTTGTTAT GT
New miRNA19	GACTGAATTGAGCGACGGACAACTCGAAGCATTCAAGAGCCTCTCAGTTGCAACCTGAAAACCCT CTTAGCTATCAATAAGAAGCAAATCAACCCAGGAGATAATTTACAGAAGATGGAAGATCTGTTCC TACAGACCCGGATCAAACTTGAGATTTAATGAATCTTTCCCCGAGGTGTGTGGATCAGAAAATGT TTTGGAGTTGTAACATGCCAAATGTGTATAATTGTTTTTCTATATAAGGCTGGGTTTCAAGTTTGA AGAACAATGCCACAAAAAAAGAGAGGGGTCTTGTCTTTTTTTTTTTCTTGTTTTTTTTCGTTTTGTT CTCTCCACACTTTAAAGAGGATTTCTCTGATGTTGTACCCTTTTATCTCCTTCTTTTTTCGGTTTCTCC ACATTATTTATTTTCTCCATTACAATGTTTAAGCTTCTTGACATCTCTGGAAAATTTTGATTTGACG AGCCTACCTTTGAGT
New miRNA29	CCCGTGAACAGCCAAGTAAGAGTAAAAAAAAAAGACTGGCTCTTCTGCGGTATAATATCGAACAA TGTATGAAGTTGTGTGTAATTTTGTTTAAAAAGTGATTAGATAGATATAGGGGAAAAGCCTTTTG GCTTTGCTCTAAGCTTCTTTTTGCTCCTTTCACTATGTAATACCTTTTGCCTTCCACCTAAAGGCTTA AAGGTTAATTACTTAGTGGAAGCAGCAACGAGAAGCTGGCAAAGGAGGCTTCTGTTCTGTTTCT GTTTCTGTATGTAAACCATGCAAATTCCGGTTTGACTATGAAATATCCGGTTATGTTTCTGTTCTTT ACCTTGTTTATATCCGTCTTCATGCAATTGCAAATGATCCTTATTACATCTTCGGTTTTCCTTATGAG AAGTGAAACTTAGTTCTCTTGGTTATGCTCTGTT
New miRNA35	CGCGGTTTATCCTTCTTCGAGGATCGTTTTATTGGTCCCTCGAACAAAATAAATAATAGAGACGAT TTTAAGTGGATAACACTCTAAATTTTTGGGGGATTGACATGAATGACACTAAAAAATTTTATTATT TTAAAAATAATAATTTCTTTTCAAACTTCCAAAAATAGCATATAAAATTTTTAATTCTCTATAAATT TATTAAATAATAAAGTTTTTTTATTACCATTTGTATCAATCTCCCAAATAAAAATGCGAGCAAACA CTAATCTAATATTAAATTGAATCAAAAACTTTTAAATAGTAAAAATATATACCCTAAATTAGAAAT AAAAAATGTAATATTACAAACAAATTTTAAAATAAAAAATCTCTTTTAAATGGTGAAAATATATA CCTTAAATTTTTTTTGGTAGAAAAATATATACCTTAAATTGGAAATAGGAAACTCAAATATAATAC CATAAACTAATATTAAAAGAAAAATTCTTTCTAATGCCCTTTTCATGATGTCCTT
New miRNA36	GTGATCCGAGAGCTTTGTGGAATTCGACCCTTATGTGCTACAATCAATCTCTTATTTGAGACAGTG GTCTTAAGGCAATTTGACCTATATCAAATTTGGCACCGTTGCCGGGGATCGACCCGGGTCACCCG CGTGACAGGCGGGAATACTTACCACTTAGTACAACGACCCAATATAGTGGTAAGTATTCTCGCAT GTCACTCGGGTTCGATCCCCGGCAACGGCGCCAAATTTGATATAGCTCAAATCGCCTTAAGCTTAC TCCCTCAAATAAGAGTTGTCGTTAGCACTTAAGTGTCGAATTCCACTGAGCTCTCGATGTTCACAC AATAGAATTATGATGTTATTAAATAATCTAGACAAAGTAATTGAATGTAAAAGACAAATAACCA AGTAAACGAAGTGTAGATTCAAGTGATTAAAGCGTCGGGTCTAAGGAATTGTCTCGGGAGATAG ATAAATTGTAGATCTAGCTAATATAAGGATTGTTATGGCACCGTTC
New miRNA39	GGTTGTTTTGGTGTCACGTTGAGATGACCTTAAAATTGTTTATATACCATTTTTGTTTGATTTTTAT GCCTTTTCATTTTGTGGTATTTGTTATTATTTAATTTATTGTCAAAATACATTTTATTTGATATTTTT AAATGGTTTACAATATAATGGACTGAGACCGCCTAACAAGATGGTTGTAGATATTTAAAACAAAT AAAACAAAAAGGCCGCAAGGCCCACAACCCAACAAAGAAATCGAGGCCCAAAATGGTTGTAGA AATTTGGTTTAGCTATCAAAAGGTTCAAGCGGCTGACAATAGAATTTGTTGTCTGTCCAAAACAA ACTACCATAAGCTTATATTGCCACGAGTACTCATTCG
New miRNA78	GACTTCTTAGACCGTGCCATGTTAATACCATTTAATGTTACTGGAAACAAATACTAATAAATAATA CGTCACTTAAAACAAAATGTCTATACAAATACTATTAACATAAAACAAAATCTCTAGACCTGTATT ACATCACATGCATCTCTTACTTATAAACAATAAACTTCTTTATTGTATGTTTTAGGGAATAAATATA TGAGTTTGATGTCATATCATGTTTTTTTTTTTGTTATCTTCATACTTTTGTGTTGCGTTTCTGTTGATT CTCCTTTTATGTATCCTCTCGGCTCTCTCGTTAATCTTTTATTACATAAACTCGTGCATCAACATAAA CACAAAACAGAAGCATGAAGATGAATAACAACGACGGCCAACATAATAAAACGGTGACATTTCC TCGAAAAAGCATCGTGAAGGCTACTGTTTTCATCGTCGTTCTATTCATTTCAGCTGCTATCTTAGAC TTACTGGGCTATCTCGATTTCGTGGGTTTTAATATCACCTTACATTCTTGGTTTATCTTTAATATAA ATCGCAGTAGTTAACGTTAATCATAATCATAATTTGTCTTATTGTTTTGTTTTATAGAATGCTTTTG CAGGTTTAAAGCTAACAACAAAGACAAAAGAACCAAATCCTTACGGATGCGACTTCGTACAGAA

	CCAATCAAGTCAAACTCCTATCTCACAAAACCGCAAATCGAGATTAAACCCCAACAACAGTTCAA AATCATCCACGTGTCCTTCTTATTTCCGTTGGATCCATGAAGATCTACGGCCA
New miRNA117	TCGTTGTTCCTCTGTTTCGTTTAGCTAACATTTACGAATCTCTTCGTTTTTCTTTGTTTATGATGTGT TGGTCTGATTGTTATCGATTTTTGGGGGTAAAAATTTGTATTTTGTTTAATCTGGTTTGGTTCTATT GATGAATCGTTTTTCTCGTTTATCTTCAAGCTTAGCTGGATTCGTCTCTGGTGCTGTAGATTTTCGA GATCCGGTCAGAGTCGTTCGTGAAAAGCTCCTTGTTTAAGTGTTACGGGCTACATCGGACGCTGA CTGGCTTTTCATGGAATTTGAGTGCCAAGAGAGACTTGAAGCTGCGAACGATCAAAGAAGTGAG TCTGGTCTTTTGAATACCGATTTGAGAATAGGTAATGATGGGTCTGTTGAGATTCCAAATGTGAA GTTGTGTTGTCAGAAAAAGAAGGGAACTTTAGTGCCCAGTGGCTCGAAACAGTTACTCTCTGACA AGGATTTGTCTACTACCATTATTGATTTACCTCAGGCATTGATATCTGAGATTCTTAATTGCCTTGA CCCAAAAGAGTTGGGTCTAGTGTCTTGTGTTTCAACTTATCTGCATAGGTTAGCATCTGAGCATCA CGCTTGGAAGGAGTTCTATCGTGAGAGATGGGGACTC
New miRNA120	TATGGGGTCTCAGTTGGACTGGTAAGACTGATCATGCTTCTTTATTGTTTGACATCTAAAGACTTT CTTATGCATGCCAATGGCTCAAACGTTATATATTATGCATTTGGCGGTGTACTAAGAATCGCACAT CTGTTAGGAGGGGAGCAAATTAATGCTTTATAGGACTTTAGGGCACCCACACTCTTAATTTATAG TGGTGACTACTTTCCAAGGTCTTATAATATTTTAACTTGTTCCCCTCCTAAAGGATGTGGAGTTCTT ATCATGTATATAATCAACCTAACTTAACCTTTCCAATGAATCTAACAGTTGATGCAAATATATTTCC ATCTATTGCTGAGCATGGGTTGTTATCCAGTGCATCATCATAGAAACATAATTGATGTCTATTCTA TGTACATGTATACAGAGTACACATATATTGTGACTGTATTCTTCGTCCTTCGTATTTTTTATGTTTTT CCAGCTTAAAAGCCAGTGACTAGAAATCTTGTTTACCTTTCGAAAATATTGTTTATCCATGTGCTTT TGCTGTGGATTACATATAGCAGTACTCTTGCTACGGGAACTGA
New miRNA141	TGCTTCCTCTGGTCGCAAGACTCGAAGGAACTGATGAAGGACCAGTAAAATCAAAATCCTGTTCC TTCTTTATGTTATTTATTTTTCTTTCGTTTTTTTGCGAATGTAGCTGTTGTTTTTATGAGAGTAGTGC TAATCACGCTCCACTCTAAAACCCGGAGAAGAAGGTAGGAGAAGGAAATTAGGGTTTCCCTATT CAAGAAAATTTTCAAAAAGAGTGGAGCGCGAATAGCACTAACTTTTCTTTTTCTTTTCTCTCAACA TTTTTCTACCATTTGAGTTCTCTAATGGAAGAAGAATTTTTAAATATATGTGAAAACATGTGTCAT TGTCTTCGTATTCCTCTACTCTATCTAATTTCTCATACTCTGTTTCGGATGCTCTGTTTCTAGGGCAT CTCTGTTTTTTGGATGTGTTTGTAAAAAGAAAATAGAGGTTTTGTGTGTACGACTATAACAAAAG TAGACATTCTTAGCGTAGCAGAGAGATTGCTT

Table A5: *Sequences taken for novel microRNA overexpression lines including sequence either-side of predicted precursor sequence. Green sequences represent primers used for sequence isolation. Red sequence represents the predicted precursor. Purple sequence represents the predicted mature microRNA sequence*

Novel miRNA	Precursor length	Upstream	Downstream	Total insert length
11	148	28	416	592
19	67	123	289	478
29	86	145	199	430
35	151	94	272	517
36	131	98	273	502
39	101	67	196	365
78	102	246	435	783
117	77	202	353	632
120	141	122	312	575
141	106	130	257	494

Table A6: *Lengths of novel microRNA precursor sequences with upstream and downstream lengths taken for attempted construction of overexpression lines.*

A.9. NOVEL MICRORNA SEQUENCING EXPRESSION

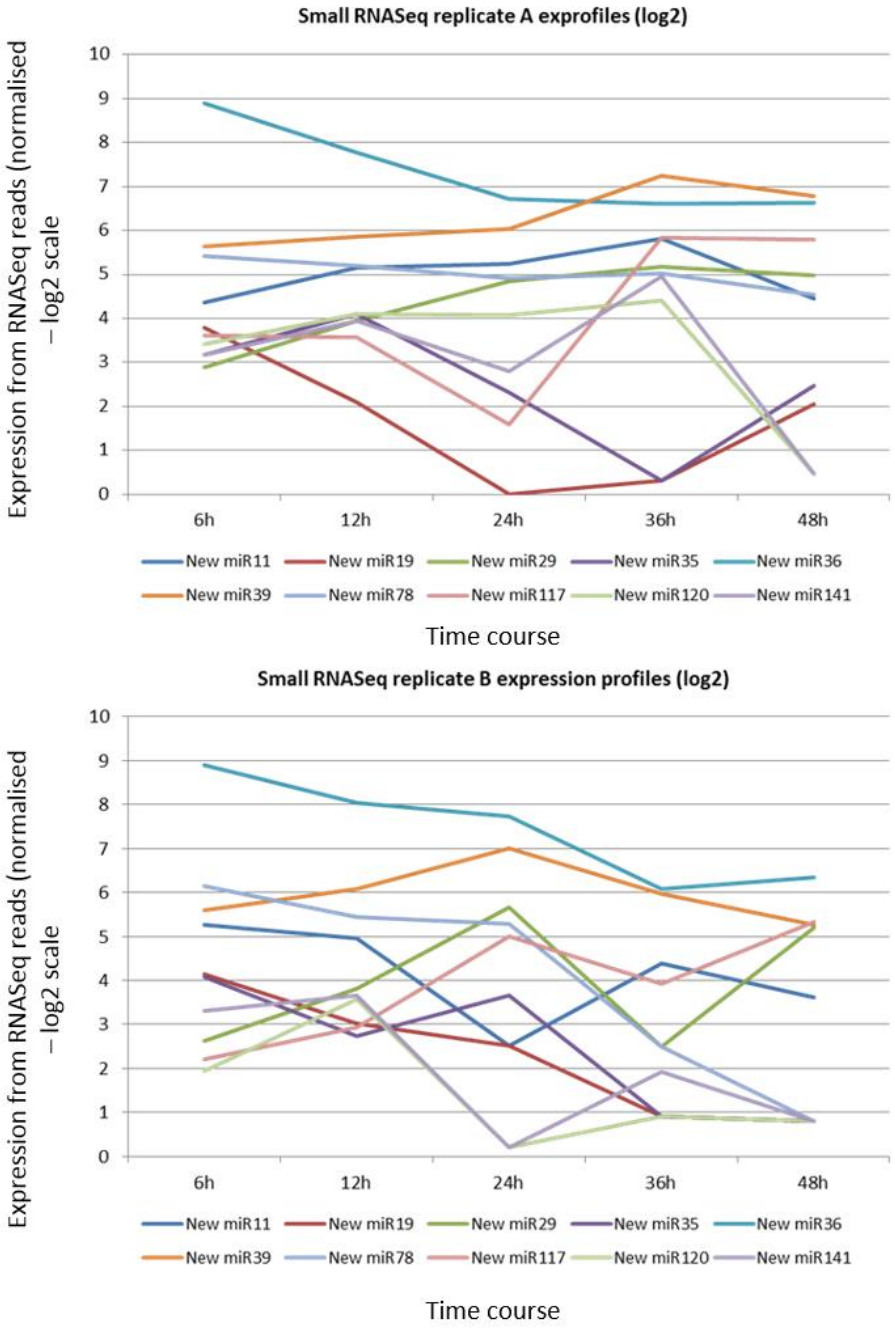


Figure A7: *Expression profiles of novel microRNAs identified in both smallRNA replicates (log2 scale). Normalised read counts – y-axis, time course – x-axis*

A.10. GENOTYPING

Prior to any studies, all lines were genotyped as necessary. For tDNA-insert mutant studies, homozygous lines were obtained and genotyped using the generic 'SALK LB1.3' primer and a gene-specific reverse primer alongside the gene-specific primer set on both the mutant and wild-type (Col-0). Novel microRNA overexpression lines and miRNA Target mimic lines were genotyped using a forward primer from the 35S promoter region and a gene-specific reverse primer. In the case of the novel microRNA overexpression lines, this was the same gene-specific primer used for initial isolation of the predicted miRNA precursor. In the case of the target mimics, this was the reverse of the mature miRNA mimic sequence itself. Only homozygous lines were used for results in any part of this study.

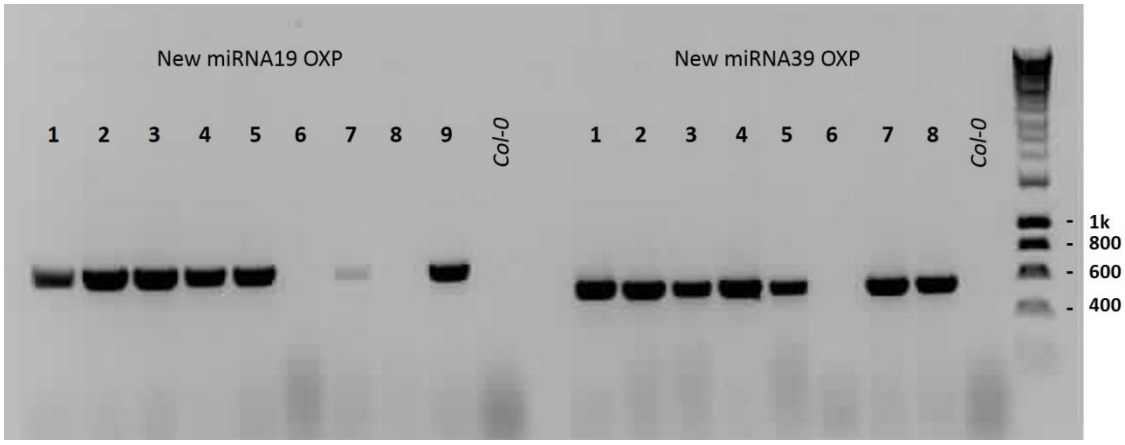


Figure A8: *Genotyping of novel microRNA19 and novel microRNA39 overexpression lines from T2-generation plants. Genotyping undertaken using a forward primer form the 35S promoter and a reverse primer from the far end of the precursor sequence being overexpressed. Lines with strong bands taken forward for study.*

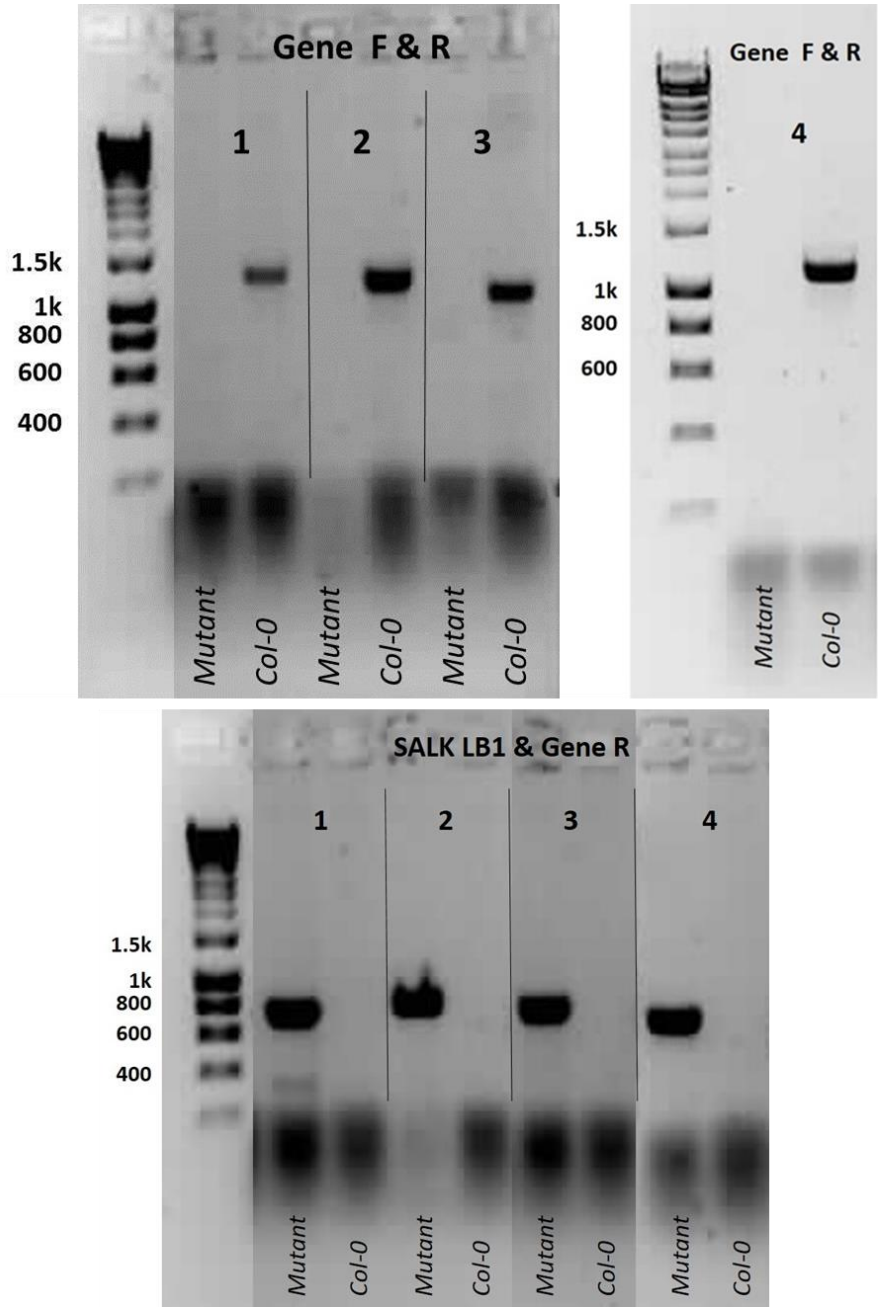


Figure A9: *Genotyping of potential novel microRNA target plants (1 - SALK_097650, 2 - SALK_057686, 3 - SALK_042297, 4 - SALK_055098). Genotyping undertaken using both the par of gene-specific primers (see section A1) and the SALK 'LB1.3' forward primer alongside the reverse gene-specific primer.*

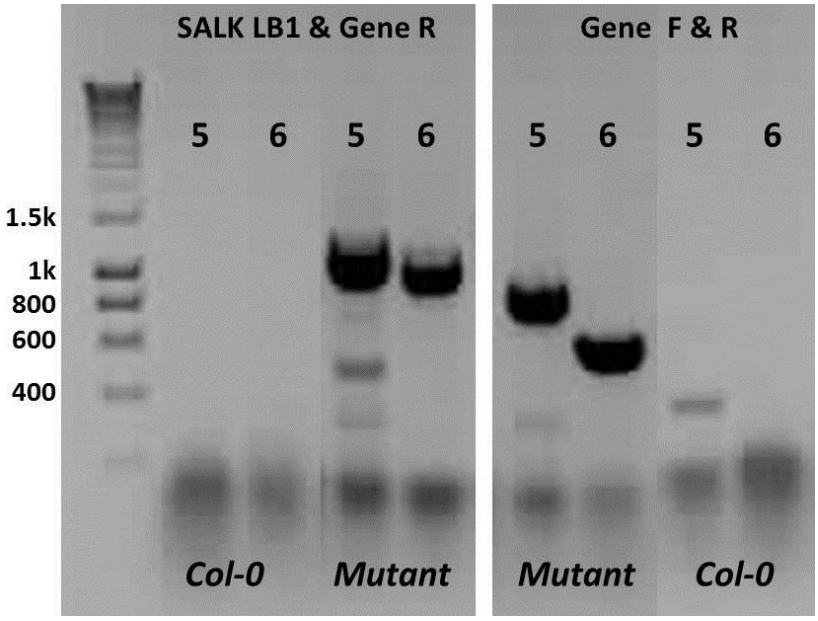


Figure A10: *Genotyping of potential novel microRNA target plants (5 - SALK_029936, 6 - SALK_068624). Genotyping undertaken using both the par of gene-specific primers (see section A1) and the SALK 'LB1.3' forward primer alongside the reverse gene-specific primer.*

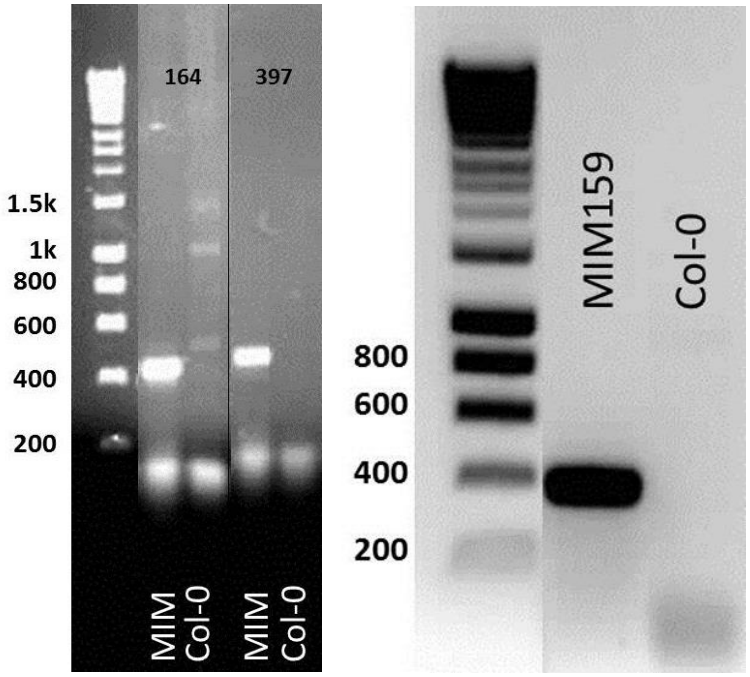


Figure A11: *Genotyping of selected Target Mimic lines (T2 generation plants) - MIM164, MIM397, and MIM159. Genotyping undertaken using a forward primer form the 35S promoter and the reverse of the Target mimic sequence.*