



ELUCIDATING METABOLIC AND
GENETIC REGULATORS OF
PHOSPHATE UPTAKE AND
UTILISATION IN BRASSICA

A thesis submitted by

Dahlia Shahbuddin

for the degree of Doctor of Philosophy

School of Agriculture

University of Reading

October 2019

TABLE OF CONTENTS

TABLE OF CONTENTS	ii
TABLE OF FIGURES	viii
TABLE OF TABLES	xvi
TABLE OF APPENDICES	xix
DECLARATION	xx
DEDICATION	xxi
ACKNOWLEDGEMENTS	xxii
LIST OF ABBREVIATIONS	xxiii
ABSTRACT	1
CHAPTER 1 INTRODUCTION AND RESEARCH OBJECTIVES	2
1.1 PHOSPHORUS	3
1.2 PHOSPHORUS IN THE ENVIRONMENT	3
1.3 PHOSPHORUS IN THE SOIL	5
1.4 PLANT UPTAKE OF P FROM THE SOIL SOLUTION	7
1.5 PLANT ADAPTATIONS TO IMPROVE SOIL PI AVAILABILITY AND ACQUISITION	8
1.5.1 Changes in root system architecture (RSA).....	8
1.5.1.1 Reduction of primary root growth and formation of lateral roots.	9
1.5.1.2 Development of root hairs	11
1.5.2 Formation of cluster roots (proteoid roots).....	12
1.5.3 Root carboxylate exudation	13
1.5.4 Secretion of acid phosphatases.....	13
1.5.5 Symbiotic associations with mycorrhizal fungi	14
1.5.6 Genes encoding proteins with transport activity.....	14
1.5.7 Phosphate transporters (PHT).....	14
1.6 REGULATION OF PLANT RESPONSES TO Pi AVAILABILITY	15
1.6.1 Types of signalling response to Pi deficiency.....	15
1.6.2 Transcription factors	17
1.6.2.1 SPX-domain containing proteins	19
1.6.3 Other intracellular/systemic signals in the regulation of Pi deficiency responses	20

1.7 CHANGES IN PLANT METABOLISM.....	24
1.7.1 Membrane lipid remodelling.....	24
1.8 EXPRESSION QUANTITATIVE TRAIT LOCI (eQTL).....	32
1.9 BRASSICAS.....	33
1.9.1 Triangle of U theory.....	33
1.9.2 Whole genome triplication event (WGT).....	35
1.9.3 Brassicas: <i>B. rapa</i>	36
1.9.4 Brassicas: <i>B. napus</i>	37
1.10 RESEARCH OBJECTIVES.....	37
CHAPTER 2 GENERAL MATERIALS AND METHODS	39
2.1 PLANT MATERIALS.....	40
2.1.1 <i>B. rapa</i>	40
2.1.2 <i>B. napus</i>	40
2.2 GROWTH CONDITIONS.....	41
2.3 GROWTH MEDIA AND NUTRIENT SOLUTIONS	41
2.3.1 Pot experiments	41
2.3.2 Hydroponics experiments	43
2.4 PLANT P ANALYSES.....	47
2.4.1 Pi analysis	47
2.4.2 Total P analysis	49
2.5 DATA ANALYSIS	51
2.6 DNA ISOLATION.....	51
2.6.1 DNA quantification and integrity.....	52
2.7 RNA ISOLATION.....	52
2.7.1 RNA quantification and integrity.....	52
2.8 PRIMER DESIGN.....	53
CHAPTER 3 REGULATION AND REMOBILISATION OF PHOSPHATE (Pi) WITHIN BRASSICAS DURING DEVELOPMENT AND THE IMPLICATIONS FOR SEED PHOSPHORUS USE EFFICIENCY (PUE).....	54
3.1 AIMS AND OBJECTIVES.....	55
3.2 BACKGROUND.....	55
3.3 MATERIALS AND METHODS.....	57
3.3.1 <i>B. rapa</i> physiological responses to Pi availability (six treatments) in compost mixture.....	57
3.3.2 <i>B. napus</i> physiological responses to Pi availability (two treatments) in different growth stages in compost mixture	57
3.3.3 <i>B. rapa</i> transcriptional responses to Pi availability in hydroponics	60
3.3.3.1 Hydroponics growth conditions and Pi analysis	60

3.3.3.2 RNA extraction and cDNA synthesis	60
3.3.3.3 Candidate reference gene selection, target genes and qPCR primer design	61
3.3.3.4 Quantitative real-time PCR (qPCR)	65
3.3.3.5 Quantitative real-time PCR (qPCR) data analysis	65
3.4 RESULTS.....	66
3.4.1 <i>B. rapa</i> plant growth responses to six different compost P concentrations	66
3.4.2 Changes in shoot Pi concentration with increasing compost Pi concentration in <i>B. rapa</i>	69
3.4.3 Changes in biomass and total P content of <i>B. napus</i> tissues grown with two different rates of P during development	71
3.4.4 Changes in Pi concentration of <i>B. napus</i> tissues grown with two different rates of P availability during development	74
3.4.5 Transcriptional responses to Pi availability	76
3.4.5.1 RNA quantification and integrity	76
3.4.5.2 Primer specificity of target reference genes and target genes	78
3.4.5.3 Gene expression changes of <i>B. rapa</i> under Pi deficiency.....	80
3.5 DISCUSSION	83
3.5.1 Growth response of <i>B. rapa</i> to increasing Pi availability	83
3.5.2 Evaluation of <i>B. napus</i> response and development under contrasting Pi availabilities.....	83
3.5.3 Transcriptional responses to Pi availability in <i>B. rapa</i>	85
CHAPTER 4 IDENTIFICATION OF THE MAJOR REGULATORS UNDERLYING TRANS- EXPRESSION QUANTITATIVE TRAIT LOCI (e-QTL) HOTSPOTS IN <i>B. RAPA</i>	89
4.1 AIMS AND OBJECTIVES	90
4.2 BACKGROUND.....	90
4.3 DNA AND RNA SEQUENCING TECHNOLOGIES	92
4.4 OXFORD NANOPORE SEQUENCING MinION	94
4.4.1 The principle of MinION sequencing.....	95
4.5 MATERIALS AND METHODS.....	96
4.5.1 Candidate genes and primer design	96
4.5.2 Bacterial artificial chromosome (BAC) sequencing	97
4.5.2.1 BAC DNA sub-culture and single colony preparation	98
4.5.2.2 BAC DNA isolation	98
4.5.2.3 BAC DNA confirmation and validation.....	99
4.5.2.4 BAC DNA precipitation	99
4.5.2.5 Sequencing library preparation for MinION sequencing.....	100

4.5.2.6 MinION flowcell preparation	102
4.5.2.7 MinION sample loading	103
4.5.2.8 MinION sequencing	103
4.5.2.9 MinION data analyses	104
4.5.3 Identification and characterisation of <i>PHO1</i> homolog genes.....	106
4.5.3.1 RNA extraction and cDNA synthesis	106
4.5.3.2 Reverse transcription polymerase chain reaction (RT-PCR)..	108
4.5.3.3 Cloning of <i>PHO1</i> transcripts	110
4.5.3.4 Restriction analysis	110
4.5.3.5 Plasmid sequencing	111
4.6 RESULTS.....	111
4.6.1 Candidate genes	111
4.6.2 BAC confirmation and validation.....	118
4.6.3 BAC extraction using large construct DNA extraction kit.....	119
4.6.4 BAC DNA sequencing using Oxford Nanopore Technologies MinION.....	121
4.6.4.1 MinION sequencing: Data analysis	121
4.6.5 Identification and characterisation of <i>PHO1</i> homolog transcripts	125
4.6.5.1 Reverse transcription polymerase chain reaction (RT-PCR)..	125
4.6.5.2 Cloning <i>PHO1</i> transcripts	126
4.6.5.3 Colony PCR and restriction/digestion analysis	126
4.7 DISCUSSION	138
4.7.1 Potential for alternatively spliced <i>PHO1</i> transcripts.....	138
4.7.2 Potential structural variation in <i>PHO1</i> paralogs.....	139
4.7.3 Conclusion	141

CHAPTER 5 VARIATION IN LIPID REMODELLING IN CULTIVARS OF <i>B. NAPUS</i> AND <i>B. RAPA</i> (R-O-18) UNDER Pi SUFFICIENT (P+) AND Pi DEFICIENT (P-) CONDITIONS.	143
5.1 BACKGROUND.....	144
5.2 AIMS AND OBJECTIVES	147
5.3 MATERIALS AND METHODS.....	147
5.3.1 Plant materials and stress treatment	147
5.3.2 Lipid extraction	149
5.3.3 Quantitative lipid analysis	149
5.4 RESULTS.....	152
5.4.1 Changes of lipid molecular species of 24 <i>B. napus</i> genotypes and <i>B. rapa</i> R-o-18 during Pi deficiency	152
5.4.2 Lipid analysis of seven genotypes in detail	155
5.4.2.1 Changes of lipid molecular species: Comparison between DGDG and PC	156
5.4.2.2 Changes of lipid molecular species: Comparison between SQDG and PC	160
5.5 DISCUSSION.....	164

5.5.1 Changes in DGDG varied between Brassica lines under Pi deficiency	164
5.5.2 Sulfolipids increased in Pi deficient Brassicas	165

CHAPTER 6 VARIATION OF GENE EXPRESSION IN CULTIVARS OF *B. NAPUS* AND *B. RAPA* (R-o-18) UNDER Pi SUFFICIENT (P+) AND Pi DEFICIENT (P-) CONDITIONS. 167

6.1 BACKGROUND.....	168
6.2 AIMS AND OBJECTIVES.....	169
6.3 MATERIALS AND METHODS.....	169
6.3.1 Plant materials and stress treatment	169
6.3.2 Total RNA extraction	170
6.3.3 RT-PCR: Pi responsive target genes and primer design.....	171
6.3.4 RNA-sequencing library preparation.....	176
6.3.5 Data analysis.....	176
6.4 RESULTS.....	177
6.4.1 Determination of RNA quantity and quality of 24 lines of <i>B. napus</i> for subsequent RT-qPCR	177
6.4.2 Determination of total RNA quantity and quality extracted from selected lines of <i>B. napus</i> and <i>B. rapa</i> R-o-18 for RNA-seq.....	180
6.4.3 RT-PCR of lipid and other Pi responsive genes.....	184
6.4.3.1 Quantification of Pi responsive genes in <i>B. napus</i>	184
6.4.3.2 Identification of lipid genes.....	186
6.4.4 Syntenic genes of <i>B. napus</i>	186
6.4.5 RNA-seq analyses.....	189
6.4.5.1 Differential gene expression analysis.....	189
6.4.5.2 MA plot.....	191
6.4.5.3 Pi response-regulated genes in genotypes under P+ and P- conditions.....	192
6.4.5.4 Variation in the expression of lipid metabolism genes across five cultivars of <i>B. napus</i>	193
6.5 DISCUSSION.....	198

CHAPTER 7 GENERAL DISCUSSION AND FUTURE RESEARCH PERSPECTIVE **204**

7.1 GROWTH AND REMOBILISATION OF Pi DURING DEVELOPMENT.....	205
7.2 KEY Pi RESPONSIVE GENES ARE ALSO DIFFERENTLY EXPRESSED IN <i>B. RAPA</i>	206
7.3 EVIDENCE OF QUADRUPLICATION OF A KEY REGULATOR OF PLANT Pi TRANSLOCATION IN <i>B. RAPA</i>	207
7.4 SIGNIFICANT VARIATION WAS OBSERVED IN LIPID PROFILES AND THE	

GENES INVOLVED IN LIPID METABOLISM BETWEEN DIFFERENT CULTIVARS OF <i>B. NAPUS</i>	208
7.5 CONCLUSION.....	211
REFERENCES	212
APPENDICES	240
APPENDIX 1	240
APPENDIX 2	246
APPENDIX 3	248
APPENDIX 4	286
APPENDIX 5	295

TABLE OF FIGURES

Figure 1.1. Phosphorus (P) cycle in the soil – plant continuum. The interactions between different forms of P in the soil. Organic P (Po) will undergo mineralisation process to form inorganic P (Pi). Then, Pi will go through desorption processes into the soil solution pool. Some of the Pi in the soil solution will be brought to the rhizosphere by diffusion process and it is available for uptake by the plant root (Adapted from Shen et. al, 2011).

..... 6

Figure 1.2. Plant adaptations to increase Pi availability in soil to maintain cellular functions (extracted from Lopez-Arredondo, 2014). Strategies including modification of root system architecture (RSA) by developing lateral roots and root hairs to increase the surface area for P uptake. The involvement of several genes (*tir1*, *pdr2*, *scr*, *lpr1/2*), sugar, and hormones contributed in controlling root branching and meristem responses. In addition, the secretion of organic acids (OAs) and phosphatases from roots and arbuscular mycorrhizal (AM) symbiosis interaction with the root system in some plants enhance Pi availability and uptake by roots. PHO1 is a protein that plays a key role in loading Pi into the xylem to maintain Pi translocation from the root to the shoot 10

Figure 1.3. The complexity of Arabidopsis P signalling pathways. Arrows linking boxes indicate positive regulation, and blunt ends indicate negative regulation. Dashed arrows indicate potential regulation. Arrows within boxes indicate increase or decrease in hormone. Extracted from Hammond and White (2011) 18

Figure 1.4. Types of SPX-domain containing proteins and the structure of the SPX domain. Extracted from Puga et al., 2017. (a) SPX-domain containing proteins of representative fungi (*Saccharomyces cerevisiae*), plants (*Arabidopsis thaliana*) and mammals (*Homo sapiens*) denoted as Sc, At and Hs, respectively. SPX domain present at the N-terminal either alone or associated with another domain; MFS, EXS, RING, VTC/DUF202, ANK, ANK/GDPD or CitMHS. (b) Model of SPX domain; ribbon model (left) and simplified model (right). The tyrosine (Y) and lysine (K) residues at $\alpha 2$ and $\alpha 4$ denoted by PBC (red) and KSC (yellow). PBC and KSC together function as high-affinity InsP binding region 23

Figure 1.5. Membrane glycerolipid classes and structure in Arabidopsis. The highly conserved (through evolution from cyanobacteria to higher plants) glycerolipid composition with 2-carbon glycerol backbone esterified to fatty acids (denoted as sn 1 and sn-2), and the polar head is at the position sn-3. They are categorised as a very low phosphoglycerolipid content, mainly phosphatidylglycerol (PG), phosphatidylcholine (PC) and phosphatidylinositol (PI), and very high non-phosphorus glycolipids contents, which are galactoglycerolipids (MGDG, DGDG) and sulfolipid (SQDG). Illustration extracted from (Boudière et al., 2014)..... 30

Figure 1.6. The Triangle of U theory. Extracted from Nagaharu U (1935). The figure describes the genomic relationships between six species in genus Brassica (n refers to the haploid number of chromosomes). Three diploid species, *Brassica rapa* (2n=20, AA), *Brassica nigra* (2n=16, BB) and *Brassica oleracea* (2n=18, CC) have interspecific hybridised to produce the three new allotetraploid species *Brassica juncea* (2n=36, AABB), *Brassica napus* (2n= 38, AACC) and *Brassica carinata* (2n=34, CCBB) 34

Figure 1.7. Phylogeny of Brassica relatives marking the relative placement of lineage divergence and polyploidy events. Polyploidies are named according to the Arabidopsis convention of α most recent in the lineage of Arabidopsis, β (second most recent) and γ (eudicot paleohexaploidy). Extracted from Tang and Lyons, 2012. 35

Figure 1.8. Whole genome triplication event in *Brassica rapa* is evident by comparing the ten chromosomes of *Brassica rapa* genome (horizontal axis) and the five chromosomes of *Arabidopsis thaliana* genome (vertical axis). The colour code show the triplication copies of Arabidopsis genes in Brassica chromosomes (Wang et al., 2011b). 36

Figure 2.1. A. *Brassica rapa* plants grown in 1 L plastic pots containing compost and sand media (75%/25%) in the glasshouse. **B.** *Brassica napus* plants grown in 1 L plastic pots containing compost and sand media (75%/25%) in the controlled environment room. 42

Figure 2.2. Image of the hydroponics tanks used for the hydroponic culture of *Brassica napus* at **A.** week 1, **B.** week 2 and **C.** week 3 after sowing. 46

Figure 2.3. Pi and total P analysis using molybdate-blue assay according to protocol by Ames (1966). **A.** Plant freeze-dried sample **B.** The free P was extracted using TissueLyser (Qiagen) **C.** The samples were kept on ice to avoid phosphatase activity during extraction **D.** Colorimetric analysis on plates **E.** The Thermo Scientific Multiscan (Thermo Scientific) used to measure the P content at λ_{620} wavelength. 48

Figure 2.4. Plant-acid digestion for total P analysis. **A.** Ground plant samples **B.** The samples were kept in air tight tubes **C.** Acid digestion procedures were conducted in the fume-hood. **D.** The recommended arrangement of MARSXpress vessels on the turntable **E.** MARS 6 microwave digestion system used to digest the samples. 50

Figure 3.1. *Brassica napus* images at different stages of growth development. **A.** Two-leaf stage **B.** Four-leaf stage **C.** and **D.** First-flower opens stage **E.** and **F.** Middle-flowering stage **G.** Seed stage **H.** Mature plants. 59

Figure 3.2. Shoot dry weights of *Brassica rapa* plants growing in compost with increasing additions of phosphate harvested after 14 days (**A**), 28 days (**B**), and 42 days (**C**). Plants were grown under glasshouse conditions. Each data point represents mean \pm SEM (n=5). Data were analysed using a REML procedure, followed by Fisher's unprotected least significant difference (LSD) test, so that data points with different letters are significantly different ($P < 0.05$). Equations for fitted lines are (A) $y = -0.0086x^2 + 0.0215x + 0.0127$, $R^2 = 0.7748$ (B), $y = -0.3704x^2 + 0.5771x + 0.0786$, $R^2 = 0.8677$ (C) $y = -1.8931x^2 + 3.0857x + 0.7368$, $R^2 = 0.482$ 67

Figure 3.3. Growth response of *Brassica rapa* plants to different additions of phosphate (Pi) to the compost (P1= 0.0, P2= 0.075, P3= 0.15, P4= 0.225, P5=0.45 and P6= 1.35 g L⁻¹ of compost mix) at 28 days after sowing. Scale bar: 10 cm. 68

Figure 3.4 Shoot phosphate (Pi) concentration of *Brassica rapa* plants growing in compost with increasing additions of Pi harvested after 14 days (A), 28 days (B), and 42 days (C). Plants were grown under glasshouse conditions. For harvest after 14 days, data are a single value due to low biomass for measuring shoot Pi. Data points for all other data represent mean \pm SEM (n=5). Data were analysed using a REML procedure, followed by Fisher's unprotected least significant difference (LSD) test, so that data points with different letters are significantly different (P<0.05) 70

Figure 3.5. Total biomass per plant of *Brassica napus* at six growth stages; two-leaf stage (1), four-leaf stage (2), flowering stage (3), first flower opens stage (4), seed filling stage (5) and maturity stage (6) at four different tissues; leaf (red), stem (orange), flower (blue) and pod silique and seed (green) grown in compost with no added phosphate (P1) and 0.225 g L⁻¹ (P4). Plants were grown under control environment laboratory conditions. At the early stage of harvest (harvest 1 and 2), whole plant was used due to low biomass. Data points for all other data represent mean \pm SEM (n=5)..... 72

Figure 3.6. Total P content and distribution between tissues throughout the growth cycle of *Brassica napus* plants; two-leaf stage (1), four-leaf stage (2), flowering stage (3), first flower opens stage (4), seed filling stage (5) and maturity stage (6) at four different tissues; leaf (red), stem (orange), flower (blue) and pod silique and seed (green) grown in compost with no added phosphate (P1) and 0.225 g L⁻¹ (P4). Plants were grown under control environment laboratory conditions. At the early stage of harvest (harvest 1 and 2), whole plant was used due to low biomass. Data points for all other data represent mean \pm SEM (n=5)..... 73

Figure 3.7. Tissue phosphate (Pi) concentration in *Brassica napus* plants; leaf (blue), stem (brown), flower (grey), and pod, silique and seed (orange) grown in compost with no added Pi (P1) and .225 g L⁻¹ (P4). Plants were grown under controlled environment conditions and harvested at two-leaf stage (1), four-leaf stage (2), flowering stage (3), first flower opens stage (4), seed filling stage (5), and maturity stage (6). For harvest 1 and 2, data are not provided due to low biomass for measuring Pi. Data points for all other data represent mean \pm SEM (n=5), except harvest 4 in treatment P4, (n=3). 75

Figure 3.8. Gel electrophoresis of RNA samples from pooled leaf tissues of *Brassica rapa* plants grown under P+ and P- conditions. RNA was loaded into a 1.2% agarose gel and run at 100 V for 60 minutes. RNA was visualised using ethidium bromide (EtBr) under UV light. RNA image of all the samples showed two bands, 28S rRNA and 18S rRNA. Column 1: 1 kb ladder, Column 2-7: Leaf RNA of *Brassica rapa* under P+ and P- conditions with replicate number. 77

Figure 3.9. Gel electrophoresis of PCR products for nine reference genes from *Brassica rapa* leaf tissues. PCR products was loaded into a 2% agarose gel and run at 100 V for 60 min. PCR products were visualised using ethidium bromide (EtBr) under UV light. PCR products of all samples showed nine reference genes with predicted amplicon size; *TIP41* (128bp), *Tubulin* (135 bp), *Cyclin* (144 bp), *TATA Box* (82 bp), *SAND* (126 bp), *PP2A* (104 bp), *GAPCP1* (95 bp), *18s rRNA* (87 bp) and *Actin* (144 bp). L= 100 bp ladder 79

Figure 3.10 Melt curves from qPCR of nine *Brassica rapa* reference genes performed on an Applied Biosystem OneStep+ PCR system to identify single products 80

Figure 3.11. Fold changes in gene expression for 14 target genes under P deficiency treatment in *Brassica rapa* relative to full nutrient control. Data represent the fold change difference between P deficient and full nutrient control plants. The error bars represent the standard error of the means of three replicates. *Genes whose expression is significantly ($P < 0.05$) different between treatment and control 82

Figure 4.1. Comparison between genetic and physical positions of 14,257 eQTL. **A.** Physical positions (bp) of individual genes (dots) against genetic locations (cM). **B.** The number of eQTL associated with individual gene expression markers across the *Brassica rapa* A genome. Horizontal dotted lines represent empirical thresholds for the upper 99% and 95% confidence interval for the Poisson distribution, assuming an equal distribution of GEMs and eQTL. Markers with values higher than these lines are defined as eQTL hotspots. (Hammond et al., 2011)..... 91

Figure 4.2. MinION sequencer device; **A.** consumable flowcell contains sensing chemistry, nanopore ad electronics; **B.** sample added to flow cell; **C.** sensor chip with multiple nanopores; **D.** Sensor chip works with custom ASIC for control and data processing **E.** USB power device. 96

Figure 4.3. Overview of the BAC DNA sequencing workflow. Ligation Sequencing Kit (SQK-LSK108) used for a library preparation, a 60 min protocol producing 1D reads (one strand of the fragment)..... 102

Figure 4.4. MinION device on the run for 48 hours. 103

Figure 4.5. Workflow for Canu software to generate read data including three stages: correction (green), trimming (red), and assembly (purple). Picture extracted from Koren et al., 2017. 105

Figure 4.6. Location of four candidate genes at the *trans* eQTL hotspots on Chromosome A06 of *Brassica rapa* (Hammond et al., 2011). Source: *Brassica rapa* clone finder (www.ncbi.nlm.nih.gov/clone/). **A.** *PHO1_A* (Bra019686) **B.** *PHO1_B* (Bra019688) **C.** *PHO1_C* (Bra019689) **D.** *PHO1_D* (Bra019690) 109

Figure 4.7. Sequence alignment of four putative PHO1 in *Brassica rapa* to determine region of homology from the sequences. The high similarity of the sequences is shown in green, yellow is for lower similarity and red refers to very low similarity of the sequences. Alignment was performed using Geneious software. *Brassica rapa* PHO1 IDs are **1.** Bra019690, **2.** Bra019689, **3.** Bra019688, **4.** Bra019686 116

Figure 4.8. Electrophoresis gel of PCR fragments of *Brassica rapa* leaf and root sample. Left and right lane represents a size of 100 bp ladder. **1.** Leaf gDNA **2.** Root gDNA **3.** Leaf cDNA **4.** Root cDNA **5.** Non-template control (NTC). **a.** Bra019686 (gDNA= 3668, cDNA=192 bp) **b.** Bra019688 (gDNA=231 bp, cDNA=147 bp) **c.** Bra019689 (gDNA and cDNA have the same size 207 bp) **d.** Bra019690 (gDNA=242 bp, cDNA=152 bp) **e.** Bro18004703 (160 bp) positive control (+ve). 117

Figure 4.9. Validation of BAC DNA using PCR with BAC end specific primers. PCR products were run on 1.2% agarose gel electrophoresis. The front (F) and rear (R) BAC ends were verified for each BAC and their expected sequence lengths are given in the parentheses. **A.** 1F: KBrB-029J08F (226 bp), 1R: KBrB-029J08R (169 bp), 2F: KBrB-003E10F (150 bp), 2R: KBrB-003E10R (198 bp), 3F: KBrB-063F11F (244 bp), 3R: KBrB-063F11R (226 kb) **B.** 4F: KBrH-038K12F (177 bp), 4R: KBrH-038K12R (228 bp), 5F: KBrH-102C10F (185 bp), 5R: KBrH-102C10R (233 bp)..... 118

Figure 4.10. BAC DNA single colony on agar plates with 12.5 µg µL⁻¹ chloramphenicol after incubation at 37 °C overnight. From top left; **A.** KBrB-063F11, **B.** KBrH102C10 **C.** KBrB029J08 **D.** KBrH038K12 **E.** KBrB003E10..... 120

Figure 4.11. Illustration of the plasmid sequenced from the MinION and which was cut to get the actual BAC sequence..... 122

Figure 4.12. The locations of *PHO1* homolog genes on CLC Sequence Viewer 7.8.1 (Qiagen). **A.** (*PHO1_A*), **B.** (*PHO1_B*), **C.** (*PHO1_C*) and **D.** (*PHO1_D*) on BAC 1 and BAC 3 sequence..... 124

Figure 4.13. Gel image of *PHO1* Reverse Transcription Polymerase Chain Reaction (RT-PCR) of cDNA from *Brassica rapa* R-o-18 leaf samples. **A.** Lane 1 from left: FastRuler Middle Range DNA Ladder (L), lane 2 to 9 is RT-PCR for cloning of *PHO1* transcripts. **B.** Gel of an aliquot of the RT-PCR products after PCR optimisation targeting *PHO1_F*+ *PHO_2224_R*..... 125

Figure 4.14. **A.** LB agar with 50 mg L⁻¹ of Kanamycin plate containing 20 µL transformed culture. **B.** LB agar with 50 mg L⁻¹ of Kanamycin plate containing 200 µL transformed culture..... 126

Figure 4.15. Gel image of colony PCR and HindIII restriction analysis of colony PCR of 30 randomly selected colonies containing *PHO1* pCR4 BLUNT TOPO vector. Band <80 bp are not shown. 127

Figure 4.16. Restriction maps of targeted region of **A.** XM_009150437 (*PHO1_A*) and XM_018652610 (*PHO1_B*), **B.** XM_018652651 (*PHO1_C*) and XM_009150438 (*PHO1_D*) = pCR4 BLUNT TOPO sequencing vector..... 129

Figure 4.17. Alignment of five *PHO1_A* splice variants in CLC Sequence Viewer 7.8.1 (Qiagen). Row from top: splice variants 1,2,3,4,5 for plasmid 2,7,23,26 and 22, respectively..... 132

Figure 4.18. Final alignment of *PHO1* paralog genes in CLC Sequence Viewer 7.8.1 (Qiagen). Row from top: XM_018652651.1 (*PHO1_C*), XM_009150437.2 (*PHO1_A*), Plasmid 2, Plasmid 12, and XM_018652610.1 (*PHO1_B*)..... 136

Figure 4.19. Homology tree of *Brassica rapa* PHO1 plasmids and NCBI paralogs produced from CLC Sequence Viewer 7.8.1 (Qiagen)..... 137

Figure 5.1. A-F: Lipid extraction procedure using acidic-chloroform-methanol method. **G.** The samples were filtered before running on the electrospray ionisation tandem mass spectrometry (ESI-MS/MS). 151

Figure 5.2. Digalactosyldiacylglycerol/phosphatidylcholine (DGDG/PC) ratio based on DGDG and PC concentrations measured by electrospray ionisation tandem mass spectrometry (ESI-MS/MS) in leaf tissue samples from 24 lines of *Brassica napus* and one line of *Brassica rapa* (R-o-18) (Table 5.1) grown hydroponically in the glasshouse under sufficient (P+) and deficient (P-) Pi conditions. Bars are based on ratios of three independent biological replicates and five technical replicates per treatment. 154

Figure 5.3. Delta ratio digalactosyldiacylglycerol/phosphatidylcholine (DGDG/PC) based on DGDG and PC concentrations measured by electrospray ionisation tandem mass spectrometry (ESI-MS/MS) on leaf tissue samples from 24 lines of *Brassica napus* and one line of *Brassica rapa* (R-o-18) (Table 5.1) grown hydroponically in the glasshouse under sufficient (P+) and deficient (P-) Pi conditions. Bars are based on ratios of three independent biological replicates and five technical replicates per treatment 155

Figure 5.4. Leaf digalactosyldiacylglycerol (DGDG) and phosphatidylcholine (PC) concentrations measured by electrospray ionisation tandem mass spectrometry (ESI-MS/MS) on six lines of *Brassica napus* (BA-099, BA-224, BA-044, BA-102, BA-218 and BA-210) and one line of *Brassica rapa* (R-o-18) (Table 5.1) grown hydroponically in the glasshouse under sufficient (P+) and deficient (P-) Pi conditions. Bars are based on ratios of five independent biological replicates and five technical replicates per treatment. Error bars represent mean \pm SEM (n=5). * DGDG or PC concentrations whose significantly different (P<0.05) using one-way ANOVA test between P+ and P-.. 157

Figure 5.5. Digalactosyldiacylglycerol/phosphoglycerol (DGDG/PC) ratio measured by electrospray ionisation tandem mass spectrometry (ESI-MS/MS) on six lines of *Brassica napus* (BA-099, BA-224, BA-044, BA-102, BA-218 and BA-210) and one line of *Brassica rapa* (R-o-18) (Table 5.1) grown hydroponically in the glasshouse under sufficient (P+) and deficient (P-) Pi conditions. Bars are based on the ratios of five independent biological replicates and five technical replicates per treatment. Error bars represent mean \pm SEM (n=5)..... 158

Figure 5.6. Delta digalactosyldiacylglycerol/phosphatidylcholine (DGDG/PC) ratio based on DGDG and PC concentrations measured by electrospray ionisation tandem mass spectrometry (ESI-MS/MS) on six lines of *Brassica napus* (BA-099, BA-224, BA-044, BA-102, BA-218 and BA-210) and one line of *Brassica rapa* (R-o-18) (Table 5.1) grown hydroponically in the glasshouse under sufficient (P+) and deficient (P-) Pi conditions. Bars are based on the ratios of five independent biological replicates and five technical replicates per treatment 159

Figure 5.7. Leaf sulfoquinovosyldiacylglycerol (SQDG) and phosphatidylcholine (PC) concentrations measured by electrospray ionisation tandem mass spectrometry (ESI-MS/MS) six lines of *Brassica napus* (BA-099, BA-224, BA-044, BA-102, BA-218 and BA-210) and one line of *Brassica rapa* (R-o-18) (Table 5.1) grown hydroponically in the glasshouse under sufficient (P+) and deficient (P-) Pi conditions. Bars are based on the ratios of five independent biological replicates and five technical replicates per treatment. Error bars represent mean \pm SEM (n=5). * SQDG or PC concentrations whose significantly different (P<0.05) using one-way ANOVA test between P+ and P- 161

Figure 5.8. Sulfoquinovosyldiacylglycerol/phosphatidylcholine (SQDG/PC) ratio measured by electrospray ionisation tandem mass spectrometry (ESI-MS/MS) on six lines of *Brassica napus* (BA-099, BA-224, BA-044, BA-102, BA-218 and BA-210) and one line of *Brassica rapa* (R-o-18) (Table 5.1) grown hydroponically in the glasshouse under sufficient (P+) and deficient (P-) Pi conditions. Bars are based on the ratios of five independent biological replicates and five technical replicates per treatment. Error bars represent mean \pm SEM (n=5)..... 162

Figure 5.9. Delta sulfoquinovosyldiacylglycerol/phosphatidylcholine (SQDG/PC) ratio based on SQDG and PC concentrations measured by electrospray ionisation tandem mass spectrometry (ESI-MS/MS) on six lines of *Brassica napus* (BA-099, BA-224, BA-044, BA-102, BA-218 and BA-210) and one line of *Brassica. rapa* (R-o-18) (Table 5.1) grown hydroponically in the glasshouse under sufficient (P+) and deficient (P-) Pi conditions. Bars are based on the ratios of five independent biological replicates and five technical replicates per treatment. 163

Figure 6.1. Determination of RNA integrity using 1.5% denaturing formaldehyde gel electrophoresis for RNA samples isolated from the leaves of 24 different *Brassica napus* lines grown hydroponically under Pi deficient and Pi sufficient conditions. 1-48 = sample number indicated in Table 6.5, L= NEB ssRNA ladder. 179

Figure 6.2. Determination of RNA integrity using 1.5% denaturing formaldehyde gel electrophoresis of RNA samples isolated from the leaves of five different *Brassica napus* lines and one *Brassica rapa* grown hydroponically under Pi sufficient (P+) and Pi deficient (P-) conditions. Lanes 1-42 = samples 1-42 defined in Table 6.6, L= NEB ssRNA ladder..... 183

Figure 6.3. Electrophoresis gel image of PCR products to test the specificity of primers designed to Pi responsive genes in *Brassica napus*. **A.** Gel of PCR products for PHT1:5_A, PHT1:5_C, SSP4_C, GPP1_A2, GPP1_C, and SQD2_A1 with their predicted product sizes. **B.** Gel of PCR products for SQD2_A2, SQD2_C, SPX2_A and SPPX2_C with their predicted product sizes. Column 1 and 2 represent different leaf cDNA samples BA-099 and BA-101, respectively; NTC = no template control (water); L= 100 base pair (bp) NEB ladder..... 185

Figure 6.4. Electrophoresis gel image of PCR products to test the specificity of primers designed to Pi responsive genes in *Brassica napus*. **A.** Gel of PCR products for PP2A1_C, PP21_A and SSP4_C with their predicted sizes in leaf cDNA sample of BA-099. **B.** Gel of PCR products for SQD2_A3 with its predicted size in leaf cDNA sample

of BA-099. **C.** Gel of PCR product for GPP1_A1 with its predicted size. Column 1 and 2 represent different leaf cDNA samples BA-099 and BA-101, respectively, NTC= Non-template control column (water), L= 100 base pair (bp) NEB ladder 185

Figure 6.5. Electrophoresis gel image of PCR products to test the specificity of primers designed to Pi responsive genes *DGD2* and *MDG2* of A and C genomes (1-4) using gDNA (A and B) and cDNA (C and D) from two lines of *Brassica napus* (31= BA-099, 32= BA-101) with their expected sizes. L= 100 base pair (bp) NEB ladder, NTC= No template control (water). 187

Figure 6.6. Electrophoresis gel image of PCR products to test the specificity of primers designed to lipid metabolism genes from *Brassica napus*. **A.** *DGD1_A* with expected size 130 bp **B.** *DGD2_C*, *MDG1_A1*, *MDG1_A2* and *MDG1_C*, with expected size of 229 bp, 118 bp, 214 bp and 137 bp, respectively. Column 1 and 2 represent different leaf cDNA samples BA-099 and BA-101, respectively, NTC = No template control (water), +ve = positive control (*Brassica rapa* Bro18004703 with size of 160 bp), L= 100 base pair (bp) NEB ladder..... 188

Figure 6.7. Gene expression distribution boxplot of \log_2 (FPKM+1) transformed values against cultivars/lines. The X-axis represents the different samples; same sample with the same colour-code (black, red, green, blue, turquoise represents BA-044, BA-099, BA-210, BA-218, BA-224, respectively). The Y-axis represents the relative abundance of \log_2 (FPKM+1)..... 190

Figure 6.8. Gene expression distribution boxplot of \log_2 (FPKM+1) transformed values against treatment. The X-axis represents the different treatments; same treatment with the same colour-code (Black, red represents P-sufficient, P-deficient, respectively). The Y-axis represents the relative abundance of \log_2 (FPKM+1)..... 190

Figure 6.9. The MA plot; the relationship between the expression change (M) and average expression strength (A); genes with adjusted p-values <0.05 are marked in red. 191

TABLE OF TABLES

Table 1.1. List of <i>Arabidopsis</i> lipid biosynthetic genes with possible involvement in lipid remodelling. NA, data not available; others, see abbreviation lists. Table extracted from Nakamura, 2013	27
Table 2.1. Stock nutrient solution composition. Seven stock solutions were prepared for the hydroponics experiment	44
Table 2.2. The protocol of inducing the Pi deficiency for the hydroponic plants. The stock solutions were changed more frequently after day 21 to ensure there was no residual Pi in the tanks	45
Table 3.1. Calculations for different measurement of Phosphorus use efficiency (PUE)	56
Table 3.2. List of potential reference genes for qPCR analysis in <i>Brassica rapa</i>	62
Table 3.3. Selected Pi responsive genes from <i>Brassica rapa</i> R-o-18, primer sequence (5' to 3'), amplicon size and primer efficiency	63
Table 3.4. Efficiency of reference gene primer pairs used for qPCR amplification	64
Table 3.5. Restricted Maximum Likelihood analysis (REML analysis) of <i>Brassica rapa</i> growth at six different Pi treatments	68
Table 3.6. Quantity and quality data for RNA extracted from leaf of <i>Brassica rapa</i> plants grown hydroponically under glasshouse conditions. RNA quantity and quality were determined using a Nanodrop 2000 spectrophotometer	77
Table 4.1. List of the primers, sequence, and length of expected genomic DNA (gDNA) and complementary DNA (cDNA) to amplify genes of interest; <i>PHO1_A</i> (<i>Bra019686</i>), <i>PHO1_B</i> (<i>Bra019688</i>), <i>PHO1_C</i> (<i>Bra019689</i>) and <i>PHO1_D</i> (<i>Bra019690</i>).....	97
Table 4.2. BAC libraries constructed by insertion of DNA fragments from <i>Brassica rapa</i> ssp. <i>pekinensis</i> cv. Chiifu into a vector <i>Escherichia coli</i> (<i>E. coli</i>) as a host.....	98
Table 4.3. List of the primers (forward and reverse) for BAC DNA end sequences and expected length of the PCR products.....	100
Table 4.4. Oxford Nanopore Technologies (ONT) barcodes used for BAC sequencing Total BAC DNA yield after DNA extraction and amount of nuclease free water (NFW)	

added to the stock DNA after DNA precipitation to prepare 1 µg in 45 µL for MinION sequencing. 101

Table 4.5 A. Primers designed to amplify cDNA. **B.** The expected sizes of *PHO1* cDNA. 107

Table 4.6. Internal primers sequenced to get complete sequence of plasmid 16 111

Table 4.7. Genes whose expression is significantly differentially altered by low Pi availability in *Brassica rapa* and co-locate to an eQTL hotspot on Chromosome A06 (Hammond et al., 2011)114

Table 4.8. BAC DNA extraction using large construct DNA extraction kit (Qiagen) quantification using Nanodrop 2000 and Qubit fluorometer. Total DNA were calculated based on Qubit reading and total volume of sample..... 120

Table 4.9. Total BAC DNA yield after DNA extraction and amount of nuclease free water (NFW) added to the stock DNA after DNA precipitation to prepare 1 µg in 45 µL for MinION sequencing 121

Table 4.10. Comparison of predicted and sequenced BAC sizes..... 123

Table 4.11. Transcript length, expected cut sites and fragment length of *PHO1* homologs restricted with HindIII enzymes 128

Table 4.12. Plasmid sequenced with the transcript number, nucleotide length, peptide length, closest BLASN and BLASTP hits, and predicted PHO1 on chromosome A06 *Brassica rapa* 133

Table 5.1. Twenty-four *Brassica napus* genotypes from the ASSYST (BA) population used to quantify genetic variation in gene expression responses to low Pi availability. Seeds were provided by the University of Nottingham 148

Table 5.2. Ranking of delta digalactosyldiacylglycerol/phosphatidylcholine (DGDG/PC) ratio in common samples from the initial experiment and the detailed experiment. Rank highest to the lowest of six lines of *Brassica napus* and *Brassica rapa* R-o-18..... 160

Table 6.1. List of the target Pi responsive genes identified from the literature in *Arabidopsis thaliana* and the syntenic analysis of orthologous copies in *Brassica napus* and the paralogous copies within the *Brassica napus* sub-genomes (LF, MF1 and MF2). 172

Table 6.2. List of the primers used in this study, the location of Pi responsive target genes in chromosome and the length of the amplified fragments (cDNA and gDNA). 173

Table 6.3. List of genes involved in lipid metabolism in *Brassica napus*. Primer sequence from 5' to 3', associated Arabidopsis gene ID, chromosome location, predicted amplified cDNA and gDNA fragments, and melting temperature in PCR174

Table 6.4. List of genes involved in lipid metabolism in *Brassica rapa*. Primer sequence from 5' to 3', associated Arabidopsis gene ID, chromosome location, predicted amplified cDNA and gDNA fragments, and melting temperature in PCR175

Table 6.5. RNA quantity and quality of 48 leaf RNA samples from 24 *Brassica napus* genotypes, BnASSYST (BA) grown hydroponically under Pi deficient (P-) or Pi sufficient (P+) conditions. RNA quantity and quality were determined using a Qubit 3.0 fluorometer and a NanoDrop 2000 spectrophotometer 178

Table 6.6. RNA quantity and quality of total RNA extracted from the leaves of five lines of *Brassica napus* and *Brassica rapa* R-o-18 plants grown hydroponically for 30 days under P sufficient (P+) or deficient conditions (P-). Samples were analysed using a NanoDrop 2000 spectrophotometer and Qubit 3.0 fluorometer181

Table 6.7. Expression levels of genes associated with lipid metabolism from leaf samples of five lines of *Brassica napus* (BA-044, BA-099, BA-210, BA-218, and BA-224) and levels of delta ratio digalactosyldiacylglycerol/phosphatidylcholine (DGDG/PC) high (H) and low (L) in response to Pi availability. The expression ratios were calculated as fold changes in transcript abundance between P+ and P-. Expression ratio of 1 indicates no difference in expression, >1 and <1 indicate up-regulated and down-regulated, respectively. Where there are no values, no transcript was detected. 195

TABLE OF APPENDICES

Appendix 1. Transcript of seven plasmid sequenced; Five splice variants of PHO1_A (Plasmid 2, 7, 23, 26, 22), PHO1_B (Plasmid 12), PHO1_D (Plasmid 24).	240
Appendix 2. Script used to process BAM files and HISAT in RNA-seq analysis.....	246
Appendix 3. List of upregulated genes of <i>Brassica napus</i> in response to Pi availability. Genes with q value < 0.05 and fold change value > 2.0.....	248
Appendix 4. List of down-regulated genes of <i>B. napus</i> in response to Pi availability. Genes with q value < 0.05 and fold change value < 0.5	286
Appendix 5. List of top five upregulated genes of <i>Brassica napus</i> BA-044, BA-099, BA-210, BA-218 and BA-224 in response to Pi availability. Genes with q value <0.05 and fold change.....	295

DECLARATION

I confirm this is my own work and the use of all materials from other sources has been properly and fully acknowledged.

Dahlia Shahbuddin

Reading, October 2019.

DEDICATION

To my parents, Shahbuddin Mohd Fiah and Latifah Abd Kahar.....

ACKNOWLEDGEMENTS

Thank you, Allah, for giving me the strength and capability to do my PhD research. I would like to express my thank you to my honourable supervisor, John Hammond, for the continuous support throughout my PhD study especially in molecular research for his patience, motivation, sharing great ideas and immense knowledge. I would like to thank Andrew Goodall, for his endless support and guidance right from the beginning of my PhD journey. I wish to thank Ihsan Ullah for his input and advice to the project. I am so grateful for the technical support from Val Jasper, George Gibbings, and Liam Doherty. It has been a great experience to be a student at Faculty of Agriculture University of Reading. I would like to thank Richard Haslam, Harrie and Peter Easmond for helping me on lipid analysis at Rothamsted Research. A very special gratitude to my sponsor, University Science of Malaysia and Ministry of Higher Education, Malaysia for providing me the funding for the research. I must express my very profound gratitude to my parents, Shahbuddin Mohd Fiah and Latifah Abd Kahar for the love and encouragement throughout my PhD journey. I would like to thank my husband, Hazni Hafiz Mohd Tony for providing me with continuous support mentally, emotionally and financially for the past four years. To all my family members and friends, thank you for always be with me in good or bad times. Finally, thank you to my wonderful children Fahry, Faris and Aurora for always keeping me busy and sane. This PhD would never have succeeded without them.

LIST OF ABBREVIATIONS

A	Adenine
ADP	Adenosine Diphosphate
Al	Aluminium
AMF	Arbuscular Mycorrhizal fungi
ASIC	Application-Specific Integrated Circuit
ATP	Adenosine Triphosphate
BAC	Bacterial Artificial Chromosome
bHLH	Basic helix-loop-helix
C	Cytosine
Ca	Calcium
CGS	Candidate Gene Sequencing
DAG	Diacylglycerol
ddNTPs	Dideoxynucleotides
DEG	Differentially Expressed Genes
DGDG	Digalactosyldiacylglycerol
DNA	Deoxyribonucleic acid
dNTPs	Deoxyribonucleotides
EMF	Ectomycorrhizal fungi
eQTL	expression Quantitative Trait Loci
Fe	Iron
FPKM	Fragments Per Kilobase
G	Guanine
GDPD	Glycerophosphodiester Phosphodiesterase
InsP	Inositol Phosphates
IPK	Inositol Polyphosphate Kinase
LAH	Lipid Acyl Hydrolase
LB	Luria-Bertani
LF	Least Fractionated
lpr1	Low Phosphate Root 1
MeS	Methylation Sequencing
MF1	Most Fractionated 1
MF2	Most Fractionated 2
MGDG	Monogalactosyldiacylglycerol
miRNA	microRNAs
MOPS	3-(N-morpholino) propanesulfonic acid
MS	Mass Spectrometry
NFW	Nuclease Free Water
NGS	Next Generation Sequencing
NLA	Nitrogen Limitation Adaptation
NPC	Non-specific phospholipase
nt	Nucleotide
OA	Organic Acid
ONT	Oxford Nanopore Technology
P	Phosphorus
PA	Phosphatidic Acid
PAH	Phosphatidate Phosphohydrolase
PAP	Purple Acid Phosphatases

PBC	Phosphate Binding Cluster
PC	Phosphatidylcholine
PCR	Polymerase Chain Reaction
PE	Phosphatidylethanolamine
PG	Phosphatidylglycerol
PHT	Phosphate Transporter
PI	Phosphatidylinositol
Pi	Phosphate
PLC	Phospholipase C
PLD	Phospholipase D
Po	Organic Phosphate
polyP	Inorganic Polyphosphate
PS	Phosphatidylserine
PSI	Pi Starvation Induced
PSR	Pi-Starvation-Responsive
PUE	P Use Efficiency
PUpE	P Uptake efficiency
PUtE	P Utilisation efficiency
QTL	Quantitative Trait Loci
RBF	Running Buffer with Fluid mix
RCC	Relative Crowding Coefficient
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
RSA	Root System Architecture
SBS	Sequencing by Synthesis
SMRT	Single-Molecule Real-Time
SOC	Super Optimal broth with Catabolite
SOLiD	Sequencing by Oligo Ligation Detection
SQDG	Sulfoquiovosyldiacylglycerol
T	Thymine
TAE	Tris-Acetate-EDTA
TF	Transcription Factor
U	Uracil
WES	Whole-Exome Sequencing
WGD	Whole Genome Duplication
WGS	Whole Genome Sequencing
WGT	Whole Genome Triplication

ABSTRACT

Phosphorus (P) is one of the six essential macronutrients in plants. Plants only assimilate P in the form of phosphate (Pi). Plants have developed various adaptations to enhance Pi acquisition and its efficient use. The low availability of Pi in the soil, means many crop plants rely on Pi fertilisers to maintain their growth and development. The overuse of fertilisers pollutes the environment. To overcome this problem, it is crucial to understand the physiological and molecular mechanism of plant Pi uptake and use.

Revealing candidate genes underlying *trans* e-QTL hotspots on chromosome A06 of *B. rapa*, Phosphate transporter PHO1 (*PHO1*) genes provide insights on the putative genes which are responsible for Pi transportation and remobilisation in plants. DNA sequencing of five Bacterial Artificial Chromosome (BAC) clones of *B. rapa* ssp. *pekinensis* cv. Chiifu revealed the occurrence of four *PHO1* paralogues in tandem in the genome sequence. However, only three *PHO1* paralogue transcripts could be confirmed by cloning.

Membrane lipid remodelling occurred in 24 cultivars of *Brassica napus* and *Brassica rapa* R-o-18 photosynthetic membranes under Pi deficient conditions. Plants change their phospholipid compositions to non-phosphorus galactolipid and sulfolipid to release Pi for other cellular functions. Analysis of lipid profiles and expression of lipid metabolism genes in *Brassica rapa* R-o-18 grown under Pi starvation showed highly responsive lipid metabolism genes are Glycerophosphodiester phosphodiesterase 1 (*GDPD1*), Monogalactosyldiacylglycerol synthase 3 (*MGD3*) and Sulfoquinovosyl transferase 2 (*SQD2*).

Analysis of transcript abundance in leaf samples of *B. rapa* R-o-18 under Pi deficiency through qPCR showed increased transcript abundance for Aluminium activated malate transporter 1 (*ALMT*) by 7.6-fold, which is involved in organic acid (OA) exudation to improve phosphorus uptake efficiency (PUPE). Whole transcriptome sequencing using RNA-seq revealed 630 transcripts whose expression changed during Pi deficiency across five investigated *B. napus* lines; 481 genes were upregulated, and 148 genes were down-regulated in response to Pi deficiency. High throughput RNA-seq contributed to molecular identification and regulation underlying biochemical and physiological adaptations to Pi deficiency. These advances provide information of on candidate genes which might be useful in developing future crops with tolerance to low Pi availability.

CHAPTER 1 INTRODUCTION AND RESEARCH OBJECTIVES

1.1 PHOSPHORUS

Phosphorus (P) is an essential element for life. The name phosphorus is derived from the Greek 'phos' meaning "light", and 'phoros' meaning "bearer". It was first discovered in 1669 by Hennig Brandt, a German chemist who isolated P from urine. He discovered that pure P was colourless, solid and glowed in the dark. It was declared an element by the founder of modern chemistry, Antoine Lavoisier, in 1775 (Emsley, 2000). Phosphorus is a major component of all living cells. It is involved in several key cellular functions, including energy transfer by adenosine triphosphate (ATP). Adenosine diphosphate (ADP), which consists of an adenosine molecule bonded with two phosphate groups uses the energy derived from cellular respiration to bond with a third phosphate group, to become ATP. The stored energy in the bonds between the second and third phosphate molecules in ATP is used to drive cellular functions, by releasing the third phosphate group and converting back to ADP after releasing the energy. Phosphorus is a vital component in plant cell membranes, as the part of phospholipids making up the phospholipid bilayer. Another important role of P in cells is in the sugar-phosphate backbone of deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). DNA is an essential part of the process of carrying the genetic code from one generation to the next, providing the blueprint of all aspects of plant growth and development (Cooper, 2000; Malboobi et al., 2014).

1.2 PHOSPHORUS IN THE ENVIRONMENT

Phosphorus, which does not exist in a gaseous phase, has a purely terrestrial biogeochemical cycle. It is the eleventh most abundant element and makes up 0.012% of the earth's crust. Phosphorus is present in rocks and found associated with other minerals, but not as free element, due to its high chemical reactivity. Most rich deposits of P were formed in the ocean through sediments and decomposition of dead organisms (Smill, 2000b). Through weathering, P gradually becomes unavailable to plants due to

losses through runoff or absorption to soil particles. Phosphorus is not only vital to life, it also the main ingredient to produce fertilisers, detergents, matches, and pesticides. The main use of P is for the manufacture of the fertilisers which accounts more than 80% of global P use.

Phosphate (Pi) fertilisers are used in agricultural production systems to maintain crop yield and quality. In addition, other human activities that result in sewage and industrial wastes contribute to excessive P concentrations in the environment and can eventually cause pollution and eutrophication in freshwater aquatic systems (Schindler et al., 2016). Eutrophication in these P limited environments causes oxygen concentrations to be depleted in the water and reduces the ability of animals to live in these habitats (White & Hammond, 2009).

Food and feed production consumes 90% of the world demand for phosphorus (Smill, 2000a; Smill, 2000b; Childers et al., 2011). There is no substitute for P in growing crops and hence in food production. In modern agriculture, the main source of Pi fertiliser is from mined rock phosphate. The three largest producers of rock phosphate are Morocco, China and United States of America. Currently, Morocco controls the majority of world's high-quality phosphate rock reserves (Cordell, 2010). Based on the current usage, the existing rock phosphate reserves could be exhausted in the next 300-400 years (Cooper et al., 2011). The extraction of phosphorus reserves is likely to reach a maximum, also known as peak phosphorus, in the next 20-80 years when supply will no longer be able to keep up with demand (Steen, 1998; Cordell & White, 2011). This will result in production costs increasing and potentially the quality of phosphate rock decreasing. The duration of P stock availability is still in debate (Cordell & White, 2013; Walan et al., 2014). However, to mitigate the depletion of Pi reserves, P recovery and recycling of a larger proportion of the Pi passing through society and the environment is required (Cordell et al., 2009; Koppelaar & Weikard, 2013). By using Pi fertilisers

efficiently, we can ensure crops can grow optimally and provide a high yield with minimal use of Pi fertilisers.

1.3 PHOSPHORUS IN THE SOIL

Although the total amount of P can be high in some soils, plant roots only acquire P from the rhizosphere solution as inorganic phosphate (Pi), also known as orthophosphate, mainly in the form of H_2PO_4^- (Vance et al., 2003). The majority of soil P exists in the form of organic phosphate (Po), generally about 65% of the total P (Shen et al., 2011). Soil Po forms vital soil properties and its concentration can be affected by the changes in the land use and management. The Po contents significantly differs across land use, where arable soils have significantly lower amount of Po concentration (Stutter et al., 2015). Generally, P in soil can be divided into four different forms, (1) inorganic Pi, (2) Po, (3) adsorbed Pi and (4) mineral Pi. Only inorganic Pi is available for plants. While Po, adsorbed Pi and mineral Pi, need to go through transformation processes such as weathering, mineralisation and desorption to make them available for plants (Fig. 1.1). In the phosphorus cycle, available P will become unavailable through processes such as precipitation, immobilisation and adsorption (Schachtman et al., 1998; Shen et al., 2011).

Mineral Pi is made available through a weathering process over long periods of time to convert it to available Pi. While precipitation turns the available Pi to unavailable Pi by reaction with dissolved calcium (Ca) in alkaline soils, iron (Fe), and aluminium (Al) in acidic soils, depending on the size of mineral particles (Hinsinger, 2001; Oelkers & Valsami-Jones, 2008,). Soil Po is in stabilised forms which need to be mineralised through enzymatic processes involving soil organisms and plant roots to convert it to H_2PO_4^- or HPO_4^{2-} . Many factors like soil moisture, pH, temperature, and surface physical-chemical properties greatly influence mineralisation processes (Turner et al., 2007). Plants take up Pi from the soil solution, and then return Po back to the soil in

dead plant material or after consumption by animals and returned through their waste products as Po residues. Concentration of available Pi in the soils is also governed by the adsorption and desorption reactions. Adsorption is the process of fixing the soluble Pi on solid surface of soil particles to form a stable form of unavailable Pi, while desorption is the slow release of adsorbed Pi from its stable state to the soil solution (Fig. 1.1).

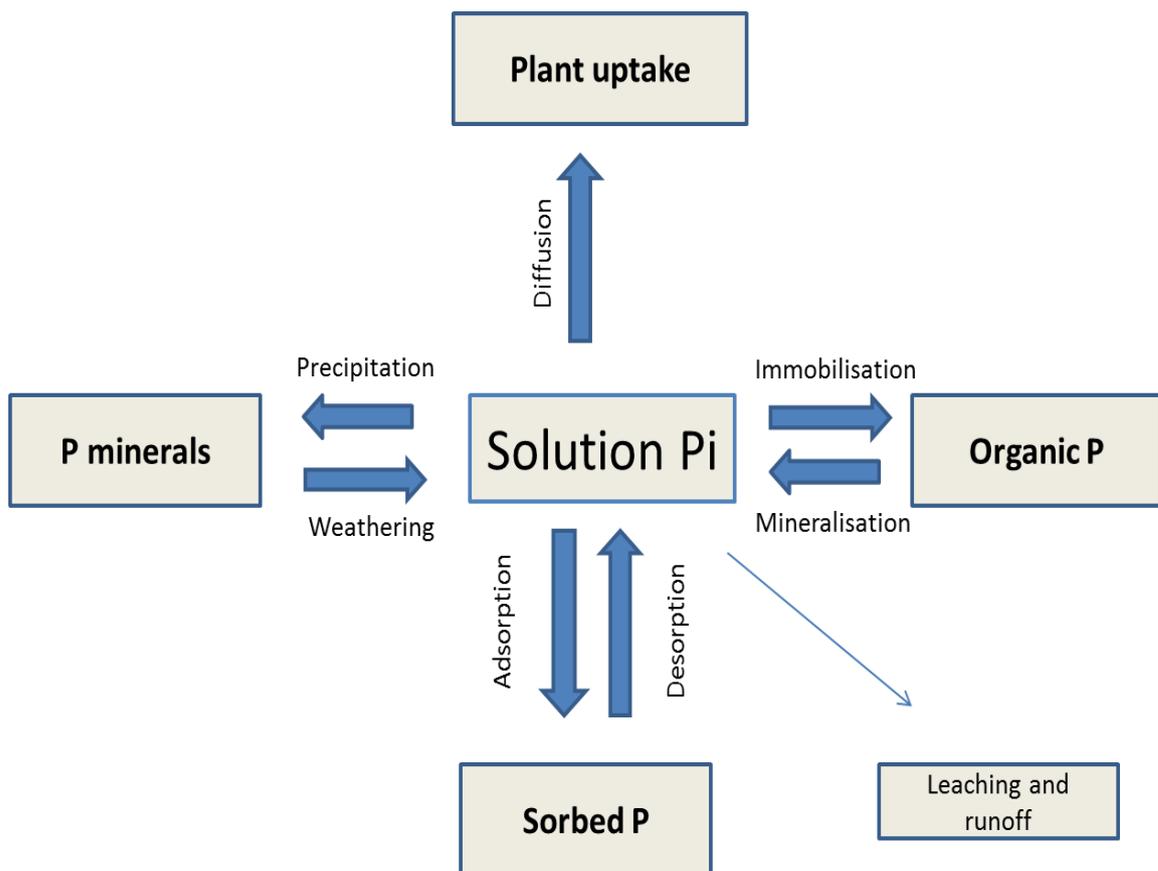


Figure 1.1. Phosphorus (P) cycle in the soil – plant continuum. The interactions between different forms of P in the soil. Organic P (Po) will undergo mineralisation process to form inorganic P (Pi). Then, Pi will go through desorption processes into the soil solution pool. Some of the Pi in the soil solution will be brought to the rhizosphere by diffusion process and will be available for uptake by the plant root (Adapted from Shen et al, 2011).

Soil solution Pi concentrations rarely exceed 10 μM (Bielecki, 1973) depending on the land uses and soil properties (Stutter et al., 2015), and are much lower than in plant tissues and intracellular compartments which are approximately 5 to 20 mM

(Raghothama, 1999). Therefore, plants must take up Pi from the soil against a strong concentration gradient. The low concentration and poor mobility of plant available Pi in soil has resulted in plants evolving a series of metabolic and developmental adaptations focused at increasing the acquisition of Pi (Wang et al., 2004; Chiou & Lin, 2011).

1.4 PLANT UPTAKE OF P FROM THE SOIL SOLUTION

Rock phosphate (apatites, strengite, and variscite) when added directly to the soil will undergo low dissolution rate due to weathering processes and their stable condition. Although the application of rock Pi has proved relatively efficient for plant growth in acidic soils, the release of available P is generally too slow to match the crop demand (Shen et al., 2011). Commonly in developed agricultural systems, inorganic P fertilisers are used. Processed inorganic fertilisers derived from rock phosphates are mixed with phosphoric acid to become a soluble powder, such as superphosphate and mono-ammonium phosphate. The availability of Pi from these inorganic fertilisers is greater, with improved solubility, allowing them to meet the demands of the growing crop.

Manure fertiliser can also be applied to enhance soil fertility. The content of total P in manure is variable, where up to 70% of P is labile. The large amount of Po in manure can be mineralised to increase soil Pi concentration (Shen et al., 2011).

Since low Pi concentrations are present in the soil solution, the supply of Pi to the root surface occurs by diffusion (slow) rather than mass flow (fast) (Suriyagoda et al., 2014). Phosphate anions move from areas of high concentration to low concentration in the soil. Due to low concentrations of Pi in the soil solution, zones of depletion often occur around plant roots as Pi is taken up from the rhizosphere solution, but is not replaced quick enough due to the slow diffusion of Pi, with diffusion coefficients ranging from 10^{-12} to $10^{-15} \text{ m}^2 \text{ s}^{-1}$ (Schachtman et al., 1998).

1.5 PLANT ADAPTATIONS TO IMPROVE SOIL PI AVAILABILITY AND ACQUISITION

Higher plant species differ in their efficiency of Pi uptake from the soil solution, especially during low Pi availability (Shenoy & Kalagudi, 2005). Phosphorus in the form of Pi has a low mobility in the soil. Therefore, plants have developed phenotypic adaptation strategies to adapt to low P environments, including (i) maximising the soil exploration through root proliferation and elongation (Niu et al., 2013), and (ii) mutual association with mycorrhizal fungi.

1.5.1 Changes in root system architecture (RSA)

Generally, the topsoil contains more Pi compared to other soil layers due to the low mobility of Pi and the deposition of plant litter over time (Lopez-Arredondo et al., 2014). However, root architecture responses to low Pi availability differ significantly between each plant species. Studies by Lynch and Brown (2001) demonstrated that low soil P availability caused changes in their morphology, topology and distribution patterns in common bean roots (*Phaseolus vulgaris*). Root length density is the most important trait in wheat (*Triticum aestivum*) to improve the P absorption in the upper soil layers (Manske et al., 2000). In Arabidopsis, plants will develop more shallow roots, enhancing the growth of lateral roots and root hairs (Bates & Lynch, 1996). Phosphate starvation in shoots induces the formation of cluster (proteoid) roots in lupins (*Lupinus albus*) (Zhou et al., 2008) and affects the root angle in common bean (*Phaseolus vulgaris*) (Bonser et al., 1996). These adaptations broaden the root surface area, place the roots in the vicinity of the soil phosphate and increase the capacity of the plant to explore the topsoil.

1.5.1.1 Reduction of primary root growth and formation of lateral roots

In *Arabidopsis*, the first obvious effect of Pi deficiency is a reduction of primary root growth after plants are transferred from higher Pi concentration to a low Pi concentration (Sánchez-Calderón et al., 2006; Péret et al., 2011). Key genes involved in this process can be categorized in three groups. First, mutants hypersensitive to low Pi concentrations such as the Pi deficiency response 2 (*pdr2*) mutant. Second, mutants which are able to maintain primary root growth under low Pi concentrations, such as low Pi root (*lpr*) mutant, and third, mutants that are low P insensitive, for example *lpi1*, *lpi2*, *lpi3* and *lpi4* which showed long primary root despite low Pi availability (Reymond et. al., 2006; Sánchez-Calderón et al., 2006). Other plants, for example common bean, will increase their lateral root formation and elongation as well as increase adventitious root formation, especially in the upper layer of soil where Pi is usually more available (Lynch & Brown, 2001) (Fig. 1.2).

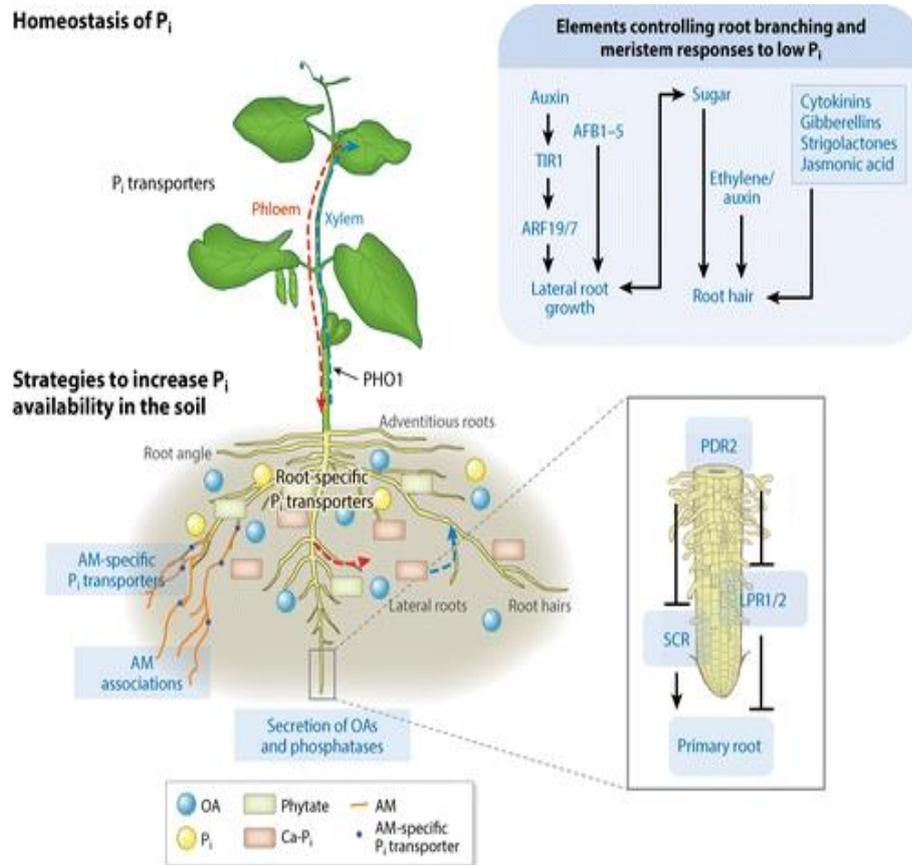


Figure 1.2. Plant adaptations to increase P_i availability in soil to maintain cellular functions (extracted from Lopez-Arredondo, 2014). Strategies including modification of root system architecture (RSA) by developing lateral roots and root hairs to increase the surface area for P_i uptake. The involvement of several genes (*tir1*, *pdr2*, *scr*, *lpr1/2*), sugar, and hormones contributed in controlling root branching and meristem responses. In addition, the secretion of organic acids (OAs) and phosphatases from roots and arbuscular mycorrhizal (AM) symbiosis interaction with the root system in some plants enhance P_i availability and uptake by roots. PHO1 is a protein that plays a key role in loading P_i into the xylem to maintain P_i translocation from the root to the shoot.

Changes in root architecture under P_i starvation are related to changes in plant growth regulators or phytohormones (Chiou & Lin, 2011; Peret et al., 2011). Phytohormones such as auxin, cytokinin, gibberellic acid, and ethylene play a role in P_i deficient root development (Jung & McCouch, 2013). In *Arabidopsis* and clover (*Trifolium repens*), auxin plays a role in lateral root development (Al-Ghazi et al., 2003; Dinh et al., 2012), while ethylene production induces changes in root hair density and length (Lynch &

Brown, 2001; Vance et al., 2003). The detailed studies of the root adaptation to Pi deficiency suggests that the root responses are genetically controlled rather than only metabolically controlled (Wang et al., 2010; Niu et al., 2013).

In the Arabidopsis accession Columbia (Col-0), the identification of Low Phosphate Root 1 (*lpr1*) as the major quantitative trait loci (QTL) controlling the root development under Pi starvation demonstrated the existence of genetic control (Reymond et al., 2006). The reduction of primary root growth correlates with the reduction in cell differentiation at the root meristem and also inhibits cell proliferation in the root tip zone (Ticconi et al., 2004; Svistoonoff et al., 2007). During this process, many genes and hormone-related genes are activated. Arabidopsis mutant *lpr1* and its paralog *lpr2* greatly reduce the inhibition of root growth (Svistoonoff et al., 2007).

1.5.1.2 Development of root hairs

Proliferation of root-hairs during low Pi conditions contributes to enhancing the total root-surface area by up to 70% and can be responsible for up to 90% of Pi uptake by plants as observed in barley and Arabidopsis (Fig 1.2) (Bates & Lynch, 2001; Gahoonia & Nielsen, 2003). The role of root hairs in enhancing Pi acquisition was determined in Arabidopsis mutants with low root hair numbers. The Pi uptake rate was drastically reduced compared to the wild-type under Pi deficiency. A relative crowding coefficient (RCC) value (K) was also used to measure competitive ability between wild-type and low root hair mutants. The K values in the wild-type plants were greater than one showing that root hairs increase the plants' competitiveness under Pi deficiency, whereas the Arabidopsis mutants with low root hair numbers had K values of less than one (Bates & Lynch, 2001). The higher K values demonstrated that Arabidopsis wild-type is highly competitive and dominant against the mutant (Bates & Lynch, 2001). Plants develop a high root hair density by the formation of shorter cells and increase the expression of master regulators of cell fate ENHANCER OF TRY AND CPC 1 (*ETC1*) gene (Savage et al., 2013). There are many gene mutants in Arabidopsis that promote

higher root hair density during Pi starvation such as the transcription factor genes *WRKY75*, *phl1/phr1*, SUMO E3 ligase gene *SIZ1* and the F-Box protein with WD40 domain (*fbx2*), inositol polyphosphate kinase (*ipk*) and Raf like Kinase (*hsp2*) (Miura et al., 2005; Chen et al., 2007; Bustos et al., 2010; Lei et al., 2011).

1.5.2 Formation of cluster roots (proteoid roots)

Another adaptation in response to Pi deficiency can be observed in most members of the Proteaceae family (Vance et al., 2003). Proteoid roots are bottle-brush like structures made of closely packed modified lateral roots which produce large root hairs to enhance Pi acquisition especially in the topsoil layer where small nutrient-rich patches are found (Skene, 1998; Dinkelaker et al., 1995; Adams et al., 2002). The super-abundance of root hairs in proteoid roots increases the surface area by 100-fold as compared to normal roots (Vance et al., 2003). Plants that develop proteoid roots normally do not form symbiotic mycorrhizal associations to enhance Pi availability (Skene, 1998; Lambers et al., 2013). Proteoid roots also produce a burst of organic acids to increase the availability of poorly available Pi in the rhizosphere (Lambers et al., 2003; Shane et al., 2004). In white lupin (*Lupinus albus*) and members of Proteaceae, the exudation occurs in the proteoid roots tips and elongation zones (Ryan et al., 2001). Proteoid roots secrete organic acids in a high concentration (10-100-fold higher than the 1-50 μM concentrations usually found in the soil solution) into the rhizosphere in the form of citrate and malate in white lupins (Lambers et al., 2013, Hunter et al., 2014). Citric acids will release P from iron to provide a rapid transport of P to the root to be taken up by a high-affinity transporter within the plant. In addition, the exudation of malate is more related to the proton release. A plasma-membrane-bound H^+ -pumping ATPase is involved in proton release to stimulate or inhibit the citrate production (Lambers et al., 2013). These specialised roots are able to 'mine' P from sorbed forms (Lambers et al., 2008). The exudation of organic acids anions also increases in the roots of other Pi-deficient plants like bean (*Phaseolus vulgaris*) and barley (*Hordeum vulgare*) (Hernandez et al., 2007;

Huang et al., 2008). The organic acid anions solubilise Pi from the metal (Al-, Fe-, and Ca- phosphates) by chelation process and desorb Pi from mineral surfaces, consequently increase soil Pi concentration (Ryan et al., 2001).

1.5.3 Root carboxylate exudation

Root exudate carboxylates (organic acids) such as citrate, malate and oxalate can increase P availability by releasing P from strongly sorbed forms for example iron phosphate and rock phosphate, in the rhizosphere during Pi deficiency (Lambers et al., 2008). These carboxylates release Pi bound to Al³⁺, Fe³⁺, and Ca²⁺ in the soil to increase its availability for uptake by roots (Ryan et al., 2001; Lopez-Arredondo et al., 2014).

1.5.4 Secretion of acid phosphatases

The production and secretion of different types of enzymes by roots, such as acid phosphatases and nucleases, into the rhizosphere also contributes to increase the Pi availability from the organic forms to make it available to the plant (Brinch-Pedersen et al., 2002; Lopez-Arredondo et al., 2014). Purple acid phosphatases (PAP) are one of the important enzymes involved in hydrolysis and mobilisation and recycling of P for cell metabolism (Tran et al., 2010). These enzymes can act on various organic molecules and are thought to be relatively stable in a wide range of pH and temperatures (Fig. 1.2) (Lopez-Arredondo et al., 2014). Studies of PAP in various plants, including tobacco (*Nicotiana tabacum*), tomato (*Solanum lycopersicum*) and Arabidopsis showed an increase of PAP activity under Pi starvation (Kaida et al., 2003; Bozzo et al., 2002; Wang & Liu, 2012). Some studies showed Pi starvation in the soil increased secretion of PAP in the rhizosphere. For example, in Arabidopsis, *Purple Acid Phosphatase 10* (*AtPAP10*) produces AtPAP10 protein that plays a significant role in plant adaptation to Pi deficiency (Wang et al., 2011a). However, some studies showed no significant response of PAP

activities in plants under Pi deficiency, for instance in common bean (*Phaseolus vulgaris*) (Yan et al., 2001).

1.5.5 Symbiotic associations with mycorrhizal fungi

The majority of plant species form symbiotic associations with soil microorganisms, especially with arbuscular mycorrhizal fungi (AMF) or ectomycorrhizal fungi (EMF), which can enhance Pi uptake from the soil, by expanding the volume of soil that can be explored and allowing interchange of nutrients in both directions (Fig. 1.2). The AMF and EMF play important roles in mobilisation of Pi from new sites in soil to Pi depletion zones that form around the root surface by extending their hyphae far beyond the Pi depletion zone, while the fungi in-turn receive photosynthetic carbon from the host plant (Smith & Read, 2008; Baker et al., 2015). The majority of the plants especially from the tropics form associations with AMF. On the other hand, EMF colonise only on some trees or shrubs in boreal and temperate regions (Lambers et al., 2008; Finlay, 2008).

1.5.6 Genes encoding proteins with transport activity

Under Pi deficiency, Pi is redistributed from mature, fully expended source leaves towards young, sink tissues within the plant (Ramaiah et al., 2011), and additional capacity to take up Pi from the rhizosphere is required. Therefore, many genes will be induced to enhance Pi acquisition and utilisation.

1.5.7 Phosphate transporters (PHT)

Phosphate uptake and transport is controlled by high-affinity Phosphate Transporters (PHT). The PHTs are proton/phosphate (H^+ , $H_2PO_4^-$) symporters, moving Pi ions cross the cell membrane against the gradient through proton-transporter protein. Proteins of the PHT family are grouped into five classes (PHT1, PHT2, PHT3, PHT4 and PHT5) based on their sequence identity, sub-cellular localisation and putative function. PHT1

members are localised at the plasma membrane and most of the *pht1* genes display an increase in gene expression during Pi deficiency (Lin et al., 2013). They are involved in Pi uptake and translocation (Nussaume et al., 2011; Gu et al., 2016; Liu et al., 2015). All the regulation takes place at the transcriptional level (Muchhal et al., 1996). Many of the PHT1 transporters are expressed in roots, but some are specifically expressed in shoot tissues (e.g. PHT1;1) and pollen grains (e.g. PHT1;4) (Karthikeyan et al., 2002; Mudge et al., 2002). Within the roots, the expression of Pi transporters is mostly concentrated in root epidermis and central cylinder, particularly in the root hair zone (Nussaume et al., 2011). PHTs have been identified and isolated from many species, including *Arabidopsis*, *Solanum tuberosum*, *Oryza sativa*, *Catharanthus roseus*, *Lycopersicon esculentum* and *Medicago truncatula* (Ren et al., 2014). In *Arabidopsis* and rice, the PHT1 family is comprised of nine and 13 members, respectively (Mudge et al., 2002; Zhao et al., 2015). The other PHT members involved in coordinating Pi transport through the organelles in the membrane have received less attention than PHT1. Other PHT members (PHT2/3/4) are localised to plastid (PHT2/4), mitochondria (PHT3) and Golgi (PHT4) membranes (Gu et al., 2016). Recent studies have reported the discovery of the first transporter PHT5;1/VPT1 involved in loading Pi into vacuoles (Pi storage) (Liu et al., 2015; Bucher & Fabianska, 2016).

1.6 REGULATION OF PLANT RESPONSES TO Pi AVAILABILITY

Under Pi starvation, plants will restrict their use of Pi only to the most important cellular functions and enhance their root system to acquire more Pi from the soil under Pi stress. This requires a complex set of signalling and regulatory processes.

1.6.1 Types of signalling response to Pi deficiency

Plants have evolved various responses to cope with Pi deficiency. These responses involve local and systemic signalling pathways (Thibaud et al., 2010; Chiou and Lin,

2011) (Figure 1.3). Local signals are local regulatory signals which are produced in the cell and move to the neighbouring cells through plasmodesmata (symplast) and the intercellular space (apoplast). These signals trigger responses to a local confined area. Many characteristic changes in RSA are controlled by the local Pi supply independently of the internal Pi content (Bates & Lynch, 1996; Thibaud et al., 2010; Shen et al., 2011), including; a size reduction of the root cells (Reymond et al., 2006), increase in the length and density of the root hairs (Bates & Lynch, 1996), inhibition of primary root growth, developed by a sensing mechanism located surrounding the root tip (Ticconi et al., 2004; Arnaud et al., 2010) and altered development to increase elongation of lateral roots in regions of high Pi availability to enhance Pi acquisition (Reymond et al., 2006; Haling et al., 2013). All these changes contribute to broaden the root surface area and enable roots to explore the soil to enhance Pi acquisition.

In contrast, systemic or long distance regulatory signals can carry information and move between roots and shoots or source and sink tissues in response to changes in internal Pi status in whole plants (Lin et al., 2014). Several molecules that are considered systemic signals to regulate responses in roots or shoots include photosynthates, phytohormones, microRNAs (miRNA), sugars, mobile mRNAs, small peptides, inositol phosphates (InsP) and P (Hammond & White, 2011; Chiou & Lin, 2011; Puga et al., 2017). These signals are involved in many processes such as Pi signalling, Pi recycling, Pi uptake and recovery, lipid metabolism, and response to metal ions and their absorption, binding, and translocation (Thibaud et al., 2010). This systemic regulation has been revealed in barley (*Hordeum vulgare*) and Arabidopsis, in which Pi uptake activity and expression of Pi transporters in local Pi depleted roots, were repressed by Pi supply to other parts of roots (Drew & Saker, 1984; Franco-Zorrilla et al., 2005). In systemic pathways, the vascular system including xylem and phloem provides a pathway for signal transportation throughout the plant (Turnbull & Lopez-Cobollo, 2013).

For optimal responses to Pi deficiency at various levels in different tissues, cross-talk among different systemic signalling pathways is important (Lin et al., 2014).

1.6.2 Transcription factors

Transcription factors (TF) are proteins that bind to specific DNA sequences and control the rate of the transcription from DNA to RNA (Jakoby et al, 2002). Transcription factors regulate the gene expression by binding to specific DNA promoter sequences or directly to the RNA polymerase. Phosphate Starvation Response 1 (PSR1), is thought to be the first TF that acts as a control regulator and has been shown to regulate various types of Pi-starvation-responsive (PSR) genes in Arabidopsis (Rubio et al., 2001). *AtPHR1* encodes a MYB-related TF and plays a central role in the regulatory network (Chiou & Lin, 2011) (Fig. 1.3). Similarly, homologs of *AtPHR1* have been characterised in other crop species, for example rice (*OsPHR2*), wheat (*TaPHR1*) and common bean (*PvPHR1*) (Valdes-Lopez et al., 2008; Wu et al., 2013; Wang et al., 2013). *AtPHR1* binds to a DNA imperfect palindromic sequence (GNATATNC) named P1BS, which is present in many PSR genes (Rubio et al., 2001; Bustos et al., 2010).

Another TF family involved in responses to Pi stress is the WRKY family. Studies on WRKY75, WRKY45 and WRKY6 have also shown them to regulate PSRs (Devaiah et al., 2007a; Chen et al., 2009; Wang et al., 2014a). In Arabidopsis, WRKY6 is involved in Pi stress response by inducing PHOSPHATE1 (*PHO1*) expression. Under low Pi supply, WRKY6 binds to a W-box consensus motif, and inhibits its binding to the *PHO1* promoter, this demonstrates that *PHO1* expression may reduce under Pi-stress environment (Chen et al., 2009).

Basic helix-loop-helix (bHLH) domain-containing TF have also been characterised in responses to Pi stress in Arabidopsis, AtbHLH32, and rice, OsPTF1 (Yi et al., 2005; Chen et al., 2007). *OsPTF1* and *bHLH32* genes are both upregulated under Pi deficiency whereas bHLH32 plays a negative role by repressing the expression of *PPCK*

phosphoenolpyruvate carboxylase kinase) genes, which are involved in glycolysis (Chen et al., 2007), and OsPTF1 serves as a positive regulator in roots and shoots to enhance tolerance to Pi stress in rice (Yi et al., 2005).

A cysteine-2/histidine-2 zinc finger (ZAT6) TF is regulated under Pi deficiency and plays a crucial role in plant growth and development (Zhang et al., 2014). In Arabidopsis, the overexpression of *ZAT6* under Pi-stress caused alteration of root architecture in older plants and Pi homeostasis (Devaiah et al., 2007b).

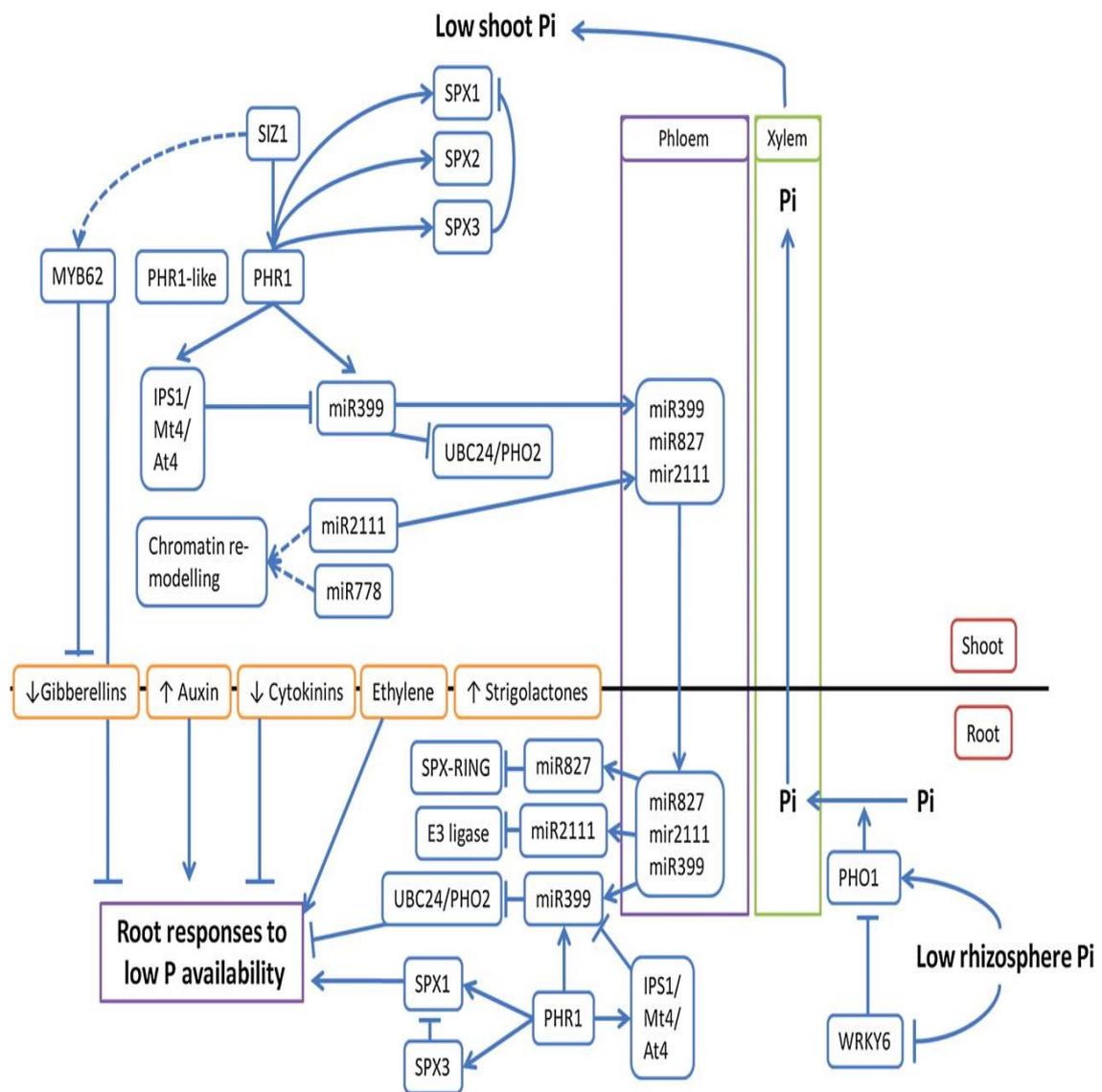


Figure 1.3. The complexity of Arabidopsis P signalling pathways. Arrows linking boxes indicate positive regulation, and blunt ends indicate negative regulation. Dashed arrows indicate potential regulation. Arrows within boxes indicate increase or decrease in hormone. Extracted from Hammond and White (2011).

1.6.2.1 SPX-domain containing proteins

Recent studies showed that several proteins containing the SPX-domain are essential in Pi transport, Pi-stress sensing and signalling in plants (Secco et al., 2012). In plants, proteins harbouring the SPX-domain can be grouped into four categories based on the existence of additional protein domains: (1) SPX, which exclusively harbours the SPX domain, and SPX proteins containing different domains at the C terminus (2) SPX-EXS, (3) SPX-MFS, and (4) SPX-RING (Chiou & Lin, 2011).

Proteins containing only the SPX domain can be classified into four groups in Arabidopsis named AtSPX1- AtSPX4 and six groups in rice named OsSPX1-OsSPX6 (Duan et al., 2008; Secco et al., 2012). Transcription analysis showed that with the exception of SPX4, the expression of all the SPX genes responds positively to Pi deficiency in Arabidopsis (Duan et al., 2008; Wang et al., 2014b; Puga et al., 2014). In contrast, SPX4 degrades under Pi deficiency, consequently releasing PHR2 (homolog to Arabidopsis PHR1) into the nucleus and regulates the expression of PSR genes as well as maintaining Pi homeostasis (Lv et al., 2014).

An essential pathway for Pi involving the sensing and transport of Pi from roots to shoots through the vascular cylinder of various tissues was initially described by the identification of *PHO1* (Hamburger et al., 2002; Wang et al., 2004). *PHO1*, which belongs to the SPX-EXS protein family, contains an SPX (SYG/Pho81/XPR1) tripartite domain in its N terminal and an ERD1/XPR1/SYG1 domain in the C-terminal region, and was first identified in yeast (*Saccharomyces cerevisiae*) (Wang et al., 2004; Secco et al., 2012). In Arabidopsis, *PHO1* localised to the endo-membranes, mainly the Golgi, in root cells, mediating Pi loading into the xylem from root to shoot either by directly regulating the Pi transporter or through signal transduction (Chen et al., 2009). The *PHO1* family is the only protein that contains the SPX and EXS domains in eukaryotes. The major role of SPX is to modulate the activities of PHR1 to regulate PSR genes, while the EXS domain is important for exporting Pi, although the EXS domain cannot

function by itself, at least in tobacco (*Nicotiana benthamiana*) (Wege et al., 2016). In Arabidopsis, the *PHO1* gene family contains 11 members, namely *PHO1*, and *PHO1;H1-H10* (Wang et al., 2004; Arpat et al., 2012). Studies on Arabidopsis revealed that *PHO1*, *PHO1;H1* and *PHO1;H3* have the same expression pattern and all were expressed in the root cell and localise to the Golgi and trans-Golgi network (Stefanovic et al., 2007; Arpat et al., 2012). Characterisation of *AtPHO1;H4* revealed its role in hypocotyl elongation under blue light (SHB1), and in seed development and flowering in Arabidopsis (Kang & Ni, 2006; Zhou & Ni, 2009; Zhou et al., 2009). *B. rapa* has the largest *PHO1* gene family, represented by 23 genes, whereas soybean (*Glycine max*) has 14 putative *PHO1* homologs, and Brachypodium (*Brachypodium distachyon*) and maize (*Zea mays*) have only two *PHO1* homologs each (He et al., 2013). Members of the *PHO1* family showed important roles in addition to the role of exporting Pi to the vascular system in different biological conditions in plants (Khan et al., 2014).

A recent discovery revealed that SPX-MFS domain protein plays a crucial role in Pi transport into the vacuole in Arabidopsis (Liu et al., 2015). The SPX-RING is another SPX subfamily containing the RING finger domain responsible for protein-protein interactions. The only characterised member of the SPX-RING family is the Nitrogen Limitation Adaptation (NLA) protein. Studies on the physiology and genetics of Arabidopsis showed that NLA regulates nitrate-dependent Pi homeostasis (Peng et al., 2007; Kant et al., 2011).

1.6.3 Other intracellular/systemic signals in the regulation of Pi deficiency responses

Plants must acquire sufficient nutrient from the soil to ensure their growth and reproduction. The interaction between leaf and root, through the vascular system allows systemic signals to operate. Specific nutrient-responsive micro RNAs (miRNA) were detected in the phloem sap as a signal of Pi stress (Buhtz et al., 2010). miRNAs are non-coding small RNAs, 20-24 nucleotides long, and produced from endogenous

hairpin-like primary transcripts that control gene expression and respond to changes in the environment and development (Rogers & Chen, 2013). MiR399 and miR827 are systemic signals which are strongly upregulated under Pi starvation (Liu et al., 2014). MiR399 recognise their target mRNAs by binding to P1BS *cis*-element, consequently regulating the target genes expression at a post-transcriptional level (Carthew & Sontheimer, 2009). In Arabidopsis, miR399 species are highly induced by Pi starvation (Chiou et al., 2006). Overexpression of miRNAs enhanced Pi uptake and translocation of Pi from the root to the leaf to maintain cellular Pi concentrations (Bari et al., 2006; Liu et al., 2012). Micrografting experiments revealed that mature miRNA primary transcripts are strongly induced and move from leaf to root via *Phosphate overaccumulator (PHO2)* regulatory system. *PHO2* encodes a ubiquitinating-conjugating E2 enzyme (UBC24) and functions as a repressor to prevent Pi over-accumulation by controlling Pi uptake and translocation (Bari et al., 2006). The results of genetic screening of *pho2* mutants in Arabidopsis showed that PHO2 mediates the protein degradation of PHO1 and transporter PHT1 at endomembranes (Liu et al., 2012). The role of miRNA as a PHO2 suppressor contributes to stabilise PHO1 protein and PHT1 and therefore improve Pi uptake under Pi starvation.

Another type of miRNA, miRNA827 targets NLA encoding SPX-RING- type ubiquitin ligase (Peng et al., 2007). Molecular mechanisms underlying NLA through the analysis of two *nla* mutations revealed that their suppressors are the Pi transport related genes *PHOSPHATE TRANSPORTER TRAFFIC FACILITATOR (PHF1)* and *PHT1*, therefore it plays an important role in regulating cellular Pi status as well as cellular N status (Kant et al., 2011; Liu et al., 2016).

A recent discovery of the SPX domain proteins (SPX1, SPX2 and SPX4) structure and function has revealed that they can bind InsP with high affinity as a signal in response to low Pi availability (Wild et al., 2016; Puga et al., 2017). SPX domain crystal structure consists of six alpha helices (α 1- α 6) in their functional region. Four of the α helices form

a tri-helix bundle in $\alpha 3$ - $\alpha 6$ (Wild et al., 2016). Tyrosine (Tyr) and Lysine (Lys) residues in $\alpha 2$ and $\alpha 4$ provide a Pi binding cluster (PBS) which allows low-affinity Pi binding. Another three Lys residues in $\alpha 4$ provide the Lys surface cluster (KSC). Both PBS and KLC clusters together form a very high positively charged surface believed to be an InsP binding site with high affinity via hydrogen bonds (Azevodo & Saiardi, 2017; Puga et al., 2017) (Fig 1.4).

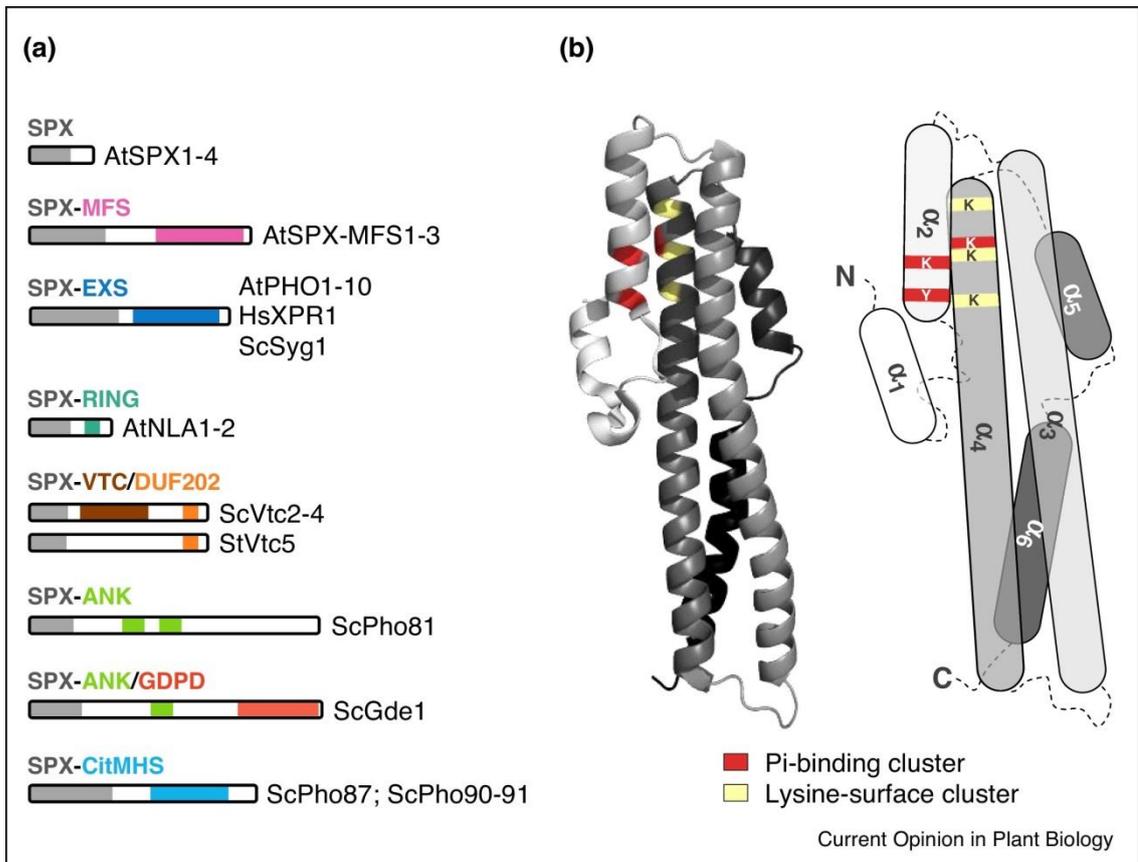


Figure 1.4. Types of SPX-domain containing proteins and the structure of the SPX domain. Extracted from Puga et al., 2017. (a) SPX-domain containing proteins of representative fungi (*Saccharomyces cerevisiae*), plants (*Arabidopsis thaliana*) and mammals (*Homo sapiens*) denoted as Sc, At and Hs, respectively. SPX domain present at the N-terminal either alone or associated with another domain; MFS, EXS, RING, VTC/DUF202, ANK, ANK/GDPD or CitMHS. (b) Model of SPX domain; ribbon model (left) and simplified model (right). The tyrosine (Y) and lysine (K) residues at $\alpha 2$ and $\alpha 4$ denoted by PBC (red) and KSC (yellow). PBC and KSC together function as high-affinity InsP binding region.

1.7 CHANGES IN PLANT METABOLISM

To maintain internal Pi-balance under Pi starvation, plants change several metabolic processes including changes to glycolysis process. The glycolysis changes vary between plant species (Huang et al., 2008). Under Pi limiting conditions, the utilisation of Pi in phosphorylating sugar containing metabolites were reduce and the low molecular weight phosphorylated metabolites were transformed to non-P containing di- and tri-saccharides such as sucrose, maltose, raffinose and 6-kestose consequently reduced levels of organic acids and produce larger amount of carbon in the tricarboxylic acid (TCA) cycle (Uhde-Stone et al., 2003; Ganie et al., 2015). Induction of enzymes that do not need ATP and ADP + Pi, such as P_{Pi}-dependent phosphofructokinase and pyruvate Pi dikinase involved in the glycolytic bypass and were secreted in the proteoid roots of white lupin (Plaxton & Tran, 2011). In common beans, reduction of organic acids of TCA were also reported in Pi-deficient roots (Hernandez et al., 2007). In contrast, the increased levels of di- and tri-saccharides suggest that glycolysis might be lowered in shoot and root of Pi-deficient barley plants (Huang et al., 2008).

1.7.1 Membrane lipid remodelling

Plants have adapted to Pi deficiency to minimise compromising photosynthesis efficiency and have developed mechanisms to ensure Pi supply for other Pi requiring cellular functions (Andersson et al., 2003; Lambers et al., 2012; Nakamura, 2013; Shimojima et al., 2015). Membrane lipid remodelling is one of these mechanisms to increase efficient use of Pi by extensively changing the membrane composition which is rich in phospholipids to non-phosphorus galactolipids and sulfoquinovosyldiacylglycerol (SQDG) sulfolipids during leaf development (Andersson et al., 2005; Lambers et al., 2012; Shimojima et al., 2013). Membrane phospholipids hold one-third of the Pi reserve; therefore the degradation of phospholipid releases Pi to increase PUE (Nakamura, 2013). There are two types of galactolipids; (i) monogalactosyldiacylglycerol (MGDG);

and (ii) digalactosyldiacylglycerol (DGDG). These galactolipids are the major lipids in the thylakoid membrane in chloroplasts. There are three types of MGDG isoforms in Arabidopsis; (i) MDG1, which is important for the majority of galactolipid synthesis in chloroplast and is localised to the inner membrane of chloroplast; (ii) MDG2 and (iii) MDG3; both provide substrates for DGDG accumulation under Pi starvation and target the outer membrane of chloroplasts (Kobayashi et al., 2007; Kobayashi et al., 2013). Similar to MGDG synthesis, there are two types of DGDG isoforms, DGDG synthase 1 (DGD1) which is involved in DGDG synthesis, and DGDG synthase 2 (DGD2) which plays a crucial role under Pi starvation (Shimajima et al., 2013). Under Pi replete conditions, MGDG accounts for approximately 50 mol% and DGDG accounts for 25 mol% of thylakoid membrane lipids in chloroplasts (Holzl & Dormann, 2007). Under Pi deficient conditions, DGDG is exported to extraplastidial membranes to compensate for reductions in phosphoglycerolipid contents in the membranes (Jouhet et al., 2004). A previous study of a null mutant in Arabidopsis (*dgd1*), lacking in its DGD1 showed a reduction in the amount of DGDG by 90% under normal growth condition, and only accumulated DGDG up to 60% of the amount present in the wild type under Pi starvation (Härtel et al., 2000). Both MGDG and DGDG are believed to play important roles in photosystem II (PSII) in the photosynthesis process (Mizusawa & Wada, 2012).

The membrane remodelling process involves two steps: first, hydrolysis of phospholipids to yield diacylglycerol (DAG) and the release of Pi; the second step is the synthesis of sulfolipids and galactolipids using DAG as a substrate (Nakamura, 2013). Genes involved in galactolipid (MGDG/DGDG) and sulfolipid (SQDG) synthesis are strongly induced, while those involved in synthesis of phospholipids are repressed in many plant species during low Pi availability (Hammond et al., 2003; Woo et al., 2012).

In Arabidopsis, all three MGDG genes (*MGD1-MGD3*) are expressed in photosynthetic tissue under Pi deficient conditions (Kobayashi et al., 2007). *MDG2* and *MDG3* showed four to ten-fold higher expression under short low Pi exposure, while *DGD1* and *DGD2*

were expressed under long-term low Pi exposure (Misson et al., 2005). In addition, MDG2 and MDG3 showed a significant role in membrane lipid remodelling (Kobayashi et al., 2009). MDG2 and MDG3 are involved in the production of MGDG and act as a substrate for DGDG synthesis. Therefore, high accumulation of DGDG was detected in Arabidopsis leaves during Pi deficiency (Kobayashi et al., 2009).

Transcripts encoding all three enzymes in the sulfolipid biosynthetic pathway have been identified, (*UDP-GLUCOSE PYROPHOSPHORYLASE3 (UGP3)*, *UDP-sulfoquinovose synthase SQD1*, and *sulfoquinovosyldiacylglycerol (SQDG) synthase SQD2*) and are coordinately induced in Arabidopsis plants with low Pi status (Hammond et al., 2003; Shimojima et al., 2013; Nakamura, 2013; Murakawa et al., 2014; Pant et al., 2015). Arabidopsis lipid metabolism genes are listed in Table 1.1.

Table 1.1. List of Arabidopsis lipid biosynthetic genes with possible involvement in lipid remodelling. NA, data not available; others, see abbreviation lists. Table extracted from Nakamura, 2013.

Gene	AGI code	Pi starvation	Expression tissues	Substrates	Subcellular localisation	Mutant phenotype under Pi starvation	Reference
<i>NPC</i>							
<i>NPC4</i>	At3g03530	Induced	Stamen, late embryo	PC>PE	Plasma membranes	Diminished upregulation of total PC-PLC activity	Nakamura et al., 2005
<i>NPC5</i>	At3g03540	Induced	Stamen, late embryo	PE>PC	Soluble fraction	Reduced DGDG contents, growth defect	Gaude et al., 2008
<i>PLD</i>							
<i>PLDz1</i>	At3g16790	Induced	NA	PC>>PE>PS	NA		Qin & Wang., 2002
<i>PLDz2</i>	At3g05630	Induced	Senescent leaf, late embryo	NA	Tonoplast	Defect in root elongation, decreased root PA and DGDG levels	Cruz-Ramirez et al., 2006; Ymaryo et al., 2008; Li et al., 2006a; Li et al., 2006b
<i>PAH</i>							
<i>PAH1</i>	At3g09560	NA	Dry seed	PA	Cytosol	Reduced DGDG accumulation, defects in overall growth in pah1 pah2	Nakamura et al., 2009
<i>PAH2</i>	At5g42870	NA	Dry seed	PA	Cytosol	Reduced DGDG accumulation, defects in overall growth in pah1 pah2	Nakamura et al., 2009

Table 1.1. continued

Gene	AGI code	Pi starvation	Expression tissues	Substrates	Subcellular localisation	Mutant phenotype under Pi starvation	Reference
<i>GDPD</i>							
<i>GPD1 (GDPD6)</i>	At5g08030	Induced	Young buds, pollen, embryo	NA	NA	No changes in lipid composition	Gaude et al., 2008; Cheng et al., 2011
<i>GDPD1</i>	At3g02040	Induced	Pollen, embryo, dry seed	GPC, GPE, GPG	Plastids	Low GDPD activity, G3P level, free Pi, seedling growth	Cheng et al., 2011
<i>LAH</i>							
<i>PLAIIA (PLP2)</i>	At2g26560	Induced	Rosette leaf	PC, PE, PG	NA	NA	Sanda et al., 2001; Yang et al., 2012
<i>MGD</i>							
<i>MGD1</i>	At4g31780	Not induced	Rosette leaf	DAG and UDP-Gal	Plastid inner envelope	NA	Jarvis et al., 2000; Awai et al., 2001; Kobayashi et al., 2007
<i>MGD2</i>	At5g20410	Induced	Late pollen	DAG and UDP-Gal	Plastid outer envelope	Slight defect in root growth	Kobayashi et al., 2004, 2006, 2007; Narise et al., 2010
<i>MGD3</i>	At2g11810	Induced	Senescent leaf, early pollen	DAG and UDP-Gal	Plastid outer envelope	Reduced DGDG accumulation	Kobayashi et al., 2004, 2006, 2007; Narise et al., 2010

Table 1.1. continued

Gene	AGI code	Pi starvation	Expression tissues	Substrates	Subcellular localisation	Mutant phenotype under Pi starvation	Reference
<i>DGD</i> <i>DGD1</i>	At3g11670	Induced	Cauline leaf, senescent leaf, stigma, guard cell	MGDG and UDP-Gal	Plastid envelope	Reduced DGDG accumulation	Dörmann et al., 1995, 1999; Kelly et al., 2003; Froehlich et al., 2001
<i>SQD</i>							
<i>UGP3</i>	At3g56040	Induced	Leaf	Glc-1-P> Gal-1-P	Chloroplasts	No detectable SQDG, slight PG decrease and MGDG increase	Okazaki et al., 2009
<i>SQD1</i>	At4gg33030	Induced	Leaf	1-P	Chloroplasts	No detectable SQDG	Essigmann et al., 1998; Okazaki et al., 2009; Sanda et al., 2001;
<i>SQD2</i>	At5g01220	Induced	Senescent leaf	UDP-Glc	Chloroplasts	No detectable SQDG, growth defect	Yu et al., 2002

The composition of glycerolipid membrane is characterised by phosphoglycerolipids, mostly phosphatidylglycerol (PG), phosphatidylcholine (PC) and phosphatidylinositol (PI), as well as a high content of non-phosphorus glycerolipids, mainly MGDG, DGDG and SQDG (Boudière et al., 2014) (Fig. 1.5).

Lipid composition changes under Pi deficiency cannot be explained by the molecular studies alone. It would be best to combine with the analysis of cellular lipid composition through lipidomics analysis. Lipidomics is the study of complete lipid composition within a cell, tissue or organism by mass spectrometry (MS) (Tenenboim et al, 2016). The comparison of both molecular mechanism and lipidomics will give a clear picture of the molecular processes establishing and regulating lipid characteristics under Pi deficiency.

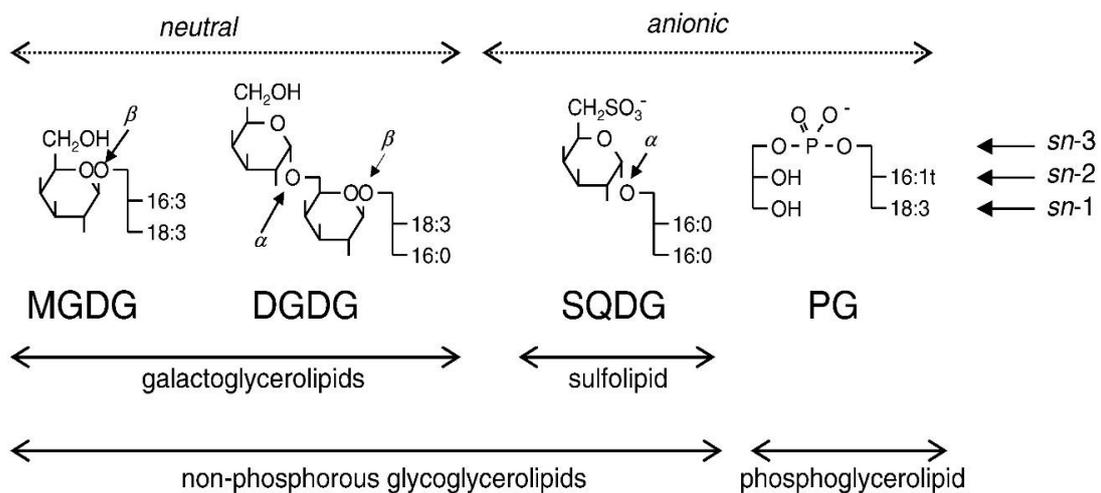


Figure 1.5. Membrane glycerolipid classes and structure in Arabidopsis. The highly conserved (through evolution from cyanobacteria to higher plants) glycerolipid composition with 2-carbon glycerol backbone esterified to fatty acids (denoted as sn 1 and sn-2), and the polar head is at the position sn-3. They are categorised as a very low phosphoglycerolipid content, mainly phosphatidylglycerol (PG), phosphatidylcholine (PC) and phosphatidylinositol (PI), and very high non-phosphorus glycerolipid contents, which are galactoglycerolipids (MGDG, DGDG) and sulfolipid (SQDG). Illustration extracted from (Boudière et al., 2014).

A null-mutant in *Arabidopsis* (*dgd1*) lacking the DGDG synthase (DGDG1) showed a big reduction (90%) in the amount of DGDG under normal growth conditions, however accumulated DGDG increased up to 60% of the amount present in the wild type under Pi deficiency (Härtel et al., 2000).

Lipidomic profiling together with transcriptomic analysis could reveal how Pi deficiency induces the eukaryotic metabolic pathway of membrane lipid remodelling as well as elucidate their molecular pathway in changing their lipid composition by switching from phospholipids to galactolipid to maintain growth under Pi-deprived conditions.

Although a lot is already known about how plants respond to low Pi availability, it is still unclear how the genetic regulatory network underlying the phenotypic variation in plant responses to Pi availability varies. The identification of genes underlying biological traits is a complex process. One of the potential mechanisms to identify genes underlying quantitative characters is through the use of expression quantitative trait loci (eQTL) studies (Druka et al., 2010).

1.8 EXPRESSION QUANTITATIVE TRAIT LOCI (eQTL)

Expression quantitative trait loci (eQTL) are genomic regions that contribute to variation in gene expression among individuals due to sequence polymorphisms in target genes or their regulatory regions (Kliebenstein, 2009). eQTL that map to the approximate physical location of their gene of origin are referred to *cis*-eQTL. In contrast, those that map far from their physical position, often on different chromosomes, are referred to as *trans*-eQTL. Identifying variation in gene expression within a segregating mapping population is potentially of immense use, in particular within the plant and crop sciences (Druka et al., 2010). Identifying eQTL under Pi deficiency will increase our understanding of P use efficiency (PUE) in plants and its regulation. The advanced understanding of the genetics of the PUE at the individual gene level may provide good opportunities for crop improvement based on candidate gene and marker identification at a scale that is much faster than one based on trait quantitative trait loci (QTL) approaches alone (Hammond et al., 2009).

Since a single gene could be associated with one or many eQTL, categorising eQTL into *cis*- or *trans*-eQTL effects relative to the physical location of a gene can reveal regulatory networks controlling gene expression without *a priori* models. Under Pi starvation, Pi acquisition, translocation and remobilisation are mediated by Pi transporter (PHT1 family). mRNA expression levels were quantified in a segregating population. Therefore, the genomic regions which control the expression level could be identified using eQTL analysis. By combining eQTL analysis with trait analysis, putative candidate genes associated with traits can be identified indirectly by correlation analysis (Kliebenstein, 2009; Hammond et al., 2011). Identification of eQTL under Pi-starvation responses has previously been conducted in the BraIRRI mapping population of *B. rapa*, a cross between rapid cycling IMB 211 (female) and yellow sarson R500 (male), the progeny

was self-cross for eight generations to generate recombinant inbred lines (RILs) (Iniguez-Lui, et al., 2009; Hammond et al., 2011).

1.9 BRASSICAS

The genus Brassica is one of the members in the Brassicaceae (Cruciferae) family and includes many economically important crops like vegetables (cabbage, broccoli, turnip and mustard) and oilseeds (oilseed rape) (Snowdon et al, 2007). There are six commercially important species in the Brassica genus, namely *Brassica oleracea* (*B. oleracea*), *Brassica napus* (*B. napus*), *Brassica rapa* (*B. rapa*), *Brassica nigra* (*B. nigra*), *Brassica juncea* (*B. juncea*) and *Brassica carinata* (*B. carinata*). Brassica can be distinguished from other families by the distinct features of their flowers. The flowers have cruciform corolla (cross shaped), six stamens where the outer two stamens are shorter than the remaining inner stamens. All of the six species in the genus Brassica are closely related through evolution, sharing many similar traits, but developing independently has made them have a great diversity of leaf and flower architecture (morphotype) (Tang & Lyons, 2012).

The Brassicas are the close relatives of the model plant *Arabidopsis thaliana*, which has a small genome size; around 140 Mb (Bennett et al., 2003) compared to Brassica genomes which are three to five times larger, 529 Mb for *B. rapa*, 696 Mb for *B. oleracea* and 976 Mb for *B. napus* (Bennett et al., 2003; Lysak et al., 2009; NCBI, 2017).

1.9.1 Triangle of U theory

Brassicas split from their common ancestor with *Arabidopsis* 20-24 Mya (Ziolkowski et al., 2006). The three diploid species, *B. rapa*, *B. nigra* and *B. oleracea* are denoted as genome AA ($2n=20$), genome BB ($2n=16$) and genome CC ($2n=18$), respectively, forming the 'Triangle of U' (Nagaharu, 1935) (Fig. 1.6). This theory describes how the

combinations of these diploid species hybridise with each other creating amphidiploid species *B. juncea*, *B. napus* and *B. carinata*, denoted as AABB (2n=36), AACC (2n=38) and CCBB (2n=34), respectively. All the Brassica species share another common characteristic in that they all experienced a whole genome triplication (WGT) event subsequent to their divergence circa 9-28 million years ago (Cheng et al., 2014).

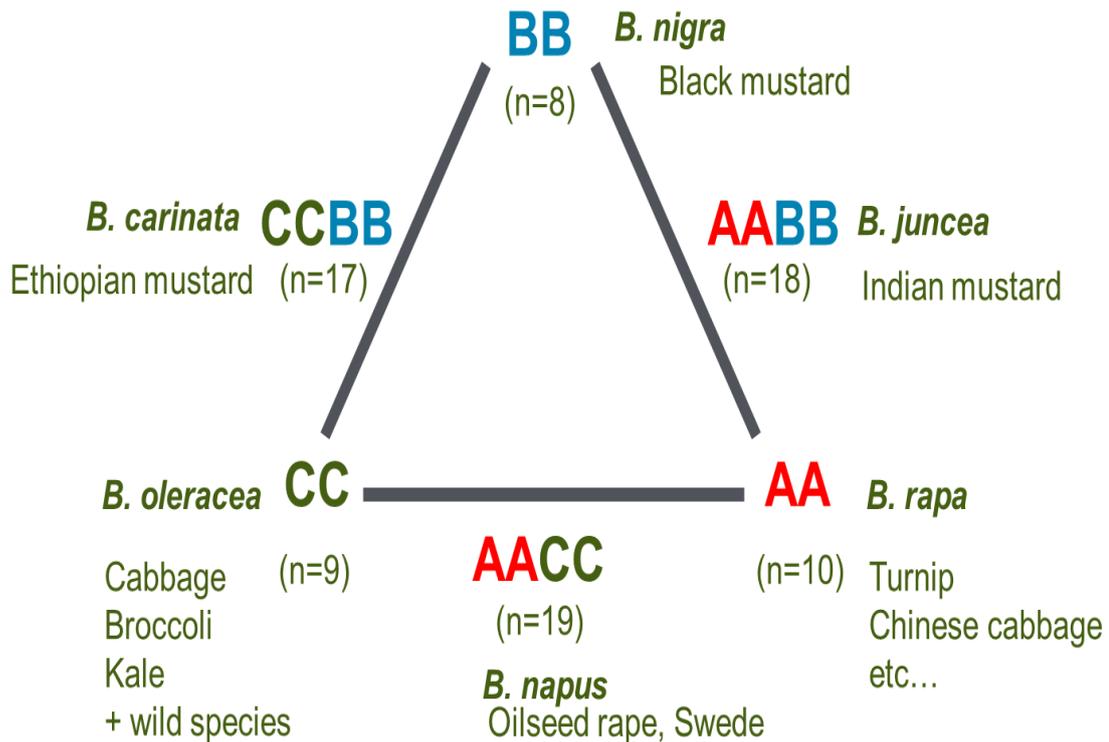


Figure 1.6. The Triangle of U theory. Extracted from Nagaharu U (1935). The figure describes the genomic relationships between six species in genus *Brassica* (n refers to the haploid number of chromosomes). Three diploid species, *Brassica rapa* (2n=20, AA), *Brassica nigra* (2n=16, BB) and *Brassica oleracea* (2n=18, CC) have interspecific hybridised to produce the three new allotetraploid species *Brassica juncea* (2n=36, AABB), *Brassica napus* (2n= 38, AACC) and *Brassica carinata* (2n=34, CCBB).

1.9.2 Whole genome triplication event (WGT)

Many plant species, including all flowering plant species have experienced whole genome duplication (WGD) or polyploidization. WGD is an event which creates extra copies of the entire genome in an organism. WGD occurred frequently during evolution and can be caused by hybridisation of genomes of different species (also called allopolyploidy) or by multiplication of the same genome (also called autopolyploidy). This event helped plants to adapt to the environment changes during evolution and improved their tolerance to mutations as well as provided a lot of genetic diversity for the evolution of new characteristics (Kellis et al., 2004). Brassicas experienced an additional WGT event after they shared a common ancestor with Arabidopsis (Cheng et al., 2015) (Fig. 1.7).

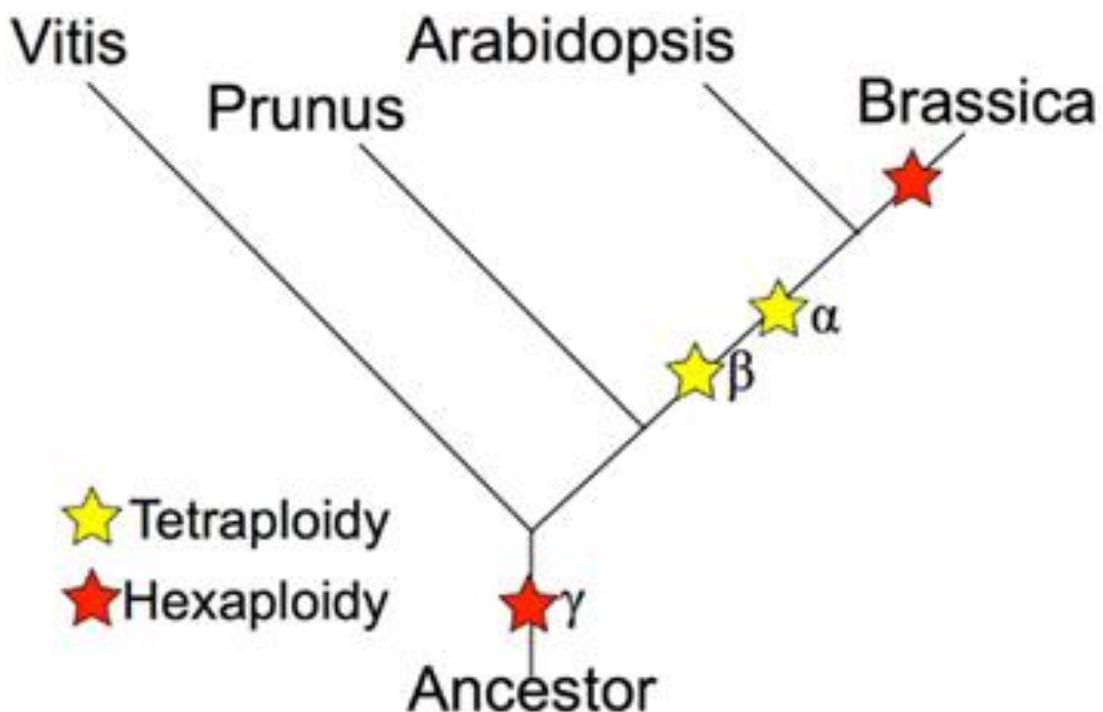


Figure 1.7. Phylogeny of Brassica relatives marking the relative placement of lineage divergence and polyploidy events. Polyploidies are named according to the Arabidopsis convention of α most recent in the lineage of Arabidopsis), β (second most recent) and γ (eudicot paleohexaploidy). Extracted from Tang and Lyons, 2012.

The whole genome triplication event that occurred subsequent to the species divergence created three extra copies of their ancestral genome through fractionation (Fang et al., 2012).

Comparative genomic analysis by mapping genetic markers was conducted between Arabidopsis genome and *B. rapa* genome (Fig. 1.8). Previous studies found that genomes in *Brassica* spp. were divided into 24 genomic blocks (GBs; labelled from A to X). Since *B. rapa* experienced WGT, there are three corresponding extra copies of orthologous genomic regions of Arabidopsis (Parkin et al., 2005; Cheng et al., 2015).

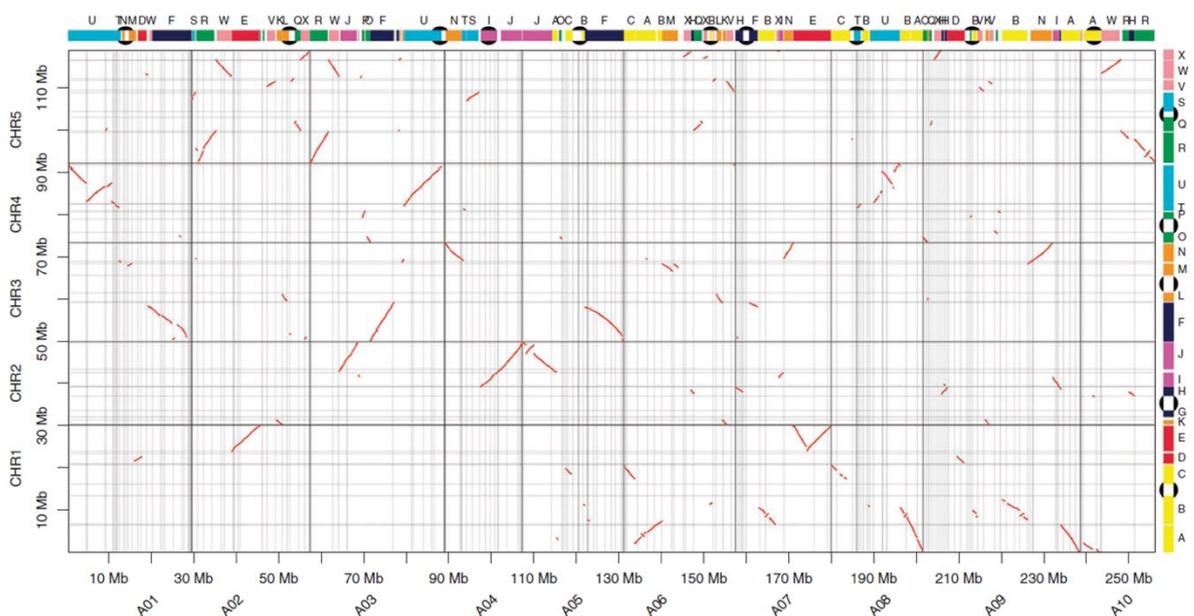


Figure 1.8. Whole genome triplication event in *Brassica rapa* is evident by comparing the ten chromosomes of *Brassica rapa* genome (horizontal axis) and the five chromosomes of *Arabidopsis thaliana* genome (vertical axis). The coloured blocks show the triplication copies of Arabidopsis genomic regions in Brassica chromosomes (Wang et al., 2011b).

1.9.3 Brassicas: *B. rapa*

B. rapa (genome AA, $2n=20$) is an important crop with various morphotypes, including leafy vegetables (pak choi, Chinese cabbage), root vegetables (turnip), oilseed and fodder crops (fodder turnip). Morphological variation of *B. rapa* occurred due to long

history of breeding for selected traits as well as natural selection for adaptation in different habitats (Kumar et al., 2015). *B. rapa* is cultivated worldwide from the tropics to the Arctic Circle (Fahey, 2003). Leafy vegetables, such as Chinese cabbage (*B. rapa* subsp. *pekinensis*) are grown widely in East Asia, while the oilseed-Brassica (*B. rapa* subsp. *oleifera*) is grown in China, Canada, India and northern Europe (Kumar et al., 2015).

1.9.4 Brassicas: *B. napus*

Oilseed rape (*B. napus*; genome AACCC, $2n=38$) is the most widely cultivated crop species in the Brassicaceae family (Snowdon et al., 2007). It is the third largest source of vegetable oil in the world after palm oil and soy oil and is the most important oil crop in Europe. About 13% of the world's vegetable oil is provided by oilseed rape (Raymer, 2002). The European Union, China, Canada and India are the largest producers of oilseed rape in the world (USDA, 2015). This species originated through interspecific hybridization of two ancestral species *B. rapa* (genome AA, $2n=20$) and *B. oleracea* (genome CC, $2n=18$), creating an amphidiploid genome (Nagaharu, 1935).

1.10 RESEARCH OBJECTIVES

Data from previous studies have shown characteristic transcriptional events in response to Pi deficiencies across the plant species. These include changes in the expression of genes involved in remobilising, decreasing, or replacing Pi in non-essential cellular compounds, the exudation of metabolites and enzymes into the rhizosphere to increase soil Pi availability, changes in root morphology and/or associations with microorganisms and increased capacity to acquire Pi more effectively from the soil (Fig. 1.2).

Gene expression is a quantitative trait that can be mapped genetically in structured populations to identify eQTL. Analysing the gene expression will help to elucidate the

molecular mechanism underlying complex traits. Genes and regulatory networks underlying complex traits can subsequently be inferred. The specific genes of interest might be useful for breeding plants with high tolerance to low Pi availability and subsequently improve the crop growth in a large scale with low dependence to Pi fertilisers.

The four main objectives of this study are:

1. To investigate the regulation and remobilisation of Pi within the crop during development and the implications for seed P use efficiency (PUE).
2. Quantify diversity in the expression of Phosphorus (P) responsive genes in *Brassica* spp. under low Pi availability.
3. To identify the major regulators underlying *trans*-eQTL hotspots in *B. rapa*.
4. Identify and quantify lipid metabolism genes in selected *B. napus* species under low Pi availability.

CHAPTER 2 GENERAL MATERIALS AND METHODS

2.1 PLANT MATERIALS

2.1.1 *B. rapa*

B. rapa R-o-18, an inbred line of the *B. rapa* subsp. *trilocularis* (Yellow Sarson) with transparent seed coat (Rusholme et al., 2007), was selected for the study. Seeds used in this experiment were provided by Prof. Martin Broadley, University of Nottingham and were used for both Pi response experiments (pot and hydroponics). Seeds were sown directly into pots containing a peat compost (Attgrow Horticulture Limited, Claygate, UK) and sand mixture (75%/25%) for the pot experiment. In the hydroponics experiment, seeds were pre-germinated in petri dishes with filter paper soaked in deionised water and sealed. Seeds were incubated in the dark at 4 °C overnight and left at room temperature under natural light intensity until they germinated. Seedlings were transferred to seed trays filled with sand for a week to let the roots elongate. Individual seedlings were placed in foam collars (Brickfill movement joint filler for bricks and blocks, Fillcrete, Kent, UK) and inserted into the lid of the hydroponic tank (size; 40 cm x 60 cm, HDPE, Allibert) containing 30 L of hydroponic solution (see section 2.3) and aerated with a pump and air-stone (Charles Austen Pumps LTD, UK; model F65DE).

2.1.2 *B. napus*

Twenty-four cultivars/lines of *B. napus* or oil seed rape were used in the hydroponics experiment; BA-040 (Apex), BA-044 (Laser), BA-048 (NK Bravour), BA-067 (Nugget), BA-089 (AMBER X COMMANCHE DH LINE), BA-099 (TAPIDOR DH), BA-101 (EUROL), BA-102 (Lesira), BA-113 (Samourai), BA-154 (Liporta), BA-158 (Major), BA-177 (Victor), BA-185 (CANARD), BA-202 (Chuosenhu), BA-204 (COUVE NABICA), BA-208 (Q100), BA-209 (RAGGED JACK), BA-210 (RED RUSSIAN), BA-216 (E94197), BA-218 (GROENE GRONINGER SNIJMOES), BA-224 (WILD ACCESSION), BA-410 (HUGUENOT), BA-418 (Altasweet), and BA-426 (Drummonds Purple Top). The line

selections were from the ERANET-ASSYST consortium diversity population (Thomas et al., 2016). Seeds of *B. napus* lines were obtained from Prof. Ian Bancroft, University of York, UK. Seeds were pre-germinated and transplanted using the method described in 2.1.1.

2.2 GROWTH CONDITIONS

Plants were grown in a glasshouse or a controlled environment growth room at the University of Reading (51°26'13.4"N 0°56'30.3"W). Glasshouse temperatures were maintained between 20-25 °C during the day and 12-15 °C during the night with supplementary lighting (SON-T bulbs) activated for 16 hours/day. Humidity was not controlled. Plants grown in the controlled environment growth room were exposed to a 16 h day at 21 °C and an 8 h night at 18 °C. Relative humidity was kept constant at 60% and the light intensity was 420 $\mu\text{mol m}^{-2} \text{sec}^{-1}$ (Skye Instrument, UK).

2.3 GROWTH MEDIA AND NUTRIENT SOLUTIONS

2.3.1 Pot experiments

In the pot experiments, plants were grown in a 1 L plastic pot containing a compost (Attgrow) (75%), and horticultural sand (25%) mix (Fig. 2.1). The following nutrients were added to the mix: 0.4 g L⁻¹ ammonium nitrate (Yara UK Limited, Grimsby, UK), 0.75 g L⁻¹ potassium nitrate (Yara), 2.25 g L⁻¹ ground limestone (Attgrow), 2.25 g L⁻¹ ground magnesian limestone (Attgrow), and 0.4 g L⁻¹ fritted trace elements (wm255; Librel ® BMX, Bradford, UK). To alter the Pi concentration in the compost mix, single superphosphate (J. Arthur Bowers, Lincoln, UK) was added to aliquots of the above mix at 0.0, 0.075, 0.15, 0.225, 0.45 and 1.35 g L⁻¹ of compost mix. Based on preliminary experiments, the plants were watered with deionised water three times a week to ensure no other nutrients could affect the growth of the plants.

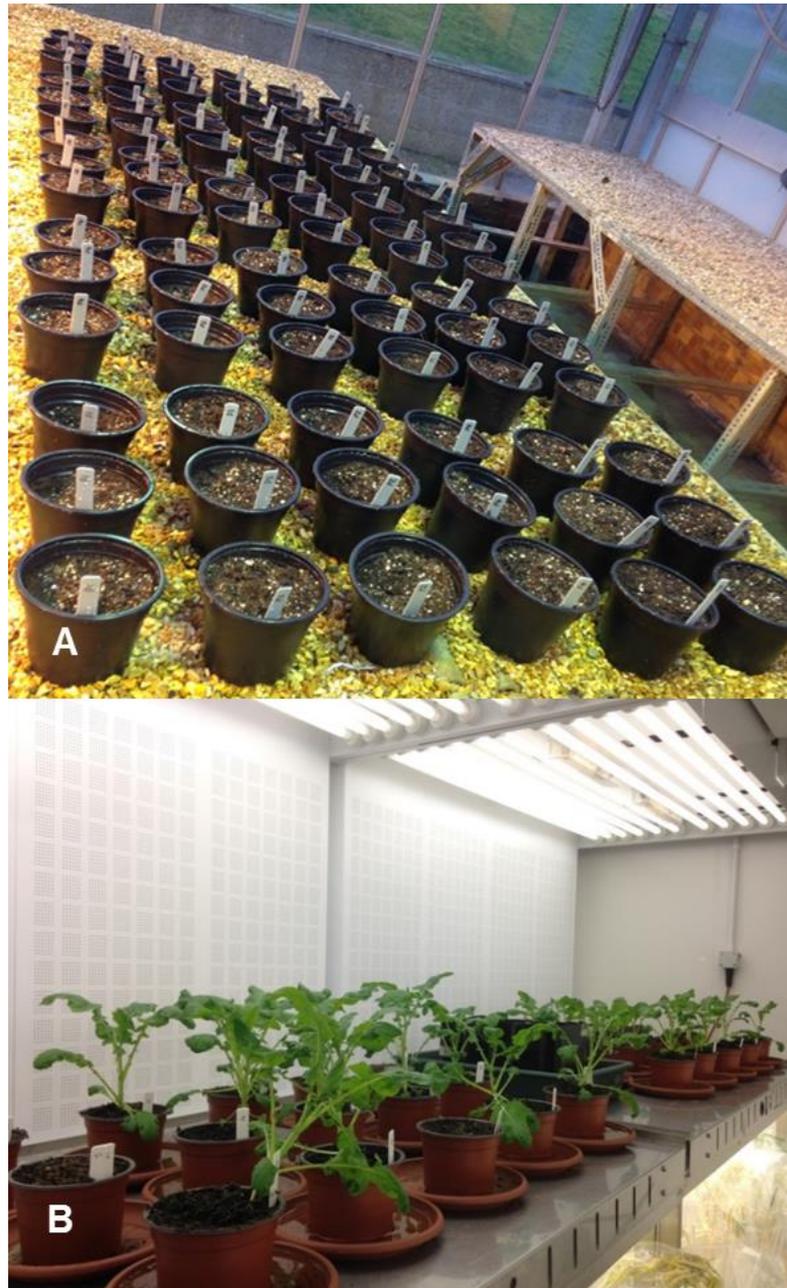


Figure 2.1. **A.** *Brassica rapa* plants grown in 1 L plastic pots containing compost and sand media (75%/25%) in the glasshouse. **B.** *Brassica napus* plants grown in 1 L plastic pots containing compost and sand media (75%/25%) in the controlled environment room.

2.3.2 Hydroponics experiments

Hydroponics experiment was conducted in the glasshouse at University of Reading to look at Pi-response to Pi-deficient conditions (Fig. 2.2). Seven hydroponic stock solutions were prepared (Table 2.1) and these were added to deionised water to make the final hydroponic solution for growing plants. All chemicals were supplied by Sigma-Aldrich, Dorset, UK. The pH of the solution was adjusted to 5.6 with 6 M NaOH and 3 M HCl and the solution was renewed completely every week. Once plants were established on day 21 at 9.30 a.m, Pi was removed from the nutrient solution supplying half the plants by replacing stock solution 1 with stock solution 7 (Table 2.1), resulting in the Pi-deplete nutrient solution (P-), whilst maintaining the K concentration in the nutrient solution. The other plants continued to receive complete nutrient solution (P+) (Table 2.2). The plants were harvested on day 30 at similar time point (9.30 a.m). Samples were taken depending on the experiment's requirement; whole shoots, representing aboveground tissue, while other plant parts like leaf, pooled leaf, leaf-disc, stem, flower and pod silique and seed were harvested accordingly.

Table 2.1. Stock nutrient solution composition. Seven stock solutions were prepared for the hydroponics experiment.

Stock Solution	Ingredient Main Solution	Molecular Weight	Concentration in Stock Solution		Volume Required for 1 L of Final Solution mL	Concentration in Final Solution mM
			M	g L ⁻¹		
1	KH ₂ PO ₄	136.09	0.2667	36.295	0.375	0.10
	KOH	56.10	0.5333	29.918		0.50
	MgSO ₄ ·7H ₂ O	246.47	0.3750	92.426		0.75
2	CaCl ₂ ·2H ₂ O	147.02	0.0125	1.841	2.000	0.03
	FeNaEDTA	367.05	0.0500	18.353		0.10
3	Ca (NO ₃) ₂ ·4H ₂ O	236.15	1.0000	236.150	2.000	2.00
5	NH ₄ NO ₃	80.00	1.0000	80.000	2.000	2.00
	Micronutrients		mM	g L ⁻¹		μM
	H ₃ BO ₃	61.83	30.0	1.855		30.00
	MnSO ₄ ·4H ₂ O	223.06	10.0	2.231		10.00
6	ZnSO ₄ ·7H ₂ O	287.55	1.0	0.288	1.000	1.00
	CuSO ₄ ·5H ₂ O	249.68	3.0	0.749		3.00
	Na ₂ MoO ₄ ·2H ₂ O	241.95	0.5	0.121		0.50
	Zero P		mM	g L⁻¹		μM
7	K ₂ SO ₄	174.25	0.133	23.228	0.938	0.13
	KOH	56.10	0.533	29.918		0.50

Table 2.2. The protocol of inducing the Pi deficiency for plants grown hydroponically. The stock solutions were changed more frequently after day 21 to ensure there was no residual Pi in the tanks.

Day	Stock Solutions Change
0	All are in complete solutions (solution 1-6).
7	All are in complete solutions (solution 1-6).
14	All are in complete solutions (solution 1-6).
21	For half of the plants KH_2PO_4 (P+) was replaced with K_2SO_4 (P-) (solution 2-7). The solution for the other half of the plants was changed but maintained in P+ condition (solution 1-6).
22	Wash K_2SO_4 (P-) tank.
25	Change KH_2PO_4 (P+) tank and K_2SO_4 (P-) tank.
28	Change KH_2PO_4 (P+) tank and K_2SO_4 (P-) tank.
30	Harvest.



Figure 2.2. Image of the hydroponics tanks used for the hydroponic culture of *Brassica napus* at **A.** week 1, **B.** week 2 and **C.** week 3 after sowing.

2.4 PLANT P ANALYSES

2.4.1 Pi analysis

Two drying methods were used for analysing free Pi concentration, oven-dried and freeze-dried. In the oven-dried method, plants were cut just above the compost to obtain whole shoots. Shoots of the plant were collected and weighed to obtain fresh weight before being dried at 60 °C and reweighed for the dry weight. In the freeze-dried method, shoots or specific plant parts were cut and directly flash frozen in liquid nitrogen. The samples were kept in a -80 °C freezer before being dried in a freeze dryer (Applied Vacuum Engineers, Bristol, UK). The samples were ground using laboratory mill (Perten Instruments Warwick, UK) to ensure representative plant samples were used for Pi analysis. The Pi was extracted using 1% acetic acid and homogenised 3 times with a TissueLyser (Qiagen, Manchester, UK) at a frequency of 30 Hz for 60 secs (Fig. 2.3). The quantification of Pi concentration in tissues was measured by releasing the cellular contents into water by repeated freeze-thaw cycles followed by quantifying Pi by the molybdate-blue assay using Thermo Scientific Multiscan (Thermo Scientific, UK) according to the procedure by Ames (1966) with some modifications. A total of 300 µL of reagent was added; 10 µL of sample, 80 µL of deionised water and 210 µL of colour substrate, incubated at 37 °C for 30 min. The Pi concentration was measured at 620 nm using Ascent Software Version 2.6.

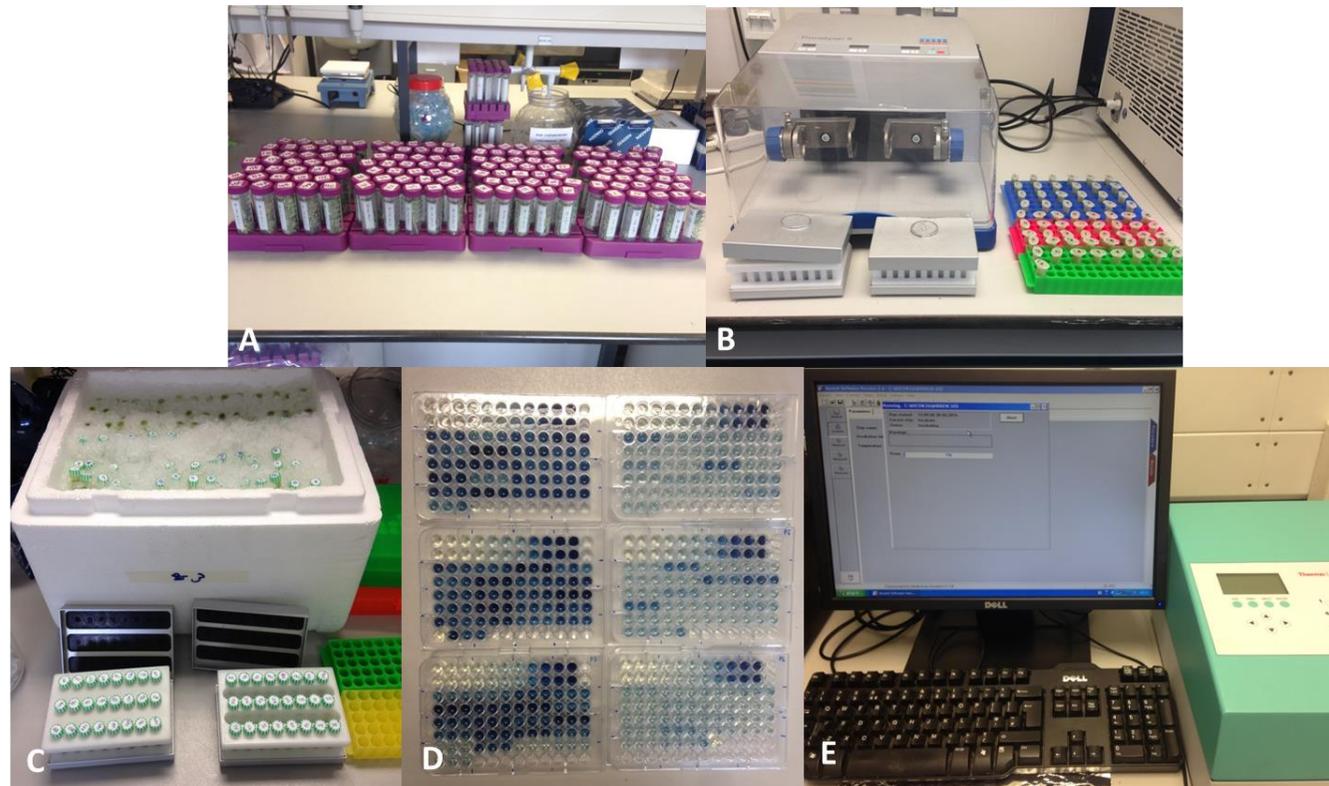


Figure 2.3. Pi and total P analysis using molybdate-blue assay according to protocol by Ames (1966). **A.** Plant freeze-dried sample **B.** The free P was extracted using TissueLyser (Qiagen) **C.** The samples were kept on ice to avoid phosphatase activity during extraction **D.** Colorimetric analysis on plates **E.** The Thermo Scientific Multiscan (Thermo Scientific) used to measure the P concentration at A_{620} wavelength.

2.4.2 Total P analysis

Plant tissues were sampled, freeze-dried and ground as described in 2.4.1. Tissue samples were digested in MARSXpress digestion tubes (CEM Corporation, Buckingham, UK) using trace element grade concentrated nitric acid (Fisher Scientific) and ultra-pure water ($\text{HNO}_3\text{-H}_2\text{O}$). Digestions were accomplished by MARS6 microwave digestion system (CEM Corporation) at 140 °C for 45 min (Fig 2.4). Total P concentration was quantified using a molybdate-blue assay according to the procedure by Ames (1966) with modifications described in 2.4.1.

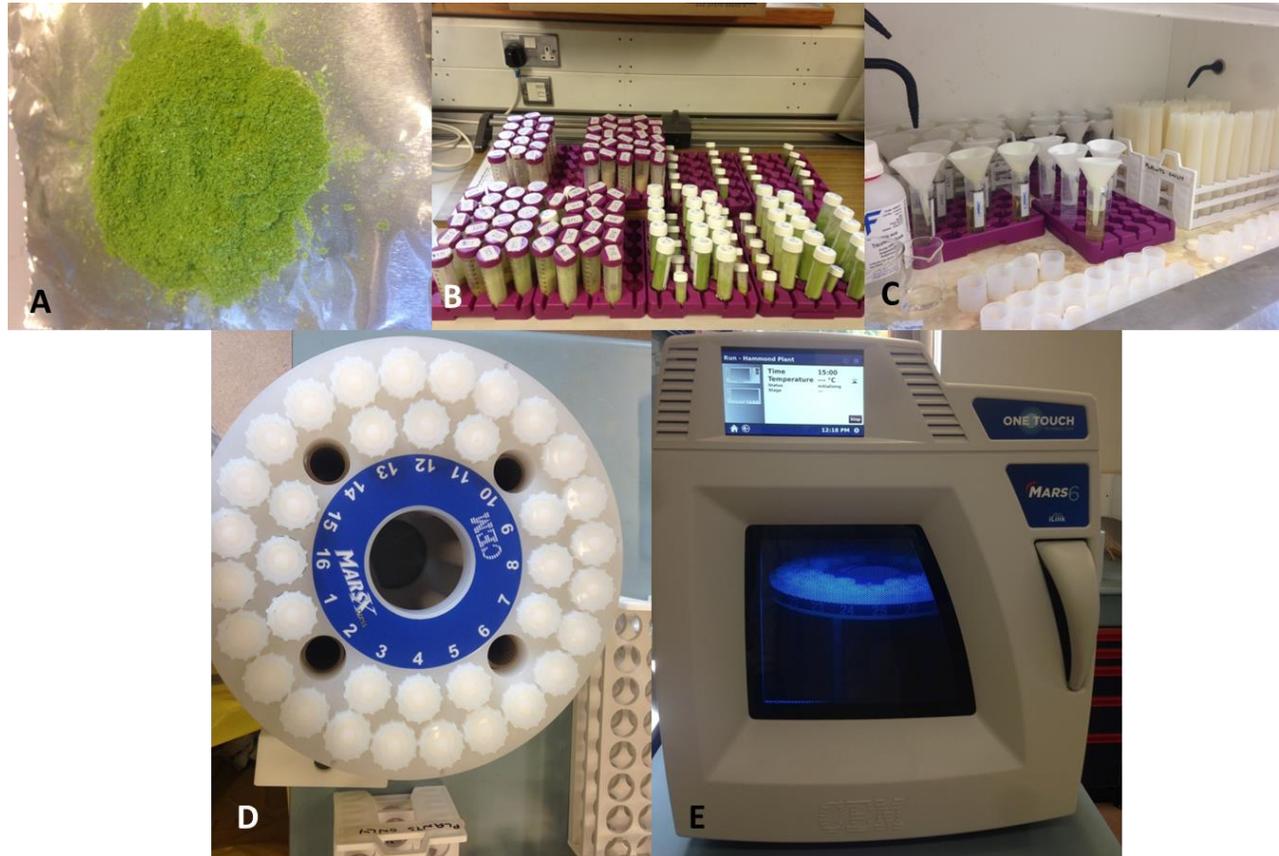


Figure 2.4. Plant-acid digestion for total P analysis. **A.** Ground plant samples **B.** The samples were kept in air tight tubes **C.** Acid digestion procedures were conducted in the fume-hood. **D.** The recommended arrangement of MARSXpress vessels on the turntable **E.** MARS 6 microwave digestion system used to digest the samples.

2.5 DATA ANALYSIS

Statistical data analyses were performed using Genstat (17th Edition, VSN International, Oxford, UK). First, the distribution and variance of the data were plotted to determine whether the data were distributed normally, and variance was evenly distributed across the data set. Data that had even variances and normal distribution were subjected to further analysis. Data that were not normally distributed and/or had unevenly distributed variance were subjected to \log_{10} transformation.

The data were analysed using a Residual Maximum Likelihood (REML) procedure (Patterson and Thompson, 1971, Welham and Thompson, 1997). REML analysis was used instead of ANOVA because it is more suitable for unbalanced data due to experimental conditions and constraints such as the failure of some seeds to germinate. Data were first analysed with all the factors (replicate, treatment, block, plot etc.) as random terms to obtain the variance components of each experiment. Then, the analysis was repeated with the component of interest (e.g *B. napus* line or P treatment) as a fixed factor to get the predicted means and standard errors. Further least significant difference (LSD) analysis was performed and the data were plotted using Microsoft Excel 2010.

2.6 DNA ISOLATION

Tissue samples were taken from leaves or shoots (aboveground sample) of *Brassica* sp. for DNA isolation. Approximately 8-10 g fresh weight of tissue was taken, wrapped in aluminium foil and flash frozen in liquid nitrogen. The samples were stored at -80 °C until analysed. Tissues were ground to a fine powder in liquid nitrogen using a mortar and pestle. DNA was isolated from tissues using the DNeasy® Plant Mini kit according to manufacturer's instructions (Qiagen, Manchester, UK.). The DNA was eluted by the addition of 20 μ L of nuclease free water twice to make the total sample volume of 40 μ L.

2.6.1 DNA quantification and integrity

The concentration of DNA and purity of the sample was assessed using a Nanodrop 2000 spectrophotometer (Thermo Scientific, Wilmington, USA) by measuring the nucleic acid quantity in $\text{ng } \mu\text{L}^{-1}$ along with a reading of $A_{260/280}$ and $A_{260/230}$ ratio. The high purity of DNA was commonly in the range of 1.8-2.2 for $A_{260/280}$ ratio and only samples with highly purified DNA were selected for further analysis. The $A_{260/230}$ reading is a secondary measure of purity indicating contamination by protein or phenol. A Qubit 3.0 fluorometer (Life Technologies, Paisley, UK) was also used to quantify the DNA by using a Broad Range ds-DNA kit (ThermoFisher Scientific) according to manufacturer's guidelines. The integrity of DNA was determined through electrophoresis on 1.2% agarose gel in TAE buffer at 110 V for approximately 90 min. The gel was visualised in a UV Transilluminator and the intensity of the DNA bands was compared to the reading from the Nanodrop and Qubit to confirm the quantification of DNA present in the sample.

2.7 RNA ISOLATION

Shoot or leaf tissues (depending on the experiment) were ground to a fine powder in liquid nitrogen using a mortar and pestle. Total RNA was isolated using 50 mg of ground tissues using the Spectrum™ Plant Total RNA kit according to manufacturer's instructions (Sigma-Aldrich, Dorset, UK). Genomic DNA was removed from the RNA using On-Column DNase 1 Digestion set (Sigma-Aldrich). The RNA was eluted by the addition 20 μL of nuclease free water twice to make the total sample volume of 40 μL .

2.7.1 RNA quantification and integrity

The concentration and purity of RNA was determined using a Nanodrop 2000 spectrophotometer (Thermo Scientific). 1 μL of RNA (isolated in 2.7) was loaded after being blanked with nuclease-free water. This gave the reading of nucleic acid concentration in ng

μL^{-1} and the ratio of absorbance at 260/280 nm and 260/230 nm. The ratio ranging from 1.8-2.0 was considered as highly purified RNA, the 260/230 nm reading is a secondary measure of purity where the ratio is often higher than the 260/230 nm values. A Qubit 3.0 fluorometer (Life Technologies) was used to quantify the amount of RNA by using a high sensitivity RNA kit (ThermoFisher Scientific) according to manufacturer's guidelines. The integrity of the RNA was determined by loading the RNA samples together with RNA loading dye (5 μL) onto 1% denaturing formaldehyde gel in 3-(N-morpholino) propanesulfonic acid (MOPS) buffer. Loading dye was prepared by adding 12.5 mL of 62.5% deionised formamide, 1.85 mL of 1.14 M formaldehyde, 2.5 mL of 1.25 X MOPS-EDTA buffer, 4 mg of bromophenol blue ($200 \mu\text{g mL}^{-1}$), 1 mg of $50 \mu\text{g mL}^{-1}$ of ethidium bromide and 3.15 mL of RNase-free water to make it up to 20 mL. When the samples were ready for electrophoresis, loading buffer was added to the required amounts of RNA in the ratio of 1:5 into microtubes. The samples were heated to 65 °C to denature the RNA, cooled down and the samples was loaded into wells of a gel. Electrophoresis was run at 120 V for 30 min (Fig 2.5). The gel was visualised in a UV Transilluminator and the intensity of the RNA bands was compared to the reading from the Nanodrop and Qubit to confirm the quantification of RNA present in the sample and its integrity.

2.8 PRIMER DESIGN

Primers were designed using primer analysis software Primer3Plus (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>) by considering the following parameters: (a) product size range: 80-250 bp; (b) primer size: 18-22 bp; (c) max self-complementary: 2; (d) GC clamp: 1.

**CHAPTER 3 REGULATION AND REMOBILISATION OF
PHOSPHATE (PI) WITHIN *BRASSICAS* DURING
DEVELOPMENT AND THE IMPLICATIONS FOR SEED
PHOSPHORUS USE EFFICIENCY (PUE)**

3.1 AIMS AND OBJECTIVES

The aims and objectives in this chapter are;

1. To investigate Pi response in Brassica under various soil Pi treatments and to determine Pi concentration to represent high Pi and low Pi treatment for Brassica growth conditions.
2. To assess the nutrient status of *B. napus* grown under Pi-sufficient and Pi-deficient conditions through its life cycle.
3. To determine the transcriptional changes in leaf sample of *B. rapa* under low Pi availability.

3.2 BACKGROUND

Low phosphate (Pi) availability in soil is a limiting factor to a growing plant due to slow diffusion and high fixation in soil. In this study, experiments were undertaken to look at the variation in Pi uptake by *Brassica* spp. under different soil Pi availabilities. First, *B. rapa* plants were grown in compost with a range of Pi fertiliser treatments to investigate the response to Pi availability. Second, two soil Pi concentrations were selected to represent the low Pi and high Pi treatment. *B. napus* plants were grown in compost with these two treatments and harvested at different parts of the plant and at six different growth stages. Free Pi concentration and total P content were determined to assess the nutrient status of the plant. Third, *B. rapa* plants were grown hydroponically under sufficient (P+) and deficient (P-) conditions to look at the variation in the expression of selected Pi responsive genes. The expression of these genes was determined using quantitative real-time PCR (qPCR). This study was conducted to provide a general overview of P effects on plant growth processes and analyse some specific experimental data on *B. rapa* and *B. napus* responses to Pi deficiency.

One strategy to reduce the amount of Pi fertilisers used in agriculture is breeding crops that acquire and use P more efficiently. Breeding crops that could produce the high yields with lower inputs of Pi fertilisers or have reduced P demand for physiological activity and tissue Pi concentration is important to ensure P use efficiency (PUE). Several approaches for PUE measurements have been proposed (Table 3.1; Hammond et al., 2009).

Table 3.1. Calculations for different measurements of Phosphorus use efficiency (PUE) .

Name	Abbreviation	Calculation	Units
Agronomic P use Efficiency	APE	$(Y_{\text{high}} - Y_{\text{low}}) / \Delta P_{\text{app}}$	g DM g ⁻¹ P _f
P uptake efficiency	PUpE	$[(P_{\text{high}} \times Y_{\text{high}}) - (P_{\text{low}} \times Y_{\text{low}})] / \Delta P_{\text{app}}$	g P g ⁻¹ P _f
P utilisation efficiency	PUtE	$(Y_{\text{high}} - Y_{\text{low}}) / [(P_{\text{high}} \times Y_{\text{high}}) - (P_{\text{low}} \times Y_{\text{low}})]$	g DM g ⁻¹ P
Physiological P use efficiency	PPUE	$Y_{\text{high}} / P_{\text{high}}$ or $Y_{\text{low}} / P_{\text{low}}$	g DM g ⁻¹ P

Y_{high} = yield on a high P soil

Y_{low} = yield on a low P soil

ΔP_{app} = difference in amount of P applied as fertiliser between high and low P treatments

DM= dry matter, P_f= fertiliser P.

P utilisation efficiency (PUtE) of plants under Pi deficiency enhances internal PUE in response to internal changes in metabolic pathways, Pi remobilisation to younger tissues and membrane lipid remodelling by replacing phospholipids with non-phosphorus galactolipids and sulfolipids. P uptake efficiency (PUpE) requires enhanced Pi uptake from the soil to the plant. In contrast to PUtE, PUpE is more desirable in low-input systems to increase the uptake of Pi from the soil. In contrast, enhanced PUtE will be more desirable in high-input systems as Pi uptake from the soil will be replenished regularly (Heuer et al., 2016). Plants can increase PUtE during Pi deficiency by Pi remobilisation in tissues, recycling P through hydrolysis of Pi from extracellular Pi-esters and exploring Pi through

advanced root system (Plaxton & Tran, 2011). Manipulation of genes and phenotypic variations involved in these processes might be useful in developing better crops for low-input agricultural systems.

3.3 MATERIALS AND METHODS

3.3.1 *B. rapa* physiological responses to Pi availability (six treatments) in compost mixture

Seeds of *B. rapa* (R-o-18) were sown in 1 L pots containing a peat (Attgrow) and sand mix (75:25). The nutrients were added according to section 2.3. After the seeds germinated, they were thinned to reduce them to one per pot. The experiment was conducted in a glasshouse at the University of Reading and the growth condition was set up as per section 2.2. Each of the six treatments, P1 (0 g L⁻¹), P2 (0.075 g L⁻¹), P3 (0.15 g L⁻¹), P4 (0.225 g L⁻¹), P5 (0.45 g L⁻¹), and P6 (1.35 g L⁻¹) had five replicates. A split plot design was used as the experimental design and the plants were harvested at three different time points (14, 28, and 42 days after sowing). The total number of pots was 90. Plants were harvested by cutting just above the level of the compost (shoot). Samples were oven dried and Pi concentrations were quantified as described in section 2.4.1. Data were analysed using Genstat (section 2.5).

3.3.2 *B. napus* physiological responses to Pi availability (two treatments) in different growth stages in compost mixture

Seeds of *B. napus* (Canard) were sown in 1 L pots containing peat and sand mix (75:25). The nutrients were added according to section 2.3. Initially several seeds were sown per pot to ensure enough material for harvesting at the early growth stages; harvest 1 (four seedlings) and harvest 2 (two seedlings). For all other harvests, only one seed was allowed to grow per pot (harvest 3-6). The experiment was conducted in a controlled environment

growth room at the University of Reading and the growth conditions were set as per section 2.2. Each of the two treatments, which represent the low and high external Pi concentrations (P1= 0 g L⁻¹ and P4= 0.225 g L⁻¹) had five replicates. A split plot design was used as the experimental design and the plants were harvested at six different growth stages; i) two-leaves, ii) four-leaves, iii) flowering, iv) first flower opens, v) seed, and vi) maturity (Fig. 3.1). Samples were freeze dried as described in section 2.4.1. Samples were analysed for Pi and total P concentrations as described in section 2.4.1 and 2.4.2. Data were analysed using Genstat (Section 2.5). For total P analysis, 0.1 g of ground dry sample was pre-digested with 2 mL of ultra-pure water and 8 mL of trace element grade concentrated nitric acid (HNO₃). Then the samples were heated in the microwave at 200 °C for 20-25 min and then were held at 200 °C for 10 min. Next, the samples were filtered through filter paper into a 50 mL centrifuge tube. The samples were weighed and diluted prior to further analysis (Moly-blue assay; section 2.4.1).



Figure 3.1. *Brassica napus* plants at different stages of growth. **A.** Two-leaf stage **B.** Four-leaf stage **C.** and **D.** First-flower opens stage **E.** and **F.** Middle-flowering stage **G.** Seed stage **H.** Mature plants.

3.3.3 *B. rapa* transcriptional responses to Pi availability in hydroponics

3.3.3.1 Hydroponics growth conditions and Pi analysis

B. rapa (R-o-18) seeds were first pre-germinated in petri dishes with filter paper soaked in deionised water and sealed with parafilm. The seeds were incubated in the dark at 4 °C overnight and left at room temperature under natural light intensity until the plants germinated. Seedlings were transferred to sand to let the roots elongate. After one week, the seedlings were transferred to a hydroponic nutrient solution (Section 2.3.2 and Table 2.1). The experiment was designed using a split plot design in two tanks, with five replicates per treatment. The Pi deficiency treatment was induced to half of the tanks according to Table 2.2. Pooled leaves from several plants were harvested after 30 days, blotted dry and wrapped in aluminium foil before being flash frozen in the liquid nitrogen (N₂). The samples were kept in the -80 °C freezer before being freeze dried and analysed for Pi (section 2.4.1). Data were analysed using Genstat (section 2.5).

3.3.3.2 RNA extraction and cDNA synthesis

Total RNA was extracted from the *B. rapa* pooled leaf samples according to section 2.7 followed by RNA quantification and integrity check according to section 2.7.1. The integrity of the RNA was examined using gel electrophoresis.

SensiFAST cDNA synthesis kit (Bioline, London, UK) was used to produce cDNA from RNA according to the manufacturer's instructions with oligo d(T) primers and 2 µg of total RNA per sample. For quantitative real-time PCR (qPCR) analysis, cDNA samples were diluted 5x with nuclease-free water prior to qPCR analysis.

3.3.3.3 Candidate reference gene selection, target genes and qPCR primer design

Nine candidate reference genes and 16 phosphate responsive genes were identified from the literature and the protein sequences of their corresponding homologs in Arabidopsis were derived from the NCBI (www.ncbi.nlm.nih.gov/) (Ligaba et al., 2006, Tran et al., 2010, Meyer et al., 2010). *B. rapa* protein sequences were obtained through sequence similarity to the Arabidopsis sequence using the Basic Local Alignment Search Tool (BLAST) using BLASTp searching against the *B. rapa* genome database (BRAD database <http://brassicadb.org/brad>). All sequences were loaded into Geneious software (<http://www.geneious.com/>; Biomatters, Auckland, New Zealand) to facilitate further analysis. The size of the coding sequences was determined by BLASTn searches against the *B. rapa* genome database. Primers for the nine reference genes were designed according to section 2.8. The selected target reference genes were *TATA box*, *SAND*, *PP2A*, *Actin*, *TIP41*, *Cyclin*, *Tubulin*, *GAPCP1*, and *18s rRNA* (Table 3.2). The 16 target genes were selected from *B. rapa* R-o-18 accessions based on their putative roles in plant responses to Pi availability. Primers for the 16 target genes had already been designed by Dr Goodall in the laboratory (Table 3.3). The selected genes were *Purple acid phosphatase (PAP)*, *PAP12*, *Citrate synthase (CS4)*, *Aluminium activated malate transporter1 (ALMT1)*, *Mitochondrial malate dehydrogenase (MMDH1)*, *MATE efflux protein (MATE)*, and *Sucrose transporter 2 (SUC2)*.

Accurate analysis of qPCR data requires that all primer pairs used should have the same amplification efficiency (Ma et al., 2013). The rate of amplification efficiency at 100% or close to it indicates that all the reaction conditions are optimal, and the results obtained should be highly repeatable. Experimentally, the amplification efficiency of each primer pair between 90-110% is acceptable to be used for the next step of experiment. From the results (Table 3.3), all the target genes showed primer efficiency between 90- 110%, indicating a good quality and highly specific primers. For the reference genes, *SAND* and

TATA box showed primer efficiency of 103.90% and 100.20%, respectively (Table 3.4). Regression coefficients (R^2) were larger than 0.98, indicating that both SAND and TATA box primer pairs are the most highly specific, efficient and can be used as reference genes in this experiment.

Table 3.2. List of potential reference genes for qPCR analysis in *Brassica rapa*.

Target genes	Sequence (5' to 3')	Size amplicon (bp)
<i>Actin</i> _F	GGTTCCAGTTCAGACCATTG	144
<i>Actin</i> _R	TTCTCCTGCGTTTCCCATAC	
<i>GAPCP1</i> _F	AACGGGAAAAAGGTCAACG	95
<i>GAPCP1</i> _R	ACCCCTGAAGACTCAACGAC	
<i>18srRNA</i> _F	TTTCAGCCTTGCGACCATAC	87
<i>18srRNA</i> _R	GCGGATGTTGCTTTTAGGAC	
<i>Cyclin</i> _F	TCGTCTCGGTTGCTTCATC	144
<i>Cyclin</i> _R	TCGTCTCGGTTGCTTCATC	
<i>TIP41</i> _F	ATTCTCACTTCCCTCGCTCAC	128
<i>TIP41</i> _R	GCATTCTCGCCAAACACC	
<i>PP2A</i> _F	CAATCCCTCATCCCCATAGTC	104
<i>PP2A</i> _R	GCAAATAACGAACATCAACATC	
<i>Tubulin</i> _F	AGCAATACCAAGACGCAACC	135
<i>Tubulin</i> _R	GTCACCAACACTTACCAAACC	
<i>SAND</i> _F	GCCCATTTATTCCAGATACGG	126
<i>SAND</i> _R	TGGTGTTCCTGCCTTGAC	
<i>TATA box</i> _F	ATACTCTCACGCCGCTTTCTC	82
<i>TATA box</i> _R	CACAATCTTAGGCACTTTCATCC	

Table 3.3. Selected Pi responsive genes from *Brassica rapa* R-o-18, primer sequence (5' to 3'), amplicon size and primer efficiency.

<i>B. rapa</i> R-o-18 ID	Arabidopsis homolog ID	Gene name	Primer sequence (5'-3') F/R	Amplicon length (bp)	Primer efficiency E (%)
Bro18 039999	At5g34850	<i>Purple acid phosphatase (PAP)</i>	CTAAACACCAAAGGCAGAAATG CGAAAACAGAGGCGGAAG	174	98.60
Bro18 047862	At2g27190	<i>PAP12</i>	AAGGTTGATGTGGTGTTTGC CCGTCCCCGATTGTTATG	143	98.70
Bro18 009479	At2g44350	<i>Citrate synthase 4 (CS4)</i>	AATCTTCGCTCGGCAGTTC CGACTTTGATGGCACTTTGG	198	91.50
Bro18 049018	At2g44350	<i>Citrate synthase 4 (CS4)</i>	ATGAAAACAGACGAGCAG CGACTTTGATGGCACTTTGG	143	91.00
Bro18 014187	At1g08430	<i>Aluminium activated malate transporter 1 (ALMT1)</i>	ACACACTCTCCCACTTCCTTC TCTTTCTCCTCTTCTCCACCTC	93	96.10
Bro18 042204	At1g08430	<i>Aluminium activated malate transporter 1 (ALMT1)</i>	ACACACTCTCGCACTTCCTTC TCCTTCTCCTCTTCTCAACCAC	93	96.50
Bro18 006954	At1g08430	<i>Aluminium activated malate transporter 1 (ALMT1)</i>	TGGGGCTATCTTGGTATTGG TGGTGTGCTGGTCTATTTG	151	95.90
Bro18 038513	At1g 53240	<i>Mitochondrial malate dehydrogenase (MMDH1)</i>	CATCCATCCATCCATCCATC AGACGAGGGGGTTGAGTTTC	192	95.30
Bro18 035465	At1g 53240	<i>Mitochondrial malate dehydrogenase (MMDH1)</i>	TCAAGATGGAGGCACAGAAG TGAGGCAAAGAAAGGAAGC	181	102.60
Bro18 013770	At1g 53240	<i>Mitochondrial malate dehydrogenase (MMDH1)</i>	AGAACCCTTTGCTCTGCCATC ACATTAGCCTTTCCAGCGTAG	196	100.20
Bro18 038689	At1g51340	<i>MATE efflux protein (MATE)</i>	GGGACTGGTTCTTGGGTTTC AGGCGTTGATTGGTTGTG	146	102.90

Table 3.3. continued

<i>B. rapa</i> R-o-18 ID	Arabidopsis homolog ID	Gene name	Primer sequence (5'-3') F/R	Amplicon length (bp)	Primer efficiency E (%)
Bro18 035630	At1g51340	<i>MATE efflux protein (MATE)</i>	TGGGACTGGTTCTTGGTTTTC AGGCATTGATTGGTTGTGTC	147	107.30
Bro18 004703	At1g22710	<i>Sucrose transporter 2 (SUC2)</i>	CTCCATAACGCTTCTCCTCTTG GCATCCACATCGGTCTTTTTC	161	97.40
Bro18 025411	At1g22710	<i>Sucrose transporter 2 (SUC2)</i>	CTTCCTCTCCATAACGCTTCTC GCATCCACATCGGTCTTTTTC	167	90.90

64

Table 3.4. Efficiency of reference gene primer pairs used for qPCR amplification.

Gene name	Primer sequence (5'-3') F/R	Amplicon length (bp)	Melt curve Peak Tm (°C)	Primer efficiency E (%)	Regression coefficient (R ²)
<i>Actin</i>	GGTTCAGTTCAGACCATTTG TTCTCCTGCGTTTCCCATAC	144	84.09	134.50	0.9653
<i>SAND</i>	GCCCATTTATTCCAGATACGG TGGTGTTCCTGCCTTGAC	126	81.71	103.90	0.9930
<i>TATA box</i>	ATACTCTCACGCCGCTTTCTC CACAATCTTAGGCACTTTCATCC	82	81.71	100.20	0.9734

3.3.3.4 Quantitative real-time PCR (qPCR)

To accurately quantify the abundance of target gene transcripts, qPCR was performed. Reactions were carried out in 96- well optical reaction plates (Applied Biosystems, Warrington, UK) sealed with ultra-clear sealing film (Applied Biosystems) with an Applied Biosystems Real-Time One Step PCR system (Applied Biosystems). The reactions were performed in a 10 μ L total volume containing 1 μ L of 5x diluted cDNA, 0.5 μ L of each primer, and 5 μ L of 2x PowerUp SYBR Green Master Mix (Bioline). The reaction conditions were 50 °C (2 min) and 95 °C (10 min) for one cycle, and 95 °C (15 s) and 60 °C (1 min) for 40 cycles. A dissociation step of 95 °C for 15 s, 60 °C for 15 s and 95 °C for 15 s was used for melting curve analyses so that primer dimers and non-specific products could be detected. All qPCR reactions, including the no template control, were performed in biological and technical triplicates. The specificity of the amplicons was confirmed by the presence of a single peak in the dissociation step. The mean values obtained from the nine values (triplicates of each biological triplicate) were used to calculate the final quantification cycle threshold (Ct).

3.3.3.5 Quantitative real-time PCR (qPCR) data analysis

The threshold cycle (Ct) values were recorded using the qPCR system default settings in which the baseline was automatically corrected, and threshold values were estimated using the noise band mode. Statistical analysis (mean and CV) of the Ct values was performed using Excel 2010 (Microsoft). Primer efficiencies (E) for candidate genes were evaluated using the dilution series method and pooled cDNA samples. Five serially diluted cDNA samples were used as templates for the construction of standard curves for each primer pair using qPCR. Standard curves were generated using linear regression based on the Ct values for the dilution series. The correlation coefficients (R^2) and slope values were obtained from the standard curves, and the PCR primer efficiencies (E) were

calculated according to the following equation: $E = (10^{-1/\text{slope}} - 1)$. Primer efficiencies (E) of the genes are shown in Table 3.4 and 3.5, with only genes with E values between 90% to 110% used for further analysis.

3.4 RESULTS

3.4.1 *B. rapa* plant growth responses to six different compost P concentrations

The shoot biomass of plants significantly increased as the amount of added Pi increased from 0 to 0.225 g L⁻¹ of compost, after which there was no significant change in shoot biomass (Fig. 3.2). For all three harvests, the effect of Pi treatment was significant (P<0.05) (Table 3.5). In the first harvest (day-14), the overall difference in shoot biomass was not as great as with the later harvests (Fig. 3.2A). This might be due to the plants still using their internal resources to support their early stage of growth. On the second harvest (day-28), the shoot biomass of plants significantly increased from P1 to P2 (P<0.05). Treatments P2, P3, P4 and P6 were not significantly different from each other, but P5 gave the highest shoot biomass with 0.249 g (Fig 3.2B). The plants showed different growth response phenotypically, where the P1 plant was the smallest and P6 plant was the biggest compared to plants from other treatments (Fig. 3.3). For the third harvest (day-42), the shoot biomass increased significantly from P1 to P2 (P<0.05) (Fig. 3.2C). There was no significant difference between the remaining treatments in the third harvest. Overall, the relationship between yield (shoot biomass) and added Pi showed a quadratic curve divided into three segments, ascent, peak and descent representing three zones of deficiency, sufficiency and toxicity, respectively.

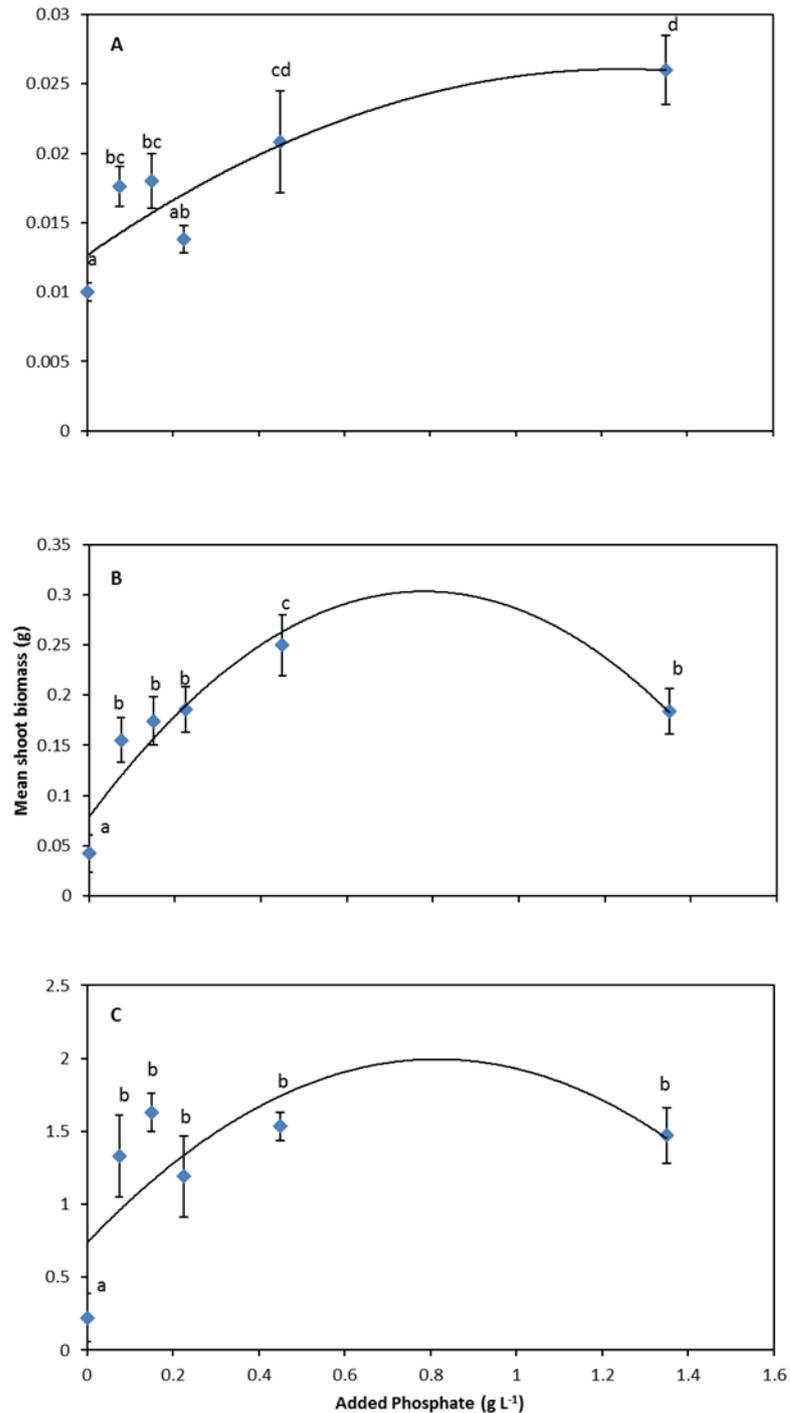


Figure 3.2. Shoot dry weights of *Brassica rapa* plants growing in compost with increasing additions of phosphate harvested after 14 days (A), 28 days (B), and 42 days (C). Plants were grown under glasshouse conditions. Each data point represents mean \pm SEM (n=5). Data were analysed using a REML procedure, followed by Fisher's unprotected least significant difference (LSD) test, so that data points with different letters are significantly different ($P < 0.05$). Equations for fitted lines are (A) $y = -0.0086x^2 + 0.0215x + 0.0127$, $R^2 = 0.7748$ (B), $y = -0.3704x^2 + 0.5771x + 0.0786$, $R^2 = 0.8677$ (C) $y = -1.8931x^2 + 3.0857x + 0.7368$, $R^2 = 0.482$.

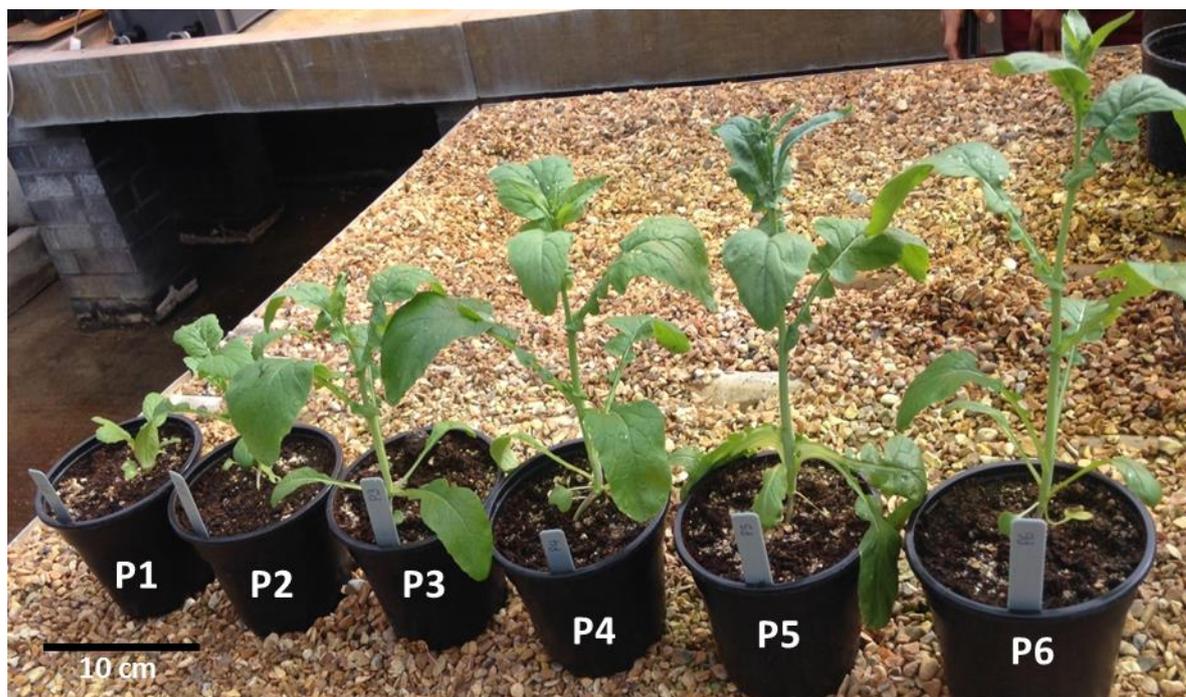


Figure 3.3. Growth response of *Brassica rapa* plants to different additions of phosphate (Pi) to the compost (P1= 0.0, P2= 0.075, P3= 0.15, P4= 0.225, P5=0.45 and P6= 1.35 g L⁻¹ of compost mix) at 28 days after sowing. Scale bar: 10 cm.

Table 3.5. Restricted Maximum Likelihood analysis (REML analysis) of *Brassica rapa* growth at six different Pi treatments.

Harvesting period	Fixed term	Wald statistic	n.d.f	F statistic	d.d.f	Prob. > F
1 st Harvest	Pi treatment	27.45	5	5.49	18.9	0.003
2 nd Harvest	Pi treatment	56.11	5	11.22	19.1	<0.001
3 rd Harvest	Pi treatment	27.99	5	5.6	20	0.002

n.d.f= numerator degrees of freedom

d.d.f= denominator degrees of freedom

3.4.2 Changes in shoot Pi concentration with increasing compost Pi concentration in *B. rapa*

Shoot Pi concentration was significantly affected by the application of increasing amounts of Pi in *B. rapa* (Fig. 3.4). For first harvest 14 days after sowing, data were recorded as a single value because of insufficient amount of sample biomass for measuring Pi concentration. Shoot Pi concentrations were at the highest value on the third harvest for every Pi treatment. The application of Pi treatments generally increased the shoot Pi concentration, but the highest value was observed under treatment P5 (Fig. 3.4). The lowest shoot Pi concentration was observed at P1 for all harvested periods, as expected due to no added Pi (0 g L^{-1}) in the compost mixture.

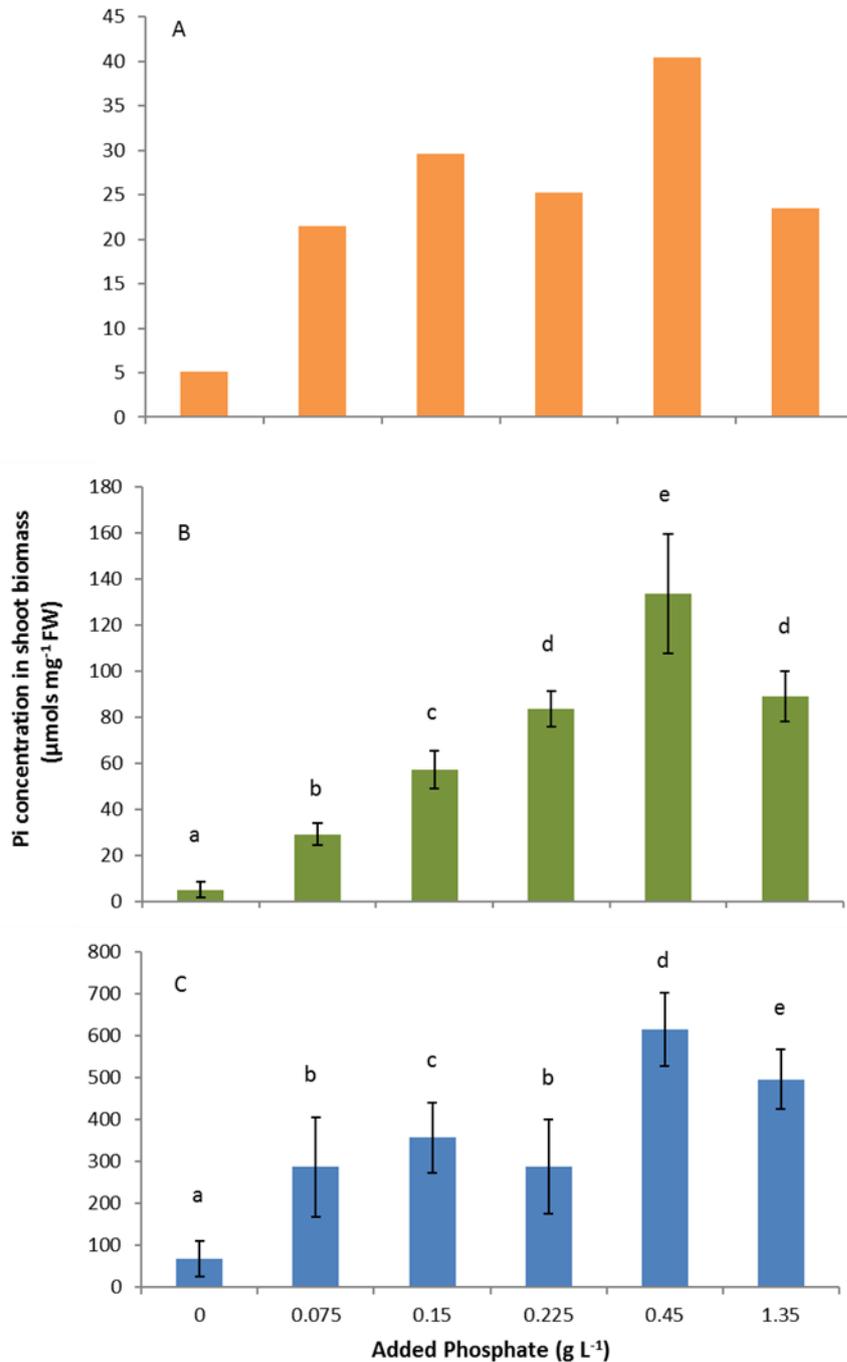


Figure 3.4. Shoot phosphate (Pi) concentration of *Brassica rapa* plants growing in compost with increasing additions of Pi harvested after 14 days (A), 28 days (B), and 42 days (C). Plants were grown under glasshouse conditions. For harvest after 14 days, data are a single value due to low biomass for measuring shoot Pi. Data points for all other data represent mean \pm SEM (n=5). Data were analysed using a REML procedure, followed by Fisher's unprotected least significant difference (LSD) test, so that data points with different letters are significantly different ($P < 0.05$).

3.4.3 Changes in biomass and total P content of *B. napus* tissues grown with two different rates of P during development

The effects of the additional Pi on *B. napus* growth and Pi related traits during development were evaluated using the line Canard grown in compost with no added Pi (P1) and with 0.225 g L⁻¹ added P (P4). Samples were harvested at six different growth stages; i) two-leaves (harvest 1), ii) four-leaves (harvest 2), iii) flowering (harvest 3), iv) first flower opens (harvest 4), v) seed filling stage (harvest 5), and vi) maturity (harvest 6), with the plants broken down into four tissues; i) leaf, ii) stem, iii) flower and iv) pod silique and seed (Fig. 3.5 and 3.6). No plants were harvested at day 7 (harvest 1) and day 9 (harvest 2) due to insufficient amount of sample for measuring Pi concentration.

The yield of filled pod, silique and seed of *B. napus* in the pot study was approximately 38.1 (± 0.5)% of the total plant biomass at maturity while the combination of leaf and stem tissue constituted roughly 60% of plant biomass at maturity, where leaf and stem recorded about 16% and 42% of plant biomass, respectively (Fig. 3.5). The maximum individual tissue biomass accumulation in *B. napus* occurred at maturity stage in both treatments P1 and P4 with 4.561 and 4.952 g, respectively in pod silique and seed tissues of the plant (Fig. 3.5).

Total P contents of plants treated with P1 and P4 both showed similar trends, where pod silique and seed of *B. napus* contained over 50% of plant P at harvest 5 (seed stage). Leaf, stem and flower showed a loss of P during senescence at maturity stage due to redistribution of P to the seed (Fig 3.6).

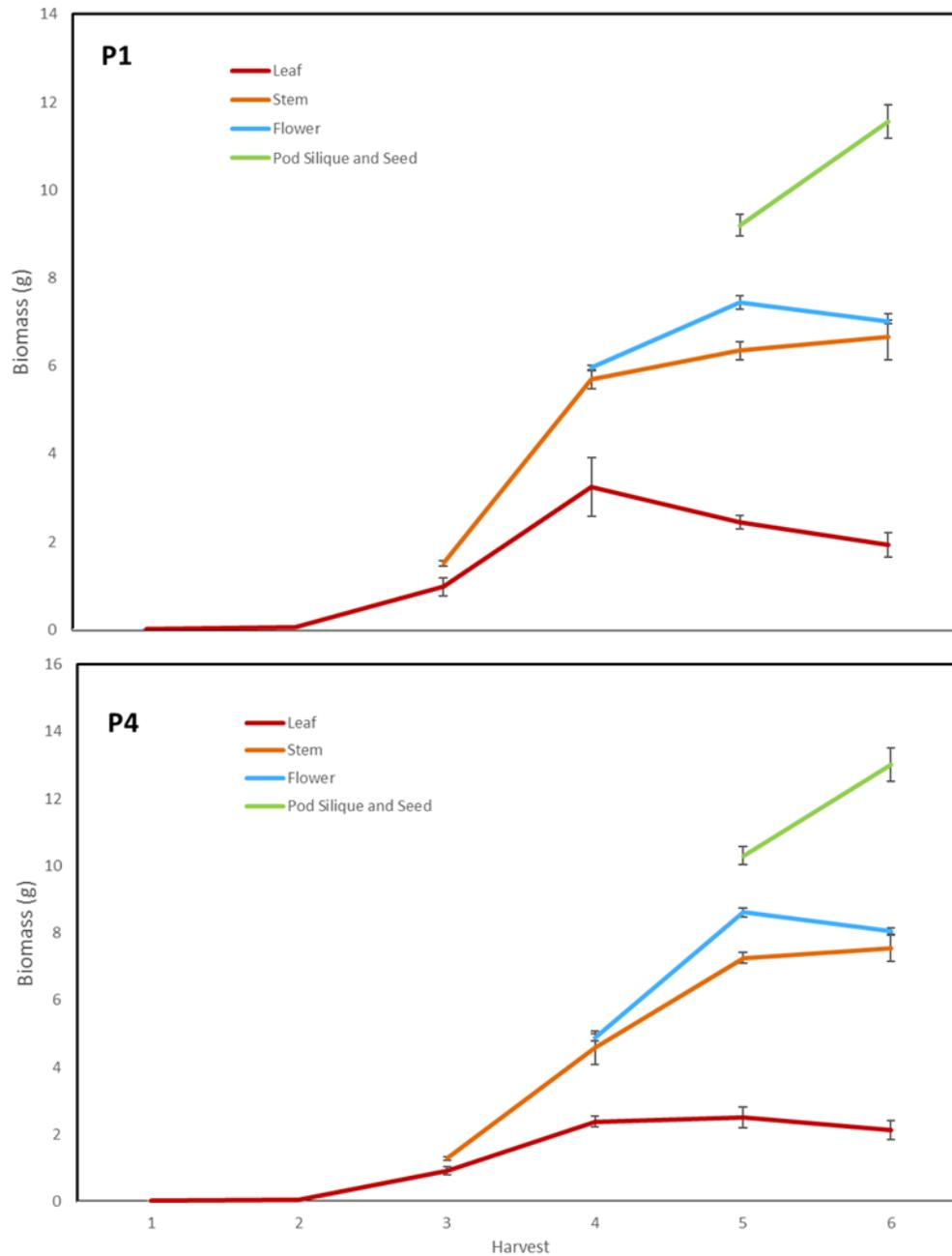


Figure 3.5. Total biomass per plant of *Brassica napus* at six growth stages; two-leaf stage (1), four-leaf stage (2), flowering stage (3), first flower opens stage (4), seed filling stage (5) and maturity stage (6) at four different tissues; leaf (red), stem (orange), flower (blue) and pod silique and seed (green) grown in compost with no added phosphate (P1) and 0.225 g L⁻¹ (P4). Plants were grown under control environment laboratory conditions. At the early stage of harvest (harvest 1 and 2), whole plant was used due to low biomass. Data points for all other data represent mean \pm SEM (n=5).

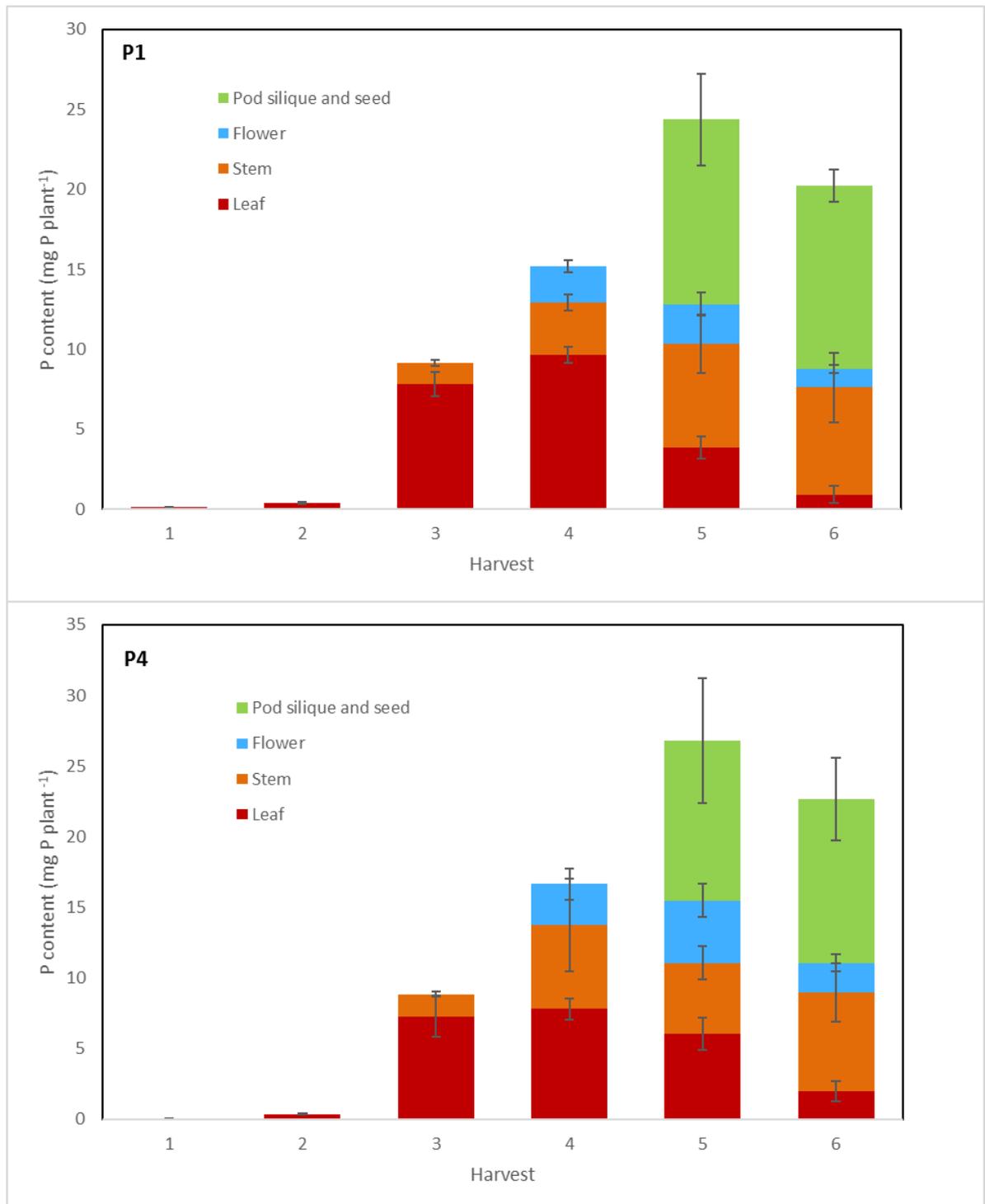


Figure 3.6. Total P content and distribution between tissues throughout the growth cycle of *Brassica napus* plants; two-leaf stage (1), four-leaf stage (2), flowering stage (3), first flower opens stage (4), seed filling stage (5) and maturity stage (6) at four different tissues; leaf (red), stem (orange), flower (blue) and pod silique and seed (green) grown in compost with no added phosphate (P1) and 0.225 g L⁻¹ (P4). Plants were grown under control environment laboratory conditions. At the early stage of harvest (harvest 1 and 2), whole plant was used due to low biomass. Data points for all other data represent mean ± SEM (n=5).

3.4.4 Changes in Pi concentration of *B. napus* tissues grown with two different rates of P availability during development

Overall, Pi concentration of plants growing in P4 were significantly higher than P1 ($P \leq 0.001$). Leaf and stem Pi concentration increased at early stages of growth until harvest 4 for both treatments (P1 and P4). The concentration of Pi decreased in all plant parts until maturity stage (harvest 6). The highest Pi concentration was recorded in P4 stem with $6.085 \mu\text{mol g}^{-1}$ DW at flowering stage (harvest 3). At first flower open stage (harvest 4), tissue Pi concentration declined in the leaf and stem tissues by about 30- 40 % as the plant remobilised P from leaf, stem and flower to the developing seed, moving from source to sink tissues (Fig. 3.7).

Tissue Pi concentration in P1 and P4 treatment showed a similar trend for leaf, stem, flower, and pod silique and seed tissues (Fig.3.7). In flowering stage (harvest 3), stem Pi concentration was highest with $5.476 \text{ mg P g}^{-1}$ DW in P1 treatment, the Pi concentration then began to decline at first flower opens stage (harvest 4), seed filling stage (harvest 5) and maturity stage (harvest 6) with 1.841, 0.966, and $0.475 \text{ mg P g}^{-1}$ DW, respectively. While in P4, stem Pi concentration was higher than P1 with 6.085, 3.429, 1.576, and $0.771 \text{ mg P g}^{-1}$ DW in harvest 3, 4, 5 and 6, respectively. Tissues Pi concentrations declined as plants reach maturity (Fig 3.7).

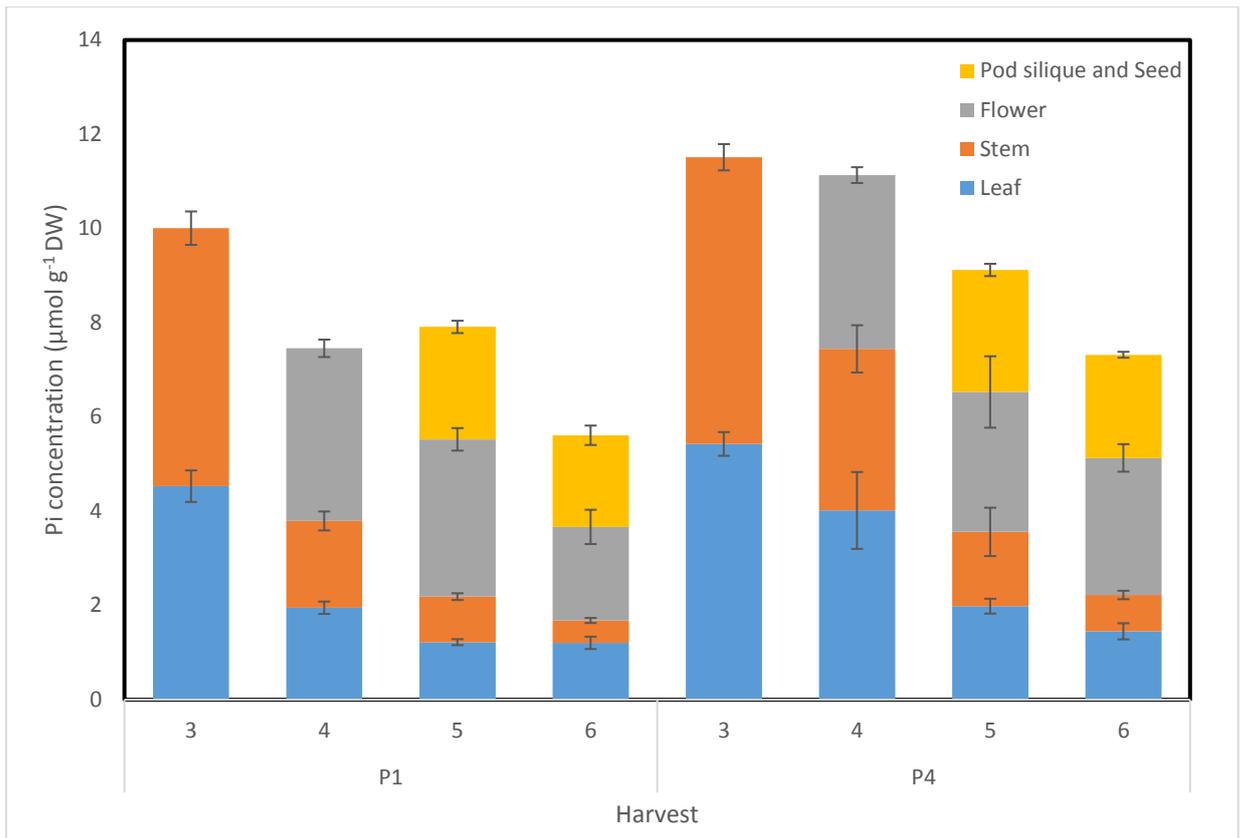


Figure 3.7. Tissue phosphate (Pi) concentration in *Brassica napus* plants; leaf (blue), stem (brown), flower (grey), and pod, silique and seed (orange) grown in compost with no added Pi (P1) and .225 g L⁻¹ (P4). Plants were grown under controlled environment conditions and harvested at two-leaf stage (1), four-leaf stage (2), flowering stage (3), first flower opens stage (4), seed filling stage (5), and maturity stage (6). For harvest 1 and 2, data are not provided due to low biomass for measuring Pi. Data points for all other data represent mean \pm SEM (n=5), except harvest 4 in treatment P4, (n=3).

3.4.5 Transcriptional responses to Pi availability

To investigate the response to Pi deficiency of *B. rapa* at the transcriptional level, qPCR analysis was conducted using pooled leaf samples from plants grown hydroponically under Pi sufficient (P+) and Pi-deficient (P-) conditions. RNA was extracted from the samples and were used to look at the variability of the target genes expression under Pi-starvation.

3.4.5.1 RNA quantification and integrity

Total RNA was extracted from the *B. rapa* pooled leaf samples according to section 2.7 followed by RNA quantification and integrity check according to section 2.7.1. (Table 3.6). The integrity of the RNA was examined using gel electrophoresis (Fig. 3.8). Each sample showed two bands of rRNA (28S and 18S), indicating a high quality of total RNA, which is sufficient for further downstream analysis, with no form of contamination or degradation in the samples. Assessment of the purity of the RNA suggested that it was suitable to proceed with cDNA preparation for six samples (P+, replicates 1, 2, 3 and P-, replicates 2, 3, 4) with $Abs_{260/230}$ and $Abs_{260/280}$ within the range of recommended values ($Abs_{260/280}=1.7-2.0$ and $Abs_{260/230}=2.0-2.2$). The results of $Abs_{260/230}$ for samples P+, replicate 4 and P-, replicate 1 were lower than the recommended values; these two samples were omitted from further analysis. The remaining samples were used for cDNA preparations.

Table 3.6. Quantity and quality data for RNA extracted from leaf of *Brassica rapa* plants grown hydroponically under glasshouse conditions. RNA quantity and quality were determined using a Nanodrop 2000 spectrophotometer.

Treatment	Replicate	RNA concentration (ng μL^{-1})	Abs _{260/230}	Abs _{260/280}
P+	1	827.4	2.04	2.03
P+	2	620.4	2.10	1.80
P+	3	842.3	2.08	2.19
P+	4	440.3	2.07	1.70
P-	1	818.3	2.05	1.45
P-	2	729.9	2.01	1.76
P-	3	1273.8	2.01	1.85
P-	4	1333.3	2.06	2.13

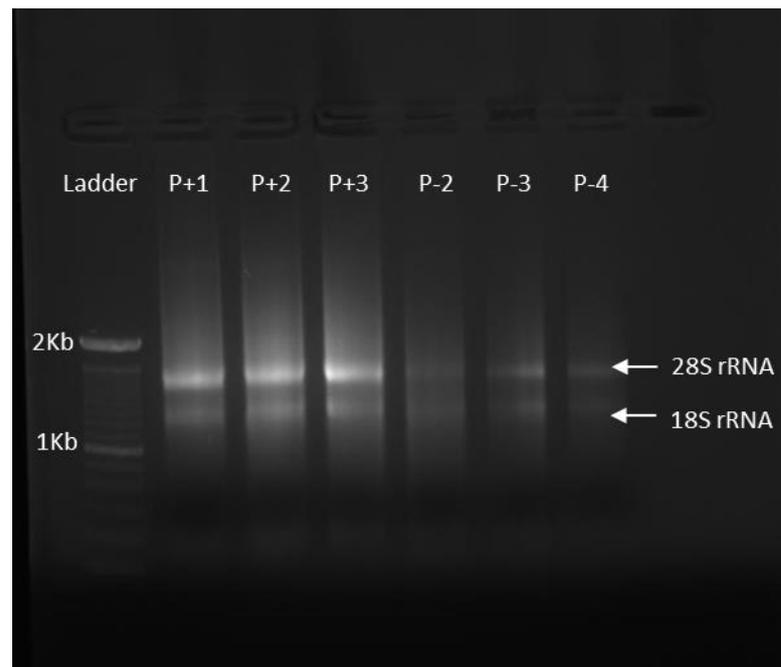


Figure 3.8. Gel electrophoresis of RNA samples from pooled leaf tissues of *Brassica rapa* plants grown under P+ and P- conditions. RNA was loaded into a 1.2% agarose gel and run at 100 V for 60 minutes. RNA was visualised using ethidium bromide (EtBr) under UV light. RNA image of all the samples showed two bands, 28S rRNA and 18S rRNA. Column 1: 1 kb ladder, Column 2-7: Leaf RNA of *Brassica rapa* under P+ and P- conditions with replicate number.

3.4.5.2 Primer specificity of target reference genes and target genes

The selected target reference genes were *TATA box*, *SAND*, *PP2A*, *Actin*, *TIP41*, *Cyclin*, *Tubulin*, *GAPCP1*, and *18s rRNA*. The 14 target genes were *Purple acid phosphatase (PAP)*, *PAP12*, *Citrate synthase (CS4)* (x2), *Aluminium activated malate transporter1 (ALMT1)* (x3), *Mitochondrial malate dehydrogenase (MMDH1)* (x3), *MATE efflux protein (MATE)* (x2), and *Sucrose transporter 2 (SUC2)* (x2). The specificity of these primer pairs was confirmed by PCR amplification and electrophoresis in 2% agarose gel (Fig. 3.9).

Target reference genes were selected based on the literature or genes validated as appropriate for studying biotic stress and non-stress conditions (Nicot et al., 2005, Hu et al., 2009, Garg et al., 2010, Gu et al., 2011). Some of the reference genes were frequently used in quantitative PCR studies like *Actin*, *Tubulin* and *18S rRNA*. The abundance of nine reference genes and 14 target genes in *B. rapa* were evaluated using qPCR in leaf tissues. Analyses of melting curves of the nine reference genes was also used to determine their suitability by the presence of a single peak, indicating no primer dimers or amplification of other products. Primers for *TATA box*, *SAND*, and *Actin* all showed single peaks and were chosen as reference genes for leaf tissues in *B. rapa* (Fig. 3.10).

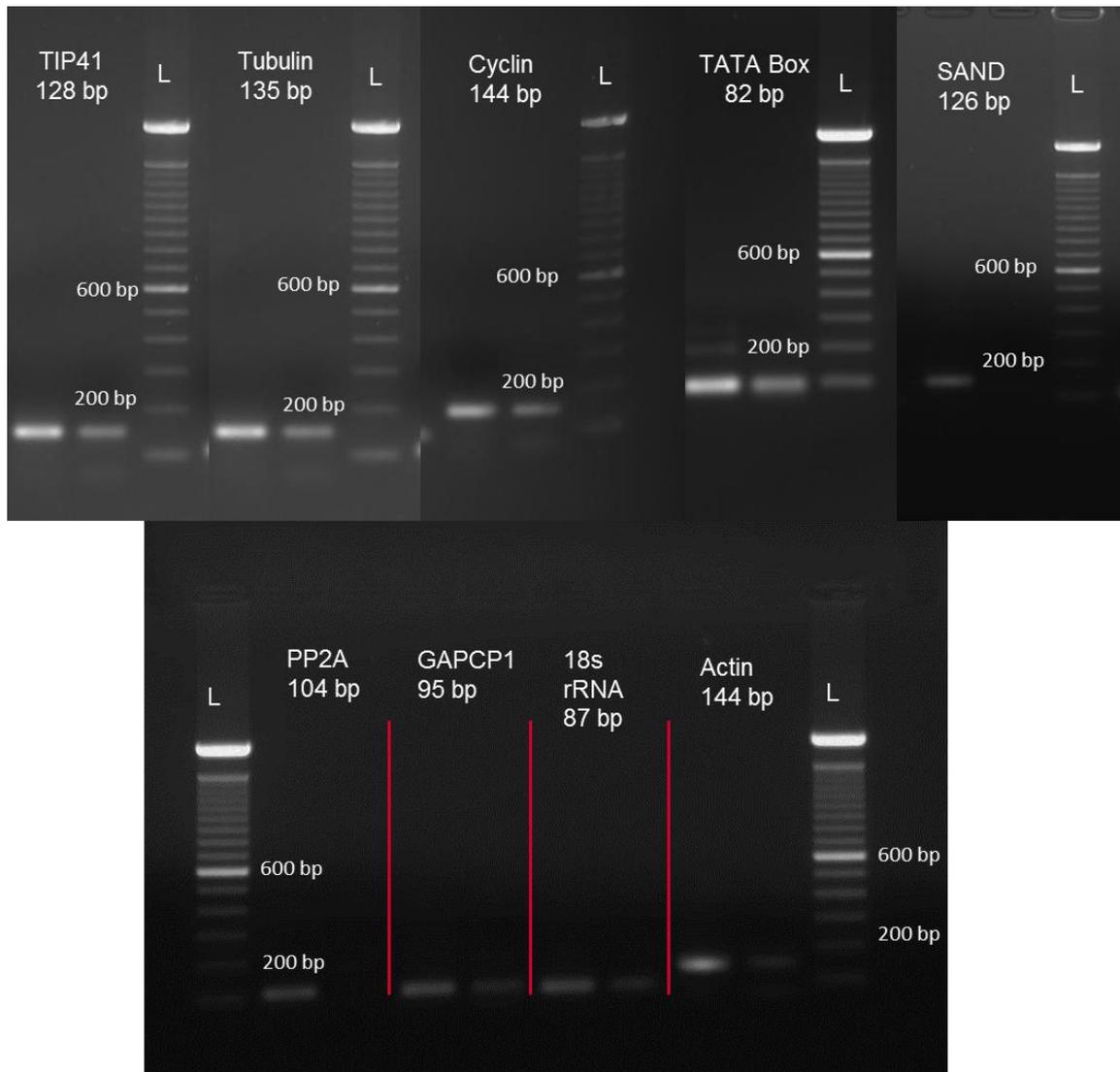


Figure 3.9. Gel electrophoresis of PCR products for nine reference genes from *Brassica rapa* leaf tissues. PCR products were loaded into a 2% agarose gel and run at 100 V for 60 min. PCR products were visualised using ethidium bromide (EtBr) under UV light. PCR products of all samples showed nine reference genes with predicted amplicon size; *TIP41* (128bp), *Tubulin* (135 bp), *Cyclin* (144 bp), *TATA Box* (82 bp), *SAND* (126 bp), *PP2A* (104 bp), *GAPCP1* (95 bp), *18s rRNA* (87 bp) and *Actin* (144 bp). L= 100 bp ladder.

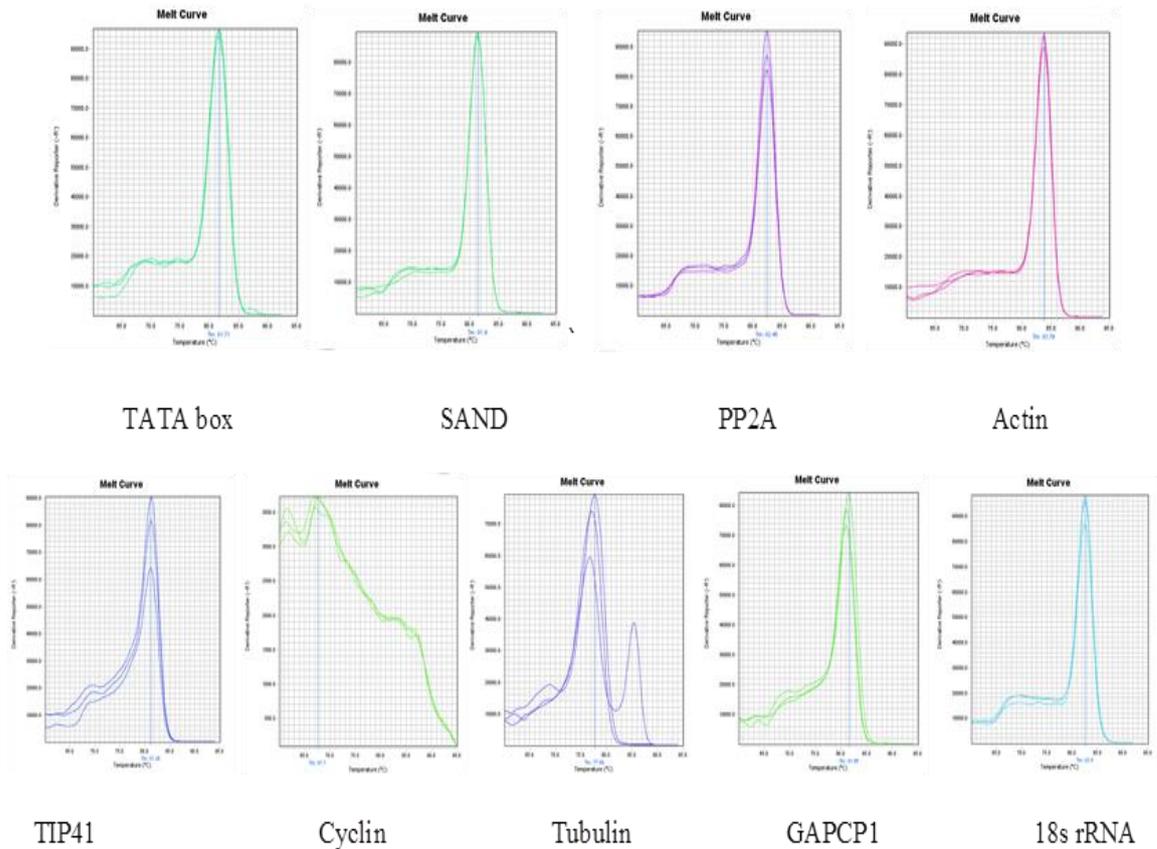


Figure 3.10. Melt curves from qPCR of nine *Brassica rapa* reference genes performed on an Applied Biosystem OneStep+ PCR system to identify single products.

3.4.5.3 Gene expression changes of *B. rapa* under Pi deficiency

Under Pi deficiency, there were differential expression patterns for all 14 target genes compared to the control (P+) (Fig. 3.11). Six of the target genes were up-regulated; Bro18 047862 (*PAP12*), Bro18 014187 (*ALMT1*), Bro18 038689 (*MATE*), Bro18 038513 (*MMDH1*), Bro18 004703 (*SUC2*), and Bro18 025411 (*SUC2*) with 3.69-fold, 7.62-fold, 1.9-fold, 1.4-fold, 3.03-fold and 1.8-fold, respectively. With three of these target genes significantly up-regulated in response to Pi stress. The highest increase in expression in response to Pi stress was Bro18 014187 (*ALMT1*), with a 7.6-fold increase relative to the full nutrient control. Three of the melt genes, Bro18 009479 (*CS4*), Bro18 042204 (*ALMT1*),

and Bro18 035630 (*MATE*), showed down-regulated responses and five of the genes Bro18 039999 (*PAP12*), Bro18 049018 (*CS4*), Bro18 006954 (*ALMT1*), Bro18 035465 (*MMDH1*) and Bro18 013770 (*MMDH1*) remained the same under Pi deficiency.

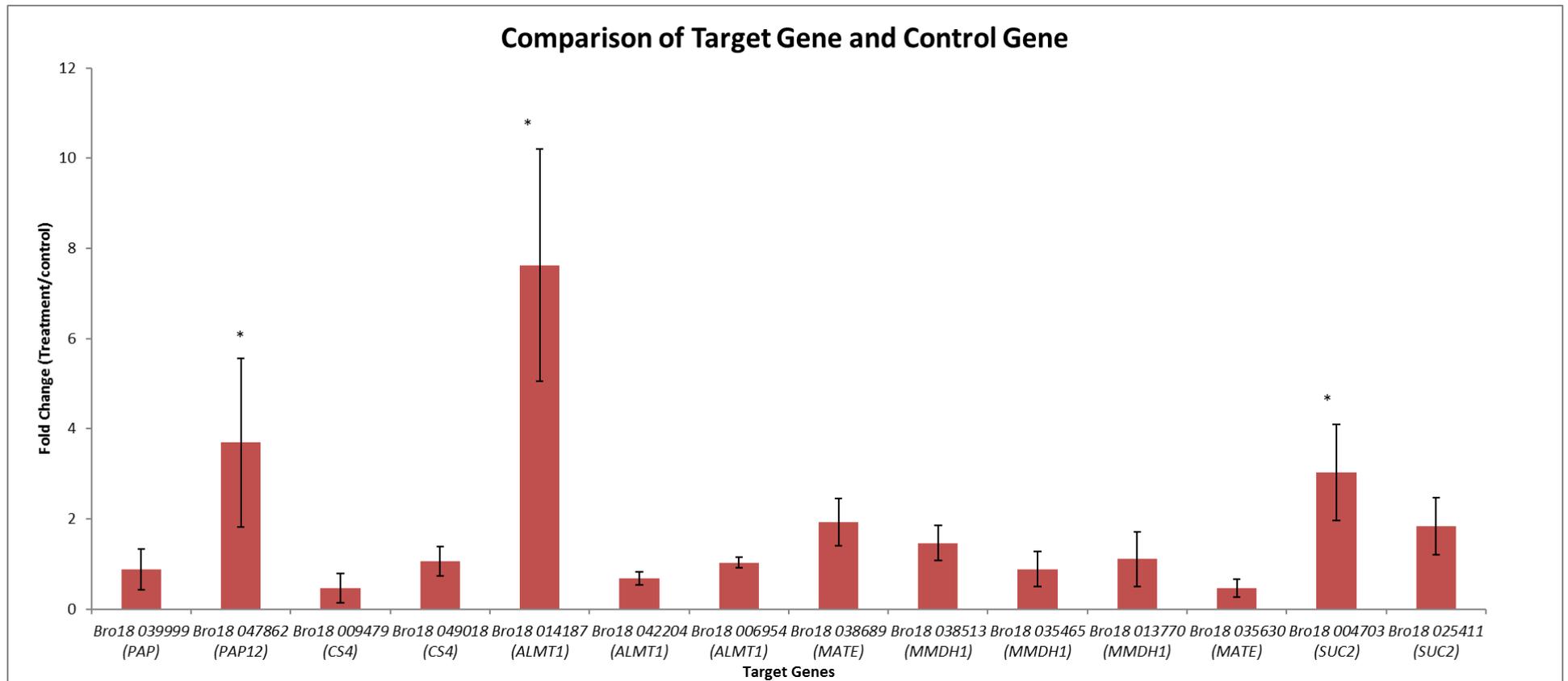


Figure 3.11. Fold changes in gene expression for 14 target genes under Pi deficiency treatment in *Brassica rapa* relative to full nutrient control. Data represent the fold change difference between P deficient and full nutrient control plants. The error bars represent the standard error of the means of three replicates. *Genes whose expression is significantly ($P < 0.05$) different between treatment and control.

3.5 DISCUSSION

3.5.1 Growth response of *B. rapa* to increasing Pi availability

This study has shown that *B. rapa* grow differentially in response to external Pi availability (Fig. 3.2). The effects elevated Pi availability on *B. rapa* growth were determined using the R-o-18 growth on compost mixture subjected to a range of increasing Pi availabilities. Plants were harvested after 14 days, 28 days, and 42 days. Shoot dry weights and Pi concentration in shoot biomass increased with time and Pi availability up to 0.45 g L⁻¹ of added Pi (Figs. 3.2 and 3.4). From these results, significant differences in biomass and shoot Pi concentrations were observed between treatments P1 and P4, providing contrasting growth conditions with which to study responses. Measurements at these two soil Pi concentrations are suitable for assessing physiological measurements and modelling the response without the need to grow the plants at six different P concentrations (Greenwood et al., 2006). This in line with previous studies in Brassicaceae where the P treatments can be determined based on growth response curves under certain conditions, providing a high P treatment represents the plant with a high and sufficient P, and low P treatment, where the P availability is not sufficient and low growth response is observed (Greenwood et al., 2005,2006; Hammond et al., 2009).

3.5.2 Evaluation of *B. napus* response and development under contrasting Pi availabilities

The growth and developmental responses of *B. napus* (Canard) to low and high Pi availability were conducted under controlled environmental conditions. The plants grown in pots were harvested at six different growth stages; i) two-leaves, ii) four-leaves, iii) flowering, iv) first flower opens, v) seed, and vi) maturity. Biomass, total P content and Pi concentration was recorded for individual plant tissues (Figs. 3.5, 3.6 and 3.7).

Analysis of plant internal Pi concentration and total P content to assess plant nutrient status is important for crop optimal production with low P input.

Under Pi deficiency (P1), leaf Pi concentration was reduced by 57% from harvest 3 to harvest 4 (Fig 3.7). This showed that Pi was redistributed from older leaves to young organs. Internal Pi redistribution during Pi deficiency is consistent with previous work demonstrating that Pi is redistributed from source to sink tissues under these conditions (Smith et al., 2003). In contrast, only a 26% reduction was observed in plants grown under the P4 treatment (Fig. 3.7). Pi remobilisation was determined by the amount of the total nutrient present in different organs of the plant at different growth development (Hermans et al., 2006; Mailard et al, 2015). In this study, *B. napus* total P content accounted over 50% in pod, silique and seed at maturity stage consistent with a study in wheat where remobilisation of P accounted for 56-63% of the grain P content (Masoni et al., 2007).

In *B. napus*, total P accumulation continued until flowering stage but started to decline at maturity (Fig. 3.7). This is in contrast with previous work on *B. napus*, where P uptake continued until late maturity (Rose et. al, 2007). Total P content of *B. napus* peaked at harvest 5 (seed filling stage) (Fig. 3.7), in contrast with reports by Holmes (1980) and Barraclough (1989) who showed the maximum P accumulation occurred during late seed stage or at maturity. The differences results obtained might be due to natural variation in the species. Total P content was significantly lower in plants grown in this study compared to those grown by Rose et al. (2007), where total P contents were between 60 and 80 mg plant⁻¹ compared to 25-30 mg plant⁻¹ (Fig. 3.6). This might be associated with different soils (sandy loam v peat-based compost), pot size (2.5 L v 1 L) or fertiliser application rates used between the studies.

3.5.3 Transcriptional responses to Pi availability in *B. rapa*

B. rapa transcriptional responses to Pi efficiency were analysed using q-PCR. Pooled leaf RNA from *B. rapa* grown hydroponically under P sufficient (P+) and deficient (P-) conditions was extracted and quantified. Fourteen target genes consistently triggered by Pi deficiency from other plants such as *B. napus* and wheat were selected (Ligaba et al., 2006, Delhaize et al., 2009). The target genes were analysed in q-PCR using SAND and TATA box as reference genes and P+ as control, and the result showed robust responses in their expression profile (Fig. 3.11). Bro18 014187 (*ALMT1*) gave the highest response to Pi deficiency, increasing by nearly 8-fold. Other *ALMT1* paralogues showed variable changes under Pi deficiency; *Bro18 006954* was up-regulated by 1.3-fold, while *Bro18 042204* was down-regulated under Pi deficiency (Fig. 3.11). *ALMT1* is an Aluminium tolerance gene commonly related to Pi deficiency reported in many crops such as *B. napus* (Ligaba et al., 2006) and wheat (Sasaki et al., 2004). In Arabidopsis, *ALMT1* showed a complex pattern of regulation and expression in response to abiotic stress (Kobayashi et al., 2013). *ALMT1* is involved in organic acid exudation (OA) in the root system to improve PUE (Ryan et al., 2001). OA exudates, such as malate and citrate provide organic anions to mobilise Pi either by exchanging anions to release Pi in acidic soil, or through metal ion chelation which immobilises Pi in the soil (Sas et al., 2001). In wheat, *ALMT1* was induced at the plasma membrane and it facilitates the release of organic acids, to increase the solubility and availability of Pi anions to be taken up by the plant (Yamaguchi et al., 2005, Palmer et al., 2016). In wheat, *TaALMT1* encoding protein channel at the plasma membrane providing the efflux of malate from the root (Delhaize et al., 2009; Delhaize et al., 2012). When overexpressed in transgenic barley plants, the plants accumulated more P on acidic soils in comparison to the wild type (Delhaize et al., 2009). However, this was not observed when the soil was limed. The high response of this gene supporting report in the other member of the *ALMT* family in controlling stomatal movement (Meyer et al.,

2010). In Arabidopsis, ALMT members were shown to be localised in the tonoplast (AtALMT9 and AtALMT6) and membrane (AtALMT12) of the guard cell (Meyer et al., 2010; De Angeli et al., 2013). The stomatal movement was initiated by the release of Pi anions which in turn depolarise the membrane potential (Meyer et al., 2010).

Another family of membrane protein involved in OA exudation is transporter of the multidrug and toxic compound extrusion type (MATE) family (Baker et al., 2015). Members of this family have been shown to localise at the plasma membrane and contribute to citrate efflux transport activities in many plants such as Arabidopsis, sorghum, and soybean (Rogers et al., 2009; Carvalho et al., 2016). Gene expression data in *B. napus* showed contrasting regulation of *MATE* paralogues under Pi deficiency; *Bro18 038689* (*MATE*) was upregulated whereas *Bro18 035630* (*MATE*) was down-regulated in leaf tissues (Fig. 3.11). *MATE* genes have been shown to have high expression in root tissues compared to leaf tissues in plants under Pi-stress (Ryan et al., 2001).

MMDH1 gene encodes a mitochondrial malate dehydrogenase and plays a role in plant and root growth, being involved in physiological functions and metabolic pathways (Imran et al., 2016). *MMDH1* has the capacity to increase malate production in roots, leaves and root exudates to increase Pi acquisition (Wang et al., 2015). In this study, two of three *MMDH1* paralogues showed high expression (*Bro18 038513* and *Bro18 0133770*), and one unchanged (*Bro18 035465*) under Pi deficiency. The expression profiles showed *MMDH1* highly responsive to Pi availability in *B. rapa* in line with other research on cotton (*Gossypium hirsutum*) (Wang et al., 2015).

The transcript abundance of a purple acid phosphatase (*PAP*) remained unchanged in *B. rapa* leaf tissues under Pi deficiency (Fig. 3.11). However, the transcript abundance of *PAP12* was significantly higher (3.7-fold) under Pi deficiency. *PAP12* is involved in extracellular Pi scavenging in Arabidopsis (Haran et al., 2000). *PAP12* is an acid

phosphatase (APase) secreted from the surface of the root epidermal layer and hydrolyse Pi from extracellular Pi monoesters to increase the availability of Pi for uptake by the roots (Tran et al., 2010; Wang et al., 2011a).

In this study, Citrate Synthase CS4 paralogues showed different responses during Pi deficiency. *Bro18 009479* (CS4) was down-regulated and *Bro18049018* (CS4) was up-regulated during Pi deficiency. Citrate Synthase regulation is essential in tricarboxylic acid (TCA) cycle for energy metabolism in plants. During Pi deficiency, roots can excrete organic acid (citrate) to improve PUE by the chelation of metal cations that immobilise Pi and could increase the Pi concentration by up to 1000-fold (Plaxton & Tran, 2011; Schmidtman et al., 2014). The effectiveness of Pi remobilisation depends on different forms of organic acid secreted into the rhizosphere soil. Citrate is more effective at solubilising cation-phytate salts compared to another organic anions such as malate and oxalate (Tang et al., 2006).

The transcript abundance of two *SUC2* paralogues in *B. rapa* leaves were both up-regulated more than 1.5-fold in this study, although not significantly. Sugar sensing and signalling component such as *SUC2* gene is an important gene involved in almost all plants under Pi deficiency and encodes a sucrose-proton symporter which is essential for loading sucrose into the phloem (Gottwald et al., 2000; Lei et al., 2011). *SUC2* gene is also responsible for the *pho3* mutant in Arabidopsis and accumulates large amounts of sugars and starch (Lloyd & Zakhleniuk, 2004). This result supports previous experimental results showing the expression of many genes involved in sucrose synthesis, translocation and degradation are altered during Pi deficiency in other plants such as soy bean (*Glycine max*), Arabidopsis, and barley (*Hordeum vulgare*) (Wu et al., 2003; Hammond & White, 2008; Lei et al., 2011).

The different or contrasting gene expression patterns observed in this study may be in part the result of the WGT event, where a single homologous gene in Arabidopsis, may be represented by up to three copies in *B. rapa*. Neo-functionalisation of paralogue genes can then occur when the gene remains functional but may take on a new role or be silenced to avoid any functional redundancy (Tang & Lyons, 2012). Furthermore, in comparison with Arabidopsis, *B. rapa* has more developed tissues regarding the size and composition of the seed as well as seed dormancy and the genes might express in the specific cell types or tissues after the polyploidy process between species. This gives a higher spatial resolution of gene expression. This is supported by Zhou et al. (2017) who showed that the duplicated genes had spatially different expression patterns in *B. napus* after allopolyploidisation.

During Pi deficiency, the general response of plant is the up regulation of intracellular (vacuolar) and increase secretion of APases which hydrolyse Pi from Pi monoesters under optimum acidic pH. Pi remobilising to maintain Pi homeostasis by enhancing external Pi acquisition, limit its internal Pi consumption as well as redistribution and recycling Pi internally to other plant parts or cellular compartments (Ragothama, 1999; Poirier & Bucher, 2002). Pi remobilisation was facilitated by alteration of metabolic pathways and reallocation of internal Pi to different organs and cellular compartments. These responses are in line with the findings from this chapter; the plant experienced physiological changes through the biomass allocation and redistribution of P to the developing organ in matured plants. The expression profiles also showed high expression of Pi-responsive genes under Pi starvation, such as *SUC2*, *MATE*, *ALMT1*, *MMDH1* and *PAP12* which are involved in signalling, transport, extracellular Pi scavenging, increasing Pi availability to be taken up by the plant as well as involved in plant physiological function and metabolic pathways. The analysis of plant P nutrient status and expression profile is important for crop optimal growth and production under low Pi supply.

**CHAPTER 4 IDENTIFICATION OF THE MAJOR
REGULATORS UNDERLYING *TRANS*- EXPRESSION
QUANTITATIVE TRAIT LOCI (E-QTL) HOTSPOTS IN *B.*
*RAPA***

4.1 AIMS AND OBJECTIVES

The aims and objectives in this chapter are;

1. Identification of candidate genes at a *trans*-eQTL hotspot of *B. rapa*.
2. To validate the *B. rapa* genome sequence through DNA sequencing of the target region of interest.
3. Identification and characterisation of *PHO1* homolog genes.

4.2 BACKGROUND

The advent of next generation sequencing technologies and genotyping approaches has allowed genome-wide expression profiling within natural accessions and structured populations to be carried out in plant systems, many of which are industrially or economically relevant, such as *B. rapa* (Simon et al., 2008; Gan et al., 2011). Expression quantitative trait loci (eQTL) analysis has been used to locate and identify the underlying genes that control the phenotype. A framework genetic map of the whole genome could be developed through the combination of the QTL analysis method and the data of transcript abundance of particular genes in samples taken from genotypes in a segregated population (Druka et al., 2010) (section 1.8). Data from a previous eQTL study on *B. rapa* under altered soil Pi supply have shown characteristic transcriptional events in response to Pi deficiency and revealed higher numbers of eQTLs than expected based on physical and genetic maps on two chromosomes, chromosome A06 and A01; these were enriched with P metabolism related Gene Ontology terms (A06) as well as chloroplast and photosynthesis related terms (A01) (Hammond et al., 2011) (Fig. 4.1). This suggested *trans*-eQTL hotspots occurred in *B. rapa* and identifying candidate genes at these *trans*-eQTL hotspots may be good targets for future crop improvement. In this experiment, the aim is to identify candidate genes at *trans*-eQTL hotspots that might be responsible for the

regulation of a large number of genes in response to low Pi availability. Identification of candidate genes at eQTL hotspots will increase our understanding of the genomic architecture and genetics of phosphorus use efficiency (PUE) at the individual gene level. In addition, it may generate more opportunities for crop improvement based on candidate gene and marker identification at a scale that is much more rapid than one based on trait QTL approaches alone.

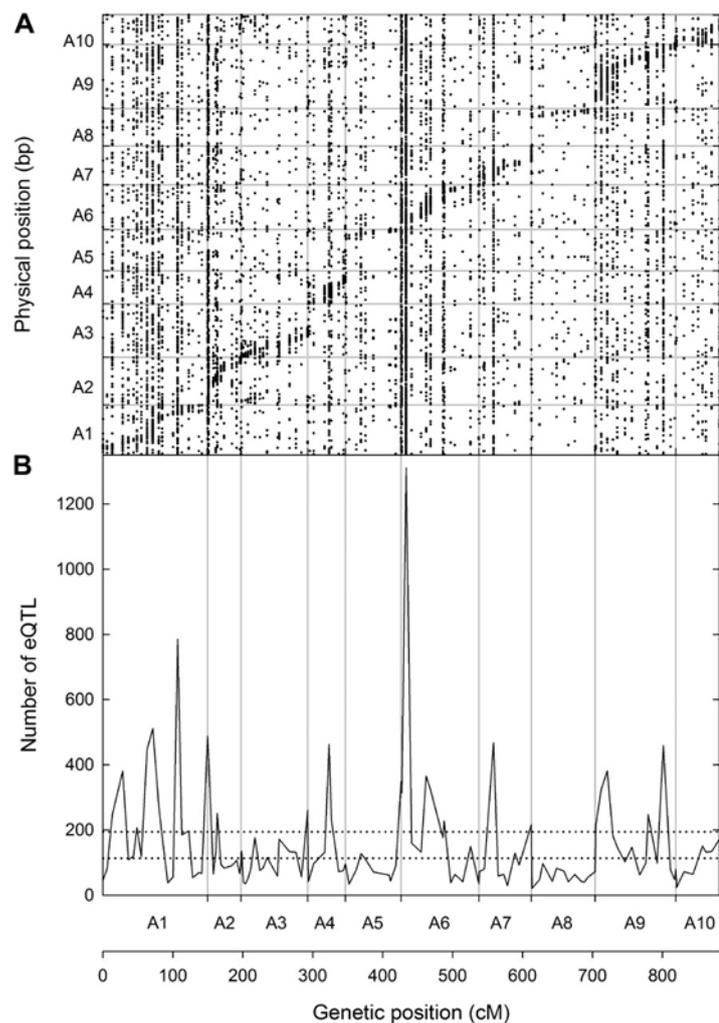


Figure 4.1. Comparison between genetic and physical positions of 14,257 eQTL. **A.** Physical positions (bp) of individual genes (dots) against genetic locations (cM). **B.** The number of eQTL associated with individual gene expression markers across the *Brassica rapa* A genome. Horizontal dotted lines represent empirical thresholds for the upper 99% and 95% confidence interval for the Poisson distribution, assuming an equal distribution of GEMs and eQTL. Markers with values higher than these lines are defined as eQTL hotspots. (Hammond et al., 2011).

4.3 DNA AND RNA SEQUENCING TECHNOLOGIES

Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are the genetic material that define every organism. DNA/RNA sequencing is the process for determining the exact order of nucleotides ((adenine (A), guanine (G), cytosine (C), thymine (T) and uracil (U)) within a DNA/RNA molecule. It is very important for research into the understanding of cell or organism development and function, and is used in a broad range of applications such as finding cancerous genes in medicine, plant breeding in agriculture and genetic and evolution studies in humans (Heather & Chain, 2016).

The first breakthrough of DNA sequence platform was obtained through the Sanger sequencing method developed by Frederick Sanger in 1977, also known as first generation sequencing (Sanger et al., 1977). Sanger developed 'chain-termination' technique using the chemical analogues of the deoxyribonucleotides (dNTPs), the monomers of DNA, called dideoxynucleotides (ddNTPs). Dideoxynucleotides (ddNTP) lack the 3' hydroxyl group that is required for extension of DNA sequence, therefore the replication stops as they cannot form a bond with the 5' phosphate of the other dNTP (Chidgeavadze et al., 1984). The products were then separated by length using gel electrophoresis and autoradiography was used to determine the nucleotide sequence. Sanger sequencing is a simple method; however it requires extensive labour, time and cost to sequence a whole genome of an organism. This method was the common method used for sequencing DNA until 2005, when the second-generation sequencing or next generation sequencing (NGS) emerged. The system developed by 454 Life Sciences adopts pyrosequencing technology to produce large amounts of sequence from multiple samples with high-throughput and high sequence coverage (Margulies et al., 2005). The basic characteristics of NGS are the generation of random short fragments of genomic DNA or cDNA (reverse transcribed from RNA). Adapter sequences are then ligated to these short fragmented genomic DNA or cDNA sequences for the construction of libraries

(Kulski, 2016). Other commercial NGS platforms, such as the Genome Analyzer by Solexa was launched in 2006 using the technology of sequencing by synthesis (SBS), followed by Sequencing by Oligo Ligation Detection (SOLiD) by Agencourt where the sequencer uses the technology of two-base sequencing based on ligation sequencing (Mardis, 2008). These companies were later purchased by other companies; Applied Biosystems took over Agencourt in 2006, Roche took over 454 in 2007, while Solexa was purchased by Illumina (Barba et al., 2014).

The NGS technology and rapid development of bioinformatics tools enable many researchers to generate *de novo* (sequencing a novel genome without reference sequence) genome assemblies as well as constructed whole genome sequencing (WGS) data for many organisms (Kulski, 2016). High throughput technologies in NGS also have been used for many other types of research such as for whole transcriptome profiling or RNA sequencing (RNA-seq), whole-exome sequencing (WES), candidate gene sequencing (CGS) as well as methylation sequencing (MeS) (Wang et al., 2009; Pelizzola and Ecker, 2011; Leo et al., 2015; Bertier et al., 2016). The use of NGS technologies have proven their capabilities in generating more accurate, detailed and thorough analysis (McGinnis et al., 2016). However, a few drawbacks of the NGS systems are their error-prone nature as well as the extensive needs in computation time and storage (Daber et al., 2013).

Third generation sequencing is a new advancement in DNA sequencing technologies. There are two main criteria of third generation sequencing; first, it could generate direct sequencing of a single DNA molecule in real-time. Second, it saves DNA preparation time and cost since it does not require polymerase chain reaction (PCR) before sequencing (Thompson & Milos, 2011).

Currently, there are three third-generation sequencing technologies available commercially. First, single molecule real-time (SMRT) sequencing developed by Pacific Bioscience (PacBio) using PacBio RS II device. This method uses a modified enzyme and direct observation of the enzymatic reaction through fluorescence detection in real-time (Eid et al., 2009). Second, the Moleculo protocol also known as Illumina TruSeq Synthetic Long Reads which could generate very accurate long reads of up to 18.5 Kbp from the short sequences (Voskoboynik et al., 2013). Third, nanopore based sequencing devices, developed by Oxford Nanopore Technologies (ONT), which use the electronic disruption in a nanopore as the DNA molecule passes through it to determine the DNA sequence (Branton et al., 2008).

4.4 OXFORD NANOPORE SEQUENCING MinION

MinION from Oxford Nanopore Technologies (ONT) is a portable device that can be used to directly sequence DNA molecules (Fig. 4.2). The MinION device is able to sequence long DNA molecules, such as BACs, using nanopores as biosensors to detect ionic current changes as the DNA passes through the nanopore one base at a time in real-time. Using these data could potentially improve the *de novo* assembly of genomes as well as elucidate the structural variants and complex rearrangement with a high accuracy (Magi et al., 2017). As with other third generation sequencing technologies, MinION can produce long reads (up to 200 Kb) and the data can be generated on a real-time basis (Leggett & Clark, 2017). It is compact in size, weigh only about 90 g, portable and easy to use only by connecting to the USB 3.0 port of a laptop (Fig. 4.2).

4.4.1 The principle of MinION sequencing

Nanopores are biological pores, embedded in a high resistant artificial membrane. The prepared DNA molecule contains a leader adaptor and motor protein (to unzip double-stranded DNA) attached to one strand of the DNA molecule. During sequencing, a single strand of DNA passes through the pore one base at a time. The presence of DNA in the pore causes a disruption to the current across the pore which is specific the base in the centre of the pore and can be detected by a sensor to determine the exact bases present at the pore at that moment (Jain et al., 2016).

The core of the MinION is a flow cell containing an array of 512 sensors, where each of the sensors is connected to four nanopores (one of which is in use at any one time) and controlled by an application-specific integrated circuit (ASIC) (Fig. 4.2). MinKNOW control software is used in the computer connected to the MinION to record the signal (current) data from all channels, and subsequently store all the data in the local hard drive. The data stored is then converted to a nucleotide sequence by a process called base-calling either locally or through a cloud-based service.

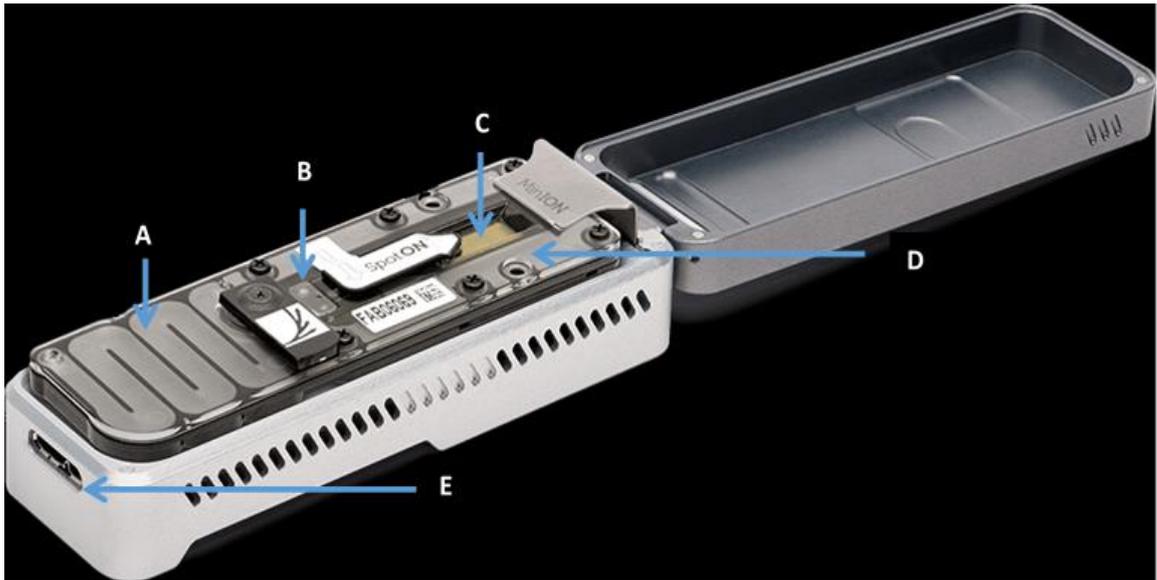


Figure 4.2. MinION sequencer device; **A.** consumable flowcell contains sensing chemistry, nanopores and electronics; **B.** sample added to flow cell; **C.** sensor chip with multiple nanopores; **D.** Sensor chip works with custom ASIC for control and data processing **E.** USB power device.

4.5 MATERIALS AND METHODS

4.5.1 Candidate genes and primer design

Microarray gene expression data from *B. rapa* R-o-18 under Pi deficient conditions was used to produce a list of genes whose expression was significantly differentially expressed in response to low Pi availability (Hammond et al., 2011). This list of genes was compared to a list of genes in the genomic region around the eQTL hotspot on Chromosome A06 using Microsoft Access to identify genes in this region whose expression was altered by Pi availability. Identification of candidate genes underlying major regulatory *trans*-eQTL hotspot revealed four candidates, all belonging to PHO1 family. Primers were designed for the genes of interest in *B. rapa* using the same method as the previous experiment (section 2.8) and tested using the PCR and gel electrophoresis to check the specificity of the primers. The genes were *Bra019686*, *Bra019688*, *Bra019689* and *Bra019690* (Table 4.1).

Table 4.1. List of the primers, sequence, and length of expected genomic DNA (gDNA) and complementary DNA (cDNA) to amplify genes of interest; *PHO1_A* (*Bra019686*), *PHO1_B* (*Bra019688*), *PHO1_C* (*Bra019689*) and *PHO1_D* (*Bra019690*).

Primer	Sequence (5' to 3')	Expected size (bp)
Bra019686_F	GACTTCCAACATCGCTACTTCC	gDNA= 3668
Bra019686_R	CACTTCATCGGGCTCTTCTC	cDNA= 193
Bra019688_F	AACCGCTGGCTTCGTGATAA	gDNA= 231
Bra019688_R	AGTAACCATTGTCTGTCTGTGC	cDNA= 147
Bra019689_F	GGATGAGAAAGGAGGAGGAG	gDNA= 208
Bra019689_R	AGCGATACCAGAAGCGAATG	cDNA= 208
Bra019690_F	CGCTGGCTTCGTGATAAACT	gDNA= 243
Bra019690_R	AGCAACAAGAGTAACCATCGT	cDNA= 152

4.5.2 Bacterial artificial chromosome (BAC) sequencing

Five BAC clones, KBrB029J08, KBrB003E10, KBrB063F11, KBrH038K12 and KBrH102C10 (Table 4.2) from BAC libraries of *B. rapa* ssp. *pekinensis* cv. Chiifu, which contain the region of interest were identified from CloneFinder (NCBI, www.ncbi.nlm.nih.gov/clone). The clones were kindly supplied by Prof. Yong Pyo Lim at the Molecular Genetics and Genomics Laboratory, Department of Horticulture, Chungnam University, South Korea. The BAC's were sub-cultured in 10 mL of Luria-Bertani (LB) media containing 12.5 µg mL⁻¹ chloramphenicol by vigorously shaking for 6-12 h at 37 °C. Glycerol (50:50) was added to a sub-sample and these were kept at -80 °C for long-term storage.

Table 4.2. BAC libraries constructed by insertion of DNA fragments from *Brassica rapa* ssp. *pekinensis* cv. Chiifu into a vector *Escherichia coli* (*E. coli*) as a host.

BAC clone	Chromosome	Size	Host	Vector	Restriction site
KBrB063F11	A6	134,576	<i>E. coli</i>	pCUGIBAC1	BamHI
KBrH102C10	A6	152,113	<i>E. coli</i>	pCUGIBAC1	HindIII
KBrB029J08	A6	148,972	<i>E. coli</i>	pCUGIBAC1	BamHI
KBrH038K12	A6	155,059	<i>E. coli</i>	pCUGIBAC1	HindIII
KBrB003E10	A6	136,929	<i>E. coli</i>	pCUGIBAC1	BamHI

4.5.2.1 BAC DNA sub-culture and single colony preparation

The BAC clones were sub-cultured in 10 mL of LB with 12.5 µg mL⁻¹ chloramphenicol under sterile conditions and were shaken vigorously for 6-12 h in the incubator at 37 °C. BAC clones were then spread on separate LB agar plates with 12.5 µg mL⁻¹ chloramphenicol to enable single colonies for each BAC to be selected. The agar plates of each BAC were incubated overnight at 37 °C.

4.5.2.2 BAC DNA isolation

Starter cultures were prepared by inoculating 10 mL LB with 12.5 µg mL⁻¹ chloramphenicol under sterile conditions with a single colony for each BAC from the agar plates (section 4.5.2.1). Starter cultures were shaken vigorously for 6-12 h at 37 °C. A 5 mL aliquot of the starter culture was then added to 500 mL LB media and was shaken overnight at 37 °C for subsequent large-construct BAC DNA isolation. BAC DNA isolation was performed using the Qiagen Large-Construct Kit (Qiagen, Manchester, UK) according to manufacturer's protocol. BAC DNA was quantified using a Nanodrop 2000 spectrophotometer and Qubit fluorometer. The quality of DNA was confirmed with ratios

of $Abs_{260/280}$ and $Abs_{260/A280}$ above 1.8. An aliquot of 1 μ g of each BAC was used as the initial BAC DNA concentration for nanopore sequencing by MinION.

4.5.2.3 BAC DNA confirmation and validation

Primers were designed from BAC ends as described previously (Section 2.8; Table 4.3). BAC end sequences were then amplified to confirm their identity. The following cycling parameters were used for the amplification: 1 cycle of 1 min at 95 °C; 40 cycles of 15 s at 95 °C, 15 s at 65 °C, 15 s at 70 °C; one cycle of 5 min at 72 °C. The PCR products were verified through electrophoresis as described earlier (section 2.7.1).

4.5.2.4 BAC DNA precipitation

To each DNA sample a 0.1 volume of 3 M sodium acetate pH 5.2 (Thermo Scientific) and 2.5 volumes of ice cold 100% ethanol were added. The resulting solution was vortexed briefly to mix and stored at -20 °C overnight to precipitate the DNA. DNA was pelleted by centrifugation at 12 000 g for 20 min at 4 °C. The supernatant was poured off and 500 μ L of 70% ethanol added and inverted several times to wash the pellet before being briefly vortexed. This step was repeated twice. The DNA was pelleted in a centrifuge at a 12 000 g for 5 min at 4 °C. The supernatant was decanted, and the pellet air-dried briefly at room temperature. The DNA was re-suspended in 50 μ L nuclease free water (NFW).

Table 4.3. List of the primers (forward and reverse) for BAC DNA end sequences and expected length of the PCR products.

BAC end primer	Sequence (5'to 3')	Length (bp)
KBrB-063F11F_F	TCCCCCGTTTAATTCTCTCTC	
KBrB-063F11F_R	GCGCGGGTCAAATCTAGTA	244
KBrB-063F11R_F	TGGACCATTCTTCAGGAACC	
KBrB-063F11R_R	ACATTGCACATCCTCCCATT	208
KBrH-102C10F_F	GCGCTATCAAAGGGTTCAAG	
KBrH-102C10F_R	GAAGTGGCAGGACCAACATT	185
KBrH-102C10R_F	GCCCTTCGACAAAATCCATA	
KBrH-102C10R_R	CTCTAGCGCGTTGGTTAAGG	233
KBrB-029J08F_F	TCGTGAACTTCTCCATCGTG	
KBrB-029J08F_R	GTGGGAAACTGTGGAAAGGA	226
KBrB-029J08R_F	AATGGGAGGATGTGCAATGT	
KBrB-029J08R_R	GTTCTCCATCGCTGATAGGC	169
KBrH-038K12F_F	CTGACACCAGCACTGAAGGA	
KBrH-038K12F_R	ATGGAGGAATGTGGGATTGA	177
KBrH-038K12R_F	TCGAGAGGGAGTAACCAGGA	
KBrH-038K12R_R	CAAGAGGGCAACAGAGAACC	228
KBrB-003E10F_F	GTGGACGGTCATTGGATCAT	
KBrB-003E10F_R	TTCCTTGGGAGAGAAGCTGA	150
KBrB-003E10R_F	TGGACCATTCTTCAGGAACC	
KBrB-003E10R_R	TCCTCCATTGCTCCATATC	198

4.5.2.5 Sequencing library preparation for MinION sequencing

A 1 µg subsample of BAC DNA in 45 µL NFW was fragmented by loading the sample into a G-tube (Covaris, Brighton, UK) and centrifuged at 6,000 g for 2 min before inverting the tube and centrifuging again for 1 minute. The fragmented DNA was repaired by adding 8.5 µL NFW, 6.5 µL FFPE repair buffer and 2 µL FFPE repair mix (NEBNext FFPE RepairMix, NEB, Hitchin, UK). The resulting repaired DNA was cleaned-up using 62 µL AMPureXP beads (Beckman Coulter, High Wycombe, UK) according to manufacturer's instructions at room temperature, and was eluted in 46 µL NFW. The cleaned FFPE repaired DNA was quantified using a Qubit fluorometer with expected recovery of greater than 1000 ng of material. Ligation library preparation was performed using SQK-LSK108

1D Ligation Sequencing Kit (Oxford Nanopore Technologies, Oxford, UK) with EXP-NBD103 Native Barcoding kit (Oxford Nanopore Technologies) (Table. 4.4; Fig. 4.3). The end-repair and dA tailing were performed using NEBNext Ultra II End-Repair kit (NEB) in a total volume of 60 μ L according to manufacturer's instructions. The end-repaired DNA was cleaned-up by adding 60 μ L AMPure XP beads (NEB) according to manufacturer's instructions. The DNA was then quantified again using Qubit fluorometer with recovery target of 700 ng of material. The protocol for native barcoding genomic DNA was performed according to manufacturer's instructions with some modifications to maximise DNA recovery (Oxford Nanopore Technology). Following the barcode ligation reaction, the DNA was cleaned again with AMPure XP beads and eluted in 26 μ L of NFW. For library pooling, the amount of DNA added per BAC clone was calculated based on 1 μ L of DNA from the sample with the lowest concentration. All other samples were added accordingly to produce an equimass pool of BAC DNA in volume of 65.45 μ L. A subsample of 1 μ L was used for Qubit fluorometer quantification, and 64.45 μ L of this pooled DNA sample was used for adapter ligation using E6056 NEBNext Quick Ligation Module (NEB) according to manufacturer's instructions.

Table 4.4. Oxford Nanopore Technologies (ONT) barcodes used for BAC sequencing.

Barcode ID	Primer Sequence	BAC ID
NB01	AAGAAAGTTGTCGGTGTCTTTGTG	KBrB-029J08
NB02	TCGATTCCGTTTGTAGTCGTCTGT	KBrB-003E10
NB03	GAGTCTTGTGTCCCAGTTACCAGG	KBrB-63F11
NB04	TTCGGATTCTATCGTGTTCCCTA	KBrH-038K12
NB05	CTTGTCAGGGTTTGTGTAACCTT	KBrH-102C10

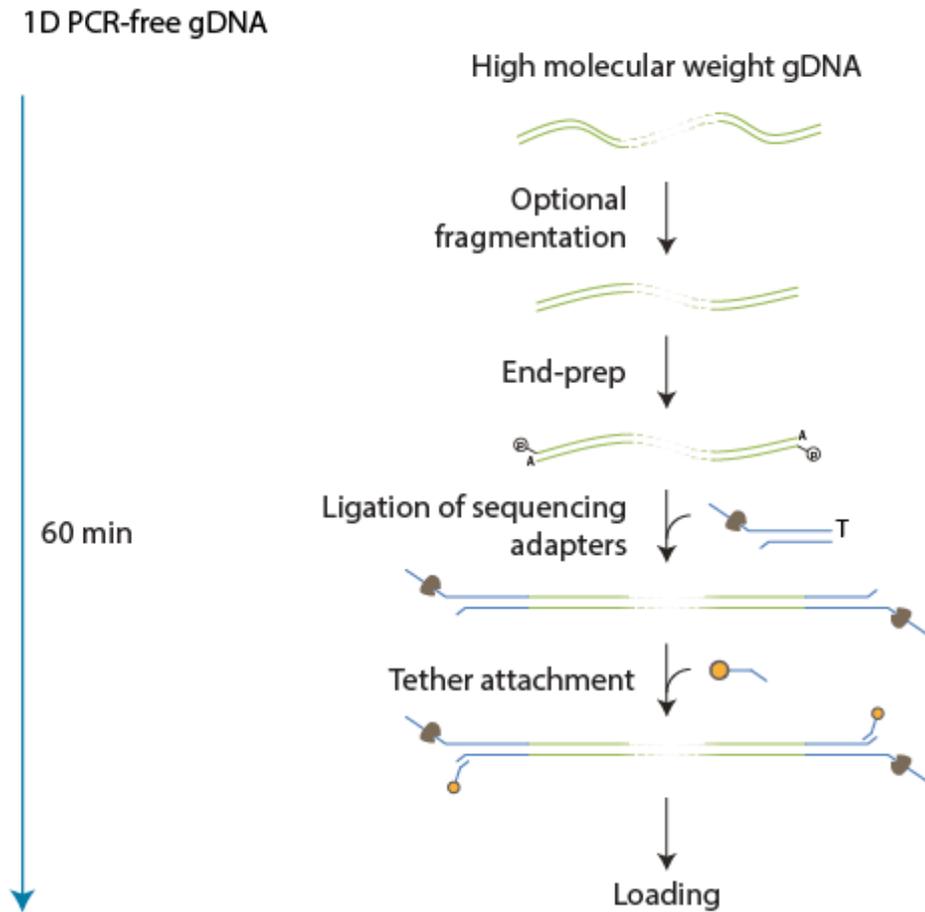


Figure 4.3. Overview of the BAC DNA sequencing workflow. Ligation Sequencing Kit (SQK-LSK108) used for a library preparation, a 60 min protocol producing 1D reads (one strand of the fragment).

4.5.2.6 MinION flowcell preparation

A new flowcell was removed from storage at 4 °C. The flowcell was inserted into the MinION device and MinKNOW software was used to perform Platform QC. The number of active pores was recorded after the QC completed. The flowcell was then prepared by adding a mixture of 504 μ L Running buffer with fluid mix (RBF) buffer and 546 μ L NFW into the sample loading port to prime the flow cell and left for 5 min.

4.5.2.7 MinION sample loading

The library was quantified using Qubit fluorometer ($0.27 \text{ ng } \mu\text{L}^{-1}$). A $14 \text{ } \mu\text{L}$ aliquot of the adapted and tethered (providing passage between the fragment and the nanopore) library was loaded into the sample loading port of the flowcell.

4.5.2.8 MinION sequencing

A 48-hour sequencing protocol was initiated using the MinION control software, MinKNOW (Oxford Nanopore Technologies). Read event data were base-called by the standalone Albacore software (Oxford Nanopore Technologies) due to computer re-setting after 48 h (Fig. 4.4). The Albacore software was also used to separate sequences based on barcodes.

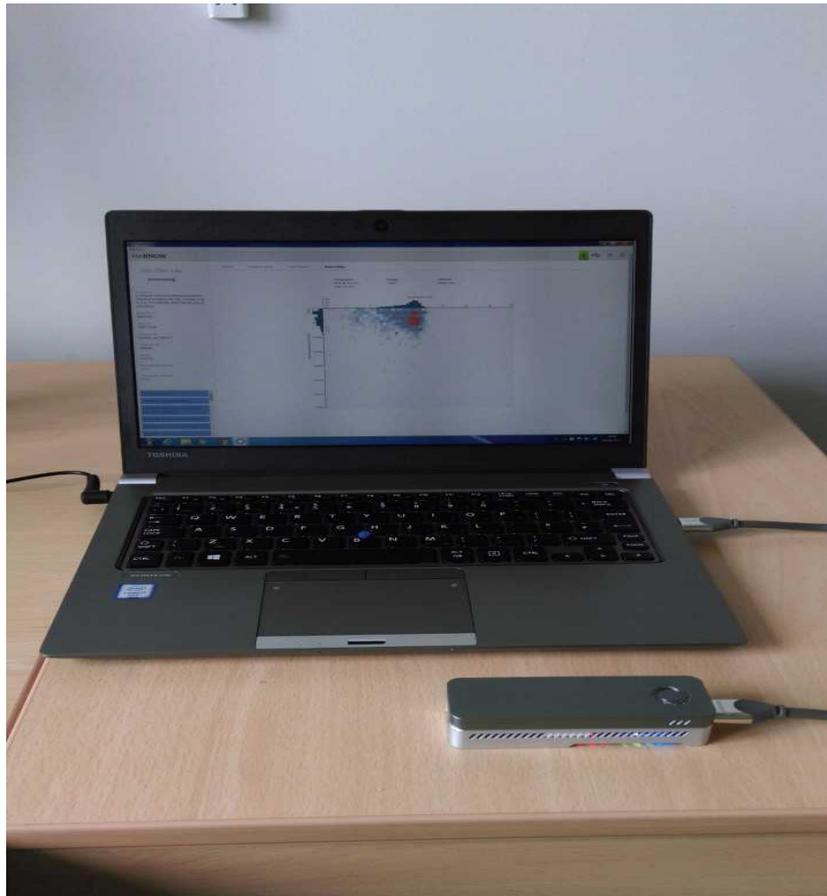


Figure 4.4. MinION device on the run for 48 hours.

4.5.2.9 MinION data analyses

Canu software was used to *de novo* assemble individual BAC sequences (Koren et al., 2017) (Fig. 4.5). Process of assembling data involved three phases; 1. correction, 2. trimming and 3. assembly. The quality of the bases in reads is checked and adjusted in the correction phase. Second, the high-quality sequences were produced through trimming process; and finally, in assembly phase the reads are used to generate contigs and consensus sequences. Canu supports nanopore sequencing by enhancing the quality of sequence assembly compared to other software (Koren et al., 2017). The sequences were then aligned to look at the plasmid backbone and the BAC ends using CLC Sequence Viewer software 7.8.1 (Qiagen). The plasmid sequence was removed to obtain the actual size of BAC DNA sequenced and the size of each BAC was determined and compared to their predicted size. The target genes (*PHO1*) were aligned to each BAC using CLC Sequence Viewer software 7.8.1 (Qiagen).

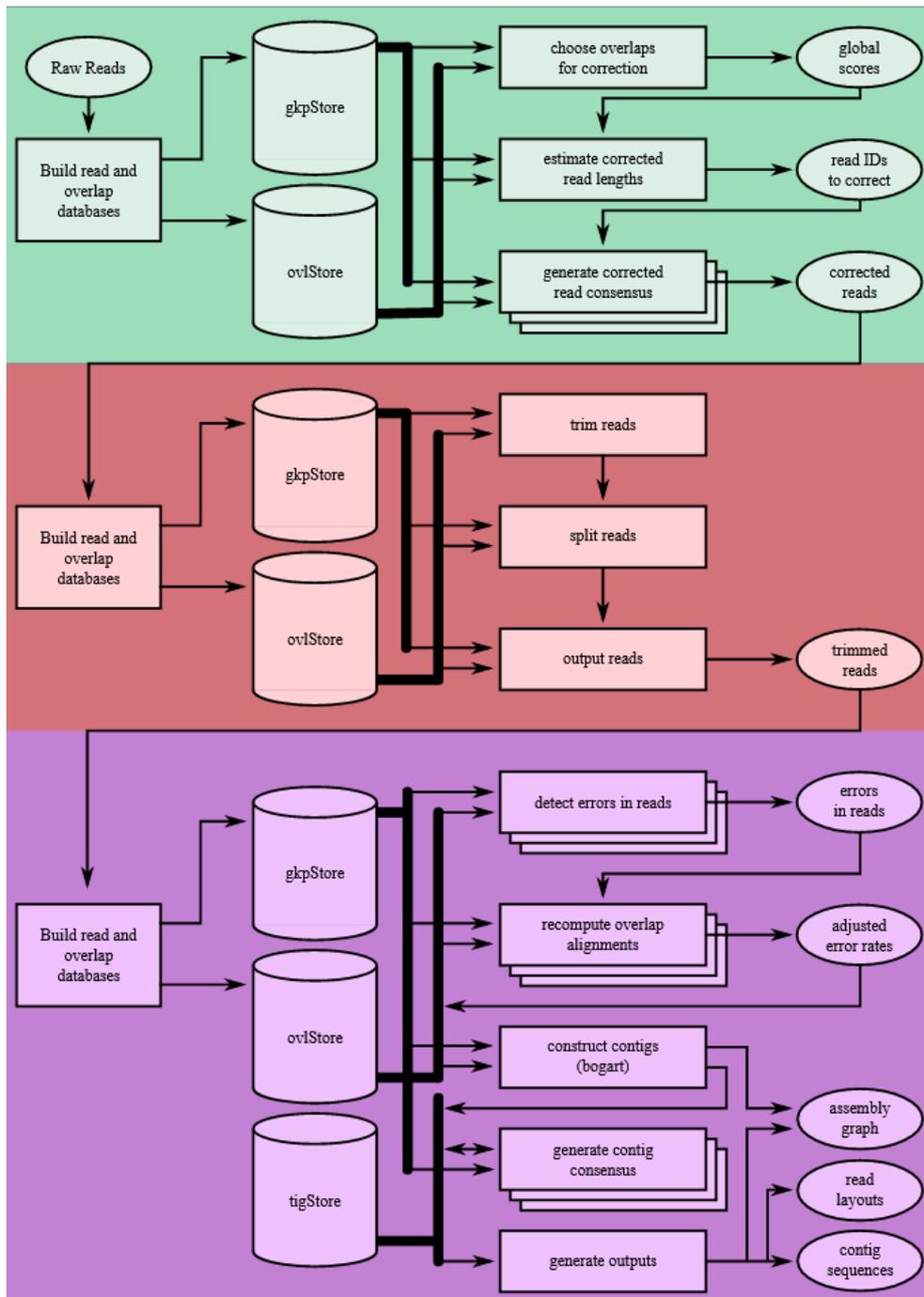


Figure 4.5. Workflow for Canu software to generate read data including three stages: correction (green), trimming (red), and assembly (purple). Picture extracted from Koren et al., 2017.

4.5.3 Identification and characterisation of *PHO1* homolog genes

4.5.3.1 RNA extraction and cDNA synthesis

Total RNA of pooled leaf samples of *B. rapa* grown hydroponically was extracted and quantified according to the method described in 2.7. First strand cDNA was synthesized using SuperScript™ III First-Strand Synthesis SuperMix (ThermoFisher Scientific, UK) following the manufacturer's instructions. Briefly, 1 µL each of total RNA (1 µg), annealing buffer and oligo (dT) 20 primer (50 µM), and 5 µL of NFW were added to a sterile 0.2 mL thin-walled PCR tube on ice. The mixture was incubated at 65 °C for 5 min in a thermal cycler (Veriti, Applied Biosystems, UK).

The tube was immediately placed on ice after incubation and 10 µL of 2X first-strand reaction mix and 2 µL of SuperScript III/RNaseOUT enzyme mix were added to yield a total sample volume of 20 µL. The sample was then briefly vortexed to mix and centrifuged to collect the contents. The mixture was incubated at 50 °C for 50 min and the reaction was terminated by heating at 85 °C for 5 min and chilled on ice. The cDNA was diluted to 1:10 with NFW and stored at -20 °C.

Previous research revealed candidate genes at *trans*-eQTL hotspots on chromosome A06 *B. rapa*. Analysis of microarray data (not shown) revealed the *B. rapa* genome contained four copies of *PHO1*, an identified regulator of plant responses to Pi availability (Stefanovic et al., 2007). However, only three copies are evident from short read transcriptome data (Fig. 4.6). Efficiency of short read sequencing for annotation of highly similar copies is needed to reveal the transcript existence. Therefore, to gain complete insight of the *B. rapa PHO1* gene, *PHO1* transcripts were amplified and cloned. Alignment of nucleotide (nt) sequences of *B. rapa PHO1* homologs XM_009150437 (*PHO1_A*), XM_018652610 (*PHO1_B*), XM_018652651 (*PHO1_C*) and XM_009150438 (*PHO1_D*) revealed three conserved regions which offered efficient primer design. A 21 nt long left primer containing

2 degenerate bases was designed incorporating the start codon and two right primers were designed to capture >2 Kb of PHO1 transcript(s), one of which is designed from complete consensus region among the four PHO1 mRNAs.

Nucleotide sequences of *B. rapa* PHO1 homologs XM_009150437 (*PHO1_A*), XM_018652610 (*PHO1_B*), XM_018652651 (*PHO1_C*) and XM_009150438 (*PHO1_D*), were retrieved from GenBank database (<https://www.ncbi.nlm.nih.gov/genbank/>). The retrieved sequences were aligned using CLC Sequence Viewer (QIAGEN Bioinformatics). Two sets of primers (Table 4.5) were designed from conserved regions in the alignment to harbour coding sequence of the four *PHO1* homologs using Primer3 plus software (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>).

Table 4.5. A. Primers designed to amplify cDNA. B. The expected sizes of *PHO1* cDNA.

A. Primer	Sequence (5' to 3')
PHO_1_F	ATGAMGTTTCGGKAAGGAGTTT
PHO_2360_R	ACATTGTTCAAATGCTYGTCT
PHO_2224_R	ATGCAAATCTCAGCAGGACA
B. Target genes	Expected size of cDNA (bp)
<i>PHO1_A</i>	2224
<i>PHO1_B</i>	2044
<i>PHO1_C</i>	2152
<i>PHO1_D</i>	2227

4.5.3.2 Reverse transcription polymerase chain reaction (RT-PCR)

Amplification was performed in 20 μL reaction containing 10 μL of Phusion Flash high – fidelity PCR master mix (Thermo Scientific, UK), 2 μL each of 5 μM forward and reverse primers, 3 μL of first strand cDNA, diluted 1:10. The PCR amplification was carried out with the following gradient cycle: 1 cycle of 1 min at 98 $^{\circ}\text{C}$; 35 cycles of 1 s at 98 $^{\circ}\text{C}$, 5 s at 65 $^{\circ}\text{C}$, 45 s at 72 $^{\circ}\text{C}$; one cycle of 1 min at 72 $^{\circ}\text{C}$. After PCR amplification, the products were run on a 1.2% agarose TAE gel along with Fast Ruler Middle Range DNA Ladder (Thermo Scientific) for 2 h at 80 V. PCR products were cleaned using GeneJET Gel Extraction and DNA Cleanup Micro Kit (Thermo Scientific) and subsequently quantified using a NanoDrop 2000 Spectrophotometer (ThermoFisher Scientific, UK).

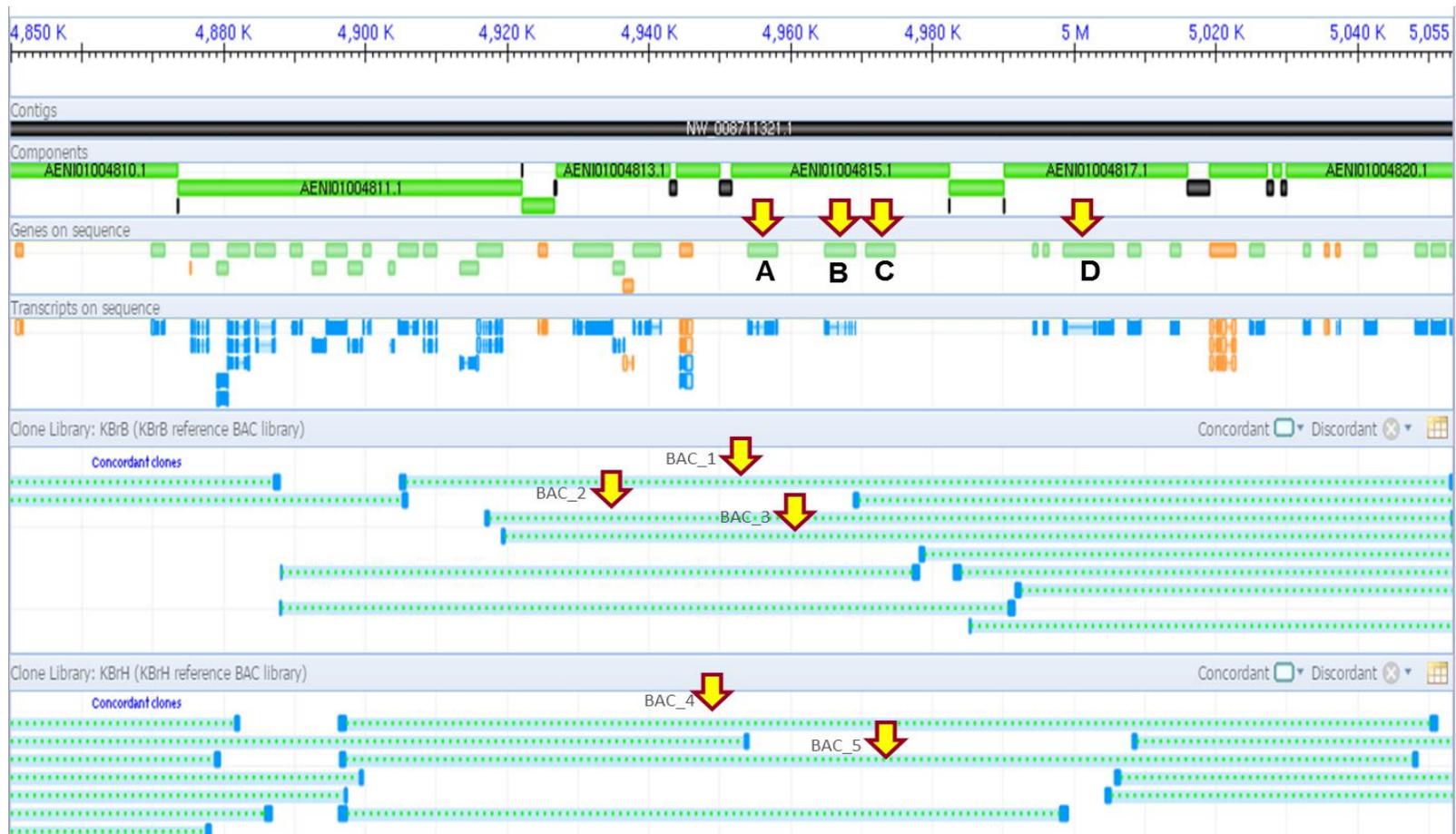


Figure 4.6. Location of four candidate genes at the *trans* eQTL hotspots on Chromosome A06 of *Brassica rapa* (Hammond et al., 2011). Source: *Brassica rapa* clone finder (www.ncbi.nlm.nih.gov/clone/). **A.** *PHO1_A* (Bra019686) **B.** *PHO1_B* (Bra019688) **C.** *PHO1_C* (Bra019689) **D.** *PHO1_D* (Bra019690).

4.5.3.3 Cloning of *PHO1* transcripts

The pool of *PHO1* transcripts was cloned into pCR4 Blunt TOPO vector using ZeroBlunt TOPO Kit (Invitrogen) following the supplier's instructions. Briefly, for ligation, a total of 1 μL (24 ng) of purified PCR product was added to a 0.2 mL PCR tube along with 1 μL of vector, 1 μL of salt solution (1.2 M NaCl and 0.06 M MgCl_2) and 3 μL of NFW. The contents were mixed gently and centrifuged shortly, incubated at 22 °C for 30 min and immediately chilled on ice. For transformation, 2 μL of ligation mixture was added to a single tube of chemically competent Top10 cells (ThermoFisher Scientific, UK). The tube was incubated on ice for 30 min. The transformation was performed by a heat shock at 42 °C for 50 s in a dry bath. The tube was chilled on ice for 2 min and supplemented with 250 μL of Super Optimal broth with Catabolite repression (SOC) medium (Sigma-Aldrich) to get a higher transformation efficiency of plasmids. The culture was incubated at 37 °C for an hour while shaking horizontally on an orbital shaker at 225 rpm. The culture was spread at 20 and 200 μL on two LB agar plates containing 50 mg L^{-1} of Kanamycin. The plates were incubated overnight at 37 °C. A total of 30 colonies were picked for colony PCR and subsequently cultured in LB liquid supplemented with 50 mg L^{-1} Kanamycin at 37 °C for 14-16 h with gentle shaking at 250 rpm. The plasmids were isolated using GeneJET plasmid miniprep kit (ThermoFisher, UK).

4.5.3.4 Restriction analysis

Isolated plasmids were analysed using HindIII restriction enzyme. For restriction, 1 μg of plasmid DNA was digested in 20 μL reaction containing 1 μL of enzyme and 2 μL of 10X restriction buffer. The restriction reaction was mixed and incubated at 37 °C for 1 h. Digested plasmids were run on 1.2% agarose TAE gel at 80 V until fully resolved.

4.5.3.5 Plasmid sequencing

The plasmids selected on the basis of restriction fingerprints were sent for sequencing using Sanger sequencing method to Source Bioscience (Nottingham, UK). The sequencing data were analysed using SeqMan (DNASTAR, Madison, USA). M13F and M13R primers were used to sequence the plasmid, and two internal primers were designed to capture the full-length sequence of plasmid 16 (Table 4.6).

Table 4.6. Internal primers sequenced to get complete sequence of plasmid 16.

A. Primer	Sequence (5' to 3')
PHO1_16_seq1	CAAAAGCTTCGGCTTCTCAA
PHO1_16_seq2	G TTCCTTCCTTGAACCCAAA

4.6 RESULTS

4.6.1 Candidate genes

Using previously published data (Hammond et al., 2011), genes whose expression was significantly differentially altered by low Pi availability were located in the region of e-QTL hotspot on Chromosome A06 (Table 4.7). One gene, (Arabidopsis ID *AT1G14040*) was of interest since it encodes PHO1, a previously identified regulator of plant responses to Pi availability (Stefanovic et al., 2007). Analysis of sequence data revealed four copies of *PHO1* genes located at this location of chromosome A06 in the *B. rapa* genome. The four candidate genes were *Bra019686* (*PHO1_A*), *Bra019688* (*PHO1_B*), *Bra019689* (*PHO1_C*) and *Bra019690* (*PHO1_D*). The alignment in Geneious software showed high similarities among all four candidate genes (Fig.4.7). Primers were designed to amplify the genes. PCR amplification using *B. rapa* gDNA and cDNA from leaf and root samples showed a variety of products for the different primer tissue combinations (Fig. 4.8). No amplification of *Bra019686* from gDNA was observed, likely due to the large product size

(3668 bp) and primers for this gene only amplified a product from the root cDNA sample. Analysis on *Bra019688* showed products from both gDNA and cDNA at the right size of 231 bp and 147 bp for both leaf and root samples. Only root cDNA amplified a product with the expected size of 207 bp for *Bra019689*, Root and leaf gDNA amplified products for *Bra019690* at 242 bp. No products were detected in leaf and root cDNA for *Bra019690*, with leaf cDNA showing some gDNA contamination.

From microarray data, transcript abundance of *PHO1* in Pi deficient *B. rapa* had a fold change value of 4.51 (Table 4.7), the value was down regulated under Pi deficient conditions compared to Pi replete conditions. In most studies, *PHO1* is up regulated in Pi deficient conditions (Hamburger et al., 2002). There are many reasons why the microarray data doesn't match. The microarray design was based on an earlier version of the genome, so the probe might be a closer match to another member of the *PHO1* family, which is down regulated (e.g. *SPX4*), or the Pi deficiency was not yet sufficient to increase its expression. There are four possible transcripts in this region of the genome, *PHO1_A*, *PHO1_B*, *PHO1_C* and *PHO1_D*. From the RNA-seq data in the databases, there is no transcript evidence of *PHO1_C*. Other highly differentially expressed transcripts in this region of the genome, like *AT1G13300*, *AT1G12200*, and *At1G14560* encodes MYB transcription factor, flavin-containing monooxygenase family protein and mitochondrial substrate carrier family protein, respectively. Only one *PHO1* represented in the microarray (Table 4.7). Given its central role in regulating plant adaptations to low Pi and the identification of four copies in this region, it is still an important target to investigate further.

The *B. rapa* genome sequence revealed four copies of *PHO1* located in tandem on chromosome A06 in the *B. rapa* genome. To get a better understanding of these gene copies and their transcripts, BAC clones, which contain the gene of interest region, were sequenced using MinION (ONT). Using Clone Finder (NCBI

www.ncbi.nlm.nih.gov/clonefinder), five BAC clones, named as KBrB-063F11 (BAC 1), KBrH102C10 (BAC 2), KBrB029J08 (BAC 3), KBrH038K12 (BAC 4), and KBrB003E10 (BAC 5) were identified as spanning the region of interest. All BACs selected contained the four candidate genes region at the eQTL hotspot on chromosome A06 in *B. rapa*. The size of BACs ranged from 134,576 bp (BAC 1) to 155,059 bp (BAC 4) following sequencing.

Table 4.7. Genes whose expression is significantly differentially altered by low Pi availability in *Brassica rapa* and co-locate to an eQTL hotspot on Chromosome A06 (Hammond et al., 2011).

<i>B. rapa</i> chromosome	Physical location (bp)	Microarray Probe ID	Fold change	Up or down regulated	Arabidopsis ID	Description
Chr_A06	4007798	B_1047079	2.10	down	AT1G80740	CMT1 (CHROMOMETHYLASE 1); DNA binding / chromatin binding
Chr_A06	4007798	B_1047079	2.10	down	AT1G80740	CMT1 (CHROMOMETHYLASE 1); DNA binding / chromatin binding
Chr_A06	4109805	B_1063524	2.29	up	AT1G11720	ATSS3 (starch synthase 3)
Chr_A06	4109807	B_1081284	2.34	up	AT1G11720	ATSS3 (starch synthase 3)
Chr_A06	4183860	B_1071971	4.68	up	AT1G13300	myb family transcription factor
Chr_A06	4504093	B_1010249	2.37	up	AT1G12540	basic helix-loop-helix (bHLH) family protein
Chr_A06	4662648	B_1049226	4.68	down	AT1G12200	flavin-containing monooxygenase family protein / FMO family protein
Chr_A06	4727134	X_1058205	2.06	down		
Chr_A06	4803270	X_1038811	2.21	up	AT3G04120	GAPC1 (GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE C SUBUNIT 1)
Chr_A06	4945809	B_1057552	2.49	down		
Chr_A06	4977447	B_1031390	4.51	down	AT1G14040	PHO1
Chr_A06	5133107	B_1078178	2.09	down	AT1G14300	binding
Chr_A06	5133107	B_1078178	2.09	down	AT1G14300	binding
Chr_A06	5133111	B_1072667	2.06	down	AT1G14300	binding
Chr_A06	5133111	B_1072667	2.06	down	AT1G14300	binding
Chr_A06	5135029	B_1010529	2.14	down	AT1G14300	binding
Chr_A06	5135029	B_1010529	2.14	down	AT1G14300	binding
Chr_A06	5296871	B_1013003	5.79	up	AT1G14560	mitochondrial substrate carrier family protein
Chr_A06	5368364	B_1067834	2.64	up	AT1G14710	hydroxyproline-rich glycoprotein family protein
Chr_A06	5369801	B_1065928	2.44	up	AT1G14710	hydroxyproline-rich glycoprotein family protein
Chr_A06	5480815	B_1086630	2.26	down	AT1G14980	CPN10 (CHAPERONIN 10); chaperone binding
Chr_A06	5646032	B_1084467	2.43	down	AT1G15440	transducin family protein / WD-40 repeat family protein
Chr_A06	5672230	X_1033514	2.15	down	AT1G15510	pentatricopeptide (PPR) repeat-containing protein

Table 4.7. continued.

<i>B. rapa</i> chromosome	Physical location (bp)	Microarray Probe ID	Fold change	Up or down regulated	Arabidopsis ID	Description
Chr_A06	5880871	B_1069604	2.30	down	AT1G15870	mitochondrial glycoprotein family protein / MAM33 family protein
Chr_A06	5672230	X_1033514	2.15	down	AT1G15510	pentatricopeptide (PPR) repeat-containing protein
Chr_A06	5747223	B_1066708	2.35	down	AT1G15730	PRLI-interacting factor L, putative
Chr_A06	5892970	B_1090675	2.07	down		
Chr_A06	5893942	B_1072394	2.15	down	AT2G32060	40S ribosomal protein S12 (RPS12C)
Chr_A06	5919506	B_1074644	2.24	up		
Chr_A06	6017873	B_1053879	2.29	up	AT1G16410	CYP79F1 (CYTOCHROME P450 79F1)
Chr_A06	6019599	B_1041769	2.85	up		



Figure 4.7. Sequence alignment of four putative PHO1 in *Brassica rapa* to determine region of homology from the sequences. The high similarity of the sequences is shown in green, yellow is for lower similarity and red refers to very low similarity of the sequences. Alignment was performed using Geneious software. *Brassica rapa* PHO1 IDs are 1. Bra019690, 2. Bra019689, 3. Bra019688, 4. Bra019686.

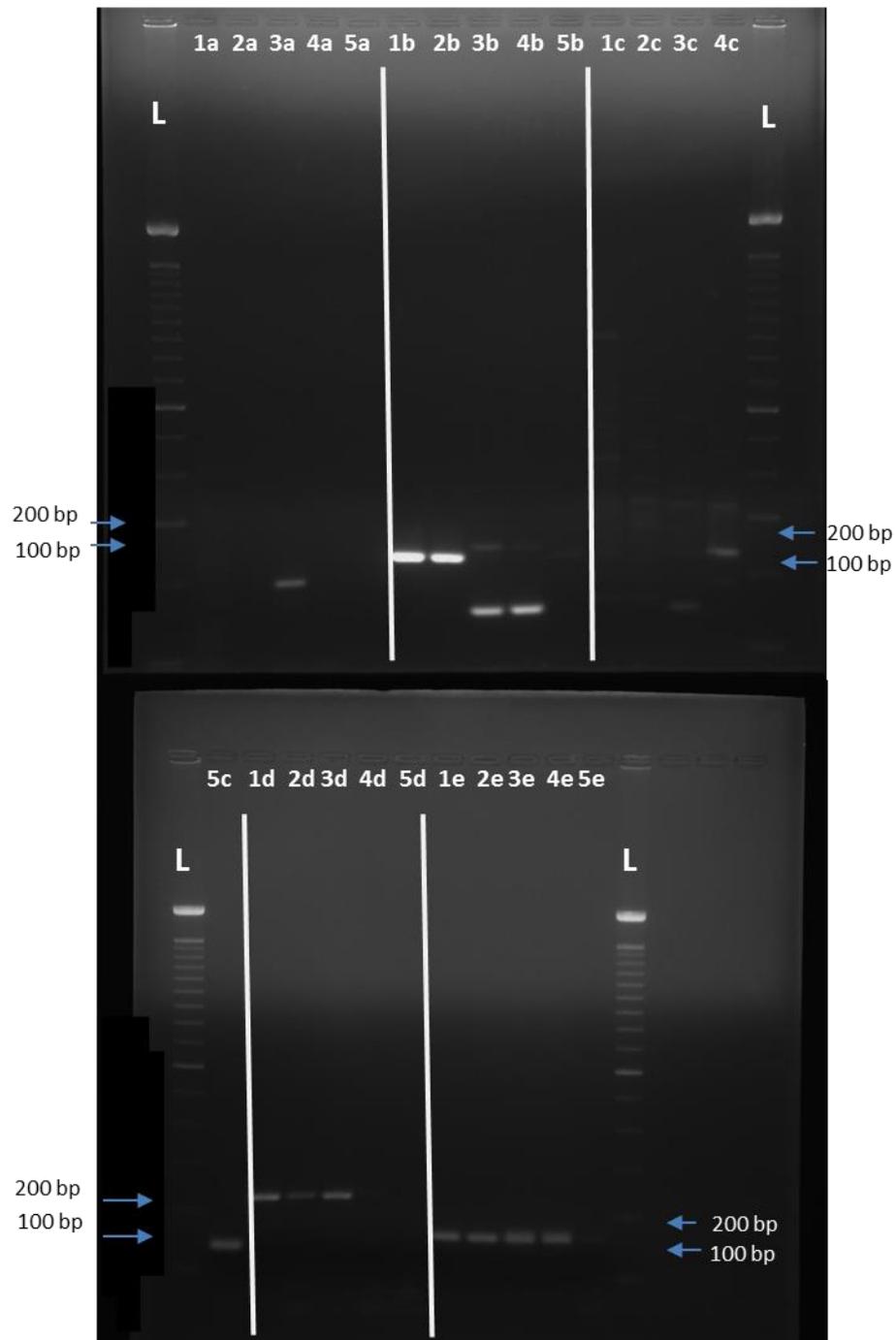


Figure 4.8. Electrophoresis gel of PCR fragments of *Brassica rapa* leaf and root sample. Left and right lane represents a size of 100 bp ladder (L). **1.** Leaf gDNA **2.** Root gDNA **3.** Leaf cDNA **4.** Root cDNA **5.** Non-template control (NTC). **a.** Bra019686 (gDNA= 3668, cDNA=192 bp) **b.** Bra019688 (gDNA=231 bp, cDNA=147 bp) **c.** Bra019689 (gDNA and cDNA have the same size 207 bp) **d.** Bra019690 (gDNA=242 bp, cDNA=152 bp) **e.** Bro18004703 (160 bp) positive control (+ve).

4.6.2 BAC confirmation and validation

All five BACs were verified through reverse transcription PCR (RT-PCR) with the specific primers designed from the ends of each BAC. All the primers amplified their target correctly (Fig. 4.9). Forward (F) and Reverse (R) ends of KBrB-029J08F amplified the target sequence at 226 bp and 169 bp, respectively. In KBrB-029J08R, amplification showed slightly lower amount of product 198 bp. BAC KBrB-063F11, BAC KBrH-038K12 and BAC KBrH-102C10 showed good amplification of their target sequence at their F and R ends.

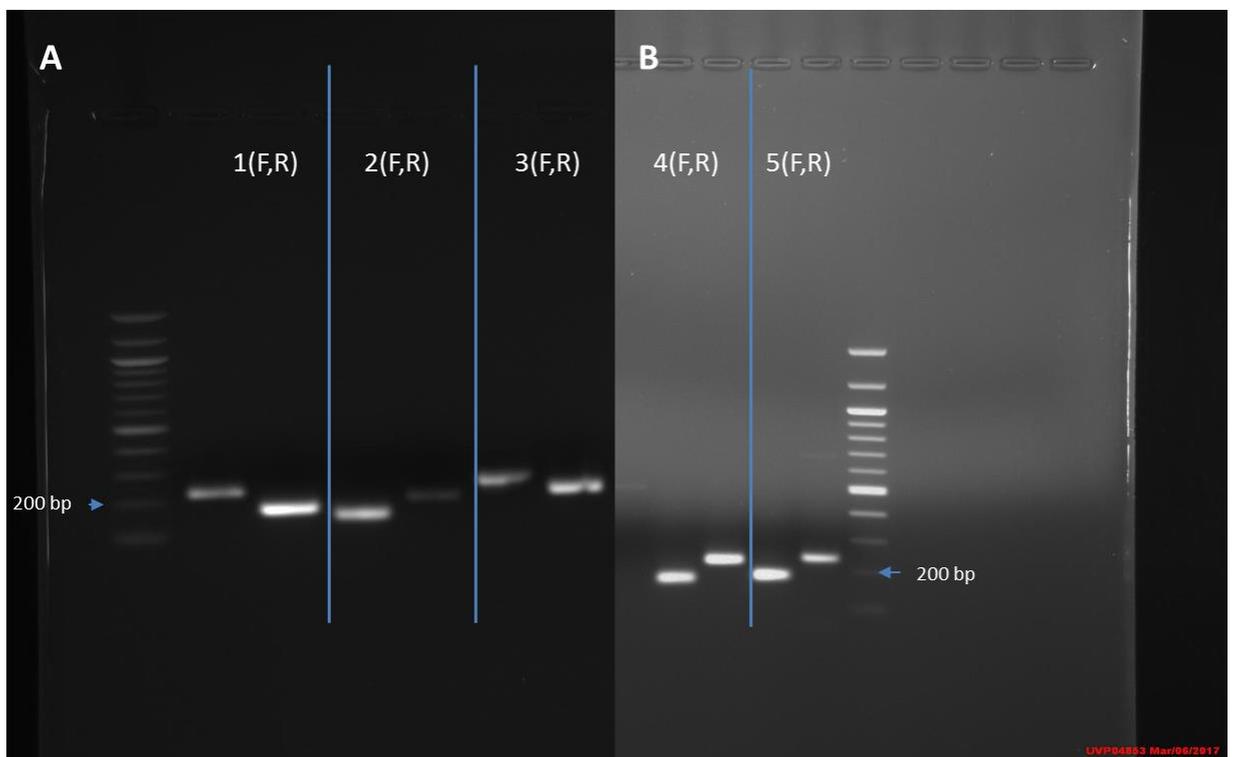


Figure 4.9. Validation of BAC DNA using PCR with BAC end specific primers. PCR products were run on 1.2% agarose gel electrophoresis. The front (F) and rear (R) BAC ends were verified for each BAC and their expected sequence lengths are given in the parentheses. **A.** 1F: KBrB-029J08F (226 bp), 1R: KBrB-029J08R (169 bp), 2F: KBrB-003E10F (150 bp), 2R: KBrB-003E10R (198 bp), 3F: KBrB-063F11F (244 bp), 3R: KBrB-063F11R (226 kb) **B.** 4F: KBrH-038K12F (177 bp), 4R: KBrH-038K12R (228 bp), 5F: KBrH-102C10F (185 bp), 5R: KBrH-102C10R (233 bp).

4.6.3 BAC extraction using large construct DNA extraction kit

Single colonies were produced by streaking the bacteria on agar plate containing $12.5 \mu\text{g mL}^{-1}$ chloramphenicol (Fig. 4.10). The starter culture was prepared in 500 L LB media for BAC DNA extraction using a Large Construct DNA extraction kit (Qiagen). The quantification of BAC DNA using Nanodrop 2000 spectrophotometer and Qubit fluorometer showed the high quality of DNA where the ratios of $\text{Abs}_{260/280}$ and $\text{Abs}_{260/A280}$ are above 1.8 (Table 4.8). The total yield after DNA extraction is shown in Table 2.3. The highest total DNA was recorded for BAC 2 (KBrB-3E10) with $11.6 \mu\text{g } \mu\text{L}^{-1}$, followed by BAC 5 (KBrH-102C10), BAC 1 (KBrB-029J08), BAC 3 (KBrB-63F11) with 9.55, 8.94 and $2.67 \mu\text{g } \mu\text{L}^{-1}$, respectively. The lowest was recorded in BAC 4 (KBrH-38K12) with $2.02 \mu\text{g } \mu\text{L}^{-1}$. It shows that, after DNA precipitation was carried out, the total DNA was reduced by about 25-40%. $1 \mu\text{g}$ in $45 \mu\text{L}$ of each BAC was used as the initial BAC DNA concentration for nanopore sequencing by MinION (Table 4.9).

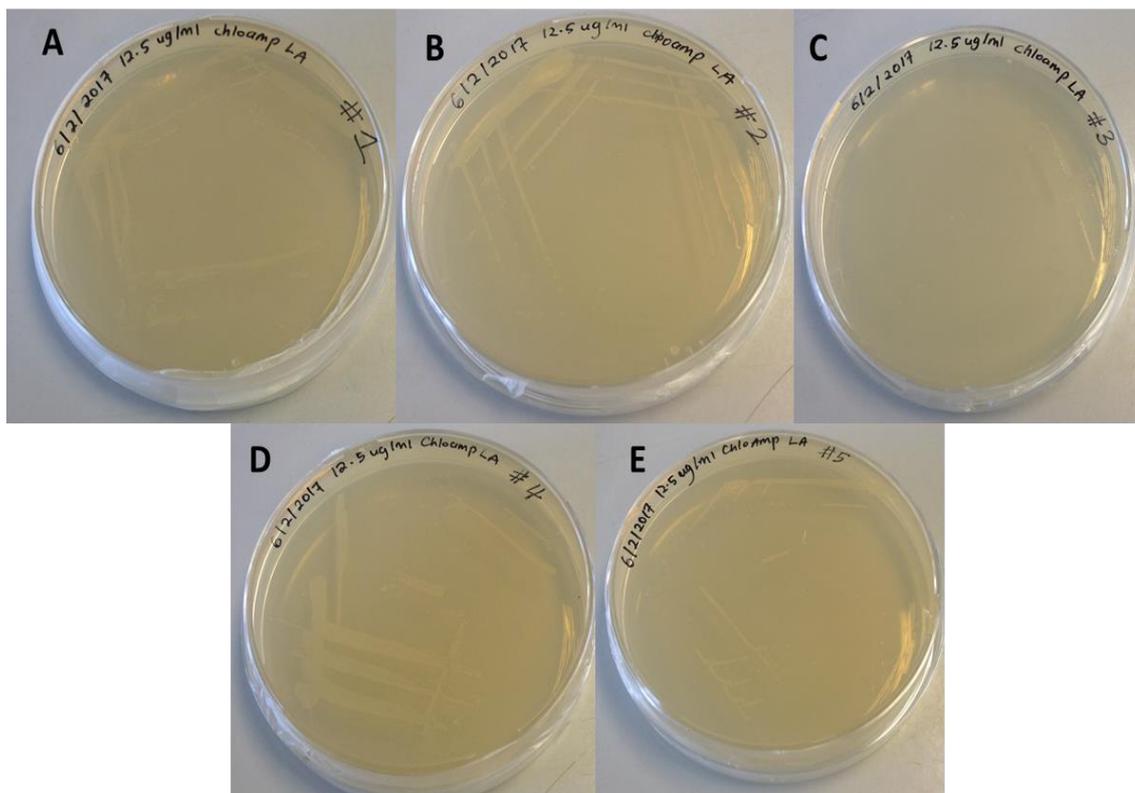


Figure 4.10. BAC DNA single colony on agar plates with $12.5 \mu\text{g } \mu\text{L}^{-1}$ chloramphenicol after incubation at 37°C overnight. From top left; **A.** KBrB-063F11, **B.** KBrH102C10 **C.** KBrB029J08 **D.** KBrH038K12 **E.** KBrB003E10.

Table 4.8. BAC DNA extraction using large construct DNA extraction kit (Qiagen) quantification using Nanodrop 2000 and Qubit fluorometer. Total DNA was calculated based on Qubit reading and total volume of sample.

ID	Sample	Nanodrop ($\text{ng } \mu\text{L}^{-1}$)	$A_{260/280}$	$A_{260/230}$	Qubit ($\text{ng } \mu\text{L}^{-1}$)	Total Volume of Sample (mL)	Total DNA ($\mu\text{g } \mu\text{L}^{-1}$)
1	KBrB-63F11	8.5	2.19	12.68	5.34	500	2.67
2	KBrH-102C10	29.8	1.82	2.67	19.1	500	9.55
3	KBrB-029J08	13.1	2.02	9.71	8.94	1000	8.94
4	KBrH-38K12	4.9	2.11	12.09	4.04	500	2.02
5	KBrB-3E10	20.3	1.89	3.88	11.6	1000	11.6

Table 4.9. Total BAC DNA yield after DNA extraction and amount of nuclease free water (NFW) added to the stock DNA after DNA precipitation to prepare 1 µg in 45 µL for MinION sequencing.

BAC ID	Nanodrop (ng µL⁻¹)	A 260/280	A 260/230	Qubit (ng µL⁻¹)	Stock Vol. (µL)	Total DNA (µg)	Total DNA (µg µL⁻¹)	Stock DNA used (µL)	NFW added (µL)
KBrB-63F11	52.4	1.80	2.11	39.2	50	1.96	0.040	25.00	20.0
KBrH-102C10	267.1	1.77	2.02	136	50	6.80	0.139	7.21	37.8
KBrB-029J08	156.1	1.86	2.45	100	50	5.00	0.102	9.80	35.2
KBrH-38K12	57.0	1.54	1.37	28.6	50	1.43	0.029	34.27	10.7
KBrB-3E10	241.1	1.86	2.53	171	50	8.55	0.175	5.73	39.3

4.6.4 BAC DNA sequencing using Oxford Nanopore Technologies MinION

4.6.4.1 MinION sequencing: Data analysis

A total of 991,846 sequences were sequenced, generating approximately 6.4 Gb total bases sequenced with an average sequence length of 6,446 bp. The number of sequences produced from quality sequencing summary was 860,433 with about 5.7 Gb of total bases sequenced. The average sequence length was 6,642 bp and the longest read was 97,111 bp.

During MinION sequencing the plasmid sequence was also sequenced. To obtain the full BAC sequence, the plasmid sequence was removed using CLC and CLC alignments were used to align the plasmid backbone and BAC ends (Fig. 4.11). Comparison of the predicted size of BAC (from NCBI clone finder) with the actual size from MinION sequencing showed some differences (Table 4.10). BAC 1 (KBrB-63F11) showed an

increase in size, from 134,576 bp to 184,651 bp. In BAC 2 (KBrH-102C10) the actual size from MinION sequencing was shown to be slightly higher (1 Kb) from the predicted size of 152,113 bp. BAC 3 (KBrB-029J08) and BAC 4 (KBrH-38K12) showed an increase in the actual size to 185,935 bp and 165,146 bp, respectively. BAC 5 (KBrB-3E10) showed a shorter BAC length compared to the predicted size, recording only 45,485 bp. This might be attributed to the MinION device stopping during the 48h run. Consequently, the incomplete sequence of BAC 5 was not used in further analysis. Interestingly, from the five BAC's sequenced and analysed, there was only a consensus between BAC 1 (KBrB-63F11) and BAC 3 (KBrB-029J08). Both of these BAC's are from the same library KBrB and have the same BAC ends. For this reason, only these BAC's were chosen for further studies. After aligning all our four target genes (*PHO1_A*, *PHO1_B*, *PHO1_C* and *PHO1_D*) to each BAC, the results showed all four target genes were present and located in tandem (Fig. 4.12).

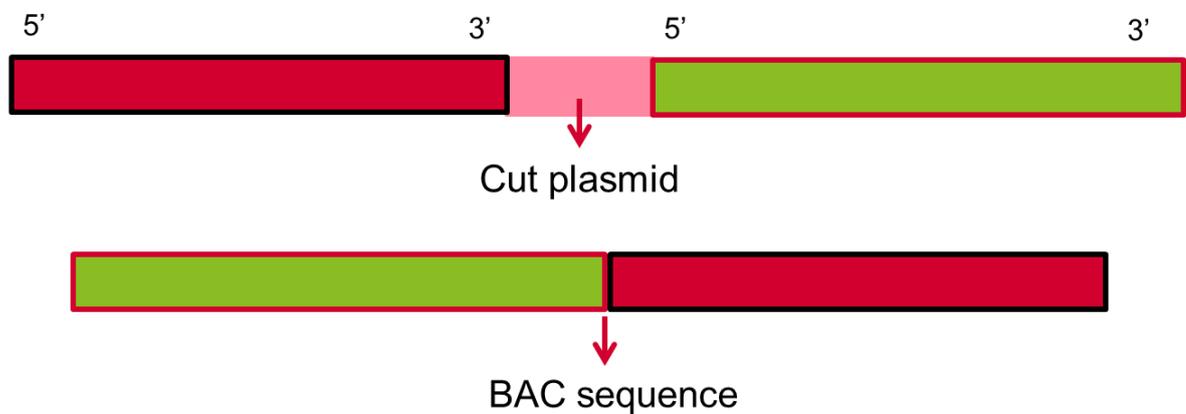


Figure 4.11. Illustration of the plasmid sequenced from the MinION and which was cut to get the actual BAC sequence.

Table 4.10. Comparison of predicted and sequenced BAC sizes.

ID	Sample (BAC)	Predicted BAC size (Kb)	BAC length (Kb)
1	KBrB-63F11	134,576	184,651
2	KBrH-102C10	152,113	153,335
3	KBrB-029J08	148,972	185,935
4	KBrH-38K12	155,059	165,146
5	KBrB-3E10	136,929	45,485

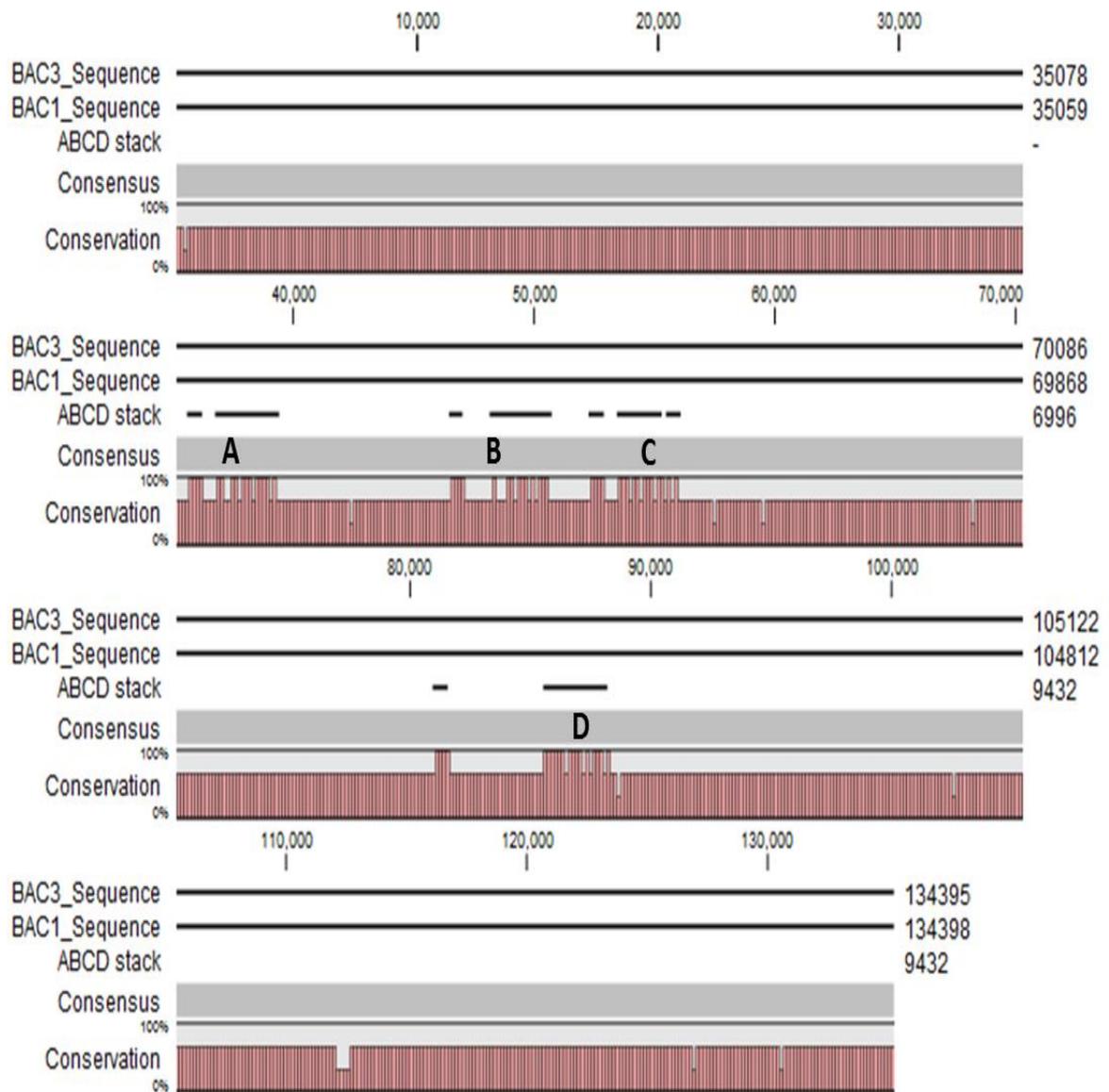


Figure 4.12. The locations of *PHO1* homolog genes on CLC Sequence Viewer 7.8.1 (Qiagen). **A.** (*PHO1_A*), **B.** (*PHO1_B*), **C.** (*PHO1_C*) and **D.** (*PHO1_D*) on BAC 1 and BAC 3 sequence.

4.6.5 Identification and characterisation of *PHO1* homolog transcripts

4.6.5.1 Reverse transcription polymerase chain reaction (RT-PCR)

Total RNA from leaf tissues of *B. rapa* was transcribed into cDNA using a kit capable of synthesising long amplicons. A gradient RT-PCR was conducted with both primer sets at 60 and 65 °C annealing temperature using cDNA from P+ and P- treatments for optimisation. Both primer sets amplified multiple bands in expected size ranges at 60 °C (Fig. 4.13). However, primer set PHO_1_F and PHO_2224_R produced brighter and distinct bands. Fig. 4.13 A). A band of unexpected size of around 4 Kb was also amplified that might have been originated from alternative splicing of *PHO1* or be an entirely non-related transcript.

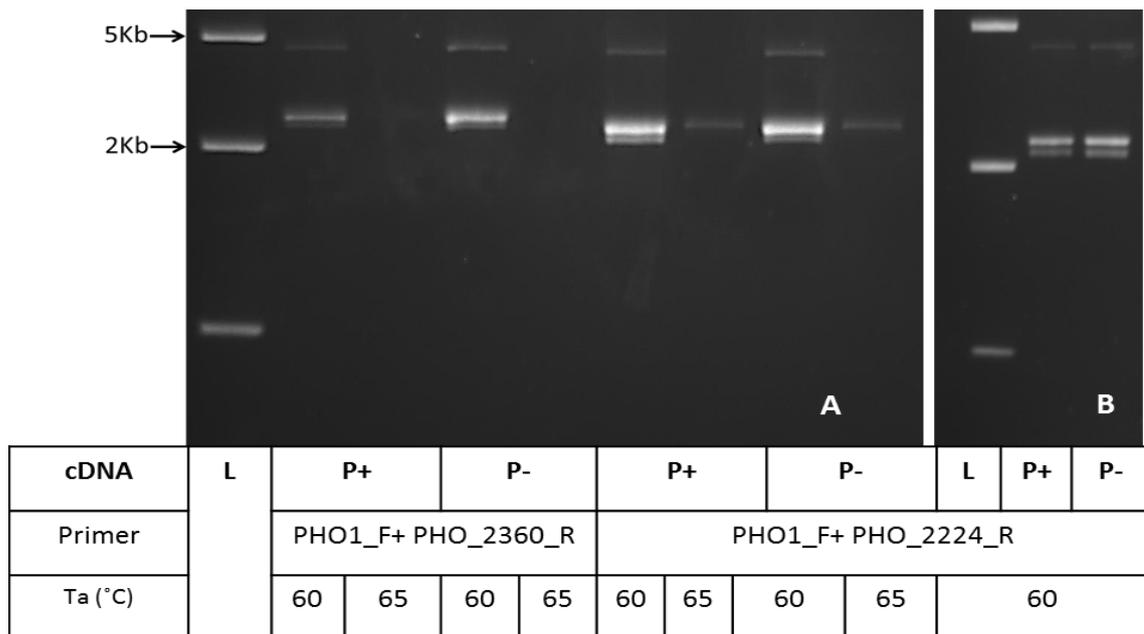


Figure 4.13. Gel image of *PHO1* Reverse Transcription Polymerase Chain Reaction (RT-PCR) of cDNA from *Brassica rapa* R-o-18 leaf samples. **A.** Lane 1 from left: FastRuler Middle Range DNA Ladder (L), lane 2 to 9 is RT-PCR for cloning of *PHO1* transcripts. **B.** Gel of an aliquot of the RT-PCR products after PCR optimisation targeting PHO1_F+PHO_2224_R.

4.6.5.2 Cloning *PHO1* transcripts

Two PCR reactions were performed using P+ and P- cDNA and PHO1_F+ PHO_2224_R using the optimised PCR protocol. An aliquot of the RT-PCR products was run on gel which verified the sizes found in previous PCR (Fig. 4.13). Both RT-PCR products were pooled for representation of both treatments, purified and ligated into pCR4 BLUNT TOPO vector. Topo10 *E. coli* cells were transformed with the ligation mixture and showed good growth (Fig. 4.14).

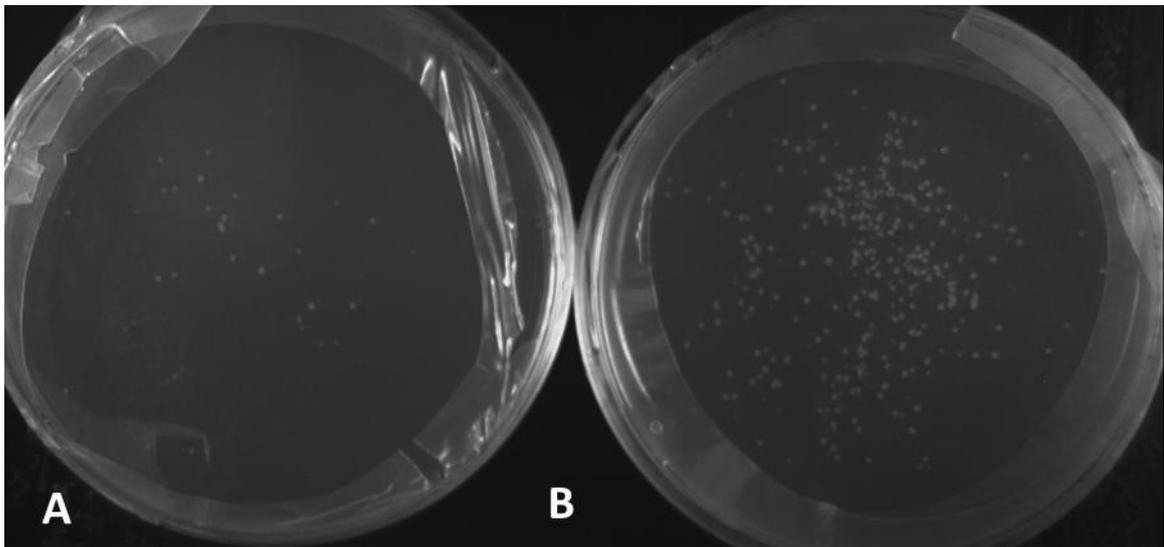


Figure 4.14. A. LB agar with 50 mg L⁻¹ of Kanamycin plate containing 20 µL transformed culture. B. LB agar with 50 mg L⁻¹ of Kanamycin plate containing 200 µL transformed culture.

4.6.5.3 Colony PCR and restriction/ digestion analysis

Thirty colonies were picked and analysed using colony PCR. The gel image of the colony PCR products revealed several size variants (Fig. 4.15). For precise size analysis, colony PCR products were subjected to HindIII restriction enzyme digestion to differentiate among the four *PHO1* homologs (Table 4.11). The colony PCR restriction analysis verified the size variants found in colony PCR (Fig. 4.15). The variant type 1 (colony 1-6, 8-11, 13,

15, 17-21, 25, 26-30) was prevalent among the colonies analysed. Other variants included, type 2 (7 and 14), 12, 16, 22, 23, and 24. Selected colonies were cultured overnight, plasmid DNA was isolated and subjected to restriction confirmation and subsequent Sanger sequencing.

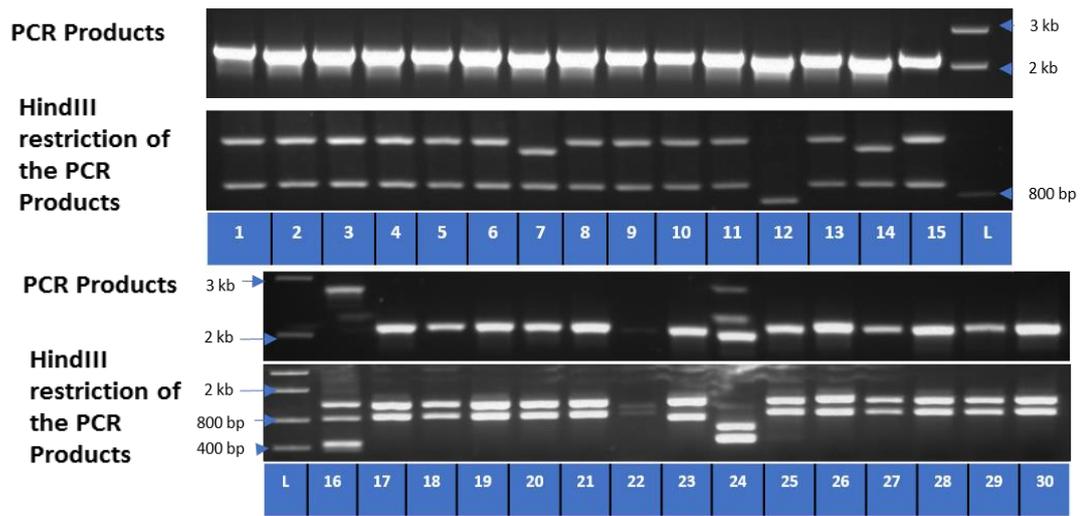


Figure 4.15. Gel image of colony PCR and HindIII restriction analysis of colony PCR of 30 randomly selected colonies containing PHO1 pCR4 BLUNT TOPO vector. Band <80 bp are not shown.

Plasmid DNA was isolated from ten clones on the basis of colony PCR analysis. Restriction maps of the targeted region of XM_009150437 (*PHO1_A*), XM_018652610 (*PHO1_B*), XM_018652651 (*PHO1_C*) and XM_009150438 (*PHO1_D*) in pCR4 BLUNT TOPO vector were constructed using Vector NTI software (Fig. 4.16). Restriction of 1 µg plasmid DNA of the ten selected clones confirmed the findings of the colony PCR (Data not shown), hence these were sent for sequencing.

Table 4.11. Transcript length expected cut sites and fragment length of *PHO1* homologs restricted with HindIII enzymes.

Gene	Transcript length (bp)	Cut sites	Fragment Size (bp)
<i>PHO1_A</i>	2224	50, 970	50, 920, 1254
<i>PHO1_B</i>	2044	80, 1593, 1677	80, 84 367, 1513,
<i>PHO1_C</i>	2152	46, 966	46, 920, 1186
<i>PHO1_D</i>	2227	50, 479, 938, 961, 1772	23, 50, 429, 455, 459, 805

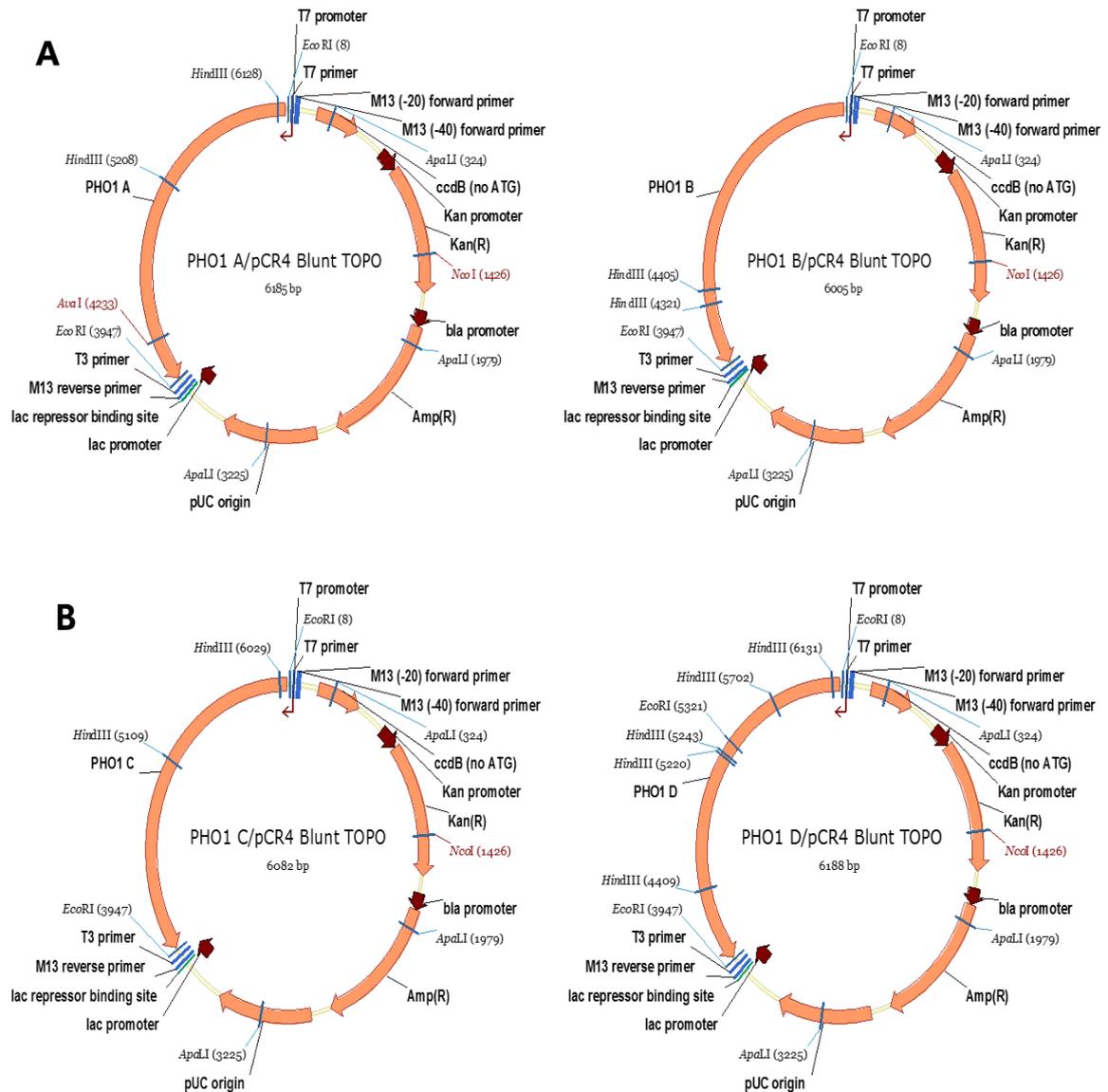


Figure 4.16. Restriction maps of targeted region of **A.** XM_009150437 (*PHO1_A*) and XM_018652610 (*PHO1_B*), **B.** XM_018652651 (*PHO1_C*) and XM_009150438 (*PHO1_D*) = pCR4 BLUNT TOPO sequencing vector.

Plasmid numbers 2, 7, 12, 16, 22, 23, 24 and 26 were sequenced from both ends using M13F and M13R primers located in the vector upstream and downstream of the insert. The results showed good quality single read long enough to extract full sequence. However, sequencing of both ends of plasmid 16, which is around 4 kb on gel, could not produce a contig. Therefore, two internal primers were designed for complete coverage of plasmid 16. Vector sequence was trimmed and the sequences of all eight transcripts (transcript of plasmid) can be viewed Appendix 1.

BLASTN analysis (using *B. rapa* cultivar Chiifu 401-402 nucleotide sequence) revealed that transcript 2 had 96% similarity with *B. rapa* phosphate transporter PHO1 homolog 3 (XM_009150437). Both sequences differed for 94 SNPs and two indels. The transcript 2 sequence was 2236 bp long and was predicted to encode a peptide 744 aa long. BLASTP analysis of the protein encoded by transcript 2 found 98% identity with *B. rapa* phosphate transporter PHO1 homolog 3 isoform X1 (XP_009148685) (Table 4.12).

BLASTN search (using *B. napus*, cultivar ZS11 nucleotide sequence) for 2092 nucleotide long sequence of transcript 12 revealed 99% identity with *B. napus* phosphate transporter PHO1 homolog 3-like mRNA (XM_013785900) which differed by just 17 SNPs. Phosphate transporter PHO1 homolog 3 transcripts variant X3 mRNA (XM_018652610.1) emerged as closest hit in *B. rapa* having 80% coverage and 90% homology. The transcript 12 sequence may encode a 697 aa long peptide which had very high conservation of 99% with *B. napus* phosphate transporter PHO1 homolog 3-like (XP_013641354). However, surprisingly had just 66% homology in *B. rapa* with phosphate transporter PHO1 homolog 3 isoform X2 (XP_018508125).

Transcript 24 was found to be 2042 bases long. BLASTN search identified 97% similarity of 100% coverage with *B. rapa* PHO1 homolog 3-like mRNA (XM_009150438). However, this sequence only encodes a protein of 516 amino acid (aa) compared to 811 aa protein

(XP_009148686) encoded by XM_009150438. Multiple alignment of both protein sequences showed that first 415 aa shared 97% similarity. Thorough analysis of exon to exon alignment showed that part of the exon 7 of XM_009150438 is missing in transcript 7, which implies that the alternative splicing may have caused a truncated protein (data not shown). However, on the basis of highly significant similarity of DNA and protein sequence found between XM_009150438 and transcript 24, it was concluded that transcript 24 is *PHO1_D* homolog.

Multiple alignments of eight selected transcript sequences portray an interesting picture. Transcript 2 which proved to be *PHO1_A/C* has 4 other splice variants including transcripts 7, 22, 23 and 26 which are 2150, 2073, 2130 and 2317 bp long, respectively and the alignment is shown in Fig. 4.17. Alternative splicing resulted in a truncated protein of 552, 490 and 629 aa encoded by transcript 7, 22 and 23, respectively. Transcript 26 was 2317 bp long and may encode a 746 aa long protein. All the sequences produced can be seen in Appendix 1.

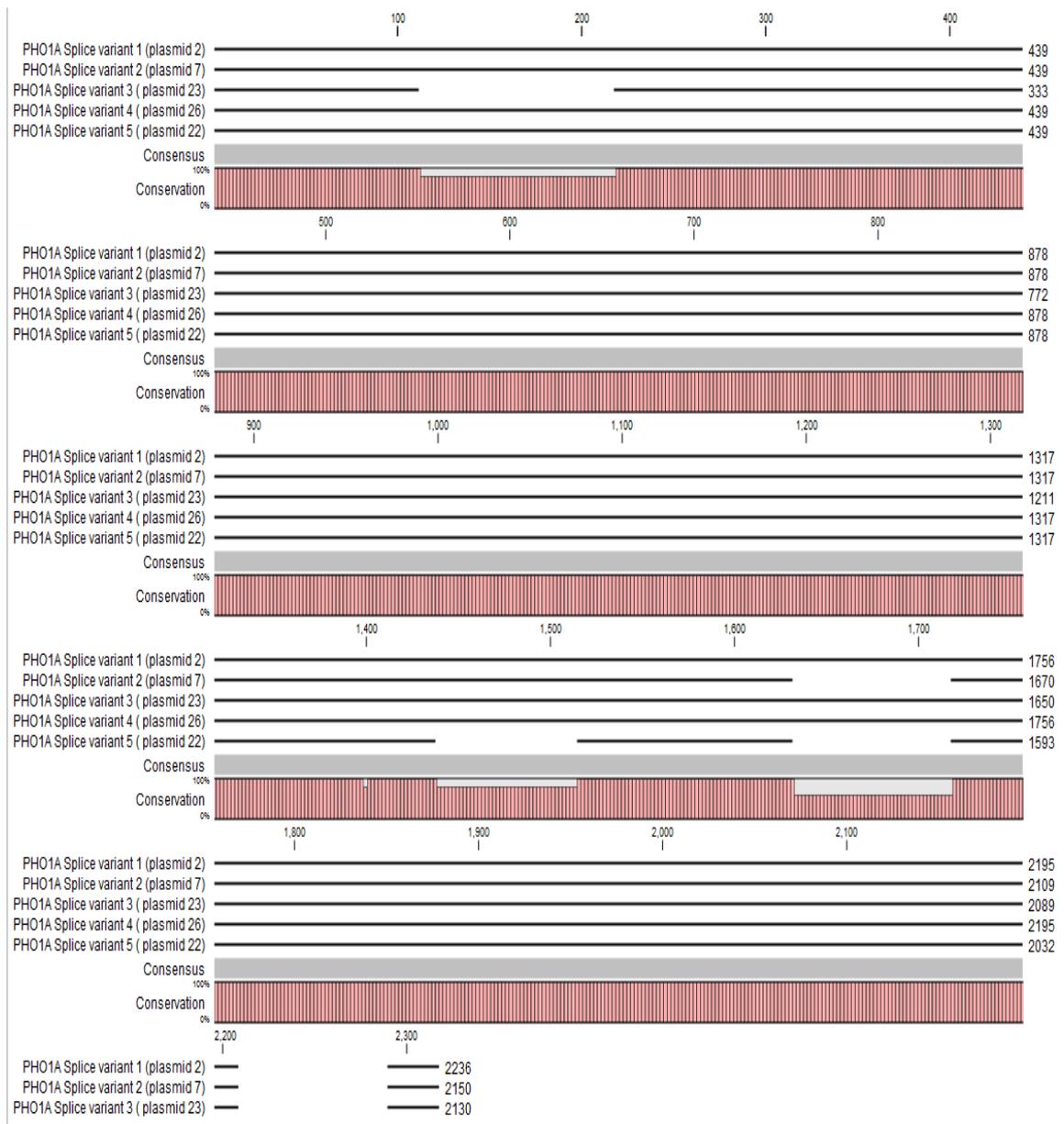


Figure 4.17. Alignment of five *PHO1_A* splice variants in CLC Sequence Viewer 7.8.1 (Qiagen). Row from top: splice variants 1,2,3,4,5 for plasmid 2,7,23,26 and 22, respectively.

Table 4.12. Plasmid sequenced with the transcript number, nucleotide length, peptide length, closest BLASTN and BLASTP hits, and predicted PHO1 on chromosome A06 *Brassica rapa*.

Transcript no. From NCBI	Plasmid	Nucleotide length	Peptide length	Closest BLASTN hits	Closest BLASTP hits	Predicted <i>PHO1</i>
XM_009150437.2	2	2236	744	XM_009150437.2	XP_009148685.1	A
XM_009150437.2	7	2150	552	XM_009150437.2	XP_018508167.1	A
XM_009150437.2	23	2130	629	XM_009150437.2	XP_009148685.1	A
XM_009150437.2	26	2317	746	XM_009150437.2	XP_009148685.1	A
XM_009150437.2	22	2073	490	XM_009150437.2	XP_018508167.1	A
XM_018652610	12	2092	697	XM_018652610	XP_018508125	B
XM_009150438	24	2042	516	XM_009150438	XP_009148686	D
XM_009150438	16	513	170	XM_009150438	-	-

Conserved domain search in the *PHO1_A/C* splice variants identified SPX (pfam03105) and EXS (pfam03124) domains in proteins encoded by transcript 2, 23 and 26, however, EXS (pfam03124) is absent in 7 and 22. Both domains were present in *PHO1_B* (Transcript 12) whereas *PHO1_D* (Transcript 24) truncated protein contained SPX (pfam03105) domain only.

At least three groups of *PHO1* paralogs expressed in leaf tissue of *B. rapa* in P+ and P- conditions have been cloned and sequenced. Transcript 2 along with four other splice variants, transcript 7, 23, 22 and 26 may be designated as *PHO1_A*. *PHO1_A* (XM_009150437) and *PHO1_C* (XM_018652651) are almost identical sequences except for seven SNPs. However, *PHO1_C* contains an open reading frame of 2025 bp, compared to *PHO1_A* which is 2445 bp long. As all five transcripts designated as *PHO1_A* ranged in size from 2130 to 2317 without any SNP, the presence of *PHO1_C* transcripts, at least in *B. rapa* leaf cDNA library, can be ruled out. This is consistent with previous evidence (Fig 4.7).

The prevalence of *PHO1_A* full length transcripts along with a number of splice variants, as determined by colony PCR and restriction analysis of 30 colonies and sequencing of eight selected plasmids from leaf samples, implies a role of *PHO1_A* in P regulation. The splice variants may have a role in gene regulation rather than encoding a protein based on some of the splice variants not encoding full proteins.

Strong homology of transcript 12 in *B. napus* and *B. oleracea*, and poor coverage and identity in *B. rapa* genome is surprising. It may be attributed to poor efficiency of short read sequencing in annotation of duplicated genes with considerable homology.

As mentioned before transcript 24 has been attributed *PHO1_D*. However, the full length could not be captured. The transcript may be in low abundance in leaf. Screening of cDNA library from other tissues like roots may confirm this hypothesis.

The final alignment and phylogenetic tree were generated and clearly show the relationship of all related transcripts to the target genes (Fig. 4.18 and Fig.4.19).

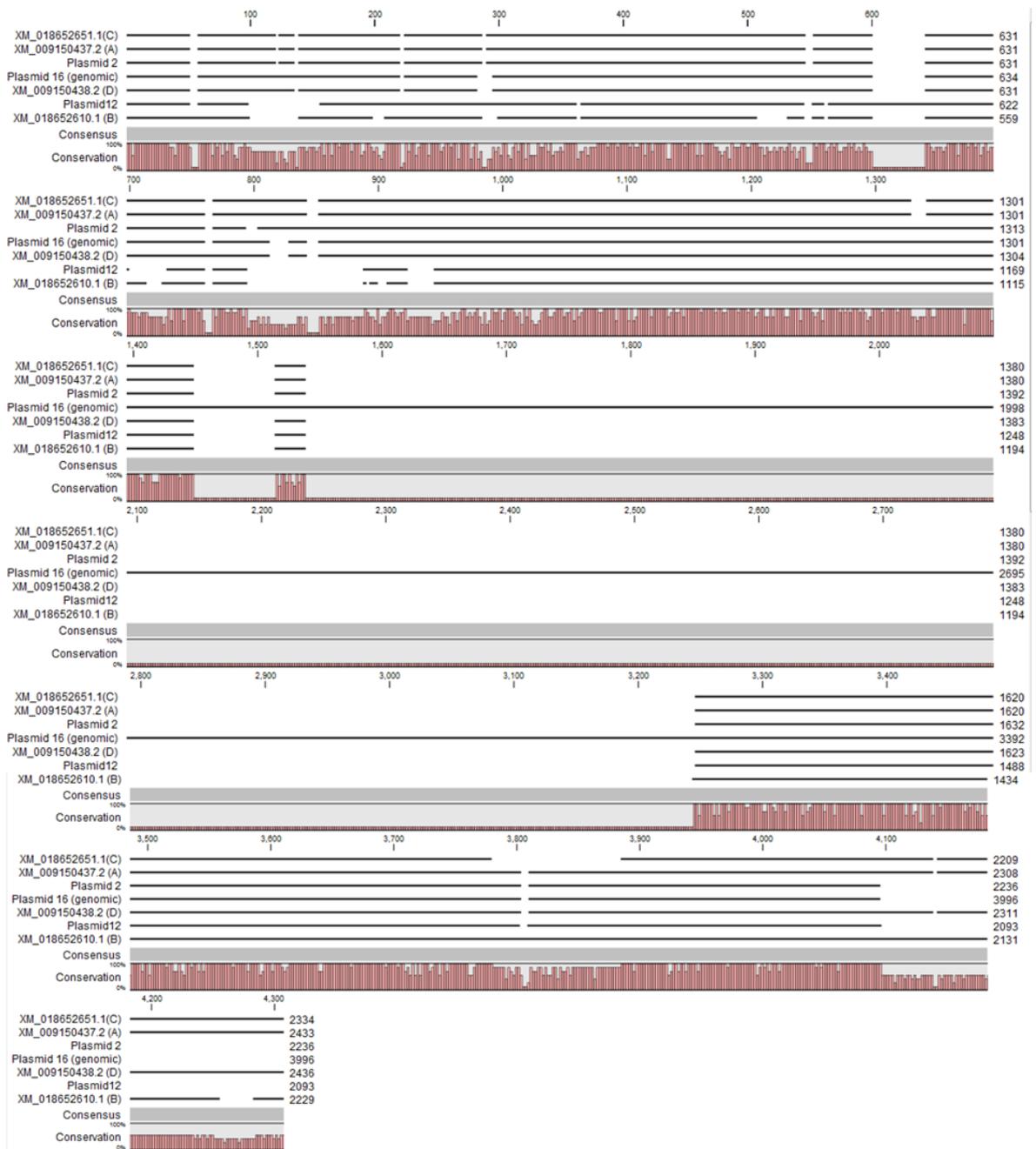


Figure 4.18. Final alignment of *PHO1* paralog genes in CLC Sequence Viewer 7.8.1 (Qiagen). Row from top: XM_018652651.1 (*PHO1_C*), XM_009150437.2 (*PHO1_A*), Plasmid 2, Plasmid 12, and XM_018652610.1 (*PHO1_B*).

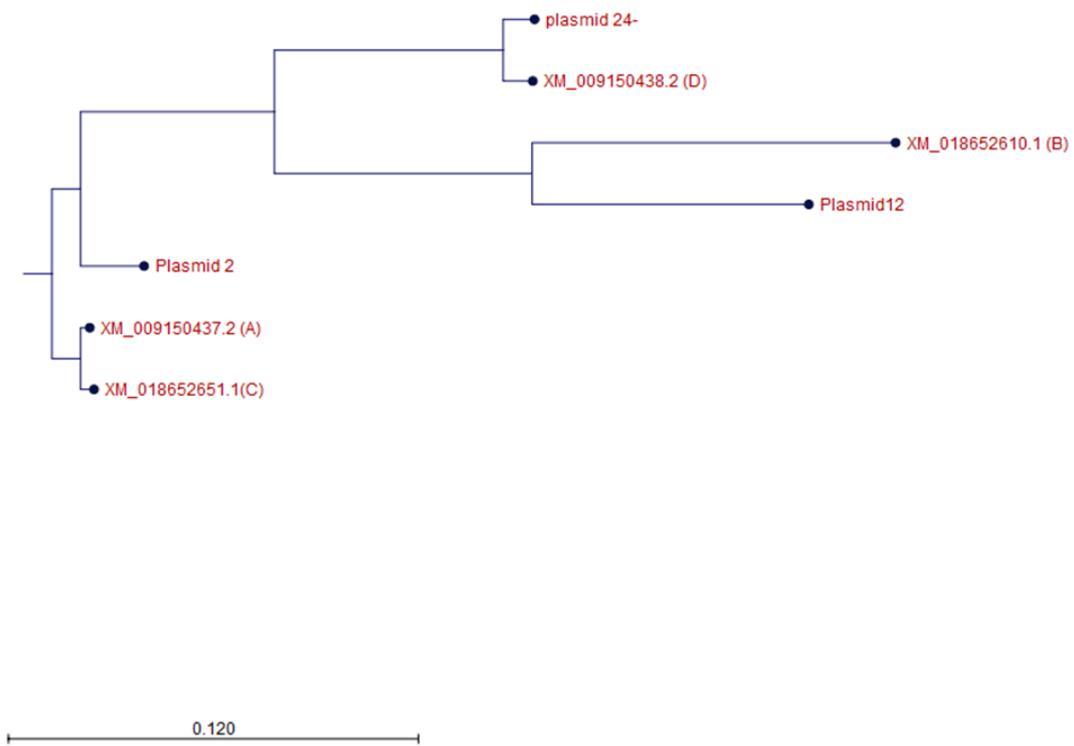


Figure 4.19. Homology tree of *Brassica rapa* PHO1 plasmids and NCBI paralogs produced from CLC Sequence Viewer 7.8.1 (Qiagen).

4.7 DISCUSSION

Analysis of *B. rapa* *PHO1* genes located on chromosome A06 in this study showed strongest homolog to *PHO1:H3* (AT1G14040) in Arabidopsis (Tair, www.Arabidopsis.org/info). Four *PHO1* paralogs were identified located in tandem on chromosome A06 of *B. rapa* (Fig. 4.7). The *de novo* assembly of BAC sequences from this region using long read MinION sequencing technology confirmed the presence of these four paralog copies, namely *PHO1_A*, *PHO1_B*, *PHO1_C* and *PHO1_D* (Fig. 4.12). Analysis of colony PCR and restriction analysis of 30 colonies, together with sequencing of eight selected plasmids showed only three transcripts present in leaf samples of *B. rapa* R-o-18 Pi-deficient plants; *PHO1_A*, *PHO1_B*, and *PHO1_D*. Expression evidence could not be done in this experiment due to the highly similarity of the sequences preventing the design of paralog specific primers for qPCR analysis, as well as time constraints.

4.7.1 Potential for alternatively spliced *PHO1* transcripts

From the results, *PHO1_A* appeared to have five splice variants (Fig. 4.17). The post-transcriptional changes of *PHO1_A* produced potential alternative splicing or structural divergence which could play a significant role in the Pi-gene regulation as well as contribute to protein diversity. Further work is required to validate these findings, including identification of alternatively spliced transcripts and proteins. In addition, the presence of these four paralogs are the product of gene duplication, where initially they would have similar sequences and functions, but eventually differ in their regulatory and coding regions (Xu et al., 2012).

Three process have been identified that lead to alternative splicing or structural divergence based on the intron-exon structures present in the genes. First, intron-exon gain/loss; the process where entire/partial sections of introns/exons were obtained by duplication of the intron and exon or by reshuffling of intron and/or exon. Second, exonisation and

pseudoexonisation; these processes involved interchanges between exonic and nonexonic sequences. Third, intraexonic insertion or deletion; this may shift the open reading frame and biochemical function (Xu et al.,2012). Comparison of the genomic sequence with five paralog transcripts of *PHO1_A* indicated that the differences in intron-exon structure happened due to exonisation/pseudoexonisation, with the changes occurring at nucleotide 3494 of genomic sequence aligned with sequence of plasmid 26 (data not shown).

Alternative splicing occurred due to nutrient or particularly Pi-stress, consistent with other studies suggesting alternative splicing is necessary for plants to adapt environmental stress (Shang et al., 2017). For example, in rice, 33% of all rice genes are alternatively spliced with 58% of these alternatively spliced genes experiencing multiple alternative splicing events, thus producing variety of transcripts (Zhang et al., 2010).

4.7.2 Potential structural variation in PHO1 paralogs

Analysis of the *B. rapa* PHO1 paralog structures showed the proteins comprise two main domains and share the same features found in all members of the PHO1 family in Arabidopsis. The first half of the protein contains an SPX domain harbouring hydrophobic N-terminal and the second half of the protein contains an EXS domain harbouring the hydrophilic C-terminal. Compared to other eukaryotes, PHO1 family members are the only proteins which contain both SPX-EXS domains (Wang et al., 2004). SPX-EXS proteins are involved in many biological processes such as Pi uptake, transport and storage to maintain Pi homeostasis (Hamburger, 2002, Liu et al., 2018). In this study, evidence of different proteins being produced from a single gene was observed.

Analysis of PHO1_A protein sequences using BLASTP revealed five potential splice variants that resulted in altered domain structures. Splice variant (sv) 1, sv 2, sv 3, sv 4,

and sv 5 encode proteins of 744, 552, 629, 746, and 490 amino acids, respectively. Three proteins contain both SPX-EXS domains, while two others (sv 2 and sv 5) contain only an SPX domain (pfam03105) and lack an EXS domain (pfam03124) in the C-terminus. In Arabidopsis, four proteins were identified that only contain a SPX domain, namely SPX1-SPX4 (Duan et al., 2008). In rice, six proteins have been identified that only contain a SPX domain (OsSPX1-OsSPX6) (Secco et al., 2012). Analysis of PHO1_B showed both SPX and EXS domains were present at both protein terminal, while PHO1_D did not encode a full protein, lacking the EXS domain in the C-terminal. The variety of domain harbouring proteins gives differences in the structure, and therefore potentially the expression, function and subcellular localisation of these proteins (Du et al., 2017).

SPX domains are known to be involved in Pi metabolism and could be found in a variety of evolutionary unrelated proteins (Secco et al., 2012). Proteins which exclusively harbour SPX-domains are called SPX proteins. The majority of SPX genes, except AtSPX4 and OsSPX4, are phosphate starvation induced (PSI), and are important for signal transduction of inorganic Pi status in plants and strongly up-regulated under Pi starvation (Duan et al., 2008). SPX acts as a sensor to activate the activity of PHO1 signal to mediate the Pi export to the xylem. Previous studies showed that the SPX domain was not essential for Pi export activity, but important in binding Pi to its target sequence (Puga et al., 2014; Wege et al., 2016). In Arabidopsis, AtSPX1 and AtSPX2 are localised in the nucleus and bind with PHR1 transcription factor depending on tissue Pi status, thus affecting PHR1 binding to its target sequence. Similar results were obtained with the rice ortholog OsPHR2, confirming that the SPX domain may act as a Pi sensor and activate the activity of PHO1 as Pi exporter or Pi signal transduction pathway (Rouached et al., 2011; Wang et al., 2014).

EXS domains have several functions; and is not only important in Pi export activity but have been shown to facilitate long-distance signals from root to shoot to support growth and localise to the Golgi and *trans*-Golgi network in *Nicotiana benthamiana* (Wege et al., 2016). *Pho1* mutant expressing EXS domain showed significant improvement of shoot growth, but reduced expression of many genes associated with Pi deficiency (Wege et al., 2016). The genes involved included genes encoding enzymes which are involved in improving Pi acquisition, Pi recycling and conservation including purple acid phosphatase *PAP5* (APase activity), monogalactosyldiacylglycerol synthase *MGD3* (lipid modification) and *SPX1* and *SPX3* (signalling cascade) (Hammond et al., 2003; Misson et al., 2005, Morcuende et al., 2007; Wege et al., 2016). The EXS domain of PHO1 is important for Pi export as well as playing a role in modulating Pi response through a long-distance signalling from root to shoot (Arpat et al., 2012; Wege et al., 2016). The EXS domain of PHO1 is also essential for proper localisation to the Golgi and *trans*-Golgi network although EXS alone could not function to export Pi (Wege et al., 2016).

4.7.3 Conclusion

PHO1 is predominantly and overly expressed in the root and most of PHO1 homologs are expressed in vascular systems of various tissues such as leaves, stem, roots and flowering tissues under Pi deficiency (Wang et al., 2004). The PHO1 family in Arabidopsis comprises a total of 11 exporters that mainly function to mediate Pi export from the root cells and distribute it to other plant organs (Hamburger et al., 2002; Stefanovic et al., 2007; Kisko et al., 2018). In Arabidopsis mutants, PHO1 and PHO1:H1 showed severe phenotype changes by decreased Pi accumulation under Pi deficiency (Stefanovic et al., 2007). The uptake and translocation of Pi into and out of cells are crucial for Pi homeostasis in plants. The stability of Pi homeostasis depends on the ion movement between cells, subcellular compartments and organs (Kisko et al., 2018).

Results obtained here can be used to shed new light on potential evolution of splice variants with altered domain structures with the *B. rapa PHO1* gene family. The data will facilitate future work examining the expression profiles of these variants in different tissue types and environmental conditions. This is important since the alternative splicing events vary in their differential regulation between tissues, as well as environmental conditions (Blencowe, 2006). Moreover, the clone transcripts contained can be used in transgenic studies to find their specific role. The role of SPX and EXS domains can be studied in PHO1_D and PHO1_A as three splice variants contained both domains whereas EXS is depleted in the other two splice variants.

**CHAPTER 5 VARIATION IN LIPID REMODELLING IN
CULTIVARS OF *B. NAPUS* AND *B. RAPA* (R-O-18)
UNDER PI SUFFICIENT (P+) AND PI DEFICIENT (P-)
CONDITIONS.**

5.1 BACKGROUND

Phospholipids are the main molecules that make up the biological membranes in organisms and are highly conserved from bacteria to mammals and higher plants (Nakamura, 2017). In plants, membrane lipid composition is not symmetrical in the cell, for instance, phospholipids are the primary lipids in extra-plastidial membranes, while in photosynthetic membranes, non-phosphorus galactolipids are the major constituents (Nakamura, 2013). Phospholipid composition consists of phosphatidylcholine (PC), phosphatidylethanolamine (PE), and phosphatidylglycerol (PG) and these lipid classes are classified based on the structure of specific polar head groups and hydrophobic acyl tails (Schwertner & Biale, 1973; Nakamura, 2017). Changes in the membrane lipid composition can occur due to the exposure to different kinds of stress, such as Pi starvation. During the Pi starvation, the plant makes use of membrane phospholipids to cope with the internal Pi demand by converting phospholipid to non-phosphorus galactolipid and sulfolipid and releasing free Pi for use in critical cellular functions (Lambers et al., 2012). Two major galactolipids, monogalactosyldiacylglycerol (MGDG) and digalactosyldiacylglycerol (DGDG), and a sulfolipid, sulfoquinovosyldiacylglycerol (SQDG), are synthesized exclusively in the chloroplasts (Moellering & Benning, 2011). MGDG and SQDG comprise approximately 50 and 30% of chloroplast membrane lipids, respectively. There are two types of MGDG enzyme, type-A (MGD1) and type-B (MGD2 and MGD3). In Arabidopsis, under normal growth condition, MDG1 is responsible for MGDG (single-layer lipids) synthesis and subsequent DGDG (bilayer lipids) synthesis. MDG1 synthase activity localizes on the inner envelop membrane and synthesizes MGDG in the thylakoid membrane (Awai et al., 2001). Type-B MGDG synthases, MDG2 and MDG3 localise on the outer envelope membrane (Awai et al., 2001; Murakawa et al., 2014). Under Pi deficiency, phospholipids are degraded to release Pi. To compensates for decreases in PC, DGDGs are exported to extraplastidial membranes to maintain the membrane lipid-

bilayer structure of PC in shoot and root tissues (Essigmann et al., 1998; Andersson et al., 2003; Jouhet et al., 2004).

The mechanism of membrane lipid metabolism under Pi deficiency begins with the induction of DGDG production in the plastid. The excess of DGDG is exported to plasma, tonoplast and mitochondria membranes (Jouhet et al., 2004; Andersson et al., 2005; Shimojima, 2011). High P content phospholipids are hydrolysed to produce diacylglycerol (DAG) and free Pi. DAG acts as a substrate for galactolipid (MGDG and DGDG) and SQDG synthesis (Nakamura, 2013). Two types of DAG synthases, first, DGD1 and DGD2 synthesise DGDG through either one of two pathways; i) one-step reaction by phospholipase C (PLC); ii) two-step reactions by phospholipase D (PLD). Second, breakdown of PC to produce DAG (Shimojima & Ohta, 2010). Analysis of tissue extract of Arabidopsis showed the majority of PC to DAG conversion involved PLC. (Kuppusamy et al., 2014). Pi starvation induced PLD ζ 2 and PLD ζ 1, isoforms of enzymes involved in PC hydrolysis in roots and increased DGDG accumulation (Cruz-Ramirez et al., 2006; Li et al., 2006). Transcriptional responses to P stress showed another isoform in root and leaf tissues are NONSPECIFIC PHOSPHILIPASE C4 (NPC4) and NPC5 (Nakamura et al., 2005; Gaude et al., 2008). Glycerophosphodiester will also be hydrolysed by GLYCEROPHOSPHODIESTER PHOSPHODIESTERASE (GDPD), for example glycerophosphocholine into glycerol-3-phosphate and choline (Cheng et al., 2011).

Another type of non-phosphorus lipid, sulfolipid may replace phospholipids in plants under Pi deficiency (Essigmann et al., 1998). Three enzymes encoded by the genes of the sulfolipid biosynthetic pathway are UDP-GLUCOSEPYROPHOSPHORYLASE3 (UGP3), THE UDP-sulfoquinovose synthase SQD1, and the sulfoquinovosyldiacylglycerol (SQDG) synthase SQD2 (Hammond et al., 2003; Shimojima, 2011).

The genes involved in MGDG type-B synthesis, as well as phospholipid degradation, are highly upregulated under Pi starvation (Kobayashi et al., 2009; Kuppusamy et al., 2014). Although *MGD1* transcript was unchanged, MGD1 synthase still contributed to the increasing DGDG concentration during Pi deficiency (Kobayashi et al., 2009). *MDG2* and *MGD3* genes were highly induced in non-photosynthetic tissues especially in the roots in *Arabidopsis* (Kobayashi et al., 2004). Membrane lipid remodelling under Pi deficiency has also been reported in other plants including soy beans (*Glycine max*), oat (*Avena sativa*), bean (*Phaseolus vulgaris*), and Protaceae species (Andersson et al., 2003; Gaude et al., 2004; Lambers et al., 2012).

Recent advances in mass spectrometry (MS) technology now enable the determination of lipid profiles and the study of lipidomics in plants through spatiotemporal phospholipid molecular profiles (Han & Gross, 2003; Samarakoon et al., 2012; Nakamura, 2017). Lipidomics can also define the isobaric species and structural isomers (Shulaev & Chapman, 2017). Lipid species are determined using a sensitive electrospray ionisation tandem mass spectrometry (ESI-MS/MS) based approach. ESI-MS/MS allows the direct biological extract to pass through a triple quadrupole tandem mass spectrometer which collects mass spectra in series. Each spectrum denotes specific lipid classes with similar polar lipid head groups, subsequently determining the lipid composition of the sample (Welti & Wang, 2004).

5.2 AIMS AND OBJECTIVES

The aim of this study was to understand the within species variation in lipid remodelling under low Pi availability and was achieved through the following objectives:

1. Determine the species variation in membrane lipid composition 24 different cultivars/lines of *B. napus* and *B. rapa* R-o-18.
2. Determine variation in the changes in lipid profiles in response to low Pi availability in 24 different cultivars/lines of *B. napus* and *B. rapa* R-o-18 grown hydroponically under Pi replete and Pi deficient conditions.

5.3 MATERIALS AND METHODS

5.3.1 Plant materials and stress treatment

Twenty-four cultivars/lines of oil seed rape (*B. napus*) (Table 5.1) and *B. rapa* R-o-18 seeds were pre-germinated on a petri dish and grown hydroponically in nutrient media (Table 2.1) in a glass house at the University of Reading. After 21 days, half of the seedlings were subjected to Pi stress treatment (P-) by changing the media to one with zero Pi concentrations (by replacing KH_2PO_4 with K_2SO_4) as described in section 2.3 (Table 2.2), and the other half treated with a full nutrient (P+). After 30 days, leaf samples from the third leaf were collected at approximately 9.30 a.m, frozen in liquid nitrogen, and stored at -80 °C until subsequent analyses. Initially pooled leaf samples of 24 lines of *B. napus* and *B. rapa* R-o-18 with three replicates were used for screening. A subsequent experiment was conducted using leaf discs taken from the same spot next to the mid-rib of six lines of *B. napus* and one line of *B. rapa* R-o-18 with five replicates. The lines were selected for detailed lipid analysis based on ratio DGDG/PC response from the first experiment.

Table 5.1. Twenty-four *Brassica napus* genotypes from the ASSYST (BA) population used to quantify genetic variation in gene expression responses to low Pi availability. Seeds were provided by the University of Nottingham.

No./Line ID	Line Name	Line Morphotype
1. BA-040	Apex	Modern winter OSR
2. BA-044	Laser	Modern winter OSR
3. BA-048	NK Bravour	Modern winter OSR
4. BA-067	Nugget	Modern winter OSR
5. BA-089	AMBER X COMMANCHE DH	Winter OSR
6. BA-099	TAPIDOR DH	Winter OSR
7. BA-101	EUROL	Winter OSR
8. BA-102	Lesira	Winter OSR
9. BA-113	Samourai	Winter OSR
10. BA-154	Liporta	Winter OSR
11. BA-158	Major	Winter OSR
12. BA-177	Victor	Winter OSR
13. BA-185	CANARD	Winter fodder
14. BA-202	Chuosenshu	"Exotics": Winter
15. BA-208	COUVE NABICA	"Exotics": cauve nabica
16. BA-204	Q100	"Exotics":synthetic
17. BA-209	RAGGED JACK	"Exotics": rape kale
18. BA-210	RED RUSSIAN	"Exotics": Siberian kale
19. BA-216	E94197	"Exotics": unspecified
20. BA-218	GROENE GRONINGER	"Exotics": unspecified
21. BA-224	WILD ACCESSION	"Exotics": wild accession
22. BA-410	HUGUENOT	swede
23. BA-418	Altasweet	swede
24. BA-426	Drummonds Purple Top	swede

5.3.2 Lipid extraction

The lipid extractions were performed using the acidic-chloroform-methanol method under the fume hood. First, the frozen leaf disc was added to 7 mL glass tube containing 1 mL chloroform/methanol/formic acid (1:1:0.1) and the sample was homogenised with a polytron homogeniser. Another 2 mL chloroform/methanol/formic acid (1:1:0.1, v/v) and 1 mL aqueous 0.2 M phosphoric acid (H_3PO_4), 1 M potassium chloride (KCl) were added. The tube was then briefly shaken, centrifuged for 3 min at 1500 g to allow the phases to separate. The lower phase, which contains the lipids, was transferred using a glass pipette to a new glass tube. The extraction process was repeated twice. A 3 mL volume of chloroform was added to the remaining tissue in the tube, vortexed and centrifuged for 2 min at 1500 g. After this, the bottom phase was removed and subsequently combined with the other chloroform phase. The samples were stored under N_2 in sealed tube at -20°C . Prior to running the samples on the electrospray ionisation tandem mass spectrometry (ESI-MS/MS), the samples were filtered (to remove debris) using 4 μm Millex Syringe Filters (Millipore, MA, USA) and diluted in 250 μL chloroform (Fig 5.1).

5.3.3 Quantitative lipid analysis

ESI-MS/MS (4000, QTRAP, SCIEX, CA, USA) was used to quantify and identify the major type lipids particularly phospholipid (e.g. PC), galactolipid (e.g. DGDG) and sulfolipid (SQDG) in the leaf sample of *Brassica* spp. under Pi stress. The internal standards of PC and PE were obtained from Avanti (AL, USA). PC was incorporated as 0.857 nmol of 24:1 PC and 0.08 nmol of 14:0 PG. Standards were dissolved in chloroform. 10 μL of sample extract in 250 μL chloroform was combined with standard and polar spray solvent (chloroform/methanol/400 nM ammonium acetate; 300:665:3.5, v/v/v) to make a final volume (1 mL). Samples were infused at 15 $\mu\text{L min}^{-1}$ with an autosampler (HTS-xt PAL,

CTC-PAL Analytics AG, Switzerland). Data were analysed according to method by Gonzalez-Thuillier et al., (2015) with some modifications.

Relative abundance of lipids measured by ESI-MS/MS was used to determine lipid species in different lines of *B. napus*. The lipid species were measured by precursor or neutral loss scanning and the lipids in each head-group class were quantified by a correction curve determined between internal standards. The measurements and analysis were conducted by Dr Richard Haslam at Rothamsted Research, Harpenden, UK.

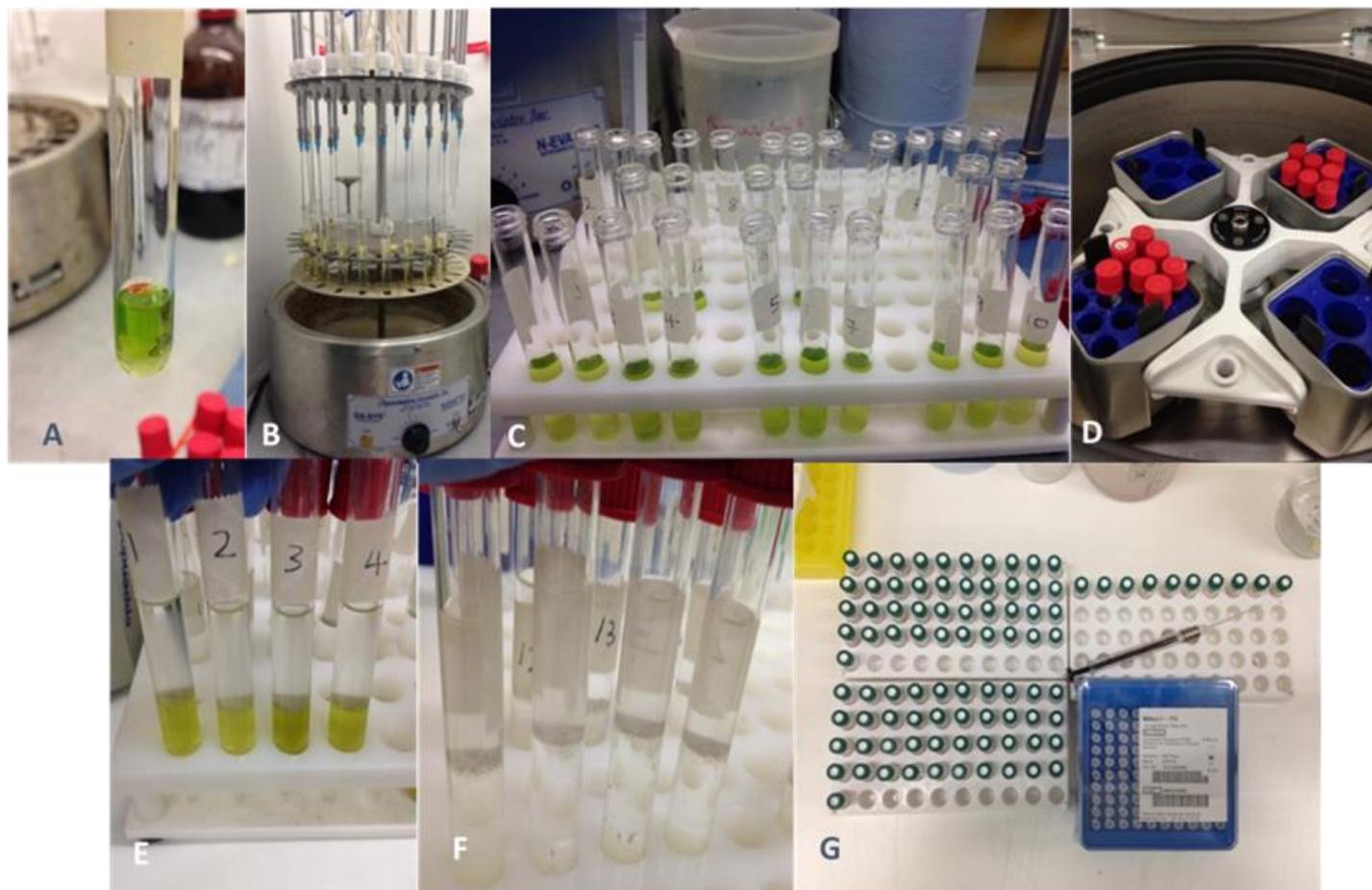


Figure 5.1. A-F: Lipid extraction procedure using acidic-chloroform-methanol method. G. The samples were filtered before running on the electro-spray ionisation tandem mass spectrometry (ESI-MS/MS).

5.4 RESULTS

Under normal and Pi sufficient growth conditions, the digalactosyldiacylglycerol/phosphatidylcholine ratio is stable. This stable condition will face significant changes under different environmental conditions or developmental stages to maintain their homeostasis. Under Pi deficiency, regulation of lipid composition changes occurs. These changes can be quantified using the DGDG/PC ratio as the replacement of phosphatidylcholine (PC) by digalactosyldiacylglycerol (DGDG) in plasma membranes during Pi deficiency occurs.

5.4.1 Changes of lipid molecular species of 24 *B. napus* genotypes and *B. rapa* R-o-18 during Pi deficiency

Twenty-four lines were selected in this study as a subset of a larger diversity of *B. napus* lines (about > 300 lines in total) representing the diversity of the species morphological and biochemical properties. Lipid profile screening of leaf samples of *B. napus* grown hydroponically in the glasshouse under sufficient Pi (P+) and deficient Pi (P-) conditions for nine days were analysed (Fig. 5.2). The breakdown of the membrane lipid phosphatidylcholine (PC) is replaced with non-phosphorus galactolipid DGDG and SQDG. To assess the rate of PC to DGDG and SQDG conversion in plants grown hydroponically under P+ and P- for nine days before harvesting, DGDG/PC ratio as well as SQDG/PC ratio were calculated. Only DGDG/PC ratio results were shown for screening of 24 lines of *B. napus* and *B. rapa* R-o-18 because DGDG is the predominant replacement of phospholipid in Pi-deficient plants (Andersson et al., 2003; Russo et al., 2007; Tjellström et al., 2008) (Fig. 5.2).

Under P+ conditions, genotype BA-113 (Samourai) showed the highest DGDG/PC ratio (6-fold), while BA-218 (GROENE GRONINGER SNIJMOES) recorded the highest under

P- conditions (18-fold). The lowest DGDG/PC ratio was recorded in BA-210 (RED RUSSIAN), both under P+ and P- conditions. Under P+ conditions there was a 6-fold variation in DGDG/PC ratio between the extreme phenotypes, whereas a 9-fold variation amongst the 24 genotypes was observed under P- conditions (Fig. 5.2).

Delta ratio of DGDG/PC was calculated to look at the response on DGDG/PC changes in all lines between P+ and P- conditions (Fig 5.3). The changes in this ratio within a line were varied, ranging from 1-fold up to 13-fold. BA-224 (WILD ACCESSION) recorded the highest delta ratio, similar to BA-218 (GROENE GRONINGER SNIJMOES), while BA-099 (TAPIDOR DH) recorded the lowest (Fig. 5.3). These results were used to select six lines of *B. napus* and *B. rapa* R-o-18 were selected for further lipidomics and RNA-seq analysis (Chapter 6) based on their extreme responses in lipid profiles under P+ and P- conditions. Based on this, the six *B. napus* lines selected were, BA-224 (WILD ACCESSION) which showed the most response in DGDG/PC ratio, followed by BA-218 (GROENE GRONINGER SNIJMOES), BA-102 (Lesira), BA-044 (Laser), BA-210 (RED RUSSIAN) and BA-099 (TAPIDOR DH) in addition to *B. rapa* R-o-18.

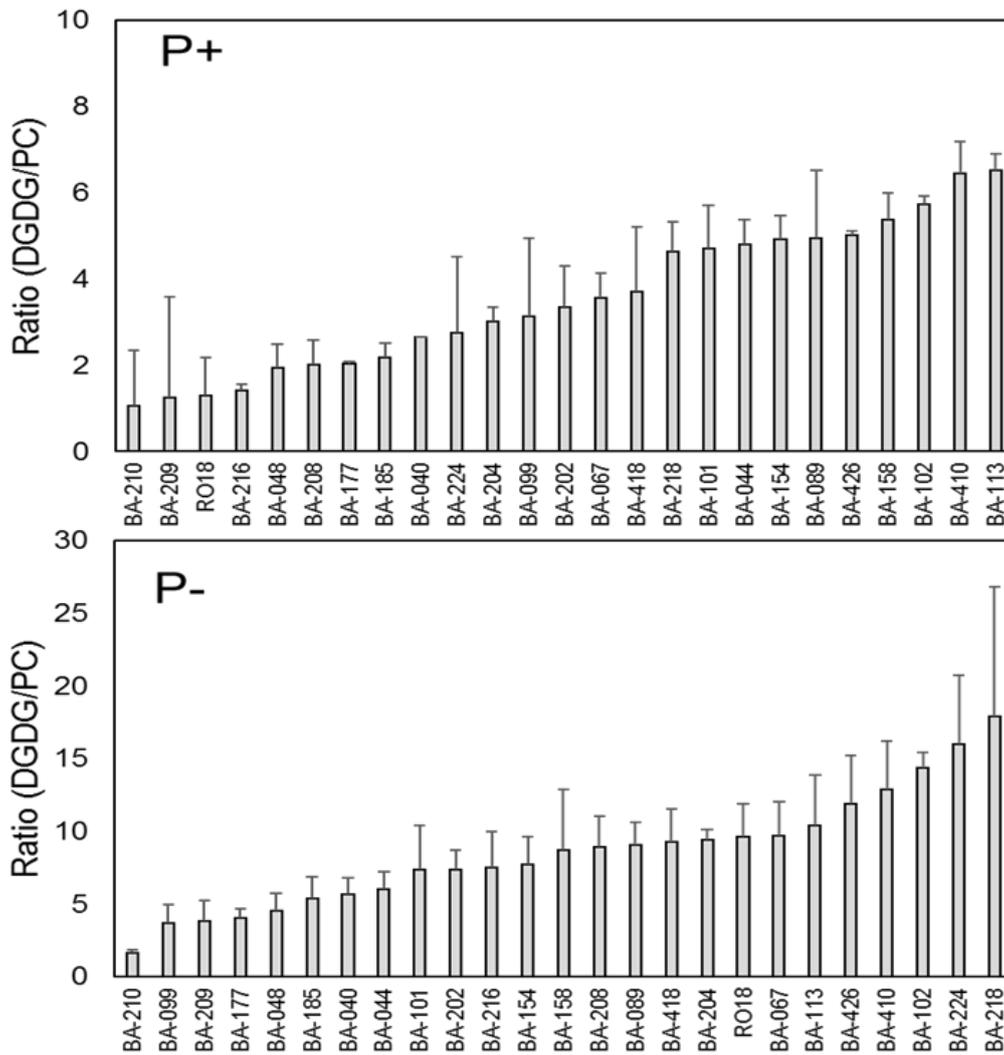


Figure 5.2. Digalactosyldiacylglycerol/ phosphatidylcholine (DGDG/PC) ratio based on DGDG and PC concentrations measured by electrospray ionisation tandem mass spectrometry (ESI-MS/MS) in leaf samples from 24 lines of *Brassica napus* and one line of *Brassica rapa* (R-o-18) (Table 5.1) grown hydroponically in the glasshouse under sufficient (P+) and deficient (P-) Pi conditions. Bars are based on ratios of three independent biological replicates and five technical replicates per treatment.

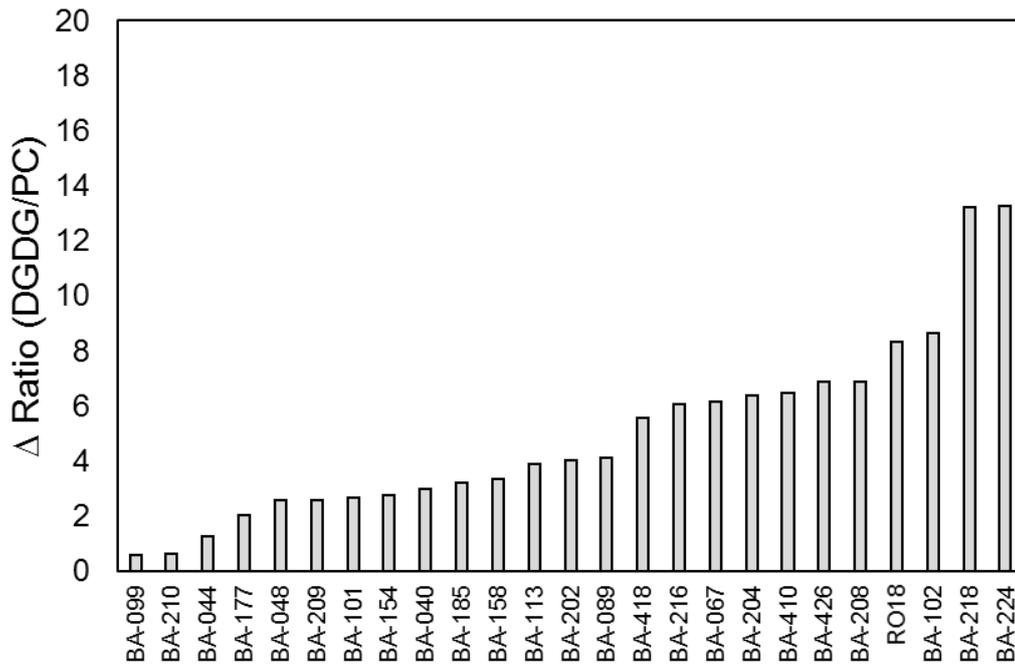


Figure 5.3. Delta ratio digalactosyldiacylglycerol/ phosphatidylcholine (DGDG/PC) based on DGDG and PC concentrations measured by electrospray ionisation tandem mass spectrometry (ESI-MS/MS) on leaf tissue samples from 24 lines of *Brassica napus* and one line of *Brassica rapa* (R-o-18) (Table 5.1) grown hydroponically in the glasshouse under sufficient (P+) and deficient (P-) Pi conditions. Bars are based on ratios of three independent biological replicates and five technical replicates per treatment.

5.4.2 Lipid analysis of seven genotypes in detail

Analysis of the lipid profiles was conducted on six lines of *B. napus* (BA-210, BA-102, BA-044, BA-099, BA-218 and BA-224) and on *B. rapa* R-o-18 leaf disc samples from plants grown hydroponically in the glasshouse. These lines were selected for detailed lipid remodelling analysis based on their extreme response in the initial lipid profile screening on 24 lines experiment (Fig. 5.2 and 5.3).

5.4.2.1 Changes of lipid molecular species: Comparison between DGDG and PC

The total lipid concentration for individual lipids was calculated by the accumulation all the quantities of each lipid class for that lipid species (measured as nmol mg⁻¹ DW), for example DGDG 34:1-6 36:3-6. Leaf PC concentration accounted for a lower quantity of the lipids in the P- condition compared to P+ condition across all cultivars/lines.

Under P- condition, DGDG concentration increased in BA-102, BA-218, R-o-18 and BA-224 compared to DGDG concentration under P+ condition. While three lines showed slightly lower concentrations of DGDG by 0.32, 0.19 and 0.26 µmol mg⁻¹ DW in BA-099, BA-044 and BA-210, respectively in comparison to P+ condition (Fig. 5.4). Although some lines showed slightly reduced concentrations of DGDG under P- condition compared to P+ condition, the overall DGDG concentrations were significantly higher ($P < 0.05$) than PC concentrations under Pi deficiency. It can be concluded that the concentration of DGDG was highly responsive and responded to external Pi concentration, but there was significant variation in the absolute amounts of lipids and the changes in these within the Brassica lines examined here.

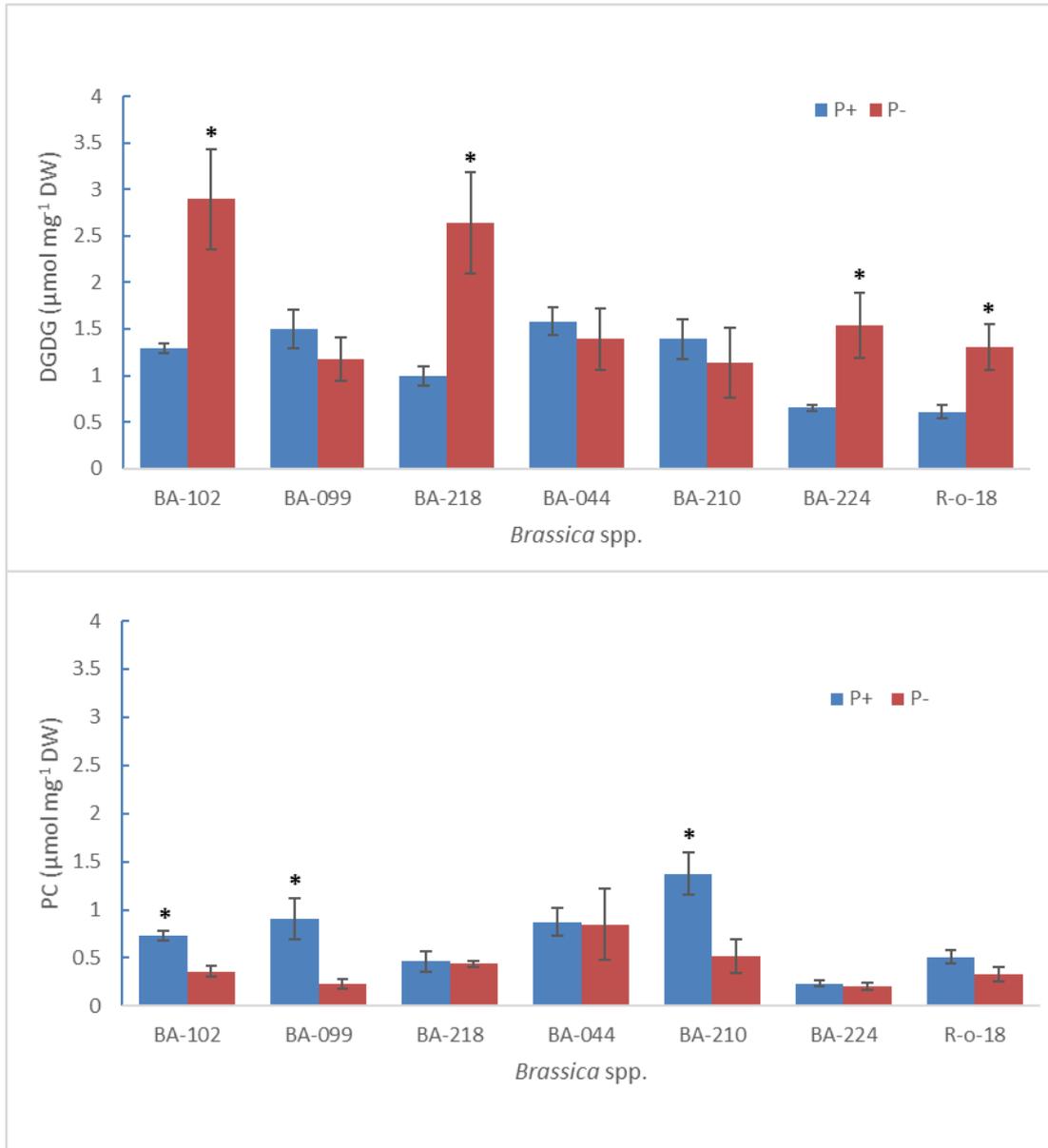


Figure 5.4. Leaf digalactosyldiacylglycerol (DGDG) and phosphatidylcholine (PC) concentrations measured by electrospray ionisation tandem mass spectrometry (ESI-MS/MS) on six lines of *Brassica napus* (BA-099, BA-224, BA-044, BA-102, BA-218 and BA-210) and one line of *Brassica rapa* (R-o-18) (Table 5.1) grown hydroponically in the glasshouse under sufficient (P+) and deficient (P-) Pi conditions. Bars are based on ratios of five independent biological replicates and five technical replicates per treatment. Error bars represent mean \pm SEM (n=5). * DGDG or PC concentrations whose significantly different ($P < 0.05$) using one-way ANOVA test between P+ and P-.

Converting these absolute lipid quantities to ratios showed that BA-224 recorded the highest DGDG/PC ratio under the P+ condition (Fig. 5.5). While under the P- condition, BA-210 recorded highest DGDG/PC ratio. The lowest DGDG/PC ratio was recorded in BA-210 and BA-099 in the P+ and P- conditions, respectively.

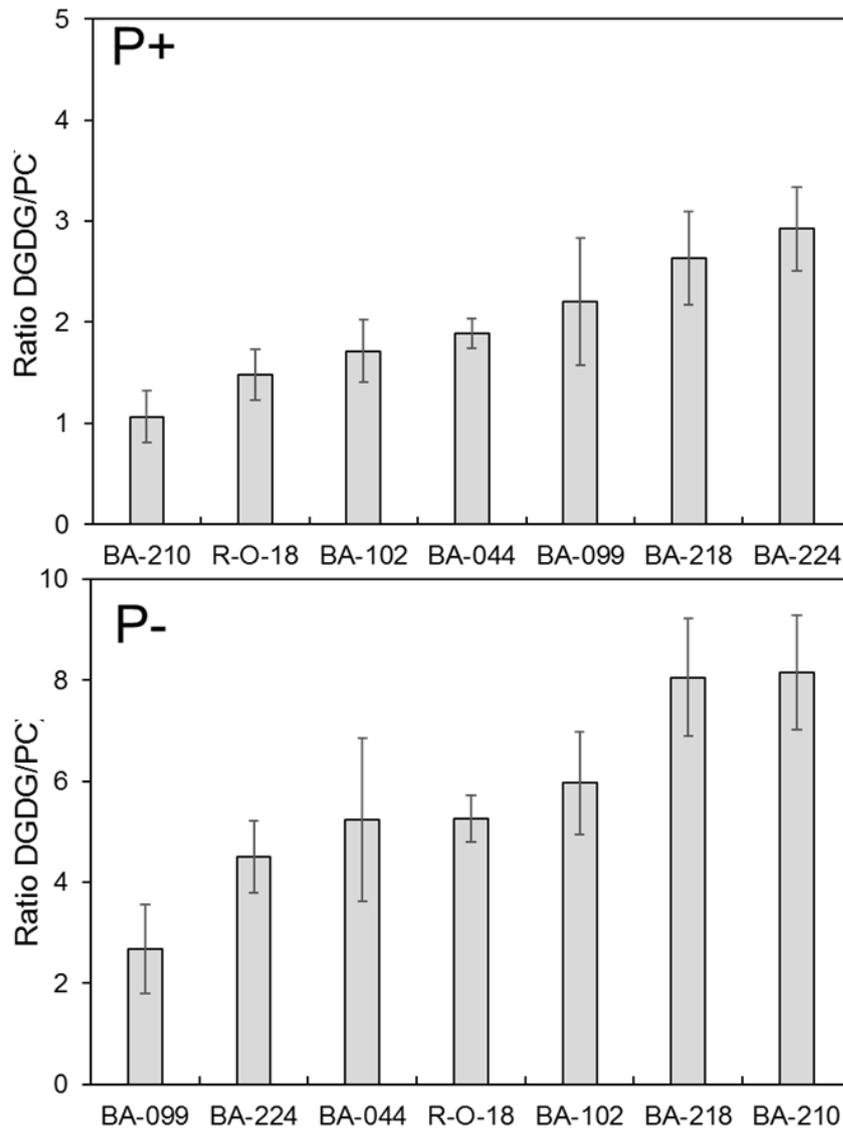


Figure 5.5. Digalactosyldiacylglycerol/phosphatidylcholine (DGDG/PC) ratio measured by electrospray ionisation tandem mass spectrometry (ESI-MS/MS) on six lines of *Brassica napus* (BA-099, BA-224, BA-044, BA-102, BA-218 and BA-210) and one line of *Brassica rapa* (R-o-18) (Table 5.1) grown hydroponically in the glasshouse under sufficient (P+) and deficient (P-) Pi conditions. Bars are based on the ratios of five independent biological replicates and five technical replicates per treatment. Error bars represent mean \pm SEM (n=5).

Considering the response to Pi availability, BA-210 recorded the most responsive Δ DGDG/PC ratio, followed by BA-218, BA-102, R-o-18, BA-044, BA-224 and BA-099 (Fig. 5.6). The variation of response rate across all the samples indicated that membrane lipid remodelling response varies depending of lines of *B. napus* (BA-102, BA-099, BA-218, BA-044, BA-210 and BA-224) and species (*B. napus* and *B. rapa* R-o-18). Interestingly, the range in response was lower than that observed in the initial experiment (Fig. 5.3 and 5.3).

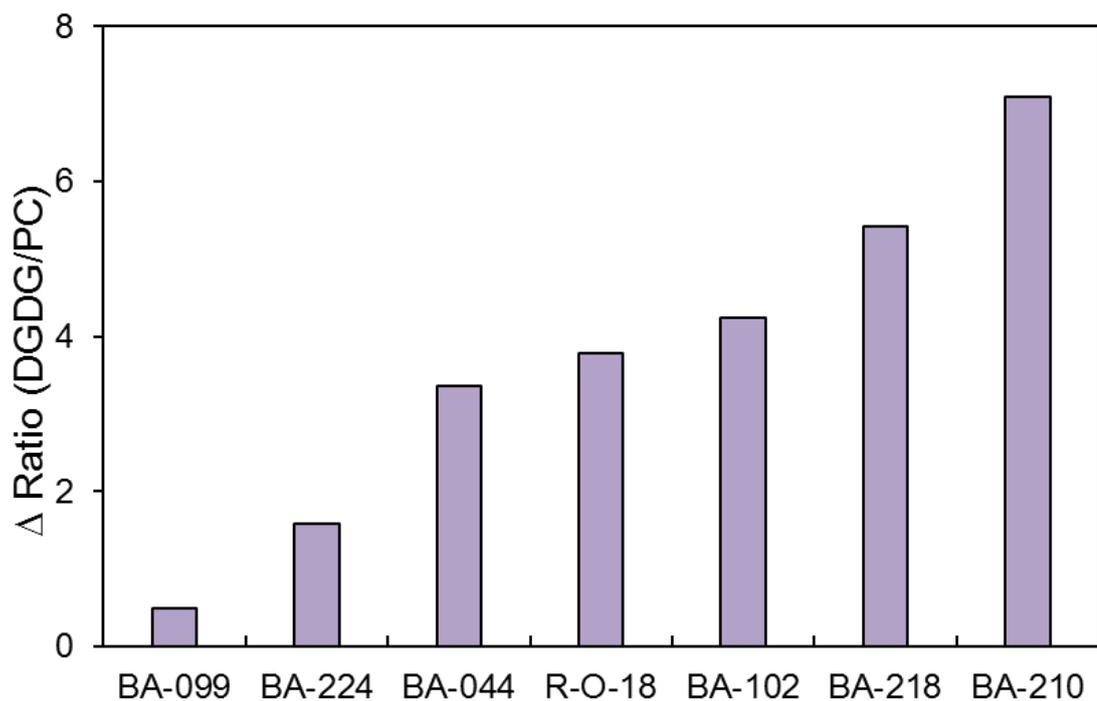


Figure 5.6. Delta digalactosyldiacylglycerol/phosphatidylcholine (DGDG/PC) ratio based on DGDG and PC concentrations measured by electrospray ionisation tandem mass spectrometry (ESI-MS/MS) on six lines of *Brassica napus* (BA-099, BA-224, BA-044, BA-102, BA-218 and BA-210) and one line of *Brassica rapa* (R-o-18) (Table 5.1) grown hydroponically in the glasshouse under sufficient (P+) and deficient (P-) Pi conditions. Bars are based on the ratios of five independent biological replicates and five technical replicates per treatment.

Table 5.2. Ranking of delta digalactosyldiacylglycerol/phosphatidylcholine (DGDG/PC) ratio in common samples from the initial experiment and the detailed experiment. Rank highest to the lowest of six lines of *Brassica napus* and *Brassica rapa* R-o-18.

Initial experiment	Detailed experiment
BA-224	BA-210
BA-218	BA-218
BA-102	BA-102
R-o-18	R-o-18
BA-044	BA-044
BA-210	BA-224
BA-099	BA-099

5.4.2.2 Changes of lipid molecular species: Comparison between SQDG and PC

Another effect of P availability on membrane lipid composition can be seen through the replacement of phospholipids/ phosphatidylcholine (PC) by sulfolipids/ sulfoquinovosyldiacylglycerol (SQDG) in the photosynthetic membranes. Results showed variation between lines and P treatment in SQDG and PC concentrations (Fig. 5.7). Under the P+ condition, PC concentrations in leaf tissues were significantly higher ($P < 0.05$) than SQDG across all the lines (BA-102, BA-099, BA-218, BA-044, BA-210, BA-224 and R-o-18). Leaf PC concentrations ranged from 0.4-1.2 $\mu\text{mol mg}^{-1}$ DW across lines under P+ treatment, while the leaf SQDG concentrations ranged between 0.07-0.2 $\mu\text{mol mg}^{-1}$ DW.

The SQDG and PC composition changed significantly under the P- condition. Leaf PC concentrations were lower across all the lines in response to low P. In contrast, SQDG concentrations increased under the P- condition compared to P+ condition (Fig. 5.7). Three of the lines, BA-102, BA-099 and BA-224 recorded slightly higher concentrations of SQDG than PC, indicating that SQDG plays a significant role in phospholipid replacement under Pi starved conditions. The leaf PC concentrations in four lines (BA-214, BA-044, BA-210 and R-o-18) were significantly higher ($P < 0.05$) than SQDG content under the P- condition with 0.4, 0.55, 0.6 and 0.3 $\mu\text{mol mg}^{-1}$ DW, respectively.

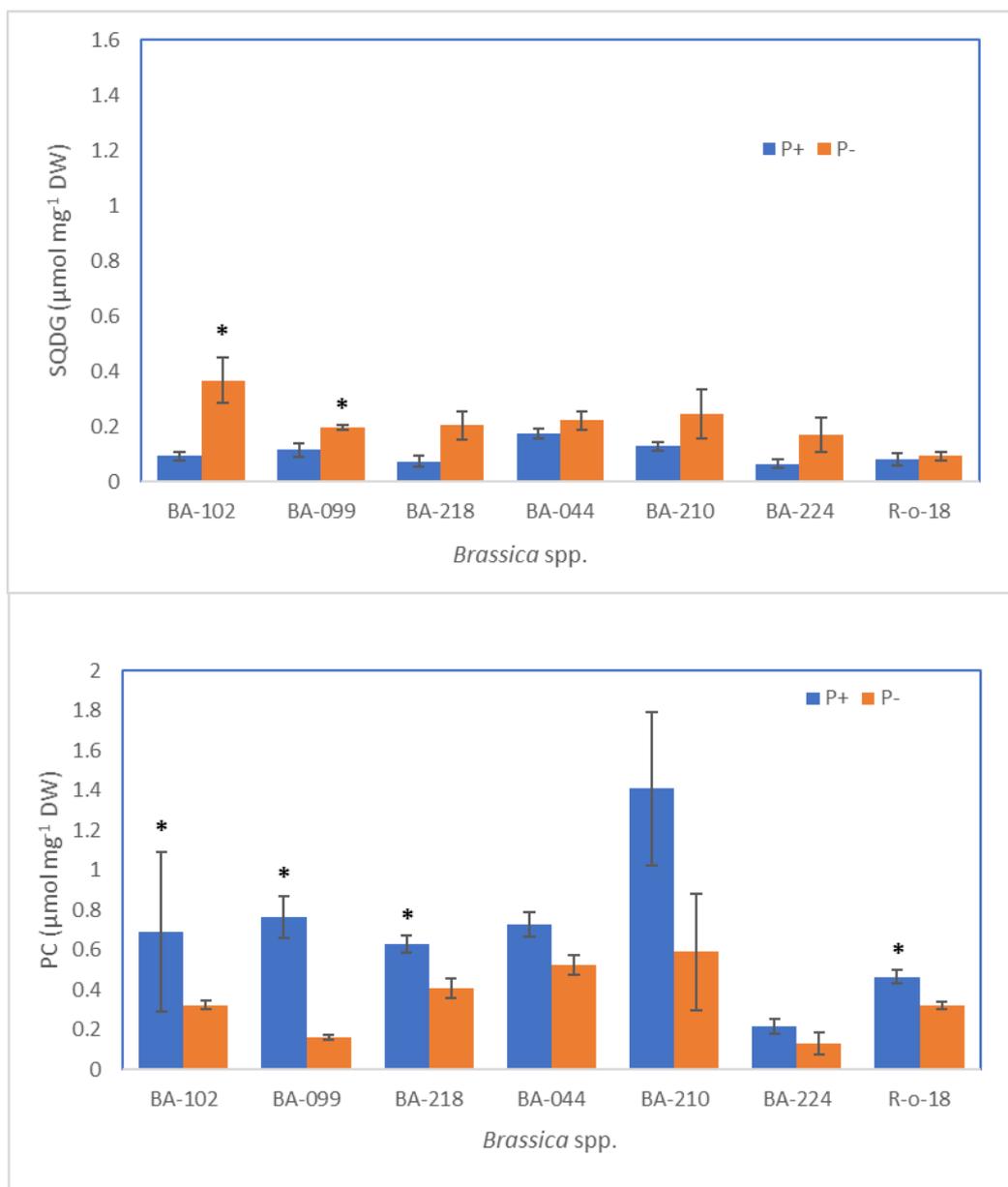


Figure 5.7. Leaf sulfoquinovosyldiacylglycerol (SQDG) and phosphatidylcholine (PC) concentrations measured by electrospray ionisation tandem mass spectrometry (ESI-MS/MS) six lines of *Brassica napus* (BA-099, BA-224, BA-044, BA-102, BA-218 and BA-210) and one line of *Brassica rapa* (R-o-18) (Table 5.1) grown hydroponically in the glasshouse under sufficient (P+) and deficient (P-) Pi conditions. Bars are based on the ratios of five independent biological replicates and five technical replicates per treatment. Error bars represent mean \pm SEM (n=5). * SQDG or PC concentrations whose significantly different ($P < 0.05$) using one-way ANOVA test between P+ and P-.

Analysis of the SQDG/PC ratio was conducted to look at the effect of phospholipid changes to sulfolipids across all investigated lines under different P treatments (P+ and P-) (Fig.5.8). The SQDG/PC ratios were significantly higher in the P- condition across all

the lines. The highest SQDG/PC ratio was recorded in BA-224 both in the P+ and P- conditions. BA-210 had the lowest SQDG/PC ratio under the P+ conditions, while R-o-18 had the lowest under the P- conditions. The variation in SQDG/PC ratio across the *Brassica* lines and species showed different lines and/or species influenced the lipid metabolism process differently under Pi deficiency.

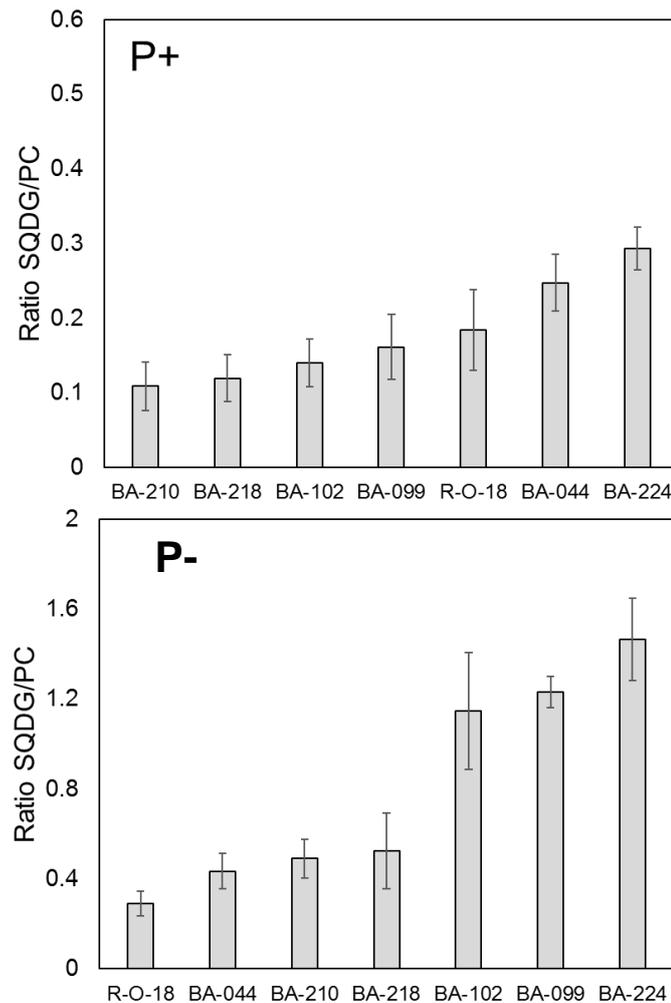


Figure 5.8. Sulfoquinovosyldiacylglycerol/phosphatidylcholine (SQDG/PC) ratio measured by electrospray ionisation tandem mass spectrometry (ESI-MS/MS) on six lines of *Brassica napus* (BA-099, BA-224, BA-044, BA-102, BA-218 and BA-210) and one line of *Brassica rapa* (R-o-18) (Table 5.1) grown hydroponically in the glasshouse under sufficient (P+) and deficient (P-) Pi conditions. Bars are based on the ratios of five independent biological replicates and five technical replicates per treatment. Error bars represent mean \pm SEM (n=5).

The delta SQDG/PC ratio showed BA-224 recorded the highest phospholipid to sulfolipid response in leaf tissues under Pi deficiency, followed by BA-044, R-o-18, BA-99, BA-102, BA-218 and the smallest ratio was recorded in BA-210 (Fig. 5.9).

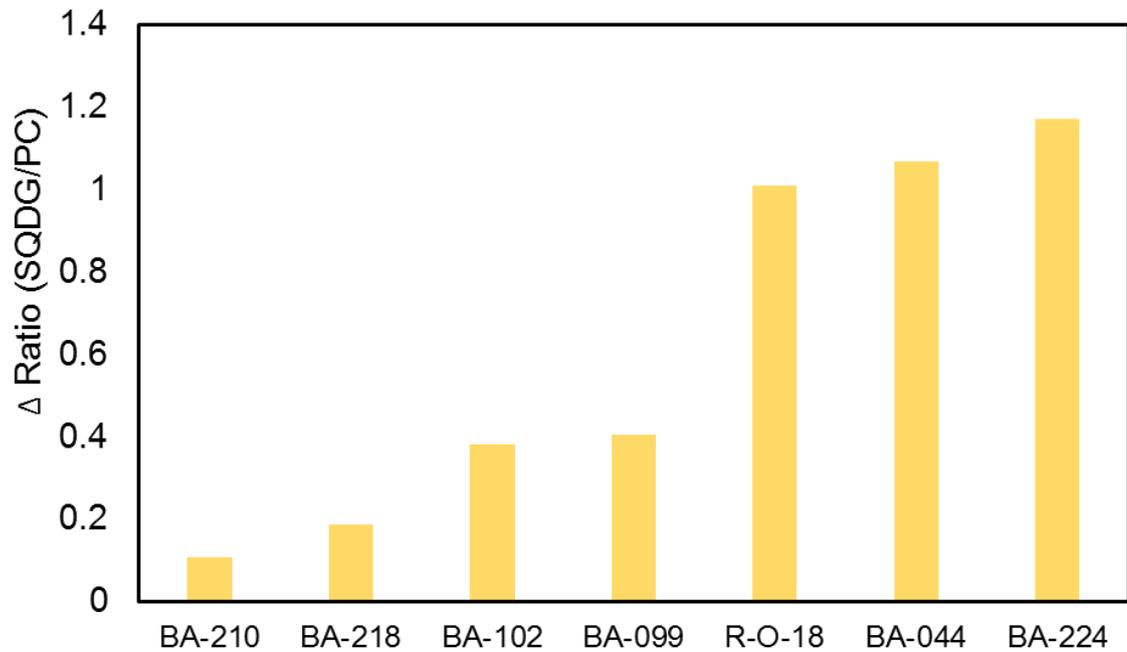


Figure 5.9. Delta sulfoquinovosyldiacylglycerol/phosphatidylcholine (SQDG/PC) ratio based on SQDG and PC concentrations measured by electrospray ionisation tandem mass spectrometry (ESI-MS/MS) on six lines of *Brassica napus* (BA-099, BA-224, BA-044, BA-102, BA-218 and BA-210) and one line of *Brassica. rapa* (R-o-18) (Table 5.1) grown hydroponically in the glasshouse under sufficient (P+) and deficient (P-) Pi conditions. Bars are based on the ratios of five independent biological replicates and five technical replicates per treatment.

5.5 DISCUSSION

Changes in leaf lipid profiles were determined for 24 cultivars/lines of *B. napus* and *B. rapa* grown hydroponically under low P conditions. Polar glycerolipids including polar head groups in phospholipids (PC, PE, PI, PS, PA and PG), two head-group classes in glycerolipids (MGDG and DGDG) and sulfolipid (SQDG) were identified using ESI-MS/MS. ESI-MS/MS allows rapid analysis of membrane lipid composition and determines lipid changes with only small quantities of sample extract (Welti et al., 2002). The quantification of the classes was determined using the comparison with internal standards of the class. ESI-MS/MS enables the determination of lipid species which could act as a substrate and product of specific enzymes that are involved in many plant responses to external and internal environmental changes (Welti et al., 2002; Nandi et al., 2003).

5.5.1 Changes in DGDG varied between Brassica lines under Pi deficiency

From the results, the amount of phospholipid PC was significantly reduced under Pi starvation in all of *B. napus* lines and *B. rapa* R-o-18. This result is consistent with previous studies indicating that phospholipids are degraded through hydrolysis to release Pi and are replaced by galactolipids and sulfolipids under Pi deficiency (Essigmann et al, 1998; Jouhet et al., 2004; Andersson et al., 2005; Nakamura, 2013). The results showed a significant increase in galactolipid DGDG concentration in leaf tissues under the P-condition of all lines. These results are consistent with previous research that reported galactolipid MGDG and DGDG synthesis at the outer membrane of plastids during Pi-depleted conditions will replace phospholipids to maintain membrane function (Awai et al., 2001; Kobayashi et al., 2004). Phospholipid degradation at the plasma membrane releases Pi to be used for other essential biological processes. MGDG and DGDG are produced in the plastid at the thylakoid membrane under normal conditions (Benning & Ohta, 2005; Shimojima & Ohta, 2011). The synthesis of galactolipids is induced

extensively under Pi depleted conditions and exported to the extraplastidial membrane to replace the loss of PC (Fig. 5.5). Similar responses of changes in lipid composition have been observed in other plants including *Arabidopsis*, soy bean (*Glycine max*), oat (*Avena sativa*) and proteaceae species (*Hakea prostrata*) (Andersson et al., 2003, Gaude et al., 2004; Jouhet et al., 2007; Lambers et al., 2012).

Comparison between the six lines of *B. napus* and *B. rapa* R-o-18 showed a consistent response between DGDG/PC ratio in both initial and the detailed experiment except for BA-224 and BA-210 (Table 5.2). The increase in DGDG/PC ratios was the result of increased in DGDG and decreased of PC concentration. The differences of results on these two lines particularly may be due to different methods used for collecting tissue samples, the initial experiment using whole leaf tissues, while in the detailed experiment, the protocol was altered to take the leaf disc sample from the same place on the leaf and increase the number of replicates from three to five replicates to make the comparison more valuable. Okazaki et al., 2017 reported the membrane lipid metabolism responded differently and varied in different tissue samples and growth stages. The differences in absolute values obtained for lipid concentrations observed between the detailed and the initial experiments occurred due to spatial variation among the samples taken. The detailed experiment with more replicates also reduced the biological variation between the samples.

5.5.2 Sulfolipids increased in Pi deficient Brassicas

Another non-phosphorus lipid species analysed in this study, SQDG showed a significant increase in leaf tissues under the P- treatment, while the high membrane PC phospholipid concentration decreased (Fig. 5.7). SQDG is the smallest lipid class comprising only 4-7% of total lipids (Shimajima, 2011). This is consistent with results obtained here where the SQDG concentration was much lower than DGDG, ranging from 0.06 to 0.19 $\mu\text{mol mg}^{-1}$

¹ for SQDG compared to DGDG from 0.6 to 2.89 $\mu\text{mol mg}^{-1}$. The changes are similar to Arabidopsis, where DGDG and SQDG increased by 23.4%, and 4.4%, respectively under Pi deficiency (Yu et al., 2002). This suggests that SQDG is involved in the membrane lipid replacement, but it is not enough to fully compensate loss of Pi in the thylakoid membrane under Pi starvation. This also supports research by Shimojima, 2011 who found SQDG was the least prevalent component of photosynthetic membrane lipids. From the results, SQDG and PC composition changed significantly under the P- condition, with SQDG recorded increase while PC recorded decrease in concentration with different amount across all the lines (Fig 5.7). The level of variation observed between lines indicate that each line responded differently to Pi deficiency. These profiling data provide information on the metabolic process underlying the changes of membrane lipid composition.

To elucidate complex molecular mechanism of lipid metabolism in low-P plants, analysis of transcripts encoding lipid remodelling enzyme is important to assess lipid and associated transcript. The data produced in lipid analysis across seven lines which have contrasting lipid profile can be used in the next chapter to elucidate gene function and expression.

**CHAPTER 6 VARIATION IN THE EXPRESSION OF
GENES BETWEEN CULTIVARS OF *B. NAPUS* AND *B.*
RAPA (R-O-18) UNDER PI SUFFICIENT (P+) AND PI
DEFICIENT (P-) CONDITIONS.**

6.1 BACKGROUND

B. napus (Oilseed rape, AACC genome) is formed through allopolyploidy between two diploid parents, *B. rapa* (AA genome) and *B. oleracea* (CC genome) (Allender & King, 2010). Since all Brassicas underwent a genome triplication event after they shared a common ancestor with *Arabidopsis* some 13-17 mya, within each Brassica genome there are potentially three paralogs of each *Arabidopsis* gene. The amphidiploid *B. napus*, therefore potentially contains six paralogs for a given gene in *Arabidopsis*. These are found within the least fractionated (LF) and more fractionated (MF1 and MF2) sub-genomes. The LF retains more orthologous genes with *Arabidopsis* compared to the MF1 and MF2 sub-genomes (Wang et al., 2011b).

B. napus is the third world's most important oil crop after palm oil and soy bean and its protein meal is used for livestock feed (Momoh et al., 2002; Tan et al., 2011). Due to its value and significance as a temperate oil crop, its production has increased significantly over the last 30 years and now it is ranked as the third largest oil producing crop (Tan et al., 2011). Since Pi availability can be a major factor limiting plant growth including in *B. napus*, it is crucial to overcome this problem to increase phosphate acquisition efficiency (PAE) from the root and phosphorus use efficiency (PUE) within the plant. Plants have developed a series of genetic, metabolic and physiological responses to Pi deficiency. Further understanding of Pi regulation in *B. napus*, such as changes in the expression of genes and how this varies within the species will enhance our understanding of the molecular mechanisms and processes involved in adaptations to low Pi availability. To improve our insight into the genetic variation in *B. napus* and its association with variation in leaf lipid profiles, the expression profile of the genes involved in lipid metabolism was undertaken.

6.2 AIMS AND OBJECTIVES

The aim of this study was to identify genetic variation associated with metabolic variation in leaf lipid profiles observed in Chapter 5.

1. Investigate the natural genetic variability in gene expression responses in response to low Pi availability in five different cultivars/lines of *B. napus* and *B. rapa* R-o-18 grown hydroponically under Pi sufficient and Pi deficient conditions.
2. Analyse the transcriptome data to reveal variation in the expression of genes for phospholipid and non-phospholipid metabolism.

6.3 MATERIALS AND METHODS

6.3.1 Plant materials and stress treatment

Twenty-four lines of oilseed rape (*B. napus*) and *B. rapa* R-o-18 were grown hydroponically for RNA extraction and quantification analysis. Seeds were pre-germinated on a petri dish and grown hydroponically in nutrient media (Table 2.1) in a glasshouse. After 21 days, the seedlings were subjected to Pi stress treatment by changing the media with one containing no Pi (by replacing KH_2PO_4 with K_2SO_4) as described in section 2.3 (Table 2.2). Leaf samples (from the third leaf) of 30-day old plant were collected seven days after the removal Pi from the hydroponic nutrient solution at approximately 9.30 a.m, frozen in liquid nitrogen, and stored at $-80\text{ }^\circ\text{C}$ until subsequent analyses.

Five lines of *B. napus* were selected (BA-044 (Laser), BA-099 (TAPIDOR DH), BA-210 (RED RUSSIAN), BA-218 (GROENE GRONINGER SNIJMOES) and BA-224 (WILD ACCESSION)) for detailed RNA sequencing based on their high responsive of DGDG/PC ratio in lipid analysis (Chapter 5). Another hydroponics similar set up was conducted on five selected lines of *B. napus* and *B. rapa* R-o-18 with three biological replications for each line.

6.3.2 Total RNA extraction

Total leaf RNA was extracted from 24 lines of *B. napus* and *B. rapa* and purified using Spectrum™ Plant Total RNA kit (Sigma-Aldrich). On-Column DNase 1 Digestion set (Sigma-Aldrich) was used to remove any contaminating DNA. The quality of extracted RNA was checked and quantified using a Nanodrop 2000 spectrophotometer (Thermo Scientific) and Qubit 3.0 fluorometer (Life Technologies). The integrity of RNA was assessed through electrophoresis 1.5% denaturing formaldehyde gel.

6.3.3 RT-PCR: Pi responsive target genes and primer design

Target Pi responsive genes from *B. napus* (*Brassica_napus_assembly_v1.0*) (Table 6.1) and lipid metabolism genes from *B. napus* (*Brassica_napus_assembly_v1.0*) and *B. rapa* (*Brassica_rapa_Chiifu_assembly_v2.0*) were identified using the synteny search function in the BRAD database using the Arabidopsis gene ID (Cheng et al., 2011). This search identified the paralogous copies in the LF, MF1 and MF2 sub-genomes for *B. napus*. The target sequences were aligned using ClustalW2 software for each gene to identify regions with sequence variation. Primers were then designed specific to either the A or the C genome of *B. napus* using Primer 3 software and the secondary structure of the primer (hairpin and self- dimer) was checked and analysed using Net Primer software. All the primers were designed to distinguish between cDNA and gDNA.

Since there are potentially six paralogous copies of individual *Arabidopsis* genes in *B. napus*, it is critical to design primers that are specific to individual paralogs to ensure accurate quantification of gene expression. Fifteen sets of primers were designed for six target genes (Table 6.2), while nine sets for lipid genes for *B. napus* and *B. rapa* (Table 6.3 and 6.4, respectively). Several attempts have been necessary to optimise the primer design to ensure single products are produced from each primer pair.

The PCR amplification was carried out using the following gradient cycle conditions, the reactions were performed in a 10 μ L total volume containing 1 μ L of cDNA, 0.5 μ L of each primer, and 5 μ L of Taq Mix (Sigma Aldrich). The reaction conditions were 95 °C (60 s) for one cycle of initial denaturation, followed by denaturation, annealing, and extension step at 98 °C (15 s), 65 °C (15 s) and 72 °C (15 s), respectively for 40 cycles. A final extension step of 72 °C for 5 min, 4 °C until turned off.

Table 6.1. List of target Pi responsive genes identified from the literature in *Arabidopsis thaliana* and the syntenic analysis of orthologous copies in *Brassica napus* and the paralogous copies within the *Brassica napus* sub-genomes (LF, MF1, and MF2).

Gene	<i>Arabidopsis thaliana</i> Gene ID	<i>B. napus</i> A genome ID			<i>B. napus</i> C genome ID		
		Least Fractionated (LF)	Most Fractionated 1 (MF1)	Most Fractionated 2 (MF2)	Least Fractionated (LF)	Most Fractionated 1 (MF1)	Most Fractionated 2 (MF2)
<i>PHT1;5</i>	AT2G32830			GSBRNA2 T00141984001			GSBRNA2 T00132254001
<i>PP2A-1</i>	AT1G59830			GSBRNA2 T00081566001			GSBRNA2 T00046442001
<i>SSP4</i>	AT5G46410	GSBRNA2 T00054716001			GSBRNA2 T00128612001		
<i>GPP1</i>	AT4G25840		GSBRNA2 T00075333001_1 GSBRNA2 T00075332001_2			GSBRNA2 T00154555001	
<i>SQD2</i>	AT5G01220	GSBRNA2 T00069492001 GSBRNA2 T00069491001 _3	GSBRNA2 T00049228001_1			GSBRNA2 T00134155001	
<i>SPX2</i>	AT2G26660			GSBRNA2 T00138787001			GSBRNA2 T00066167001

Table 6.2. List of the primers used in this study, the location of Pi responsive target genes in chromosome and the length of the amplified fragments (cDNA and gDNA).

Symbol	<i>B. napus</i> gene	Sequence (5' to 3')	Chromosome	Length cDNA (bp)	Length gDNA (bp)
PHT1;5_A	GSBRNA2T00141984001F	AACTTATTATTGGCGGAT	A03/MF2	130	1872
	GSBRNA2T00141984001R	AATCTCCGTCCCATCACCTC			
PHT1;5_C	GSBRNA2T00132254001F	CTGCCGCCGATATGTGTAAG	C03/MF2	201	1659
	GSBRNA2T00132254001R	TTTGGGAAGAGGTTGCTGCTG			
PP2A1_A	GSBRNA2T00046442001F	TCACGGACCTCTTTGATTACC	C09/MF2	239	469
	GSBRNA2T00046442001R	TGCAATGTCCTGTCCAAACG			
PP2A-1_C	GSBRNA2T00081566001F	AGTGCTTGAGGAAATACGGC	A09/MF2	198	425
	GSBRNA2T00081566001R	AATCGCACATTGGTCCTTCG			
SSP4_A	GSBRNA2T00054716001F	AGATGCTTCTCAGTCTCAGG	A06/LF	219	358
	GSBRNA2T00054716001R	ATCTTGTTGCTGCTCTCGTC			
SSP4_C	GSBRNA2T00128612001F	ATTCTGACATAGAGAGCCA	C07/LF	152	152
	GSBRNA2T00128612001R	CTGTGGTTGTGATGCTTTCC			
GPP1_A1	GSBRNA2T00075333001F	TCCCATACAGCCAGACTACAG	A03/MF1	289	407
	GSBRNA2T00075333001R	CTGGTGTTTGAAGATGCTC			
GPP1_A2	GSBRNA2T00075332001F	GCCTCTCTGCTACATTTT	A03/MF1	129	129
	GSBRNA2T00075332001R	TAAAACCGAGCCAGTCAAG			
GPP1_C	GSBRNA2T00154555001F	CGAATCGTAGCAGCTATGTC	C07/MF1	200	457
	GSBRNA2T00154555001R	TTCTCCCATCATCTTCGC			
SQD2_A1	GSBRNA2T00069492001F	ACCAAGTAACAACAAGCCA	A10/LF	271	271
	GSBRNA2T00069492001R	CGTGACAAAGGTGAGAAGT			
SQD2_A2	GSBRNA2T00069491001F	GCCAAAGTCCAAAACCACA	A10/LF	274	274
	GSBRNA2T00069491001R	GTGGTTTCTGCTTCGTCTGG			
SQD2_A3	GSBRNA2T00049228001F	CTACTTGTTCCCCCAGGTGTT	A03/MF1	201	201
	GSBRNA2T00049228001R	TCTCCGATTTCGTACCTTCC			
SQD2_C	GSBRNA2T00134155001F	GAAGGTGAGGAATCGGAG	C03/MF1	123	240
	GSBRNA2T00134155001R	ATCTGTTTTTGTATCCTGA			

Table 6.2. continued

Symbol	<i>B. napus</i> gene	Sequence (5' to 3')	Chromosome	Length cDNA (bp)	Length gDNA (bp)
SPX2_C	GSBRNA2T00066167001F	TCAACTCAGAAACCAACG	C03/MF2	112	112
	GSBRNA2T00066167001R	AAAACCCTAAGAGCGGAGAC			

Table 6.3. List of genes involved in lipid metabolism in *Brassica napus*. Primer sequence from 5' to 3', associated Arabidopsis gene ID, chromosome location, predicted amplified cDNA and gDNA fragments, and melting temperature in PCR.

No	<i>B. napus</i> gene	Sequence (5' to 3')	Symbol	<i>Ath</i> gene ID	Chromosome	Length cDNA (bp)	Length gDNA (bp)	Tm (°C)
1	GSBRNA2T00000279001_F	GTCGTCTTCGTCGTCGTC	DGD1_A	AT3G11670	A05	130	130	53.42
	GSBRNA2T00000279001_R	CCACCTCTCCATCACCTTCC						58.82
2	GSBRNA2T00111084001_F	GTCGTGTCGTCGTCCTCCG	DGD1_C	AT3G11670	C05	229	229	57.1
	GSBRNA2T00111084001_R	TCATCTCCAAAACCCCTTCC						58.75
3	GSBRNA2T00121040001_F	CAGGCACACACTCTCCCATC	MGD1_A1	AT4G31780	A01	118	118	58.09
	GSBRNA2T00121040001_R	CCAATCCGCTACAATCTTCG						58.33
4	GSBRNA2T00077751001_F	CTCTACCCGCCCGGAGAACGA	MGD1_A2	AT4G31780	A03	214	214	68.13
	GSBRNA2T00077751001_R	CCCAAACCCCATCAGTC						55.17
5	GSBRNA2T00130492001_F	ACTCTCCCTCTCCACCAATC	MGD1_C	AT4G31780	C01	137	137	55.06
	GSBRNA2T00130492001_R	CCAATCCGCTACAATCTTCG						58.33

Table 6.4. List of genes involved in lipid metabolism in *Brassica rapa*. Primer sequence from 5' to 3', associated Arabidopsis gene ID, chromosome location, predicted amplified cDNA and gDNA fragments, and melting temperature in PCR.

No	Gene	Sequence (5' to 3')	Symbol	Arabidopsis gene ID	<i>B. rapa</i> Chromosome	Length (cDNA)	Length (gDNA)	Tm (°C)
1	Bra037346_F_1	TCACITTCGCTTCCCCTTC	DGD2_A	AT4G00550	A09	312	1083	57.53
	Bra037346_R_1	GCCTTGTTTCTCTCTTTACG						55.9
2	Bra020104_F_1	GCATTCCACACCACATCTCC	MDG2_A	AT5G20410	A02	762	1084	57.93
	Bra020104_R_1	CATACGGCACATTCCCTTTC						57.93
3	Bol011504_F	TAATCCTCTCTCCGTGCTG	DGD2_C	AT4G00550	C09	1000	1717	55.32
	Bol011504_R	GGCTCTAACAAAACCCTTCC						55.84
4	Bol036141_F	TGTGAAGGATGTGTGGAAGG	MDG2_C	AT5G20410	C02	1083	1393	55.97
	Bol036141_R	GCAAAGGAAGAGGTAAGAGTGT						55.07

6.3.4 RNA-sequencing library preparation

Thirty-six leaf RNA samples consisting of three replicates from six lines (BA-044, BA-099, BA-210, BA-218, BA-224 and R-o-18) under P+ and P- conditions were extracted using the same method as in the lipid profile study described previously (Section 6.3.2). RNA samples were checked for contaminants, quality and quantity as described previously (section 6.3.4). High quality total RNA samples were sent to The Wellcome Trust Centre for Human Genetics (WTCHG) at Oxford Genomics Centre to sequence the transcriptome of the 36 samples. The RNA-seq libraries were prepared using the TruSeq RNA library preparation kit according to the manufacturer's instructions (Illumina Inc., San Diego, USA). Briefly, the mRNA fraction was selected using poly-A magnetic beads and then fragmented. First strand cDNA is then generated using reverse transcriptase and random hexamers. Next, the second strand of cDNA synthesis incorporated dUTP, followed by fragment end-repair and adapter-ligated to the cDNA. The samples are then purified and amplified to generate a library. The prepared libraries were size selected, multiplexed and quality checked by WTCHG before paired end sequencing was conducted over five lanes of a flow cell using HiSeq4000 with a 75 bp read length according to the manufacturer's instructions (Illumina Inc.). Around 36-46 Gb of data was generated from each lane.

6.3.5 Data analysis

All data were quality checked and aligned to the *B. napus* genome sequence version 2.0 using HiSAT2 with default settings and the following files available from NCBI ftp.ncbi.nlm.nih.gov (accessed on 10th October 2017); Genome FASTA file: GCF_000686985.2_Bra_napus_v2.0_genomic.fna.gz and ii) Genome GFF file: GCF_000686985.2_Bra_napus_v2.0_genomic.gff.gz. The "Tuxedo" suite of programs was used to further process the data and define differentially expressed genes (Pertea et al., 2016). Briefly, sorted BAM files from HiSAT were processed using the script as shown

in Appendix 2. The StringTie program was used to generate transcript count data for aligned sequences and these were subsequently used to generate a merged GTF file of *B. napus* transcripts in the data set. The merged GTF was then used to generate a final read count for transcripts normalised to Fragments Per Kilobase of transcript per Million mapped reads (FPKM). These data were then read in to R and Ballgown programme was used to identify significantly differentially expressed genes and transcripts, between treatments across the whole data set and for individual lines. Genes or transcripts with q value <0.05 and fold change value <0.5 or >2 were defined as significantly downregulated or upregulated respectively. A q value or False Discovery Rate (FDR) is the corrected p value to adjust for false positive results that may happen by chance as a result of performing multiple statistical tests. The annotations for genes were obtained from Entrez Batch program using the NCBI identifier.

6.4 RESULTS

6.4.1 Determination of RNA quantity and quality of 24 lines of *B. napus* for subsequent RT-qPCR

Leaf RNA was extracted, the concentration was quantified, and the purity assessed using a Nanodrop 2000 spectrophotometer and Qubit 3.0 fluorometer (Table 6.5). RNA concentration varied, ranging from 144 ng μL^{-1} to 964 ng μL^{-1} . The integrity of the RNA was examined using gel electrophoresis (Fig. 6.1). Each sample showed two bands of rRNA (28S and 18S), indicating that the RNA had not been degraded and was suitable for downstream analysis. Assessment of the purity of the RNA suggested that it was suitable to proceed with RT-PCR analysis or RNA sequencing, with $\text{Abs}_{260/280}$ and $\text{Abs}_{260/230}$ within the range of valid recommendations ($\text{Abs}_{260/280} = 1.7\text{-}2.2$ and $\text{Abs}_{260/230} \geq 2.0$). All RNA samples were used for further downstream experiment. However, only selected lines were chosen for RNA-seq analysis.

Table 6.5. RNA quantity and quality of 48 leaf RNA samples from 24 *Brassica napus* genotypes, BnASSYST (BA) grown hydroponically under Pi deficient (P-) or Pi sufficient (P+) conditions. RNA quantity and quality were determined using a Qubit 3.0 flourometer and a NanoDrop 2000 spectrophotometer.

Gel sample	Line	Treatment	Concentration (ng μL^{-1})	Abs _{260/280}	Abs _{260/230}	Gel sample	Line	Treatment	Concentration (ng μL^{-1})	Abs _{260/280}	Abs _{260/230}
1	BA-040	P+	294	2.08	2.20	25	BA-040	P-	472	2.10	2.24
2	BA-044	P+	334	2.03	2.17	26	BA-044	P-	452	2.10	2.23
3	BA-048	P+	572	2.09	2.26	27	BA-048	P-	342	2.10	2.14
4	BA-067	P+	780	2.09	2.27	28	BA-067	P-	656	2.07	2.21
5	BA-089	P+	144	2.08	2.26	29	BA-089	P-	840	2.14	2.25
6	BA-099	P+	552	2.07	2.27	30	BA-099	P-	964	2.13	2.35
7	BA-101	P+	152	2.09	2.26	31	BA-101	P-	744	2.10	2.28
8	BA-102	P+	394	2.09	2.28	32	BA-102	P-	512	2.10	2.09
9	BA-113	P+	362	2.08	2.27	33	BA-113	P-	346	2.09	2.19
10	BA-154	P+	584	2.09	2.21	34	BA-154	P-	572	2.10	2.22
11	BA-158	P+	780	2.11	2.28	35	BA-158	P-	256	2.10	2.14
12	BA-177	P+	374	2.09	2.22	36	BA-177	P-	250	2.10	2.22
13	BA-185	P+	680	2.11	2.28	37	BA-185	P-	396	2.11	2.28
14	BA-202	P+	956	2.12	2.27	38	BA-202	P-	620	2.03	2.07
15	BA-204	P+	804	2.09	2.25	39	BA-204	P-	232	2.05	1.76
16	BA-208	P+	748	2.11	2.29	40	BA-208	P-	222	2.11	2.22
17	BA-209	P+	824	2.12	2.28	41	BA-209	P-	302	2.10	2.22
18	BA-210	P+	224	2.09	2.20	42	BA-210	P-	284	2.09	2.27
19	BA-216	P+	556	2.10	2.27	43	BA-216	P-	452	2.10	2.25
20	BA-218	P+	262	2.08	2.23	44	BA-218	P-	210	2.10	2.29
21	BA-224	P+	220	2.05	2.19	45	BA-224	P-	552	2.10	2.21
22	BA-410	P+	342	2.06	2.21	46	BA-410	P-	166	2.08	2.22
23	BA-418	P+	488	2.06	2.19	47	BA-418	P-	460	2.10	2.26
24	BA-426	P+	412	2.07	2.16	48	BA-426	P-	200	2.10	2.37

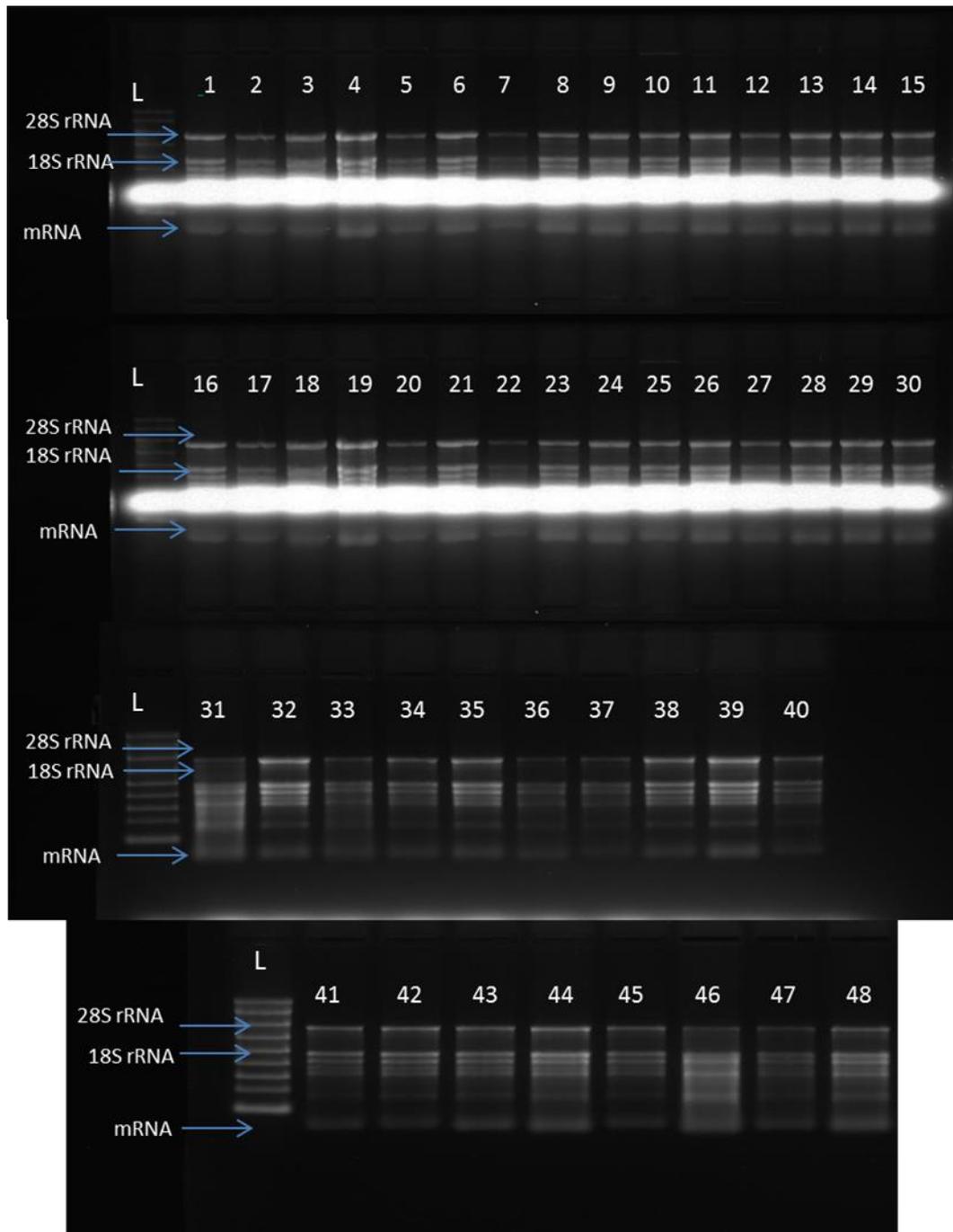


Figure 6.1. Determination of RNA integrity using 1.5% denaturing formaldehyde gel electrophoresis for RNA samples isolated from the leaves of 24 different *Brassica napus* lines grown hydroponically under Pi deficient and Pi sufficient conditions. 1-48 = sample number indicated in Table 6.5, L= NEB ssRNA ladder.

6.4.2 Determination of total RNA quantity and quality extracted from selected lines of *B. napus* and *B. rapa* R-o-18 for RNA-seq

A total of 42 leaf RNA samples were extracted from initial six lines of *B. napus* based on highly membrane lipid metabolism response. Due to limited seed availability for further analysis, BA-113 was omitted for further analysis (RNA sample number 3,10,17,24,31 and 28; Fig.6.2). These samples were consequently not analysed for RNA quantity. Leaf RNA concentration was quantified, and the purity assessed using Nanodrop 2000 spectrophotometer and Qubit 3.0 fluorometer (Table 6.6). High concentration of RNA was extracted from all samples and ranged from 221 to 1544 ng μL^{-1} . The integrity of the RNA was examined using 1.5% denaturing formaldehyde gel electrophoresis (Fig. 6.2). Each sample shows two bands of rRNA (28S and 18S), indicating no degradation of the sample and suitable for further downstream analysis. Assessment of the purity of the RNA suggested that it was suitable to proceed with RNA sequencing, with $\text{Abs}_{260/280}$ and $\text{Abs}_{260/230}$ within the range of valid recommendations ($\text{Abs}_{260/280}$ = 1.7-2.2 and $\text{Abs}_{260/230}$ \geq 2.0).

Table 6.6. RNA quantity and quality of total RNA extracted from the leaves of five lines of *Brassica napus* and *Brassica rapa* R-o-18 plants grown hydroponically for 30 days under P sufficient (P+) or deficient conditions (P-). Samples were analysed using a NanoDrop 2000 spectrophotometer and Qubit 3.0 fluorometer.

Sample number	Line name	Treatment	Replicate	Nanodrop concentration (ng μL^{-1})	Abs _{260/280}	Abs _{260/230}	Qubit concentration (ng μL^{-1})
1	BA-044	P+	1	594.9	2.12	2.31	644
2	BA-099	P+	1	335.5	2.10	2.23	366
4	BA-210	P+	1	1544.0	2.13	2.28	1428
5	BA-218	P+	1	592.4	2.13	2.3	740
6	BA-224	P+	1	926.0	2.12	2.27	1112
7	R-o-18	P+	1	819.7	2.12	2.25	888
8	BA-044	P-	1	750.5	2.13	2.31	664
9	BA-099	P-	1	512.5	2.15	2.38	620
11	BA-210	P-	1	364.3	2.08	2.24	408
12	BA-218	P-	1	615.3	2.11	2.32	604
13	BA-224	P-	1	597.7	2.12	2.35	744
14	R-o-18	P-	1	672.2	2.14	2.28	796
15	BA-044	P+	2	221.0	2.10	2.25	290
16	BA-099	P+	2	679.8	2.10	2.23	948
18	BA-210	P+	2	851.2	2.11	2.31	1116
19	BA-218	P+	2	705.7	2.10	2.26	780
20	BA-224	P+	2	734.0	2.09	2.24	1348
21	R-o-18	P+	2	547.8	2.05	2.18	596
22	BA-044	P-	2	808.0	2.09	2.2	920
23	BA-099	P-	2	515.3	2.05	2.19	664
25	BA-210	P-	2	555.0	2.09	2.32	616
26	BA-218	P-	2	507.4	2.07	2.21	660

Table 6.6. continued

Sample number	Line name	Treatment	Replicate	Nanodrop concentration (ng μL^{-1})	Abs_{260/280}	Abs_{260/230}	Qubit concentration (ng μL^{-1})
27	BA-224	P-	2	725.5	2.10	2.33	944
28	R-o-18	P-	2	697.4	2.09	2.28	576
29	BA-044	P+	3	553.1	2.10	2.23	612
30	BA-099	P+	3	817.8	2.10	2.21	888
32	BA-210	P+	3	831.4	2.08	2.25	884
33	BA-218	P+	3	741.2	2.09	2.25	876
34	BA-224	P+	3	776.0	2.10	2.26	660
35	R-o-18	P+	3	704.6	2.10	2.26	788
36	BA-044	P-	3	466.2	2.09	2.22	556
37	BA-099	P-	3	404.1	2.10	2.22	516
39	BA-210	P-	3	649.5	2.12	2.27	700
40	BA-218	P-	3	257.3	2.12	2.27	338
41	BA-224	P-	3	416.7	2.09	2.22	516
42	R-o-18	P-	3	375.6	2.10	2.26	460

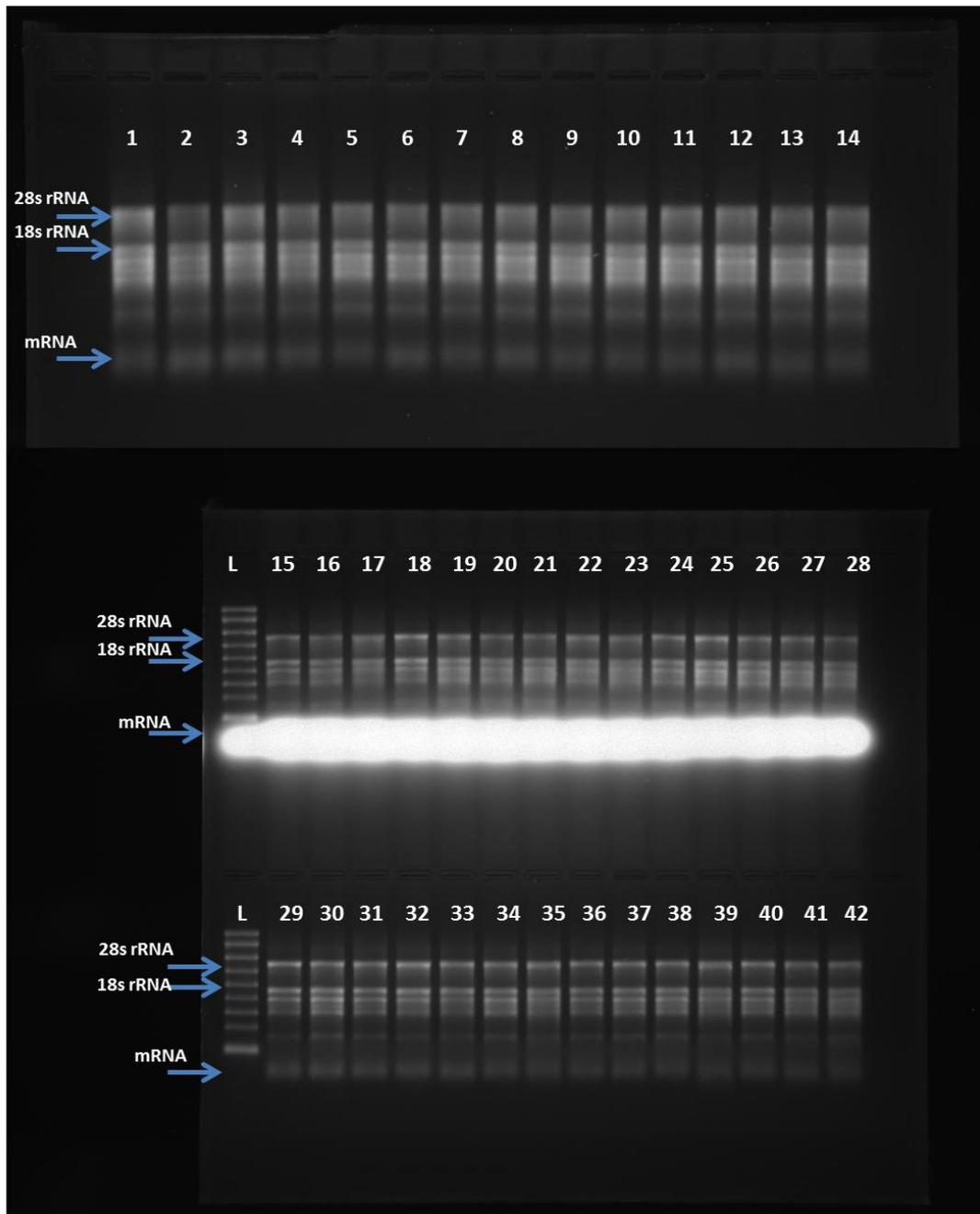


Figure 6.2. Determination of RNA integrity using 1.5% denaturing formaldehyde gel electrophoresis of RNA samples isolated from the leaves of five different *Brassica napus* lines and one *Brassica rapa* grown hydroponically under Pi sufficient (P+) and Pi deficient (P-) conditions. Lanes 1-42 = samples 1-42 defined in Table 6.6, L= NEB ssRNA ladder.

6.4.3 RT-PCR of lipid and other Pi responsive genes

6.4.3.1 Quantification of Pi responsive genes in *B. napus*

Specificity of each primer designed was tested through PCR amplification of a target sequence using leaf cDNA of *B. napus*. Primer pairs should be unique and only bind to a specific location in the genome. The PCR products were then run on an electrophoresis gel for visualisation. From the gel image (Fig. 6.3), primers designed to PHT1:5_A showed a potential primer dimer with a length that slightly was lower than the target of 130 bp. For primers designed to PHT1:5_C, multiple bands were amplified, with the occurrence also of a potential primer dimer. Primers designed to SSP4_C gave an accurate PCR product (152 bp) for sample 2 (column 2) but also formed a primer dimer (Fig. 6.3). Correct PCR product sizes were formed for GPP1_A2, GPP1_C and SQD2_A1 with bands at 129 bp, 200 bp and 271 bp, respectively (Fig. 6.3). No clear PCR product was observed for SQD2_A2, SQD2_C. with an incorrect band size or multiple bands produced respectively. No clear PCR products, other than potential primer dimers, were produced for SPX2_A. A clear PCR product of the correct size (112 bp) did amplify for SPX2_C (Fig 6.3).

The other five primer pairs were tested using the same PCR amplification method (Fig. 6.4). PP2A1_C, PP2A1_A, SSP4_C and SQD2_A3 showed the right size of amplification on six shoot cDNA samples. GPP1_A1 showed primer dimer occurred during PCR amplification and not binding to the predicted size of 289 bp. Several attempts have been made to redesign and rerun the primer on PCR to get the specific target sequence. Only eight from fifteen primer pairs work and bind accordingly.

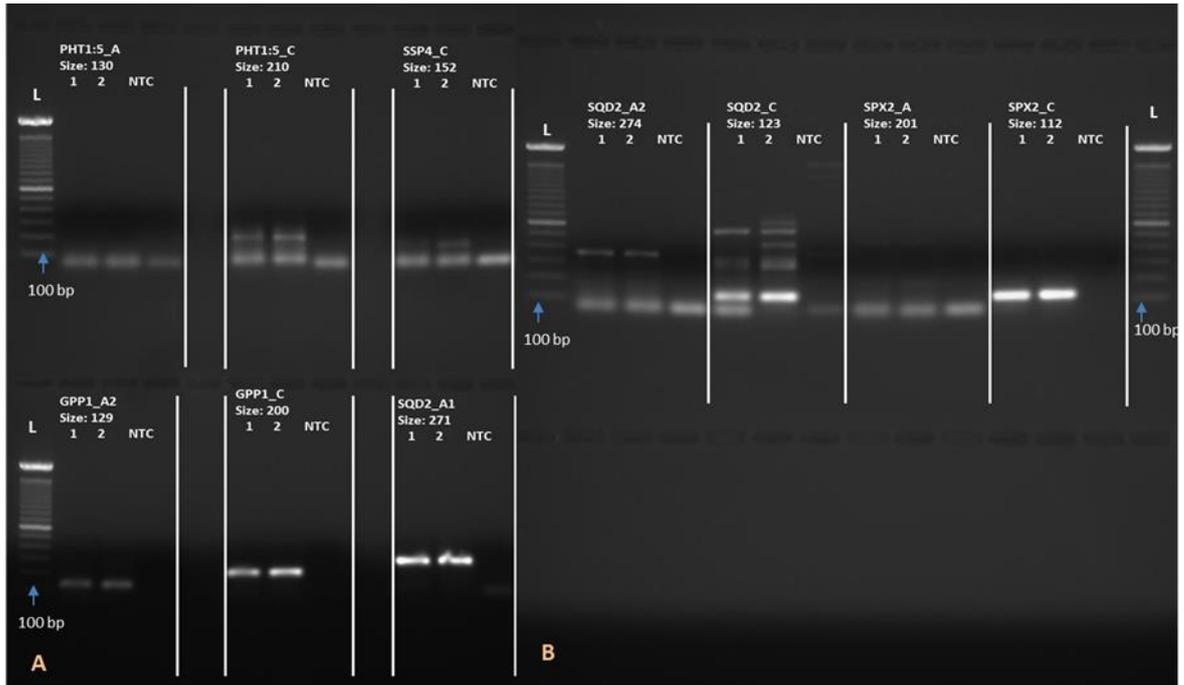


Figure 6.3. Electrophoresis gel image of PCR products to test the specificity of primers designed to Pi responsive genes in *Brassica napus*. **A.** Gel of PCR products for PHT1:5_A, PHT1:5_C, SSP4_C, GPP1_A2, GPP1_C, and SQD2_A1 with their predicted product sizes. **B.** Gel of PCR products for SQD2_A2, SQD2_C, SPX2_A and SPPX2_C with their predicted product sizes. Column 1 and 2 represent different leaf cDNA samples BA-099 and BA-101, respectively; NTC = no template control (water); L= 100 base pair (bp) NEB ladder.

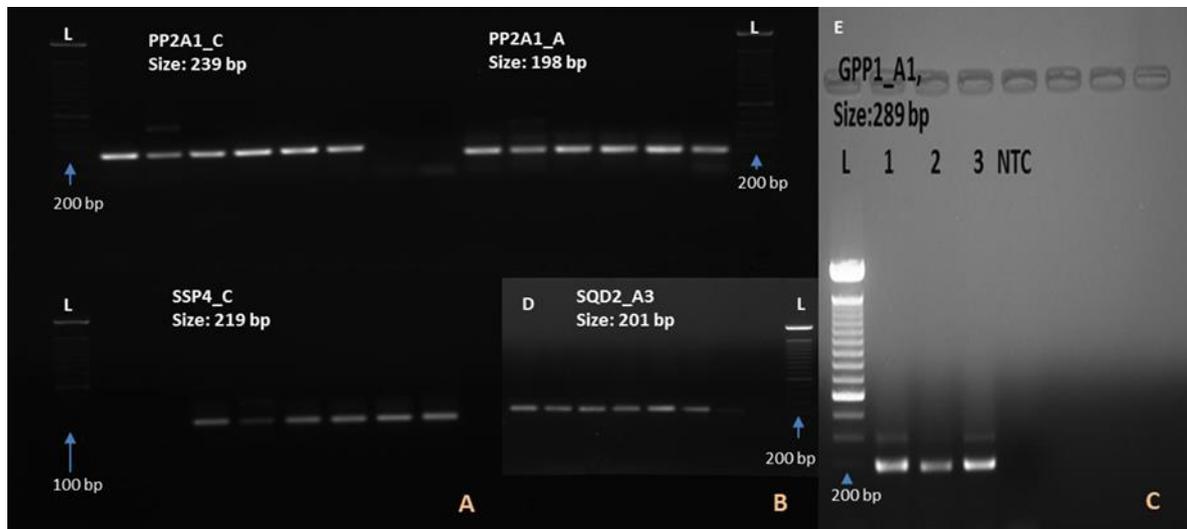


Figure 6.4. Electrophoresis gel image of PCR products to test the specificity of primers designed to Pi responsive genes in *Brassica napus*. **A.** Gel of PCR products for PP2A1_C, PP2A1_A and SSP4_C with their predicted sizes in leaf cDNA sample of BA-099. **B.** Gel of PCR products for SQD2_A3 with its predicted size in leaf cDNA sample of BA-099. **C.** Gel of PCR product for GPP1_A1 with its predicted size. Column 1 and 2 represent different leaf cDNA samples BA-099 and BA-101, respectively, NTC= Non-template control column (water), L= 100 base pair (bp) NEB ladder.

6.4.3.2 Identification of lipid genes

Five sets of primers were also designed to specifically amplify cDNAs from additional genes involved in lipid metabolism during Pi starvation in *B. napus*, and four sets for *B. rapa*. PCR amplification using primers designed to *B. rapa* genes did not work and were unable to bind to a specific gene with many reactions containing primer dimers and multiple PCR products (Fig.6.5). PCR primers designed to amplify DGDG1_A did produce a single suitably sized (130 bp) product (Fig. 6.6). PCR primers designed to specifically amplify DGD1_C, MDG1_A1, MDG1_A2 and MDG1_C however showed multiple products on the gel (Fig. 6.6).

6.4.4 Syntenic genes of *B. napus*

The whole genome triplication (WGT) events in *Brassica* spp. resulted in subgenome dominance, where one subgenome retained more genes and/or more highly expressed gene copies of paralogous pairs through fractionation (Fang et al., 2012; Cheng et al., 2015). Therefore, there are potentially six copies of an individual Arabidopsis gene in the allopolyploid *B. napus* caused by hybridization of A and C genome and WGT (Cheng et al., 2014). Three possibly come from subgenome A and another three from subgenome C of *B. napus*. Syntenic gene analysis of five selected target genes, PHT1;5, PP2A-1, SSP4, GPP1, SQD2, and SPX2 showed the occurrence of multiple copies of these genes (Table 6.1). PHT1;5 and PP2A-1 showed two copies; MF2 of genome A and genome C. SSP4 showed two copies in LF of genome A and genome C. GPP1 showed one copy in MF1 genome A, in SQD2 there was one copy in LF genome A and SPX2 showed two copies in MF2 of genome A and genome C. These genes are paralogs and have many similarities in their sequences. After several attempts to design specific primers to these target genes for use in quantitative PCR to evaluate the transcript abundance of these genes in

response to low Pi availability, it was decided not to continue. The sequences used to design the primers are from different accessions to the cDNA used for testing the PCR amplification, which may explain the challenges incurred whilst trying to develop gene specific primers. Another reason for this might be that at the time experiment was undertaken, the *B. napus* and *B. rapa* genomes were both in their first drafts (*B._napus_assembly_v1.0*) and could contain errors in their assembly compared to the latest assembly (*B._napus_assembly_v2.0*). Subsequent analysis of gene expression was conducted using RNA-seq instead of qPCR analysis.

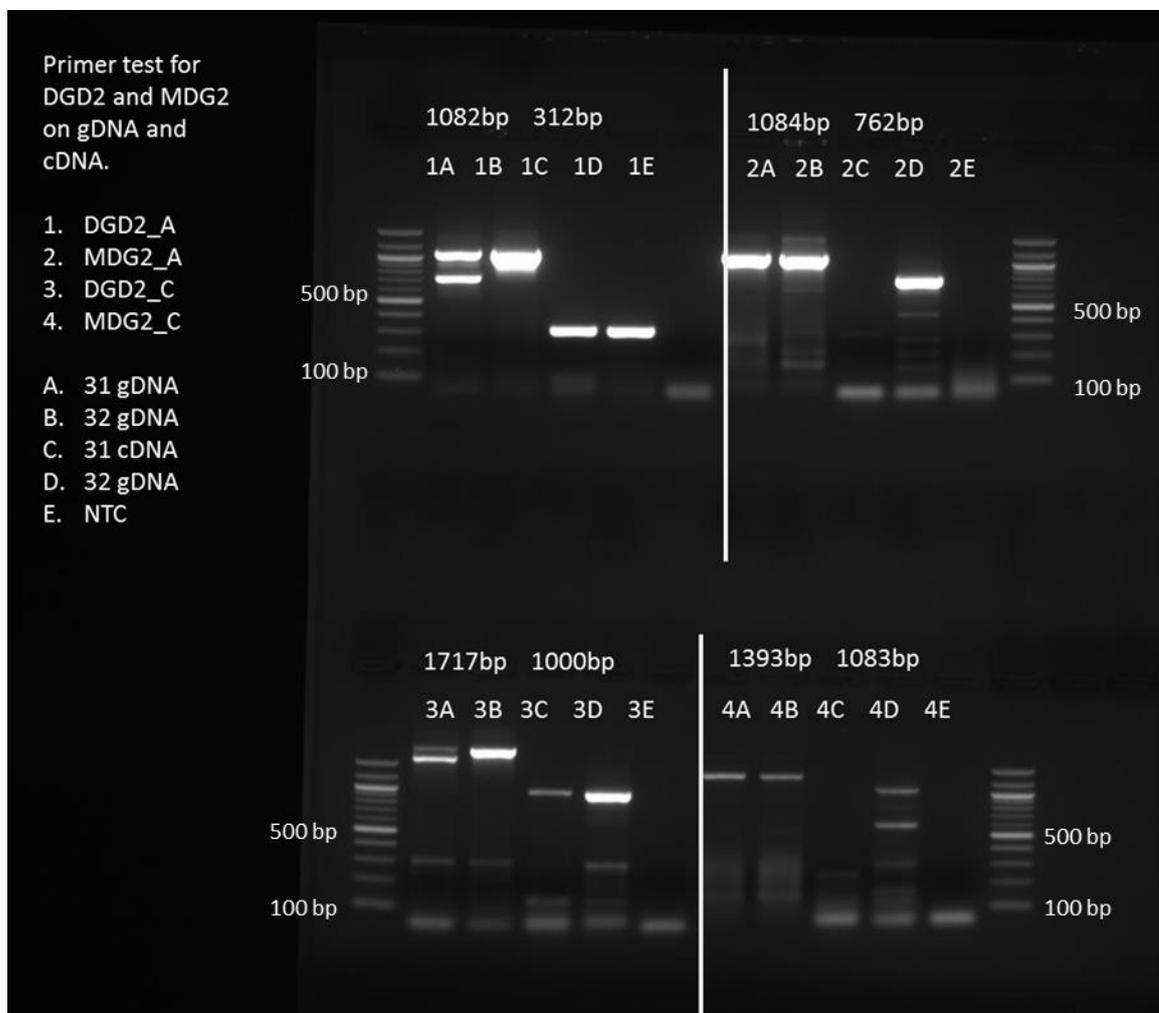


Figure 6.5. Electrophoresis gel image of PCR products to test the specificity of primers designed to Pi responsive genes *DGD2* and *MDG2* of A and C genomes (1-4) using gDNA (A and B) and cDNA (C and D) from two lines of *Brassica napus* (31= BA-099, 32= BA-

101) with their expected sizes. L= 100 base pair (bp) NEB ladder, NTC= No template control (water).

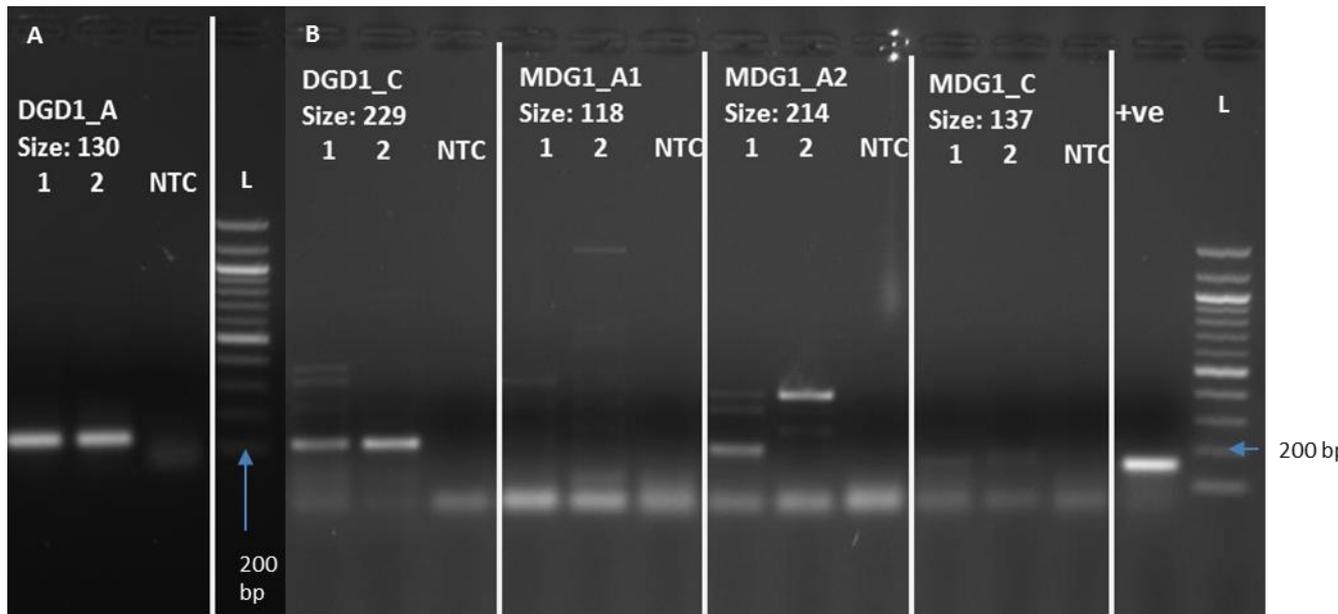


Figure 6.6. Electrophoresis gel image of PCR products to test the specificity of primers designed to lipid metabolism genes from *Brassica napus*. **A.** DGD1_A with expected size 130 bp **B.** DGD2_C, MDG1_A1, MDG1_A2 and MDG1_C, with expected size of 229 bp, 118 bp, 214 bp and 137 bp, respectively. Column 1 and 2 represent different leaf cDNA samples BA-099 and BA-101, respectively, NTC = No template control (water), +ve = positive control (*Brassica rapa* Bro18004703 with size of 160 bp), L= 100 base pair (bp) NEB ladder.

6.4.5 RNA-seq analyses

6.4.5.1 Differential gene expression analysis

RNA-seq transcriptome analysis of genotypes under P+ and P- conditions was conducted. A total of 36 RNA libraries of leaf sample from six lines of (BA-218, BA-102, BA-044, BA-210, BA-099 and R-o-18) were generated at the Oxford Genomics Centre (Table 6.6). Three biological replicates of two treatments (P+ and P-) were obtained from hydroponically grown *Brassica* spp. The libraries were sequenced by Illumina high-throughput sequencing technology. The results from the library were mapped to *B. napus* genome (only five genotypes of *B. napus* were analysed from this point). The abundance of each gene was determined by FPKM (Fragments Per Kilobase of transcript per Million mapped reads) using StringTie. The median values of $\text{Log}_2(\text{FPKM}+0.0004)$ among different lines, and different P conditions, were used for differential expression (Fig 6.7 and 6.8, respectively) using Ballgown in R. Box plot analysis showed that 30 samples of five lines in *B. napus* had similar expression profiles after normalisation. It suggests that all the selected sequencing data are suitable for further transcriptomic variation analysis.

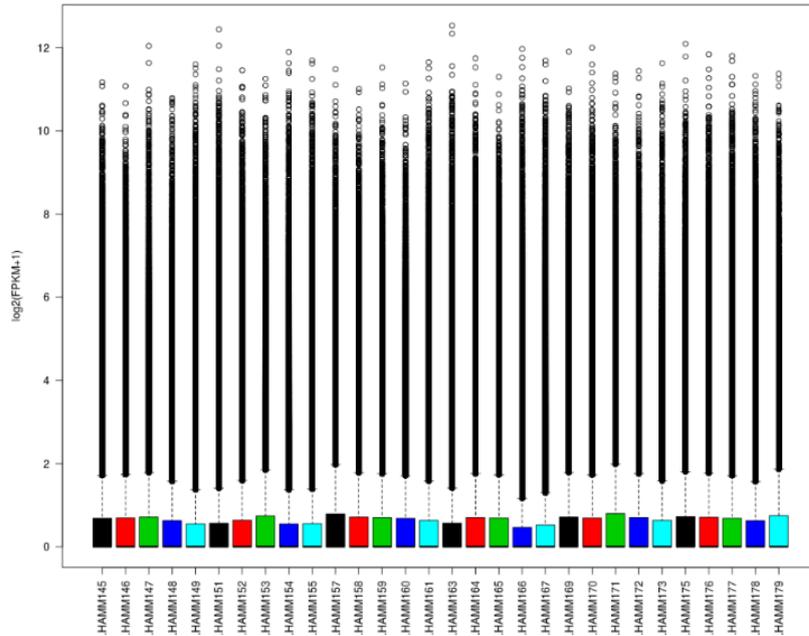


Figure 6.7. Gene expression distribution boxplot of \log_2 (FPKM+1) transformed values against cultivars. The X-axis represents the different samples; same sample with the same colour-code (black, red, green, blue, turquoise represents BA-044, BA-099, BA-210, BA-218, BA-224, respectively). The Y-axis represents the relative abundance of \log_2 (FPKM+1).

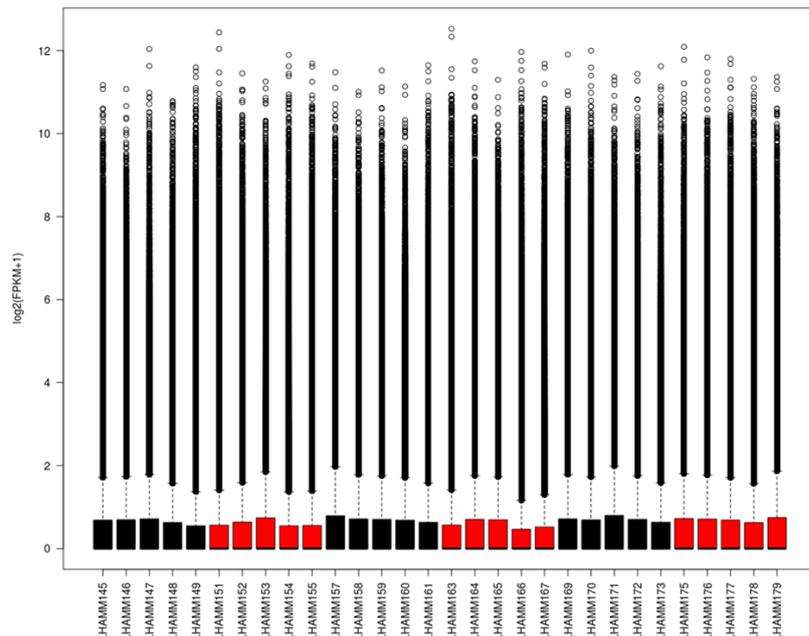


Figure 6.8. Gene expression distribution boxplot of \log_2 (FPKM+1) transformed values against treatment. The X-axis represents the different treatments; same treatment with the same colour-code (Black, red represents P-sufficient, P-deficient, respectively). The Y-axis represents the relative abundance of \log_2 (FPKM+1).

6.4.5.2 MA plot

The MA plot shows the relationship between the change in the expression between Pi replete and Pi deplete conditions (log ratio, M), against the average expression of the genes (average mean, A) (Fig. 6.9). The differential gene expression with adjusted p value <0.05 . This plot suggested P stress caused significant changes in *B. napus* leaf samples.

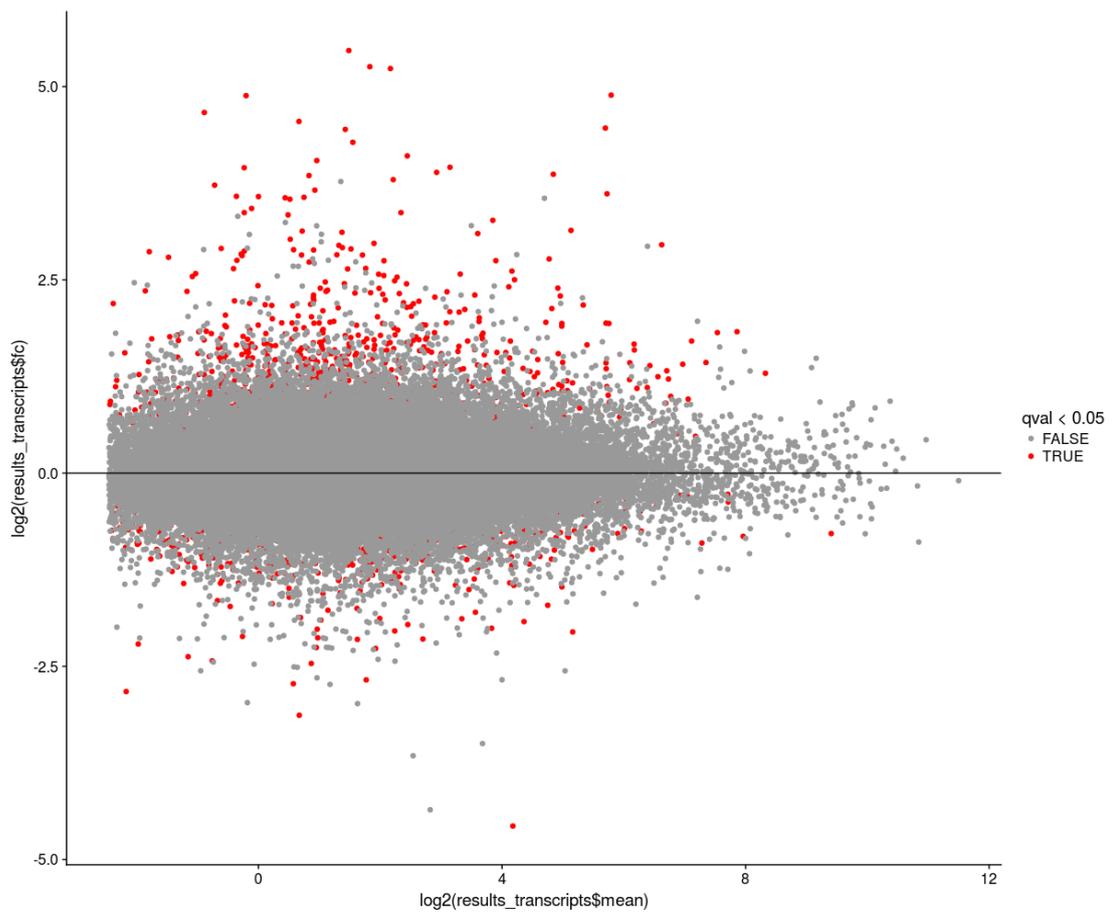


Figure 6.9. The MA plot; the relationship between the expression change (M) and average expression strength (A); genes with adjusted p-values <0.05 are marked in red.

6.4.5.3 Pi response-regulated genes in genotypes under P+ and P- conditions

To compare transcriptome differences between Pi-sufficient and Pi-deficient conditions in five lines of *B. napus*, a total of 629 genes were differentially expressed in response to Pi availability. 481 genes were up-regulated (Appendix 3), while 148 genes were down-regulated across all *B. napus* samples (Appendix 4).

The most upregulated gene encoded an inorganic pyrophosphatase 2-like (43.73-fold) was recorded in the leaf tissues of all investigated genotypes of *B. napus*. Other upregulated genes encoded for an uncharacterized protein (37.97-fold), followed by four types of inorganic pyrophosphatase 1 (37-fold, 31-fold, 26-fold and 24-fold) and sulfoquinovosyl transferase SQD2-like (23-fold). While the most downregulated gene encoded a copper chaperone for superoxide dismutase (SOD) with only 0.14-fold change in expression in P- conditions compared to P+ conditions. The next most down regulated gene encoded a pathogenesis-related protein-5 like with 0.16-fold reduction in expression in P- conditions compared to P+ conditions. Other downregulated genes encoded two uncharacterised proteins BNAC01G03020D (0.16-fold) and BNAC04G02100D (0.22-fold), a gene encoding an AAA-ATPase At2g18193-like protein (0.18-fold), a gene encoding for DETOXIFICATION 1-like protein (0.18-fold), a gene encoding 40S ribosomal protein (0.19-fold) and a vacuolar amino acid transporter YPQ3 (0.19-fold). Variation in the expression of different genes was observed between the different *B. napus* lines (Appendix 5). The highest upregulated genes in BA-044 encoded a remorin like protein with a 773-fold increase in expression in P- condition compared to the P+ condition. While the highest upregulated gene in BA-099, BA-210, BA-218 and BA-224 encoded a probable galactinol-sucrose galactosyltransferase 2 (100-fold), an inorganic pyrophosphatase 2 like (53-fold), a 50S ribosomal protein L35 chloroplastic like (6828-fold), and a glutathione S-

transferase U20 (127-fold), respectively. At the individual line level, many of these increase in expression were no significant however.

6.4.5.4 Variation in the expression of lipid metabolism genes across five cultivars of *B. napus*

Differential expression of transcripts of genes associated with lipid metabolism in response to Pi availability were calculated. A total of 41 lipid-metabolism related transcripts were differentially expressed in response to Pi availability in leaf samples of the five lines of *B. napus* (Table 6.7). Generally, results showed an upregulation of lipid metabolism genes under Pi-stress, including *DGD1*, *DGD2*, *MGD1*, *MGD2*, *MGD3*, non-specific phospholipase C4 (*NPC4*), phosphatidate phosphatase 1 (*PAH1*), phosphatidate phosphatase 2 (*PAH2*), phospholipase D zeta 1-like (*PLDZ1*), phospholipase D zeta 2 (*PLDZ2*), *SQD2*, chloroplastic glycerophosphodiester phosphodiesterase 1 (*GDPD1*), glycerophosphodiester phosphodiesterase 6 (*GDPD6*), UDP-sulfoquinovose synthase (*SQD1*) and UTP--glucose-1-phosphate uridylyltransferase 3 across all the lines.

The expression of transcripts encoding DGD1 all had higher expression under P- compared to P+ conditions in most lines, but the average increase in expression was only 1.24-fold. Interestingly in line BA-210, one transcript (LOC106452493) encoding DGD1 was significantly down regulated, even though the lipid profile suggests this line has a high DGDG/PC ratio.

Overall, all *GDPD1* transcripts were significantly upregulated across all lines of *B. napus* under Pi-stress, with the highest increase in expression recorded in BA-218, with a 164-fold increase in expression under P- conditions compared to P+ conditions (Table 6.7). Based on lipid profile data, BA-218 also recorded the highest level in DGDG/PC ratio under P- conditions, suggesting that *GDPD1*'s high expression in this line might contribute to this high ratio by reducing the PC component of the cellular membranes. Another *GDPD*

family, *GDPD6*, also showed significant upregulation in BA-218, and also BA-044, which had a low DGDG/PC ratio. BA-044 also had the second highest upregulation of one of the *GAPD1* transcripts (Table 6.7), but also maintained one of the highest PC concentrations under P- condition and showed no significant change in leaf PC concentration between P- and P+ conditions (Fig 5.4).

Whilst several transcripts were identified as encoding MGD1, the overall change in expression between P- and P+ showed no consistent pattern and the highest increase in expression was for transcript LOC106423395 in BA-210, with a 4.6-fold increase in expression. Transcripts encoding MGD2 were only detected in two lines (BA-044 and BA-218) with a 13-fold upregulation of expression occurring between P- and P+ conditions (Table 6.7). Both *MDG3* transcripts were highly upregulated under P- conditions in all five lines, with one transcript (LOC106392218) more than 40-fold upregulated in BA-044 and BA-218.

Transcripts for *NCP4* were only detected in one line (BA-218) and transcripts for *PLDZ1* were only detected in two lines (BA-210 and BA-224). Transcripts encoding *PLDZ2* were detected in only one line, but this showed a 30-fold upregulation in P- conditions compared to P+ conditions (Table 6.7).

SQD1 transcripts were mainly upregulated in response to P- treatment, however, only two of the four transcripts were detected in all lines. One transcript showed a 56-fold increase in expression under P- compared to P+ in BA-044. Under P+ conditions, BA-044 had the highest *SQDG* leaf concentration, but this was not maintained under P- (Fig 5.7). *SQD2* transcripts were also highly upregulated in P- conditions compared to P+ conditions (Table 6.7). Four transcripts encoding *SQD2* showed significant upregulation, ranging from 2 to 54-fold, with exception of one line, BA-218, where one *SQD2* transcript (LOC106433829) was 0.3-fold downregulated.

Table 6.7. Expression levels of genes associated with lipid metabolism from leaf samples of five lines of *Brassica napus* (BA-044, BA-099, BA-210, BA-218, and BA-224) and levels of delta ratio galactoclycerolipid/phosphoglycerolipid (DGDG/PC) high (H) and low (L) in response to Pi availability. The expression ratios were calculated as fold changes in transcript abundance between P+ and P-. Expression ratio of 1 indicates no difference in expression, >1 and <1 indicate up-regulated and down-regulated, respectively. Where there are no values, no transcript was detected

			Line_High (H)/Low (L) ΔDGDG/PC ratio				
Symbol	Aliases	Description	BA-044_L	BA-099_L	BA-210_H	BA-218_H	BA-224_L
LOC111215733	<i>DGD1</i>	digalactosyldiacylglycerol synthase 1, chloroplastic	1.420	1.268	1.049	2.342	1.316
LOC106372022	<i>DGD1</i>	digalactosyldiacylglycerol synthase 1, chloroplastic	1.557	1.154	1.286	1.374	1.267
LOC106452493	<i>DGD1</i>	digalactosyldiacylglycerol synthase 1, chloroplastic-like	0.749	1.112	0.255	1.338	1.159
LOC106365523	<i>DGD2</i>	digalactosyldiacylglycerol synthase 2, chloroplastic	12.872	0.614	1.366	5.501	0.719
LOC106434959	<i>DGD2</i>	digalactosyldiacylglycerol synthase 2, chloroplastic-like	2.599	2.172	1.166	1.006	
LOC106393734	<i>GDPD1</i>	glycerophosphodiester phosphodiesterase GDPD1, chloroplastic	9.131	11.941	9.857	164.531	6.872
LOC106436516	<i>GDPD1</i>	glycerophosphodiester phosphodiesterase GDPD1, chloroplastic-like	81.457	8.129	4.284	4.142	36.877
LOC106419105	<i>GDPD6</i>	glycerophosphodiester phosphodiesterase GDPD6	10.436	1.693		6.211	
LOC106372755	<i>MGD1</i>	monogalactosyldiacylglycerol synthase 1, chloroplastic	0.875	0.715	1.475	0.258	
LOC106420497	<i>MGD1</i>	monogalactosyldiacylglycerol synthase 1, chloroplastic	2.843	1.329	2.104	1.407	2.027
LOC106423395	<i>MGD1</i>	monogalactosyldiacylglycerol synthase 1, chloroplastic	1.082	0.859	4.655	1.253	4.242

Table 6.7 continued

Symbol	Aliases	Description	Line_High (H)/Low (L) Δ DGDG/PC ratio				
			BA-044_L	BA-099_L	BA-210_H	BA-218_H	BA-224_L
LOC106354743	<i>MGD1</i>	monogalactosyldiacylglycerol synthase 1, chloroplastic-like					1.798
LOC106375399	<i>MGD1</i>	monogalactosyldiacylglycerol synthase 1, chloroplastic-like	1.079	0.983	2.775	0.702	1.024
LOC106425766	<i>MGD2</i>	monogalactosyldiacylglycerol synthase 2, chloroplastic	2.007			13.576	
LOC106392218	<i>MGD3</i>	monogalactosyldiacylglycerol synthase 3, chloroplastic	40.553	5.233	8.370	43.929	19.475
LOC106366315	<i>MGD3</i>	monogalactosyldiacylglycerol synthase 3, chloroplastic-like	12.101	5.461	7.016	23.589	14.703
LOC106387555	<i>NPC4</i>	non-specific phospholipase C4				2.862	
LOC106432877	<i>NPC4</i>	non-specific phospholipase C4-like					
LOC106452750	<i>PAH1</i>	phosphatidate phosphatase PAH1	1.957	4.254	1.413	0.810	1.688
LOC106427967	<i>PAH1</i>	phosphatidate phosphatase PAH1-like	1.964	0.459	1.324	2.229	
LOC111215781	<i>PAH1</i>	phosphatidate phosphatase PAH1-like					
LOC106349531	<i>PAH2</i>	phosphatidate phosphatase PAH2	2.360	1.496	2.018	0.855	2.217
LOC106447071	<i>PAH2</i>	phosphatidate phosphatase PAH2-like					

Table 6.7 continued

Symbol	Aliases	Description	Line_High (H)/Low (L) Δ DGDG/PC ratio				
			BA-044_L	BA-099_L	BA-210_H	BA-218_H	BA-224_L
LOC106452070	<i>PLDZ1</i>	phospholipase D zeta 1-like					
LOC106401673	<i>PLDZ1</i>	phospholipase D zeta 1-like			1.202		0.652
LOC106382211	<i>PLDZ2</i>	phospholipase D zeta 2					30.277
LOC106427391	<i>SQD2</i>	sulfoquinovosyl transferase SQD2-like	9.653	1.762	2.078	4.143	4.375
LOC106435170	<i>SQD2</i>	sulfoquinovosyl transferase SQD2-like	20.580	25.039	16.359	30.281	7.437
LOC106433829	<i>SQD2</i>	sulfoquinovosyl transferase SQD2-like	0.381		2.559	2.979	2.308
LOC106435863	<i>SQD2</i>	sulfoquinovosyl transferase SQD2-like	19.691	12.428	14.515	43.123	54.905
LOC106364202		UDP-sulfoquinovose synthase, chloroplastic	56.487	2.098	0.479		
LOC106410744		UDP-sulfoquinovose synthase, chloroplastic	4.212	2.566	2.027	6.718	3.973
LOC106423381		UDP-sulfoquinovose synthase, chloroplastic-like					3.796
LOC106368764		UDP-sulfoquinovose synthase, chloroplastic-like	2.383	1.988	1.534	7.224	3.298
LOC106360814		UTP--glucose-1-phosphate uridylyltransferase 3, chloroplastic	4.244	3.517	2.127	5.388	4.288
LOC106368116		UTP--glucose-1-phosphate uridylyltransferase 3, chloroplastic-like	4.031	14.085	2.914	0.001	3.775

6.5 DISCUSSION

Plant adaptations to cope with Pi deficiency involve the coordinated regulation and differential expression of a large number of genes. The expression and genetic variation in these genes vary between species and also within species (Ramaiah et al., 2011). As part of these adaptations to low Pi availability, plants undertake membrane lipid remodelling to conserve Pi by replacing membrane phospholipids with galactolipids and sulfolipids (Nakamura, 2013). Therefore, in this study variations in the expression of lipid metabolism genes across different *B. napus* lines were determined. RNA-seq transcriptome analysis of five lines of *B. napus* grown hydroponically under P+ and P- conditions was conducted by Illumina high-throughput sequencing technology. The five lines which showed high and low lipid changes (DGDG/PC ratio) were selected from earlier experiment in Chapter 5. Analysis of the transcriptomes showed differentially expressed genes with 481 up-regulated and 148 down-regulated in response to low Pi availability (Appendix 3 and 4). Functional classification of some of these genes indicated their roles in many physiological and metabolic processes related to the plant growth and development. The remarkably high upregulation of genes encoding inorganic pyrophosphatase 2-like and inorganic pyrophosphatase 1 proteins indicate a potential increase in the irreversible hydrolysis of pyrophosphate to release inorganic phosphate during Pi deficiency (Plaxton & Tran, 2011). The second highest upregulated gene encoded a sulfoquinovosyl transferase SQD2-like protein, which is consistent with membrane lipid remodelling under Pi deficiency (Yu et al., 2002; Okazaki et al., 2013).

The down regulation of a gene encoding a copper chaperone for superoxide dismutase (SOD) localised at the chloroplast and/or cytosolic compound demonstrate the potential for the reduction of SOD activity during Pi deficiency (Wintz & Vulpe, 2002). Down regulation of gene encoding AAA-ATPase At2g18193-like suggests the potential reduction

in ATP breakdown. Interestingly, the downregulation of genes encoding for protein DETOXIFICATION 1-like and protein DETOXIFICATION 1 suggesting the reduction of transmembrane transport activities of toxic compounds, including Cd (Li et al., 2002). Genes encoded 40S ribosomal protein were also downregulated suggesting the reduction of synthesis of proteins under Pi deficiency (Caroll, 2013).

Orthologs of known membrane lipid remodelling genes in Arabidopsis were identified in *B. napus*. The analysis of genes associated with membrane lipid metabolism under Pi stress revealed genes involved in synthesis of non-phosphorus galactolipid and sulfolipid were strongly induced in response to Pi stress in leaf tissues, consistent with previously studies in Arabidopsis, *Hakea prostrata* (Proteaceae), rice (*Oryza sativa*), and chickpea (*Cicer arietinum*) (Essigmann et al., 1998; Hammond et al., 2003; Kobayashi et al., 2009; Lambers et al., 2012; Mehra & Giri, 2016). These results confirmed that transcriptional regulation of membrane lipid remodelling occurs extensively in the leaves of *B. napus* under Pi-starvation (Table 6.7). This includes the upregulation of genes encoding proteins involved in the replacement of membrane phospholipids with galactolipid and sulfolipid and the downregulation of genes involved in the phospholipid synthesis. These observations are consistent with those observed in other species (Lan et al., 2012; Kuppusamy et al., 2014).

Variation in lipid metabolism genes was observed between cultivars. For example, *DGD1* was induced less than two-fold, except for LOC111215733 of BA-218 which showed an increase of 2.3-fold (Table 6.7). Pi starvation increases the accumulation of DGDG by DGD1 synthase in the plastid envelope membranes and is transported to the extraplastidial membranes such as plasma membrane, tonoplast membrane and mitochondria membrane (Andersson et al., 2003; Jouhet et al., 2005). Lipid analysis showed significant increases in DGDG for BA-218 (Fig 5.4) and the higher increase in the

expression of *DGD1* (LOC111215733) suggests that the DGDG increase in this line may in part be related to this higher expression.

Phospholipids are degraded by phospholipase D (PLD), phospholipase C (PLC) and phosphatidate phosphatase (PAP), releasing diacylglycerol to enable the synthesis of glycolipid sulfoquinovosyldiacylglycerol (SQDG) and digalactosyldiacylglycerol (DGDG) which could enhance Pi remobilisation during Pi-stress. These glycolipids are found in the chloroplast, with DGDG also found in the extraplastidic membranes (Jouhet et al., 2004; Andersson et al., 2005). Free phosphate released from these processes are used for essential cellular processes, including DNA and RNA synthesis. Three remarkably high up-regulated lipid genes revealed in this study are glycerophosphodiester phosphodiesterase 1 (*GDPD1*) up to 164-fold, sulfoquinovosyldiacylglycerol 2 (*SQD2*) up to 54-fold and monogalactosyldiacylglycerol synthase (*MGD3*) up to 43-fold. This result consistent with study by Misson et al., 2005, reported that the orthologues of these genes had high fold changes in Arabidopsis grown in a low Pi-media.

The transcripts of *GDPD1* were highly up-regulated in Arabidopsis and rice, subsequently total activity of GDPD activity increases suggesting that GDPD plays a significant role during phospholipid degradation in plants under Pi starvation. (Wang et al., 2006; Cheng et al., 2011). Another study revealed this gene was also transcriptionally active in flowers and siliques (Cheng et al., 2011). GDPD hydrolyses glycerophosphodiester into glycerol-3-phosphate (G-3-P) and alcohols. The G-3-P may then be dephosphorylated by glycerol-3-phosphatase (GPP) to release Pi to maintain Pi homeostasis. GDPD1 was localised in plastids (Cheng et al., 2011). *GDPD1* is involved in hydrolysing glycerophosphodiester in phospholipids to release Pi. For example, breaking down glycerophosphocholine into glycerol-3-phosphate and choline. These results suggested *GDPD1* genes were

transcriptionally active in leaf of *B. napus*, while in Arabidopsis, this gene was highly upregulated in flowers and siliques (Cheng et al., 2011).

Membrane phospholipid contains about 30% of total organic phosphorus (Poirier et al., 1991). In plastids, phosphatidylglycerol (PG) is the only type of phospholipid present in the thylakoid and inner membrane and contributes up to 10% of the total membrane (Block et al., 1983). The bilayer galactolipid, DGDG usually substitutes phospholipids during Pi-starved condition, while the replacement of PG to SQDG is more limited because the functionality of PG in the photosynthetic complexes cannot be fully substituted by SQDG (Cheng et al., 2011). Regulation of *GDPD1* were highly up-regulated across all investigated five lines, BA-044, BA-099, BA-210, BA-218 and BA-224 (Table 6.7). The highest *GDPD1* recorded in BA-218 which showed high rate of DGDG/PC ratio in its lipid profile. These results indicating *GDPD1* were transcriptionally active in leaf tissues and play an important role in Pi recycling processes to maintain Pi homeostasis of P-starved plants.

MGD3 is a type B MGD synthase gene involved in accumulation of non-phosphorus galactolipid MGDG and DGDG in chloroplasts under Pi deficiency (Kobayashi et al., 2009). *MDG3* is involved in cellular response to P starvation, galactolipid metabolic processes as well as glycolipid biosynthetic processes (Kobayashi et al., 2004). The results show that *MDG3* was highly upregulated in the leaf tissue of *B. napus* under Pi deficiency. Other reports in Arabidopsis suggest that most *MDG3* was also highly induced in non-photosynthetic tissue, especially in roots (Kobayashi et al., 2004), however, further work is required to assess that in Brassicas.

The *de novo* pathway of DGDG synthesis starts with the conversion of G-3-P into PA by two acyltransferases. Subsequently the PA is converted into DAG by phosphatidate phosphatase (PAP). Through the recycling pathway, the DAG will be incorporated into

DGDG by the coupled reaction by MGDG synthase and DGDG synthase (Gaude et al., 2008; Kobayashi et al., 2009). The expression levels of *MGD3* were up-regulated across all five lines. The fold-change varied between the lines tested may be due to natural variation among the samples.

SQD2 encoding sulfolipid synthase is involved in sulfolipid biosynthesis. Arabidopsis mutants of *sqd2* showed reduced growth under Pi starvation (Yu et al., 2002). Unlike galactolipids, sulfolipids are anionic in nature because they contain sulphur in their 6-deoxy-6-sulfonate-glucose head group. Consequently, they are a smaller component in the photosynthetic membrane and the least anionic glycerolipid containing sulphur in the membrane lipids (Shimojima, 2011; Boudière et al., 2014). The process of sulfolipid synthesis proceeds in two steps involving two genes *SQD1* and *SQD2* in plastids. First, the forming of UDP-sulfoquinovose from UDP-glucose and sulphate, catalysed by the *SQD1* protein. Second, UDP-sulfoquinovose is transferred to diacylglycerol, catalysed by the *SQD2* protein. In this study, two genes, LOC106435863 and LOC106435170 encoding *SQD2* were highly up-regulated in all five lines (with high and low delta SQDG/PC ratios (Fig 5.9)) tested showing *SQD2* could be a potential gene for further investigation for *B. napus* to cope with Pi deficiency problems. The large amount of enzyme produced by *SQD2* could not only produced diacylglycerol for glycolipid synthase but also essential in the conversion process of diacylglycerol into SQDG in the membrane lipid of *B. napus* under Pi starvation.

Results from this study showed the availability of genetic variation in *B. napus*, together with the expression profile of the genes involved in lipid metabolism. Their transcript profiles in leaves were analysed together with changes in polar lipid fractions under Pi deficiency. Selection of the highly responsive Pi genes could give a good opportunity to breed for crop adaptation to grow in Pi limited conditions. However, there was no clear

pattern of expression in the lipid metabolism genes associated with their lipid profiles. Since the abundance of different lipids reflects a balance between synthesis and degradation, and the variation observed in the expression of these genes, different lines may be achieving their lipid profiles through different combinations of phospholipid degradation and synthesis of alternative lipids. It is also important to remember, that changes in the expression of genes, might not always reflect protein abundance and activity. Therefore, the enzyme activity and abundance of lipid remodelling proteins, would be an important area to investigate in the future.

CHAPTER 7 GENERAL DISCUSSION AND FUTURE RESEARCH PERSPECTIVE

Plants have developed various and complex responses and adaptive mechanisms for acquisition, remobilisation and recycling to maintain adequate cellular concentrations of Pi (Heuer et al., 2017). For example, plants increase the exudation of enzymes, such as phosphatases, and organic acids to enhance the availability of Pi in the rhizosphere and subsequent Pi acquisition under Pi deficiency (Wang et al., 2010). In addition, plants enhance Pi uptake through interactions with microbial species, especially symbiotic arbuscular mycorrhizal (AM) fungi (Suriyagoda et al., 2014). Under Pi deficiency, plants will also experience sophisticated morphological changes, especially in roots, including inhibition in primary root elongation, increase lateral root length and increases in root hair length and density (Bouain et al., 2016). The aim of this thesis was to advance our understanding of some of these adaptations and the genetic regulation of them in the economically important Brassica family.

7.1 GROWTH AND REMOBILISATION OF Pi DURING DEVELOPMENT

To define suitable assay conditions with which to study plants under Pi stress, initial studies determined suitable external compost Pi concentrations to represent low P (0.0 g L⁻¹) and high P (0.225 g L⁻¹) to proceed with experiment to assess *B. napus* physiological responses to Pi availability (Fig 3.5). P1 and P4 in compost showed a significant effect of Pi treatment (Fig 3.7). Generally, Pi treatment and harvesting time affected all growth and yield parameters (biomass, total P content and Pi concentration) (Fig 3.8, Fig 3.9, Fig 3.10). Accumulation of total P content in plants increased until the seed filling stage, which was attributed to the plants remobilising Pi from leaves and stem tissues to developing seeds to maintain growth and development and increase utilisation efficiency (Fig. 3.9). The reduction in the leaf Pi concentration at the maturity stage occurred as the translocation of Pi from leaves to the other plant parts such as grain or seed (Fig. 3.10). Similar observations have been made previously (Rose et al., 2007). The analysis of

biomass and Pi concentration in plant parts, improves our understanding of the plant's nutrient status and the pattern of remobilisation of Pi within the plant. These data are important to understand the nutritional requirements of the plant and consequently the amount and timing of fertilisers that could be required by the plant.

7.2 KEY Pi RESPONSIVE GENES ARE ALSO DIFFERENTIALLY EXPRESSED IN *B. RAPA*

Transcriptional responses in *B. rapa* grown hydroponically in different Pi treatments (high and low-P) showed variable changes in expression in response to low Pi availability (Figure 3.11). Fourteen target genes were investigated in this study. A homologue of *ALMT1* (Bro18 014187) showed the highest increase in expression in response to P-deficiency, followed by *SUC2* (Bro18 004703) and *PAP12* (Bro18 047862) (Fig 3.11). All three genes have previously been shown to increase their expression in response to low Pi availability (Gottwald et al., 2000; Haran et al., 2000; Ligaba et al., 2006; Sasaki et al., 2004; Lei et al., 2011; Wang et al., 2011a). *ALMT1* is an Al³⁺ tolerance gene related to P deficiency, which encodes a malate transporter. Exudation of malate by the roots can release Pi in the rhizosphere subsequently improve the availability of Pi for uptake from the soil (Sas et al., 2001). *SUC2* is involved in loading sucrose into the phloem and may contribute to signalling low Pi availability or concentrations in the shoot (Lei et al., 2011). Secretion phosphatases, such as *PAP12*, are involved in enhancing Pi acquisition by hydrolysing Pi from extracellular Pi monoesters during Pi deficient conditions (Tran et al., 2010).

7.3 EVIDENCE OF QUADRUPLICATION OF A KEY REGULATOR OF PLANT Pi TRANSLOCATION IN *B. RAPA*

Trans-eQTL hotspots, identified previously (Hammond et al., 2011), occurred on chromosome A06 in *B. rapa* and work here revealed four PHO1 homolog 3 (PHO1:H3) genes located in tandem at this location. PHO1 is important in the translocation of Pi from the root to the shoot via the vascular cylinder. In Arabidopsis mutants of *pho1*, only 3-10% of the Pi translocated to the shoot in the wild-type was translocated to the shoots in the *pho1* mutant (Delhaize & Randall, 1995). In Arabidopsis, *PHO1* has ten homologs (*PHO1*; *H1-10*) indicating they may have many other functions other than mediating translocation of Pi. This is supported by other research showing variation in the expression pattern of *PHO1* homologues (Wang et al., 2004).

Analysis from CloneFinder showed the existence of four copies of *PHO1* genes in tandem, however, there was only transcriptional evidence for three of them from short read transcriptomic data (Fig 4.7). To confirm the presence of the four copies, five BACs spanning this region of the genome were obtained from South Korea; KBrB-063F11 (BAC 1), KBrH102C10 (BAC 2), KBrB029J08 (BAC 3), KBrH038K12 (BAC 4), and KBrB003E10 (BAC 5). Two BAC sequences, BAC 1 and BAC 3 showed a consensus with sequence sizes of 184,651 and 185,935, respectively following long read sequencing by nanopore technology (Figure 4.10). Alignment of the sequences confirmed the four copies of genes located in tandem on chromosome A06 *B. rapa* *PHO1* genes. These were cloned to get better insight of the gene and transcript. Thirty plasmid colonies were analysed on PCR. Subsequent detailed analysis of eight plasmid sequences confirms three groups of *PHO1* homolog (or ortholog AT1G14040 in Arabidopsis) are expressed in leaf tissues of *B. rapa*; *PHO1_A*, *PHO1_B* and *PHO1_C*. Five splice variants were also detected for *PHO1_A*, plasmid 2, 7, 23, 26, and 22 with transcript ID XM_009150437.2. Plasmid 12 proved to be *PHO1_B* (XM_018652610) and plasmid 24 proved to be *PHO1_D* (XM_009150438)

(Table 4.12). The information generated here can be used to design PHO1 homolog/splice variants specific primers to establish their expression through PCR for future research. Another future prospect of this study is to functionally characterize the cloned PHO1 ORFs using a suitable expression system.

All PHO1 family harbour SPX and EXS domain in N-terminal and C-terminal, respectively. Both domains are involved in Pi transport and sensing to internal membrane (Hamburger, 2002; Liu et al., 2018). The expression patterns of PHO1 family in plants are diverse (Wang et al., 2004). The predominant expression is in the vascular tissues of leaves, roots, stems and flowers (Wang et al., 2004; Arpat et al., 2012). In this research, at least seven types of PHO1 genes were expressed in the leaf of *B. rapa* R-o-18 under Pi deficiency, including five splice variants of PHO1_A (Table 4.12).

7.4 SIGNIFICANT VARIATION WAS OBSERVED IN LIPID PROFILES AND THE GENES INVOLVED IN LIPID METABOLISM BETWEEN DIFFERENT CULTIVARS OF *B. NAPUS*

Under Pi-stress conditions, plants can undergo membrane lipid remodelling by replacing phospholipids with non-phosphorus galactolipids (DGDG and MGDG) and sulfolipids (SQDG) (Essigmann et al., 1998). The changes in membrane lipid composition release Pi for other essential cellular functions (Lambers et al., 2012). Remodelling pathways mainly involve the hydrolysis of phospholipids to diacylglycerol (DAG), with the DAG then able to act as a substrate for the biosynthesis of galactolipids (Andersson et al., 2003; Jouhet et al., 2004). These processes have been observed in many plant species, but the extent to which this varies within a species has yet to be investigated.

Identification of lipid species, particularly DGDG, SQDG, and PC in *B. napus* lines and *B. rapa* grown hydroponically under P+ and P- conditions was conducted using ESI-MS/MS. ESI-MS/MS works by collecting mass spectra in series, where each spectrum produces

each specific lipid class with a similar polar lipid head group. Twenty-four lines were selected for initial screening of the lipid species and responses to Pi deficiency based on ratio DGDG/PC. Seven lines (six of *B. napus* and *B. rapa* R-o-18) were subsequently selected based on contrasting responses in their DGDG/PC ratio between P+ and P- treatments (Figure 5.3). The increasing amount of DGDG under P- condition indicates an increase in phospholipid PC degradation and/or an increase in DGDG synthesis. Analysis of lipid changes across seven lines of *Brassica* spp. revealed the concentration of lipid types was highly responsive and controlled by Pi availability (Fig. 5.4 and 5.7). The lipid profiles provide clues to many metabolic and biological process underlying lipid composition changes, but it is not possible to understand how these changes are being regulated at the genetic level.

Investigations in to the variation in the gene expression of (i) Pi responsive genes and (ii) membrane lipid metabolism genes were conducted in five lines of *B. napus* and *B. rapa* R-o-18 grown hydroponically under P+ and P- conditions using quantitative PCR (qPCR) and high-throughput RNA-sequencing (RNA-seq). Identification of Pi target genes based on their homology to previously identified Pi responsive genes in Arabidopsis identified five target Pi responsive genes (Mudge et al., 2002; Yu et al., 2002; Lan et al., 2012; Wang et al., 2017). Several time-consuming attempts were required to amplify target genes for qPCR analysis since there are potentially six paralogs for every gene in amphidiploid *B. napus* genes due to whole triplication event that occurred million years ago (Wang et al., 2011b). Therefore, it was not possible to amplify single products for many of the genes, so a different approach was undertaken.

For the five selected lines of *B. napus*, which were selected based on contrasting (highest or lowest) DGDG/PC ratios, RNA samples from leaf tissues were collected and subjected to RNA-seq. Transcriptome differences between P+ and P- was compared. A total of 630

differentially expressed genes were produced, with 481 genes were upregulated and 149 genes were down regulated (Appendix 3 and 4). Analysis of membrane lipid metabolism genes under P stress showed *GDPD1* and *MGD3* were highly upregulated on all five *B. napus* lines, but there was significant variation in expression (Table 6.7). The variation might be attributed to real changes in the expression of these genes as indicated by the RNA-seq analysis but could also result from variation in the underlying sequences of the genes between the *B. napus* lines or a combination of both (Shukla et al., 2015). It is interesting that not all transcripts were detected in some lines (Table 6.7). Further work to sequence these genes across the lines and confirm via qPCR their expression is required to confirm how the variation in expression of these gene might relate to the observed variation in lipid profiles. Gene Ontology (GO) data associated with *GDPD1* confirms it is involved in cellular Pi homeostasis as well as glucophospholipid catabolic processes (Cheng at al., 2011). GO associated with Pi responsive gene, *MGD3*, confirms its involvement in cellular responses to Pi starvation and galactolipid metabolic process (Kobayashi et al., 2009). Another gene which showed a high response to Pi stress in all five lines except in BA-044 is *SQD2*. *SQD2* is involved in synthesis of SQDG to replace PG in the chloroplast (Essigmann et al.,1998).

Overall gene expression patterns in five investigated lines of *B. napus* under Pi-deficient showed the highest abundance and expression, upregulated up to 43-fold was in Inorganic pyrophosphatase (PPase) which is involved in hydrolysing inorganic pyrophosphate to orthophosphate (Appendix 3). Due to time constraint and limitation, detailed analysis of variation of phosphate responsive genes in *B. napus* and their molecular mechanisms will require further work beyond the PhD. Further analysis is required to investigate genes with function related to Pi activities and biological processes in *Brassica* spp.

7.5 CONCLUSION

Results obtained from this thesis could give a better insight in elucidating physiological and molecular mechanism of Pi uptake and utilisation in Brassicas. Highlighting the complex nature of working with polyploid genomes with the added complexity genome duplication, this work highlights the role these have played in generating variation within the species and the evolution of potentially new functions for genes in responses to stress. Manipulating the candidate genes could be used by the plant breeders to develop plants with the ability to use Pi more efficiently and grow at low Pi, a key requirement of crops going forward with the long-term availability of Pi fertilisers still at risk.

REFERENCES

- ADAMS, M. A., BELL, T. L. & PATE, J. S. 2002. Phosphorus sources and availability modify growth and distribution of root clusters and nodules of native Australian legumes. *Plant, Cell & Environment*, 25, 837-850.
- AL-GHAZI, Y., MULLER, B., PINLOCHE, S., TRANBARGER, T. J., NACRY, P., ROSSIGNOL, M., TARDIEU, F. & DOUMAS, P. 2003. Temporal responses of Arabidopsis root architecture to phosphate starvation: Evidence for the involvement of auxin signalling. *Plant, Cell & Environment*, 26, 1053-1066.
- ALLENDER, C. J. & KING, G. J. 2010. Origins of the amphidiploid species *Brassica napus* L. investigated by chloroplast and nuclear molecular markers. *BMC Plant Biology*, 10, 54. Doi: 10.1186/1471-2229-10-54.
- AMES, B. N. 1966. Assay of inorganic phosphate, total phosphate and phosphatases. *Methods in Enzymology*, 8, 115-118.
- ANDERSSON, M. X., LARSSON, K. E., TJELLSTROM, H., LILJENBERG, C. & SANDELIUS, A. S. 2005. Phosphate-limited oat. The plasma membrane and the tonoplast as major targets for phospholipid-to-glycolipid replacement and stimulation of phospholipases in the plasma membrane. *Journal of Biological Chemistry*, 280, 27578-27586.
- ANDERSSON, M. X., STRIDH, M. H., LARSSON, K. E., LILJENBERG, C. & SANDELIUS, A. S. 2003. Phosphate-deficient oat replaces a major portion of the plasma membrane phospholipids with the galactolipid digalactosyldiacylglycerol. *FEBS Letters*, 537, 128-132.
- ARNAUD, C., BONNOT, C., DESNOS, T. & NUSSAUME, L. 2010. The root cap at the forefront. *Comptes Rendus Biologies*, 333, 335-343.
- ARPAT, A. B., MAGLIANO, P., WEGE, S., ROUACHED, H., STEFANOVIC, A. & POIRIER, Y. 2012. Functional expression of *PHO1* to the golgi and trans-Golgi network and its role in export of inorganic phosphate. *The Plant Journal*, 71, 479-491.
- AWAI, K., MARECHAL, E., BLOCK, M. A., BRUN, D., MASUDA, T., SHIMADA, H., TAKAMIYA, H., OHTA, H. & JOYARD, J. 2001. Two types of MGDG synthase genes, found widely in both 16:3 and 18:3 plants, differentially mediate galactolipid syntheses in photosynthetic and nonphotosynthetic tissues in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences USA*, 98, 10960-10965.
- AZEVEDO, C. & SAIARDI, A. 2017. Eukaryotic phosphate homeostasis: The inositol pyrophosphate perspective. *Trends in Biochemical Sciences*, 42, 219-231.
- BAKER, A., CEASAR, S. A., PALMER, A. J., PATERSON, J. B., QI, W., MUENCH, S. P. & BALDWIN, S. A. 2015. Replace, reuse, recycle: Improving the sustainable use of phosphorus by plants. *Journal of Experimental Botany*, 66, 3523-3540.

- BARBA, M., CZOSNEK, H. & HADIDI, A. 2014. Historical perspective, development and applications of Next-Generation Sequencing in plant virology. *Viruses*, 6, 106-136.
- BARI, R., DATT PANT, B., STITT, M. & SCHEIBLE, W. R. 2006. PHO2, microRNA399, and PHR1 define a phosphate-signaling pathway in plants. *Plant Physiology*, 141, 988-999.
- BARRACLOUGH, P. B. 1989. Root growth, macro-nutrient uptake dynamics and soil fertility requirements of a high-yielding winter oilseed rape crop. *Plant and Soil*, 119, 59-70.
- BATES, T. R. & LYNCH, J. P. 1996. Stimulation of root hair elongation in *Arabidopsis thaliana* by low phosphorus availability. *Plant, Cell & Environment*, 19, 529-538.
- BATES, T. R. & LYNCH, J. P. 2001. Root hairs confer a competitive advantage under low phosphorus availability. *Plant and Soil*, 236, 243-250.
- BENNETT, M. D., LEITCH, I. J., PRICE, H. J. & JOHNSTON, J. S. 2003. Comparisons with *Caenorhabditis* (~100 Mb) and *Drosophila* (~175 Mb) using flow cytometry show genome size in *Arabidopsis* to be ~157 Mb and thus ~25% larger than the *Arabidopsis* genome initiative estimate of ~125 Mb. *Annals of Botany*, 91, 547-557.
- BENNING, C. & OHTA, H. 2005. Three enzyme systems for galactoglycerolipid biosynthesis are coordinately regulated in plants. *Journal of Biological Chemistry*, 280, 2397-2400.
- BERTIER, G., HÉTU, M. & JOLY, Y. 2016. Unsolved challenges of clinical whole-exome sequencing: A systematic literature review of end-users' views. *BMC Medical Genomics*, 9, 52. Doi: 10.1186/s12920-016-0213-6.
- BIELESKI, R. L. 1973. Phosphate pools, phosphate transport, and phosphate availability. *Annual Review of Plant Physiology*, 24, 225-252.
- BLENCOWE, B. J. 2006. Alternative splicing: New insights from global analyses. *Cell*, 126, 37-47.
- BLOCK, M., DORNE, A., JOYARD, J. & DOUCE, R. 1983. Preparation and characterization of membrane fractions enriched in outer and inner envelope membranes from spinach chloroplasts: II. Biochemical characterization. *Journal of Biochemical Chemistry*, 258, 13281-13286.
- BONSER, A. M., LYNCH, J. & SNAPP, S. 1996. Effect of phosphorus deficiency on growth angle of basal roots in *Phaseolus vulgaris*. *New Phytologist*, 132, 281-288.
- BOUAIN, N., DOUMAS, P. & ROUACHED, H. 2016. Recent advances in understanding the molecular mechanisms regulating the root system response to phosphate deficiency in *Arabidopsis*. *Current Genomics*, 17, 308-314.
- BOUDIÈRE, L., MICHAUD, M., PETROUTSOS, D., RÉBEILLÉ, F., FALCONET, D., BASTIEN, O., ROY, S., FINAZZI, G., ROLLAND, N., JOUHET, J., BLOCK, M. A. & MARÉCHAL, E. 2014. Glycerolipids in photosynthesis: Composition, synthesis

and trafficking. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1837, 470-480.

- BOZZO, G. G., RAGHOTHAMA, K. G. & PLAXTON, W. C. 2002. Purification and characterization of two secreted purple acid phosphatase isozymes from phosphate-starved tomato (*Lycopersicon esculentum*) cell cultures. *European Journal of Biochemistry*, 269, 6278-6286.
- BRANTON, D., DEAMER, D. W., MARZIALI, A., BAYLEY, H., BENNER, S. A., BUTLER, T., DI VENTRA, M., GARAJ, S., HIBBS, A., HUANG, X., JOVANOVIĆ, S. B., KRSTIĆ, P. S., LINDSAY, S., LING, X. S., MASTRANGELO, C. H., MELLER, A., OLIVER, J. S., PERSHIN, Y. V., RAMSEY, J. M., RIEHN, R., SONI, G. V., TABARD-COSSA, V., WANUNU, M., WIGGIN, M. & SCHLOSS, J. A. 2008. The potential and challenges of nanopore sequencing. *Nature Biotechnology*, 26, 1146-1153.
- BRINCH-PEDERSEN, H., SORESEN, L. D. & HOLM, P. B. 2002. Engineering crop plants: Getting a handle on phosphate. *Trends in Plant Science*, 7, 118-125.
- BUCHER, M. & FABIANSKA, I. 2016. Long-sought vacuolar phosphate transporters identified. *Trends in Plant Science*, 21, 463-466.
- BUHTZ, A., PIERITZ, J., SPRINGER, F. & KEHR, J. 2010. Phloem small RNAs, nutrient stress responses, and systemic mobility. *BMC Plant Biology*, 10, 64. Doi: 10.1186/1471-2229-10-64.
- BUS, A., KÖRBER, N., PARKIN, I. A. P., SAMANS, B., SNOWDON, R. J., LI, J. & STICH, B. 2014. Species- and genome-wide dissection of the shoot ionome in *Brassica napus* and its relationship to seedling development. *Frontiers in Plant Science*, 5. Doi: 10.3389/fpls.2014.00485.
- BUSTOS, R., CASTRILLO, G., LINHARES, F., PUGA, M. I., RUBIO, V., PEREZ-PEREZ, J., SOLANO, R., LEYVA, A. & PAZ-ARES, J. 2010. A central regulatory system largely controls transcriptional activation and repression responses to phosphate starvation in *Arabidopsis*. *PLoS Genetics*, 6. Doi: 10.1371/journal.pgen.1001102.
- CAROLL, A. J. 2013. The *Arabidopsis* cytosolic ribosomal proteome: From form to function. *Frontiers in Plant Science*, 4. Doi: 10.3389/fpls.2013.00032.
- CARTHEW, R. W. & SONTHEIMER, E. J. 2009. Origins and mechanisms of miRNAs and siRNAs. *Cell*, 136, 642-655.
- CARVALHO, G., SCHAFFERT, R. E., MALOSETTI, M., VIANA J. H. M., MENEZES, C. B., SILVA, L. A., GUIMARAES, C. T., COELHO, A. M., KOCHIAN, L. V., VAN EUWIJK, F. A. & MAGALHAES, J. V. 2016. Back to acid soil fields: The citrate transporter SbMATE is a major asset for sustainable grain yield for sorghum cultivated in acid soils. *Genes Genome Genetics*, 6, 475-484.
- CHEN, Y. F., LI, L. Q., XU, Q., KONG, Y. H., WANG, H. & WU, W. H. 2009. The WRKY6 transcription factor modulates *PHOSPHATE1* expression in response to low Pi stress in *Arabidopsis*. *The Plant Cell*, 21, 3554-3566.

- CHEN, Z. H., NIMMO, GILLIAN A., JENKINS, GARETH I. & NIMMO, HUGH G. 2007. BHLH32 modulates several biochemical and morphological processes that respond to P(i) starvation in *Arabidopsis*. *Biochemical Journal*, 405, 191-198.
- CHENG, F., WU, J. & WANG, X. 2014. Genome triplication drove the diversification of *Brassica* plants. *Horticulture Research*, 1, 14024. Doi: 10.1038/hortres.2014.24.
- CHENG, F., WU, J., LIU, B. & WANG, X. 2015. Genome evolution after whole genome triplication: the subgenome dominance in *Brassica rapa*. In: WANG, X. & KOLE, C. (eds.) *The Brassica rapa Genome*. Berlin, Heidelberg: Springer-Verlag.
- CHENG, Y., ZHOU, W., EL SHEERY, N. I., PETERS, C., LI, M., WANG, X. & HUANG, J. 2011. Characterisation of the *Arabidopsis* glycerophosphodiester phosphodiesterase (GDPD) family reveals a role of the plastid-localized AtGDPD1 in maintaining cellular phosphate homeostasis under phosphate starvation. *The Plant Journal*, 66, 781-795.
- CHIDGEAVADZE, Z. G., BEABEALASHVILLI, R. S., ATRAZHEV, A. M., KUKHANOVA, M. K., AZHAYEV, A. V. & KRAYEVSKY, A. A. 1984. 2',3'-Dideoxy-3' aminonucleoside 5'-triphosphates are the terminators of DNA synthesis catalyzed by DNA polymerases. *Nucleic Acids Research*, 12, 1671-1686.
- CHILDERS, D. L., CORMAN, J., EDWARDS, M. & ELSER, J. J. 2011. Sustainability challenges of phosphorus and food: Solutions from closing the human phosphorus cycle. *BioScience*, 61, 117-124.
- CHIOU, T. J. & LIN, S. I. 2011. Signaling network in sensing phosphate availability in plants. *Annual Review of Plant Biology*, 62, 185-206.
- CHIOU, T. J., AUNG, K., LIN, S. I., WU, C. C., CHIANG, S. F. & SU, C. L. 2006. Regulation of phosphate homeostasis by microRNA in *Arabidopsis*. *Plant Cell*, 18, 412-421.
- COOPER, G. M. 2000. *The cell: A molecular approach*. 2nd edition ed. Sinaur Associates Inc., Sunderland (MA), U.S.A.
- COOPER, J., LOMBARDI, R., BOARDMAN, D. & CARLIELL-MARQUET, C. 2011. The future distribution and production of global phosphate rock reserves. *Resources, Conservation and Recycling*, 57, 78-86.
- CORDELL, D. 2010. *Sustainability implications of global phosphorus scarcity for food security*. Linköping University.
- CORDELL, D. & NESET, T. S. S. 2014. Phosphorus vulnerability: A qualitative framework for assessing the vulnerability of national and regional food systems to the multi-dimensional stressors of phosphorus scarcity. *Global Environmental Change*, 24, 108-122.
- CORDELL, D. & WHITE, S. 2011. Peak phosphorus: Clarifying the key issues of a vigorous debate about long-term phosphorus security. *Sustainability*, 3, 2027-2049.

- CORDELL, D. & WHITE, S. 2013. Sustainable phosphorus measures: Strategies and technologies for achieving phosphorus security. *Agronomy*, 3, 86-116.
- CORDELL, D., DRANGERT, J., O. & WHITE, S. 2009. The story of phosphorus: Global food security and food for thought. *Global Environmental Change*, 19, 292-305.
- CRUZ-RAMIREZ, A., OROPEZA-ABURTO, A., RAZO-HERNANDEZ, F., RAMIREZ-CHAVEZ, E., & HERERA-ESTRELLA, L. 2006. Phospholipase DZ2 plays an important role in extraplastidic galactolipid biosynthesis, and phosphate recycling in *Arabidopsis* roots. *Proceedings of the National Academy of Sciences USA*, 103, 6765-6770.
- DABER, R., SUKHADIA, S. & MORRISSETTE, J. J. D. 2013. Understanding the limitations of next generation sequencing informatics, an approach to clinical pipeline validation using artificial data sets. *Cancer Genetics*, 206, 441-448.
- DE ANGELI, A., ZHANG, J., MEYER, S. & MARTINOIA, E. 2013. AtALMT9 is a malate activated vacuolar chloride channel required for stomatal opening in *Arabidopsis*. *Nature Communications*, 4, 1804. Doi: 10.1038/ncomms2815.
- DELHAIZE, E. & RANDALL, P. J. 1995. Characterization of a phosphate-accumulator mutant of *Arabidopsis thaliana*. *Plant Physiology*, 107, 201-213.
- DELHAIZE, E., MA, J. F. & RYAN, P. R. 2012. Transcriptional regulation of aluminium tolerance genes. *Trends in Plant Science*, 17, 341-348.
- DELHAIZE, E., TAYLOR, P., HOCKING, P. J., SIMPSON, R. J., RYAN, P. R. & RICHARDSON, A. E. 2009. Transgenic barley (*Hordeum vulgare* L.) expressing the wheat aluminium resistance gene (*TaALMT1*) shows enhanced phosphorus nutrition and grain production when grown on an acid soil. *Plant Biotechnology Journal*, 7, 391-400.
- DEVAIAH, B. N., KARTHIKEYAN, A. S. & RAGHOTHAMA, K. G. 2007a. WRKY75 transcription factor is a modulator of phosphate acquisition and root development in *Arabidopsis*. *Plant Physiology*, 143, 1789-1801.
- DEVAIAH, B. N., NAGARAJAN, V. K. & RAGHOTHAMA, K. G. 2007b. Phosphate homeostasis and root development in *Arabidopsis* are synchronized by the zinc finger transcription factor ZAT6. *Plant Physiology*, 145, 147-159.
- DINH, P. T. Y., ROLDAN, M., LEUNG, S. & MCMANUS, M. T. 2012. Regulation of root growth by auxin and ethylene is influenced by phosphate supply in white clover (*Trifolium repens* L.). *Plant Growth Regulation*, 66, 179-190.
- DINKELAKER, B., HENGELER, C. & MARSCHNER, H. 1995. Distribution and function of proteoid roots and other root clusters. *Botanica Acta*, 108, 183-200.
- DÖRMANN, P., BALBO, I., & BENNING, C. 1999. *Arabidopsis* galactolipid biosynthesis and lipid trafficking mediated by *DGD1*. *Science*, 284, 2181-2184.

- DÖRMANN, P., HOFFMAN-BENNING, S., BALBO, I. & BENNING, C. 1995. Isolation and characterization of an *Arabidopsis* mutant deficient in the thylakoid lipid digalactosyl diacylglycerol. *Plant Cell*, 7, 1801-1810.
- DREW, M. C. & SAKER, L. R. 1984. Uptake and long-distance transport of phosphate, potassium and chloride in relation to internal ion concentrations in barley: Evidence of non-allosteric regulation. *Planta*, 160, 500-507.
- DU, H., YANG, C., DING, J., SHI, L. & XU, F. 2017. Genome wide identification and characterisation in SPX domain-containing member and their responses to Pi deficiency in *Brassica napus*. *Frontiers in Plant Science*, 8, 35. Doi: 10.3389/fpls.2017.00035.
- DUAN, K., YI, K., DANG, L., HUANG, H., WU, W. & WU, P. 2008. Characterization of a sub-family of *Arabidopsis* genes with the SPX domain reveals their diverse functions in plant tolerance to phosphorus starvation. *The Plant Journal*, 54, 965-975.
- DRUKA, A., POTOKINA, E., LUO, Z., JIANG, N., CHEN, X., KEARSEY, M. & WAUGH, R. 2010. Expression quantitative trait loci analysis in plants. *Plant Biotechnology Journal*, 8, 10-27.
- EID, J., FEHR, A., GRAY, J., LUONG, K., LYLE, J., OTTO, G., PELUSO, P., RANK, D., BAYBAYAN, P., BETTMAN, B., BIBILLO, A., BJORNSON, K., CHAUDHURI, B., CHRISTIANS, F., CICERO, R., CLARK, S., DALAL, R., DEWINTER, A., DIXON, J., FOQUET, M., GAERTNER, A., HARDENBOL, P., HEINER, C., HESTER, K., HOLDEN, D., KEARNS, G., KONG, X., KUSE, R., LACROIX, Y., LIN, S., LUNDQUIST, P., MA, C., MARKS, P., MAXHAM, M., MURPHY, D., PARK, I., PHAM, T., PHILLIPS, M., ROY, J., SEBRA, R., SHEN, G., SORENSON, J., TOMANEY, A., TRAVERS, K., TRULSON, M., VIECELI, J., WEGENER, J., WU, D., YANG, A., ZACCARIN, D., ZHAO, P., ZHONG, F., KORLACH, J. & TURNER, S. 2009. Real-time DNA sequencing from single polymerase molecules. *Science*, 323, 133-138.
- EMSLEY, J. 2000. *The 13th Element: The Sordid Tale of Murder, Fire, and Phosphorus*, John Wiley & Sons, New York.
- ESSIGMANN, B., GÜLER, S., NARANG, R. A., LINKE, D. & BENNING, C. 1998. Phosphate availability affects the thylakoid lipid composition and the expression of *SQD1*, a gene required for sulfolipid biosynthesis in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences USA*, 95, 1950-1955.
- FAHEY, J. W. 2003. Brassicas. In: CABALLERO, B. (ed) *Encyclopedia of Food Sciences and Nutrition (Second Edition)*. Oxford: Academic Press.
- FANG, L., CHENG, F., WU, J. & WANG, X. 2012. The impact of genome triplication on tandem gene evolution in *Brassica rapa*. *Frontiers in Plant Science*, 3. Doi: 10.3389/fpls.2012.00261.
- FINLAY, R. D. 2008. Ecological aspects of mycorrhizal symbiosis: With special emphasis on the functional diversity of interactions involving the extraradical mycelium. *Journal of Experimental Botany*, 59, 1115-1126.

- FRANCO-ZORRILLA, J. M., MARTIN, A. C., LEYVA, A. & PAZ-ARES, J. 2005. Interaction between phosphate-starvation, sugar, and cytokinin signaling in *Arabidopsis* and the roles of cytokinin receptors CRE1/AHK4 and AHK3. *Plant Physiology*, 138, 847-857.
- FROEHLICH, J. E., BENNING, C. & DÖRMANN, P. 2001. The digalactosyldiacylglycerol (DGDG) synthase DGD1 is inserted into the outer envelope membrane of chloroplasts in a manner independent of the general import pathway and does not depend on direct interaction with monogalactosyldiacylglycerol synthase for DGDG biosynthesis. *Journal of Biological Chemistry*, 276, 31806-31812.
- GAHOONIA, T. S. & NIELSEN, N. E. 2003. Phosphorus (P) uptake and growth of a root hairless barley mutant (*bald root barley*, *brb*) and wild type in low- and high-P soils. *Plant, Cell & Environment*, 26, 1759-1766.
- GAN, X., STEGLE, O., BEHR, J., STEFFEN, J. G., DREWE, P., HILDEBRAND, K. L., LYNGSOE, R., SCHULTHEISS, S. J., OSBORNE, E. J., SREEDHARAN, V. T., KAHLES, A., BOHNERT, R., JEAN, G., DERWENT, P., KERSEY, P., BELFIELD, E. J., HARBERD, N. P., KEMEN, E., TOOMAJIAN, C., KOVER, P. X., CLARK, R. M., RÁTSCH, G. & MOTT, R. 2011. Multiple reference genomes and transcriptomes for *Arabidopsis thaliana*. *Nature*, 477, 419-423.
- GANIE, A. H., AHMAD, A., PANDEY, R., AREF, I. M., YOUSUF, P. Y., AHMAD, S. & IQBA, M. 2015. Metabolite profiling of low-P tolerant and low-P sensitive maize genotypes under phosphorus starvation and restoration conditions. *PLoS ONE*, 10, 6. Doi: 10.1371/journal.pone.0129520.
- GARG, R., SAHOO, A., TYAGI, A. K. & JAIN, M. 2010. Validation of internal control genes for quantitative gene expression studies in chickpea (*Cicer arietinum* L.). *Biochemical and Biophysical Research Communications*, 396, 283-288.
- GAUDE, N., NAKAMURA, Y., SCHEIBLE, W. R., OHTA, H., & DORMANN, P. 2008. Phospholipase C5 (NPC5) is involved in galactolipid accumulation during phosphate limitation in leaves of *Arabidopsis*. *The Plant Journal*, 56, 28-39.
- GAUDE, N., TIPPMANN, H., FLEMETAKIS, E., KATINAKIS, P., UDVARDI, M. & DÖRMANN, P. 2004. The galactolipid digalactosyldiacylglycerol accumulates in the peribacteroid membrane of nitrogen-fixing nodules of soybean and lotus. *Journal of Biological Chemistry*, 279, 34624-34630.
- GONZALEZ-THUILLIER, I., SALT, L., CHOPE, C., PENSON, S., SKEGGS, P., TOSI, P., POWERS, S. J., WARD, J. L., WILDE, P., SHEWRY, P. R. & HASLAM, R. P. 2015. Distribution of lipids in the grain of wheat (cv. Hereward) determined by lipidomic analysis of milling and pearling fractions. *Journal of Agricultural and Food Chemistry*, 63, 10705-10716.
- GOTTWALD, J. R., KRYSAN, P. J., YOUNG, P. J., EVERT, R. F. & SUSSMAN, M. R. 2000. Genetic evidence for the in planta role of phloem-specific plasma membrane sucrose transporters. *Proceedings of the National Academy of Sciences USA*, 97, 13979-13984.

- GREENWOOD, D. J., STELLACCI, A. M., MEACHAM, M. C., BROADLEY, M. R. & WHITE, P. J. 2005. Components of P response of different *Brassica oleracea* genotypes are reproducible in different environments. *Crop Science*, 45, 1728-1735.
- GREENWOOD, D. J., STELLACCI, A. M., MEACHAM, M. C., BROADLEY, M. R. & WHITE, P. J. 2006. Relative values of physiological parameters of P response of different genotypes can be measured in experiments with only two P treatments. *Plant and Soil*, 281, 159-179.
- GU, C., CHEN, S., LIU, Z., SHAN, H., LUO, H., GUAN, Z. & CHEN, F. 2011. Reference gene selection for quantitative real-time PCR in chrysanthemum subjected to biotic and abiotic stress. *Molecular Biotechnology*, 49, 192-197.
- GU, M., CHEN, A., SUN, S. & XU, G. 2016. Complex regulation of plant phosphate transporters and the gap between molecular mechanisms and practical application: What is missing? *Molecular Plant*, 9, 396-416.
- HALING, R. E., BROWN, L. K., BENGOUGH, A. G., YOUNG, I. M., HALLETT, P. D., WHITE, P. J. & GEORGE, T. S. 2013. Root hairs improve root penetration, root-soil contact, and phosphorus acquisition in soils of different strength. *Journal of Experimental Botany*, 64, 3711-3721.
- HAMBURGER, D., REZZONICO, E., MACDONALD-COMBER, PETETOT, J., SOMERVILLE, C. & POIRIER, Y. 2002. Identification and characterization of the Arabidopsis *PHO1* gene involved in phosphate loading to the xylem. *Plant Cell*, 14, 889-902.
- HAMMOND, J. P. & WHITE, P. J. 2008. Sucrose transport in the phloem: Integrating root responses to phosphorus starvation. *Journal of Experimental Botany*, 59, 93-109.
- HAMMOND, J. P. & WHITE, P. J. 2011. Sugar signaling in root responses to low phosphorus availability. *Plant Physiology*, 156, 1033-1040.
- HAMMOND, J. P., BENNETT, M. J., BOWEN, H. C., BROADLEY, M. R., EASTWOOD, D. C., MAY, S. T., RAHN, C., SWARUP, R., WOOLAWAY, K. E. & WHITE, P. J. 2003. Changes in gene expression in Arabidopsis shoots during phosphate starvation and the potential for developing smart plants. *Plant Physiology*, 132, 578-596.
- HAMMOND, J. P., BROADLEY, M. R., CRAIGON, D. J., HIGGINS, J., EMMERSON, Z. F., TOWNSEND, H. J., WHITE, P. J. & MAY, S. T. 2005. Using genomic DNA-based probe-selection to improve the sensitivity of high-density oligonucleotide arrays when applied to heterologous species. *Plant Methods*, 1, 1-10.
- HAMMOND, J. P., BROADLEY, M. R., WHITE, P. J., KING, G. J., BOWEN, H. C., HAYDEN, R., MEACHAM, M. C., MEAD, A., OVERS, T., SPRACKLEN, W. P. & GREENWOOD, D. J. 2009. Shoot yield drives phosphorus use efficiency in *Brassica oleracea* and correlates with root architecture traits. *Journal of Experimental Botany*, 60, 1953-1968.

- HAMMOND, J. P., MAYES, S., BOWEN, H. C., GRAHAM, N. S., HAYDEN, R. M., LOVE, C. G., SPRACKLEN, W. P., WANG, J., WELHAM, S. J., WHITE, P. J., KING, G. J. & BROADLEY, M. R. 2011. Regulatory hotspots are associated with plant gene expression under varying soil phosphorus supply in *Brassica rapa*. *Plant Physiology*, 156, 1230-1241.
- HAN, X. & GROSS, R. W. 2003. Global analyses of cellular lipidomes directly from crude extracts of biological samples by ESI mass spectrometry: A bridge to lipidomics. *Journal of Lipid Research*, 44, 1071-1079.
- HARAN, S., LOGENDRA, S., SESKAR, M., BRATANOVA, M. & RASKIN, I. 2000. Characterization of Arabidopsis acid phosphatase promoter and regulation of acid phosphatase expression. *Plant Physiology*, 124, 615-626.
- HÄRTEL, H., DÖRMANN, P. & BENNING, C. 2000. DGD1-independent biosynthesis of extraplasmidic galactolipids after phosphate deprivation in *Arabidopsis*. *Proceedings of the National Academy of Sciences USA*, 19, 10649-10654.
- HE, L., ZHAO, M., WANG, Y., GAI, J. & HE, C. 2013. Phylogeny, structural evolution and functional diversification of the plant *PHOSPHATE1* gene family: A focus on *Glycine max*. *BMC Evolutionary Biology*, 13, 103. Doi: 10.1186/1471-2148-13-103.
- HEATHER, J. M. & CHAIN, B. 2016. The sequence of sequencers: The history of sequencing DNA. *Genomics*, 107, 1-8.
- HERNANDEZ, G., RAMÍREZ, M., VALDÉS-LÓPEZ, O., TEFAYE, M., GRAHAM, M. A., CZECHOWSKI, T., SCHLERETH, A., WANDREY, M., ERBAN, A., CHEUNG, F., WU, H. C., LARA, M., TOWN, C. D., KOPKA, J., UDVARDI, M. K. & VANCE, C. P. 2007. Phosphorus stress in common bean: Root transcript and metabolic responses. *Plant Physiology*, 144, 752-767.
- HEUER, S., GAXIOLA, R., SCHILLING, R., HERERA-ESTRELLA, L., LOPEZ-ARREDONDO, D., WISSUWA, M., DELHAIZE, E. & ROUACHED, H. 2016. Improving phosphorus use efficiency: A complex trait with emerging opportunities. *The Plant Journal*, 90, 868-885.
- HINSINGER, P. 2001. Bioavailability of soil inorganic P in the rhizosphere as affected by root-induced chemical changes: A review. *Plant and Soil*, 237, 173-195.
- HOLMES, M. R. J. 1980. *Nutrition of the oilseed rape crop*. Applied Science Publishers, London.
- HOLZL, G. & DÖRMANN, P. 2007. Structure and function of glycoacylglycerolipids in plants and bacteria. *Progress in Lipid Research*, 46, 225-243.
- HU, R., FAN, C., LI, H., ZHANG, Q. & FU, Y.-F. 2009. Evaluation of putative reference genes for gene expression normalization in soybean by quantitative real-time RT-PCR. *BMC Molecular Biology*, 10, 1-12.
- HUANG, C. Y., ROESSNER, U., EICKMEIER, I., GENC, Y., CALLAHAN, D. L., SHIRLEY, N., LANGRIDGE, P. & BACIC, A. 2008. Metabolite profiling reveals distinct

changes in carbon and nitrogen metabolism in phosphate-deficient barley plants (*Hordeum vulgare* L.). *Plant Cell Physiology*, 49, 691-703.

HUNTER, P. J., TEAKLE, G. R. & BENDING, G. D. 2014. Root traits and microbial community interactions in relation to phosphorus availability and acquisition, with particular reference to *Brassica*. *Frontiers in Plant Science*, 5, 27. Doi: 10.3389/fpls.2014.00027.

IMRAN, M., TANG, K. & LIU J. 2016. Comparative genome wide analysis of the malate dehydrogenase gene families in cotton. *PLoS ONE*, 11, 1-22. Doi: 10.1371/journal.pone.0166341.

INIGUEZ-LUI, F. L., LUKENS, L., FARNHAM, M. W., AMASINO, R. M., & OSBORN, T. C. 2009. Development of public immortal mapping populations, molecular markers, and linkage maps for rapid cycling *Brassica rapa* and *Brassica oleracea*. *Theoretical Applied Genetics*, 120, 31-43.

JAIN, M., OLSEN, H. E., PATEN, B. & AKESON, M. 2016. The Oxford Nanopore MinION: Delivery of nanopore sequencing to the genomics community. *Genome Biology*, 17, 239. Doi:10.1186/s13059-016-1103-0.

JAKOBY, M., WEISSHAAR, B., DROGE-LASER, W., VICENTE-CARBAIOSA, J., TIEDEMANN, J., KROJ, T. & PARCY, F. 2002. bZIP transcription factors in *Arabidopsis*. *Trends in Plant Science*, 7, 106-111.

JARVIS, P., DÖRMANN, P., PETO, C.A., LUTES, J., BENNING, C. & CHORY, J. 2000. Galactolipid deficiency and abnormal chloroplast development in the *Arabidopsis* *MGD synthase 1* mutant. *Proceedings of the National Academy of Sciences USA*, 97, 8175-8179.

JOUHET, J., MARECHAL, E. & BLOCK, M. A. 2007. Glycerolipid transfer for the building of membranes in plant cells. *Progress in Lipid Research*, 46, 37-55.

JOUHET, J., MARECHAL, E., BALDAN, B., BLIGNY, R., JOYARD, J. & BLOCK, M. A. 2004. Phosphate deprivation induces transfer of DGDG galactolipid from chloroplast to mitochondria. *The Journal of Cell Biology*, 167, 863-874.

JUNG, J. K. H. & MCCOUCH, S. 2013. Getting to the roots of it: Genetic and hormonal control of root architecture. *Frontiers in Plant Science*, 4, 186. Doi:10.3389/fpls.2013.00186.

KAIDA, R., SAGE-ONO, K., KAMADA, H., OKUYAMA, H., SYONO, K. & KANEKO, T. S. 2003. Isolation and characterization of four cell wall purple acid phosphatase genes from tobacco cells. *Biochimica et Biophysica Acta*, 1625, 134-140.

KANG, X. & NI, M. 2006. *Arabidopsis* SHORT HYPOCOTYL UNDER BLUE1 contains SPX and EXS domains and acts in cryptochrome signaling. *Plant Cell*, 18, 921-934.

KANT, S., PENG, M. & ROTHSTEIN, S. J. 2011. Genetic regulation by NLA and microRNA827 for maintaining nitrate-dependent phosphate homeostasis in *Arabidopsis*. *PLoS Genetics*, 7. Doi: 10.1371/journal.pgen.1002021.

- KARTHIKEYAN, A. S., VARADARAJA, D. K., MUKATIRA, U. T., D'URZO, M. P., DAMSZ, B & RAGOTHAMA, K. G. 2002. Regulated expression of Arabidopsis phosphate transporters. *Plant Physiology*, 130, 221-233.
- KELLIS, M., BIRREN, B. W. & LANDER, E. S. 2004. Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature*, 428, 617-624.
- KELLY, A. A. & DÖRMANN, P. 2002. *DGD2*, an *Arabidopsis* gene encoding a UDP-galactose-dependent digalactosyldiacylglycerol synthase is expressed during growth under phosphate-limiting conditions. *Journal of Biological Chemistry*, 277, 1166-1173.
- KELLY, A. A., FROEHLICH, J. E. & DÖRMANN, P. 2003. Disruption of the two digalactosyldiacylglycerol synthase genes *DGD1* and *DGD2* in *Arabidopsis* reveals the existence of an additional enzyme of galactolipid synthesis. *Plant Cell*, 15, 2694-2706.
- KHAN, G. A., BOURAINE, S., WEGE, S., LI, Y., DE CARBONNEL, M., BERTHOMIEU, P., POIRIER, Y. & ROUACHED, H. 2014. Coordination between zinc and phosphate homeostasis involves the transcription factor PHR1, the phosphate exporter *PHO1*, and its homologue *PHO1;H3* in *Arabidopsis*. *Journal of Experimental Botany*, 65, 871-884.
- KISKO, M., SHUKLA, V., KAUR, M., BOUAIN, N., CHAIWONG, N., LACOMBE, B., PANDEY, A. K. & ROUACHED, H. 2018. Phosphorus transport in *Arabidopsis* and wheat: Emerging strategies to improve P pool in seeds. *Agriculture*, 8, 27.
- KLIEBENSTEIN, D. 2009. Quantitative genomics: Analyzing intraspecific variation using global gene expression polymorphisms or eQTLs. *Annual Review of Plant Biology*, 60, 93-114.
- KOBAYASHI, K., AWAI, K., NAKAMURA, M., NAGATANI, A., MASUDA, T. & OHTA, H. 2009. Type-B monogalactosyldiacylglycerol synthases are involved in phosphate starvation-induced lipid remodeling, and are crucial for low-phosphate adaptation. *The Plant Journal*, 57, 322-331.
- KOBAYASHI, K., AWAI, K., TAKAMIYA, K. & OHTA, H. 2004. *Arabidopsis* type B monogalactosyldiacylglycerol synthase genes are expressed during pollen tube growth and induced by phosphate starvation. *Plant Physiology*, 134, 640-648.
- KOBAYASHI, K., KONDO, M., FUKUDA, H., NISHIMURA, M. & OHTA, H. 2007. Galactolipid synthesis in chloroplast inner envelope is essential for proper thylakoid biogenesis, photosynthesis, and embryogenesis. *Proceedings of the National Academy of Sciences USA*, 104, 17216-17221.
- KOBAYASHI, K., MASUDA, T., TAKAMIYA, K. & OHTA, H. 2006. Membrane lipid alteration during phosphate starvation is regulated by phosphate signaling and auxin/cytokinin cross-talk. *The Plant Journal*, 47, 238-248.

- KOBAYASHI, K., NARISE, T., SONOIKE, K., HASHIMOTO, H., SATO, N., KONDO, M., NISHIMURA, M., SATO, M., TOYOOKA, K., SUGIMOTO, K., WADA, H., MASUDA, T. & OHTA, H. 2013. Role of galactolipid biosynthesis in coordinated development of photosynthetic complexes and thylakoid membranes during chloroplast biogenesis in *Arabidopsis*. *The Plant Journal*, 73, 250-261.
- KOBAYASHI, Y., KOBAYASHI, Y., SUGIMOTO, M., LAKSHMANAN, V., LUCHI, S., KOBAYASHI, M., BAIS, H. P. & KOYAMA, H. 2013. Characterization of the complex regulation of *AtALMT1* expression in response to phytohormones and other inducers. *Plant Physiology*, 162, 732-740.
- KOPPELAAR, R. H. E. M. & WEIKARD, H. P. 2013. Assessing phosphate rock depletion and phosphorus recycling options. *Global Environmental Change*, 23, 1454-1466.
- KOREN, S., WALENZ, B. P., BERLIN, K., MILLER, J. R., BERGMAN, N. H. & PHILLIPPY, A. M. 2017. Canu: Scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Research*, 27, 722-736.
- KULSKI, J. K. 2016. Next-Generation Sequencing - An overview of the history, tools, and "Omic" Applications. In: KULSKI, J. K. (ed.) *Next Generation Sequencing - Advances, Applications and Challenges*, Rijeka: InTech.
- KUMAR, A., SALISBURY, P. A., GURUNG, A. M. & BARBETTI, M. 2015. *Brassica oilseeds: Breeding and Management*. CAB International.
- KUPPUSAMY, T., GIAVALISCO, P., ARVIDSSON, S., SULPICE, R., STITT, M., FINNEGAN, P. M., SCHEIBLE, W.-R., LAMBERS, H. & JOST, R. 2014. Lipid biosynthesis and protein concentration respond uniquely to phosphate supply during leaf development in highly phosphorus-efficient *Hakea prostrata*. *Plant Physiology*, 166, 1891-1911.
- LAMBERS, H., CAWTHRAY, G. R., GIAVALISCO, P., KUO, J., LALIBERTE, E., PEARSE, S. J., SCHEIBLE, W. R., STITT, M., TESTE, F. & TURNER, B. L. 2012. Proteaceae from severely phosphorus-impooverished soils extensively replace phospholipids with galactolipids and sulfolipids during leaf development to achieve a high photosynthetic phosphorus-use-efficiency. *New Phytologist*, 196, 1098-1108.
- LAMBERS, H., CLEMENTS, J. C. & NELSON, M. N. 2013. How a phosphorus-acquisition strategy based on carboxylate exudation powers the success and agronomic potential of lupines (*Lupinus*, Fabaceae). *American Journal of Botany*, 100, 263-288.
- LAMBERS, H., CRAMER, M. D., SHANE, M.W., WOUTERLOOD, M., POOT, P. & VENEKLAAS, E. J. 2003. Structure and functioning of cluster roots and plant responses to phosphate deficiency. *Plant and Soil*, 248, ix-xix.
- LAMBERS, H., RAVEN, J. A., SHAVER, G. R. & SMITH, S. E. 2008. Plant nutrient-acquisition strategies change with soil age. *Trends in Ecology and Evolution*, 23, 95-103.

- LAN, P., LI, W., & SCHMIDT, W. 2012. Complementary proteome and transcriptome profiling in phosphate-deficient *Arabidopsis* roots reveals multiple levels of gene regulation. *Molecular & Cellular Proteomics*, 11, 1156-1166.
- LEGGETT, R. M. & CLARK, M. D. 2017. A world of opportunities with nanopore sequencing. *Journal of Experimental Botany*, 69, 5419-5429.
- LEI, M., LIU, Y., ZHANG, B., ZHAO, Y., WANG, X., ZHOU, Y., RAGHOTHAMA, G. & LIU, D. 2011. Genetic and genomic evidence that sucrose is a global regulator of plant responses to phosphate starvation in *Arabidopsis*. *Plant Physiology*, 156, 1116-1130.
- LEI, M., ZHU, C., LIU, Y., KARTHIKEYAN, A. S., BRESSAN, R. A., RAGHOTHAMA, K. G. & LIU, D. 2011. Ethylene signalling is involved in regulation of phosphate starvation-induced gene expression and production of acid phosphatases and anthocyanin in *Arabidopsis*. *New Phytologist*, 189, 1084-1095.
- LEO, V. C., MORGAN, N. V., BEM, D., JONES, M. L., LOWE, G. C., LORDKIPANIDZÉ, M., DRAKE, S., SIMPSON, M. A., GISSEN, P., MUMFORD, A., WATSON, S. P., DALY, M. E. & THE, U. K. G. S. G. 2015. Use of next-generation sequencing and candidate gene analysis to identify underlying defects in patients with inherited platelet function disorders. *Journal of Thrombosis and Haemostasis*, 13, 643-650.
- LI, L., HE, Z., PANDEY, G. K., TSUCHIYA, T. & LUAN, S. 2002. Functional cloning and characterization of a plant efflux carrier for multidrug and heavy metal detoxification. *Journal of Biological Chemistry*, 277, 5360-5368.
- LI, M., QON, C., WELTI, R. & WANG, X. 2006a. Double knockouts of phospholipase D ζ 1 and D ζ 2 in *Arabidopsis* affect root elongation during phosphate-limited growth but do not affect root hair patterning. *Plant Physiology*, 140, 761-770.
- LI, M., WELTI, R. & WANG, X. 2006b. Quantitative profiling of *Arabidopsis* polar glycerolipids in response to phosphorus starvation. Roles of phospholipases D ζ 1 and D ζ 2 in phosphatidylcholine hydrolysis and digalactosyldiacylglycerol accumulation in phosphorus-starved plants. *Plant Physiology*, 142, 750-761.
- LIGABA, A., KATSUHARA, M., RYAN, P. R., SHIBASAKA, M. & MATSUMOTO, H. 2006. The *BnALMT1* and *BnALMT2* genes from rape encode Aluminum-Activated Malate Transporters that enhance the aluminum resistance of plant cells. *Plant Physiology*, 142, 1294-1303.
- LIN, W. Y., HUANG, T. K. & CHIOU, T. J. 2013. NITROGEN LIMITATION ADAPTATION, a target of MicroRNA827, mediates degradation of plasma membrane-localized phosphate transporters to maintain phosphate homeostasis in *Arabidopsis*. *Plant Cell*, 10, 4061-4074.
- LIN, W. Y., HUANG, T. K., LEONG, S. J. & CHIOU, T. J. 2014. Long-distance call from phosphate: Systemic regulation of phosphate starvation responses. *Journal of Experimental Botany*, 65, 1817-1827.
- LIU, J., YANG, L., LUAN, M., WANG, Y., ZHANG, C., ZHANG, B., SHI, J., ZHAO, F. G., LAN, W. & LUAN, S. 2015. A vacuolar phosphate transporter essential for

- phosphate homeostasis in *Arabidopsis*. *Proceedings of the National Academy of Sciences USA*, 112, 6571-6578.
- LIU, N., SHANG, W., LI, C., JIA, L., WANG, X., XING, G. & ZHENG, W. 2018. Evolution of the *SPX* gene family in plants and its role in the response mechanism to phosphorus stress. *Open Biology*, 8, 1, 170231. Doi: 10.1098/rsob.170231.
- LIU, T. Y., LIN, W. Y., HUANG, T. K. & CHIOU, T. J. 2014. MicroRNA-mediated surveillance of phosphate transporters on the move. *Trends in Plant Science*, 19, 647-655.
- LIU, T. Y., HUANG, T. K., TSENG, C. Y., LAI, Y. S., LIN, S. I., LIN, W. Y., CHEN, J. W. & CHIOU, T. J. 2012. *PHO2*-Dependent degradation of *PHO1* modulates phosphate homeostasis in *Arabidopsis*. *The Plant Cell*, 24, 2168-2183.
- LLOYD, J. C. & ZAKHLENIUK, O. V. 2004. Responses of primary and secondary metabolism to sugar accumulation revealed by microarray expression analysis of the *Arabidopsis* mutant, *pho3*. *Journal of Experimental Botany*, 55, 1221-1230.
- LOPEZ-ARREDONDO, D. L., LEYVA-GONZALEZ, M. A., GONZALEZ-MORALES, S. I., LOPEZ-BUCIO, J. & HERRERA-ESTRELLA, L. 2014. Phosphate nutrition: Improving low-phosphate tolerance in crops. *Annual Review of Plant Biology*, 65, 95-123.
- LV, Q., ZHONG, Y., WANG, Y., WANG, Z., ZHANG, L., SHI, J., WU, Z., LIU, Y., MAO, C., YI, K. & WU, P. 2014. *SPX4* negatively regulates phosphate signaling and homeostasis through its interaction with *PHR2* in rice. *Plant Cell*, 26, 1586-1597.
- LYNCH, J. P. & BROWN, K. M. 2001. Topsoil foraging – an architectural adaptation of plants to low phosphorus availability. *Plant and Soil*, 237, 225-237.
- LYSAK, M. A., KOCH, M. A., BEAULIEU, J. M., MEISTER, A. & LEITCH, I. J. 2009. The dynamic ups and downs of genome size evolution in Brassicaceae. *Molecular Biology and Evolution*, 26, 85-98.
- MA, S., NIU, H., LIU, C., ZHANG, J., HOU, C. & WANG, D. 2013. Expression stabilities of candidate reference genes for RT-qPCR under different stress conditions in soybean. *PLoS ONE*, 8, 1-7. Doi: 10.1371/journal.pone.0075271.
- MAGI, A., SEMERARO, R., MINGRINO, A., GIUSTI, B. & D'AURIZIO, R. 2017. Nanopore sequencing data analysis: State of the art, applications and challenges. *Briefings in Bioinformatics*, 1-17.
- MALBOOBI, M. A., ZAMANI, K., LOHRASEBI, T., SARIKHANI, M. R., SAMAIAN, A. & SABET, M. S. 2014. Phosphate: The silent challenge. *Progress in Biological Sciences*, 4, 1-32.
- MANSKE, G. G. B., ORTIZ-MONASTERIO, J. I., VAN GINKEL, M., GONZÁLEZ, R. M., RAJARAM, S., MOLINA, E. & VLEK, P. L. G. 2000. Traits associated with improved P-uptake efficiency in CIMMYT's semi-dwarf spring bread wheat grown on an acid Andisol in Mexico. *Plant and Soil*, 221, 189-204.

- MARDIS, E. R. 2008. The impact of next-generation sequencing technology on genetics. *Trends in Genetics*, 24, 133-141.
- MARGULIES, M., EGHOLM, M., ALTMAN, W. E., ATTIYA, S., BADER, J. S., BEMBEN, L. A., BERKA, J., BRAVERMAN, M. S., CHEN, Y. J., CHEN, Z., DEWELL, S. B., DU, L., FIERRO, J. M., GOMES, X. V., GODWIN, B. C., HE, W., HELGESEN, S., HO, C. H., IRZYK, G. P., JANDO, S. C., ALENQUER, M. L., JARVIE, T. P., JIRAGE, K. B., KIM, J. B., KNIGHT, J. R., LANZA, J. R., LEAMON, J. H., LEFKOWITZ, S. M., LEI, M., LI, J., LOHMAN, K. L., LU, H., MAKHIJANI, V. B., MCDADE, K. E., MCKENNA, M. P., MYERS, E. W., NICKERSON, E., NOBILE, J. R., PLANT, R., PUC, B. P., RONAN, M. T., ROTH, G. T., SARKIS, G. J., SIMONS, J. F., SIMPSON, J. W., SRINIVASAN, M., TARTARO, K. R., TOMASZ, A., VOGT, K. A., VOLKMER, G. A., WANG, S. H., WANG, Y., WEINER, M. P., YU, P., BEGLEY, R. F. & ROTHBERG, J. M. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, 437, 376-380.
- MASONI, A., ERCOLI, L., MARIOTTI, M. & ARDUINI, I. 2007. Post-anthesis accumulation and remobilization of dry matter, nitrogen and phosphorus in durum wheat as affected by soil type. *European Journal of Agronomy*, 26, 179-186.
- MCGINNIS, J., LAPLANTE, J., SHUDT, M. & GEORGE, K. S. 2016. Next generation sequencing for whole genome analysis and surveillance of influenza A viruses. *Journal of Clinical Virology*, 79, 44-50.
- MEHRA, P. & GIRI, J. 2016. Rice and chickpea GDPDs are preferentially influenced by low phosphate and *CaGDPD1* encodes an active glycerophosphodiester phosphodiesterase enzyme. *Plant Cell Reports*, 35, 1699-1717.
- MEYER, S., DE ANGELI, A., FERNIE, A. R. & MARTINOIA, E. 2010. Intra-and extra-cellular excretion of carboxylates. *Trends in Plant Science*, 15, 40-47.
- MEYER, S., MUMM, P., IMES, D., ENDLER, A., WEDER, B., AL-RASHEID, K. A. S., GEIGER, D., MARTEN, I., MARTINOIA, E., & HEDRICH, R. 2010. *AtALMT12* represents an R-type anion channel required for stomatal movement in *Arabidopsis* guard cells. *The Plant Journal*, 63, 1054-1062.
- MISSON, J., RAGHOTHAMA, K. G., JAIN, A., JOUHET, J., BLOCK, M. A., BLIGNY, R., ORTET, P., CREFF, A., SOMERVILLE, S., ROLLAND, N., DOUMAS, P., NACRY, P., HERRERRA-ESTRELLA, L., NUSSAUME, L. & THIBAUD, M.-C. 2005. A genome-wide transcriptional analysis using *Arabidopsis thaliana* Affymetrix gene chips determined plant responses to phosphate deprivation. *Proceedings of the National Academy of Sciences USA*, 102, 11934-11939.
- MIURA, K., RUS, A., SHARKHUU, A., YOKOI, S., KARTHIKEYAN, A. S., RAGHOTHAMA, K. G., BAEK, D., KOO, Y. D., JIN, J. B., BRESSAN, R. A., YUN, D.-J. & HASEGAWA, P. M. 2005. The *Arabidopsis* SUMO E3 ligase SIZ1 controls phosphate deficiency responses. *Proceedings of the National Academy of Sciences USA*, 102, 7760-7765.
- MIZUSAWA, N. & WADA, H. 2012. The role of lipids in photosystem II. *Biochimica et Biophysica Acta*, 1817, 194-208.

- MOELLERING, E. R. & BENNING, C. 2011. Galactoglycerolipid metabolism under stress: A time for remodeling. *Trends in Plant Science*, 16, 98-107.
- MOMOH, E. J. J., ZHOU, W. J. & KRISTIANSSON, B. 2002. Variation in the development of secondary dormancy in oilseed rape genotypes under conditions of stress. *Weed Research*, 42, 446-455.
- MUCHHAL, U. S., PARDO, J. M. & RAGHOTHAMA, K. G. 1996. Phosphate transporters from the higher plant *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences USA*, 93, 10519-10523.
- MUDGE, S. R., RAE, A. L., DIATLOFF, E. & SMITH, F. W. 2002. Expression analysis suggests novel roles for members of the PHT1 family of phosphate transporters in *Arabidopsis*. *The Plant Journal*, 31, 341-353.
- MULLER, R., MORANT, M., JARMER, H., NILSSON, L. & NIELSEN, T. H. 2007. Genome-wide analysis of the *Arabidopsis* leaf transcriptome reveals interaction of phosphate and sugar metabolism. *Plant Physiology*, 143, 156-171.
- MURAKAWA, M., SMIMOJIMA, M., SHIMOMURA, Y., KOBAYASHI, K., AWAI K., & OHTA, H. 2014. Monogalactosyldiacylglycerol synthesis in the outer envelope membrane of chloroplasts is required for enhanced growth under sucrose supplementation. *Frontiers in Plant Science*, 5. Doi: 10.3389/fpls.2014.00280.
- NAGAHARU, U. 1935. Genome analysis in Brassica with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Journal of Japanese Botany*, 7, 389-452.
- NAKAMURA, Y. 2013. Phosphate starvation and membrane lipid remodeling in seed plants. *Progress in Lipid Research*, 52, 43-50.
- NAKAMURA, Y. 2017. Plant phospholipid diversity: Emerging functions in metabolism and protein–lipid interactions. *Trends in Plant Science*, 22, 1027-1040.
- NAKAMURA, Y., AWAI, K., MASUDA, K., YOSHIOKA, Y., TAKAMIYA, K., & OHTA, H. 2005. A novel phosphatidylcholine-hydrolyzing phospholipase C induced by phosphate starvation in *Arabidopsis*. *Journal of Biological Chemistry*, 280, 7469-7476.
- NAKAMURA, Y., KOIZUMI R., SHUI, G., SHIMOJIMA, M., WENK, M. R. & ITO, T. 2009. *Arabidopsis* lipins mediate eukaryotic pathway of lipid metabolism and cope critically with phosphate starvation. *Proceedings of the National Academy of Sciences USA*, 106, 20978-20983.
- NANDI, A., KROTHAPALLI, K., BUSEMAN, C. M., LI, M., WELTI, R., ENYEDI, A. & SHAH, J. 2003. *Arabidopsis* *sfd* mutants affect plastidic lipid composition and suppress dwarfing, cell death, and the enhanced disease resistance phenotypes resulting from the deficiency of a fatty acid desaturase. *Plant Cell*, 15, 2383-2398.

- NARISE, T., KOBAYASHI, K., BABA, S., SHIMOJIMA, M., MASUDA, S., FUKAKI, H. OHTA, H. 2010. Involvement of auxin signaling mediated by IAA14 and ARF7/19 in membrane lipid remodeling during phosphate starvation. *Plant Molecular Biology*, 72, 533-544.
- NER-GAON, H., HALACHMI, R., SAVALDI-GOLDSTEIN, S., RUBIN, E., OPHIR, R. & FLUHR, R. 2004. Intron retention is a major phenomenon in alternative splicing in *Arabidopsis*. *The Plant Journal*, 39, 877-885.
- NICOT, N., HAUSMAN, J.-F., HOFFMANN, L. & EVERS, D. 2005. Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. *Journal of Experimental Botany*, 56, 2907-2914.
- NIU, Y. F., CHAI, R. S., JIN, G. L., WANG, H., TANG, C. X. & ZHANG, Y. S. 2013. Responses of root architecture development to low phosphorus availability: A review. *Annals of Botany*, 112, 391-408.
- NUSSAUME, L., KANNO, S., JAVOT, H., MARIN, E., POCHON, N., AYADI, A., NAKANISHI, T. M. & THIBAUD, M. C. 2011. Phosphate import in plants: Focus on the PHT1 transporters. *Frontiers in Plant Science*, 2, 83. Doi: 10.3389/fpls.2011.00083.
- OELKERS, E. H. & VALSAMI-JONES, E. 2008. Phosphate mineral reactivity and global sustainability. *Elements*, 4, 83-87.
- OKAZAKI, Y., OTSUKI, H., NARISAWA, T., KOBAYASHI, M., SAWAI, S., KAMIDE, Y., KUSANO, M., AOKI, T., HIRAI, M. Y. & SAITO, K. 2013. A new class of plant lipid is essential for protection against phosphorus depletion. *Nature Communications*, 4, 1510. Doi: 10.1038/ncomms2512.
- OKAZAKI, Y., SHIMOJIMA, M., SAWADA, Y., TOYOOKA, K., NARISAWA, T., MOCHIDA, K., TANAKA, H., MATSUDA, F., HIRAI, A., HIRAI, M. Y., OHTA, H. & SAITO, K. 2009. A chloroplastic UDP-glucose pyrophosphorylase from *Arabidopsis* is the committed enzyme for the first step of sulfolipid biosynthesis. *Plant Cell*, 21, 892-909.
- OKAZAKI, Y., TAKANO, K. & SAITO, K. 2017. Lipidomic analysis of soybean leaves revealed tissue-dependent difference in lipid remodeling under phosphorus-limited growth conditions. *Plant Biotechnology*, 34, 57-63. Doi: 10.5511/plantbiotechnology.17.0113a.
- O'ROURKE, J. A., YANG, S. S., MILLER, S. S., BUCCIARELLI, B., LIU, J., RYDEEN, A., BOZSOKI, Z., UHDE-STONE, C., TU, Z. J., ALLAN, D., GRONWALD, J. W. & VANCE, C. P. 2013. An RNA-Seq transcriptome analysis of orthophosphate-deficient white lupin reveals novel insights into phosphorus acclimation in plants. *Plant Physiology*, 161, 705-724.
- PALMER, A., J., BAKER, A., & MUENCH, S. P. 2016. The varied functions of aluminium-activated malate transporters-much more than aluminium resistance. *Biochemical Society Transactions*, 44, 856-862.

- PANT, B. D., BURGOS, A., PANT, P., CUADROS-INOSTROZAL, A., WILLMITZER, L., & SCHEIBLEL, W. 2015. The transcription factor PHR1 regulates lipid remodelling and triacylglycerol accumulation in *Arabidopsis thaliana* during phosphorus starvation. *Journal of Experimental Botany*, 66, 1907- 1918.
- PARIASCA-TANAKA, J., KANJI, S., TERRY, R., RAMIL, M. & MATTHIAS, W. 2009. Stress response versus stress tolerance: A transcriptome analysis of two rice lines contrasting in tolerance to phosphorus deficiency. *Rice*, 2, 167-185. Doi: 10.1007/S12284-009-9032-0.
- PARKIN, I. A., GULDEN, S. M., SHARPE, A. G., LUKENS, L., TRICK, M., OSBORN, T. C. & LYDIATE, D. J. 2005. Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. *Genetics*, 171, 765-781.
- PATTERSON, H. D. & THOMPSON, R. 1971. Recovery of inter-block information when block sizes are unequal. *Biometrika*, 58, 545-554.
- PELIZZOLA, M. & ECKER, J. R. 2011. The DNA methylome. *FEBS letters*, 585, 1994-2000.
- PENG, M., HANNAM, C., GU, H., BI, Y. M. & ROTHSTEIN, S. J. 2007. A mutation in *NLA*, which encodes a RING-type ubiquitin ligase, disrupts the adaptability of *Arabidopsis* to nitrogen limitation. *The Plant Journal*, 50, 320-337.
- PERET, B., CLÉMENT, M., NUSSAUME, L. & DESNOS, T. 2011. Root developmental adaptation to phosphate starvation: Better safe than sorry. *Trends in Plant Science*, 16, 442-450.
- PERTEA, M., KIM, D., PERTEA, G.M., LEEK, J. T. & SALZBERG, S. L. 2016. Transcript-level expression analysis of RNA-seq experiments with HISAT, Stringtie and Ballgown. *Nature Protocols*, 11, 1650-1667.
- PLAXTON, W. C. & TRAN, H. T. 2011. Metabolic adaptations of phosphate-starved plants. *Plant Physiology*, 156, 1006-1015.
- POIRIER, Y. & BUCHER, M. 2002. Phosphate transport and homeostasis in *Arabidopsis*. *The American Society of Plant Biologists*, 1-35.
- POIRIER, Y., THOMA, S., SOMERVILLE, C. & SCHIEFELBEIN, J. 1991. A mutant of *Arabidopsis* deficient in xylem loading of phosphate. *Plant Physiology*, 97, 1087-1093.
- PUGA, M. I., MATEOS, I., CHARUKESI, R., WANG, Z., FRANCO-ZORRILLA, J. M., DE LORENZO, L., IRIGOYEN, M. L., MASIERO, S., BUSTOS, R., RODRÍGUEZ, J., LEYVA, A., RUBIO, V., SOMMER, H. & PAZ-ARES, J. 2014. SPX1 is a phosphate-dependent inhibitor of PHOSPHATE STARVATION RESPONSE 1 in *Arabidopsis*. *Proceedings of the National Academy of Sciences USA*, 111, 14947-14952.
- PUGA, M. I., ROJAS-TRIANA, M., DE LORENZO, L., LEYVA, A., RUBIO, V. & PAZ-ARES, J. 2017. Novel signals in the regulation of Pi starvation responses in plants: Facts and promises. *Current Opinion in Plant Biology*, 39, 40-49.

- QIN, C. & WANG, X. 2002. The Arabidopsis phospholipase D family. Characterization of a calcium-independent and phosphatidylcholine-selective *PLD ζ 1* with distinct regulatory domains. *Plant Physiology*, 128, 1057-1068.
- RABOY, V. 2009. Approaches and challenges to engineering seed phytate and total phosphorus. *Plant Science*, 177, 281-296.
- RAGHOTHAMA, K. G. 1999. Phosphate acquisition. *Annual Review of Plant Physiology and Plant Molecular Biology*, 50, 665-693.
- RAMAIAH, M., JAIN, A., BALDWIN, J. C., KARTHIKEYAN, A. S. & RAGHOTHAMA, K. G. 2011. Characterization of the phosphate starvation-induced *Glycerol-3-phosphate permease* gene family in Arabidopsis. *Plant Physiology*, 157, 279-291.
- RAYMER, P. L. 2002. Canola: An emerging oilseed crop. *Trends in New Crops and New Uses*, 122-126.
- REN, F., GUO, Q. Q., CHANG, L. L., CHEN, L., ZHAO, C. Z., ZHONG, H. & LI, X. B. 2012. *Brassica napus* *PHR1* gene encoding a MYB-Like protein functions in response to phosphate starvation. *PLoS One*, 7. Doi: 10.1371/journal.pone.0044005.
- REN, F., ZHAO, C. Z., LIU, C. S., HUANG, K. L., GUO, Q. Q., CHANG, L. L., XIONG, H. & LI, X. B. 2014. A *Brassica napus* PHT1 phosphate transporter, BnPht1;4, promotes phosphate uptake and affects roots architecture of transgenic Arabidopsis. *Plant Molecular Biology*, 86, 595-607.
- REYMOND, M., SVISTOONOFF, S., LOUDET, O., NUSSAUME, L. & DESNOS, T. 2006. Identification of QTL controlling root growth response to phosphate starvation in *Arabidopsis thaliana*. *Plant Cell and Environment*, 29, 115-125.
- RIETZ, S., HOLK, A. & SCHERER, G. F. 2004. Expression of the patatin-related phospholipase A gene *AtPLA IIA* in *Arabidopsis thaliana* is up-regulated by salicylic acid, wounding, ethylene, and iron and phosphate deficiency. *Planta*, 219, 743-753.
- ROGERS, E. E., WU, X., STACEY, G. & NGUYEN, H. T. 2009. Two MATE proteins play a role in iron efficiency in soybean. *Journal of Plant Physiology*, 166, 1453-1459.
- ROGERS, K. & CHEN, X. 2013. Biogenesis, turnover, and mode of action of plant MicroRNAs. *The Plant Cell*, 25, 2383-2399.
- ROSE, T. J., RENGEL, Z., MA, Q. & BOWDEN, J. W. 2007. Differential accumulation patterns of phosphorus and potassium by canola cultivars compared to wheat. *Journal of Plant Nutrition and Soil Science*, 170, 404-411.
- ROUACHED, H. 2011. Multilevel coordination of phosphate and sulfate homeostasis in plants. *Plant Signaling & Behavior*, 6, 952-955.
- RUBIO, V., LINHARES, F., SOLANO, R., MARTÍN, A. C., IGLESIAS, J., LEYVA, A. & PAZ-ARES, J. 2001. A conserved MYB transcription factor involved in phosphate

- starvation signaling both in vascular plants and in unicellular algae. *Genes & Development*, 15, 2122-2133.
- RUSHOLME, R. L., HIGGINS, E. E., WALSH, J. A. & LYDIATE, D. J. 2007. Genetic control of broad-spectrum resistance to turnip mosaic virus in *Brassica rapa* (Chinese cabbage). *Journal of General Virology*, 88, 3177-3186.
- RUSSO, M. A., QUARTACCI, M. F., IZZO, R., BELLIGNO, A. & NAVARI-IZZO, F. 2007. Long and short-term phosphate deprivation in bean roots: Plasma membrane lipid alterations and transient stimulation of phospholipases. *Phytochemistry*, 68, 1564-1571.
- RYAN, P. R., DELHAIZE, E. & JONES, D. L. 2001. Function and mechanism of organic anion exudation from plant roots. *Annual Review of Plant Physiology and Plant Molecular Biology*, 52, 527-560.
- SAMARAKOON, T., SHIVA, S., LOWE, K., TAMURA, P., ROTH, M. R. & WELTI, R. 2012. *Arabidopsis thaliana* membrane lipid molecular species and their mass spectral analysis. In: NORMANLY, J. (ed.) *High-Throughput Phenotyping in Plants: Methods and Protocols*. Totowa, NJ: Humana Press.
- SÁNCHEZ-CALDERÓN, L., LÓPEZ-BUCIO, J., CHACÓN-LÓPEZ, A., GUTIÉRREZ-ORTEGA, A., HERNÁNDEZ-ABREU, E. & HERRERA-ESTRELLA, L. 2006. Characterization of *low phosphorus insensitive* mutants reveals a crosstalk between low phosphorus-induced determinate root development and the activation of genes involved in the adaptation of *Arabidopsis* to phosphorus deficiency. *Plant Physiology*, 140, 879-889.
- SANDA, S., LEUSTEK, T., THEISEN, M. J., GARAVITO, R. M. & BENNING, C. 2001. Recombinant *Arabidopsis* SQD1 converts UD-glucose and sulfite to the sulfolipid head group precursor UDP-sulfoquinovose *in vitro*. *Journal of Biological Chemistry*, 276, 3941-3946.
- SANGER, F., NICKLEN, S. & COULSON, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences USA*, 74, 5463-5467.
- SAS, L., RENGEL, Z. & TANG, Z. 2001. Excess cation uptake, and extrusion of protons and organic acid anions by *Lupinus albus* under phosphorus deficiency. *Plant Science*, 160, 1191-1198.
- SASAKI, T., YAMAMOTO, Y., EZAKI, B., KATSUHARA, M., JU AHN, S., RYAN, P. R., DELHAIZE, E. & MATSUMOTO, H. 2004. A wheat gene encoding an aluminium-activated malate transporter. *The Plant Journal*, 37, 645-653.
- SAVAGE, N., YANG, T. J., CHEN, C. Y., LIN, K. L., MONK, N. A. & SCHMIDT, W. 2013. Positional signaling and expression of *ENHANCER OF TRY AND CPC1* are tuned to increase root hair density in response to phosphate deficiency in *Arabidopsis thaliana*. *PLoS One*, 8. Doi: 10.1371/journal.pone.0075452.
- SCHACHTMAN, D. P., REID, R. J. & AYLING, S. M. 1998. Phosphorus uptake by plants: From soil to cell. *Plant Physiology*, 116, 447-453.

- SCHINDLER, D. W., CARPENTER, S. R., CHAPRA, S. C., HECKY, R. E. & ORIHEL, D. M. 2016. Reducing phosphorus to curb lake eutrophication is a success. *Environmental Science & Technology*, 50, 8923-8929.
- SCHMIDTMANN, E., KONIG, A., ORWAT, A., LEISTER, D., HARTL, M. & FINKEMEIER, I. 2014. Redox regulation of Arabidopsis mitochondrial citrate synthase. *Molecular Plant*, 7, 156-169.
- SCHWERTNER, H. A. & BIALE, J. B. 1973. Lipid composition of plant mitochondria and of chloroplasts. *Journal of Lipid Research*, 14, 235-242.
- SECCO, D., WANG, C., ARPAT, B. A., WANG, Z., POIRIER, Y., TYERMAN, S. D., WU, P., SHOU, H. & WHELAN, J. 2012. The emerging importance of the SPX domain-containing proteins in phosphate homeostasis. *New Phytologist*, 193, 842-851.
- SHANG, X., CAO, Y. & MA, L. 2017. Alternative splicing in plant genes: A means of regulating the environmental fitness of plants. *International Journal of Molecular Sciences*, 18, 432-449.
- SHANE, M. W., CRAMER, M. D., FUNAYAMA-NOGUCHI, S., CAWTHRAY, G. R., MILLAR, H., DAY, D.A. & LAMBERS, H. 2004. Developmental physiology of cluster-root carboxylate synthesis and exudation in harsh Hakea. expression of phosphoenolpyruvate carboxylase and the alternative oxidase. *Plant Physiology*, 135, 549-560.
- SHEN, J., YUAN, L., ZHANG, J., LI, H., BAI, Z., CHEN, X., ZHANG, W. & ZHANG, F. 2011. Phosphorus dynamics: From soil to plant. *Plant Physiology*, 156, 997-1005.
- SHENOY, V. V. & KALAGUDI, G. M. 2005. Enhancing plant phosphorus use efficiency for sustainable cropping. *Biotechnology Advances*, 23, 501-513.
- SHIMOJIMA, M. & OHTA, H. 2011. Critical regulation of galactolipid synthesis controls membrane differentiation and remodeling in distinct plant organs and following environmental changes. *Progress in Lipid Research*. 50, 258-266. Doi: 10.1016/j.plipres.2011.03.001.
- SHIMOJIMA, M. 2011. Biosynthesis and functions of the plant sulfolipid. *Progress in Lipid Research*, 50, 234-239.
- SHIMOJIMA, M., MADOKA, Y., FUJIWARA, R., MURAKAWA, M., YOSHITAKE, Y., IKEDA, K., KOIZUMI, R., ENDO, K., OZAKI, K. & OHTA, H. 2015. An engineered lipid remodeling system using a galactolipid synthase promoter during phosphate starvation enhances oil accumulation in plants. *Frontiers in Plant Science*, 6, 664. Doi: 10.3389/fpls.2015.00664.
- SHIMOJIMA, M., WATANABE, T., MADOKA, Y., KOIZUMI, R., YAMAMOTO, M. P., MASUDA, K., YAMADA, K., MASUDA, S. & OHTA, H. 2013. Differential regulation of two types of monogalactosyldiacylglycerol synthase in membrane lipid remodeling under phosphate-limited conditions in sesame plants. *Frontiers in Plant Science*, 4, 469. Doi: 10.3389/fpls.2013.00469.

- SHUKLA, T., KUMAR, S., KHARE, R., TRIPATHI, R. D., & TRIVEDI, P. K. 2015. Natural variations in expression of regulatory and detoxification related genes under limiting phosphate and arsenate stress in *Arabidopsis thaliana*. *Frontiers in Plant Science*, 6, 898. Doi: 10.3389/fpls.2015.00898.
- SHULAEV, V & CHAPMAN, K. D. 2017. Plant lipidomics at the crossroads: From technology to biology driven science. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, 1862, 786-791.
- SIMON, M., LOUDET, O., DURAND, S., BÉRARD, A., BRUNEL, D., SENNESAL, F.-X., DURAND-TARDIF, M., PELLETIER, G. & CAMILLERI, C. 2008. Quantitative trait loci mapping in five new large recombinant inbred line populations of *Arabidopsis thaliana* genotyped with consensus single-nucleotide polymorphism markers. *Genetics*, 178, 2253-2264.
- SKENE, K. R. 1998. Cluster roots: Some ecological considerations. *Journal of Ecology*, 86, 1060-1064.
- SMILL, V. 2000a. *Feeding the World: A Challenge for the Twenty-First Century*, Cambridge, The MIT Press.
- SMILL, V. 2000b. Phosphorus in the environment: Natural flows and human interferences. *Annual Review of Energy and Environment*, 25, 53-88.
- SMITH, F. W., MUDGE, S. R., RAE, A. L. & GGLASSOP, D. 2003. Phosphate transport in plants. *Plant and Soil*, 248, 71-83.
- SMITH, S. E. & READ, D. J. 2008. *Mycorrhizal Symbiosis*, Academic Press.
- SNOWDON, R., LÜHS, W. & FRIEDT, W. 2007. Oilseed Rape. In: KOLE, C. (ed.) *Oilseeds*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- STEEN, I. 1998. Phosphorus availability in the 21st Century: Management of a non-renewable resource. *Phosphorus and Potassium*, 217, 25-31.
- STEFANOVIC, A., RIBOT, C., ROUACHED, H., WANG, Y., CHONG, J., BELBAHRI, L., DELESSERT, S. & POIRIER, Y. 2007. Members of the *PHO1* gene family show limited functional redundancy in phosphate transfer to the shoot, and are regulated by phosphate deficiency via distinct pathways. *The Plant Journal*, 50, 982-994.
- STUTTER, M. I., SHAND, C. A., GEORGE, T. S., BLACKWELL, M. S. A., DIXON, L., BOL, R., MACKAY, R. L., RICHARDSON, A. E., CONDRON, L. M. & HAYGARTH, P. M. 2015. Land use and soil factors affecting accumulation of phosphorus species in temperate soils. *Geoderma*, 257, 29-39.
- SULTAN, M., SCHULZ, M. H., RICHARD, H., MAGEN, A., KLINGENHOFF, A., SCHERF, M., SEIFERT, M., BORODINA, T., SOLDATOV, A., PARKHOMCHUK, D., SCHMIDT, D., O'KEEFFE, S., HAAS, S., VINGRON, M., LEHRACH, H. & YASPO, M. L. 2008. A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science*, 321, 356-360.

- SURIYAGODA, D. B., RYAN, H., RENTON, M. & LAMBERS, H. 2014. Chapter four- plant responses to limited moisture and phosphorus availability: A meta-analysis. *Advances in Agronomy*, 124,143-200.
- SVISTOONOFF, S., CREFF, A., REYMOND, M., SIGOILLOT-CLAUDE, C., RICAUD, L., BLANCHET, A., NUSSAUME, L. & DESNOS, T. 2007. Root tip contact with low-phosphate media reprograms plant root architecture. *Nature Genetics*, 39, 792-796.
- TAN, H., YANG, X., ZHANG, F., ZHENG, X., QU, C., MU, J., FU, F., LI, J., GUAN, R., ZHANG, H., WANG, G. & ZUO, J. 2011. Enhanced seed oil production in canola by conditional expression of *Brassica napus* LEAFY COTYLEDON1 and LEC1-LIKE in developing seeds. *Plant Physiology*, 156, 1577-1588.
- TANG, H. & LYONS, E. 2012. Unleashing the Genome of *Brassica rapa*. *Frontiers in Plant Science*, 3. Doi: 10.3389/fpls.2012.00172.
- TANG, J., LEUNG, C. & LIM, B. M. 2006. Hydrolysis of precipitated phytate by three distinct families of phytases. *Soil Biology and Biochemistry*. 38, 1316-1324.
- TENENBOIM, H., BURGOS, A., WILMITZER, L. & BROTMAN, Y. 2016. Using lipidomics for expanding the knowledge on lipid metabolism in plants. *Biochimie*, 130, 91-96.
- THIBAUD, M. C., ARRIGHI, J. F., BAYLE, V., CHIARENZA, S., CREFF, A., BUSTOS, R., PAZ-ARES, J., POIRIER, Y. & NUSSAUME, L. 2010. Dissection of local and systemic transcriptional responses to phosphate starvation in *Arabidopsis*. *The Plant Journal*, 64, 775-789.
- THOMPSON, J. F. & MILOS, P. M. 2011. The properties and applications of single-molecule DNA sequencing. *Genome Biology*, 12, 217. Doi: 10.1186/gb-2011-12-2-217.
- TICCONI, C. A., DELATORRE, C. A., LAHNER, B., SALT, D. E. & ABEL, S. 2004. *Arabidopsis pdr2* reveals a phosphate-sensitive checkpoint in root development. *The Plant Journal*, 37, 801-814.
- TJELLSTRÖM, H., ANDERSSON, M. X., LARSSON, K. E. & SANDELIUS, A. S. 2008. Membrane phospholipids as a phosphate reserve: The dynamic nature of phospholipid-to-digalactocylglycerol exchange in higher plants. *Plant, Cell & Environment*, 31, 1388-1398.
- TOMAZ, T., BAGARD, M., PRACHAROENWATTANA, I., LINDEN, P., LEE, C. P., CARROLL, A. J., STRÖHER, E., SMITH, S. M., GARDESTRÖM, P. & MILLAR, A. H. 2010. Mitochondrial malate dehydrogenase lowers leaf respiration and alters photorespiration and plant growth in *Arabidopsis*. *Plant Physiology*, 154, 1143-1157.
- TRAN, H. T., QIAN, W., HURLEY, B. A., SHE, Y. M., WANG, D. & PLAXTON, W. C. 2010. Biochemical and molecular characterization of AtPAP12 and AtPAP26: The predominant purple acid phosphatase isozymes secreted by phosphate-starved *Arabidopsis thaliana*. *Plant, Cell & Environment*, 33, 1789-1803.

- TURNBULL, C. G. & LOPEZ-COBOLLO, R. M. 2013. Heavy traffic in the fast lane: Long-distance signalling by macromolecules. *New Phytologist*, 198, 33-51.
- TURNER, B. L., A.E., R. & MULLANEY, E. J. 2007. *Inositol Phosphates: Linking Agriculture and the Environment*, Wallingford, UK, CAB International.
- UHDE-STONE, C., GILBERT, G., JOHNSON, J. M.-F., LITJENS, R., ZINN, K. E., TEMPLE, S. J., VANCE, C. P. & ALLAN, D. L. 2003. Acclimation of white lupin to phosphorus deficiency involves enhanced expression of genes related to organic acid metabolism. *Plant and Soil*, 248, 99-116.
- USDA (UNITED STATES DEPARTMENT OF AGRICULTURE). 2015. Oilseeds: World markets and trade. Foreign agricultural service office of global analysis.
- VALDES-LOPEZ, O., ARENAS-HUERTERO, C., RAMIREZ, M., GIRARD, L., SANCHEZ, F., VANCE, C. P., LUIS REYES, J. & HERNANDEZ, G. 2008. Essential role of MYB transcription factor: PvPHR1 and microRNA: PvmiR399 in phosphorus-deficiency signalling in common bean roots. *Plant, Cell & Environment*, 31, 1834-1843.
- VANCE, C. P., UHDE-STONE, C. & ALLAN, D. L. 2003. Phosphorus acquisition and use: Critical adaptations by plants for securing a nonrenewable resource. *New Phytologist*, 157, 423-447.
- VOSKOBOYNIK, A., NEFF, N. F., SAHOO, D., NEWMAN, A. M., PUSHKAREV, D., KOH, W., PASSARELLI, B., FAN, H. C., MANTALAS, G. L., PALMERI, K. J., ISHIZUKA, K. J., GISSI, C., GRIGGIO, F., BEN-SHLOMO, R., COREY, D. M., PENLAND, L., WHITE, R. A., WEISSMAN, I. L. & QUAKE, S. R. 2013. The genome sequence of the colonial chordate, *Botryllus schlosseri*. *eLife*, 2. Doi:10.7551/eLife.00569.
- WALAN, P., DAVIDSSON, S., JOHANSSON, S. & HÖÖK, M. 2014. Phosphate rock production and depletion: Regional disaggregated modeling and global implications. *Resources, Conservation and Recycling*, 93, 178-187.
- WANG, H., XU, Q., KONG, Y. H., CHEN, Y., DUAN, J. Y., WU, W. H. & CHEN, Y. F. 2014a. Arabidopsis WRKY45 transcription factor activates *PHOSPHATE TRANSPORTER1;1* expression in response to phosphate starvation. *Plant Physiology*, 164, 2020-2029.
- WANG, J., SUN, J., MIAO, J., GUO, J., SHI, Z., HE, M., CHEN, Y., ZHAO, X., LI, B., HAN, F., TONG, Y. & LI, Z. 2013. A phosphate starvation response regulator *Ta-PHR1* is involved in phosphate signalling and increases grain yield in wheat. *Annals of Botany*, 111, 1139-1153.
- WANG, L. & LIU, D. 2012. Arabidopsis purple acid phosphatase 10 is a component of plant adaptive mechanism to phosphate limitation. *Plant Signaling & Behavior*, 7, 306-310.
- WANG, L., LI, Z., QIAN, W., GUO, W., GAO, X., HUANG, L., WANG, H., ZHU, H., WU, J.-W., WANG, D. & LIU, D. 2011a. The Arabidopsis purple acid phosphatase

- AtPAP10 is predominantly associated with the root surface and plays an important role in plant tolerance to phosphate limitation. *Plant Physiology*, 157, 1283-1299.
- WANG, X., SHEN, J., & LIAO, H. 2010. Acquisition or utilization, which is more critical for enhancing phosphorus efficiency in modern crops? *Plant Science*, 179, 302-306.
- WANG, X., WANG, H., WANG, J., SUN, R., WU, J., LIU, S., BAI, Y., MUN, J.-H., BANCROFT, I., CHENG, F., HUANG, S., LI, X., HUA, W., WANG, J., WANG, X., FREELING, M., PIRES, J. C., PATERSON, A. H., CHALHOUB, B., WANG, B., HAYWARD, A., SHARPE, A. G., PARK, B.-S., WEISSHAAR, B., LIU, B., LI, B., LIU, B., TONG, C., SONG, C., DURAN, C., PENG, C., GENG, C., KOH, C., LIN, C., EDWARDS, D., MU, D., SHEN, D., SOUMPOUROU, E., LI, F., FRASER, F., CONANT, G., LASSALLE, G., KING, G. J., BONNEMA, G., TANG, H., WANG, H., BELCRAM, H., ZHOU, H., HIRAKAWA, H., ABE, H., GUO, H., WANG, H., JIN, H., PARKIN, I. A. P., BATLEY, J., KIM, J.-S., JUST, J., LI, J., XU, J., DENG, J., KIM, J. A., LI, J., YU, J., MENG, J., WANG, J., MIN, J., POULAIN, J., WANG, J., HATAKEYAMA, K., WU, K., WANG, L., FANG, L., TRICK, M., LINKS, M. G., ZHAO, M., JIN, M., RAMCHIARY, N., DROU, N., BERKMAN, P. J., CAI, Q., HUANG, Q., LI, R., TABATA, S., CHENG, S., ZHANG, S., ZHANG, S., HUANG, S., SATO, S., SUN, S., KWON, S.-J., CHOI, S.-R., LEE, T.-H., FAN, W., ZHAO, X., TAN, X., XU, X., WANG, Y., QIU, Y., YIN, Y., LI, Y., et al. 2011b. The genome of the mesopolyploid crop species *Brassica rapa*. *Nature Genetics*, 43, 1035-1039.
- WANG, Y., RIBOT, C., REZZONICO, E. & POIRIER, Y. 2004. Structure and expression profile of the Arabidopsis *PHO1* gene family indicates a broad role in inorganic phosphate homeostasis. *Plant Physiology*, 135, 400-411.
- WANG, Z., GERSTEIN, M. & SNYDER, M. 2009. RNA-Seq: A revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 10, 57-63.
- WANG, Z., LI, Q., GE, X., YANG, C., LUO, X., ZHANG, A., XIAO, J., TIAN, Y., XIA, G., CHEN, X., LI, F. & WU, J. 2015. The mitochondrial malate dehydrogenase 1 gene *GhmMDH1* is involved in plant root growth under phosphorus deficiency conditions in cotton. *Scientific Reports*, 5, 1-14.
- WANG, Z., RUAN, W., SHI, J., ZHANG, L., XIANG, D., YANG, C., LI, C., WU, Z., LIU, Y., YU, Y., SHOU, H., MO, X., MAO, C. & WU, P. 2014b. Rice SPX1 and SPX2 inhibit phosphate starvation responses through interacting with PHR2 in a phosphate-dependent manner. *Proceedings of the National Academy of Sciences USA*, 111, 14953-14958.
- WEGE, S., KHAN, G. A., JUNG, J.-Y., VOGIATZAKI, E., PRADERVAND, S., ALLER, I., MEYER, A. J. & POIRIER, Y. 2016. The EXS Domain of PHO1 participates in the response of shoots to phosphate deficiency via a root-to-shoot signal. *Plant Physiology*, 170, 385-400.
- WELHAM, S. J. & THOMPSON, R. 1997. Likelihood ratio tests for fixed model terms using residual maximum likelihood. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 59, 701-714.

- WELTI, R. & WANG, X. 2004. Lipid species profiling: A high throughput approach to identify lipid compositional changes and determine the function of genes involved in lipid metabolism and signaling. *Current Opinion in Biology*, 7, 337-344.
- WELTI, R., WEIQI, LI., LI, M., SANG, Y., BISIADA, H., ZHOU, H., RAJASHEKAR, C. B., WILLIAMS, T. D. & WANG, X. 2002. Profiling membrane lipids in plant stress responses: Role of phospholipase D α in freezing-induced lipid changes in *Arabidopsis*. *The American Society for Biochemistry and Molecular Biology*, 277, 31994-32002.
- WHITE, P. J. & HAMMOND, J. P. 2009. The sources of phosphorus in the waters of Great Britain. *Journal of Environmental Quality*, 38, 13-26.
- WILD, R., GERASIMAITE, R., JUNG, J. Y., TRUFFAULT, V., PAVLOVIC, I., SCHMIDT, A., SAIARDI, A., JESSEN, H. J., POIRIER, Y., HOTHORN, M. & MAYER, A. 2016. Control of eukaryotic phosphate homeostasis by inositol polyphosphate sensor domains. *Science*, 352, 986-990.
- WINTZ, H. & VULPE, C. 2002. Plant copper chaperones. *Biochemical society transactions*, 30, 732-735.
- WOO, J., MACPHERSON, C. R., LIU, J., WANG, H., KIBA, T., HANNAH, M. A., WANG, X. J., BAJIC, V. B. & CHUA, N. H. 2012. The response and recovery of the *Arabidopsis thaliana* transcriptome to phosphate starvation. *BMC Plant Biology*, 12, 62. Doi: 10.1186/1471-2229-12-62.
- WU, P., MA, L., HOU, X., WANG, M., WU, Y., LIU, F. & DENG, W. 2003. Phosphate starvation triggers distinct alterations of genome expression in *Arabidopsis* roots and leaves. *Plant Physiology*, 132, 1260-1271.
- WU, P., SHOU, H., XU, G. & LIAN, X. 2013. Improvement of phosphorus efficiency in rice on the basis of understanding phosphate signaling and homeostasis. *Current Opinion in Plant Biology*, 16, 205-212.
- XU, G., GUO, C., SHAN, H. & KONG, H. 2012. Divergence of duplicate genes in exon-intron structure. *Proceedings of the National Academy of Sciences USA*, 109, 1187-1192.
- YAMAGUCHI, M., SASAKI, T., SIVAGURU, M., YAMAMOTO, Y., OSAWA, H., AHN, S. J., & MATSUMOTO, H. 2005. Evidence for the plasma membrane localization of Al-activated malate transporter (ALMT1). *Plant Cell Physiology*, 46, 812-816.
- YAMARYO, Y., DUBOTS, E., ALBRIEUX, C., BALDAN, B., & BLOCK M. A. 2008. Phosphate availability affects the tonoplasts localization of PLD ζ 2, an *Arabidopsis thaliana* phospholipase D. *FEBS Letters*, 582, 685-690.
- YAN, F., ZHU, Y. Y., MULLER, C., ZRB, C. & SCHUBERT, S. 2002. Adaptation of H⁺-pumping and plasma membrane H⁺ ATPase activity in proteoid roots of white lupin under phosphate deficiency. *Plant Physiology*, 129, 50-63.

- YAN, X., LIAO, H., TRULL, M. C., BEEBE, S. E. & LYNCH, J. P. 2001. Induction of a major leaf acid phosphatase does not confer adaptation to low phosphorus availability in Common Bean. *Plant Physiology*, 125, 1901-1911.
- YANG, W. Y., ZHENG, Y., BAHN, S. C., PAN, X. Q., LI, M. Y., ROTH, M. R., SCHEU, B., WELTI, R., HONG, Y. Y. & WANG, X. M. 2012. The patatin-containing phospholipase A pPLA1 α modulates oxylipin formation and water loss in *Arabidopsis thaliana*. *Molecular Plant*, 5, 452-460.
- YI, K., WU, Z., ZHOU, J., DU, L., GUO, L., WU, Y. & WU, P. 2005. OsPTF1, a novel transcription factor involved in tolerance to phosphate starvation in rice. *Plant Physiology*, 138, 2087-2096.
- YOKOSHO, K., YAMAJI, N., FUJII-KASHINO, M. & MA, J. F. 2016. Functional analysis of MATE gene *OsFRDL2* revealed its involvement in Al-induced secretion of citrate, but a lower contribution to Al tolerance in rice. *Plant & Cell Physiology*, 57, 976-985.
- YU, B., XU, C., & BENNING, C. 2002. Arabidopsis disrupted in *SQD2* encoding sulfolipid synthase is impaired in phosphate-limited growth. *Proceedings of the National Academy of Sciences USA*, 99, 5732-5737.
- ZHANG, G., GUO, G., HU, X., ZHANG, Y., LI, Q., LI, R., ZHUANG, R., LU, Z., HE, Z., FANG, X., CHEN, L., TIAN, W., TAO, Y., KRISTIANSEN, K., ZHANG, X., LI, S., YANG, H., WANG, J. & WANG, J. 2010. Deep RNA sequencing at single base-pair resolution reveals high complexity of the rice transcriptome. *Genome Research*, 20, 646-654.
- ZHANG, Z., LIAO, H. & LUCAS, W. J. 2014. Molecular mechanisms underlying phosphate sensing, signaling, and adaptation in plants. *Journal of Integrative Plant Biology*, 56, 192-220.
- ZHAO, Z., SHI, H.-J., WANG, M.-L., CUI, L., ZHAO, H. & ZHAO, Y. 2015. Effect of nitrogen and phosphorus deficiency on transcriptional regulation of genes encoding key enzymes of starch metabolism in duckweed (*Landoltia punctata*). *Plant Physiology and Biochemistry*, 86, 72-81.
- ZHOU, K., YAMAGISHI, M., OSAKI, M. & MASUDA, K. 2008. Sugar signalling mediates cluster root formation and phosphorus starvation-induced gene expression in white lupin. *Journal of Experimental Botany*, 59, 2749-2756.
- ZHOU, Y. & NI, M. 2009. *SHB1* plays dual roles in photoperiodic and autonomous flowering. *Developmental Biology*, 331, 50-57.
- ZHOU, Y., XU, D., JIA, L., HUANG, X., MA, G., WANG, S., ZHU, M., ZHANG, A., GUAN, M., LU, K., XU, X., WANG, R., LI, J. & QU, C. 2017. Genome-wide identification and structural analysis of *bZIP* transcription factor genes in *Brassica napus*. *Genes*, 8, 288. Doi: 10.3390/8100288.
- ZHOU, Y., ZHANG, X., KANG, X., ZHAO, X., ZHANG, X. & NI, M. 2009. SHORT HYPOCOTYL UNDER BLUE1 associates with *MINISEED3* and *HAIKU2*

promoters in vivo to regulate *Arabidopsis* seed development. *The Plant Cell*, 21, 106-117.

ZIOLKOWSKI, P. A., KACZMAREK, M., BABULA, D. & SADOWSKI, J. 2006. Genome evolution in *Arabidopsis/Brassica*: conservation and divergence of ancient rearranged segments and their breakpoints. *The Plant Journal*, 47, 63-74.

APPENDICES

APPENDIX 1. Transcript of seven plasmid sequenced; Five splice variants of PHO1_A (Plasmid 2, 7, 23, 26, 22), PHO1_B (Plasmid 12), PHO1_D (Plasmid 24).

>PHO1A_Splice_variant_1_(plasmid_2)

ATGACGTTTCGGGAAGGAGTTTTTCGTCTCAGATGGTACCGGAGTGGCAACAAGCTTAC
ATGGACTACGATTATCTCAAACCCTTCTCAAAGAAATCATCCGCTTCAAACCTCAGA
ACCAACAATGCACCTACTCGTGGTGGCGCCAAGAATCATCAAGGCGGAGGATTA
CCGGAAGATGACTCTTTACCGAGCGTTTAGCGGTTTAGTCTCAACTCCGGGAAGACA
TAGACGCGGTAACCCTCACGACGTAGAGGAAGGGATACAGCTGACAGGGACGACGA
CGACGTCAGGGCCTATTCTTGTTAAACAACACCGCGGACCGCGGCTACGAGACCACGT
TCCTTATGGCGGCGGAGGAAGGAGGAGAGTACGAGCTGGTGTTTTTCCGGAGACTAG
ACGACGAGTTCAATAAAGTAAATAAGTTTTACAAAGAGAAAGTGGACGAAGTGTTG
AAAGAAGCTGTGGTGCTTAACAACAGATGGACGCTTTGATCGCGTTTTCGTGTTAAA
GTTGAGAATCCAGCAGGGTGGGGATGGGATGAACGAGCGGTGGAGATCACTCGCTT
GGCTCCGACATCGCTACTTCCGCGGCGGCTATCTCCGCTTCTACTCCCGCCGGAGCT
AAATCCATGAAAGTTCGGAGTCAAGCTCACATGGAGGCAATACAGGAAGGAGGGTC
GAGCAAAGCTGGGCAATTAGAAGATGATGAAGAGGAAGAAGCGCAAGCAGAAATT
GTAGCTTCCGTGTCTACCGGAGCTAGTGACGTGAGCACGACCAGGATGAGAGCAGTG
AGACCAGCTCCGTTGGATATTCTTGATCGAGTGACGATCAATAACACCAAAGAGACG
CCTCGTTCCACCATCAAAGAGTTCTACAGGTATCCAAGAACACTGATTTAAAGTTC
AGCAGAGAAAATCTGATGAAGGTCGAGGAGAACTCAGGCACGCTTTCATCGTGTTT
TATCAGAAGCTTAGGCTTCTCAAGAGCTACAGCTTTTTGAATGTACTGGCGTTCTCTA
AGCTATTGAAGAAGTATGACAAGATTACTTCGAGGGATGCAACCAAGCCTTACATGA
AAGTGGTTGATAGTTCATACCTCGGAAGCTCTGATGAAGTTGTGCGACTCATGGAGC
GTGTTGAAGCTACGTTCATAAAGCATTTTGCAAATGCTAACCGAACCAAAGGAATGA
ACATTTTACGGCCTCAGGCAAACGAGAGACATAGACTTACTTTCTCCACAGGTT
TCACGGCTGGATGCGTTTTCTCTTATAGTGGCTCTTGCCGCTATCATCCGCACGCG
AAATCTCTTGACGAGGAAGGCCAGAAGCAATACATGAATACTATGTTTCCTCTTTA
TAGCTTCTTCGGTTTTATCGTGCTGCACATAAATGTATGCTGCTAATATATACTAC
TGGAGGCGATACAAAGTAAACTATTCCTTCATATTTGGGTTCAAGCAAGGAAGTAA
CTTGTTATAGACAAGTCTGCTTGTGGGTTTCAGCATTGGAGTCTTTGCGTTGCTTT
GTGTTCTTGCTAATCTTGACATGGAGGCAAATCCCAAACCAAAGACTATAAAACAT
TCACCGAACTTCTTCCTCTTTTCTACTTCTTGCCATGTTTGTGGTTCTAATCTTGCCA
TTCAACTTCTTCTACCGGTCGAGTCGCTTTTTCTTCTCACTTGTCTGTTTCATTGCCTC
GCAGTCTCTTTACAAGGTAACACTGCCTGATTTCTTCTTGGGAGATCAATTAATA
GCCAGGTTCAAGCTCTTCGAAGCATCGAGTCTACATCTGTTACTATGGTTGGGGAG
ACTTCAGACACAGACAGAACACTTGTAACAAATCTAGTGTGTACAACACTTTCTTATT
CATTGTTGCTGTCATCCCTTACGCCTCTCGCCTCCTTCAGTGCCTGAGACGTTTGTTCG
AAGAGAAAAACCGGAGCAAGGGTGAATGGACTCAAGTACTTCTTACTATAGTG
GCGGTTTGCTTGAGAACTGCTTATAGTATCCAGAAAGGTCAAATCGCTTGGAGGGTT
TTAGCAGCTATCTCCTCAGCTGCAGCTGCAGTATTCAGTACTTACTGGGACTTTATCC
ATGATTGGGGTCTTCTAAACCGGACATCAAAAAACCGCTGGCTTCGTGATAAACTCC
TCATTCCCCAAAGAAAGTATACTTCATCGCCATGATTTTGAATGTCTGCTGAGATT
TGCAT

>PHO1A_Splice_variant_2_(plasmid_7)

ATGACGTTTCGGGAAGGAGTTTTTCGTCTCAGATGGTACCGGAGTGGCAACAAGCTTAC
ATGGACTACGATTATCTCAAAACCCTTCTCAAAGAAATCATCCGCTTCAAACCTCAGA
ACCAACAATGCACCTACTCGTGGTGGCGCCAAGAATCATCAAGGCGGAGGATTA
CCGGAAGATGACTCTTTACCGAGCGTTTAGCGGTTTAGTCTCAACTCCGGGAAGACA
TAGACGCGGTAACCCTCACGACGTAGAGGAAGGGATACAGCTGACAGGGACGACGA
CGACGTCAGGGCCTATTCTTGTTAACAACACCGCGGACCGCGGCTACGAGACCACGT
TCCTTATGGCGGCGGAGGAAGGAGGAGAGTACGAGCTGGTGTTTTTCCGGAGACTAG
ACGACGAGTTCAATAAAGTAAATAAGTTTTACAAAGAGAAAGTGGACGAAGTGTTG
AAAGAAGCTGTGGTGCTTAACAACAGATGGACGCTTTGATCGCGTTTTCTGTGTTAAA
GTTGAGAATCCAGCAGGGTGGGGATGGGATGAACGAGCGGTGGAGATCACTCGCTT
GGCTTCCGACATCGCTACTTCCGCGGCGGCTATCTCCGCTTCTACTCCCGCCGGAGCT
AAATCCATGAAAGTTCGGAGTCAAGCTCACATGGAGGCAATACAGGAAGGAGGGTC
GAGCAAAGCTGGGCAATTAGAAGATGATGAAGAGGAAGAAGCGCAAGCAGAAAATT
GTAGCTTCCGTGTCTACCGGAGCTAGTGACGTGAGCACGACCAGGATGAGAGCAGTG
AGACCAGCTCCGTTGGATATTCTTGATCGAGTGACGATCAATAACACCAAAGAGACG
CCTCGTTCCACCATCAAAGAGTTCTACAGGTATCCAAGAACACTGATTTAAAGTTC
AGCAGAGAAAATCTGATGAAGGTCGAGGAGAACTCAGGCACGCTTTCATCGTGTTT
TATCAGAAGCTTAGGCTTCTCAAGAGCTACAGCTTTTTGAATGTACTGGCGTTCTCTA
AGCTATTGAAGAAGTATGACAAGATTACTTCGAGGGATGCAACCAAGCCTTACATGA
AAGTGGTTGATAGTTCATACCTCGGAAGCTCTGATGAAGTTGTGCGACTCATGGAGC
GTGTTGAAGCTACGTTCATAAAGCATTTTGCAAATGCTAACCGAACCAAGGAATGA
ACATTTTACGGCCTCAGGCAAACGAGAGACATAGACTTACTTTCTCCACAGGTT
TCACGGCTGGATGCGTTTTCTCTTATAGTGGCTCTTGCCGCTATCATCCGCACGCG
AAATCTCTTGCAGGAGGAAGGCCAGAAGCAATACATGAATACTATGTTTCCTCTTTA
TAGCTTCTTCGGTTTTATCGTGCTGCACATAATAATGTATGCTGCTAATATATACTAC
TGGAGGCGATACAAAGTAAACTATTCTTCATATTTGGGTTCAAGCAAGGAAGTAA
CTTGTTATAGACAAGTCTGCTTGTGGGTTTCAGCATTGGAGTCTTTCGCTTGCTTT
GTGTTCTTGCTAATCTTGACATGGAGGCAAATCCCAAACCAAAGACTATAAAACAT
TCACCGAACTTCTTCTCTTTTCTACTTCTTGCCATCTCCTCTTTACAAGGTAACACT
GCCTGATTTCTTCTTGGGAGATCAATTAAGTCCAGGTTCAAGCTCTTTCGAAGCATC
GAGTTCTACATCTGTTACTATGGTTGGGGAGACTTCAGACACAGACAGAACACTTGT
AACAAATCTAGTGTGTACAACACTTTCTTATTCATTGTTGCTGTATCCCTTACGCCTC
TCGCCTCCTTCAGTGCCTGAGACGTTTGTTCGAAGAGAAAAACCCGGAGCAAGGGTG
GAATGGACTCAAGTACTTCTTGACTATAGTGGCGGTTTGTCTGAGAACTGCTTATAGT
ATCCAGAAAGGTCAAATCGCTTGGAGGGTTTTAGCAGCTATCTCCTCAGCTGCAGCT
GCAGTATTCAGTACTTACTGGGACTTTATCCATGATTGGGGTCTTCTAAACCGGACAT
CAAAAAACCGCTGGCTTCGTGATAAACTCCTCATTCCTCCCAAAGAAAGTATACTTCA
TCGCCATGATTTTGAATGTCCTGCTGAGATTTGCAT

>PHO1A_Splice_variant_3_(plasmid_23)

ATGACGTTTCGGGAAGGAGTTTTTCGTCTCAGATGGTACCGGAGTGGCAACAAGCTTAC
ATGGACTACGATTATCTCAAAACCCTTCTCAAAGAAATCATCCGCTTCAAACCTCCGG
GAAGACATAGACGCGGTAACCCTCACGACGTAGAGGAAGGGATACAGCTGACAGGG
ACGACGACGACGTCAGGGCCTATTCTTGTTAACAACACCGCGGACCGCGGCTACGAG
ACCACGTTCTTATGGCGGCGGAGGAAGGAGGAGAGTACGAGCTGGTGTTTTTCCGG
AGACTAGACGACGAGTTCAATAAAGTAAATAAGTTTTACAAAGAGAAAGTGGACGA
AGTGTGAAAGAAGCTGTGGTGCTTAACAACAGATGGACGCTTTGATCGCGTTTCG
TGTTAAAGTTGAGAATCCAGCAGGGTGGGGATGGGATGAACGAGCGGTGGAGATCA
CTCGCTTGGCTTCCGACATCGCTACTTCCGCGGCGGCTATCTCCGCTTCTACTCCCGC
CGGAGCTAAATCCATGAAAGTTCGGAGTCAAGCTCACATGGAGGCAATACAGGAAG

GAGGGTCGAGCAAAGCTGGGCAATTAGAAGATGATGAAGAGGAAGAAGCGCAAGC
AGAAATTGTAGCTTCCGTGTCTACCGGAGCTAGTGACGTGAGCACGACCAGGATGAG
AGCAGTGAGACCAGCTCCGTTGGATATTCTTGATCGAGTGACGATCAATAACACCAA
AGAGACGCCTCGTTCACCATCAAAAGAGTTCTACAGGTATCCAAGAACACTGATTT
AAAGTTCAGCAGAGAAAATCTGATGAAGGTCGAGGAGAACTCAGGCACGCTTTCA
TCGTGTTTTATCAGAAGCTTAGGCTTCTCAAGAGCTACAGCTTTTTGAATGTACTGGC
GTTCTCTAAGCTATTGAAGAAGTATGACAAGATTACTTCGAGGGGATGCAACCAAGCC
TTACATGAAAGTGGTTGATAGTTCATACCTCGGAAGCTCTGATGAAGTTGTGCGACT
CATGGAGCGTGTTGAAGCTACGTTCATAAAGCATTTTTGCAAATGCTAACCGAACCAA
AGGAATGAACATTTTACGGCCTCAGGCAAACGAGAGAGACATAGACTTACTTTCTC
CACAGGTTTCACGGCTGGATGCGTTTTCTCTTATAGTGGCTCTTGCCGCTATCATC
CGCACGCGAAATCTCTTGCAGGAGGAAGGCCAGAAGCAATACATGAATACTATGTTT
CCTCTTTATAGCTTCTTCGGTTTTATCGTGCTGCACATAATAATGTATGCTGCTAATAT
ATACTACTGGAGGCGATACAAAGTAACTATTCCTTCATATTTGGGTTCAAGCAAGG
AACTGAACCTGGTTATAGACAAGTCCCTGCTTGTGGGTTTCAGCATTGGAGTCTTTGCG
TTGCTTTGTGTTCTTGCTAATCTTGACATGGAGGCAAATCCCAAACCAAAGACTATA
AAACATTCACCGAACTTCTTCCTCTTTTCTACTTCTTGCCATGTTTGTGGTTCTAATC
TTGCCATTCAACTTCTTCTACCGGTCGAGTCGCTTTTTCTTCCTCACTTGTCTGTTCA
TTGCCTCGCAGCTCCTCTTTACAAGGTAACACTGCCTGATTTCTTCTTGGGAGATCAA
TTAACTAGCCAGGTTCAAGCTCTTCGAAGCATCGAGTTCTACATCTGTTACTATGGTT
GGGAGACTTCAGACACAGACAGAACACTTGTAACAAATCTAGTGTGTACAACACTT
TCTTATTCATTGTTGCTGTCATCCCTTACGCCTCTCGCCTCCTTCAGTGCCTGAGACGT
TTGTTCCAAGAGAAAAACCCGGAGCAAGGGTGGAAATGGACTCAAGTACTTCTTGACT
ATAGTGGCGGTTTGCTTGAGAAGTCTTATAGTATCCAGAAAGGTCAAATCGCTTGG
AGGGTTTTAGCAGCTATCTCCTCAGCTGCAGCTGCAGTATTCAGTACTTACTGGGACT
TTATCCATGATTGGGGTCTTCTAAACCGGACATCAAAAAACCGCTGGCTTCGTGATA
AACTCCTCATTCCCCAAAAGAAAGTATACTTCATCGCCATGATTTTGAATGTCCTGCT
GAGATTTGCAT

>PHO1A_Splice_variant_4_(_plasmid_26)

ATGACGTTCCGGAAGGAGTTTTCTGCTCTCAGATGGTACCGGAGTGGCAACAAGCTTAC
ATGGACTACGATTATCTCAAACCCCTTCTCAAAGAAATCATCCGCTTCAAACCTCAGA
ACCAACAATGCACCTACTCGTGGTGGCGCCAAGAATCATCAAGGCGGAGGATTAAA
CCGGAAGATGACTCTTTACCGAGCGTTTAGCGGTTTAGTCTCAACTCCGGGAAGACA
TAGACGCGGTAACCCTCACGACGTAGAGGAAGGGATACAGCTGACAGGGACGACGA
CGACGTCAGGGCCTATTCTTGTTAACAACACCGCGGACCGCGGCTACGAGACCACGT
TCCTTATGGCGGCGGAGGAAGGAGGAGAGTACGAGCTGGTGTTTTTCCGGAGACTAG
ACGACGAGTTCAATAAAGTAAATAAGTTTTACAAAGAGAAAGTGGACGAAGTGTTG
AAAGAAGCTGTGGTGCTTAACAACAGATGGACGCTTTGATCGCGTTTTCGTGTTAAA
GTTGAGAATCCAGCAGGGTGGGGATGGGATGAACGAGCGGTGGAGATCACTCGCTT
GGCTTCCGACATCGCTACTTCCGCGGCGGCTATCTCCGCTTCTACTCCCGCCGGAGCT
AAATCCATGAAAGTTCGGAGTCAAGCTCACATGGAGGCAATACAGGAAGGAGGGTC
GAGCAAAGCTGGGCAATTAGAAGATGATGAAGAGGAAGAAGCGCAAGCAGAAATT
GTAGCTTCCGTGTCTACCGGAGCTAGTGACGTGAGCACGACCAGGATGAGAGCAGTG
AGACCAGCTCCGTTGGATATTCTTGATCGAGTGACGATCAATAACACCAAAGAGACG
CCTCGTTCACCATCAAAAGAGTTCTACAGGTATCCAAGAACACTGATTTAAAGTTC
AGCAGAGAAAATCTGATGAAGGTCGAGGAGAACTCAGGCACGCTTTCATCGTGTTT

TATCAGAAGCTTAGGCTTCTCAAGAGCTACAGCTTTTTGAATGTACTGGCGTTCTCTA
AGCTATTGAAGAAGTATGACAAGATTACTTCGAGGGATGCAACCAAGCCTTACATGA
AAGTGGTTGATAGTTCATACCTCGGAAGCTCTGATGAAGTTGTGCGACTCATGGAGC
GTGTTGAAGCTACGTTCATAAAGCATTGCAAAATGCTAACCGAACCAAGGAATGA
ACATTTTACGGCCTCAGGCAAACGAGAGAGACATAGACTTACTTTCTCCACAGGTT
TCACGGCTGGATGCGTTTTCTCTTTATAGTGGCTCTTGCCGCTATCATCCGCACGCG
AAATCTCTTGCAGGAGGAAGGCCAGAAGCAATACATGAATACTATGTTTCCTCTTTA
TAGCTTCTTCGGTTTTATCGTGCTGCACATAATAATGTATGCTGCTAATATATACTAC
TGGAGGCGATACAAAGTAAACTATTCCTTCATATTTGGGTTCAAGCAAGGAAGTAA
CTTGTTATAGACAAGTCTGCTTGTGGGTTTCAGCATTGGAGTCTTTGCGTTGCTTT
GTGTTCTTGCTAATCTTGACATGGAGGCAAATCCCAAACCAAGACTATAAAACAT
TCACCGAACTTCTTCCTCTTTTCTACTTCTTGCCATGTTTGTGGTTCTAATCTTGCCA
TTCAACTTCTTCTACCGGTCGAGTCGCTTTTTCTTCTCACTTGTCTGTTTCATTGCCTC
GCAGCTCCTCTTACAAGGTAACACTGCCTGATTTCTTCTTGGGAGATCAATTAAC
GCCAGGTTCAAGCTCTTCGAAGCATCGAGTCTACATCTGTTACTATGGTTGGGGAG
ACTTCAGACACAGACAGAACACTTGTAAACAATCTAGTGTGTACAACACTTTCTTATT
CATTGTTGCTGTCATCCCTTACGCCTCTCGCCTCCTTCAGTGCCTGAGACGTTTGTTCG
AAGAGAAAAACCCGGAGCAAGGGTGAATGGACTCAAGTACTTCTTACTATAGTG
GCGGTTTGTGTTGAGAACTGCTTATAGTATCCAGAAAGGTCAAATCGCTTGGAGGGTT
TTAGCAGCTATCTCCTCAGCTGCAGCTGCAGTATTCAGTACTTACTGGGACTTTATCC
ATGATTGGGGTCTTCTAAACCGGACATCAAAAACCGCTGGCTTCGTGATAAACTCC
TCATTCCCAAAGAAAGTATACTTCATCGCCATGGTGAGTCTCAAACCAACAAT
ATTTTTATAACAGAACATGTTTCACAGCAACTGATGTTTGGCACATGTATCTCAACA
GATTTTGAATGTCCTGCTGAGATTTGCAT

>PHO1A_Splice_variant_5_(_plasmid_22)

ATGACGTTTCGGGAAGGAGTTTTCGTCTCAGATGGTACCGGAGTGGCAACAAGCTTAC
ATGGACTACGATTATCTCAAACCCCTTCTCAAAGAAATCATCCGCTTCAAACCTCAGA
ACCAACAATGCACCTACTCGTGGTGGCGCCAAGAATCATCAAGGCGGAGGATTA
CCGGAAGATGACTCTTTACCGAGCGTTTAGCGGTTTAGTCTCAACTCCGGGAAGACA
TAGACGCGGTAACCCTCACGACGTAGAGGAAGGGATACAGCTGACAGGGACGACGA
CGACGTCAGGGCCTATTCTTGTTAACAACACCGCGGACCGCGGCTACGAGACCACGT
TCCTTATGGCGGCGGAGGAAGGAGGAGTACGAGCTGGTGTTTTTCCGGAGACTAG
ACGACGAGTTCAATAAAGTAAATAAGTTTTACAAAGAGAAAGTGGACGAAGTGTTG
AAAGAAGCTGTGGTGCTTAACAACAGATGGACGCTTTGATCGCGTTTTCGTGTTAAA
GTTGAGAATCCAGCAGGGTGGGGATGGGATGAACGAGCGGTGGAGATCACTCGCTT
GGCTTCCGACATCGCTACTTCCGCGGCGGCTATCTCCGCTTCTACTCCCGCCGGAGCT
AAATCCATGAAAGTTCGGAGTCAAGCTCACATGGAGGCAATACAGGAAGGAGGGTC
GAGCAAAGCTGGGCAATTAGAAGATGATGAAGAGGAAGAAGCGCAAGCAGAAATT
GTAGCTTCCGTGTCTACCGGAGCTAGTGACGTGAGCACGACCAGGATGAGAGCAGTG
AGACCAGCTCCGTTGGATATTCTTGATCGAGTGACGATCAATAACACCAAAGAGACG
CCTCGTTCCACCATCAAAGAGTTCTACAGGTATCCAAGAACACTGATTTAAAGTTC
AGCAGAGAAAATCTGATGAAGGTCGAGGAGAACTCAGGCACGCTTTCATCGTGTTT
TATCAGAAGCTTAGGCTTCTCAAGAGCTACAGCTTTTTGAATGTACTGGCGTTCTCTA
AGCTATTGAAGAAGTATGACAAGATTACTTCGAGGGATGCAACCAAGCCTTACATGA
AAGTGGTTGATAGTTCATACCTCGGAAGCTCTGATGAAGTTGTGCGACTCATGGAGC
GTGTTGAAGCTACGTTCATAAAGCATTGCAAAATGCTAACCGAACCAAGGAATGA
ACATTTTACGGCCTCAGGCAAACGAGAGAGACATAGACTTACTTTCTCCACAGGTT
TCACGGCTGGATGCGTTTTCTCTTTATAGTGGCTCTTGCCGCTATCATCCGCACGCG
AAATCTCTTGCAGGAGGAAGGCCAGAAGCAATACATGAATACTATGTTTCCTCTTTA
TAGCTTCTTCGGTTTTATCGTGCTGCACATAACTGTATGCTGCTAATATATACTACT
GGAGGCGATACAAACATTGGAGTCTTTGCGTTGCTTTGTGTTCTTGCTAATCTTGACA

TGGAGGCAAATCCCAAACCAAAGACTATAAAAACATTACCCGAACCTTCTTCCTCTTT
TCCTACTTCTTGCCATCTCCTCTTTACAAGGTAACACTGCCTGATTTCTTCTTGGGAGA
TCAATTAAGTAGCCAGGTTCAAGCTCTTCGAAGCATCGAGTTCTACATCTGTTACTAT
GGTTGGGGAGACTTCAGACACAGACAGAACAACCTTGTAACAAATCTAGTGTGTACAAC
ACTTTCTTATTCATTGTTGCTGTCATCCCTTACGCCTCTCGCCTCCTTCAGTGCCTGAG
ACGTTTGTTCGAAGAGAAAAACCCGGAGCAAGGGTGAATGGACTCAAGTACTTCTT
GACTATAGTGGCGGTTTGCTTGAGAACTGCTTATAGTATCCAGAAAGGTCAAATCGC
TTGGAGGGTTTTAGCAGCTATCTCCTCAGCTGCAGCTGCAGTATTTCAGTACTTACTGG
GACTTTATCCATGATTGGGGTCTTCTAAACCCGGACATCAAAAAACCGCTGGCTTCGT
GATAAACTCCTCATTCCCAAAGAAAGTATACTTCATCGCCATGATTTTGAATGTCC
TGCTGAGATTTGCAT

>PHO1_D (Plasmid 24)

ATGACGTTCCGGTAAGGAGTTTTTCATCTCAGATGGTGCCAGAGTGGCAACAAGCTTAC
ATGGATTACGGTTCTCTCAAAAAATCTCTCAAAGAAATCATCGGGTTCAAACCTCGGA
CAACCGGCCGATGCATTGCCTACTCGTGGTGGCGCAATAATCACCACGGCGGAGGA
TTACACCGAAGAATGACTTTGTACAAGACGTTTAGCGGTTTACTCTCGATTCAGGGA
AGACAAAGACGCGGCCATTCTCACGACGTGGAGGAAGGGTTACAGATGATGGCGAC
CACTGGGCCGATTCTGGTGGAGACCAACGCTGACGGCAGCTATGAAACTACGTTTCT
GATGGTGAATGAGAAAGGAGGAGAGTACGAGTTGGTGTTTTTCCGGAGACTGGATG
AGGAGTTCAATAAAGTGAGTAAGTTTTACAAGAGGAAGTTGATAAAGTGTCGAAA
GAGGAGAAGGAGCTTAACAAACAGATGGAAGCTTTGATAGCATTTCGTGCTAAAGTT
GAGAATCAAGAAGGGAATGGATGGCGATTGCAGGAACGTGGGGTTGAGTTCACTCG
ATTGACTTCCAACATCGCTACTTCCGCAGCGGCTATCTATGCTTCTTCTCCCGCGGGT
GCTAAGTCCAAGGAAGTTGGAAGTCAAGCTCATATGGAGACAATAGAGGAAGGAGA
GTCGAGCAGAGCTGGGAACTAAAAGATGATGAAGAAGCAGAAGATAATGGAGCTA
GAGAAGAGCCCAATGAAGTGAATAAGAGCAGGACGAGAGCAGCGAGACCTGCGCC
GTTAGATATTCTGGATGGAGTGAGGATCAACAACACCATAGAGACACCTCGTTCCAC
CTTTAGAGGAATTCTTAAAGTATCAAAGCCGACCGAGTTAGTATATAACAGAGAAAA
TCTGAAGAAAGCCGAGCAGATGCTCATAGAAGCTTTCAGCGTGTTTTATCAAAGCT
TCGGCTTCTCAAAGCTACAGCTTTTTGAATGTAAAGGCGGTCTCGAAGATATTGAA
GAAGTACGATAAGATAACTTCAAGGGATGCAACAAAGCCTTACATGAAAATGGTTG
ATAGTTCATATCTAGGAAGCTCTGATGAAGTTGTACGACTCATGGAGCATGTTGAAG
CTACATTCATAAAAACATTTTGCAAATGCTAATCGAACCAAAGGAATGAACATTTTAC
GGCCTAAAGCAAACGAGAGAGACACAGACTTACATTCTCCACTGGTTTCTCGGCTG
GATGCATATTCTCTCTTATAGTGGCTCTTGCTGCGATCATTTCGCACGCGAAATTTCTT
GCAGGAGGAAGGCCAGGCGCAATACATGGATACTATGTTCCCACTTTATAGGAAGG
AACTGAACCTGGGTTACAGACAAGTCCTCCTTGTGGCGTTCAGCACTGGCGTCCTTGC
ATTGCTTTGCGTTCTTGCAAATCTTGACATGGAGTCATGCCATTCAACTTCTTTTACCG
GTCCAATCGCATATTCTTCTCCTCACTTGTCTGTTTCATTGCATGGCCGCTCCTCTTTACA
AAGTAACGTTGCCTGATTTCTTCTTGGGAGATCAATTAAGTAGCCAGGTTCAAGCTTT
ACGAAGCATCGAGTTCTACATCTGTTACTATGGTTGGGGCGACTTCAGACACAGAAA
CAACACTTGTAACAAGTCTGGTGCGTACAAAGGTTTCTTCTTCATTGTTGCTGTCATC
CCTTATGTCTCTCGCCTCCTTCAGTGCCTGAGACGTCTGTTTGAAGAGGACAATCCGG
AGCAAGGATGGAATGGGCTCAAGAACTTCTTGACTATAGTGGCAATTTGCTTGAGAA
CTGCTTACAGTATCCAGCAAGGTCAAATCGCTTGGAGGGTTTTAGCAGCTATCTCTTC
AGCTGCAGCTGCCATATTCTGCACTTACTGGGACTTTATCCATGATTGGGGTCTTCTA
AAGCGGACATCCAAAAACCGTTGGCTTCGAGATAAACTCCTCATTCTCAAAGAAA
GTATACTTCATCGCCATGATTTTGAATGTCTGCTGAGATTTGCAT

>PHO1_B (Plasmid 12)

ATGACGTTCCGGGAAGGAGTTTTTCGTCTCAGATGGTGTTCAGAGTGGCAACAAGCTTAC
ATGGACTACGATTATCTGAAAAAATGTCTTAAACACAAGGATCACCACGGCGGAGG
ATTAAACCGGAAAAAGACTCTGGACAGCCTATTTAGCTGTTTACAACCTTTTACTCCG
CGAAGACAAAGACTCGGTAATTCTCAAGACGTCGAGGGAGGGGTGCAATTGAAGAC
GAAAACACTACGACGACGACTGGGCCGATTTCAGGTGGAGACCACCGCGGATGGCCAGG
TCCAGACGAAGTTTTCTGATGGAGGAGGAAGTAGGAGAGTACGAGAGAGAGTTTTTC
GGACGACTAGATAAGGAATTCAATAAAGTCATAAAGTTTTACGAAGAGAAAAGTTGA
GAAAGTGTTGAAAGAAGAAGAAGGGCTTAAAGGAGAGCTGGAAGCTTTGATCACGG
CCGGTTCTAAGCATAACGATATAAAATATTCGCGTTGGAAAAGAGTAGCTAATATAG
CTTCTCGTGCTAAAGTTGAGAATCCAGAAAGGTGCGGACGGGAAGAAAAGGGATGAA
GAGATCAGTCCACTGGTGGCGGCTATCTCTCCTTCTAGTCCCCTCGGAGCTAAATCCG
TGATAGAGGAAGGAGAGTCGAGCAGAGTTGGTCAATCAGAAGATGGAGCCATAGAA
TCGGCCGATGCAATGGATACGAGCATGATGAGCAACAACACCAGAGGAAAATATTA
GGAATCGAAGAAGAGCGAGTTAAGTTTTAAACAAAACAAAACACTGAGGGAAAAGGAG
AAGAACTGGAACACGCTTACGTCACGTTATATCGGAAGCTTCTGGATATCCAAAAC
TACAGTTTTTCGAATGCAAAGGCGTTCTCGAAGATATTGAAGAAGTATGAGAAGATA
ACTTTGAGGGATGCAAGGAAGCCTTACATGGAAGTGGTTGATAGTTCATACCTCGGA
AGCTCTGATGAAGTCGTGCGACTCACAAAGCGTGTGGAAGAGGACTTCATAAAACAT
TTTGCAGAAGATAATCGAACCAAGGAATGAACATTCTACGGCCTAAAATGACACG
AGAGAGACACAGACTTACATTCTCCACAGGTTTCTCGGCTGGATGCGTTTTCTCTCTT
ATAGTGGCTCTTGTCTCGATCATCCGCACCCGAAATATCTTGCAGAACGATGGCCAG
GATGTATACaTGAATACTATGTTCCCCTTTATAGCTGGTTCCGTTTTGTAATGCTGCA
TATCATCGTGTATGCTGCGAATATACTACTGGAGGCGATACAGAATAAACTATTC
CAAAATATTGGATTTCAAGCAAGGAAGTGAAGTGGACACAGACAAGTCCTGCTTGT
GGCTTTCAGCATTGGCGTCTTTCATTGCTTGTGTTCTTGCAAATCTTGACATGGAG
GTGGATCCTGAAACGAAAGACTATCAAGCATTACCGAAGTCTTCTCTTTTCTTCTAC
TTATTGGTATATTTGGGGTTCTATTCCTGCCATTCAAATCTTTTATCACTCGAAACGC
CAATTCTACCTTACATGTTTGTTCATTGCATCGCCGCTCCTTTTATAAGGTATCATT
TGCTGATAGCTTCTTGGGCGATCAATTCAGTCCAGGTTCAAGTCTTTCGAAGCTTC
GAGTTCTACATATGTTACTATGGTTGGGGTGAAGTTCAAACACAGAGAGAACAGTTGT
ACCAAGTCTCATGCGTACAACAGTTTCTTCTTCATAGTTGCTGTTCATCCCTTTCGCTC
TCGCCTCCTCAGTGCCTGCGACGGTTGTTTGATGATAAAGACCCGGATCACAGGTG
GAACGGACTCAAGTACTTCTTGACAATAGTGGCGGTTTGCTTCAGAACTGCTTACAG
TATCCAGCCAGCTCAAATCGCTTGAGAGTTTTAGCTATTATCTTCTCAGTCGTAGCT
GCAATATTCGGTACTTACTGGGACTTTGTCCATGACTGGGGTCTTCTAAACCAGAAAT
CCGAAAACCGCTGGCTTCGGGATAACCTCCTCATTCTCAAAGGAAGTATACTTTA
TTGCTATGATTTTGAATGTCTGCTGAGATTTGCATA

APPENDIX 2. Script used to process BAM files and HISAT in RNA-seq analysis.

```
dir="/local/john/Tuxedo"
dir2="/local/john/Tuxedo/STRG"

for sample in local/john/Tuxedo/*.dedup.bam
do
base=$(basename ${sample} '_accepted_hits.dedup.bam')
stringtie -p 16 -G /local/john/Tuxedo/genes/GCF_000686985.2_Bra_napus_V2.0_genomic.gff \
  -0 ${dir2}/stringtie_output/${base}/${base}_Bna_v2.gtf \
  -1 ${dir2}/stringtie_output/${base}/${base} ${dir}/${base}_accepted_hits.dedup.bam
Done
1s ${dir2}/stringtie_output/*/*_Bna_v2.gtf >mergelist.txt;|
stringtie --merge -p 16 -G
/local/john/Tuxedo/genes/GCF_000686985.2_Bra_napus_v2.0_genomic.gff \
-0 ${dir2}/stringtie_output/stringtie_merged.gtf mergelist.txt
```

This generated a merged .gtf file for the sequences, which was then used to generate StringTie outputs for Ballgown using the following script;

```
dir="/local/john/Tuxedo"
dir2="/local/john/Tuxedo/STRG"
for sample in /local.john/Tuxedo/*dedup.bam
do
base=$(basename ${sample} '_accepted_hits.dedup.bam')
stringtie -B -p 16 -G ${dir2}/stringtie_output/stringtie_merged.gtf \
  -0 ${dir2}/ballgown/${base}/${base}_Bna_v2.gtf \
  -1 ${dir2}/stringtie_output/${base}/${base} ${dir}/${base}_accepted_hits.dedup.bam
done
```

The following commands were then used in R to analyse the data

```
library(ballgown)
library(RSkittleBrewer)
library(genefilter)
library(dplyr)
library(devtools)

setwd("/local/john/Tuxedo")

pheno_data = read.csv("pheno_all.csv")

bg_chrX = ballgown(dataDir = "STRG/ballgown", samplePattern = "HAMMM", pData=pheno_data)

## Filter to remove low abundance genes. One common issue with RNA-seq data is that genes
often have very few or zero counts.
bg_chrX_filt = subset(bg_chrX,"rowVars(expr(bg_chrX)) > 1",genomesubset=TRUE)

## Identify transcripts that show statistically significant differences between groups.
##The statistically test uses a cumulative upper quartile normalization.
results_transcripts = statest(bg_chrX_filt, feature="transcript", covariate="Treatment", \
```

```

adjustvars = c("Line"), getFC=TRUE, meas="FPKM")

##Identify genes that show statistically significant differences between groups
results_genes = stattest(bg_chrX_filt, feature="gene", covariate="Treatment", adjustvars = c
("Line"), getFC=TRUE, meas="FPKM")

##Add gene names and gene IDs to the results_transcripts and results_genes data frame:
results_transcripts = data.frame(geneNames=ballgown: :geneNames(bg_chrX_filt), \
geneIDs=ballgown: :geneIDs(bg_chrX_filt), results_transcripts)
indices <- match(results_genes$id, expr(bg_chrX_filt, 'all')$gene_id)
gene_names_for_result <- expr(bg_chrX_filt, 'all')$gene_name[indices]
results_genes <- data.frame(geneNames=gene_names_for_result, results_genes)

##Sort the results from smallest p-value to the largest:
results_transcripts = arrange(results_transcripts,pval)
results_genes = arrange(results_genes,pval)

## Write the results to a CSV (comma-separated values) file that can be shared and distributed:
write.csv(results_transcripts, "chrX_transcript_results.csv", row.names=FALSE)
write.csv(results_genes, "chrX_gene_results.csv", row.names=FALSE)

## Identify transcripts and genes with a q-value of less than 0.05:
subset(results_transcripts,results_transcripts$qval<0.05)
subset(results_genes,results_genes$qval<0.05)

## Show the distribution of the gene abundances (measured as FPKM values) across samples,
colored by Treatment
## The first command below accesses the FPKM (Fragments Per Kilobase of transcript per Million
mapped reads) data
## Then transform the FPKM data to easily visualise the boxplot
fpkm = expr(bg_chrX,meas="FPKM")
fpkm = log2(fpkm+1)
boxplot (fpkm, col=as.numeric(pheno_data$treatment), las=2, ylab="log2(FPKM+1)")

## create an MA plot
library(ggplot2)
library(cowplot)
results_transcripts$mean <- rowMeans(expr(bg_chrX_filt)) ggplot(results_transcripts,
aes(log2(results_transcripts$mean), \
log2(results_transcripts$fc), colour = qval<0.05)) + scale_color_manual(values=c("#999999",
#FF0000")) + geom_point() + \
geom_hline(yintercept=0)

```

APPENDIX 3. List of up-regulated genes of *Brassica napus* in response to Pi availability. Genes with q value < 0.05 and fold change value > 2.0.

	Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
	43.73	4.00553E-11	8.90979E-08	A9	rna57637	LOC106369031	gene46215	inorganic pyrophosphatase 2-like
	37.97	1.34929E-11	4.09149E-08	A2	rna8416	BNAA02G05080D	gene6677	uncharacterized protein BNAA02G05080D
	37.21	5.36354E-09	5.28332E-06	C1	rna68425	LOC111197984	gene54883	inorganic pyrophosphatase 1
	30.55	1.92122E-07	7.80235E-05	C2	rna73153	LOC106381279	gene58929	inorganic pyrophosphatase 1 BnaC02g21720D
	27.79	5.64815E-12	2.27017E-08	A2	gene5709	LOC106425289	gene5709	
248	25.85	3.75882E-08	2.2796E-05	A2	rna10456	LOC106399564	gene8350	inorganic pyrophosphatase 1 BnaAnng30500D
	24.18	5.57417E-09	5.33771E-06	C8	rna120845	LOC106356588	gene97738	inorganic pyrophosphatase 2 BnaC08g17490D
	23.29	2.25819E-13	3.10667E-09	Un	rna141842	LOC106435863	gene114947	sulfoquinovosyl transferase SQD2-like
	23.10	1.95322E-12	1.18456E-08	C7	rna112703	LOC106435115	gene90945	uncharacterized protein LOC106435115
	21.46	7.10265E-11	1.36027E-07					
	20.25	2.66454E-15	9.69571E-11	C2	gene55977	LOC106359556	gene55977	
	19.92	3.7162E-10	6.1466E-07	C2	rna70816	LOC106379199	gene56935	SPX domain-containing protein 1 BnaA02g04730D
	19.33	4.08524E-11	8.90979E-08	Un	rna146238	LOC111213425	gene118531	fructose-1,6-bisphosphatase, cytosolic- like

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
16.93	4.4665E-07	0.000151895	A8	rna49263	LOC106391533	gene39559	inorganic pyrophosphatase 2-like
14.95	6.23879E-12	2.27017E-08	C3	gene69814	LOC106356087	gene69814	
14.54	1.58996E-08	1.20532E-05	A10	rna63521	LOC106435170	gene50887	sulfoquinovosyl transferase SQD2-like
13.13	7.6302E-08	3.91053E-05	A6	rna39257	LOC106452020	gene31386	purple acid phosphatase 17 BnaA05g22460D purple acid phosphatase 17 PM
12.90	5.76142E-11	1.1647E-07	C1	rna69060	LOC106429291	gene55406	purple acid phosphatase 17-like BnaC01g34650D purple acid phosphatase 17
12.49	2.51251E-09	2.77046E-06	C9	rna126723	LOC106392218	gene102544	monogalactosyldiacylglycerol synthase 3, chloroplastic
11.70	1.71318E-06	0.000415594	Un	rna132650	LOC111210345	gene107414	sucrose-phosphate synthase 1-like
11.48	9.20483E-09	8.1694E-06	A3	rna14445	LOC106434609	gene11615	SPX domain-containing protein 1-like
10.58	5.94397E-10	8.31882E-07	C3	rna79731	LOC106386711	gene64227	SPX domain-containing protein 1-like
10.48	2.30423E-08	1.52448E-05	Un	rna143570	LOC106393743	gene116344	UPF0496 protein At3g19330-like
10.44	2.13617E-06	0.000476878	Un	rna143296	LOC106436517	gene116127	glycerophosphodiester phosphodiesterase GDPD1, chloroplastic-like
10.39	7.09877E-13	5.1662E-09	A3	rna13382	LOC111214530	gene10746	

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
9.62	7.43549E-10	1.00208E-06	A9	rna51238	LOC106366315	gene41099	monogalactosyldiacylglycerol synthase 3, chloroplastic-like
9.26	1.20205E-11	3.97638E-08	C4	rna94550	BNACNNG14700D	gene76374	uncharacterized protein BNACNNG14700D
9.14	1.95245E-08	1.34049E-05	C3	rna84201	LOC106386978	gene67852	proline-rich protein 3-like
8.83	5.07673E-10	7.68628E-07	A1	rna5282	LOC106347425	gene4208	purple acid phosphatase 17-like purple acid phosphatase 17
8.70	9.18783E-07	0.00027181	A10	rna61068	LOC106371711	gene48984	uncharacterized protein LOC106371711
8.40	0.001447802	0.029308099	A3	rna17362	LOC106438827	gene13914	ribonuclease II, chloroplastic/mitochondrial-like
7.70	0.001648607	0.031490556	Un	rna139024	LOC106407032	gene112679	probable desiccation-related protein LEA14
7.68	0.000168218	0.008229957	A7	rna41371	LOC106356366	gene33101	uncharacterized protein LOC106356366
7.37	0.000758649	0.019774872	Un	rna139438	LOC106434614	gene113030	protein CHAPERONE-LIKE PROTEIN OF POR1, chloroplastic-like
7.32	1.78846E-08	1.25151E-05	A7	rna42768	LOC106353436	gene34253	EP1-like glycoprotein 1

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
7.16	1.26684E-05	0.001559168	A9	rna51105	LOC106366144	gene40983	soluble inorganic pyrophosphatase 4 BnaA09g33870D
7.07	0.002140891	0.036591229	Un	rna134172	LOC106426623	gene108645	S-adenosylmethionine synthase
6.95	6.18751E-07	0.000199249					
6.91	0.002467329	0.039566392	C1	rna68572	LOC106376898	gene54995	synaptotagmin-1
6.83	1.6777E-07	7.26763E-05	A5	rna31638	LOC106452187	gene25404	U-box domain-containing protein 33-like
6.80	2.73667E-05	0.002641446					
6.65	4.39713E-10	6.95664E-07	A10	rna62241	LOC106372348	gene49897	elongation factor 1-beta 1
6.58	1.09367E-08	9.14375E-06					
6.28	3.41505E-13	3.10667E-09					
6.09	8.2644E-05	0.005389334	C3	rna85396	BNAC03G44620D	gene68866	uncharacterized protein BNAC03G44620D
5.95	0.000238261	0.010114195	C7	rna116774	LOC106420885	gene94302	5'-adenylylsulfate reductase 3, chloroplastic
5.93	4.30522E-06	0.000791203	C1	rna67404	LOC106376672	gene54046	phosphoglycerate mutase-like protein 2
5.85	3.44729E-08	2.1261E-05					
5.82	4.16254E-11	8.90979E-08	A8	rna46497	LOC106360679	gene37252	ATP phosphoribosyltransferase 1, chloroplastic

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
5.68	8.69023E-07	0.000259197	A8	rna49926	LOC106362324	gene40059	phosphoenolpyruvate carboxylase kinase 1-like
5.61	5.12928E-08	2.91631E-05	C5	rna99309	LOC106415678	gene80229	phosphoenolpyruvate carboxylase kinase 1-like
5.50	8.81656E-08	4.39475E-05	C9	rna129110	LOC106390526	gene104539	putative glycerol-3-phosphate transporter 1
5.46	0.000268828	0.01088338	A3	rna17991	LOC106438924	gene14394	uncharacterized protein At1g04910
5.42	6.02427E-05	0.004366754	A3	rna21047	LOC106440277	gene16835	probable mannan synthase 1
5.26	2.17152E-08	1.46328E-05					
5.15	9.46378E-07	0.000277716	C3	rna87747	LOC106386109	gene70840	protein RADIALIS-like 3 BnaA08g15280D BnaC03g76670D
5.14	0.000187779	0.008784284	A3	rna20200	BNAA03G57760D	gene16134	uncharacterized protein BNAA03G57760D
5.13	0.002040794	0.035702119	A1	rna5809	LOC106387275	gene4622	thioredoxin F-type, chloroplastic thioredoxin-f
5.13	0.001705302	0.032201628	A4	rna24680	LOC106445942	gene19768	D-glycerate 3-kinase, chloroplastic BnaC06g40610D
5.10	3.12107E-05	0.002897184	A7	rna43858	LOC106355931	gene35160	phosphate transporter PHO1 homolog 1-like
5.05	7.25772E-05	0.004964175					

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
4.96	9.49099E-08	4.667E-05	C1	rna64106	LOC106375246	gene51361	phosphoenolpyruvate carboxylase kinase 1-like
4.90	1.26831E-05	0.001559168	A3	rna14505	LOC106441542	gene11660	uncharacterized protein LOC106441542
4.89	0.000120999	0.006793541					
4.87	0.000338896	0.012469876	A4	rna23941	LOC111214990	gene19194	calmodulin-5
4.86	1.63459E-07	7.1827E-05	C5	rna99272	LOC106413829	gene80202	uncharacterized protein LOC106413829
4.79	5.24025E-12	2.27017E-08	A6	rna33062	BNAA06G01330D	gene26509	uncharacterized protein BNAA06G01330D
4.72	1.57996E-08	1.20532E-05	A3	rna16276	LOC106438156	gene13076	putative glycerol-3-phosphate transporter 1 BnaA06g17430D
4.69	2.61287E-11	7.31361E-08	C9	rna131210	LOC106424605	gene106263	elongation factor 1-beta 1 BnaC09g44570D
4.66	1.72192E-05	0.001928779	C7	rna118401	LOC106406923	gene95646	heptahelical transmembrane protein 4-like
4.64	3.31291E-08	2.07845E-05					
4.62	7.64545E-07	0.000233784	A8	rna48223	LOC106361376	gene38678	cytochrome P450 79B1-like
4.61	0.002191568	0.03705706	Un	rna145725	LOC106401032	gene118120	phosphoinositide phospholipase C 2
4.55	3.15326E-06	0.000637449					

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
4.55	4.69061E-05	0.003765476	C3	rna89087	LOC106390094	gene71925	probable sucrose-phosphatase 1
4.54	5.00419E-06	0.000878806	C8	rna121976	LOC106367810	gene98639	purple acid phosphatase 22
4.53	1.39529E-06	0.000362655	C8	rna122369	LOC106360814	gene98981	UTP--glucose-1-phosphate uridylyltransferase 3, chloroplastic
4.53	0.000502765	0.015690053	A1	rna492	LOC106454413	gene408	uncharacterized protein At2g17340-like
4.49	0.000722443	0.019291802	C4	rna98088	LOC106370198	gene79282	SPX domain-containing protein 3 BnaC04g50120D
4.44	0.000121167	0.006793541	A4	rna25239	LOC106446405	gene20244	tryptophan N-monooxygenase 2
4.42	2.31533E-05	0.002373242	A9	rna53051	LOC111200956	gene42561	
4.38	1.84686E-07	7.72454E-05	A8	rna48330	LOC106361468	gene38771	peptidyl-prolyl cis-trans isomerase CYP18-3
4.37	3.55862E-07	0.000128209	Un	rna141210	LOC106440981	gene114442	protein C2-DOMAIN ABA-RELATED 8 BnaAnng37440D BnaC05g18470D
4.33	2.48184E-07	9.50624E-05	A1	rna3723	LOC106437919	gene2995	peptidyl-prolyl cis-trans isomerase CYP18-3
4.28	3.55676E-07	0.000128209	C3	rna87416	LOC106357692	gene70558	peptidyl-prolyl cis-trans isomerase CYP18-3
4.24	4.06782E-07	0.000143708	A9	rna58094	LOC106369102	gene46563	probable inactive purple acid phosphatase 1 BnaA09g46120D
4.24	0.000530819	0.016245106	A9	gene42560	LOC106451228	gene42560	

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
4.23	1.33231E-10	2.30857E-07					
4.19	1.48469E-06	0.000380456					
4.18	1.10565E-08	9.14375E-06	A6	rna33655	LOC106346346	gene26981	uncharacterized protein LOC106346346
4.16	0.000336905	0.012446005	A1	rna2954	LOC106444872	gene2367	uncharacterized protein LOC106444872
4.14	6.88225E-06	0.001065666	A2	rna11343	LOC106445850	gene9077	uncharacterized protein LOC106445850 BnaA04g07060D
4.12	4.48467E-06	0.00081594	A4	rna25666	LOC106449964	gene20583	uncharacterized protein LOC106449964
4.10	4.27803E-07	0.000146858	A5	rna32922	LOC106453258	gene26402	mitochondrial outer membrane protein porin 1 BnaA05g37480D
4.07	4.45618E-08	2.61535E-05	C5	rna99173	LOC106411631	gene80125	
4.06	5.28078E-10	7.68628E-07	C5	rna99264	LOC106415517	gene80196	uncharacterized protein LOC106415517
4.06	1.63835E-07	7.1827E-05	A1	rna543	LOC106434737	gene448	protein NUCLEAR FUSION DEFECTIVE 4-like
4.05	2.49682E-05	0.002505842	C8	rna122091	LOC106382798	gene98730	soluble inorganic pyrophosphatase 4 BnaC08g24660D

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
3.94	0.00354487	0.048878636	A2	rna7863	LOC106367226	gene6225	3-isopropylmalate dehydrogenase 3, chloroplastic BnaA02g02020D
3.94	0.00011946	0.006718581	C2	gene60694	LOC106347046	gene60694	
3.93	1.53914E-06	0.000388932	Un	rna143541	LOC106372748	gene116320	tricyclene synthase, chloroplastic-like
3.88	0.001834567	0.033579596	Un	rna141816	LOC106395965	gene114924	delta-1-pyrroline-5-carboxylate synthase A BnaC04g05620D
3.86	0.000153126	0.007803841	A3	rna21521	LOC106440433	gene17219	5'-adenylylsulfate reductase 3, chloroplastic-like
3.85	3.55246E-05	0.00313755	A3	rna13742	LOC106430678	gene11030	amino acid permease 2 BnaA03g02650D amino acid permease 2-1
3.80	5.27272E-05	0.004039238	C3	gene64260	LOC106386991	gene64260	
3.80	0.0016844	0.031906272	C4	rna97055	LOC106391437	gene78465	BEL1-like homeodomain protein 1
3.78	1.14317E-05	0.001480343	A4	rna26417	LOC106391472	gene21206	1,4-alpha-glucan-branching enzyme 2-1, chloroplastic/amyloplastic-like
3.78	0.000674332	0.018617644	C1	rna63990	LOC106370064	gene51256	protein RADIALIS-like 3
3.77	2.73989E-05	0.002641446	A6	rna35362	LOC111198585	gene28315	uncharacterized protein LOC111198585
3.76	0.001122175	0.025128776	C8	rna125258	LOC106419858	gene101305	choline-phosphate cytidyltransferase 1 BnaAnng16260D

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
3.76	0.000981916	0.023337663	A2	rna8266	LOC106425767	gene6559	uncharacterized protein LOC106425767
3.75	0.00221164	0.037325042	A10	rna63485	LOC106372572	gene50860	uncharacterized protein At5g02240-like
3.74	8.20129E-06	0.001188959	A2	rna9943	LOC106420320	gene7936	phosphate transporter PHO1 homolog 1-like
3.68	4.32626E-05	0.003537613	C4	rna97266	LOC106402737	gene78627	haloacid dehalogenase-like hydrolase domain-containing protein Sgpp
3.66	1.20966E-09	1.46723E-06	Un	rna135359	LOC106366906	gene109653	probable tyrosine-protein phosphatase At1g05000 BnaC09g22970D
3.66	0.000344142	0.012617722	Un	rna139602	LOC106433422	gene113155	homeobox-leucine zipper protein HAT5-like
3.65	4.82181E-07	0.000162459	C7	rna117731	BNAC07G42220D	gene95086	uncharacterized protein BNAC07G42220D
3.63	2.6793E-06	0.000570143	A8	rna49208	LOC106362012	gene39512	
3.60	0.000595118	0.017296448	Un	rna140233	LOC106435288	gene113677	putative glycerol-3-phosphate transporter 1
3.59	3.09304E-05	0.00288065	C4	rna94001	LOC106440702	gene75944	soluble inorganic pyrophosphatase 4- like
3.52	8.30343E-10	1.07909E-06	C9	rna126926	LOC106375529	gene102714	threonine synthase 1, chloroplastic-like

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
3.52	2.56685E-06	0.000552677	Un	rna140190	LOC106435266	gene113646	glycerophosphodiester phosphodiesterase GDPD2-like
3.50	4.46354E-12	2.27017E-08	A9	rna51082	LOC106364108	gene40962	probable tyrosine-protein phosphatase At1g05000 BnaA09g20630D
3.49	1.22152E-05	0.001519371	A2	rna9818	BNA02G13420D	gene7829	uncharacterized protein BNA02G13420D
3.47	2.66323E-05	0.002612117	A6	rna37558	LOC106352037	gene30063	monothiol glutaredoxin-S9 BnaA04g17610D
3.47	0.003493881	0.048432507	C9	rna127991	LOC106416478	gene103592	phosphoglycolate phosphatase 2
3.45	5.17333E-05	0.003988285	A6	gene31429	LOC106365436	gene31429	
3.44	4.14291E-06	0.000769144					
3.44	0.002346619	0.038446088	A4	rna27057	LOC106391939	gene21721	SPX domain-containing protein 3
3.43	1.99601E-06	0.000462616	A3	rna19183	LOC106444048	gene15339	histidine-containing phosphotransfer protein 4-like
3.42	2.6808E-07	0.000101614	C2	rna76164	LOC106345339	gene61299	histidine-containing phosphotransfer protein 4-like
3.40	1.82236E-06	0.000433412	C8	rna123085	LOC106415565	gene99585	uncharacterized protein C24B11.05-like
3.39	2.1536E-05	0.002226288					
3.39	4.65399E-06	0.000842534	A10	rna61372	LOC106387761	gene49230	anthranilate phosphoribosyltransferase, chloroplastic BnaA10g16850D

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
3.37	0.000528076	0.016188401	A7	rna40910	LOC106357970	gene32700	probable methyltransferase PMT24
3.37	6.51001E-06	0.001048169	C9	rna128416	LOC106418248	gene103958	uncharacterized protein LOC106418248
3.35	0.002703563	0.041456913	C4	rna92453	BNAC04G16640D	gene74709	uncharacterized protein At2g27730, mitochondrial
3.34	5.59162E-07	0.000184971	A5	rna27644	LOC106454541	gene22188	uncharacterized protein LOC106454541
3.33	1.59746E-06	0.000400886	C8	rna124562	LOC106416009	gene100746	probable inactive purple acid phosphatase 1
3.32	1.8858E-07	7.79778E-05	A5	rna30064	LOC106449216	gene24105	phosphoglycerate mutase-like protein 2
3.32	1.88354E-06	0.00044129	A3	gene16922	LOC106444396	gene16922	
3.31	9.22415E-10	1.15741E-06	A3	rna16999	LOC106438407	gene13632	SPX domain-containing protein 2
3.31	0.000478259	0.015414424	C4	rna89889	LOC106396187	gene72591	insulin-degrading enzyme-like 1, peroxisomal zinc-metallopeptidase, peroxisomal
3.29	0.00018948	0.008816875	C9	rna131024	LOC106372323	gene106124	glucose-6-phosphate 1-dehydrogenase 2, chloroplastic BnaC09g43980D
3.29	2.74394E-05	0.002641446	Un	rna146270	LOC111213444	gene118561	probable tyrosine-protein phosphatase At1g05000

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
3.28	1.153E-08	9.32343E-06	A1	rna1589	BNAA01G08680D	gene1275	uncharacterized membrane protein At4g09580
3.27	9.99569E-05	0.006051968	C3	rna85005	LOC106385848	gene68536	F-box/kelch-repeat protein At3g27150 BnaC02g36670D
3.27	3.10863E-06	0.000631938	C2	rna74720	LOC111198423	gene60156	phospholipid-transporting ATPase 2-like
3.27	5.73177E-05	0.004267426	C1	rna65204	LOC106375610	gene52252	uncharacterized protein LOC106375610
3.26	6.02231E-05	0.004366754	A2	rna11873	LOC106424736	gene9490	F-box/kelch-repeat protein At3g27150 BnaA02g28650D
3.26	0.000181137	0.008604706	A6	rna36293	LOC106371409	gene29081	uncharacterized protein LOC106371409
3.24	0.002370342	0.038574241					
3.24	0.001784693	0.033015075	A8	rna46007	LOC106371877	gene36864	uncharacterized protein LOC106371877
3.23	1.79646E-07	7.69055E-05	A3	rna13276	LOC106427391	gene10653	sulfoquinovosyl transferase SQD2-like
3.21	0.002503451	0.03993659	C2	rna71242	LOC106417365	gene57295	non-specific lipid-transfer protein 4-like
3.21	0.000307081	0.01174979	A4	rna24248	LOC106449323	gene19440	soluble inorganic pyrophosphatase 4 BnaA04g04610D

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
3.21	4.12005E-07	0.000144154	A6	rna33044	LOC106345760	gene26495	beta-fructofuranosidase, insoluble isoenzyme CWINV3 cell wall invertase 3-1 cell wall invertase 3a
3.20	0.001013975	0.023651312	C4	rna97290	LOC106406370	gene78647	inorganic phosphate transporter 1-4-like
3.20	0.000285176	0.011303909	A3	rna20615	LOC106399237	gene16468	putative glycine-rich cell wall structural protein 1
3.19	2.70855E-09	2.89878E-06	C7	rna118030	LOC106410744	gene95326	UDP-sulfoquinovose synthase, chloroplastic BnaC07g44070D
3.17	1.94079E-05	0.002108099	A4	rna26670	LOC106447690	gene21423	inorganic phosphate transporter 1-4
3.15	9.9285E-05	0.006038983	A10	rna59566	LOC106423037	gene47780	monothiol glutaredoxin-S1 BnaA10g01610D
3.14	3.31415E-05	0.003003367	A6	rna39249	LOC111198756	gene31379	probable inactive receptor kinase RLK902
3.14	0.000225775	0.009792025	A1	rna6083	LOC106351805	gene4840	serine/threonine-protein kinase Nek7 BnaA01g30790D
3.12	2.05751E-05	0.002182762	C5	rna104085	LOC106452723	gene84069	phosphoglucan phosphatase LSF2, chloroplastic
3.11	0.001667718	0.031689259	A6	rna37846	LOC106348430	gene30284	uncharacterized protein LOC106348430

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
3.11	0.003530646	0.04875641	C4	rna95119	LOC106376499	gene76862	uncharacterized protein LOC106376499 BnaC04g34370D
3.11	0.000204704	0.009236664	A1	rna2500	LOC106442390	gene2023	probable inactive purple acid phosphatase 24
3.10	1.04782E-07	4.95172E-05	A3	rna22471	BNAA03G49870D	gene18021	uncharacterized protein BNAA03G49870D
3.10	0.001179218	0.025913876	A5	rna32592	LOC106452881	gene26147	protein ACTIVITY OF BC1 COMPLEX KINASE 7, chloroplastic-like
3.07	7.66944E-06	0.001136654					
3.06	0.002278102	0.037814071	C5	rna104488	LOC106410327	gene84392	sterol 3-beta-glucosyltransferase UGT80A2 BnaC05g44970D
3.05	2.11697E-05	0.002219955	Un	rna139704	LOC106434790	gene113235	protein FAF-like, chloroplastic
3.05	0.000218366	0.009596519	A7	gene35436	LOC106354202	gene35436	
3.05	3.57196E-11	8.90979E-08	C6	gene88395	LOC106403351	gene88395	
3.05	4.67344E-05	0.003762328	C9	rna126276	LOC106423955	gene102177	polyadenylate-binding protein- interacting protein 10 BnaC01g26120D
3.04	6.64462E-05	0.004649697	C3	gene70897	LOC106396255	gene70897	
3.03	0.001013417	0.023651312	C7	rna117449	LOC106410439	gene94875	probable pectate lyase 22 BnaCnng15740D
3.03	0.000904554	0.022179852	C3	rna86093	LOC106348147	gene69436	beta-galactosidase 6-like

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
3.00	0.000448545	0.014851363	A5	rna27535	LOC106450462	gene22107	pectin acetyltransferase 3-like
3.00	0.000137051	0.00728029	C4	rna91363	LOC111205058	gene73802	uncharacterized protein LOC111205058
3.00	0.00021816	0.009596519	C1	rna63847	LOC106426867	gene51138	heptahelical transmembrane protein 4
2.98	6.41763E-07	0.000203065					
2.98	0.001531365	0.03018598	A1	rna1034	LOC106354963	gene815	cellulose synthase A catalytic subunit 1 [UDP-forming]-like
2.97	2.63404E-06	0.00056381					
2.97	0.000498683	0.01558942	A9	rna51085	LOC106366128	gene40965	pyrophosphate--fructose 6-phosphate 1-phosphotransferase subunit beta 2
2.97	0.000547917	0.016477344	C1	rna65331	LOC106365923	gene52354	protein PHLOEM PROTEIN 2-LIKE A1- like
2.96	2.10528E-05	0.002214068	C4	rna97584	LOC106397671	gene78861	uncharacterized protein LOC106397671
2.96	1.0992E-06	0.000304073	C6	rna106261	LOC106348467	gene85800	beta-fructofuranosidase, insoluble isoenzyme CWINV3 cell wall invertase 3-2
2.95	4.74788E-06	0.000853149	A2	rna10287	LOC106432951	gene8199	ethylene-responsive transcription factor ERF070-like
2.95	0.001625056	0.031204502	A10	rna63455	LOC106372591	gene50832	protein PAF1 homolog

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
2.94	0.00199799	0.035207196	A3	rna20218	LOC106361822	gene16148	indole-3-glycerol phosphate synthase, chloroplastic-like
2.94	0.000819415	0.020727059	C4	rna90040	LOC106433154	gene72713	jacalin-related lectin 3-like
2.92	8.65047E-05	0.005541786	C3	rna86209	LOC106435063	gene69529	protein BROTHER of FT and TFL 1 BnaC03g52010D
2.91	0.002381087	0.038628173	A2	gene7034	LOC106387550	gene7034	
2.88	1.4759E-05	0.001726846	A10	rna61751	LOC106370510	gene49526	auxin efflux carrier component 5 BnaA10g17940D putative auxin efflux carrier component 8
2.86	0.003241693	0.046149743	C2	rna70853	LOC106427340	gene56964	chaperone protein dnaJ 3-like
2.86	0.000222723	0.009699483	A3	rna20281	LOC106439786	gene16200	autophagy-related protein 8d
2.86	0.003085838	0.044841429	A1	rna2372	LOC106414112	gene1927	probable protein phosphatase 2C 63
2.85	0.001016465	0.023651312	C1	rna67549	LOC106350723	gene54128	uncharacterized protein LOC106350723
2.85	1.34277E-05	0.00163182	A10	rna62145	LOC106372390	gene49832	glucose-6-phosphate 1-dehydrogenase 2, chloroplastic-like
2.84	1.59212E-09	1.86884E-06	C6	rna109466	BNAC06G43650D	gene88360	uncharacterized protein BNAC06G43650D
2.83	0.000734442	0.019498133	A6	rna38510	LOC106410943	gene30793	uncharacterized protein LOC106410943

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
2.83	0.000259988	0.010665658	C3	rna84715	LOC106356383	gene68278	histidine-containing phosphotransfer protein 4-like
2.82	0.000963146	0.023087582	C2	rna72465	LOC106440163	gene58357	phosphate transporter PHO1 homolog 1-like
2.81	0.00178749	0.033033621	A2	rna8225	LOC106431427	gene6521	probable feruloyl esterase A BnaAnng08480D
2.81	0.003050268	0.044509591	C7	rna116399	LOC106410860	gene93980	probable mannan synthase 1
2.80	0.001837064	0.033591496	A7	rna44073	LOC106406022	gene35328	glutamate--glyoxylate aminotransferase 2
2.80	0.000356475	0.012868458	C3	rna87560	LOC106382592	gene70700	cytochrome P450 79B1-like
2.80	0.000760661	0.019813135	C8	rna122200	BNAC07G49750D	gene98831	uncharacterized protein BNAC07G49750D
2.79	2.69879E-05	0.002632799	A10	rna61595	LOC106428730	gene49399	UDP-glycosyltransferase 90A1
2.79	0.000784237	0.020181631	A5	rna28743	LOC106451091	gene23022	LOW QUALITY PROTEIN: callose synthase 10-like
2.78	0.000663475	0.018468365	C2	rna72390	LOC106381160	gene58292	WAT1-related protein At1g70260 BnaC02g19720D
2.77	1.43089E-05	0.001690494	A8	rna48245	LOC106359053	gene38700	PLAT domain-containing protein 1-like
2.77	0.001753266	0.032678809	C1	rna63906	LOC106423366	gene51192	uncharacterized protein LOC106423366

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
2.76	0.001105775	0.024914511					
2.76	6.05532E-06	0.001006124					
2.74	2.29216E-06	0.000508581	A4	rna25592	LOC106446783	gene20518	fe(3+)-Zn(2+) purple acid phosphatase 12-like
2.73	1.0821E-05	0.001431829	A6	rna39984	LOC106349478	gene31951	peptidyl-prolyl cis-trans isomerase CYP18-3-like
2.72	1.52531E-06	0.000388133	A1	rna1847	LOC106358825	gene1493	protein NUCLEAR FUSION DEFECTIVE 4-like
2.71	2.6097E-05	0.002566537	C8	rna125160	LOC106420230	gene101230	monothiol glutaredoxin-S11 BnaA09g49730D BnaC08g44980D
2.71	0.00134665	0.028081327	C3	rna78933	LOC106358518	gene63560	pollen-specific protein SF21-like
2.71	1.76161E-06	0.000424514	C3	rna87494	LOC106390428	gene70636	PLAT domain-containing protein 1 dehydration stress-induced protein
2.69	1.70061E-08	1.21337E-05	C5	rna103284	LOC106425080	gene83426	E3 ubiquitin ligase BIG BROTHER-related BnaC05g32790D
2.68	0.001610981	0.031022722					
2.68	0.00050375	0.015694254	A1	rna4998	LOC106346384	gene4026	eukaryotic initiation factor 4A-III homolog

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
2.67	1.92979E-07	7.80235E-05	C1	rna63876	LOC106426868	gene51166	cytochrome P450 81F3
2.67	0.000122646	0.006855391	C3	rna86897	BNACNNG04910D	gene70141	uncharacterized protein BNACNNG04910D
2.67	0.00200586	0.035251938	A10	rna62588	LOC106419401	gene50172	uncharacterized aarF domain- containing protein kinase At5g05200, chloroplastic-like
2.66	3.6021E-06	0.000708503					
2.66	0.002829132	0.042734097	A8	rna49652	LOC106362162	gene39851	starch synthase 3, chloroplastic/amyloplastic-like
2.65	0.00074222	0.019592125	A10	rna61106	LOC106370443	gene49016	WAT1-related protein At5g64700 BnaA09g06980D
2.65	0.000831829	0.020932637	A9	rna51279	LOC106366343	gene41133	dehydrin Rab18-like
2.64	2.52501E-06	0.000550127	C3	rna82799	LOC106388386	gene66715	SPX domain-containing protein 2-like
2.64	6.46234E-05	0.004583856	A10	rna63167	LOC106357500	gene50616	uncharacterized protein LOC106357500
2.62	0.003512439	0.048578721	C4	rna97351	LOC106447703	gene78692	SAC3 family protein A-like
2.62	0.000165167	0.008156086	C9	rna128603	LOC106383122	gene104110	phosphoenolpyruvate carboxylase kinase 1-like
2.61	2.32012E-06	0.000511665	Un	rna146476	LOC111213560	gene118723	probable LRR receptor-like serine/threonine-protein kinase At2g23950

Appendix 3 continued

	Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
	2.60	5.78149E-05	0.004268384	A3	rna19788	LOC106389302	gene15810	transmembrane protein 53 BnaA03g35520D
	2.60	0.002332585	0.038285114	C1	rna66590	LOC106445550	gene53376	uncharacterized protein LOC106445550
	2.59	0.001784232	0.033015075	A5	rna29207	LOC106406290	gene23418	U-box domain-containing protein 9-like
	2.58	7.3376E-06	0.001121822					
	2.58	1.20333E-07	5.61367E-05	C5	rna98814	LOC111206575	gene79844	fe(3+)-Zn(2+) purple acid phosphatase 12
268	2.57	1.99485E-05	0.00214126	A4	rna26733	LOC106447459	gene21474	mannose-1-phosphate guanylyltransferase 1 BnaA04g22820D
	2.57	0.002715718	0.041590714	C8	rna119657	LOC106411649	gene96711	probable zinc transporter 10
	2.57	0.000689572	0.018809707	A6	rna36886	LOC106347667	gene29526	serine/threonine-protein kinase D6PK
	2.57	1.52175E-05	0.001769122	A5	rna28460	LOC106396536	gene22788	mannose-1-phosphate guanylyltransferase 1-like
	2.57	0.001879514	0.033989759	A3	rna18830	LOC106353244	gene15078	protein STRICTOSIDINE SYNTHASE-LIKE 9-like
	2.57	0.000458114	0.015045001	A6	rna38283	LOC111198927	gene30607	uncharacterized protein LOC111198927
	2.56	0.000168272	0.008229957	A5	rna29543	LOC106454885	gene23694	monothiol glutaredoxin-S9 BnaA05g11910D

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
2.55	0.002599459	0.040788746	A8	rna49893	LOC106362340	gene40036	protein SRC2 homolog
2.55	0.000350708	0.012729448	A3	rna20103	LOC106389664	gene16048	adenosylhomocysteinase 2-like
2.54	0.000325	0.012203491	A9	rna51177	LOC106366184	gene41044	uncharacterized protein LOC106366184
2.54	0.002576121	0.040597616	A7	rna42304	LOC106352861	gene33874	protein NCA1-like
2.53	0.001131365	0.025196178	A1	rna6553	LOC106353118	gene5222	glycine-rich cell wall structural protein
2.52	0.001555838	0.030486709	C9	rna130673	LOC106424648	gene105855	lamin-like protein
2.52	0.000188718	0.008792679	C9	rna126663	LOC106392257	gene102505	dehydrin Rab18-like
2.52	2.87572E-05	0.00271092	C5	rna101642	LOC106401581	gene82088	protein C2-DOMAIN ABA-RELATED 8- like
2.52	3.66444E-05	0.003188674	C2	rna72317	BNAC02G17680D	gene58218	uncharacterized protein BNAC02G17680D
2.52	5.0237E-06	0.000878806	C3	rna85294	LOC106439617	gene68775	LOW QUALITY PROTEIN: ATP sulfurylase 1, chloroplastic
2.52	0.000272698	0.011001042	A8	rna47435	LOC106361008	gene37996	protein NUCLEAR FUSION DEFECTIVE 4-like
2.52	5.38097E-05	0.004079226	A5	rna28423	LOC111215549	gene22763	E3 ubiquitin ligase BIG BROTHER- related-like

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
2.51	0.002253549	0.037653241	A8	rna46875	LOC106401466	gene37537	SH3 domain-containing protein 1-like
2.50	0.003034986	0.044352243	A9	rna57815	LOC106441106	gene46351	translation initiation factor IF3-4, chloroplastic BnaA03g50490D
2.49	0.000465744	0.015236486	A1	rna148	LOC106365006	gene112	probable trehalose-phosphate phosphatase H
2.49	9.65545E-05	0.005905694	C6	rna110422	LOC111207235	gene89162	U-box domain-containing protein 9-like
2.49	0.003082779	0.044816691	C5	rna103384	LOC106397121	gene83496	peroxidase 29-like
2.49	0.000569191	0.016838792	A3	rna20448	LOC106444247	gene16348	ubiquitin-conjugating enzyme E2 29 BnaA03g39290D
2.49	9.92165E-08	4.81372E-05	C7	rna117372	BNAC07G50690D	gene94805	uncharacterized protein BNAC07G50690D
2.48	0.000872823	0.021605632	C7	rna116456	LOC111198139	gene94031	calcineurin B-like protein 1
2.48	0.001393262	0.028660938	A2	rna13141	LOC106405874	gene10562	LOB domain-containing protein 37-like
2.47	0.002216245	0.037335524					
2.47	0.000303593	0.011721877	A4	rna26557	LOC106354919	gene21327	uncharacterized protein LOC106354919
2.47	1.44509E-07	6.57299E-05	A9	rna51873	LOC106366544	gene41617	protein YLS3 BnaA09g12710D
2.47	9.27512E-06	0.001293115	C3	rna81173	LOC106387853	gene65452	

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
2.46	0.003594724	0.049323085	C7	rna114464	LOC106411025	gene92409	cytochrome P450 71B26-like
2.44	0.000192588	0.008870728	C4	rna89960	LOC106391948	gene72653	phosphoenolpyruvate carboxylase 2 BnaC04g02430D
2.44	7.9671E-06	0.001165154	A2	rna12774	LOC106405403	gene10236	uncharacterized protein LOC106405403
2.43	0.001464702	0.029462455	A4	rna25608	LOC106446774	gene20532	calmodulin-7-like
2.43	0.001263185	0.027133868	A9	gene43017	LOC106447088	gene43017	
2.42	0.000118922	0.006698642	C2	rna76839	LOC106399579	gene61853	
2.42	5.65671E-05	0.004226621	A1	rna495	BNAA01G02080D	gene411	uncharacterized protein BNAA01G02080D
2.41	3.87857E-05	0.003244442	A4	rna23519	LOC106445032	gene18848	probable glycosyltransferase At5g03795
2.40	6.37512E-06	0.001034394	A6	rna38915	LOC106349269	gene31112	uncharacterized protein LOC106349269
2.39	0.000211734	0.009453463					
2.39	0.001086055	0.024741418	C5	rna101551	LOC106401656	gene82009	ethylene-responsive transcription factor ERF019 BnaC05g18050D
2.39	0.000858861	0.021403179	A7	rna40264	LOC106356169	gene32167	protein RADIALIS-like 4
2.37	0.000954555	0.022977801	C2	rna76375	LOC106347720	gene61481	annexin D1-like annexin 1

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
2.36	0.001482162	0.029648992	A3	rna18791	LOC106353228	gene15046	1-acyl-sn-glycerol-3-phosphate acyltransferase 2 Lysophosphatidyl acyltransferase 2 acyl-CoA:1-acylglycerol-3-phosphate acyltransferase lysophosphatidic acid acyltransferase 2
2.36	0.00332653	0.046935163	C8	rna123257	LOC106415278	gene99724	peptide chain release factor APG3, chloroplastic
2.36	0.000251507	0.010483206	A7	rna45043	LOC106354681	gene36097	probable gamma-glutamyl hydrolase 3
2.34	0.000259453	0.010665658	C4	rna95834	LOC106445438	gene77420	non-specific phospholipase C2 BnaA09g23040D BnaC04g38510D
2.34	2.82698E-05	0.002678861	A2	rna6938	LOC106352709	gene5486	glycerophosphodiester phosphodiesterase GDPD2-like
2.34	3.7173E-05	0.003200923	C8	rna123916	LOC106413506	gene100262	tyrosine aminotransferase-like
2.34	0.00067178	0.018589139	A3	rna22157	LOC106431127	gene17734	telomere repeat-binding protein 4
2.33	0.000136537	0.007277513	A1	rna462	LOC106346599	gene381	ganglioside-induced differentiation-associated protein 2 BnaA01g01800D
2.33	0.001395373	0.02867014	C7	rna116020	LOC106428039	gene93669	transcription factor MYB28-like high aliphatic glucosinolate 1
2.33	0.000141928	0.00739541					

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
2.32	5.79472E-05	0.004268384	C6	rna108042	LOC106391303	gene87214	EP1-like glycoprotein 1
2.32	0.00147316	0.029550913	C3	rna78034	LOC106433803	gene62801	uncharacterized protein LOC106433803
2.32	0.001736493	0.032537343	A1	rna681	LOC106357529	gene556	V-type proton ATPase subunit H AT3g42050-like protein
2.31	0.000174434	0.008409431	A6	rna35597	LOC111198594	gene28508	
2.31	0.001627381	0.031205416	A4	rna24521	LOC106445804	gene19652	high-affinity nitrate transporter 3.1 BnaA03g57010D
2.31	0.000771464	0.020008565	A6	rna36088	LOC106404797	gene28902	inositol-tetrakisphosphate 1-kinase 2- like
2.30	0.000906657	0.022210886	A8	rna48723	LOC106360120	gene39087	uncharacterized protein LOC106360120
2.30	5.11876E-05	0.003966801	C4	rna95497	LOC106446621	gene77151	probable LRR receptor-like serine/threonine-protein kinase At2g23950 BnaC04g56200D
2.30	0.000146466	0.007549037	A10	rna58999	LOC106368996	gene47312	nardilysin-like
2.29	2.46348E-05	0.002496971	C3	rna83762	BNAC03G32080D	gene67517	uncharacterized protein BNAC03G32080D
2.29	1.50766E-05	0.001758359	A5	rna27848	LOC106450613	gene22347	
2.29	0.003514024	0.048582184					

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
2.29	0.001841682	0.033592447	A3	rna22247	LOC106440963	gene17826	early nodulin-like protein 2
2.28	7.13439E-05	0.004924411	C4	rna91403	LOC106396176	gene73836	berberine bridge enzyme-like 16 BnaCnng22600D reticuline oxidase- like protein
2.28	2.32222E-05	0.002373622	C8	rna120578	LOC106362089	gene97527	nodulin-related protein 1-like
2.28	8.56301E-06	0.001217152	A3	rna18478	LOC106353048	gene14784	protein yippee-like At3g11230 BnaA03g31640D
2.28	1.82536E-07	7.72339E-05	A3	rna18059	LOC106438958	gene14436	late embryogenesis abundant protein At5g17165-like
2.28	0.001017505	0.023658133	C4	rna96939	LOC106447289	gene78361	WAT1-related protein At4g01440-like
2.27	0.001550559	0.030416035	C5	rna104293	LOC111206511	gene84247	phosphoinositide phospholipase C 2
2.27	4.94825E-08	2.85805E-05	C5	rna101120	LOC106447605	gene81673	late embryogenesis abundant protein At5g17165-like
2.26	5.92884E-05	0.00432342	C3	rna85292	LOC106444154	gene68773	stem-specific protein TSJT1-like
2.26	0.003447222	0.048060349	C3	rna80275	LOC106387449	gene64690	ATP-dependent 6-phosphofructokinase 7-like
2.26	0.000249375	0.010449677	A2	rna7992	LOC106355987	gene6332	uncharacterized protein LOC106355987

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
2.25	0.001032645	0.023933689	A9	rna57796	LOC106441096	gene46338	signal recognition particle receptor subunit alpha BnaA03g50330D
2.25	7.75044E-06	0.001141793					
2.25	0.002121028	0.036425574	A6	rna37993	LOC106348561	gene30403	
2.24	0.001327704	0.027861869	A10	rna59748	LOC106370865	gene47925	glycerol-3-phosphate 2-O-acyltransferase 4-like BnaA10g00370D sn-glycerol-3-phosphate acyltransferase 4
2.24	0.000232326	0.009969217	C1	rna64417	LOC106368756	gene51593	alpha-aminoadipic semialdehyde synthase lysine-ketoglutarate reductase/saccharopine dehydrogenase
2.24	5.65105E-07	0.000185253	A3	rna16679	LOC106438197	gene13379	UDP-glycosyltransferase 74F2-like
2.23	0.003191851	0.045726405	Un	rna138044	LOC106433855	gene111865	glucomannan 4-beta-mannosyltransferase 9-like
2.22	0.001067951	0.024440629	C1	rna63747	LOC106439198	gene51055	probable trehalose-phosphate phosphatase H
2.22	0.001127113	0.025189903	A4	gene18956	LOC106445095	gene18956	
2.22	0.000270302	0.01091648					
2.22	4.04775E-06	0.000759225	A8	gene39128	LOC106361888	gene39128	

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
2.22	0.002364751	0.038528865	C1	rna64081	LOC106351245	gene51336	ganglioside-induced differentiation-associated protein 2-like
2.22	0.001853549	0.033708396	A2	rna7498	LOC106396524	gene5931	alpha-galactosidase 2
2.21	0.001114153	0.025010367	C6	rna109028	LOC106353978	gene88018	sufE-like protein 2, chloroplastic
2.21	0.001300647	0.02754378	A1	rna5914	LOC106375091	gene4706	uncharacterized protein LOC106375091
2.21	0.000187285	0.008784284	C9	rna130823	BNAC09G43030D	gene105962	uncharacterized protein BNAC09G43030D
2.21	0.00088222	0.021734977	C4	rna89607	LOC106396363	gene72370	short-chain dehydrogenase reductase 3b
2.21	0.0002821	0.011243199	C6	rna110062	LOC106406400	gene88872	isoflavone reductase homolog P3 BnaC06g36040D
2.21	0.001401073	0.028715151	A8	rna46504	LOC106360686	gene37257	probable mediator of RNA polymerase II transcription subunit 37e
2.20	0.000365487	0.01302581	A9	rna57578	LOC106369067	gene46172	myb-related protein 330-like transcription factor MYB106-like transcription repressor MYB5-like
2.20	6.49055E-05	0.004585982	C1	rna64127	BNAC01G03230D	gene51380	uncharacterized protein BNAC01G03230D
2.20	7.97305E-06	0.001165154	A3	rna16140	LOC106437868	gene12967	fasciclin-like arabinogalactan protein 16

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
2.20	0.001503971	0.029905191	A3	rna21167	LOC106440206	gene16939	aluminum-activated malate transporter 12-like
2.20	5.01361E-05	0.00395955	C5	rna104578	LOC106397015	gene84469	hydrophobic protein RCI2A BnaC05g45940D
2.19	0.000979221	0.023319303	A1	rna5341	LOC106451562	gene4242	
2.19	0.00042254	0.014329336	A3	rna14324	LOC106433933	gene11513	monothiol glutaredoxin-S2-like
2.19	0.002073307	0.036011219	A3	rna17798	LOC106438604	gene14250	shaggy-related protein kinase theta-like
2.18	3.04333E-08	1.94282E-05	A1	rna1250	LOC106372301	gene998	RING-H2 finger protein ATL14 BnaA01g06660D
2.18	0.001035847	0.023977348	C9	rna125903	LOC106444461	gene101893	ethylene-responsive transcription factor TINY BnaC09g50850D
2.18	0.002261603	0.037715492	C9	rna126498	BNAC09G07050D	gene102360	uncharacterized protein At5g65660
2.18	0.00029048	0.011457333	C9	rna130538	LOC106370582	gene105738	auxin efflux carrier component 5 BnaCnng54140D putative auxin efflux carrier component 8
2.18	3.77746E-06	0.000738694	A9	rna56497	BNAA09G37670D	gene45334	uncharacterized protein At3g49140
2.17	0.001797939	0.033176168	C5	rna104508	LOC106397166	gene84411	zinc finger CCCH domain-containing protein 34
2.17	1.76535E-05	0.001964455	A9	rna54398	LOC106360868	gene43669	probable methyltransferase PMT14

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
2.16	2.48742E-08	1.61629E-05					
2.16	5.06856E-05	0.003966338	C5	rna98499	LOC106397629	gene79606	cyclin-A1-1-like
2.16	0.000332481	0.012332628	A5	rna31627	LOC106452180	gene25398	O-glucosyltransferase rumi homolog BnaA04g26470D
2.16	0.000223753	0.009727529	C6	rna110127	LOC106406877	gene88920	gibberellin-regulated protein 1 BnaAnng08970D
2.16	0.002326899	0.038226272					
2.16	5.21588E-08	2.91993E-05	C9	rna132905	LOC106445180	gene107637	mitogen-activated protein kinase 5-like BnaC09g24030D mitogen-activated protein kinase 5.1
2.15	0.002671076	0.041201831					
2.15	0.001928352	0.03453581	C4	rna96382	LOC106405961	gene77863	serine/threonine-protein kinase EDR1 BnaC04g41700D
2.15	6.5222E-05	0.004590515	A3	rna22560	LOC106441054	gene18080	UDP-glucuronate 4-epimerase 1-like
2.15	9.48027E-06	0.001304801	C7	rna118537	LOC106411102	gene95762	phosphoinositide phospholipase C 3-like
2.15	9.4815E-05	0.005867565	A6	rna35762	LOC106347355	gene28645	uncharacterized protein LOC106347355
2.15	0.002379678	0.038622527	C8	rna122115	LOC106415556	gene98751	protein NRT1/ PTR FAMILY 5.7-like

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
2.15	0.000109732	0.006402624	A3	rna20038	LOC111214198	gene15995	ATP sulfurylase 1, chloroplastic
2.14	0.001448169	0.029308099	A6	rna35621	LOC106399273	gene28529	
2.14	0.002580656	0.040626516	C4	rna91374	LOC111205065	gene73814	glycine cleavage system H protein 3, mitochondrial
2.14	0.000631088	0.017940645	C3	rna88821	LOC106345228	gene71713	cyclin-dependent kinase inhibitor 6 BnaA05g21020D
2.14	0.001096094	0.024805235	Un	rna132482	LOC106406440	gene107276	
2.14	0.000405354	0.01405186	A3	rna22129	LOC106440909	gene17714	tryptophan synthase beta chain 2, chloroplastic
2.14	0.002554244	0.040399532	C3	rna79465	LOC111203909	gene64000	uncharacterized protein LOC111203909
2.14	0.003316183	0.046807317	A7	rna43205	LOC106353800	gene34639	protein TIFY 7 BnaA07g23750D
2.13	0.000642256	0.018160056	C3	rna86183	LOC106424857	gene69507	protein NRT1/ PTR FAMILY 4.7-like
2.13	7.97584E-08	4.0309E-05	A1	rna1243	LOC106392805	gene992	UDP-glucuronate 4-epimerase 1
2.13	5.08669E-05	0.003966801	A2	rna10605	LOC106422271	gene8474	bZIP transcription factor 44-like
2.13	4.69824E-05	0.003765476					
2.13	1.22184E-05	0.001519371	A6	rna33130	LOC106345974	gene26567	phosphoenolpyruvate carboxylase kinase 2

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
2.13	0.001092102	0.024768506	Un	rna145542	LOC106365322	gene117972	ethylene-responsive transcription factor ERF070 BnaCnng17120D
2.13	3.82061E-05	0.0032328	A2	rna10275	LOC106395044	gene8187	MLP-like protein 31 BnaA02g15250D
2.12	0.000714328	0.019126542	A10	rna63089	LOC106419507	gene50551	histone acetyltransferase of the MYST family 2
2.12	0.002096627	0.03619168	A2	rna7967	LOC106364603	gene6315	splicing factor 3A subunit 2 BnaA02g02790D
2.12	0.000475586	0.015378693	C3	rna78799	LOC106384921	gene63444	cellulose synthase A catalytic subunit 5 [UDP-forming] BnaA03g55860D
2.12	0.000861457	0.021424372	C5	rna104392	LOC106401284	gene84329	U-box domain-containing protein 9 BnaC05g44610D
2.11	0.000698054	0.018899402	Un	rna132339	LOC106425067	gene107173	serine carboxypeptidase-like 48 BnaC03g55150D
2.11	0.000220809	0.009668843	C8	rna122722	LOC106368312	gene99297	probable pectinesterase/pectinesterase inhibitor 35 BnaCnng47940D
2.11	8.83707E-05	0.005594845	C3	rna80802	LOC106387935	gene65147	hypersensitive-induced response protein 4 BnaC03g15940D
2.11	0.000480773	0.015420282	C2	rna74039	LOC106410292	gene59605	UDP-glucuronate 4-epimerase 6 BnaC07g07730D
2.11	0.000584008	0.017102069	A3	rna13890	LOC106429877	gene11158	uncharacterized protein LOC106429877

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
2.11	0.00234488	0.038434907	A3	rna16124	LOC106437858	gene12953	putative dual specificity protein phosphatase DSP8
2.11	0.000117565	0.006642783	A6	rna34832	LOC106346782	gene27923	glyceraldehyde-3-phosphate dehydrogenase GAPCP2, chloroplastic BnaA06g10900D
2.10	3.33451E-05	0.003003367	C1	rna66141	LOC106396827	gene53018	glucose-6-phosphate isomerase 1, chloroplastic-like
2.10	0.000105648	0.006240757	C7	rna112181	LOC106374817	gene90565	uncharacterized protein LOC106374817
2.10	0.000173817	0.008409431	C1	rna65488	LOC111202187	gene52486	bifunctional riboflavin kinase/FMN phosphatase-like
2.10	0.001689257	0.031964996	A4	rna23456	LOC106449059	gene18795	U-box domain-containing protein 33-like
2.10	7.45202E-05	0.005040226	A2	rna13076	LOC106426073	gene10494	cyclin-dependent kinase inhibitor 2-like
2.10	0.000146896	0.007549806	C1	rna64347	LOC106438554	gene51541	inositol-tetrakisphosphate 1-kinase 2
2.09	9.42624E-05	0.005843308	C6	rna106979	BNAC06G14030D	gene86372	heavy metal-associated isoprenylated plant protein 30
2.09	0.001940893	0.034671198	C2	rna76714	LOC106429018	gene61748	S-adenosylmethionine synthase 2-like
2.09	0.00321529	0.04589956	A5	rna28408	LOC106446159	gene22750	jasmonic acid-amido synthetase JAR1
2.09	0.00015949	0.007983398	A3	gene12682	LOC106437570	gene12682	
2.09	0.000200404	0.009138206					

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
2.08	0.00305187	0.044509591	A2	rna12976	LOC106404924	gene10401	uncharacterized protein LOC106404924
2.07	7.51976E-06	0.001127123	A2	rna12867	LOC106435933	gene10311	late embryogenesis abundant protein At5g17165 seed specific protein Bn15D1B
2.07	1.03164E-06	0.000293276	A3	rna16181	LOC106392756	gene13007	adenylate kinase 1, chloroplastic BnaA03g17110D
2.07	0.001944427	0.034683246	Un	rna144334	LOC106419148	gene116983	amino acid permease 2 amino acid permease 2-2
2.06	0.001382564	0.028568279	A6	rna35210	LOC106346914	gene28203	cytosolic sulfotransferase 17 desulfo- glucosinolate sulfotransferase
2.06	0.002052813	0.03582626	A6	rna37796	LOC106348265	gene30252	uncharacterized WD repeat-containing protein C3H5.08c-like
2.06	0.000164595	0.008156086	A3	rna15548	LOC111213406	gene12518	BAG family molecular chaperone regulator 1-like
2.06	0.000226073	0.009793281					
2.06	0.001354015	0.028170335	C2	rna71764	LOC106380934	gene57750	cyclic nucleotide-gated ion channel 4- like
2.06	0.000109523	0.006402624	Un	rna135358	LOC106366907	gene109652	probable plastid-lipid-associated protein 1, chloroplastic
2.06	0.000564843	0.016778363	C2	rna77688	LOC106431202	gene62518	LOB domain-containing protein 37-like

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
2.06	0.000290202	0.011457333	A9	rna56967	LOC111200370	gene45665	
2.06	0.003201085	0.045765387	C3	rna82118	LOC106420561	gene66165	proline transporter 3 BnaC03g22280D
2.06	0.002691675	0.041394151	A3	rna20496	BNAA03G39580D	gene16375	uncharacterized protein BNAA03G39580D
2.05	9.9898E-05	0.006051968	C3	rna79683	LOC106386761	gene64187	uncharacterized protein LOC106386761
2.05	0.001508323	0.029958985	C8	rna118913	LOC111208462	gene96063	
2.05	0.002085441	0.036118528	C4	rna96700	LOC106427690	gene78149	nitrile-specifier protein 2 BnaCnng21640D
2.05	0.000688554	0.018796031	C6	rna109032	BNAA07G26700D	gene88022	uncharacterized protein BNAA07G26700D
2.05	1.09429E-08	9.14375E-06	A3	rna19198	LOC106439444	gene15354	probable polygalacturonase
2.05	0.000890736	0.021914884	A7	rna40265	LOC106356170	gene32168	protein RADIALIS-like 4
2.04	0.000560206	0.016695145	C8	gene100651	LOC106382506	gene100651	
2.04	0.001004898	0.023621602	A9	rna51759	LOC106436352	gene41525	threonine synthase 1, chloroplastic-like
2.04	8.32299E-06	0.001194875	C6	rna107015	LOC106452438	gene86401	pyruvate kinase 1, cytosolic BnaC06g14360D
2.04	0.001401171	0.028715151	A7	rna41257	LOC106433401	gene33006	pyruvate dehydrogenase E1 component subunit alpha-2, mitochondrial BnaA07g09080D

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
2.04	4.25767E-05	0.003505161	A4	rna26491	LOC106447390	gene21277	adenylate kinase 1, chloroplastic-like
2.04	0.002185753	0.037010328	A5	rna27688	LOC106450687	gene22226	probable E3 ubiquitin-protein ligase RHC1A
2.03	0.001615479	0.031053382	C3	rna85569	LOC106442411	gene69009	uncharacterized protein LOC106442411
2.03	0.002909338	0.043278605	C7	rna117304	LOC106409715	gene94742	early nodulin-like protein 2
2.03	0.000474977	0.015378693	A2	rna11543	LOC106401465	gene9230	protein NSP-INTERACTING KINASE 2-like
2.03	0.000172675	0.008400109	C9	rna127670	LOC106418533	gene103326	fasciclin-like arabinogalactan protein 13
2.03	0.000329433	0.012269601	Un	rna145224	LOC111212966	gene117701	hydrophobic protein RCI2A-like
2.02	0.000110675	0.006402624	C2	rna76491	LOC106375471	gene61553	probable carboxylesterase 15 BnaC09g48990D
2.02	2.48939E-05	0.002505842	C7	rna117941	LOC106407607	gene95251	uncharacterized protein LOC106407607
2.02	0.000662758	0.018468365	A5	rna31116	LOC106428891	gene24965	seco-amyrin synthase-like
2.02	0.000222363	0.009699483	C2	rna74922	LOC106391343	gene60315	casein kinase 1-like protein HD16
2.02	0.000580579	0.017050932	C3	rna80780	LOC106387919	gene65126	RNA pseudouridine synthase 7 BnaC03g16060D
2.02	2.02256E-05	0.002162069	A3	rna19614	LOC106439112	gene15686	nuclear transcription factor Y sub-unit A-2 transcription factor subunit NF-YA2

Appendix 3 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_name	Reference Gene_ID	Full Description
2.01	1.62312E-08	1.20535E-05	A4	rna24135	LOC106445480	gene19348	probable arabinose 5-phosphate isomerase BnaA04g04030D
2.01	0.003303544	0.046684239	A6	rna38524	LOC111198707	gene30804	uncharacterized protein LOC111198707
2.01	0.000880278	0.021731038	A3	rna15754	LOC106437571	gene12681	glutathione S-transferase F9
2.01	0.003271118	0.046423334	A7	rna44673	LOC106421692	gene35809	gibberellin-regulated protein 1-like
2.01	0.000468006	0.015261608	C4	rna95588	LOC106446652	gene77221	probable sarcosine oxidase
2.00	0.000653385	0.018302814	C3	rna84929	LOC111204547	gene68463	sodium/proton antiporter 1-like
2.00	0.00106014	0.02435377	A7	rna41712	LOC106358315	gene33381	uncharacterized protein LOC106358315
2.00	0.00074863	0.019663604	Un	rna144174	LOC111212542	gene116855	granule-bound starch synthase 1, chloroplastic/amyloplastic

APPENDIX 4. List of down-regulated genes of *B. napus* in response to Pi availability. Genes with q value < 0.05 and fold change value < 0.5

Fold Change	pval	qval	Chromosome	Transcript	Gene_Name	Reference Gene_ID	Full Description
0.14	5.60661E-05	0.004201744	C4	rna91542	LOC106368861	gene73951	copper chaperone for superoxide dismutase, chloroplastic/cytosolic-like
0.16	0.000484759	0.015452265	C6	rna110029	LOC106406764	gene88848	pathogenesis-related protein 5-like uncharacterized protein
0.16	0.001638329	0.031360074	C1	rna64093	BNAC01G03020D	gene51347	BNAC01G03020D
0.18	0.003663029	0.049968592	A6	rna37433	LOC106348235	gene29975	AAA-ATPase At2g18193-like
0.18	0.003196258	0.045765387	C3	rna85413	LOC106391953	gene68881	protein DETOXIFICATION 1-like
0.189	0.000803355	0.020473259	A3	rna17322	LOC111214089	gene13881	40S ribosomal protein S18
0.19	5.24674E-05	0.004027811	A1	rna341	LOC106429842	gene288	probable vacuolar amino acid transporter YPQ3
0.19	7.99058E-05	0.005276971	C7	rna117122	LOC106411335	gene94581	superoxide dismutase [Fe] 1, chloroplastic-like
0.19	0.002608675	0.040880475	C7	rna111674	LOC106406310	gene90174	protein DETOXIFICATION 1 MATE efflux family protein DTX1 uncharacterized protein
0.22	0.000724842	0.019308597	C4	rna89885	BNAC04G02100D	gene72587	BNAC04G02100D
0.23	1.38394E-08	1.09476E-05	C7	rna116129	LOC106425621	gene93763	probable calcium-binding protein CML41
0.23	0.003464953	0.048178334	C3	gene66435	LOC106384011	gene66435	probable calcium-binding protein CML41
0.24	1.10559E-06	0.000304073	A3	rna20709	LOC106440010	gene16555	probable calcium-binding protein CML41
0.24	0.000462834	0.015172603	A1	rna5370	LOC106444959	gene4267	ADP-ribosylation factor 1-like nitrate reductase [NADH], clone PBNBR1412
0.24	0.002135953	0.036541169	Un	rna142925	LOC106368948	gene115817	

Appendix 4 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_Name	Reference Gene_ID	Full Description
0.25	0.000436953	0.014669588	A9	rna51864	LOC111197895	gene41609	copper chaperone for superoxide dismutase, chloroplastic/cytosolic-like
0.25	0.001403439	0.028738518	C1	rna65997	LOC106375998	gene52908	probable inactive ATP-dependent zinc metalloprotease FTSHI 1, chloroplastic
0.25	0.00348403	0.04833278	A3	rna15385	LOC106437108	gene12384	
0.27	6.47726E-05	0.004585499	Un	rna137262	LOC106347285	gene111200	prolyl endopeptidase-like
0.27	2.85052E-07	0.000105842	A9	rna52913	LOC106430532	gene42448	peroxidase C2 BnaA09g53760D
0.27	0.002811214	0.042541183	A1	rna6466	LOC106427885	gene5158	glutathione S-transferase U8 BnaC01g39170D
0.28	1.10695E-05	0.001453927	C1	rna65541	LOC106346563	gene52533	superoxide dismutase [Fe] 1, chloroplastic-like
0.28	0.000381977	0.013457553	C3	rna88124	LOC106432169	gene71152	vesicle-associated protein 4-2-like uncharacterized protein
0.30	0.002863151	0.043015837	A1	rna3838	LOC106404667	gene3103	LOC106404667
0.31	0.000694627	0.018862747	A1	rna4939	LOC106346131	gene3975	probable inactive serine/threonine-protein kinase fnkC
0.31	9.6298E-11	1.75205E-07	C1	rna68706	LOC106374867	gene55114	monothiol glutaredoxin-S10-like
0.32	0.000754196	0.019743662	Un	rna142348	LOC106352980	gene115354	extensin-3-like
0.33	5.82663E-05	0.00428322	C9	rna126886	LOC106404682	gene102694	transcription factor bHLH146-like
0.33	0.003332969	0.046989575	C8	gene100040	LOC106415396	gene100040	
0.33	0.002841876	0.042855439	A6	rna37242	LOC106348023	gene29817	probable WRKY transcription factor 51 BnaA06g23750D
0.34	0.002791014	0.042320626	Un	rna142794	LOC111212067	gene115721	protein DJ-1 homolog A-like
0.34	0.000485285	0.015452265	C5	rna99342	LOC106415531	gene80254	sugar transporter ERD6-like 2 BnaC05g06460D

Appendix 4 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_Name	Reference Gene_ID	Full Description
0.35	0.00020134	0.009161729	C9	rna131515	LOC106446553	gene106496	HD domain-containing protein 2 BnaA04g13930D
0.35	0.003633507	0.049705287	C9	rna130940	LOC106372275	gene106060	tropinone reductase-like
0.35	8.09495E-06	0.001178236	A5	rna27424	LOC106450418	gene22020	universal stress protein PHOS34 BnaA05g00280D
0.35	0.000686404	0.018779208	C8	rna123570	LOC106365769	gene99982	transcription termination factor MTERF4, chloroplastic-like
0.36	0.003072194	0.044716393	A5	rna32482	LOC111215793	gene26065	60S acidic ribosomal protein P0-2-like
0.36	0.000999639	0.023574121	C8	rna123569	LOC106365770	gene99981	FK506-binding protein 4-like
0.36	0.001880328	0.033989759	Un	rna142209	LOC106441027	gene115249	FAM10 family protein At4g22670-like
0.37	0.003505505	0.048519715	A1	rna2138	LOC106366980	gene1733	protein At-4/1-like
0.37	9.12384E-05	0.005724108	A9	rna57020	LOC106375223	gene45707	putative 60S ribosomal protein L30-1
0.37	0.001094529	0.024799333	A9	rna52237	LOC106349079	gene41915	fibrous sheath CABYR-binding protein
0.37	0.001725315	0.032461609	Un	rna140366	LOC106376607	gene113772	uncharacterized protein LOC106376607
0.37	0.000631762	0.017945241	A9	rna55099	LOC106367546	gene44184	
0.38	0.000959935	0.023040971	C2	rna72586	LOC106389800	gene58464	basic endochitinase CHB4-like chitinase 1
0.39	0.001902789	0.034208839	C5	rna101138	BNAC03G33270D	gene81681	uncharacterized protein BNAC03G33270D
0.39	0.000713417	0.019123104	C2	rna70758	LOC106380747	gene56882	probable sulfate transporter 3.5
0.39	0.00107413	0.024511307	A4	rna24123	LOC111215001	gene19337	probable protein phosphatase 2C 9

Appendix 4 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_Name	Reference Gene_ID	Full Description
0.39	0.000616403	0.017605709	C8	rna120359	LOC106453113	gene97322	NF-X1-type zinc finger protein NFXL1-like
0.39	0.000118831	0.006698642	A9	rna58425	LOC106419842	gene46848	zinc transporter 4, chloroplastic
0.39	0.002271504	0.03781129					
0.39	0.001873003	0.03392475	A9	rna54259	LOC111200555	gene43555	EG45-like domain containing protein 2
0.39	3.73466E-05	0.003205115	C3	rna82605	LOC106438306	gene66584	calcium-binding protein KIC-like
0.39	0.002633565	0.041023189	Un	rna140309	LOC106443544	gene113721	EG45-like domain containing protein 2 BnaCnng71940D
0.39	0.0010071	0.023642812	C5	rna99306	LOC106415092	gene80227	probable low-specificity L-threonine aldolase 1 BnaC05g06260D
0.39	0.000475699	0.015378693	C3	rna79691	LOC106386755	gene64193	probable sulfate transporter 3.5
0.39	0.000602229	0.01741964	C6	rna108556	LOC106410074	gene87643	NF-X1-type zinc finger protein NFXL1-like
0.40	0.000540603	0.016392882	C7	rna110830	LOC106386060	gene89497	EG45-like domain containing protein 2
0.40	0.001552423	0.03043618	C9	rna129291	LOC106353597	gene104676	putative lipid-transfer protein DIR1
0.40	0.002640795	0.041081111	C6	rna106886	LOC106406761	gene86299	
0.41	2.19998E-05	0.002267784	A7	rna42582	LOC106356879	gene34100	monothiol glutaredoxin-S6 BnaA07g19500D
0.41	3.80266E-05	0.0032328	A4	rna27364	LOC106453235	gene21964	calcium-binding protein KIC BnaA04g27150D
0.41	0.000699233	0.018907063	A9	rna53962	LOC106430003	gene43314	transcription initiation factor TFIID subunit 1-like
0.41	0.001546267	0.030397395	A9	rna51732	LOC106374646	gene41510	transcription factor bHLH146-like

Appendix 4 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_Name	Reference Gene_ID	Full Description
0.41	0.002567414	0.040522609	A6	rna33815	LOC106346251	gene27105	U2 small nuclear ribonucleoprotein A'-like
0.41	0.000641477	0.018160056					
0.41	0.000593247	0.017255865	C7	rna114899	LOC106411198	gene92768	proline dehydrogenase 1, mitochondrial
0.42	0.000607667	0.017493508	C4	rna93857	LOC106445196	gene75807	insulin-degrading enzyme-like 2
0.42	4.82985E-06	0.000853149	A1	rna4765	LOC106379225	gene3849	monothiol glutaredoxin-S10-like
0.42	7.24531E-05	0.004964175	A8	rna47762	LOC106372088	gene38279	60S ribosomal protein L28-2-like
0.42	3.4313E-05	0.003052768	C6	rna109460	LOC106406886	gene88355	60S ribosomal protein L27a-3-like
0.43	0.000632236	0.017945241	Un	rna144297	LOC111212589	gene116951	N6-adenosine-methyltransferase MT-A70-like
0.43	1.62228E-05	0.00183899	C4	rna95158	LOC106420803	gene76896	calcium-binding protein KIC BnaC04g50920D
0.43	0.0002115	0.009453463	A7	rna41070	LOC106358023	gene32833	squamosa promoter-binding-like protein 11
0.43	2.21293E-05	0.002274691	A6	rna35100	LOC106346832	gene28117	receptor like protein 30-like
0.43	0.001270304	0.027236081	Un	rna137456	LOC106426404	gene111358	mediator of RNA polymerase II transcription subunit 36a
0.44	0.000106478	0.006279616	C6	rna105408	LOC106382071	gene85108	probable inactive purple acid phosphatase 9 BnaC02g35040D
0.44	0.001729301	0.032493552	A6	rna36277	LOC106371418	gene29065	putative lipid-transfer protein DIR1
0.44	0.000349458	0.0127079	C8	rna120386	LOC106453100	gene97347	two-component response regulator ARR4 BnaC08g14280D
0.44	4.7084E-05	0.003765476	C8	rna121911	BNAC08G23610D	gene98588	uncharacterized protein BNAC08G23610D

Appendix 4 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_Name	Reference Gene_ID	Full Description
0.44	0.001573103	0.03067636	Un	rna146207	LOC106437091	gene118505	40S ribosomal protein S14-2-like
0.44	3.87679E-06	0.000744833	A6	rna38106	LOC106391236	gene30481	F-box/LRR-repeat protein 3-like
0.44	2.73764E-05	0.002641446	C4	rna96237	LOC106450055	gene77738	glutathione S-transferase U2
0.44	0.002199304	0.037170593	A9	rna55807	LOC106381667	gene44766	urea-proton symporter DUR3-like uncharacterized protein
0.44	8.62199E-05	0.005533283	A3	rna20955	LOC106419205	gene16754	LOC106419205
0.44	0.000712731	0.019123104	A9	gene41962	LOC106420338	gene41962	
0.45	0.000159476	0.007983398	A9	rna51352	LOC106366276	gene41190	U3 small nucleolar ribonucleoprotein protein MPP10-like
0.45	0.001627056	0.031205416	A9	rna52691	LOC106434239	gene42268	60S ribosomal protein L21-2-like
0.45	0.00014247	0.00739541	A4	rna23365	LOC106449010	gene18713	glutaredoxin-C11 BnaA04g00250D
0.45	1.82959E-05	0.002011329	A9	rna52337	LOC106422065	gene41985	60S ribosomal protein L3-2-like
0.45	0.001108327	0.024952856	Un	rna146218	LOC111213415	gene118515	
0.45	0.002657351	0.041116537	C2	rna72624	LOC106435228	gene58498	expansin-A1-like
0.45	0.001869445	0.033877181	A8	rna47043	LOC106422381	gene37677	serine/threonine-protein phosphatase PP1 isozyme 6-like
0.45	0.001598853	0.030956384	C6	gene88419	LOC106404530	gene88419	
0.45	0.002824824	0.042704492	C6	rna106672	LOC106394569	gene86144	uncharacterized protein LOC106394569
0.46	1.74597E-05	0.001948846	Un	rna137404	LOC106432188	gene111312	60S ribosomal protein L27a-3-like uncharacterized protein
0.46	0.000100807	0.006083216	A9	rna54962	BNAA09G54510D	gene44086	BNAA09G54510D
0.46	0.000788316	0.020272262	Un	rna139255	LOC106382197	gene112864	40S ribosomal protein S19-3-like
0.46	0.000130087	0.007109176	C2	rna72973	LOC106353779	gene58784	60S ribosomal protein L27a- 3 BnaA07g38870D BnaC02g20030D

Appendix 4 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_Name	Reference Gene_ID	Full Description
0.46	0.000439117	0.014713258	C3	rna86547	LOC106405890	gene69838	histone H2B.7 BnaC03g55640D
0.46	0.000278567	0.011151279	A6	rna33916	LOC106346197	gene27184	two-component response regulator ARR4-like
0.46	0.003120401	0.045136686	A3	rna13354	LOC106428854	gene10722	40S ribosomal protein S23-2 calcium-transporting ATPase 12, plasma membrane- type BnaC04g20690D
0.46	0.000611073	0.017518818	C4	rna93384	LOC106396742	gene75457	uncharacterized protein LOC106353918
0.47	0.001995488	0.035197195	A7	rna43633	LOC106353918	gene34977	60S ribosomal protein L36a
0.47	5.84521E-05	0.004288212	C3	rna87388	LOC106439716	gene70533	40S ribosomal protein S21-2-like
0.47	0.003478266	0.048299081	Un	rna141332	LOC106382780	gene114525	cysteine proteinase inhibitor 5 BnaC07g20300D
0.47	3.38517E-06	0.000673112	C7	rna113772	LOC106407373	gene91823	60S ribosomal protein L17-1-like
0.47	0.0022501	0.037653241	C3	rna83090	LOC106423484	gene66943	40S ribosomal protein S24-1
0.47	0.002213796	0.037325042	C3	rna86946	LOC106361898	gene70184	40S ribosomal protein S19-3-like
0.47	4.07268E-05	0.003391229	C2	rna76742	LOC106382223	gene61771	60S ribosomal protein L27a-3-like
0.47	0.000147658	0.007570588	A9	rna50979	LOC106366072	gene40875	60S ribosomal protein L3-2-like
0.47	1.87772E-05	0.002051842	A2	rna10220	LOC106395252	gene8158	40S ribosomal protein S24-1
0.47	0.0003056	0.011741271	Un	rna139614	LOC106416543	gene113168	uncharacterized protein BNACNNG49960D
0.47	4.25662E-05	0.003505161	C7	rna115175	LOC106348644	gene93001	uncharacterized protein BNACNNG07930D
0.47	1.53308E-05	0.001776611	C9	rna129249	BNACNNG49960D	gene104643	60S ribosomal protein L9-1-like
0.47	0.001760799	0.032739881	Un	rna135173	BNACNNG07930D	gene109492	
0.47	0.001849698	0.033687089	A2	rna8620	LOC106406948	gene6864	

Appendix 4 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_Name	Reference Gene_ID	Full Description
0.47	0.002685395	0.041366568	A4	rna26687	LOC106447433	gene21438	eukaryotic translation initiation factor 2 subunit alpha homolog
0.47	1.12208E-05	0.001458229	A3	rna19944	LOC106444181	gene15924	monothiol glutaredoxin-S10
0.48	0.003287963	0.046535354	C8	rna124814	LOC106412072	gene100959	zinc transporter 4, chloroplastic-like
0.48	0.00079831	0.020421363					
0.48	0.003347309	0.047118721	A3	rna20274	LOC106444235	gene16194	cold and drought-regulated protein CORA-like
0.48	0.000282879	0.011257312	A7	rna44100	LOC106354132	gene35345	transcription factor bHLH160-like
0.48	0.001611326	0.031022722	Un	rna136425	LOC106364133	gene110522	40S ribosomal protein S28-2
							tRNA:m(4)X modification enzyme TRM13 homolog tRNA:m(4)X modification enzyme TRM13-like
0.48	0.003005944	0.044011401	Un	rna141519	LOC106389094	gene114677	
0.48	0.001157134	0.025596228	Un	rna146393	LOC106355427	gene118660	auxin response factor 10
0.48	9.17039E-05	0.00574341	A3	rna16511	LOC106438023	gene13253	40S ribosomal protein S26-3-like
0.49	0.003478946	0.048299081	A2	rna12693	LOC106404571	gene10160	high mobility group B protein 7-like
0.49	0.001300124	0.02754378	A2	rna12639	LOC106404278	gene10112	protein DOWNY MILDEW RESISTANCE 6-like
0.49	0.00099475	0.023505101	C4	rna93974	LOC106390300	gene75920	60S ribosomal protein L21-2 BnaA09g15190D
0.49	0.001015008	0.023651312	C7	rna114794	LOC106407105	gene92689	protein AIG2 A-like protein AIG2 B-like

Appendix 4 continued

Fold Change	pval	qval	Chromosome	Transcript	Gene_Name	Reference Gene_ID	Full Description
0.49	0.003421251	0.047826535	A9	rna55368	LOC106395140	gene44399	
0.49	0.001065465	0.024414457	A8	rna47543	LOC106389474	gene38099	60S ribosomal protein L39-1
0.49	0.002953603	0.043653826	A4	rna26534	LOC106407316	gene21310	60S ribosomal protein L36-2
0.49	3.66896E-05	0.003188674	A4	rna23623	LOC106445047	gene18934	40S ribosomal protein S13-2
0.49	0.001607133	0.031022722	C3	rna85642	LOC106429399	gene69075	40S ribosomal protein S4-1-like
0.49	0.002655897	0.041116537	C9	rna128228	LOC106422139	gene103797	peroxidase 37 BnaC09g25420D
0.50	0.001316932	0.027715754	A3	rna16898	LOC106438345	gene13560	histone H4
0.50	0.000209263	0.009389215	A10	rna61322	LOC106387784	gene49190	anamorsin homolog
0.50	3.44782E-07	0.000126727	C4	rna95659	LOC106429353	gene77297	vacuolar cation/proton exchanger 1
0.50	0.00017324	0.008409431	C3	rna82438	LOC111204438	gene66428	60S ribosomal protein L38-like

Appendix 5. List of top five up-regulated genes of *Brassica napus* BA-044, BA-099, BA-210, BA-218 and BA-224 in response to Pi availability. Genes with q value < 0.05 and fold change,

<i>B. napus</i> line	Gene names	Description	fc	pval	qval
BA-044	LOC106392037	remorin like	773.64	0.05456	0.34706
	LOC106372453	aspartyl protease family protein	735.80	0.05456	0.34706
	LOC106446916	inner membrane protein ALBINO3	608.49	0.05456	0.34706
	LOC106394172	peroxisomal (S)-2-hydroxy-acid oxidase GLO1-like	466.83	0.01066	0.34706
	LOC106399203	photosystem I reaction center subunit IV A, chloroplastic-like	403.54	0.18341	0.53646
BA-099	LOC106415291	probable galactinol--sucrose galactosyltransferase 2	99.87	0.05278	0.58203
	LOC106381500	protein ACTIVITY OF BC1 COMPLEX KINASE 3, chloroplastic-like	90.33	0.03834	0.54538
	LOC106405688	fructose-bisphosphate aldolase 2, chloroplastic-like	70.75	0.23068	0.83977
	LOC106368471	cystine lyase CORI3-like	59.22	0.01457	0.48013
	LOC106381104	inner membrane protein ALBINO3, chloroplastic-like	46.12	0.02190	0.48514
BA-210	LOC106369031	inorganic pyrophosphatase 2-like	52.76	0.00036	0.14964
	LOC106382350	glyceraldehyde-3-phosphate dehydrogenase, cytosolic-like	50.96	0.00012	0.10194
	LOC106345785	protein EARLY-RESPONSIVE TO DEHYDRATION 7, chloroplastic-like	35.25	0.03374	0.36471
	LOC106418331	12-oxophytodienoate reductase 3	35.04	0.01994	0.33585
	LOC106450934	aquaporin PIP1-2	34.24	0.01410	0.32101
BA-218	LOC106446542	50S ribosomal protein L35, chloroplastic-like	6828.30	0.00889	0.96079
	LOC106372522	ferredoxin-dependent glutamate synthase 1, chloroplastic/mitochondrial	3328.94	0.11801	0.96079
	LOC106402589	rhodanese-like domain-containing protein 14, chloroplastic	1048.86	0.12152	0.96079
	LOC106432966	myrosinase-like	676.22	0.11490	0.96079
	LOC106451520	phosphomethylpyrimidine synthase, chloroplastic	542.24	0.08906	0.96079
BA-224	LOC106353439	glutathione S-transferase U20	127.44	5.01E-05	0.04238
	LOC111197984	inorganic pyrophosphatase 1	80.62	0.00074	0.12845
	LOC106381279	inorganic pyrophosphatase 1	74.67	0.00067	0.12497
	LOC106399564	inorganic pyrophosphatase 1	71.34	0.00221	0.20874
	BNA02G05080D	uncharacterized BNA02G05080D	64.93	0.00068	0.12551

