REVERSE GENETIC APPROACHES TO EXPLORE NEW GENES FOR RICE POLLEN DEVELOPMENT

Hong Lin, Msc

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

September 2017

ABSTRACT

In plants, normal anther and pollen development involves many important biological events and complex molecular regulatory coordination. Understanding the gene regulatory relationships during male reproductive development is essential for fundamental biology and crop breeding.

In this work, a rice gene co-expression network was developed for anther development (RiceAntherNet) that allows prediction of gene regulatory relationships during pollen development. RiceAntherNet was generated from 57 microarrays derived from rice anther tissue across all developmental stages. The microarray datasets from 9 rice male sterile mutants, including *msp1-4, ostdl1a, gamyb-2, tip2, udt1-1, tdr, eat1-1, ptc1* and *mads3-4* mutants were used to explore and test the anther development network. Among the changed genes, three clades showing differential expression patterns were further constructed to identify genes associated with pollen formation during anther development. Consistent with the functional predictions, in Clade 1, 7 reported genes (*OsABCG15, OsLAP5, OsLAP6, DPW, CYP703A3, OsNP1* and *OsCP1*) were confirmed to participate in pollen wall formation in rice. Furthermore, Clade 1 contained 12 genes with predicted orthologs in Arabidopsis that have been reported as key genes during pollen development in Arabidopsis and may play similar roles in rice. A website is also provided to present plots of the expression profiles for gene characterisation. This will assist in determining

the key relationships between genes involved in pollen development, thus enabling characterization of critical genes associated with these regulatory networks. 9 candidate genes predicted by RiceAntherNet were chose for further study.

Mutants of these genes were obtained by CRISPR-Cas9 editing system. Phenotyping of these mutants indicate their function in rice pollen wall formation, which provides additional evidence that RiceAntherNet is a valuable tool for identifying novel genes during anther and pollen development in plants. The reliability of RiceAntherNet offers the possibility of using reverse genetics approaches for novel gene characterization.

TIP2 was shown to be involved in cell differentiation at early stages of pollen development and previous studies suggested the presence of a TIP2-mediated transcription cascade that regulates tapetum development in rice. In this study, downstream genes of *TIP2* were predicted by reverse genetics approaches including microarray data analysis, gene expression patterns and gene promoter analysis. This work not only helps to explore the function of *TIP2* and its regulatory network, but also offers the possibility of using reverse genetic approaches for novel gene characterization.

ACKNOWLEDGEMENT

There are many without whom my work on this project would not have been possible. First of all, I want to express my deepest gratitude to my supervisor Prof. Zoe A. Wilson at the University of Nottingham. She gave me lots of patient guidance during my 4-years doctoral program and offers great supports in my research design and thesis writing. Her serious attitude to science, profound knowledge and charming personality touched me so deeply. Words cannot express how lucky I feel to be her student. I'm grateful to my assessor Prof. Malcolm Bennett for giving me valuable advices on the research.

I'd also like to thank my co-supervisor Dabing Zhang from Shanghai Jiao Tong University. He gave me the opportunity to receive PhD training in this UoN-SJTU joint program. Not only does Prof. Zhang educate me with his vast knowledge as a successful biologist, but also with his wisdom of life as a venerable elder. His dedication to science and education have been greatly inspiring me during these years. I'm also grateful to Prof. Wanqi Liang, Prof. Zheng Yuan and Dr. Jie Xu for giving me valuable advices for my studies.

I want to give special thanks to Dr. Simon Pearce in Zoe's group who is of vital importance for my study. Dr. Simon is so kind and so smart. He introduced me to bioinformatics step by step with his professional knowledge and skills. I couldn't have worked on this project without his help.

I really appreciate my lovely study team of four lead by Dr. Jing Yu. Jing is a good teacher and a good friend to me. She shares excellent ideas and teaches me a lot in the laboratory. Meanwhile her particular passion in science encourages me to overcome difficulties in research. The junior fellows in this team, Yulu Zhou and Yichun Xu, supported me a lot with my experiments. Our partnership and friendship make my study life rich and colorful. College life with you will be my most precious memory.

A big thank to my dear parents who are willing to do everything for my well-being. Last, but certainly not least, I am deeply grateful to Bin He, my beloved husband, who is always my strongest supporter.

TABLE OF CONTENTS

ABSTRACT I
ACKNOWLEDGEMENT III
TABLE OF CONTENTS V
LIST OF FIGURES X
LIST OF TABLES XIV
ABBREVIATIONS XV
1 INTRODUCTION 1
1.1 RICE ANTHER DEVELOPMENT 3
1.2 TAPETUM PROGRAMMED CELL DEATH (PCD) AND REGULATORY FACTORS 11
1.3 *TIP2*-RELATED ANTHER DEVELOPMENT 23
1.4 GENE CO-EXPRESSION ANALYSIS 25
1.5 AIMS OF THE PROJECT 32
2 MATERIALS AND METHODS 34
2.1 MATERIALS 34
2.1.1 Plant materials 34
2.1.2 Bacterial strains and vectors 34
2.1.3 Culture medium, buffers and chemicals 35
2.2 MICROARRAY DATA ANALYSIS 38
2.2.1 Quality control (QC) 38
2.2.2 RMA normalization 38
2.2.3 Pearson correlation coefficient (PCC) 39
2.2.4 Gene clustering 39
2.2.5 Gene annotation 40
2.2.6 Gene ontology (GO) analysis 40
2.2.7 Phylogenetic analysis 40
2.3 GENOTYPING 41
2.3.1 DNA extraction (by CTAB method) 41
2.3.2 PCR 42
2.3.3 Agarose gel electrophoresis 42
2.4 MOLECULAR CLONING METHODS 43

2.4.1 *E. coli* transformation by heat-shock 44
2.4.2 Plasmid DNA extraction and sequencing 45
2.4.3 *Agrobacterium* (EHA 105 and GV3101) transformation by heat-shock 45
2.4.4 Storage of bacteria as glycerol stocks 46
2.5 TARGETED MUTAGENESIS BY CRISPR-CAS9 SYSTEM 46
2.5.1 Selection of CRISPR targets 46
2.5.2 Plasmid construction 47
2.5.3 Callus tissues formation 48
2.5.4 Transformation of calli 49
2.5.5 Selection of transformed calli and plant growth 50
2.5.6 Genotyping 50
2.6 EXPRESSION ANALYSIS 51
2.6.1 Total RNA extraction by the trizol method 51
2.6.2 cDNA synthesis 51
2.6.3 Quantitative Reverse Transcriptase PCR (qRT-PCR) 52
2.7 Dual-Luciferase Reporter Assay 53
2.7.1 Plasmid construction 53
2.7.2 Preparation of bacteria solution for infiltration 53
2.7.3 Tobacco plants infiltration and transient expression analysis mediated by *Agrobacterium* 53
2.7.4 Fluorescence measurement 54
2.8 CHARACTERIZING PHENOTYPE OF MUTANT 55
2.8.1 Measurement of pollen viability by iodine staining 55
2.8.2 Scanning electron microscope 55
2.8.3 Semi-thin section 56
3 GENERATION OF A RICE GENE CO-EXPRESSION NETWORK FOR IDENTIFYING GENES INVOLVED IN ANTHER AND POLLEN DEVELOPMENT 57
3.1 INTRODUCTION 57
3.2 RESULTS 57
3.2.1 Datasets selection 59
3.2.2 Data processing 60

3.2.3 Generation of gene expression profiles from rice anthers 61
3.2.4 Construction of rice anther correlation network (RiceAntherNet) for gene function analysis 63
3.2.5 Mapping differentially expressed genes (DEGs) in male sterile mutants onto the correlation network 67
3.2.6 Clade 1 includes genes for pollen wall formation 74
3.2.7 Clade 2 includes genes for anther cell division and differentiation 88
3.2.8 Clade 3 includes genes expressed at the later stages of pollen development 91
3.2.9 Database for gene expression pattern query 93
3.3 DISCUSSION 94
3.3.1 RiceAntherNet is a useful tool for gene functional study during rice anther and pollen development 94
3.3.2 Processes of sporopollenin synthesis and transport are conserved in rice and Arabidopsis 95
3.3.3 Rice has gene regulation network that is partly not conserved compared to Arabidopsis 95
3.4 THE IMPORTANCE OF THIS WORK 96
4 VALIDATION OF TARGETS FROM ANTHERNET 98
4.1 INTRODUCTION 98
4.2 RESULTS 99
4.2.1 *Calmodulin like 41 (CML41)* 103
4.2.2 *Glucose-methanol-choline oxidoreductase (GMCO)* 114
4.2.3 *Dihydroflavonol-4-reductase Like1 (OSDRL1)* and *Dihydroflavonol-4-reductase Like2 (OSDRL2)* 118
4.2.4 *GAMYB Like 2 (GAMYBL2)* 123
4.2.5 Other candidate genes that were investigated for their roles in pollen development 128
4.2.6 *LOB DOMAIN 2 (LBD2)* in Arabidopsis 129
4.3 DISCUSSION 133
4.3.1 *CML41* plays as downstream of EAT1 and TDR in regulating Ca^{2+} homeostasis during rice anther development 134

4.3.2 Lipid synthesis genes are involved in pollen wall formation by affecting lipid metabolism 135
4.3.3 GA regulated MYB family members in rice pollen development 137
4.3.4 Other genes in RiceAntherNet would be good candidates for further study 137
5 *TIP2* REGULATION NETWORK 139
5.1 INTRODUCTION 139
5.2 RESULTS 139
5.2.1 Bioinformatics investigation for prediction of downstream gene targets of TIP2 140
5.2.2 LTPs were regulated by TIP2, TDR and GAMYB 143
5.2.3 LTPs were indirect downstream genes of TIP2 144
5.3 DISCUSSION 146
6 GENERAL DISCUSSION 148
6.1 GENERATION OF GENE EXPRESSION DATA FROM PUBLIC MICROARRAY DATASETS FOR REFERENCE 149
6.1.1 Gene expression profiles in rice anther were generated 149
6.1.2 Gene fold changes in male sterile mutants 150
6.1.3 Gene co-expression network during pollen and anther development were generated 150
6.2 REVERSE GENETIC APPROACHES FOR GENE FUNCTION PREDICTION 151
6.2.1 Forward genetics in plant biology 151
6.2.2 Reverse genetics in plant biology 152
6.2.3 RiceAntherNet for prediction of genes involved in anther and pollen development 152
6.2.4 Exploration of downstream genes of TIP2 by using microarray data of *tip2* 153
6.3 GENE FUNCTION INVESTIGATION FOR NOVEL GENES PREDICTED FROM THE CORRELATION NETWORK 155
6.3.1 The function of DFR/TKPR (DIHYDROFLAVONOL 4-REDUCTASE) family in pollen development is conserved in rice

and Arabidopsis 155
6.3.2 *GMCO (Glucose-methanol-choline oxidoreductase)* may be involved in cutin component biosynthesis 157
6.3.3 *GAMYBL2* may play important role in pollen exine formation 158
6.3.4 Study of *CML41 (calmodulin like 41)* may help with understanding Ca^{2+} signalling during tapetum PCD 159
6.3.5 *LBD2* in Arabidopsis affect stamen and petal development 160
6.3.6 Other candidates would be good targets for deepening understanding of gene regulation network during pollen development in rice 161
6.4 CONCLUSIONS 161
REFERENCES 163
APPENDIX 180

LIST OF FIGURES

Figure 1-1 Hybrid breeding systems.....3

Figure 1-2 Diagrams of biological events and major regulators during anther development.....4

Figure 1-3 Schematic model of pollen wall development in *A. thaliana*. (Jiang et al., 2013).....9

Figure 1-4 Pollen wall structure of a typical angiosperm pollen grain (Ariizumi & Toriyama, 2011) 11

Figure 1-5 Principal pathway of GA metabolism in plants. (Sakamoto et al., 2004)…….....15

Figure 1-6 Model for GA Signaling in the Anther. (Aya et al., 2009).....17

Figure 1-7 Comparison of Genes/Enzymes Involved in Various Stages of Pollen Wall Development in Arabidopsis and Rice. (Shi et al., 2015).....22

Figure 1-8 Proposed Model for the TIP2-TDR-EAT1 Cascade That Regulates Anther Wall Cell Fate and Microspore Development in Rice. (Fu et al., 2014).....25

Figure 1-9 Definitions of terms describing networks (Aoki et al., 2007).....29

Figure 3-1 Workflow of predicting candidate genes by co-expression network.....59

Figure 3-2 Boxplots and density plots of log-intensity distribution for 57 arrays before and after normalization.....61

Figure 3-3 Gene expression profile of *TDR* using the generated RiceAntherNet tool (www.cpib.ac.uk/anther/riceindex.html) 62

Figure 3-4 Distribution of edge correlations.....64

Figure 3-5 Global co-expression network plot by organic layout in

Cytoscape…….....65

Figure 3-6 Network generated by TransClust..... 66

Figure 3-7 Analysis of gene Cluster 116 from RiceAntherNet.....67

Figure 3-8 Differentially Expressed Genes (DEGs) of male sterile mutants in the global RiceAntherNet co-expression network.....69

Figure 3-9 Example clusters from the RiceAntherNet.....70

Figure 3-10 Hierarchical clustering of example RiceAntherNet clusters.....71

Figure 3-11 Expression patterns of Clade 1 genes in rice anther and other tissues……74

Figure 3-12 Number of Differentially Expressed Genes (DEGs) in Clade 1……....76

Figure 3-13 Analysis of TKPR genes in Clade 1.....80

Figure 3-14 Expression patterns of Clade 2 in rice anther and other tissues.....89

Figure 3-15 Phylogenetic tree of LRR-RLK genes in Clade 1.....90

Figure 3-16 Expression patterns of Clade 3 in rice anther and other tissues..92

Figure 3-17 Functional framework of RiceAntherNet database.....94

Figure 4-1 Expression patterns of 7 candidate genes in rice anther and other tissues.....102

Figure 4-2 Molecular Identification of *CML41*.....104

Figure 4-3 The expression of *CML41* in WT anther.....105

Figure 4-4 Expression patterns of 24 CML genes in rice anther and other tissues……..107

Figure 4-5 Regulation of the *CML41* Promoter by TIP2, TDR and EAT1.....110

Figure 4-6 Schematic representation of *CML41* with CRISPR targets,

genotyping primers and sequencing information of *cml41-1* mutant.....111

Figure 4-7 Pollen viability staining, using potassium iodine, of *cml41-1* at mature pollen stage.....112

Figure 4-8 SEM pictures of anthers and pollen grains in WT and *cml41-1* at stage 12.....113

Figure 4-9 The 3000 bp Promoter/Upstream region of *GMCO*.....115

Figure 4-10 Schematic representations of *GMCO* with CRISPR targets.....116

Figure 4-11 SEM analysis of WT, *osnp1-4* and *gmco-1 osnp1-4* at stage12.....118

Figure 4-12 Schematic representations of *OsDRL1* and *OsDRL2*.....120

Figure 4-13 Pollen iodine viability staining of *osdrl1-1* and *osdrl2-1* at mature pollen stage.....121

Figure 4-14 SEM pictures of anthers and pollen grains in WT.....123

Figure 4-15 The expression of *GAMYBL2* in WT anther (RiceAntherNet: https://www.cpib.ac.uk/anther/riceindex.html).....124

Figure 4-16 Schematic representations of *GAMYBL2* with CRISPR targets, genotyping primers and sequencing information of *gamybl2-1, gamybl2-2, gamybl2-3* and *gamybl2-4* mutants.....125

Figure 4-17 SEM pictures of pollen grains in WT and the *gamybl2-3*.....127

Figure 4-18 Schematic representations of *OsLAP6, LRR1* and *LRR5* with CRISPR targets, genotyping primers and sequencing information of mutants.....129

Figure 4-19 The expression of *LBD2* in rice and Arabidopsis.....130

Figure 4-20 Schematic representations of *LBD2* with qRT-PCR primers and

T-DNA insertion line information.....131

Figure 4-21 qRT-PCR of *LBD2* in WT and *lbd2-1*.....132

Figure 4-22 Phenotypes of *lbd2-1*.....133

Figure 5-1 Gene expression analysis and E-box analysis of DEGs of *tip2*.....141

Figure 5-2 Heatmap of fold changes of LTPs in *tip2, eat1-1* and *gamyb-2* mutants……..……...…144

Figure 5-3 Regulation of the *LTPL1 and LTPL2* Promoters by TIP2.....146

LIST OF TABLES

Table 2-1 Bacterial strains.....35

Table 2-2 Vectors.....35

Table 2-3 General reaction composition.....42

Table 2-4 Standard PCR program.....42

Table 2-5 Reaction composition for GXL polymerase…….....43

Table 2-6 PCR program for GXL polymerase.....44

Table 2-7 In-fusion HD cloning reaction composition.....44

Table 2-8 Annealing reaction composition.....47

Table 2-9 Restriction Enzyme Digestion composition.....47

Table 2-10 Ligation reaction composition.....48

Table 2-11 SYBR Green PCR reaction composition.....52

Table 2-12 qRT-PCR program.....52

Table 3-1 Nine male sterile mutant datasets.....68

Table 3-2 GO enrichment analysis for 3 Clades.....73

Table 3-3 Genes in Clade 1.1.....77

Table 3-4 Genes in Clade 1.2.....83

Table 4-1 Annotations of candidate genes.....100

Table 4-2 Fold change of candidate genes in male sterile mutants.....100

Table 4-3 Phenotypes of *DIHYDROFLAVONOL 4-REDUCTASE* (*DFR)* genes in rice and Arabidopsis.....137

Table 5-1 GO enrichment analysis for 31 putative downstream genes of *TIP2*.....142

Table 5-2 Candidate genes in lipid metabolism pathway that show altered expression in the *tip2* mutant.....143

ABBREVIATIONS

µg/ml:	microgram per millilitre
µl:	microlitre
µM:	micromolar
ABCG:	*ATP Binding Cassette G Transporters*
ACOS5:	*CoA Synthetase5*
AMS:	*ABORTED MICROSPORES*
AS:	Acetosyringone
bHLH:	basic-Helix–Loop–Helix
BLAST:	Basic Local Alignment Search
bp:	base pair
Callase:	Glucan endo-1,3-b-D-glucosidase
Callose:	b-1,3-glucan
CDF:	Custom Chip Definition File
cDNA:	complementary Deoxyribonucleic Acid
cm:	centimeter
CML41:	*Calmodulin like 41*
DAP:	Days after pollination
ddH2O:	distilled water
DEG:	Differentially Expressed Gene
DEPC:	diethylpyrocarbonate
DNA:	Deoxyribonucleic Acid
DNase:	deoxyribonuclease
dNTP:	Deoxynucleotide triphosphate
DPW:	*DEFECTIVE POLLEN WALL*
DRL1:	*Dihydroflavonol 4-reductase-like1*
DYT1:	*DYSFUNCTIONAL TAPETUM1*
ER:	*ERECTA*
g:	gram
GAMYBL2:	*GAMYB Like 2*
Gen:	Gentamycin
GEO:	Gene Expression Omnibus

GMCO:	*Glucose-methanolcholine oxidoreductase*
GO:	Gene Ontology
h:	hour
HM:	Homozygous
HT:	Heterozygous
Kan:	Kanamycin
kb:	kilo base pair
l:	litre
LB:	Luria Broth
LBD2:	*LOB DOMAIN 2*
LRR-RLK:	*Leucine-rich repeat receptor-like kinase*
LTP:	*Lipid Transfer Protein precursor*
M:	Mutant
MES:	2-(N-morpholino) ethanesulfonic acid
min:	minute
ml:	millilitre
mM:	milimolar
mRNA:	messenger Ribonucleic Acid
MS:	Murashige and Skoog Basal Medium
MS2:	*MALE STERILE2*
MSP1:	*MULTIPLE SPOROCYTE 1*
NASC:	Nottingham Arabidopsis Stock Centre
NCBI:	National Center For Biotechnology Information
ng/µl:	nanogram per microlitre
ng:	nanogram
OsCP1:	*Oryza sativa Cysteine Protease1*
OSDRL2:	*Dihydroflavonol-4-reductase Like2*
OsGPAT3:	*Glycerol-3-Phosphate Acyltransferase 3*
OsNP1:	*No Pollen 1*
OsTDL1A:	*TPD1-like 1A*
PCC:	Pearson Correlation Coefficient
PCD:	Programmed Cell Death

PCR:	Polymerase Chain Reaction
PKSA:	*POLYKETIDESYNTHASEA*
PKSB:	*POLYKETIDESYNTHASEB*
PTC1:	*PERSISTANT TAPETAL CELL1*
QC:	Quality Control
QRT3:	*QUARTET 3*
qRT-PCR:	Quantitative Reverse Transcriptase PCR
Rif:	Rifampicin
RMA:	Robust Multi-array Average
RNA:	Ribonucleic Acid
RNAi:	Ribonucleic Acid Interference
ROS:	Reactive Oxygen Species
rpm:	revolutions per minute
sec:	second
SEM:	Scanning electron microscope
SKM2:	*STERILITY-REGULATING KINASE MEMBER2*
TAIR:	The Arabidopsis Information Resource
T-DNA:	Transferred-Deoxyribonucleic Acid
TDR:	*TAPETUM DEGENERATION RETARDATION*
TF:	Transcription Factor
TIP2:	*TDR INTERACTING PROTEIN2*
TKPR1:	*TETRAKETIDE ALPHA-PYRONE REDUCTASE 1*
TKPR2:	*TETRAKETIDE ALPHA-PYRONE REDUCTASE 2*
UDT1:	*UNDEVELOPED TAPETUM1*
UTR:	Untranslated Regions
UV:	Ultraviolet
v/v:	volume to volume ratio
w/v:	weight to volume radio
WT:	Wild type

1 INTRODUCTION

Rice is one of the most important food crops which feeds over half of the world's population (Londo et al., 2006). As Arabidopsis is used widely as model plant in research, cultivated rice (*Oryza sativa L.*) is taken as the model plant for monocot crops (Sasaki, 1997). There are some advantages for rice being a model species, such as small genome size, abundance of high-through put data, particularly sequence and transcriptomic data, highly efficient transformation systems to generate T-DNA insertion lines and CRISPR/Cas9 mutants (Guo et al., 2014; Sasaki, 1997).

To help meet the demand for crop breeding, yield increases and other agricultural targets, much research has focused on plant reproduction. Hybrid crops are important in this since they frequently show enhanced yield and resilience to pests and disease; however, generating hybrids is challenging particularly in autogamous species. Manipulating male-sterile female lines is therefore a key target for hybrid breeding (Zhang & Liang, 2016). Currently, three-line hybrid systems and two-line hybrid systems have been widely performed to increase rice yield (Huang et al., 2014) (Figure 1-1). Three-line hybrid systems typically involve the use of Cytoplasmic Male Sterile (CMS) and restorer lines, whereas two-line systems involve lines where male sterility can be switched by growth under different environmental conditions. The three-line system uses i) a cytoplasmic male-sterile line lacking nuclear

restorer genes (*rf*) - as the female parent, ii) a maintainer line that has normal cytoplasm but lacks nuclear recessive restorer gene (*rf*) - for maintaining the sterile line when crossed to the CMS line, and iii) the restorer line - a cytoplasmic male-fertile line carrying nuclear dominant restorer gene (*Rf*) - as a pollen donor (male parent) (Figure 1-1A). However, the poor genetic diversity of both CMS and restorer line limits the further application of three-line hybrid system (Chang et al., 2016).

In two-line hybrid systems, environment sensitive genic male sterile (EGMS) lines which cause male sterile only under a certain environmental condition act as female parent lines as well as being able to produce EGMS progeny (Figure 1-1B). Since EGMS lines are usually the mutations of nuclear genes, more genetic diversity could be achieved by identifying more environment sensitive male-sterile genes. EGMS tend to utilise temperature or day-length dependent systems (Chen and Liu, 2014), however, the control of environment for complete male sterility in the two-line hybrid can be difficult for some EGMS lines (Zhang & Liang, 2016). Therefore, identifying more nuclear male-sterile genes as well as new rice male sterility systems are needed. Understanding the gene regulatory relationships during male reproductive development is essential for fundamental biology and crop breeding.

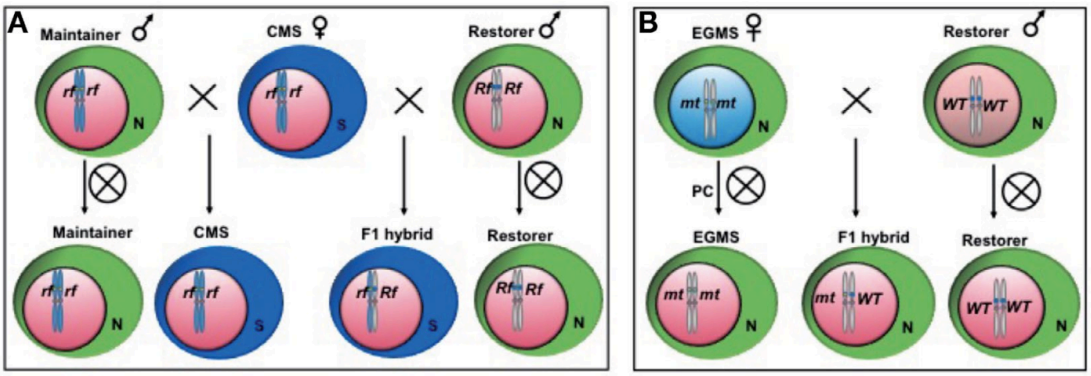


Figure 1-1 Hybrid breeding systems. A: Cytoplasmic male sterility (CMS)-based three-line hybrid breeding system. B: Environment-sensitive genic male-sterile (EGMS)-based two-line hybrid breeding system. PC: permissive condition for producing the EGMS progeny.

1.1 RICE ANTHER DEVELOPMENT

There are many important biological events that occur during anther and pollen development. In Arabidopsis, anther development is divided into 14 stages based on the distinctive cellular events visualized by light microscopy (Sanders et al., 1999). In 2005, Itoh *et al*. proposed a staging system for anther development in rice (Itoh et al., 2005). More and more reports of rice anther development staging have led to a better understanding of the whole male reproductive developmental process. A widely used staging system has been proposed with details of cytological changes of anther development in rice being well described to facilitate the study of rice male reproductive development (Zhang et al., 2011).

In Zhang's staging system (Zhang et al., 2011), there are 14 stages in rice pollen development, including 3 phases: the cell division and differentiation phase (stage 1 to 6), the meiosis and PCD initiation phase (stage 7 to 8) and late stages of pollen development (stage 9-14). The biological events and

major regulators during every stage of anther development are shown in Figure 1-2.

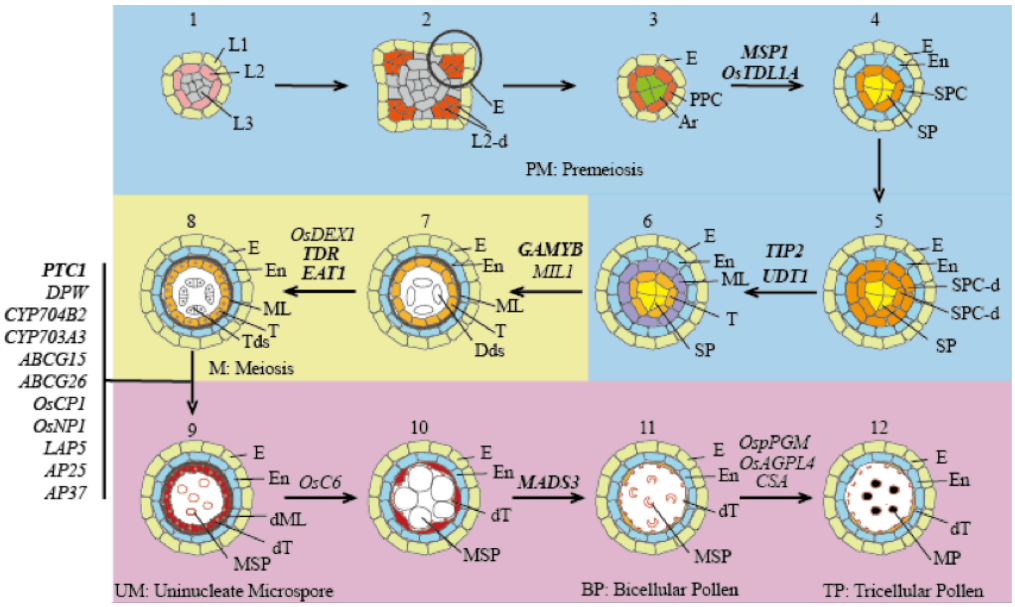


Figure 1-2 Diagrams of biological events and major regulators during anther development. Structure: Three cell layers in stamen primordia: L1 = Layer 1, L2 = Layer 2, L3 = Layer 3; E = Epidermis; L2-d = L2 derived (derive to Ar = Archesporial cell and PPC = primary parietal cell layer); En= Endothecium; SPC= Secondary parietal cell layer; SPC-d=SPC derived (derive to ML = Middle layer and T = Tapetum); SP = Sporogenous cell; Dds = Dyad cell; Tds = Tetrads; dML = Degraded middle layer; dT = Degraded Tapetum cells; MSP = microspore parietal cell. Developmental stages (Anther Stage 1-14) are divided into 3 phases. Phase 1 (blue background): PM (Premeiosis) = Anther Stage 1-6; Phase 2 (yellow background): M (Meiosis) = Anther Stage 7 and 8; Phase 3 (pink background): after meiosis, including UM (Uninucleate microspore) = Anther Stage 9 and 10; BP (Bicellular Pollen) = Anther Stage 11; TP (Tricellular Pollen) = Anther Stage 12; MP (Mature Pollen) = Anther Stage 13 and 14. Genes with mutant datasets used in this research are in bold.

1.1.1 Cell division and differentiation

At the cell division and differentiation phase (stage 1 to 6), differentiation of anther primordium occurs. At stage 1, anther primordium is observed with 3 cell layers: L1, L2 and L3. At stage 2, archesporial cell (Ar) is formed from L2 cellular layers and connective tissue is formed from L3. Later at stage 3, Ar divides and differentiates into primary parietal cells (PPCs). In stage 4, PPCs

further generate the secondary parietal cells (SPCs) and primary sporogenous cells (PSCs). During stage 5, the PPCs generate the endothecium layer, and the middle layer and the tapetum layer. Ar generates microspore mother cells at stage 6. After the cell division and differentiation during this phase, the innermost meiocytes were formed encased within four somatic anther cell layers; from inner to outer these are the tapetum, middle layer, endothecium and epidermis (Ma, 2005; Zhang et al., 2011; Zhang & Yang, 2014).

In this process, MSP1 (MULTIPLE SPOROCYTE 1) and OsTDL1A (TPD1-like 1A) interact with each other to specify the cell identity of anther wall layers and microsporocytes in rice, which is conserved in Arabidopsis (Yang et al., 2016). *MSP1* encodes a Leu-rich repeat (LRR) receptor-like protein kinase. It is an ortholog to *EXS/EMS1* in Arabidopsis. The phenotype of *msp1* mutant shows increased numbers of male and female sporocytes, and disrupted anther wall layers which lack the tapetum cell layer (Wang et al., 2006). *OsTDL1A* is reported as a TPD1-like gene, which is co-expressed with *MSP1* and binds to the LRR domain of MSP1, that plays a role as a corresponding rice ligand to limit sporocyte numbers (Zhao et al., 2008). *UNDEVELOPED TAPETUM1 (UDT1)* encodes a basic helix-loop-helix (bHLH) gene which is required for generate mature tapetum cells from secondary parietal cells (Jung et al., 2005). The homologous gene of *UDT1* in Arabidopsis, *DYT1 (DYSFUNCTIONAL TAPETUM1)* regulates a series of bHLH transcription factors and MYB transcription factors to regulate the

tapetum development (Barrett et al., 2013; Phan et al., 2012; Zhu et al., 2015). In early premeiosis stage, the mutant of *UDT1* seems normal, but during the meiotic stage, the mutant tapetum becomes vacuolated because of the failure to differentiate (Jung et al., 2005). Another bHLH gene *TIP2 (TDR INTERACTING PROTEIN2)* regulates differentiation of middle layer and tapetum in rice (Fu et al., 2014; Ko et al., 2014).

1.1.2 Meiosis and PCD initiation

Meiotic division of meiocytes occurs from stage 7 to stage 8, forming dyads, tetrads and release of free haploid microspores by stage 9. At stage 7, meiosis starts and callose (b-1,3-glucan) accumulates. At stage 8, along with meiosis, dyads and tetrads surrounded by callose are generated during meiosis I and meiosis II respectively, meanwhile, tapetum cells become vacuolated and initiate PCD.

During this meiosis and PCD initiation phase, *Microsporeless 1 (MIL1), MEL1 MEIOSIS ARRESTED AT LEPTOTENE1 (MEL1)* and *MEIOSIS ARRESTED AT LEPTOTENE 2 (MEL2)* are required for meiosis initiation; *PAIR1 HOMOLOGOUS PAIRING ABERRATION IN RICE MEIOSIS 1 (PAIR1), TOPOISOMERASE 6 SUBUNIT A1 (OsSPO11-1), TOPOISOMERASE 6 SUBUNIT A (OsSPO11-4), Completion of Meiosis 1 (OsCOM1), MEIOTIC PROTEIN DMC1A (OsDMC1), DNA REPAIR PROTEIN RAD51A2 (OsRAD51)* are involved in DNA double-strand break (DSB) formation and

processing in rice (Luo et al., 2014). In Arabidopsis, *SYNAPTIC 1 (SYN1/DIF1), SWITCH1 (SWI1/DYAD)* are essential for sister chromatid cohesion in meiosis; *meiotic asynaptic mutant 1 (ASY1), HOP1, ARABIDOPSIS HOP2 HOMOLOG (AHP2), SPORULATION 11-1 (AtSPO11-1), AtRAD50, MEIOTIC RECOMBINATION 11 (MRE11), AtDMC1, AtRAD51* are required for chromosome pairing, synapsis and recombination (Itoh et al., 2005).

Tapetum PCD is essential for microspore development and pollen wall formation (Zhang et al., 2008). This process is regulated by many factors, including phytohormones, reactive oxygen species (ROS), lipid metabolism and other signals (Zhang & Yang, 2014). The details about tapetum PCD and the regulation factors will be further described in section 1.2.

1.1.3 Late stages of pollen development

After meiosis, free haploid microspores, surrounded by primexine which is maintained by the callose wall, are released at stage 9. The primexine containing neutral and acidic polysaccharide materials is formed during the tetrad stage, under sporophytic regulation (Ariizumi et al., 2008). Callase (Glucan endo-1,3-b-D-glucosidase) is subsequently secreted from the tapetum cells to degrade the outer callose wall and release the microspores. Although not all mechanisms of pollen exine patterning are clear, sporopollenin synthesis and deposition are recognized as a major developmental processes

(Ariizumi & Toriyama, 2011). Sporopollenin is synthesized in the tapetum cells, with associated high levels of gene expression observed (section 1.2.4); primexine and callose wall formation are essential for correct sporopollenin deposition (Ariizumi & Toriyama, 2011).

Exine on the microspore surface becomes thicker at stage 10. Tapetum cells degeneration is required for pollen exine formation (Jiang et al., 2013). Figure 1-3 shows that pollen wall initiation starts at meiosis when tapetum cell degradation is initiated. Microspores undergo the first mitotic division at stage 11 (bicellular pollen stage), forming a generative cell and a vegetative cell. By then, the pollen wall is formed (Figure 1-4) and the tapetum has broken down (Figure 1-3). Then the generative cell undergoes the second mitosis in stage 12 and generates mature pollen grain. In stage 13 and 14, the flower opens and the mature pollen grains are released.

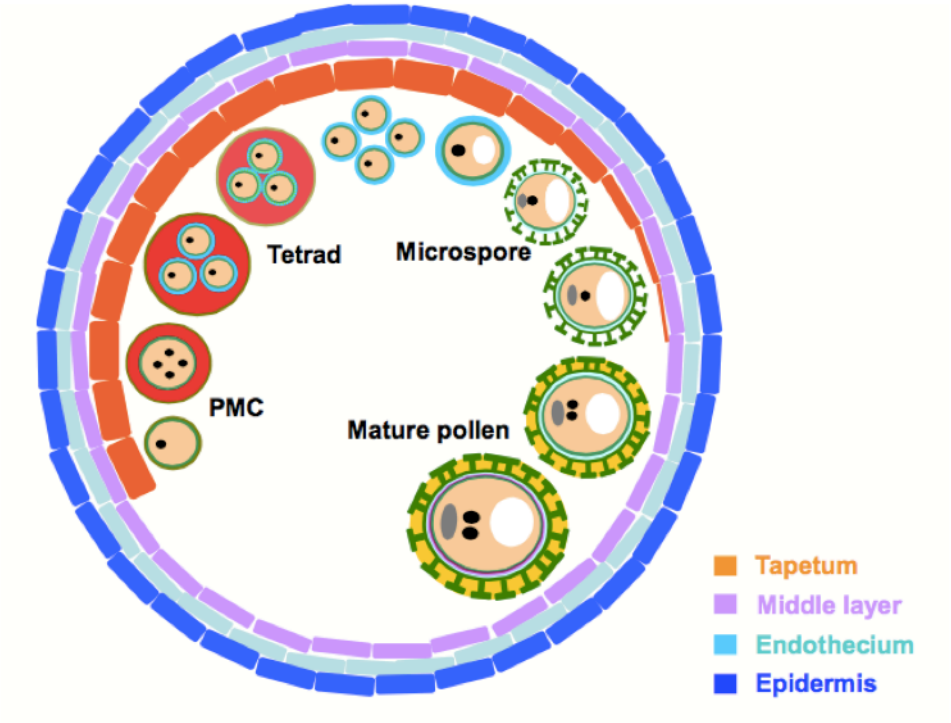


Figure 1-3 Schematic model of pollen wall development in *A. thaliana*. (Jiang et al., 2013). PMCs (pollen mother cells) surrounded by the tapetum undergoe meiosis to produce tetrads with four microspores. At the tetrad stage, pollen wall development is initiated. The microspores with primexine are released from the tetrad because the callose wall is degenerated by tapetum-produced callase. During tapetum PCD (anther stage 9 to stage 11), the thickness of the exine increases along with deposition of polymerized sporopollenin derived from the tapetum. The intine is formed at the late bicelluar stage. The tryphine, recognized as a tapetum remnant, fills the inter space of the exine at the tricelluar stage.

The mature pollen wall includes the exine as outer layer, intine as inner layer and tryphine which fills the interspaces of the exine (Figure 1-4) (Ariizumi & Toriyama, 2011). The formation of exine and tryphine is principally regulated by the tapetum, whereas intine formation is regulated by the developing microspore (Shi et al., 2015).

Exine is a complex structure contains two layers: the outer sexine layer with baculum and tectum, which is sculpted in species-specific forms; the inner

nexine bilayer, which consists of nexine I and nexine II, acts as a skeleton for exine formation (Figure 1-4). It has been found that polyhydroxylated aliphatics and phenolics are the major components of exine (Shi et al., 2015). Besides the protective role that the exine plays for pollen, it also acts as genetic barrier for fertilization (Zinkl et al., 1999).

The intine is the inner layer beneath the exine (Figure 1-4) composed of cellulose, pectin and proteins. The major component of the intine is pectin. As the innermost layer of pollen wall, the intine is required for pollen grain structure maintenance, as well as pollen germination and pollen tube growth (Shi et al., 2015).

Tryphine, also known as the pollen coat, distributes in the depressions of the pollen exine (Figure 1-4). It is hydrophobic and full of lipids and other components, such as flavonoids, proteins, pigments and aromatic substances. The lipids help to minimise water loss from the pollen grain and flavonoids protect the pollen from UV radiation damage, pathogen attack as well as photo-oxidative damage (Ariizumi & Toriyama, 2011). In addition, tryphine plays a role in pollination by attracting animal vectors. The species which are wind-pollinated as rice, usually have less tryphine on the pollen.

The formation of pollen wall is a key process at later pollen development stages. The pollen wall not only protects pollen grains from biotic and abiotic stresses but also play an important role in pollination and fertilization (Shi et

al., 2015). Normal tapetum PCD is essential for pollen wall formation since tapetum provide nutrition and sporopollenin precursors for this process (section 1.2.4).

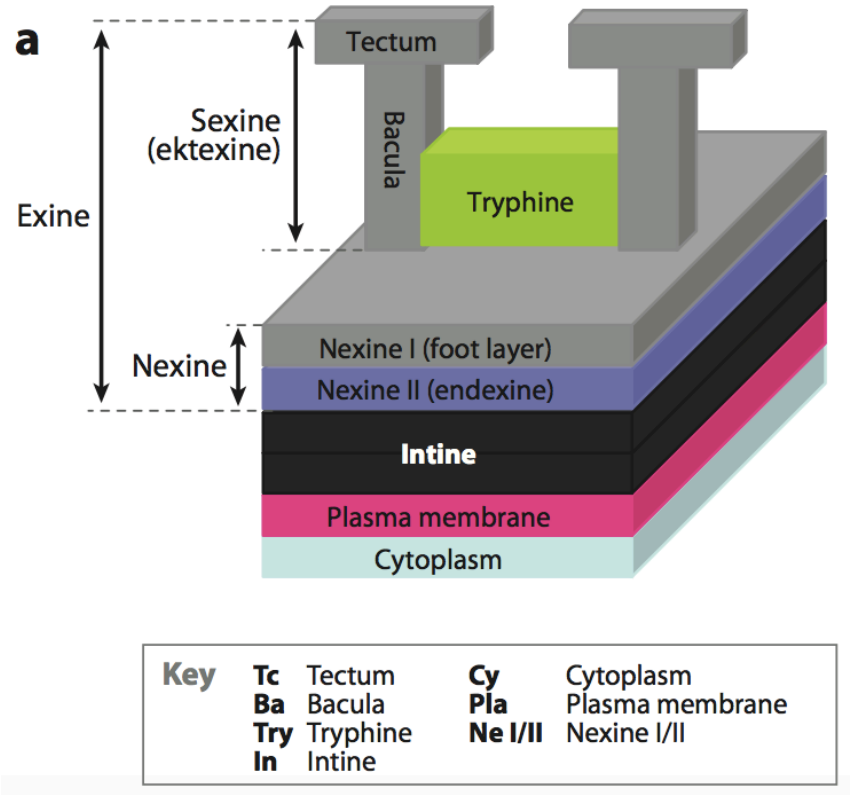


Figure 1-4 Pollen wall structure of a typical angiosperm pollen grain (Ariizumi & Toriyama, 2011)

1.2 TAPETUM PROGRAMMED CELL DEATH (PCD) AND REGULATORY FACTORS

The tapetum is the innermost layer encasing meiocytes. In addition to providing physical support for microspore development, as a transitory tissue, tapetum cells degenerate by a PCD process to provide cellular contents as nutritional support for the formation of pollen grains and pollen wall (Li et al., 2006a; Quilichini et al., 2014; Zhang & Yang, 2014). PCD is a critical process

for development and stress response in plants. There are different morphological types of cell death, including apoptosis, autophagy and non-lysosomal PCD (van Doorn & Woltering, 2005). Tapetum cell death occurs via an apoptosis, process that involves shrinkage of the cytoplasm, cleavage of DNA, rupture of the vacuole, swelling of the endoplasmic reticulum and disintegration of mitochondria and the cytoskeleton (Li et al., 2006a; van Doorn & Woltering, 2005; Parish & Li, 2010). Recent studies of plant PCD have revealed that many factors including phytohormones, ROS, Ca^{2+} signal, lipids and proteases regulate tapetum PCD programs dynamically (Li et al., 2006a; Quilichini et al., 2014; Yi et al., 2016; Yu et al., 2016; Zhang & Yang, 2014).

1.2.1 Tapetum PCD process

Tapetum development can be divided into the formation phase (anther stages 1 to 5), differentiation phase (anther stages 6 and 7), and the degradation phase (anther stages 8 to 12). The tapetum is formed from secondary parietal layer until stage 5 and disappears in stage 12. Details of tapetum degradation phase are described below according to Zhang's staging system (Zhang et al., 2011).

At stage 7, Meiocytes initiate meiosis and contact the tapetum layer; meanwhile, microsporocytes synthesize a transient callose wall.

At Stage 8, meiocytes undergo meiosis I and meiosis II, "dyads" and "tetrads" are formed because of the surrounding by callose wall. Tapetum cells becomes

condensed and vacuolated, the TUNEL (terminal deoxynucleotidyl transferase-mediated dUTP nick end-labeling) signal which is used for detecting DNA fragmentation could be observed in stage 8, indicating the initiation of tapetum cell death (Wilson & Zhang, 2009).

At stage 9, tapetum secretes callase, which leads to the degeneration of the callose wall and release of the free haploid microspores. This indicates the role of the tapetum in microspore wall formation at later stages (Quilichini et al., 2014). Callose wall breakdown at this stage is essential for producing viable pollen (Scott et al., 2004). The tapetum becomes more condensed and produces Ubisch bodies on the inner surface of tapetum layer facing towards the microspores. Ubisch bodies are thought to be responsible for transporting materials from the tapetum to microspores (Li et al., 2011a).

At stage 10, the tapetum produces abundant electron-dense Ubisch bodies and continues degenerating. During the tapetum degradation process, vesicles containing sporopollenin precursors release their cell contents into the anther locule for construction of the pollen wall. On the outer surface of the microspores, exine with nexine and sexine layers become thicker (Zhang et al., 2011).

At stage 11, the tapetum cells nearly completely degenerate, leaving cellular debris and Ubisch bodies on the internal surface.

The high level of coordination of tapetum PCD and pollen and pollen wall

formation indicates the vital importance of the tapetum. The tapetum not only promotes the meiosis of the microspore mother cell during meiosis stage, but also regulates later microspore development and pollen exine formation, by secreting cell wall products and then degrading to release more materials into the locale for deposition onto the developing pollen grains (Li et al., 2006a).

1.2.2 Phytohormones as signals in regulation of tapetum PCD

Phytohormones are involved in many key biological processes in the plant kingdom including growth, defence and reproduction (Hirano et al., 2008). Most plant developmental PCD is regulated by phytohormones through hormonal signalling pathways (Kuriyama & Fukuda, 2002).

1.2.2.1 Gibberellin (GA)

Gibberellins (GAs) are growth factors that form a large family of tetracyclic diterpenoid phytohormones (Sakamoto et al., 2004). GA signalling is important for plant reproductive development, including floral initiation, floral organ development including anther development and pollen microsporogenesis (Achard et al., 2004).

Research in 2004 showed that GA-deficient mutants have a typical phenotype of dwarf and abnormal flower development by screening 29 GA metabolic enzymes and further identifying 18 GA-defective mutants in rice (Sakamoto et al., 2004). The major pathway of GA metabolism in plants is shown in Figure 1-5 (Sakamoto et al., 2004). Among the 29 GA metabolic enzymes,

ent-copalyl diphosphate synthase (CPS) is a type of enzyme which converted trans-geranylgeranyl diphosphate (GGDP) to ent-copalyl diphosphate (CDP) (Hedden & Phillips, 2000). One of the most severe GA-defective mutants, *oscps1-1,* shows abnormal tapetum cells without degradation after anther stage 8. However, in *oscps1-1*, the defect in anther development could be rescued by treating the mutant with GA_3 at the panicle primordium stage (Aya et al., 2009). This suggests the essential role of GA in regulating tapetum cell death.

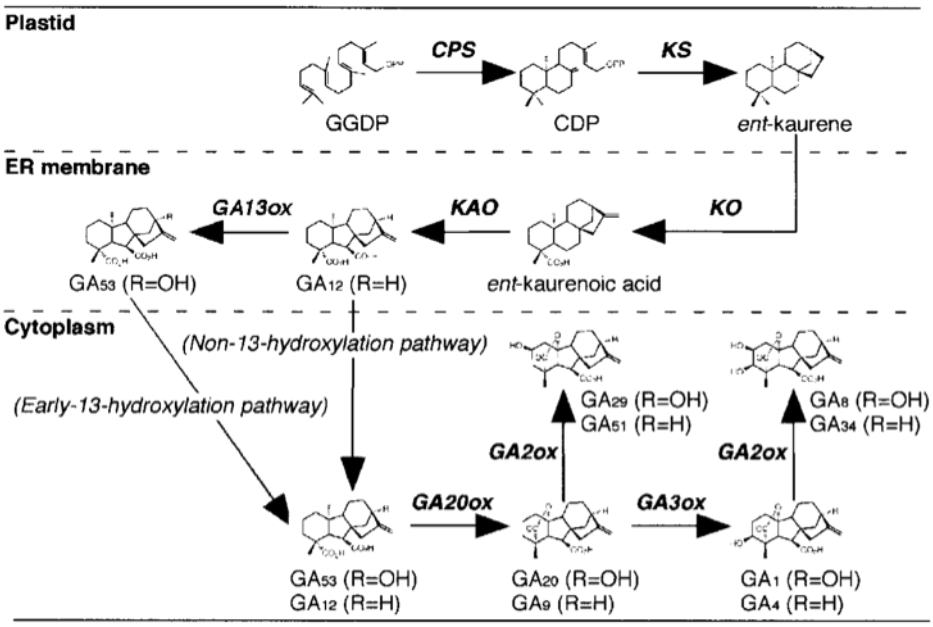


Figure 1-5 Principal pathway of GA metabolism in plants. (Sakamoto et al., 2004)

It has been also reported that GA regulates tapetum cell death and exine formation through a MYB transcriptional factor *GAMYB* in rice and the homologs *MYB33* and *MYB65* in Arabidopsis (Aya et al., 2009). Rice *GAMYB* is expressed in the anther, specifically in tapetum cells, during anther stage 7 to anther stage 10. It regulates the expression of GA dependent α-amylase and

is involved in the expression of almost all GA-regulated genes in rice anther (Aya et al., 2009). In the rice *gamyb-2* mutant, abnormal microspore mother cells and tapetum hypertrophy are observed (Wilson & Zhang, 2009); after anther stage 7, the microspores also do not adhere to the tapetum cells as in wild type (Aya et al., 2009). A regulatory model of GA regulating tapetum cell death and exine formation was further proposed by Aya *et al*. (Figure 1-6) (Aya et al., 2009). In this model, it was shown that *GAMYB* is activated by GA through the GID1/DELLA perception pathway. A number of protease and protease inhibitors, which may be involved in the cell death of the tapetum cells, are up-regulated by *GAMYB* and consequently contributes to tapetum PCD. *GAMYB* also activates the expression of the sporopollenin synthesis genes *CYP703A3* and *KAR*, and contributes to Ubisch body and pollen exine formation (Aya et al., 2009).

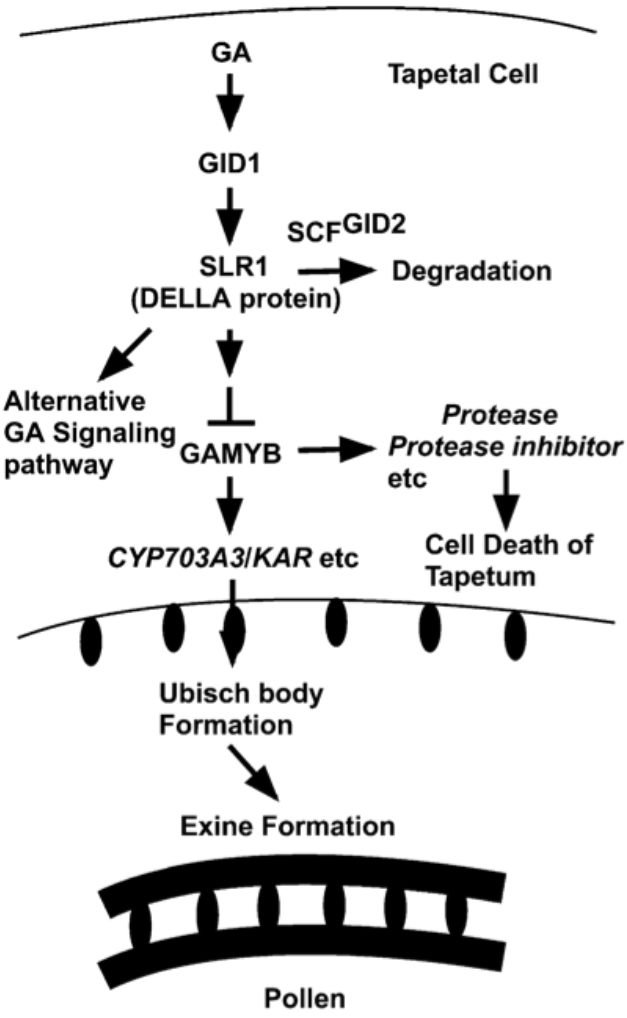


Figure 1-6 Model for GA Signalling in the Anther. (Aya et al., 2009)

1.2.2.2 Other phytohormones

Some other phytohormones are also believed to play roles in regulating PCD of tapetum cells. Ethylene may negatively regulate tapetum PCD initiation (Takada et al., 2006; Tzeng et al., 2009). In lily, when anthers were treated with 2,5-norbornadien (NBD), an inhibitor of ethylene activity, the tapetum cells become densely cytoplasmic as seen during the PCD process (Tzeng et al., 2009). What's more, when expressing an ethylene receptor gene, *Cm-ERS1/H70A,* in transgenic tobacco plants under an anther-specific promoter, a delay of tapetum cell death could be observed (Takada et al.,

2006).

Although the roles of other plant hormones in tapetum PCD remain unclear, there is evidence indicating their importance for this process. Expression of jasmonic acid (JA) synthesis genes were detected in rice anthers when tapetum PCD starts and JA has been shown to regulate leaf senescence by cooperating with ROS (Avanci et al., 2010; He et al., 2002; Hirano et al., 2008). Furthermore, analysis of plant hormone biosynthesis and the corresponding signalling factors in rice, shows that the amount of active abscisic acid (ABA) and cytokinin (CK) may increase in the tapetum at meiosis stage (when tapetum PCD starts) and decrease at the uninuclear microspore stage (when the tapetum undergoes degradation) (Hirano et al., 2008). The specific expression of ABA and CK signalling factors strongly indicates their role in tapetum PCD.

1.2.3 Ca^{2+} involvement in tapetum PCD

Calcium plays an important role in PCD. In plant, calcium-based signalling includes influx, efflux and decoding of Ca^{2+}. Three major protein families which contain the EF-hand domain for Ca^{2+} ion chelation are required for decoding. They are calcium-dependent protein kinases (CDPKs), calcineurin B-like (CBL)-interacting protein kinases (CIPKs) and calmodulin (CaM) and CaM-like proteins (CMLs) (Edel et al., 2017). Among them, CDPK contains the kinase domain, auto-inhibitory domain and four EF-hands; the CBL

contains only three EF-hands; the CaM and CML contain only four EF-hands. There are 5 CaMs and 33 CMLs in rice (Mohanta et al., 2017).

It is reported that *DEX1* is required for the primexine formation in Arabidopsis (Paxson-Sowders et al., 2001). The homologous gene in rice, *OsDEX1*, regulates Ca^{2+} homeostasis to regulate tapetum PCD, callose degradation and pollen wall formation, suggesting the coordination of tapetum PCD with callose degradation (Yu et al., 2016).

1.2.4 ROS plays a crucial role in tapetum PCD

ROS are highly reactive molecules including $O2^{\cdot -}$ (Superoxide radical), H_2O_2 (Hydrogen peroxide), $OH^{\cdot}$ (Perhydroxyl radical) and $^1O2$ (Singlet oxygen) which can damage the cells (Gill & Tuteja, 2010). ROS control many processes, such as growth, cell cycle, plant defence, stress responses and PCD, by influencing a large number of genes in corresponding pathways (Gill & Tuteja, 2010).

During anther development, a high level of ROS is accumulated in tapetum cells after stage 8, whereas the level drops down rapidly when tapetum PCD is completed during stage 10 and 11 (Hu et al., 2011). The accumulation of ROS is consistent with tapetum degradation process indicating that ROS may be one of the factors which promotes tapetum PCD (Zhang & Yang, 2014). *OsMADS3* in rice is a transcriptional factor that adjusts ROS homeostasis. It is expressed highly in the tapetum during late anther development (Zhang et al.,

2015). It plays a crucial role in late anther development and pollen formation by regulating ROS homeostasis (Hu et al., 2011). Although whether ROS acts as a signal in initiation of tapetum PCD is still unknown, some genes (*TDR (TAPETUM DEGENERATION RETARDATION)* (Li et al., 2006a), *CYP703A3* (Yang et al., 2014), *CYP704B2* (Li et al., 2010), *C6* (Zhang et al., 2010a)) which show defective tapetum PCD are largely up-regulated in *osmads3* mutant. These findings suggest that ROS play an important role in tapetum PCD as signal molecules.

1.2.5 Lipid metabolism is important for normal tapetum PCD and pollen wall formation

Lipids are thought to be involved in tapetum PCD. Loss of function of *OsGPAT3 (Glycerol-3-Phosphate Acyltransferase 3)* (Men et al., 2017), a lipid metabolism associated gene in rice, as well as *GPAT1 (Glycerol-3-phosphate Acyltransferase1)* in Arabidopsis (Zheng et al., 2003), both cause abnormal tapetum development and degradation. *PTC1 (PERSISTANT TAPETAL CELL1)* encodes a PHD-finger protein which is specifically expressed in tapetum cells and microspores during anther development stage 8 and 9 (Li et al., 2011a). The *ptc1* mutant shows normal anther development before the meiosis stage. However, after stage 9, when the tapetum cells start degradation in wild type, the mutant shows more vacuolated and less stained tapetum cells. Until anther stage 11, the abnormally enlarged tapetum could still be seen in

ptc1. 75 genes putatively related to lipid transport and metabolism and secondary metabolite biosynthesis, transport, and catabolism had expression changes in *ptc1*. These indicate that *PTC1* may controls the tapetum PCD process by influencing lipid metabolism (Li et al., 2011a). Many other rice mutants with tapetum PCD defect, such as *tdr* (Zhang et al., 2008), and *eat1* (Niu et al., 2013) show abnormal lipid metabolism and accumulation.

Figure 1-7 shows the structure of pollen wall and genes involved in pollen wall formation process (Shi et al., 2015); during pollen wall formation, sporopollenin is synthesised by the tapetum cells. Tapetum-expressed lipid synthesis genes *CYP704B2, CYP703A3* (Yang et al., 2014) and *DEFECTIVE POLLEN WALL (DPW)* (Xu et al., 2017) in rice and the corresponding orthologs in Arabidopsis (*CYP704B1* (Dobritsa et al., 2009b)*, CYP703A2* (Morant et al., 2007) and *MS2 (MALE STERILE2)* (Chen et al., 2011)) are required for the pollen wall formation. Rice *DPW (defective pollen wall)* and the probable ortholog in Arabidopsis *MS2* encode for a fatty acid reductase, which take palmitoyl-Acyl Carrier Protein to produce 1-hexadecanol. Both the *dpw* mutant in rice and *ms2* in Arabidopsis have defects in pollen wall formation (Chen et al., 2011; Shi et al., 2011).

In addition, the biochemical pathway involving *ACOS12, PKS1, TKPR1 (TETRAKETIDE ALPHA-PYRONE REDUCTASE 1)* and *TKPR2 (TETRAKETIDE ALPHA-PYRONE REDUCTASE 2)* is conserved in rice,

Arabidopsis, tobacco, *Hypericum perforatum* and canola (Jepson et al., 2014; Li et al., 2016; Qin et al., 2013; Yang et al., 2017). In Arabidopsis, the medium- to long- chain fatty acids are catalysed by ACOS5 into medium- to long- chain fatty-acyl-CoA (de Azevedo Souza et al., 2009) which can be condensed with malonyl-CoA by PKSA (POLYKETIDESYNTHASEA) and PKSB (POLYKETIDESYNTHASEB) into tri- and tetra-ketide α-pyrone (Kim et al., 2010a; Wang et al., 2013) and reduced by TKPR1 and TKPR2 (Grienenberger et al., 2010; Wang et al., 2013).

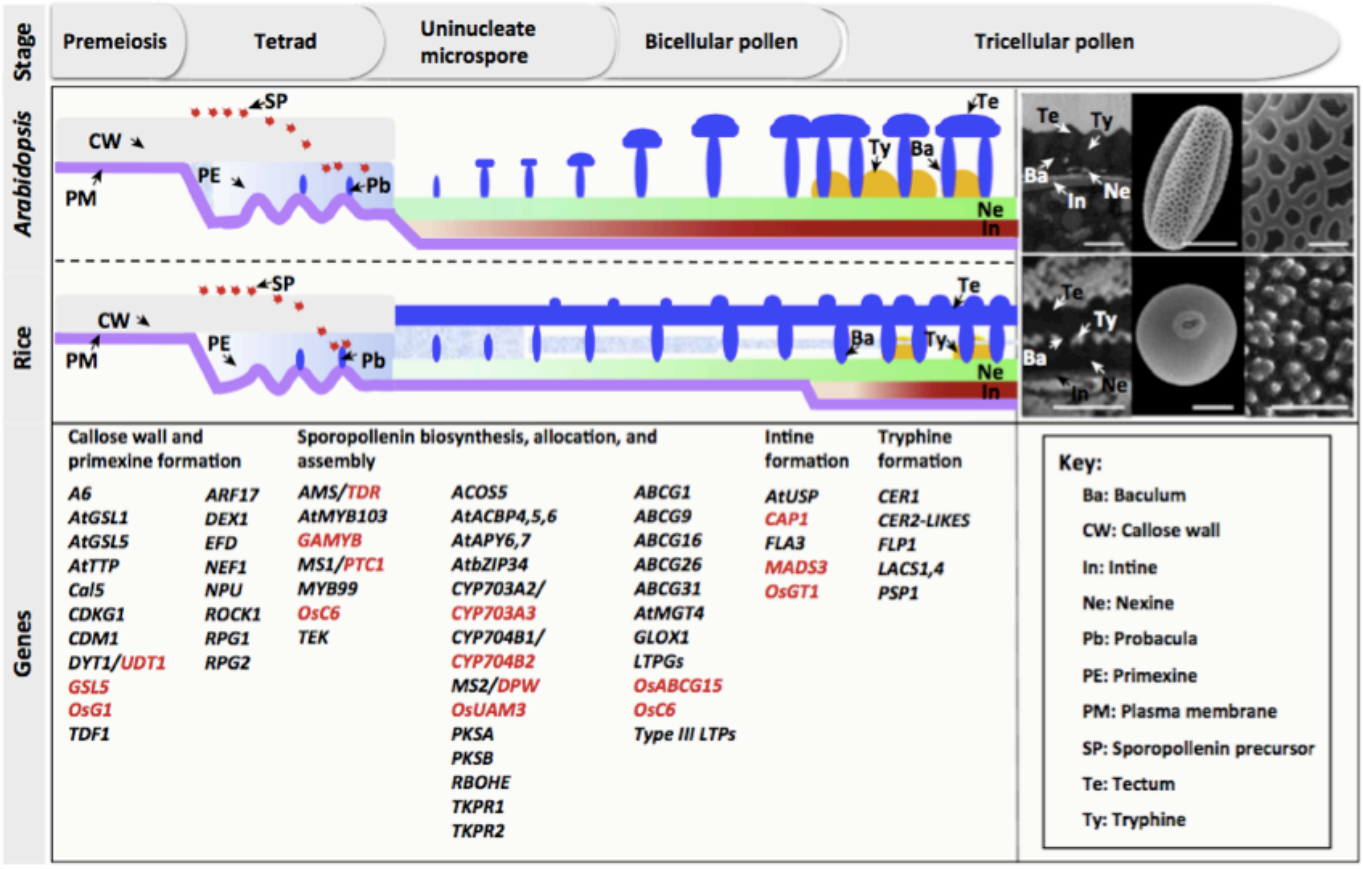


Figure 1-7 Comparison of Genes/Enzymes Involved in Various Stages of Pollen Wall Development in Arabidopsis and Rice. Pollen wall development-associated genes/enzymes identified in Arabidopsis and rice are indicated in black and red, respectively. (Shi et al., 2015).

Sporopollenin transport-related genes are also required for pollen wall formation. Two ATP Binding Cassette G Transporters, *OsABCG15* (*ATP Binding Cassette G Transporters*) and *OsABCG26* are involved in rice anther cuticle and pollen exine formation. OsABCG15 transport lipidic molecules from tapetum cells to anther locules for pollen wall formation, while OsABCG26 transport lipidic molecules from tapetum cells to anther wall layers for anther wall formation (Hu et al., 2010; Qin et al., 2013; Zhao et al., 2015). Their Arabidopsis orthologs *AtABCG11* (Panikashvili et al., 2010) and *WBC27/ABCG26* (Choi et al., 2011) are involved in transfer lipid precursors and polyketides from tapetum for pollen exine and anther wall formation.

1.3 *TIP2*-RELATED ANTHER DEVELOPMENT

TIP2 encodes for a bHLH transcription factor which normally forms homodimers or heterodimers to regulate the expression of its downstream genes (Feller et al., 2011; Fu et al., 2014; Murre et al., 1989). In plants, bHLH family genes are frequently involved in cell differentiation and division (Ranjan et al., 2017). It is widely reported that bHLH family participate in the stoma development both in rice and in Arabidopsis (Bergmann & Sack, 2007; Lampard & Bergmann, 2007). *SPCH* regulates the meristemoid mother cell to divide into meristemoids and stomatal-lineage ground cell by asymmetric cell division, and then the meristemoids divide into guard mother cell by asymmetric cell division, *MUTE* is involved in this process. After that *FAMA*

promotes the guard mother cell to divide into two guard cells. The bHLH family genes: *SPCH*, *MUTE*, *FAMA* work as a regulatory cascade to regulate stoma cell differentiation and division (Lampard & Bergmann, 2007).

Studies show that the bHLH proteins TIP2, TDR and EAT1 work together to play a central role by a similar way in regulating differentiation, morphogenesis, and degradation of anther somatic cell layers (Figure 1-8) (Fu et al., 2014; Niu et al., 2013). *TIP2* activates the expression of *TDR* and *EAT1*. It is involved in cell differentiation at early stage, and it has been shown that the presence of a TIP2-mediated transcription cascade regulates tapetum development in rice (Fu et al., 2014). The phenotype shows that *tip2* mutant has undifferentiated innermost three layers; the *tdr* mutant shows an abnormally expanded tapetum; *eat1* mutant has a persistent tapetum with delayed PCD. Therefore, it was concluded that *TIP2* promotes the specification and differentiation of anther wall cell; *TDR* plays crucial roles in tapetum differentiation and degradation; *EAT1* is required for tapetum PCD initiation. *TDR* in rice and its' homologs *AMS (ABORTED MICROSPORES)* in Arabidopsis have been reported as also involved in tapetum PCD (Sorensen et al., 2003). *AMS* regulates a series of genes required for sporopollenin synthesis and transport (Xu et al., 2014; Xu et al., 2010), *TDR* induces a Cys protease *OsCP1* to regulate tapetum PCD (Lee et al., 2004; Li et al., 2006a) and a series of lipid metabolism genes to regulate pollen wall formation (Shi et al., 2015). *OsCP1* is also regulated by two ATP-dependent RNA helicases,

AIP1 and *AIP2* (Li et al., 2011b). *EAT1* is a positive regulator in rice (Niu et al., 2013). EAT1 interacts with TDR and TIP2, to regulate two Asp proteases AP25 and AP37, controlling the tapetum PCD (Fu et al., 2014; Ko et al., 2014; Niu et al., 2013). This is illustrated in figure1-7.

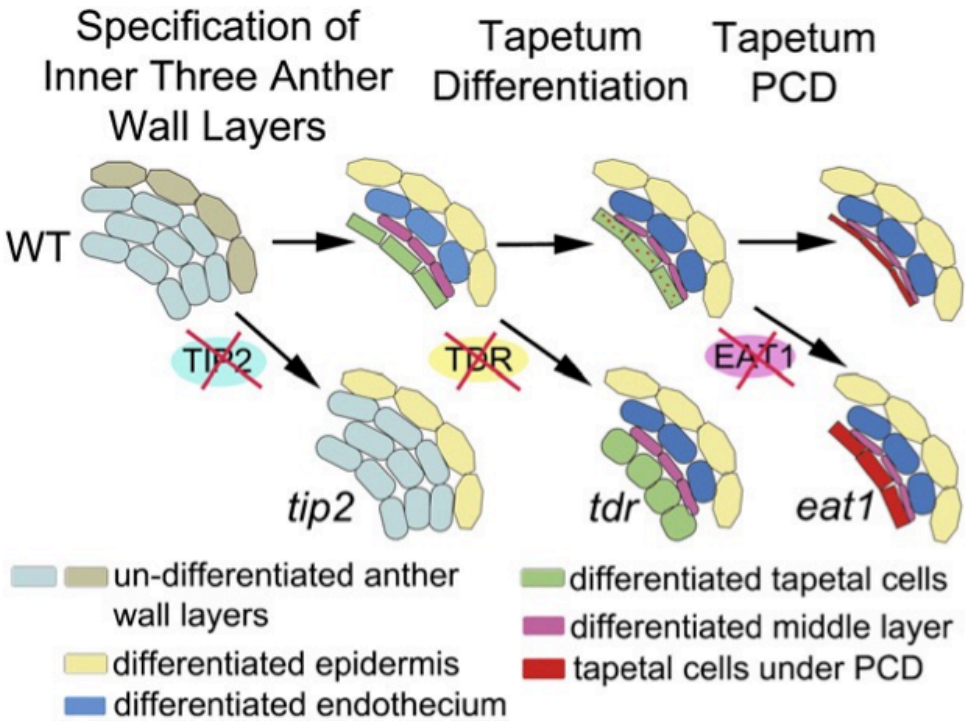


Figure 1-8 Proposed Model for the TIP2-TDR-EAT1 Cascade That Regulates Anther Wall Cell Fate and Microspore Development in Rice. (Fu et al., 2014)

1.4 GENE CO-EXPRESSION ANALYSIS

To gain a better understanding of gene networks, gene co-expression analysis has emerged as a powerful tool for gene function prediction (Usadel et al., 2009). Large numbers of gene expression datasets are publicly available on the Gene Expression Omnibus (GEO) (Barrett et al., 2013) and ArrayExpress (Kolesnikov et al., 2015) thus providing opportunities for co-expression network analysis and gene function analysis. Different clustering methods can

be used to identify co-expressed genes based on gene expression profiles (Kafieh & Mehridehnavi, 2013); genes with similar expression patterns may take part in the same pathway or have similar functions. Therefore, co-expression analysis can provide reliable characterization of gene functions especially for currently unknown genes (Usadel et al., 2009).

Note that gene co-expression network analysis is the same as gene regulation network analysis. Using gene expression data it is possible to identify co-expressed genes on a large-scale, thus those genes that have the potential to be regulated by the same transcription factor (TF), or to co-regulate the same pathway can be characterised (Bar-Joseph et al., 2003). However, genes in a regulation network usually have more complex gene expression patterns because TFs might work together simultaneously or sequentially, as well as a single gene might be regulated by multiple factors (Awad & Chen, 2014).

1.4.1 Co-expression analysis process

There are 4 steps involved in general co-expression analysis.

First of all is selection of data. There are 4 types of data that can be used based on the different objectives. Condition-independent data which contain microarray data from as many different conditions as possible, could show general gene relationships; Condition-dependent data are sets from particular samples or conditions, to view gene relationships based on the condition that one is interested in; combined data combines data from some specified

samples or conditions which is a better way to control the noise in microarray data; personal data is what you get from your own experiments for your own experimental design rather than from public database (Usadel et al., 2009). To reduce false positives in the co-expression network, sufficient microarray data is needed. It has been shown, based on yeast data, that the accuracy of gene associations increases up to 50 experiments, after which point the level of accuracy remains constant between 50 to 100 experiments(Yeung, 2004). Therefore, using data of over 50 microarray experiments is recommended in clustering analysis.

Normalization is a crucial step to make raw data into comparable data. There are some popular methods for normalization. MAS5 (Affymetrix Microarray Suite version 5) (Hubbell, et al., 2002). MAS5 algorithm is a part of the MAS 5.0 software package and can be easily implied. It works well in reducing false positives without losing many true positives (McClintick & Edenberg, 2006); RMA (Robust Multi-array Average) (Irizarry et al., 2003a) is a popular way for microarray data normalization because its outstanding performance in identifying differentially expressed genes; GC-RMA (GeneChip-RMA) (Wu, et al., 2004) is different from RMA only in the background adjustment by using probe sequence information to help the background estimation (Wettenhall, et al., 2006).

After normalization, PCC (Pearson's correlation coefficient) and SCC (Spearman's correlation coefficient) are generally used to identify the

similarity of expression patterns in the next step. PCC is widely used in measuring gene co-expression and is more sensitive to outliers (Obayashi & Kinoshita, 2009) while SCC is more robust to outliers (Obayashi & Kinoshita, 2009).

The last step is generating a network for global gene interactions. Genes in the same pathway, or with similar functions usually show very similar expression pattern in a variety of biological processes and experimental conditions (Ruan, Dean, & Zhang, 2010). As a result, co-expression networks which are generated based upon correlation can provide hypothetical gene interactions on a global scale. As with other networks, a co-expression network has a structure with nodes (gene) and links (correlation) as the basic elements. The nodes linked more densely make up the sub structures such as motifs and modules. Figure 1-9 defines the common terms describing networks (Aoki et al., 2007).

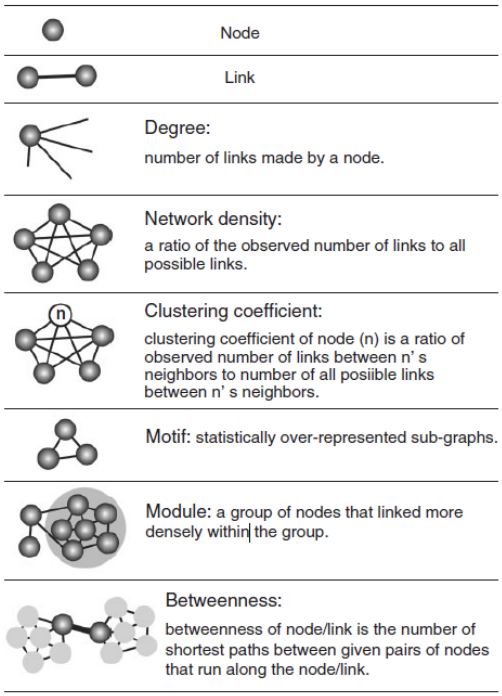


Figure 1-9 Definitions of terms describing networks (Aoki et al., 2007).

There are two kinds of strategies to generate a co-expression network, called "guide-gene" approach and non-targeted approach (Aoki et al., 2007). For "guide-gene" approach, a set of genes which the researchers are interested in or are related to some particular biological problems are selected in advance of gene correlation coefficient analysis. Then in co-expression analysis, correlation coefficients between the guide genes are extracted to show co-expression modules within the guide genes. That could also be done between guide genes and other genes. It is a good strategy to define directly which genes might be co-expressed to the target gene. The other strategy is a non-targeted approach. In this approach, the correlation coefficient of any two expressed genes is calculated and all the correlation coefficients over the threshold are used to generate the network. It is a more efficient way to show

the global interactions and predict new genes, however needs more experimental verification to confirm the relevant interactions. There are various clustering algorithms to gather genes with similar expression patterns together, for example, hierarchical average-link, hierarchical complete-link, MCLUST and IMM (Yeung, 2004).

To get a better understanding of the co-expression network, biological knowledge such as gene annotation and GO pathway should be used to interpret results (Aoki et al., 2007). Experimental verification should also be conducted to test and confirm the hypothetic biologically relevant derived from the analysis (Aoki et al., 2007).

1.4.2 Current progress in co-expression analysis for anther development

In Arabidopsis, genome-wide co-expression analysis has been used to identify 23 downstream genes directly regulated by AMS in pollen wall formation (Xu et al., 2014). Furthermore, co-expression networks can be generated based on correlations to provide hypothetical gene interactions on a global scale. Based on this approach, a network called FlowerNet has been generated in Arabidopsis to identify gene regulation networks in Arabidopsis flowers and reliably predict candidate genes associated with anther and pollen development (Pearce et al., 2015). Nevertheless, compared with Arabidopsis, fewer public co-expression network platforms are available for rice. RiceFREND is a widely used rice co-expression platform that normalises 815 microarrays and

provides a chain-like network for co-expression gene query based on “single guide gene search” and “multiple guide gene search” (Sato et al., 2013). OryzaExpress includes data from 871 microarrays and integrates Omics-Data from public databases to provide co-expression information over whole datasets (Hamada et al., 2011). However, neither of these tools allows targeting of the co-expression analysis to specific conditions or developmental stages, such as anther development, which can mean that the networks generated can be biased by highly expressed genes that are not associated with the developmental processes under analysis. This function was previously available on the Rice Oligonucleotide Array Database (ROAD) (Cao et al., 2012), which provided co-expression network analysis on samples under biotic stress and abiotic stress separately, but this tool is now no longer supported and is unavailable.

Previously a co-expression network in rice anthers was constructed by using 33 laser microdissection-microarray datasets of anthers (Aya et al., 2011). Aya et al. (2011) predicted many putative genes involved in meiosis and pollen wall synthesis by taking genes (nodes) within 2 steps from a bait gene in their meiosis and pollen wall synthesis co-expression network. This work showed that this reverse genetics approach could be used to identify novel genes. However, the performance of this supervised network relies heavily on the chosen bait genes and the double-step approach often gives sets of genes which can have very different expression patterns (Wittkop et al., 2010).

Although many genes required for viable pollen formation have been identified, the molecular mechanism regulating of pollen development remains elusive. It has been reported using a 10-k rice cDNA microarray that over 2000 genes (~22% of the genes on the array) were differentially expressed between anthers and seedlings, which indicates the complex requirements for anther-specific gene regulation for anther and pollen development (Wang et al., 2005). An un-targeted co-expression network in rice anther is needed to determine gene regulation relationships globally during anther development.

1.5 AIMS OF THE PROJECT

In plants, anther and pollen development involves many important biological events and complex molecular regulatory coordination. Understanding the gene regulatory relationships during reproductive development is essential for fundamental biology and crop breeding. The aim of this project is to predict novel genes involved in pollen wall formation by using reverse genesis approaches. In this work, we have developed a rice gene co-expression network (RiceAntherNet) that allows prediction of gene regulatory relationships during anther and pollen development. Furthermore, downstream targets of TIP2 were also predicted to study the function of TIP2 in cell division and differentiation during anther development.

In chapter 3, a rice co-expression network (RiceAntherNet) was generated to provide a global view of gene expression patterns throughout rice anther

development. 43 novel genes were predicted to be involved in pollen wall formation using mutants defective in pollen development.

In chapter 4, 7 genes which were predicted by RiceAntherNet were knocked out by CRISPR-Cas9 system and mutants of 4 genes were analysed for their phenotypes.

In chapter 5, 6 Lipid Transfer Protein precursors (LTP) were predicted as TIP2 downstream genes and analysed for their phenotypes.

2 MATERIALS AND METHODS

2.1 MATERIALS

2.1.1 Plant materials

Wild-type Arabidopsis used in this research is *Arabidopsis thaliana* ecotype Columbia. T-DNA insertion mutants of Arabidopsis were obtained from the Nottingham Arabidopsis Stock Centre (NASC). Arabidopsis plants were grown in a growth room at 22°C and 16hrs photoperiod.

Wild-type rice used in this research includes japonica rice 9522 and japonica rice DongJin. T-DNA insertion mutants of rice were obtained from Pohang University of science and technology, Korea. Rice plants were grown in the paddy field in Shanghai in the summer of 2016 and 2017.

Nicotiana benthamiana (N. benthamiana) was used for *Agrobacterium*-mediated transient expression. Plants were grown in a growth room at 22°C and 16hrs photoperiod.

2.1.2 Bacterial strains and vectors

Escherichia coli (*E. coli*) strain Dh5α and *Agrobacterium tumefaciens (A. tumefaciens)* strain EHA105 and GV3101 were obtained from Tiangen biotech (Beijing) co. ltd. The application and antibiotic type of these strains are list in Table 2-1.

Table 2-1 Bacterial strains

Strain	Application	Antibiotic selection (μg/ml)
DH5α	Routine cloning	
EHA105	Routine cloning	Rifampicin (50)
GV3101	Transient expression in *N. benthamiana*	Rifampicin (50), gentamycin (25)

pBin-sgR-Cas9-p1300 (CH) vector was used for targeted gene editing by the CRISPR/Cas9 system (Zhang et al., 2014). pGreen0800 and pGreen0229 were used for analyzing promoter sequences regulation of gene expression by the Dual-Luciferase Reporter Assay (Hellens et al., 2005). pSoup was used as a helper vector which helps the replication of other binary plasmid in the Agrobacterium cell (Hellens et al., 2005). The details of these vectors were listed in Table 2-2; the vector maps were shown in Appendix I.

Table 2-2 Vectors

Vector	Application	Antibiotic selection (μg/ml)
pBin-sgR-Cas9-p1300 (CH)	CRISPR/Cas9 system	Kanamycin (50)
pGreen0800	Transient expression for quantification of promoter activity	Kanamycin (50)
pGreen0229	Dual-Luciferase Reporter Assay	Kanamycin (50)
pSoup	Helper vector for ensuring high levels of transient gene expression	Tetracycline (50)

2.1.3 Culture medium, buffers and chemicals

1. 1.5% CTAB: 75 mM Tric-HCL (pH = 8.0), 15 mM EDTA (pH8.0), 1.05 M NaCl

2. 0.5X TBE buffer: 45 mM Tris-borate, 1 mM EDTA

3.LB medium (1L): yeast extract 5g, trypsin peptone 10g, NaCl 10g, pH = 7.2, agar powder 15g (for Solid medium)

4. NBD2 medium (1L): 9 stock solutions (according to the table below), sucrose 30g, inositol 0.1g, L-Proline 0.5g, L-glutamine 0.5g, Casein hydrolysate 0.5g, pH = 5.8, agar powder 7g

Stock solution preparation (required in 1L NBD2 medium):

Stock solution (1L)	Chemicals		Volume (ml)
1. 20× N6 macroelement 1	KNO_3	56.6g	50
	$(NH_4)_2SO_4$	9.26g	
	KH_2PO_4	8g	
2. 40× N6 macroelement 2	$CaCl_2 \cdot 2H_2O$	6.64g	25
3. 40× N6 macroelement 3	$MgSO_4 \cdot 7H_2O$	7.4g	25
4. 100× B5 microelement 1	$MnSO_4 \cdot H_2O$	781mg	10
	$ZnSO_4 \cdot 7H_2O$	200mg	
	H_3BO_3	300mg	
	KI	75mg	
5. 1000× B5 microelement 2	$Na_2MoO_4 \cdot 2H_2O$	250mg	1
	$CuSO_4 \cdot 5H_2O$	25mg	
	$CoCl_2 \cdot 6H_2O$	25mg	
6. 100× B5 vitamin 1	Aneurine hydrochloride (VB_1)	1000mg	10
	Pyridoxine hydrochloride (VB_6)	100mg	
	Niacin (VB_3)	100mg	
7. 100× B5 vitamin 2	Glycine	200mg	10
8. 100× 2,4 D	2,4 D	200mg	10
9. 200× Fe^{2+}	Na_2-EDTA	7460mg	5
	$FeSO_4 \cdot 7H_2O$	5560mg	

5. NBD2-AS (1L): 9 stock solutions (the same as NBD2 medium), sucrose 30g,

inositol 0.1g, Casein hydrolysate 0.5g, glucose 10g, pH = 5.2

6. AAM-AS medium (1L): 6 sother solutions (according to the table below), KCl 2.94g, sucrose 68.5g, inositol 0.1g, casein hydrolysate 0.5g, glucose 36g, pH = 5.2

Stock solution preparation (required in 1L AAM-AS medium):

Stock solution (1L)	Chemicals		Volume (ml)
1. 10× AA macroelement	KH_2PO_4	1700mg	100
	$MgSO_4 \cdot 7H_2O$	3700mg	
	$CaCl_2 \cdot 2H_2O$	4400mg	
2. 100× AA microelement 1	$MnSO_4 \cdot H_2O$	1690mg	10
	$ZnSO_4 \cdot 7H_2O$	860mg	
	H_3BO_3	620mg	
	KI	83mg	
3. 1000× AA microelement 2	$Na_2MoO_4 \cdot 2H_2O$	250mg	1
	$CuSO_4 \cdot 5H_2O$	25mg	
	$CoCl_2 \cdot 6H_2O$	25mg	
4. 10× AA vitamin 1	Glutamine	8770mg	10
	Aspartic acid	2660mg	
	Arginine	2880mg	
	Glycine	750mg	
5. 200× Fe^{2+}	Na_2-EDTA	7460mg	5
	$FeSO_4 \cdot 7H_2O$	5560mg	
6. 100× MS vitamin	Aneurine hydrochloride (VB_1)	1000mg	10
	Pyridoxine hydrochloride (VB_6)	100mg	
	Niacin (VB_3)	100mg	

7. MS-H medium (1 L): MS 4.4 g, sucrose 30 g, Casein hydrolysate 0.5 g, sorbitol 15 g, NAA (0.5 mg/L) 2 ml, 6-BA (2 mg/L) 4 ml, pH = 5.8, agar powder 4 g

8. MSNH rooting medium (1 L): MS 4.4 g, sucrose 30 g, agar powder 2.5 g, pH=5.8

9.FAA fixing solution (100 ml): Ethanol 50ml, Acetic acid 5ml, formaldehyde (35%) 10 ml, and ddH_2O 35 ml

2.2 MICROARRAY DATA ANALYSIS

2.2.1 Quality control (QC)

QC was performed with the R statistical analysis software (Gautier et al., 2004) using the package 'affyPLM'. Fluorescence detection was used for checking the fluorescence intensities for each probe in each sample. Random distribution of fluorescence intensities excludes the pollution of array and other systematic errors. RNA degradation analysis was used to assess RNA quality of each array. A gentle decrease of probe degradation was expected from 5'-end to 3'-end.

2.2.2 RMA normalization

RMA normalization (Irizarry et al., 2003) was conducted using the package 'affy' (Gautier et al., 2004) in the statistical programming language R. The analysis was performed using a custom chip definition file (CDF)

(riceosrefseqcdf_17.1.0) from the CustomCDF project (http://brainarray.mbni.med.umich.edu/) (Dai et al., 2005). This CDF remaps the individual probes to probe sets by using the most recent sequencing information available from RefSeq (Pruitt et al., 2005), resulting in 24,462 probe sets. This process eliminates the many-many relationship between probe sets and genes, and means that the genes may be used as the primary identifier for the generation of the correlation network.

2.2.3 Pearson correlation coefficient (PCC)

Gene coexpression relationships were measured by PCC according to the following formula:

$$r = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{Xi-\overline{X}}{Sx}\right)\left(\frac{Yi-\overline{Y}}{Sy}\right)$$

2.2.4 Gene clustering

Genes and correlations were organized as a table of nodes (genes) and edges (correlations), which were explored in Cytoscape 3.4.0 (Shannon et al., 2003; Smoot et al., 2011). The organic layout was used to display the global structure of the network. The plugin ClusterMaker (Morris et al., 2011) was used to divide the network into distinct clusters, using the Transitivity Clustering method (TransClust) (Wittkop et al., 2010) with parameters Max Subcluster Size = 400 and Max Time = 10, with the PCC values as edge weights. This method generates small, extremely well-connected clusters

where nearly all the edges are present, meaning that cluster members must have very similar expression patterns.

2.2.5 Gene annotation

Rice genome annotation file (IRGSP-1.0_representative_2016-08-05) and Arabidopsis annotation file (gene_description_20130831) were downloaded from the Rice Annotation Project Database (http://rapdb.dna.affrc.go.jp) (Sakai et al., 2013) and the Arabidopsis Information Resource (TAIR, http://arabidopsis.org) (Rhee et al., 2003a) respectively.

2.2.6 Gene ontology (GO) analysis

The orthologs between rice and Arabidopsis were predicted by blast analysis. GO analysis of rice genes were annotated according to the GO terms of orthologs in Arabidopsis.

2.2.7 Phylogenetic analysis

Protein sequences (see Appendix IV) were used to construct a Neighbor-Joining tree using Mega7 (Kumar et al., 2016). Alignment was performed on ClustalW (Pairwise Alignment: Gap Opening Penalty = 10, Gap Extension Penalty = 0.1; Multiple Alignment: Gap Opening Penalty = 10, Gap Extension Penalty = 0.2; Protein Weight Matrix = Gonnet; Gap Separation Distance = 4; Delay Divergent Cutoff = 30%). Phylogenetic tree was construct by Neighbor-Joining (Bootstrap replications = 1000).

2.3 GENOTYPING

2.3.1 DNA extraction (by CTAB method)

Fresh rice leaves (about 1g) were placed in a 2 ml Eppendorf tubes. The tube lids were left open and placed in to a drying oven (65°C) overnight to dry the tissue. Then 2 magnetic beads were added to, break the plant tissue using a Mixer Mill (MM400 Retsch, Germany) at speed 60 Hz for 3 mins.

700 µL 1.5X CTAB was added to the smashed tissue, and incubated at 60°C for 30 min. (shaking it gently once or twice to mix it thoroughly). Then 700 µL Chloroform (1:1, equal volume) was added and the tube mixed thoroughly immediately and then centrifuged at 12,000 rpm (13,000xg) for 10 min. The supernatant was placed into a new 1.5 ml Eppendorf tube and an equal volume of isoamyl alcohol added, and mixed thoroughly several times. White flocculus was often observed at this step. Samples were stored at -20°C for at least 1 hour to fully precipitate the DNA.

Samples were centrifuged at 1,2000 rpm (13,000xg) for 10 min, the precipitated DNA was then washed in 500 µL 70% (v/v) ethyl alcohol and centrifuged at 1,2000 rpm (13,000xg) for 1 min. This wash and centrifugation was repeated a further time, the ethyl alcohol removed and the pellet left to dry in the open tube. The DNA was then dissolved in 50µL ddH2O. The quality and concentration of DNA were detected using a NanoDrop Microvolume

Spectrophotometer (NanoDrop 2000C, Thermo Scientific, Shanghai). DNA samples were stored at -20°C.

2.3.2 PCR

Taq DNA polymerase (Vazyme Biotech) and rTaq DNA polymerase (Takara) were used for PCR. The PCR reaction composition and running program were as shown in Table 2-3 and Table 2-4.

Table 2-3 General reaction composition

Component	Volume (μl)
10 X Taq buffer	2
10 X GC buffer II	2
dNTPs (2 mM)	2
Forward primer (10 pmol/μl)	1
Reverse primer (10 pmol/μl)	1
Taq polymerase	0.2
Template DNA	100ng
ddH_2O	To a final volume of 20 μl

Table 2-4 Standard PCR program

Step	Temperature (°C)	Duration (min:sec)	Cycle
Initial denaturing	95	5:00	1
Denaturing	95	0:30	35
Annealing	55-58	0:30	
Extension	72	1 minute/kb	
Final extension	72	10:00	1
Hold	4		

2.3.3 Agarose gel electrophoresis

1-1.5% (w/v) agarose gel (1-1.5 g agarose dissolved in 100 ml 0.5x TBE

buffer) was made for separating PCR products. The concentration of agarose gel depended on the size of the PCR product. 6X Gel Loading Dye (NEB) was added to the PCR product and it was loaded along with a molecular weight marker onto the agarose gel, which was run at an electric field strength of 10 V/cm in 0.5x TBE buffer, until DNA fragments were sufficiently separated. DNA was visualized under long wave UV light, and the gel images were captured using the ChemiDoc MP Imaging System (Bio-Rad, USA).

2.4 MOLECULAR CLONING METHODS

In-Fusion cloning PCR primers were designed with 15bp which matched the ends of the linearized vector added to both the 3' and 5' regions. The DNA fragments were amplified using GXL DNA polymerase (Jiahao Biotech, Shanghai) according to the reaction composition in Table 2-5 and PCR program in Table 2-6.

Table 2-5 Reaction composition for GXL polymerase

Component	Volume (µl)
5 X GXL buffer	10
dNTPs (2.5 mM)	5
Forward primer (10 pmol/µl)	1
Reverse primer (10 pmol/µl)	1
GXL polymerase	1
Template DNA	200 ng
ddH_2O	To a final volume of 50 µl

Table 2-6 PCR program for GXL polymerase

Step	Temperature (°C)	Duration (min:sec)	Cycle
Denaturing	98	0:10	30
Annealing	55	0:15	
Extension	68	1 minute/kb	
Final extension	68	7:00	1
Hold	4		

The PCR products were checked by electrophoresis on 1.5% agarose (w/v) gel and correct size DNA fragments were recovered using a Gel Extraction Kit (Xieruicheng Biotech, Shanghai).

The DNA fragments were then combined into the vector using an In-fusion HD cloning kit (Jiahao Biotech, Shanghai) according to Table 2-7.

Table 2-7 In-fusion HD cloning reaction composition

Component	Volume (ul)
Vector	400 ng
Insertion DNA fragment	100 ng
In-fusion	2 µl
ddH_2O	To a final volume of 10 µl

The reaction was incubated at 50°C for 20 mins.

2.4.1 *E. coli* transformation by heat-shock

A tube with 50 µl of competent *E. coli* cells was thawed on ice, as soon as the cells had thawed, about 200 ng of plasmid DNA was added to the cells and mixed gently and incubated on ice for 30 mins. Then the cells were heat-shocked by moving the tube to a pre-heated 42°C water bath for exactly

90s. It was then immediately placed onto ice immediately after the heat-shock for at least 1 min.

200µl LB medium was then added to the tube, which was incubated on a shaker at 180 rpm for 1 hour. An aliquot (100 µl) was then spread onto the surface of a prepared LB agar plate (with required antibiotics). The plates were then incubated at 37°C overnight. The *E. coli* colonies could then be seen with round, white, smooth features.

2.4.2 Plasmid DNA extraction and sequencing

Plasmid DNA was extracted using a Mini Plasmid Kit (Tiangen Biotech, Beijing). 10 µl plasmid DNA with a concentration of 100 ng/µl was sent for sequencing. The primers for sequencing are listed in section 2.1.2.

2.4.3 *Agrobacterium* (EHA 105 and GV3101) transformation by heat-shock

Tubes with 50 µl of competent *Agrobacterium* were thawed on ice. As soon as the cells had thawed, about 200 ng of plasmid DNA was added to the cells and mixed gently, which was incubated on ice for 5 mins. Samples were then put it into nitrogen liquid for 5 mins, followed by incubation at 37°C for 5 mins.

An aliquot of 200 µl LB medium was added to each tube and incubated on a shaker at 28°C, 200 rpm for 2 hours to activate the *Agrobacterium* cells. 100µl was then spread onto the surface of a prepared LB agar plate (with required antibiotics). The plates were allowed to dry and incubated at 28°C for 2 to 3

days. The *Agrobacterium* colonies could be seen to have round, yellow, smooth features.

2.4.4 Storage of bacteria as glycerol stocks

A single colony was transferred to 1 ml of LB liquid media and incubated overnight (*E. coli* at 37°C, 180 rpm; *Agrobacterium* at 28 °C, 200 rpm). The culture was then mixed to a final concentration of 20% (v/v) glycerol and flash frozen in liquid nitrogen. The glycerol stock was kept at -80°C.

2.5 TARGETED MUTAGENESIS BY CRISPR-CAS9 SYSTEM

2.5.1 Selection of CRISPR targets

CRISPR targets in the form of N20 (20 bp DNA fragment)-NGG (PAM) from 5' to 3' end were first selected by CRISPR-P 2.0 (http://crispr.hzau.edu.cn/CRISPR2), a CRISPR/Cas9 tool for genome editing in plants.

The final targets were determined by the following rules: lower off-target score (sequence specificity), targeting at exon or 5'-UTR region; avoiding Pol III promoter terminators (avoiding spacers with poly T); GC content (40%-80%); RNA fold analysis (unafold.rna.albany.edu).

The first base of N20 was changed to G, if it was not this already a G, and the overhangs added to each pair of primers as below.

Upper Primer: 5'-TGGC-G(N19)-3'

Lower Primer: 5'-AAAC-(N19)C-3'

All the primers designed for amplifying CRISPR targets are listed in Appendix II.

2.5.2 Plasmid construction

Annealing was carried out by a slow cool-down process (37°C for 30 mins, 95°C for 5 mins, 95 °C ramp down to 25°C at 5°C/min) with the annealing reaction compositions as in Table 2-8.

Table 2-8 Annealing reaction composition

Component	Volume (µl)
10 X T4 PNK buffer (NEB)	1
Forward primer (10 pmol/µl)	1
Reverse primer (10 pmol/µl)	1
T4 Polynucleotide Kinase (T4 PNK, NEB)	0.5
ddH_2O	To a final volume of 10 µl

pBin-sgR-Cas9-p1300 (CH) vector was digested by *Bsa*I according to Table 2-9. The mixture was then incubated at 37°C for 2 hours, then keep at 65°C for 5 mins to inactivate the *Bsa*I enzyme.

Table 2-9 Restriction Enzyme Digestion composition

Component	Volume (µl)
CH	1 µg
*Bsa*I	10U
Cutsmart buffer	5
ddH_2O	To a final volume of 50 µl

The annealed product and digested vector was then incubated with T4 ligase and the components shown in Table 2-10 at 22°C for 30 mins.

Table 2-10 Ligation reaction composition

Component	Volume (μl)
Annealing product	1
*Bsa*I digested CH	1
10X T4 buffer	1
T4 ligase	0.5
ddH_2O	To a final volume of 10 μl

sgRNA-Cas9 plant expression vectors were transferred into *Agrobacterium* strain EHA105 as described in 2.4.

2.5.3 Callus tissues formation

2.5.3.1 Callus tissues of 9522 derived from mature seed

Mature seeds of 9522 were dried in the sun for 3 days and dehusked. Seeds were sterilised in 70% (v/v) ethanol for 1 min and then the ethanol discarded. The seeds were then placed in 33% (w/v) sodium hypochlorite solution (with one drop of Tween 80) and shaken for 15 mins. The liquid was then discarded and the seeds washed in sterile ddH_2O, at least 10 times. The sterilized seeds were then dried on sterile filter paper and then incubated on the surface of NBD2 medium with about 50 seeds per Petri dish.

After incubating at 28°C in the dark for about 10 days, small light yellow calli were formed. The calli were removed from the roots and endosperm and moved on to fresh NBD2 medium and then incubated at 28°C in the dark for a further 10 days. Calli were then moved onto fresh NBD2 medium for 5 days to keep the callus in good condition before transformation.

2.5.3.2 Callus tissues of 9522 derived from young panicles

Young tillers of 9522 at anther stages 2 to stage 7 were collected. The tillers were cleaned with 70% (v/v) alcohol and the covering leaves were peeled away. The panicle was then clipped with sterilized tweezers and cut it into 2 mm-wide pieces, which were placed onto the surface of NBD2 medium.

After incubating at 28°C in the dark for about 20 days, small calli were formed. These were moved on to fresh NBD2 medium, for a further 10 days, and then onto fresh plates for a further 5 days before transformation.

2.5.4 Transformation of calli

EHA105 with the target construct was placed into 2 ml LB liquid medium (selective medium with 50 μg/ml Kan and 50 μg/ml Rif) and incubated (28°C at 200 rpm) overnight. 1 ml of the bacteria solution was added to 50 ml LB liquid medium (selective medium with 50 μg/ml Kan and 50 μg/ml Rif) and incubated on a shaker at 28°C at 180 rpm until it reached an OD600=0.3 (about 4 hours). It was then centrifuged at 5,000 rpm (4,000xg) for 10 mins and the bacterial pellet resuspended in AAM-AS fluid medium (final AS concentration of 200 μM).

100 good calli were then placed into a sterile Erlenmeyer flask and the bacteria solution added and shaken gently for 30 mins. The calli were then drained and dried on sterile filter paper. The partially dried calli were then placed onto NBD2-AS medium (with a final AS concentration of 100 μM) with a piece of

filter paper under the calli and incubated for 3 days at 28°C in the dark.

2.5.5 Selection of transformed calli and plant growth

The infected calli were placed onto sterile filter paper, and dried for 1 hour, then transferred onto NBD2 medium (selective medium with 40 mg/L of hygromycin and 400 mg/L of timentin). They were incubated for 12 days at 28°C in the dark, and then moved onto fresh NBD2 selective medium under the same conditions for 12 days. It could be observed that some calluses blackened while some new calli were formed.

Healthy calli were transferred onto MS-H callus induction and differentiation medium (with 20 mg/L of hygromycin and 200 mg/L of timentin) at 28°C in the light for 10 days. Only transformed calluses survived on the selection medium, whilst untransformed calli died at this step.

The surviving calli were transferred to new MS-H medium (with 20 mg/L of hygromycin and 200 mg/L of timentin) every 10 days. Green spots were formed followed by formation of a small plant with a root and shoot. The small plants were then moved to MSNH rooting medium in a 15 cm-high jar and grown in the light. After about 15 days, when the plants were bigger with strong roots and up to 15 cm height, they were ready for transfer into the paddy field.

2.5.6 Genotyping

DNA of each plant was extracted by the CTAB method according to section

2.3.1. Primers were designed to cover the target regions (Appendix II). PCR amplification followed by sequencing was performed to check the mutation of the target regions. Sequencing results were analysed by SnapGene (http://www.snapgene.com) to confirm the mutation sites of each mutant.

2.6 EXPRESSION ANALYSIS

2.6.1 Total RNA extraction by the trizol method

Total RNA was extracted using TRIzol Reagent (Thermo Fisher Scientific, China) according to the manufacturers guide (Pub. No. MAN0001271). Purified RNA was stored at -80°C.

2.6.2 cDNA synthesis

The concentration of RNA was measured by Nanodrop fluorospectrometer (ND-1000 NanoDrop Technologies, USA) before cDNA synthesis. PrimeScriptTM RT reagent Kit was used for reverse transcription. 2 µg of total RNA was mixed with 1µl oligo (Goetz et al.) and DEPC (Diethylpyrocarbonate)-treated ddH_2O in a PCR tube to a total volume of 12 µl. The RNA mixture was incubated at 65°C for 5 mins then cooled on ice for 5 mins; 4 µl 5X RTase Buffer, 1 µl RNase Inhibitor and 2 µl dNTP were added to each tube and each was incubated at 37 °C for 5 mins. 1µl RTase was then added into the mixture and incubated at 42 °C for 70 mins. The reaction was inactivated by incubating at 70°C for 10 mins. The cDNA was then stored at -20°C until required.

2.6.3 Quantitative Reverse Transcriptase PCR (qRT-PCR)

qRT-PCR was performed to quantify gene expression levels. QuantiNova SYBR Green PCR Kit (QIAGEN, Shanghai) was used (see Table 2-11).

Table 2-11 SYBR Green PCR reaction composition

Component	Volume (μl)
2 X SYBR Green PCR Master Mix	5
Forward primer (10 pmol/μl)	0.2
Reverse primer (10 pmol/μl)	0.2
Template cDNA	10-50 ng
ddH_2O	To a final volume of 10 μl

qRT-PCR was performed according to the program in Table 2-12 and analysed using the LightCycler480 System (Roche, China).

Table 2-12 qRT-PCR program

Step	Temperature (°C)	Duration (min:sec)	Cycle
Initial denaturing	95	10:00	1
Denaturing	95	0:30	45
Annealing	58	0:30	
Elongation	72	1 minute/kb	
Dissociation	72-95	0:30	Increase 1°C per cycle

OsACTIN (Os03g0718100) was used as a reference gene (see primers in Appendix II). All samples were run in at least triplicate. For expression analysis in rice anthers, 3 replicates of anther samples were collected at anther development stage 8 to 10.

2.7 Dual-Luciferase Reporter Assay

2.7.1 Plasmid construction

The effector plasmids with TIP2, TDR1 and EAT1 genes under the control of the CaMV35S promoter were obtained from Dabing Zhang's lab (Shanghai Jiao Tong University). Promoter sequences of candidate downstream genes were fused to the reporter vector pGreen0800 by In-fusion cloning (section 0). Mixed 200ng pGreen0800 vector and 200ng helper vector pSoup and transformed the two vectors together to GV3101 as described in section 2.4.3.

2.7.2 Preparation of bacteria solution for infiltration

GV3101 was incubated with the required plasmid in 15 ml LB liquid medium (selective medium with Kan, Rif and Gen) overnight at 28 °C until an OD600 of between 1 and 2 was reached. The reaction was then centrifuged at 5,000 rpm (4,000xg) for 5 min to precipitate bacteria. The pellet was then resuspended with ½ MS and 2 % (w/v) sucrose solution (1.1g MS and 10 g sucrose, add ddH_2O to a final volume of 500 ml) to reach the concentration of OD600 = 0.6. MES and AS were then added to reach a final concentration of 10 mM and 200 μM respectively. The bacteria solution was then ready for tobacco injection after 3 hours standing.

2.7.3 Tobacco plants infiltration and transient expression analysis mediated by *Agrobacterium*

The bacteria solution was mixed with pGreen0800 vector and pGreen0229

vector in a ratio of 1:4. *Agrobacterium* was injected into the leaves of tobacco at approximately 4 weeks old when 6 leaves have formed. Empty pGreen0229 vector as effector plasmid mixed with constructed reporter vector was used as negative control. To minimize biological variation, every leaf was treated using both negative control and treatments at separate positions. Four biological repetitions were used for each treatment. After injection, tobacco plants were maintained in the dark for 48 hours. Tissue at the injection spots were collected into 1.5 ml Eppendorf tubes and frozen in liquid nitrogen and then used for fluorescence detection.

2.7.4 Fluorescence measurement

Dual-luciferase Reporter assay system (Promega, Shanghai) was used for studying the receptor activity. Fluorescence was read by a Modulus™ single tube multimode reader (TurnerBioSystems) running the Promega protocol, DLR-0-INJ. Passive lysis buffer (PLB), luciferase assay buffer (LAB) and stop buffer (SB) were stored and prepared according to the manufacturer's instructions.

Tobacco leaves were ground in liquid nitrogen and mixed with 100 µl PLB, then centrifuged at 12,000 rpm (13,000xg) for 1 min. 10 µl of supernatant was then transferred to a new 1.5 ml Eppendorf tube and 40 µl LAB carefully added and then mixed by gentle pipetting several times. Tube was then placed into the reader and firefly luciferase activity initiated and then measured (light

with an emission maximum of 560 nm).

40 µl SB was added carefully and mixed by pipetting several times to start Renilla luciferase activity measurement (light with an emission maximum of 485 nm). The ratio of firefly luciferase and Renilla luciferase (λ560nm/λ485nm) were calculated to present promoter activity.

The data of all treatments were normalized to the empty vector control and Wilcoxon test was performed to determine the regulation of the target promoter by target protein.

2.8 CHARACTERIZING PHENOTYPE OF MUTANT

2.8.1 Measurement of pollen viability by iodine staining

Six anthers from one flower were crushed on a glass slide with a drop of potassium iodide (0.05 mol/l). The images were analysed by microscopy and the percentage of pollen viability recorded.

2.8.2 Scanning electron microscope

Anthers at anther stage 8 to stage 12 were fixed in FAA and stored at 4 °C. Samples were dehydrated using an ethanol gradient series (70%，80%，95%，100% (v/v) in turn, each for 30 mins). Samples were dried at the critical point using CO2 in a Critical Point Drier (Leica EM CPD300, Germany). Then conductive coating was added using Super Cool Sputter Coater (Leica SCD050, Germany) to avoid any charge effect during scanning. After

conductive coating, the morphology of the pollen and anther wall were observed using Scanning Electron Microscope (Hitachi S-4800, Japan).

2.8.3 Semi-thin section

Anthers at anther stage 8 to stage 12 were prepared as described in (Zhang et al., 2002). Fluorescence confocal microscope (Leica TCS SP5, Germany) was used for sample examination.

3 GENERATION OF A RICE GENE CO-EXPRESSION NETWORK FOR IDENTIFYING GENES INVOLVED IN ANTHER AND POLLEN DEVELOPMENT

3.1 INTRODUCTION

In this chapter, available gene expression data were used to construct a correlation gene network to identify putative key genes involved in pollen wall formation, to help in the selection of candidates for further research. As discussed in section 1.4.2, gene co-expression analysis has proved to be a reliable tool for predicting gene functions, however there is no global co-expression gene network had been generated in rice during rice anther and pollen development. Therefore, developing a network with co-expressed genes clustered together in rice anther development is therefore of benefit for the study of rice anther development. Genes differentially expressed between wild type and male-sterile rice mutants were used for filtering and prioritizing the clusters into clades. Candidates in clades were considered of importance in rice pollen development.

3.2 RESULTS

The workflow used for the construction of the correlation network and identification of novel genes potentially involved in pollen development is shown in Figure 3-1. Based on microarray datasets from rice anther samples (section 3.2.1), co-expression analysis followed by tight clustering was used to

generate an untargeted genome-wide gene co-expression network (section 3.2.4) focusing on rice anther development. Genes with similar expression patterns and correlated functions, especially during anther and pollen development, were connected by this approach. Differentially expressed genes of 9 male sterile mutants were mapped onto the co-expression network to further focus on the clusters related to anther development (section 3.2.5). Expression patterns of other tissues were checked for those example clusters to predict novel genes which might play important roles in anther and pollen development for further research. Furthermore, a web-based resource (section 3.2.9) that capitalizes upon these approaches was built to provide a valuable targeted resource for gene discovery, to enable further characterization of anther and pollen development in rice, and by association to other monocot crops.

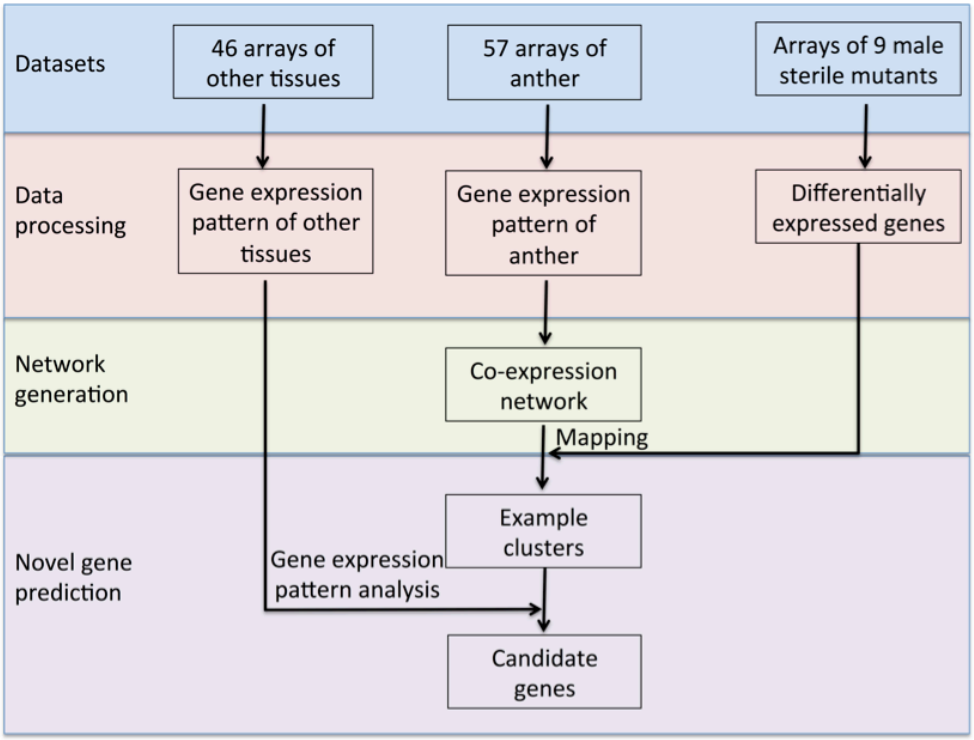


Figure 3-1 Workflow of predicting candidate genes by co-expression network.

3.2.1 Datasets selection

To globally investigate the key genes critical for rice anther development and pollen formation throughout anther development, 57 Affymetrix arrays generated only from rice anther tissue were selected from GEO (Appendix III). These samples were derived from wild-type rice anthers from five experiments that covered anther development stages 2 to 14 (Ma, 2005; Zhang & Yang, 2014), which present an overview of genes expressed during anther development in rice (Appendix III).

A further 45 Affymetrix arrays of samples from other tissues were used to characterize the global expression patterns in rice (Appendix III). The raw .cel data files of Affymetrix GeneChip Rice Genome Array were all downloaded

from NCBI's GEO (Barrett et al., 2013).

3.2.2 Data processing

Quality control (section 2.2.1) was performed to confirm the quality of each array. All the 57 arrays from anther and the 45 arrays from other tissues were of good quality and suitable for generation of the co-expression network. RMA normalization (Irizarry et al., 2003) (section 2.2.2) was conducted and overall 24,462 genes from different experiments were found to be comparable with similar average intensity values (Figure 3-2). Figure 3-2 shows a histogram of the normalized data of 57 anther arrays, showing a transition between noise and signal with a peak around 8 ($\log_2$ scale); genes above this level are above average expressed in these samples. Therefore, the set of genes was filtered by removing those which have no samples with intensity levels above 8 (on the log2 scale), leaving 14,468 genes that have at least one higher value for co-expression analysis.

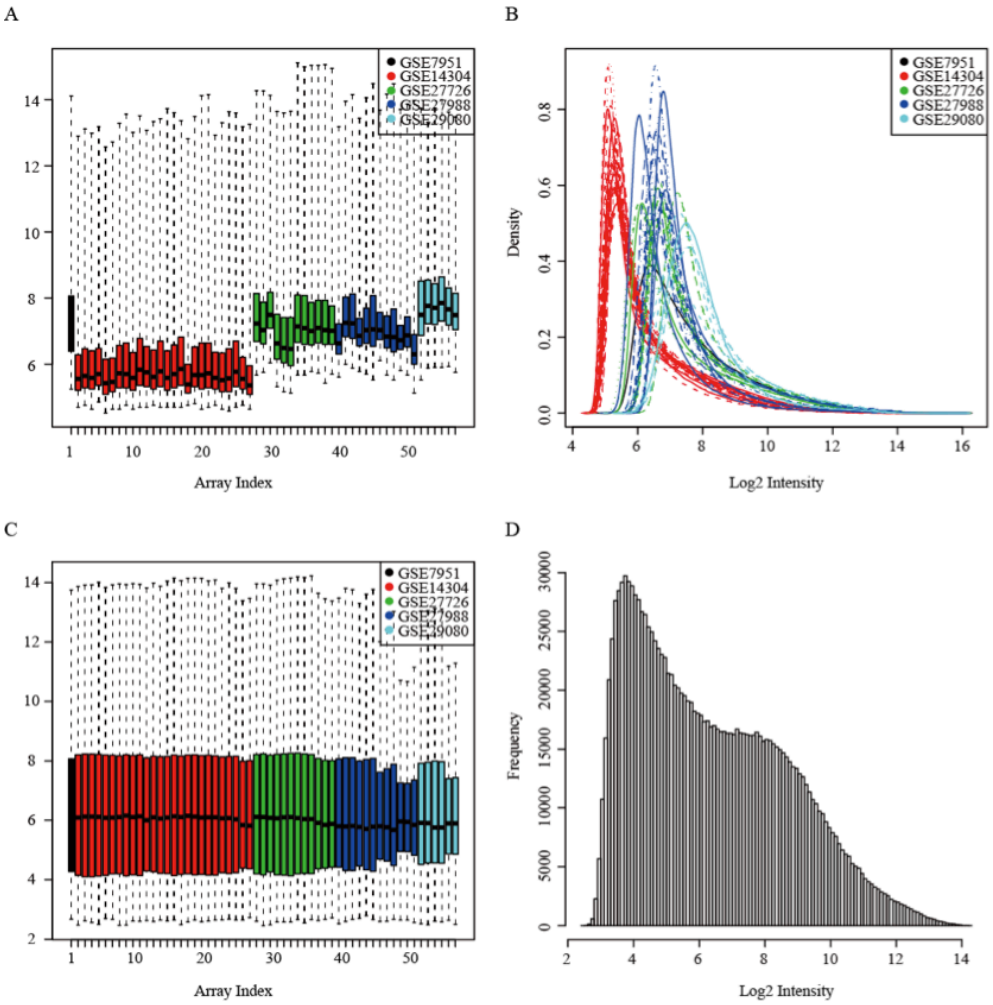


Figure 3-2 Boxplots and density plots of log-intensity distribution for 57 arrays before and after normalization. (A) Raw signal distributions are variable over 57 arrays. (B) The density distributions of raw PM (perfect match probes) log-intensities are different before normalization. (C) Normalized data had similar signal distributions over 57 arrays. (D) Total frequency distribution of gene expression intensity (y-axis) in normalized data shows a superposition of two distributions, with a first peak around log intensity =4 (noise) and a second peak around log intensity=8 (threshhold).

3.2.3 Generation of gene expression profiles from rice anthers

The 57 Affymetrix arrays generated only from anther tissue were used to globally investigate the gene expression patterns of all 24,462 genes during anther development stage 2 to 14 in rice. Both mean and SD (Standard Deviation) were calculated for each gene among replicates. Plots showing how the expression levels varied for all 24,462 genes were made,

grouped by either dataset or stage. Figure 3-3 shows *TDR*'s expression pattern in rice anthers as an example.

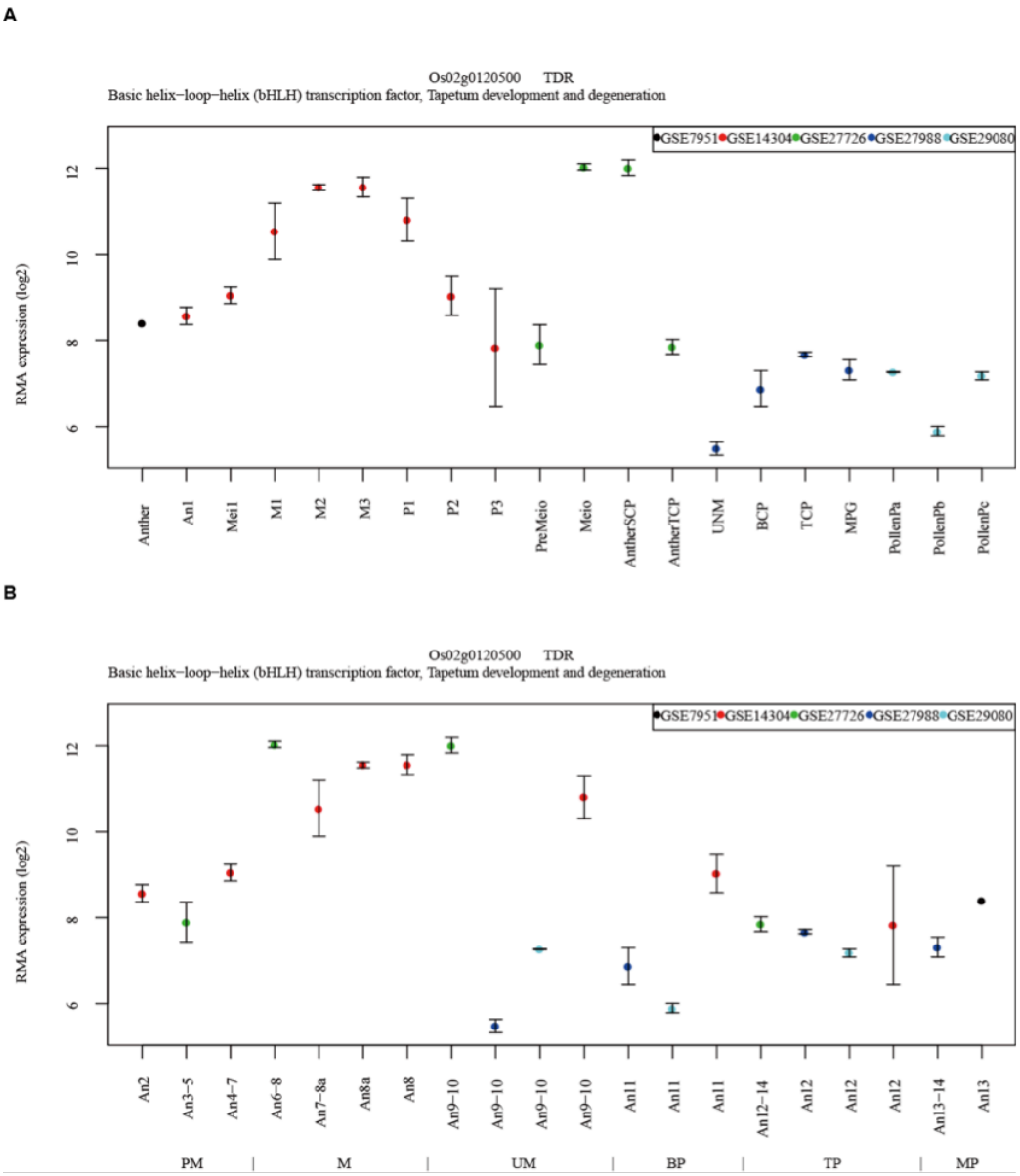


Figure 3-3 Gene expression profile of TDR using the generated RiceAntherNet tool (www.cpib.ac.uk/anther/riceindex.html). PM (Premeiosis)=Anther Stage 1-7; M(Meiosis)=Anther Stage 8; UM (Uninucleate microspore)= Anther Stage 9 and 10; BP (Bicellular Pollen)=Anther Stage 11; TP (Tricellular Pollen)=Anther Stage 12; MP (Mature Pollen)=Anther Stage 13 and 14.

3.2.4 Construction of rice anther correlation network (RiceAntherNet) for gene function analysis

Based upon wild type expression data, genes that had an expression level of over 8 in anthers (according to Figure 3-2D, there is a peak of gene expression level at 8 which was recognized as high expression) in at least one array were kept for network development, which lead to 9,358 anther-expressed genes remaining in the analysis of the 14,468 genes with intensity levels above 8 (on the log2 scale). We then used Pearson Correlation Coefficient (PCC) (section 2.2.3) to calculate the correlation between all remaining genes, giving a correlation matrix, the histogram of which gives a smooth curve (Figure 3-4). A semi-arbitrary cutoff was then applied to this matrix to give a set of edges to analyze. In line with previous studies, which found that approximately half a million edges was a reasonable number to generate clusters from (Bassel et al., 2011; Dekkers et al., 2013; Pearce et al., 2015), we chose a cutoff of 0.9, leading to 411,060 edges between 9,358 genes. Changing this cutoff would lead to more or less edges, but the general structure of the network would stay the same.

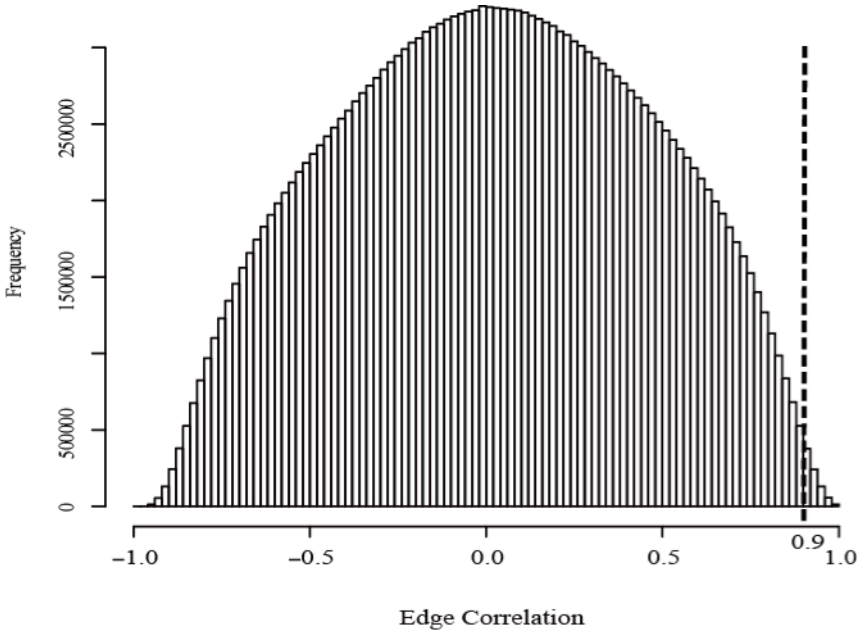


Figure 3-4 Distribution of edge correlations. PCC = 0.9 was selected as cutoff. The edges with PCC less than 0.9 were removed from network.

A correlation network was developed using the correlation matrix (see section 2.2.4), which we have called RiceAntherNet, with 411,060 edges between 9,358 genes. The full network is available as a Cytoscape file (see attached CD). Figure 3-5 shows a global view of the network, which illustrates that the gene interactions are divided into three main, densely-packed regions, with fewer genes forming transitional connections between them.

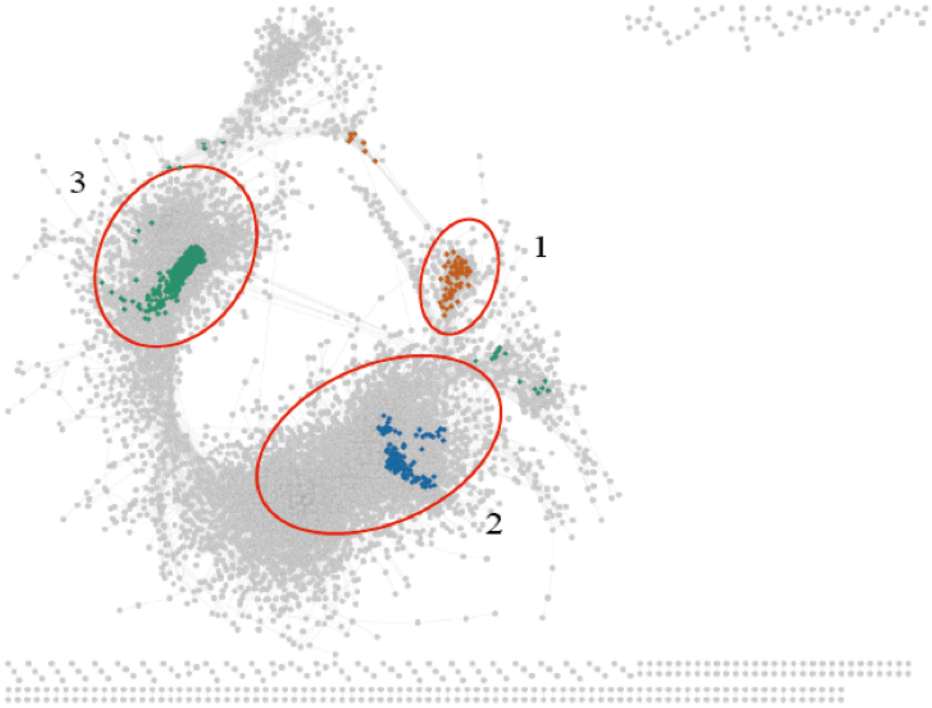


Figure 3-5 Global co-expression network plot by organic layout in Cytoscape. The three circled regions refer the three principal interaction regions. Clade 1 (red) is in region 1; Clade 2 (blue) is in region 2; Clade 3 (green) is in region 3.

To further identify genes with similar expression patterns from the global network, the plugin ClusterMaker (section 2.2.4) (Morris et al., 2011) was used to divide the network into distinct clusters, using the Transitivity Clustering method (TransClust) (Wittkop et al., 2010) with parameters Max Subcluster Size = 400 and Max Time = 10, with the PCC values as edge weights. This method generates small, extremely well-connected clusters where nearly all the edges are present, meaning that cluster members must have very similar expression patterns. When used on this network, 3422 well-connected clusters with 9358 nodes (genes) and 75853 edges (co-relations) were generated (Figure 3-6A). The cluster sizes (gene numbers

in one cluster) ranged from 2 to 202 (Figure 3-6B). 545 clusters which contain at least 4 genes were kept for further analysis. These clusters contain 4,580 genes and 73,556 intra-cluster edges, 99.74% of the possible number.

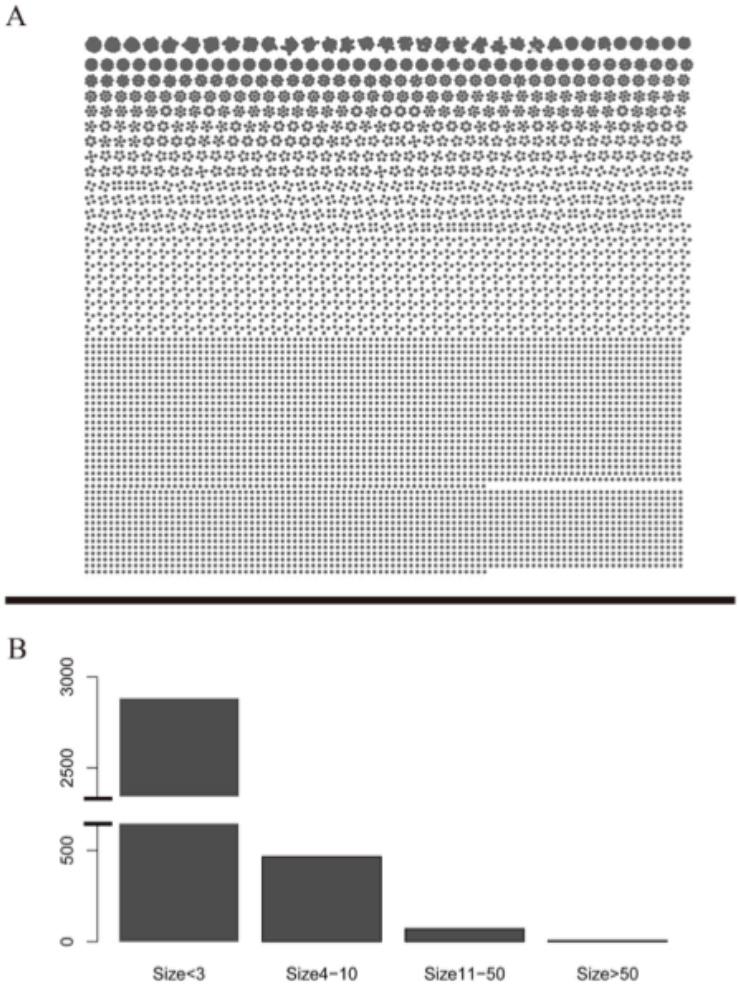


Figure 3-6 Network generated by TransClust. (A) Network with 9358 nodes (gene) and 75853 edges (co-relation). (B) Number of clusters of different sizes.

GO analysis shows genes in the same cluster had highly correlated expression pattern over the different rice anther developmental stages and many of them were annotated with the same GO terms. For example, cluster 116 contains eight genes (Figure 3-7A), all of which have very similar expression levels

over the 57 arrays (Figure 3-7B). Gene Ontology (GO) analysis using the orthologs Arabidopsis genes shows that this cluster is enriched in pollen exine formation (Figure 3-7C).

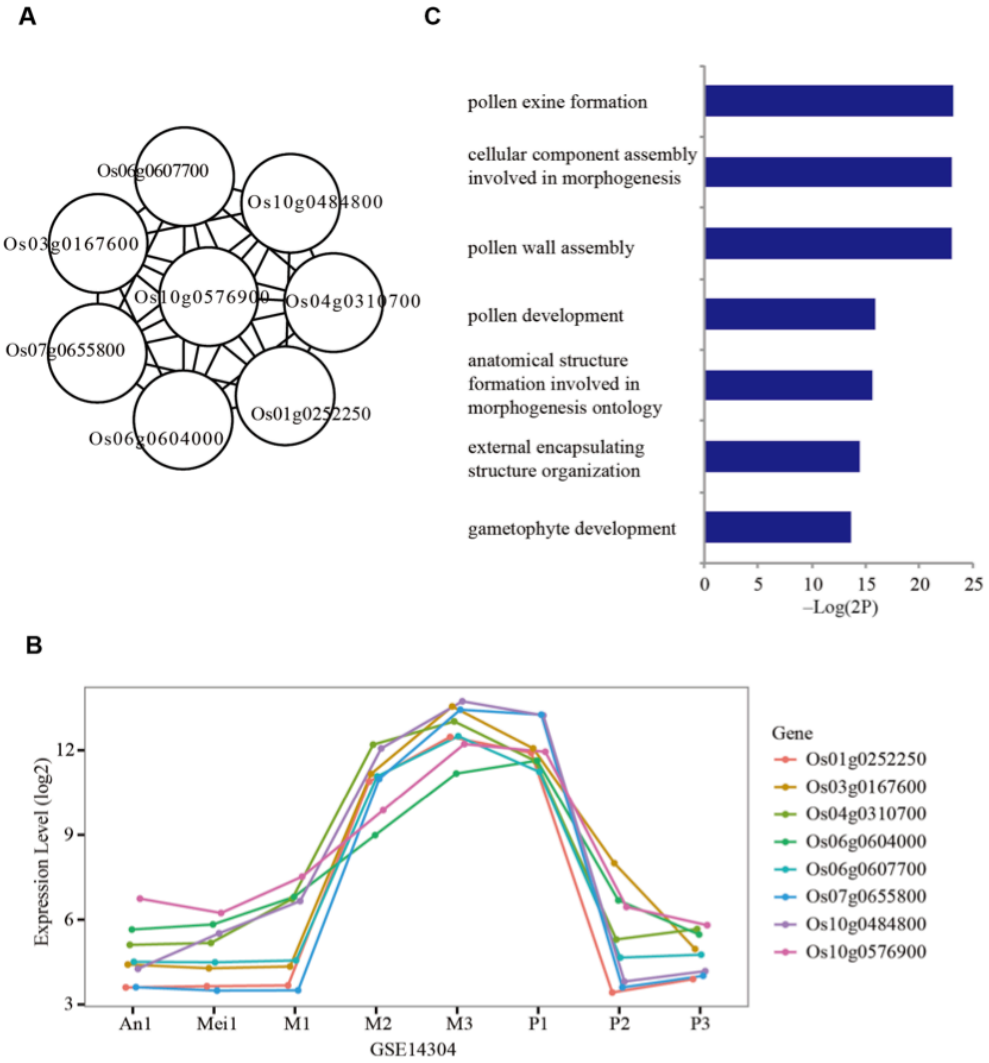


Figure 3-7 Analysis of gene Cluster 116 from RiceAntherNet. (A) 8 genes are interconnected in Cluster 116. (B) Gene expression pattern of cluster 116 in GSE14304. An1 = Anther stage 1-6, Mei1 = Anther stage 7, M1 = Anther stage 8a, M2 = Anther stage 8b, M3 = Anther stage 9, P1 = Anther stage 10, P2 = Anther stage 11, P3 = Anther stage 12. (C) Gene Ontology (GO) terms enriched in cluster 116. –Log(2P) = -log2(p-value).

3.2.5 Mapping differentially expressed genes (DEGs) in male sterile mutants onto the correlation network

Nine datasets of rice male sterile mutants (detailed in Table 3-1), including

msp1-4, ostdl1a, gamyb-2, tip2, udt1-1, tdr, eat1-1, ptc1 and *mads3-4* mutants, were selected for gene expression screening to target those genes involved in anther and pollen formation. Microarray data of mutants were processed and genes with 2 fold changes (p-value<0.05 by t test) between wild type and mutant were recognized as DEGs (Appendix III). The details of mutant datasets are listed in Table 3-1, including sample stages and number of DEGs.

Table 3-1 Nine male sterile mutant datasets. Up: the number of genes up regulated in mutants; down: the number of genes down regulated in mutants.

Gene	Mutant	Anther Sample Stages	Dataset Sources	Up	Down
MSP1	*msp1-4*	S1, S3, S5	Yang et al., 2016	157	486
OsTDL1A	*ostdl1a*	S1, S3, S5	Yang et al., 2016	243	814
GAMYB	*gamyb-2*	S9	Aya et al., 2009	568	416
TIP2	*tip2*	S6, S7	Unpublished data	104	433
UDT1	*udt1-1*	S9	Jung et al., 2005	209	712
TDR	*tdr*	S8a, S8b, S9	Unpublished data	1999	2105
EAT1	*eat1-1*	S9, S10	Unpublished data	660	503
PTC1	*ptc1*	S9	Unpublished data	243	1822
MADS3	*mads3-4*	S9	Hu et al., 2011b	833	895

DEGs in these mutant datasets mapped onto the global network (Figure 3-8). Mutants that show defects in anther cell division and differentiation (*msp1-4, ostdl1a and tip2)* (Fu et al., 2014; Hong et al., 2012; Nonomura et al., 2003; Zhao et al., 2008) had DEGs clustered in region 2 (Figure 3-8A); mutants with enlarged tapetum cells during the later stages of anther development (*udt1-1, gamyb-2* and *tdr*) (Aya et al., 2009; Hu et al., 2011; Li et al., 2006b) had a high density of DEGs in region 1 (Figure 3-8B); while mutants showing darkly stained materials, with a high density of organelles in the tapetum in later

stages (*ptc1* and *eat1-1*) (Li et al., 2011a; Niu et al., 2013) had DEGs distributed throughout all 3 regions (Figure 3-8C). Finally, *MADS3* which encodes a ROS-scavenging enzyme with associated high ROS levels in the *mads3-4* mutant (Hu *et al.*, 2011) was different from the other mutants, with most of its DEGs up-regulated in region 1 (Figure 3-8D).

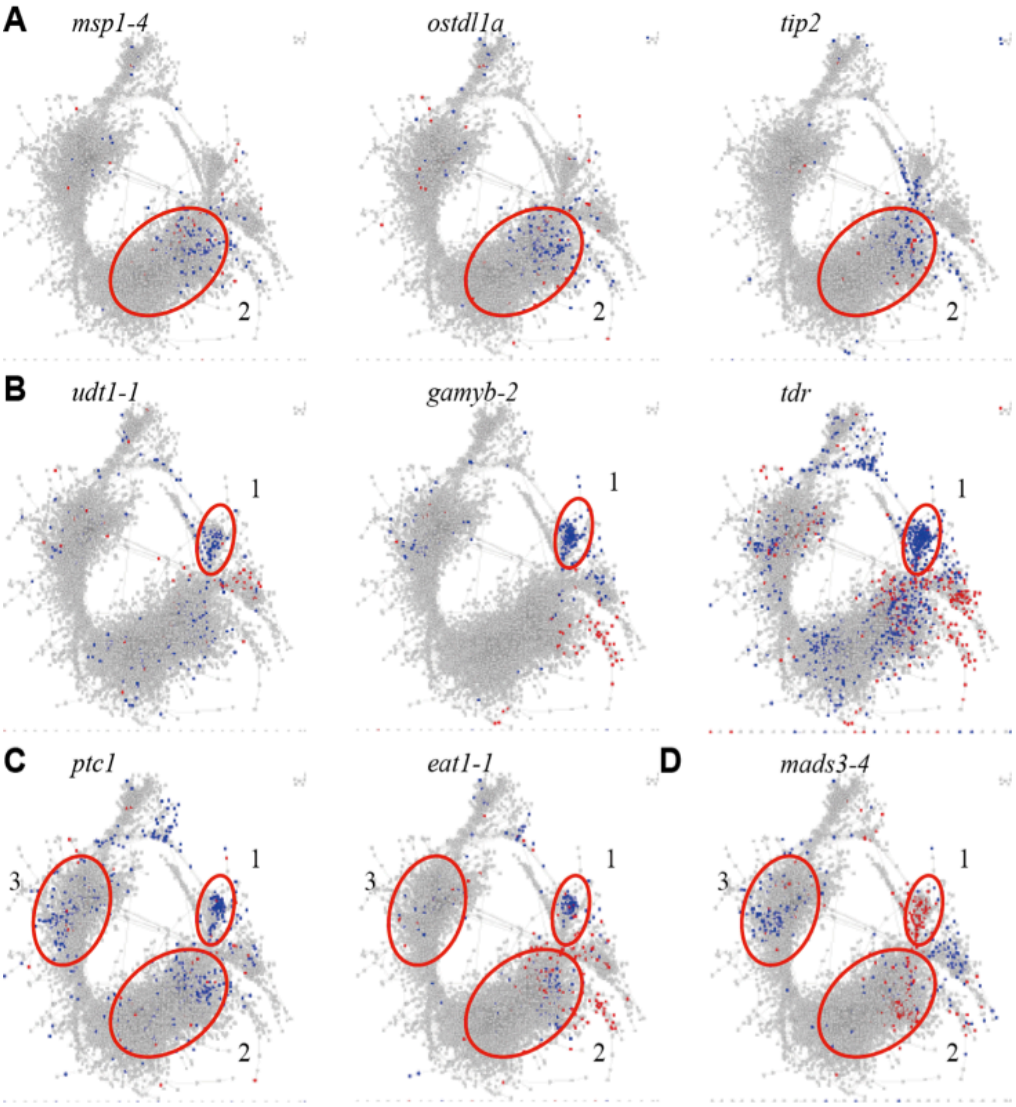


Figure 3-8 Differentially Expressed Genes (DEGs) of male sterile mutants in the global RiceAntherNet co-expression network. (A) Distribution of the DEGs of *msp1-4, ostdl1a* and *tip2* mutants. (B) Distribution of DEGs of *udt1-1, gamyb-2,* and *tdr* mutants. (C) Distribution of DEGs of *ptc1* and *eat1-1* mutants. (D) Distribution of DEGs of *mads3-4* mutants. Up-regulated genes are shown in red; down-regulated genes are shown in blue.

When DEGs of 9 male sterile mutants were mapped onto the clusters generated by TransClust, it was seen that a number of clusters are extremely rich in DEGs from different mutants while most of the clusters have none. These enriched clusters are therefore likely to be the important clusters related to anther development and were therefore 29 of these clusters were selected for further analysis (Figure 3-9). Other non-enriched clusters often represented genes involved in a particular pathway or response, for example cluster 3 is full of ribosomal proteins.

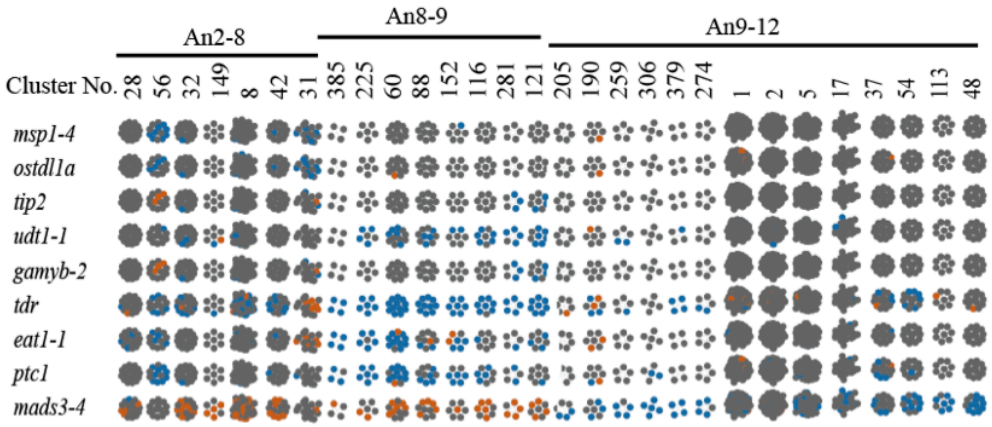


Figure 3-9 Example clusters from the RiceAntherNet. DEGs from male sterile mutants datasets were mapped onto the RiceAntherNet. Down-regulated genes are marked in blue and up-regulated genes are marked in orange.

To aggregate the clusters with similar expression trend together, the expression data for each gene was standardized, and an average expression pattern calculated for the each cluster. Using hierarchical clustering (section 2.2.4), the 29 selected clusters were grouped into three clades based upon similar expression trends. Genes in Clade 2 are predominantly expressed from Anther stage 2 to stage 8 (early anther differentiation stages and meiosis stage). Genes

in Clade 1 are highly expressed at stage 8a, stage 8b and stage 9 (the meiosis and tapetum PCD initiation stages); Genes in Clade 3 are expressed after Anther stage 9 (the late stage) (Figure 3-10). These three clades are referred to as interaction region 1, 2 and 3 respectively (Figure 3-5).

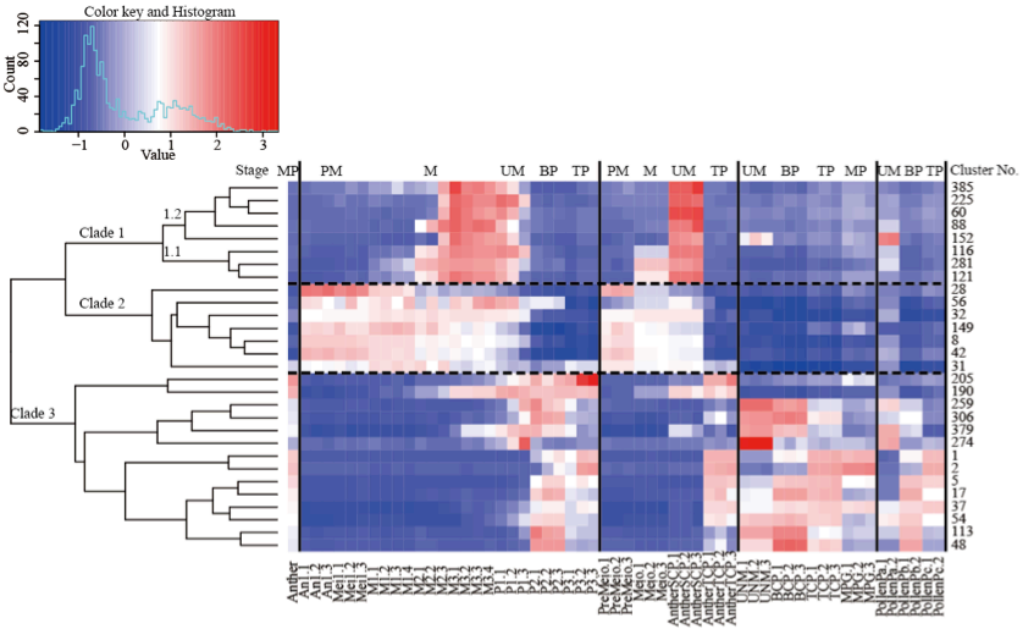


Figure 3-10 Hierarchical clustering of example RiceAntherNet clusters. Clusters that are expressed at different anther develop stages separate into three Clades. Clade 2 genes are expressed throughout anther stage 2-8 in rice anthers; Clade 1 expressed during anther stage 8 and 9 in rice anthers; Clade 3 expressed after anther stage 9 in rice pollen. PM (Premeiosis) = Anther Stage 1-6; M (Meiosis) = Anther Stage 7 and 8; UM (Uninucleate microspore) = Anther Stage 9 and 10; BP (Bicellular Pollen) = Anther Stage 11; TP (Tricellular Pollen) = Anther Stage 12; MP (Mature Pollen) = Anther Stage 13 and 14.

Most genes in Clade 1 are significantly down regulated in *udt1-1, tdr, eat1-1* and *ptc*1 mutants which have defects in pollen wall formation; some genes in Clade 2 show differential expression in anther cell division and differentiation defective mutants: *msp1-4, ostdl1a* and *tip2*; Clade 3 has less DEGs in the mutants except for *mads3-4*. Different from other male sterile mutants,

mads3-4 mutant had downstream genes up regulated in Clade 1 and Clade 2 and down regulated in Clade 3, which is consistent with its distinctive role (Figure 3-10). This consistent DEGs distribution both in global network and Clades generated by gene expression pattern indicates that genes connected to a particular function are predominantly gathered together and show co-expression patterns. GO analysis further suggested the divided functions of different clades (Table 3-2). Clade 1 shows GO enrichment of pollen exine formation, Clade2 has overrepresented GO terms in organ development and Clade3 has overrepresented GO terms in developmental growth involved in morphogenesis.

Table 3-2 GO enrichment analysis for 3 Clades.

Clade 1				
GO.ID	Term	Annotated	p-value	DEGs
GO:0010584	Pollen exine formation	55	1.40E-17	9
GO:0010208	Pollen wall assembly	58	2.30E-17	9
GO:0010927	Cellular component assembly involved in morphogenesis	58	2.30E-17	9
GO:0080110	Sporopollenin biosynthetic process	7	4.80E-14	5
GO:0009555	Pollen development	299	8.40E-11	9
Clade 2				
GO.ID	Term	Annotated	p-value	DEGs
GO:0048513	Organ development	1975	2.80E-11	14
GO:0006796	phosphate-containing compound metabolic process	1776	1.90E-10	21
GO:0006793	Phosphorus metabolic process	1777	2.00E-10	21
GO:0009887	Organ morphogenesis	503	3.20E-10	8
GO:0009653	Anatomical structure morphogenesis	1739	5.40E-10	12
GO:0050793	Regulation of developmental process	854	7.40E-10	7
GO:0048856	Anatomical structure Development	3349	5.30E-09	17
GO:0043412	Macromolecule modification	3213	5.50E-09	24
Clade3				
GO.ID	Term	Annotated	p-value	DEGs
GO:0060560	Developmental growth involved in morphogenesis	578	3.60E-29	7
GO:0016049	Cell growth	693	2.20E-28	10
GO:0009827	Plant-type cell wall modification	220	5.80E-28	5
GO:0009826	Unidimensional cell growth	483	1.20E-27	6
GO:0048589	Developmental growth	760	5.30E-27	8
GO:0071669	Plant-type cell wall organization or biogenesis	481	8.10E-27	7
GO:0042545	Cell wall modification	384	1.60E-26	6
GO:0000902	Cell morphogenesis	673	9.90E-26	9

3.2.6 Clade 1 includes genes for pollen wall formation

There are 50 genes in Clade 1; these are predominantly expressed during Anther Stages 8 and 9. Based on microarray datasets of other tissues, genes in Clade 1 also show anther-specific expression pattern (Figure 3-11).

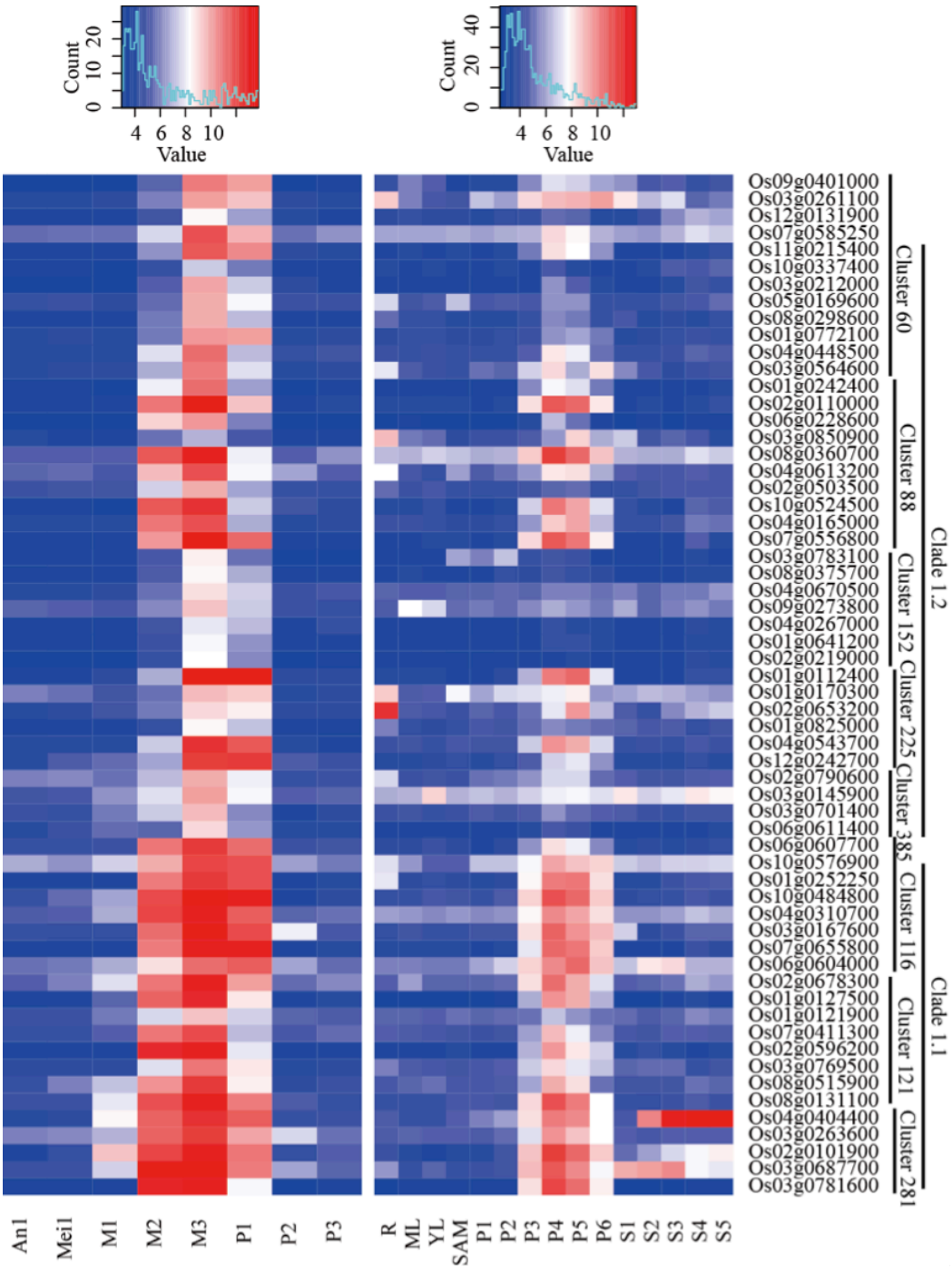


Figure 3-11 Expression patterns of Clade 1 genes in rice anther and other

tissues. Left panel: heatmap of Clade 1 in anther tissues (dataset GSE14304); Right panel: heatmap of Clade 1 in other tissues. Stages in left panel: An1 = Anther Stage 1-2; Mei = Anther Stage 3-6; M1 = Anther Stage 7; M2 = Anther Stage 8a; M3 = Anther Stage 8b; P1 = Anther Stage 9 and 10; P2 = Anther Stage 11; P3 = Anther Stage 12. Tissues in right panel: R = Root; ML = Mature Leaf; YL = Young Leaf; SAM = Shoot Apical Meristem; P1 = Inflorescence up to 3 cm; P2 = Inflorescence, 3 - 5 cm; P3 = Inflorescence, 5 - 10 cm; P4 = Inflorescence, 10 - 15 cm; P5 = Inflorescence, 15 - 22 cm; S1 = Seed, 0 - 2 DAP (Days After Pollination); S2 = Seed, 3 - 4 DAP; S3 = Seed, 5 - 10 DAP; S4 = Seed, 11 - 20 DAP; S5 = Seed, 21 - 29 DAP.

Most of the genes in Clade 1 are significantly changed in *udt1-1, tip2, gamyb-2, tdr, eat1-1, ptc1* and *mads3-4* mutants. Clade 1 was further separated into two sub-clades which show distinct gene expression patterns in the different male sterile mutants. Among all 21 genes in Clade 1.1, 20 of them are DEGs in the *tdr* mutant anthers; 13 are differentially expressed in the *mads3-4* mutant; 12 of them are DEGs in *gamyb-2* and *udt1-1* mutant; whereas among the 39 genes in Clade 1.2, 38 genes are DEGs in the *tdr* mutant; 29 genes are DEGs in *gamyb-2*. *eat1-1, mads3-4, ptc1* and *udt1-1* mutants had 26, 14, 20, 11 DEGs respectively. (Appendix III and Figure 3-12) The high percentage of DEGs in Clade 1 indicates the important role of these genes in anther and pollen development.

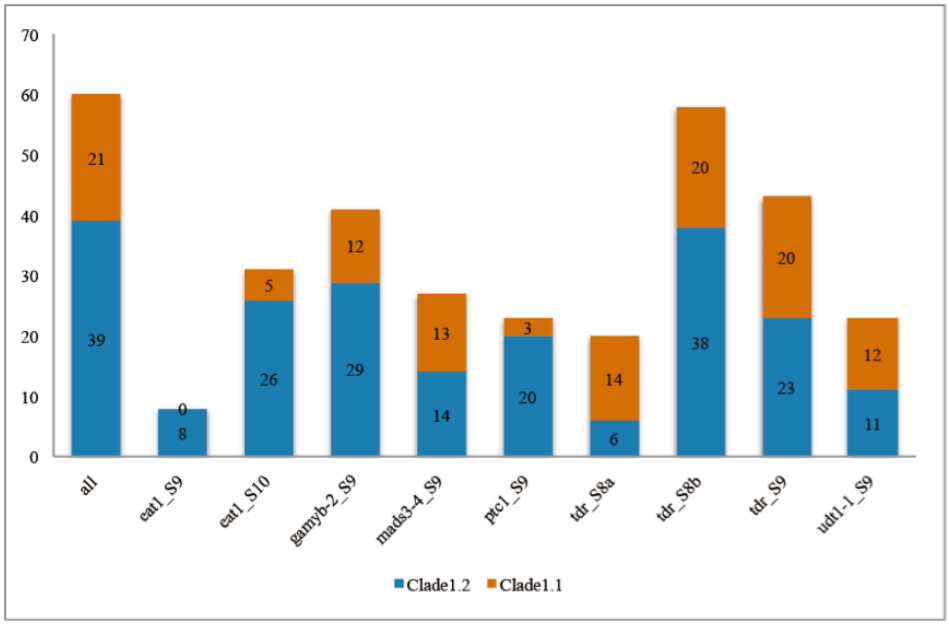


Figure 3-12 Number of Differentially Expressed Genes (DEGs) in Clade 1. The first bar shows size of Clade 1. The remaining bars are DEGs in various mutants: *eat1-1* mutant at anther stage9 and 10; *gamyb-2* mutant at anther stage 9; *mads3-4* mutant at stage anther stage9; *ptc1* mutant at stage9, *tdr* mutant at anther stage8a, stage8b and stage9; *udt1-1* mutant at stage9 respectively. Clade 1.1 is in orange; Clade 1.2 is in blue.

3.2.6.1 Characterization of genes within Clade 1.1

There are three co-expression clusters identified in Clade 1.1 (Figure 3-11), which include clusters 116, 121, and 281 which contain eight, eight and five genes respectively (Table 3-3). Genes in Clade 1.1 are specifically expressed at stage 8 and stage 9 (Figure 3-11). Their putative Arabidopsis orthologs are all found in Cluster 37 in the Arabidopsis FlowerNet (Pearce et al., 2015) and show similar expression patterns in Arabidopsis anthers. Furthermore, genes in Clade 1.1 and their orthologs are all anther-specific genes in both rice and Arabidopsis, which suggests the conservation of gene expression profile between Arabidopsis and rice during anther development.

Table 3-3 Genes in Clade 1.1. Genes with known functions related to pollen development are in bold.

RAP ID	Cluster	Arabidopsis Orthologs	Reported Function/Annotation
Os06g0607700 /*OsABCG15*	116	AT3G13220 /*WBC27*	**Involved in sporopollenin transport in rice and Arabidopsis.**
Os10g0484800 /*OsLAP6*	116	AT1G02050 /*LAP6*	**Involved in fatty acid modifications in rice and Arabidopsis.**
Os06g0604000	116	AT1G15360 /*SHN1*	**Involved in lipid synthesis in Arabidopsis.**
Os03g0167600 /*DPW*	116	AT3G11980 /*MS2*	**Involved in fatty acid modifications in rice and Arabidopsis.**
Os10g0576900	116	AT2G23910	NAD(P)-binding domain containing protein.
Os07g0655800	116	AT3G63095	Similar to Aspartic acid-rich protein.
Os04g0310700	116	NA	Putative uncharacterized protein.
Os01g0252250	116	NA	Putative uncharacterized protein.
Os03g0769500	121	AT3G50770 /*CML41*	EF-Hand type domain containing protein.
Os02g0596200	121	AT1G13130	Glycoside hydrolase, family 5 protein.
Os07g0411300 /*OsLAP5*	121	AT4G34850 /*LAP5*	**Involved in sporopollenin biosynthesis in rice and Arabidopsis.**
Os08g0131100 /*CYP703A3*	121	AT1G01280 /*CYP703A2*	**Involved in sporopollenin biosynthesis in rice and Arabidopsis.**
Os08g0515900	121	AT4G35420 /*DRL1*	**Involved in flavonoid biosynthesis in Arabidopsis.**
Os02g0678300	121	AT1G12570	Downstream of SHN1 (SHINE1) in Arabidopsis.
Os01g0127500	121	AT1G68540 /*TKPR2*	**Involved in flavonoid biosynthesis in Arabidopsis.**
Os01g0121900	121	AT5G37870	Seven-in-absentia protein, sina domain containing protein.
Os03g0263600	281	AT3G59530 /*LAP3*	**Involved in flavonoid biosynthesis in Arabidopsis.**

Os02g0101900	281	AT2G16630	Conserved hypothetical protein.
Os04g0404400	281	AT1G47980	Similar to H0502B11.4 protein.
Os03g0687700	281	NA	Ribosome-inactivating protein domain containing protein.
Os03g0781600	281	NA	Conserved hypothetical protein.

Among the 21 genes within the clade there are a number of lipid synthesis genes, such as Os03g0167600 (*DPW*), Os08g0131100 (*CYP703A3*) that are important for sporopollenin synthesis. Rice *DPW* and its putative ortholog in Arabidopsis *MS2* encode a fatty acid reductase, which utilizes palmitoyl-Acyl Carrier Protein to produce 1-hexadecanol. Both *DPW* and *MS2* are required for pollen wall formation (Chen et al., 2011; Shi et al., 2011); *CYP703A3* is required for in-chain hydroxylation in rice, while the homolog *CYP703A2* has a similar conserved function in Arabidopsis (Bush & Krysan, 2007; de Azevedo Souza et al., 2009; Dobritsa et al., 2009b; Li et al., 2010; Yang et al., 2014). Loss of function mutants of *DPW* and *CYP703A3* are defective in pollen wall formation (Xu et al., 2017; Yang et al., 2014).

In Arabidopsis, the medium- to long- chain fatty acids that are required for sporopollenin synthesis and pollen wall formation are catalyzed by ACOS5 (CoA Synthetase5) into medium- to long- chain fatty-acyl-CoA (de Azevedo Souza et al., 2009) which can be condensed with malonyl-CoA by PKSA and PKSB into tri- and tetra-ketide α-pyrone (Kim et al., 2010b; Wang et al., 2013) and reduced by DRL1 (Dihydroflavonol 4-reductase-like1)/TKPR1 and TKPR2 (Grienenberger et al., 2010). These genes (*ACOS5* (Qin et al., 2016),

LAP6/PKSA (LAP6/POLYKETIDESYNTHASEA) (Kim et al., 2010b), *LAP5/PKSB (LAP5/POLYKETIDESYNTHASEB)* (Kim et al., 2010b), *DRL1/TKPR1* and *TKPR2* (Grienenberger et al., 2010)) are all co-expressed in Arabidopsis. Os10g0484800 (*OsLAP6*), ortholog of *PKSA/LAP6*, and Os07g0411300 (*OsLAP5*), the rice orthologs of *PKSB/LAP5*, are both found in Clade 1.1. The *oslap5* RNAi mutant showed a partial male sterile phenotype (Wang et al., 2013); while oslap6 loss-of-function mutant was identified within a male sterile mutant library in Prof. Zhang's laboratory (unpublished data). Furthermore, the OsLAP5 and OsLAP6 show conserved biochemical activity in sporopollenin synthesis (Kim et al., 2009; Wang et al., 2010). These findings suggest the expression patterns of the principal sporopollenin biosynthesis genes are conserved during pollen development between rice and Arabidopsis. Unfortunately, *OsACOS12* (the ortholog of *ACOS5*) was not present in RiceAntherNet as *OsACOS12* was not included on the Affymetrix arrays, although the *osacos12* mutant shows defects in pollen development (Li et al., 2016; Yang et al., 2017). For the TKPRs in Clade 1.1, Os08g0515900 is an ortholog of *TKPR1* and Os01g0127500 is the ortholog of *TKPR2* (Figure 3-13A). In rice, *OsDFR2* has been reported to have a role in pollen development as the ortholog of *TKPR1* (Grienenberger et al., 2010). Gene expression profiles of these three genes indicate that they are co-expressed during anther development, while *OsDFR2* shows a lower, but similar expression pattern compared with the others (Figure 3-13B). For the

conservation of the co-expression pattern as well as the biochemical and biological function of genes in this pathway, we can therefore speculate that Os01g0127500 and Os08g0131100 are likely to have a key role in pollen development. These findings illustrate the importance of RiceAntherNet for identifying novel genes required for pollen development.

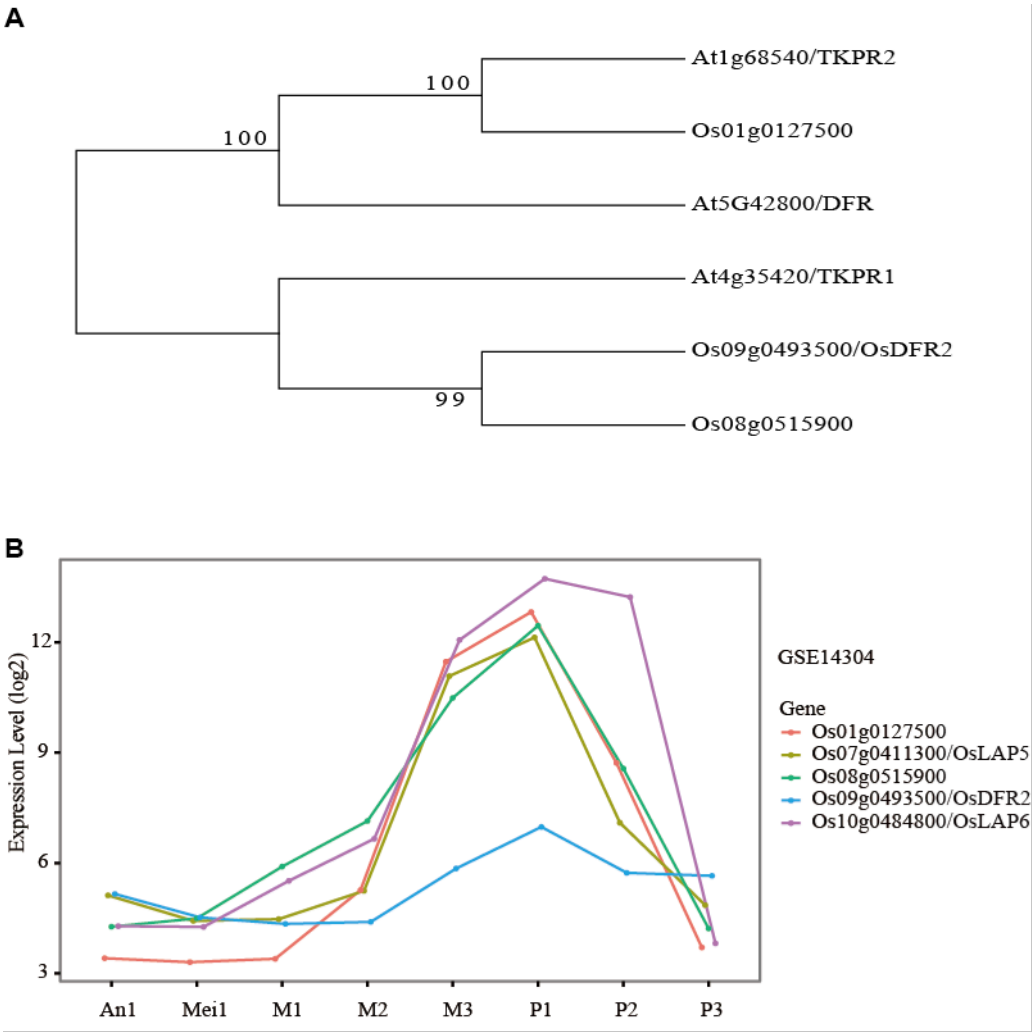


Figure 3-13 Analysis of TKPR genes in Clade 1. (A) Phylogenetic tree of selected TKPR protein sequences (details in Appendix IV) in rice and Arabidopsis. (B) Gene expression pattern of TKPR genes and co-expressed genes in GSE14303. An1 = Anther stage 1-6, Mei1 = Anther stage 7, M1 = Anther stage 8a, M2 = Anther stage 8b, M3 = Anther stage 9, P1 = Anther stage 10, P2 = Anther stage 11, P3 = Anther stage 12.

These findings suggest that sporopollenin synthesis genes are co-expressed

during anther development, and that this expression pattern is conserved in different species. This further supports the possibility that the lipid or phenolic synthesis related genes clustered in Clade 1.1 might be required for pollen wall formation.

The ABC transporter OsABCG15 that has been reported as required for pollen wall development (Qin et al., 2013) clustered in Clade 1.1 with sporopollenin synthesis related genes. The clustering of sporopollenin synthesis related genes and a sporopollenin transport related gene agrees with the observed regulation of exine formation; sporopollenin synthesis and transport have been previously shown to be spatially and temporally regulated and to work together in pollen wall development.

In Arabidopsis, *lap3* mutant plants have defective pollen exine patterning by an unknown mechanism (Dobritsa et al., 2009a). Correspondingly, the ortholog of LAP3 (LESS ADHESIVE POLLEN3) in rice, Os03g0263600, is also clustered in Clade 1.1 in RiceAntherNet.

The ortholog of Os06g0604000, WIN1 (WAX INDUCER1)/SHN1 (SHINE1), encodes for an ethylene response transcription factor and regulates a series of cutin biosynthesis related genes (Kannangara et al., 2007). Similarly, the ortholog of Os02g0678300 (At1g12570) is a WIN1/SHN1 inducible gene which belongs to the glucose–methanol–choline family, which is required for lipid synthesis (Kannangara et al., 2007). The coordinated expression pattern

between the transcription factor and the downstream target in cutin synthesis as well as the anther-specific expression pattern suggest a role for these genes in anther cutin biosynthesis (Li-Beisson et al., 2010, 2013). Both Os06g0604000 and Os02g0678300 are clustered in Clade 1.1, further validating the robust correlation of functionally related genes in RiceAntherNet.

In summary, from all of the 21 genes in Clade 1.1, five genes *(OsABCG15, OsLAP5, OsLAP6, DPW, CYP703A3)* have been reported as involved in pollen exine formation both in Rice and Arabidopsis. Five genes (Os06g0604000, Os02g0678300, Os08g0515900, Os01g0127500 and Os03g0263600) have characterized orthologs in Arabidopsis *(SHN1, AT1G12570, TKPR1/DRL1, TKPR2* and *LAP3,* respectively*)*, which function in lipid and flavonoid biosynthesis during pollen wall formation. The remaining 11 genes in Clade 1.1 have not been characterized yet, but one can hypothesis that they are also likely to be involved in pollen wall formation, since they are co-expressed with 10 key genes involved in pollen exine formation. These genes are therefore good targets for future study.

3.2.6.2 Characterization of genes within Clade 1.2

Unlike Clade 1.1, the genes clustering in Clade 1.2 show a much lower level of conservation to Arabidopsis. Five gene co-expression clusters were identified in Clade 1.2, (Clusters 60, 88, 152 225 and 385; Table 3-4).

Table 3-4 Genes in Clade 1.2. Genes with known functions related to pollen development are in bold.

RAP ID	Cluster	Arabidopsis Orthologs	Reported Function/ Annotation
Os04g0448500	60	AT2G03200	Peptidase A1 domain containing protein.
Os11g0215400	60	NA	Peptidase aspartic, catalytic domain containing protein.
Os05g0169600	60	AT5G56040/SKM2	**Regulates fertility in Arabidopsis.**
Os10g0337400	60	AT3G47570	Similar to Taxa-1.
Os01g0772100	60	AT3G03760/LBD20	Lateral organ boundaries, LOB domain containing protein.
Os09g0401000	60	AT5G16600/MYB43	Cell differentiation and regulation of secondary cell wall biogenesis.
Os03g0564600	60	AT3G22060	Protein of unknown function DUF26 domain containing protein.
Os12g0131900	60	AT1G09890	Similar to LG27/30-like.
Os03g0212000	60	NA	Hypothetical conserved gene.
Os03g0261100	60	NA	Phospholipase A2 family protein.
Os07g0585250	60	NA	Putative uncharacterized protein.
Os08g0298600	60	NA	Putative uncharacterized protein.
Os06g0228600	88	AT1G10010/AAP8	Similar to Amino acid carrier.
Os02g0110000	88	AT5G45950	Lipase, GDSL domain containing protein.
Os10g0524500/ OsNP1	88	AT1G12570	Required for tapetum degradation in rice.
Os01g0242400	88	AT1G06280/LBD2	LOB domain containing protein.
Os02g0503500	88	AT1G24360	Similar to 3-oxoacyl-[acyl-carrier-protein] reductase.
Os04g0613200	88	AT4G20050/QRT3	**Degradation of the pollen mother cell wall in Arabidopsis.**
Os03g0850900	88	AT2G02850/ARPN	**Involved in anther development and pollination in Arabidopsis.**
Os04g0165000	88	NA	Similar to thaumatin-like protein 1a.
Os07g0556800	88	NA	Ribosome-inactivating protein family protein.

Os08g0360700	88	AT3G63095	Conserved hypothetical protein.
Os04g0670500/ OsCP1	152	AT5G43060	Downstream of TDR in rice.
Os02g0219000	152	AT1G27760/SAT32	**Involved in salt tolerance in Arabidopsis.**
Os09g0273800	152	AT1G75950/SKP1	Similar to Fimbriata-associated protein.
Os08g0375700	152	AT1G20140/SK4	BTB/POZ fold domain containing protein.
Os03g0783100	152	NA	Conserved hypothetical protein.
Os04g0267000	152	NA	Putative uncharacterized protein.
Os01g0641200	152	NA	Ubiquitin-like protein-NEDD8-like protein RUB3.
Os02g0653200	225	AT3G60270	Cupredoxin domain containing protein.
Os01g0170300	225	AT5G56040/SKM2	Regulates fertility in Arabidopsis.
Os01g0825000	225	NA	Similar to LOB domain protein 12.
Os12g0242700	225	AT1G24360	Similar to 3-oxoacyl-[acyl-carrier-protein] reductase 1.
Os01g0112400	225	AT2G23910	Major intrinsic protein family protein.
Os04g0543700	225	AT3G63095	Similar to Serine proteinase.
Os06g0611400	385	AT4G18180	Pectin lyase fold/virulence factor domain containing protein.
Os03g0145900	385	AT5G22920	Similar to CHY zinc finger family protein, expressed.
Os03g0701400	385	AT2G13600	Similar to Pentatricopeptide repeat protein PPR868-14.
Os02g0790600	385	AT2G18650/MEE16	**Involved in embryo development in Arabidopsis.**

Unlike Clade 1.1 and its associated Cluster in Arabidopsis FlowerNet (Cluster 37), which have conserved expression patterns and function between them, genes in Clade 1.2 did not show similar expression pattern in rice as their putative orthologs in the Arabidopsis FlowerNet. Most of the rice genes in Clade 1.2 show anther-specific expression (Figure 3-11), while their orthologs

in Arabidopsis are expressed throughout the plant and at different developmental stages (e-FP Browser, www.arabidopsis.org). These findings suggest that although pollen regulatory networks are similar between monocots and dicots that there are still significant differences between them. It may be that pollen development involves more integrative coordination and ome different sets of genes in rice, as compared with Arabidopsis.

Os10g0524500 *(OsNP1, No Pollen 1)* encodes a putative glucose–methanol–choline oxidoreductase that is required for tapetum degradation and male fertility (Chang et al., 2016; Liu et al., 2017). Recent research has shown that there is a requirement in rice for development of Ubisch bodies in the anther tapetum for exine patterning (Liu et al., 2017). However the Arabidopsis ortholog AT1G12570, is expressed both in anther and seeds (TAIR) (Rhee et al., 2003a). This finding suggests a distinct process for pollen development in rice in which RiceAntherNet can help to enable new gene identification. Accordingly, Os12g0242700, in Cluster 225, encodes a β-ketoacyl-[ACP] reductase (KAR), using NADPH as the electron donor to reduce 3-ketoacyl-ACP into 3-hydroxyacyl-ACP, which is an important intermediate for fatty acid biosynthesis (Li-Beisson et al., 2013; Mou et al., 2000). The Arabidopsis *KAR* gene, At1g24360, shows high conservation to a *Brassica napus* seed 3-oxoacyl-ACP reductase, which has been reported as functioning as a fatty acid synthesis reductase (Sheldon et al., 1992). Os12g0242700 is specifically expressed in rice anthers, it might be that it is a fatty acid synthesis

gene required for anther development; preliminary data from a *kar* RNAi line suggests failures in pollen development (unpublished data from Zhang's lab).

OsCP1 (Oryza sativa Cysteine Protease1) encodes a cysteine protease activated by TDR, which is required for tapetum PCD (Lee et al., 2004; Li et al., 2006). A putative ortholog in Arabidopsis is reported for PCD during seed development (Andeme Ondzighi et al., 2008), however no role has been reported during pollen development. This finding further support the idea that rice has distinct and possibly more complicated aspects of its pollen development. In addition to cysteine proteases, aspartyl proteases are of vital importance for tapetum cell death (Niu et al., 2013). Os04g0448500 is a eukaryotic aspartyl protease family protein, specifically expressed during tapetum PCD stage, which is co-expressed with *OsCP1*. However, the homologous gene to Os04g0448500 in Arabidopsis (AT2G03200) is not expressed in anther. Here genes have been identified that are specifically expressed during rice anther development, which are not conserved in Arabidopsis, consistently orthologs of these genes show no evidence that they are required for pollen development in Arabidopsis. These findings suggest that although there is conservation in many regulatory aspects of pollen development in rice and Arabidopsis, but there is also clear distinction in the developmental networks. Therefore RiceAntherNet serves as a valuable tool to help identify new genes that are involved in pollen development in rice, and potentially other monocots, for future analysis.

Os05g0169600, Os01g0170300 are Leucine-rich receptor-like (LRR) protein kinases down regulated in *eat1-1, gamyb-2, tdr* and *ptc1* mutants. They share the putative ortholog AT5G56040 in Arabidopsis that encodes a Leucine-rich receptor-like protein kinase named as STERILITY-REGULATING KINASE MEMBER2 (SKM2) (Endo et al., 2013). *SKM2* specifically expressed in stipules and pollen grains. SKM2, together with SKM1, act as receptors for a CLV3/ESR-related peptide, CLE45, forming a CLE45-SKM1/SKM2 signaling pathway to regulate fertility in Arabidopsis (Endo et al., 2013). Therefore, Os05g0169600, Os01g0170300 might be required for fertility in rice.

The ortholog of Os04g0613200, *QRT3 (QUARTET 3)* is a polygalacturonase, specifically expressed in the tapetum when microspores separate from their meiotic siblings in Arabidopsis (Rhee et al., 2003b). *qrt3* microspores fail to separate and remain attached in a characteristic tetrahedral cluster of four pollen grains suggesting that *QRT3* is required for pollen mother cell wall degradation. *QRT3* is specifically expressed in the tetrad stages (stage 8) in anthers (Rhee et al., 2003b)(Figure 3-11). *QRT3* is largely reduced in *dyt1-3*, a callose degradation defective mutant (Zhu et al., 2015). Compared with Arabidopsis, Os04g0613200 shows a similar expression pattern in rice, which is highly reduced in those mutants in which callose remain undegraded at late stages such as *gamyb-2* (8.4 fold), *tdr* (27.9 fold), *udt1-1* (31.3 fold) and *ptc1* (9.95 fold) mutants (Appendix III). These findings suggest that Os04g0613200 might be a regulator for callose degradation in rice.

Of the 39 genes in Clade 1.2, 37 are currently unstudied, and 33 of them show different expression patterns between rice and Arabidopsis and no reported function in rice or Arabidopsis. Given their specific expression in rice anthers and the fact that they are highly co-expressed with the previously described important genes, they may also be good candidates for further research on anther development.

3.2.7 Clade 2 includes genes for anther cell division and differentiation

Genes in Clade 2 are expressed earlier than Clade 1, which is from anther stage 2 to stage 9, while genes in Clade 3 are expressed later than Clade 1, which is from stage10 to stage 12. Compared with Clade1, Clade 2 and 3 do not have high percentage of DEGs, however there are still some genes showing differential expression between mutants and wild type, which help prediction of the putative roles of Clade 2 and Clade 3 on rice anther development.

There are 137 genes in Clade 2, among which 87 are DEGs. 84 of these DEGs are highly expressed in root, shoot apical meristem and anther, which are the tissues with high meristem activity (Figure 3-14).

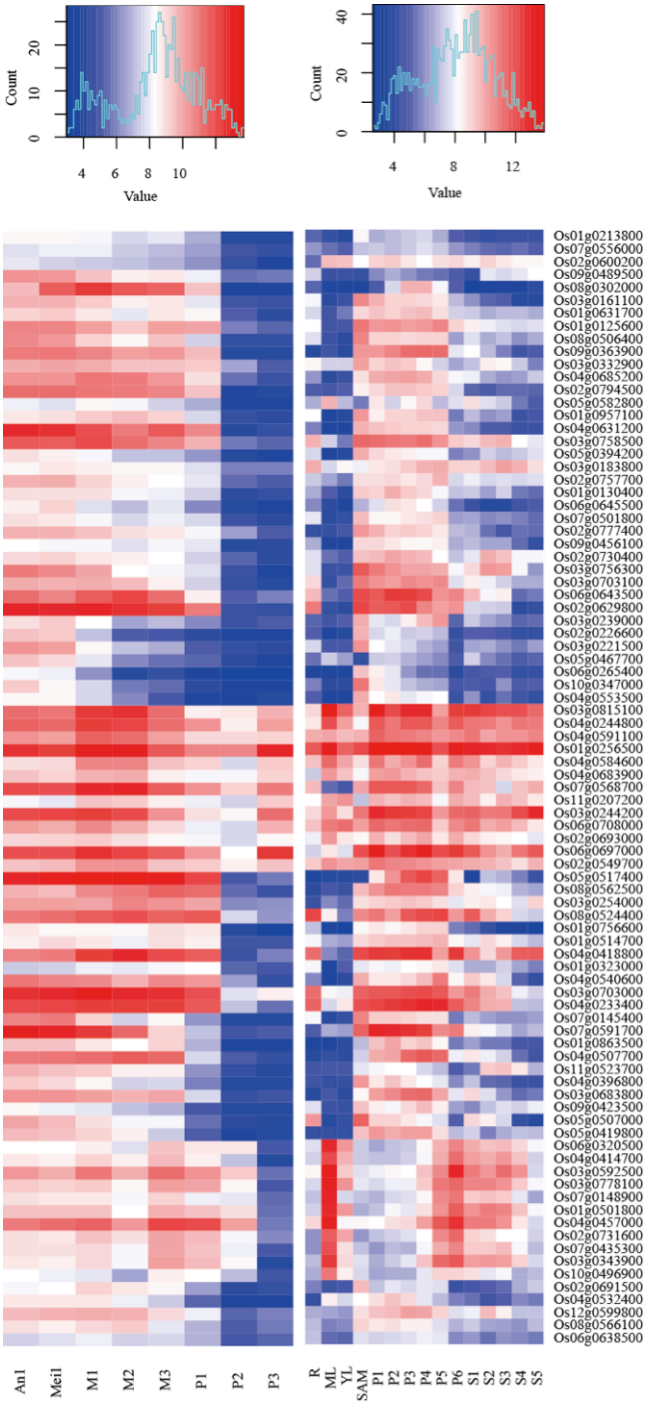


Figure 3-14 Expression patterns of Clade 2 in rice anther and other tissues. Left panel: heatmap of Clade 2 in anther (dataset GSE14304); Right panel: heatmap of Clade 2 in other tissues. Stages in left panel: An1 = Anther Stage 1-2; Mei = Anther Stage 3-6; M1 = Anther Stage 7; M2 = Anther Stage 8a; M3 = Anther Stage 8b; P1 = Anther Stage 9 and 10; P2 = Anther Stage 11; P3 = Anther Stage 12. Tissues in right panel: R = Root; ML = Mature Leaf; YL = Young Leaf; SAM = Shoot Apical Meristem; P1 = Inflorescence upto 3 cm; P2 = Inflorescence, 3-5 cm; P3 = Inflorescence, 5-10 cm; P4 = Inflorescence, 10-15 cm; P5 = Inflorescence, 15-22 cm; S1 = Seed, 0 - 2 DAP; S2 = Seed, 3-4 DAP; S3 = Seed, 5 - 10 DAP; S4 = Seed, 11 - 20 DAP; S5 = Seed, 21 - 29 DAP.

Surprisingly, cluster 8 in Clade 2 contains eight LRR-RLK (Leucine-rich repeat receptor-like kinase) genes (Figure 3-15 shows the phylogenetic analysis of these genes). Of these eight genes, Os03g0773700 and Os07g0134200 are orthologs of Arabidopsis *BAM1*, which is required for anther cell differentiation (Hord et al., 2006); while another early Arabidopsis anther development-associated gene, *ER (ERECTA)* (Hord et al., 2008) has two orthologs in rice, Os06g0203800 and Os02g0777400, which are both also in cluster 8. These findings suggest that the LRR-RLKs clustered in cluster 8 might be required for early anther development. The remaining 4 LRR-RLKs have not been reported either in rice or in Arabidopsis, but one can speculate that these genes might be required for early anther development stages. The consistent expression profile of genes in Clade 2 supports that they might have conserved function in anther cell division and differentiation.

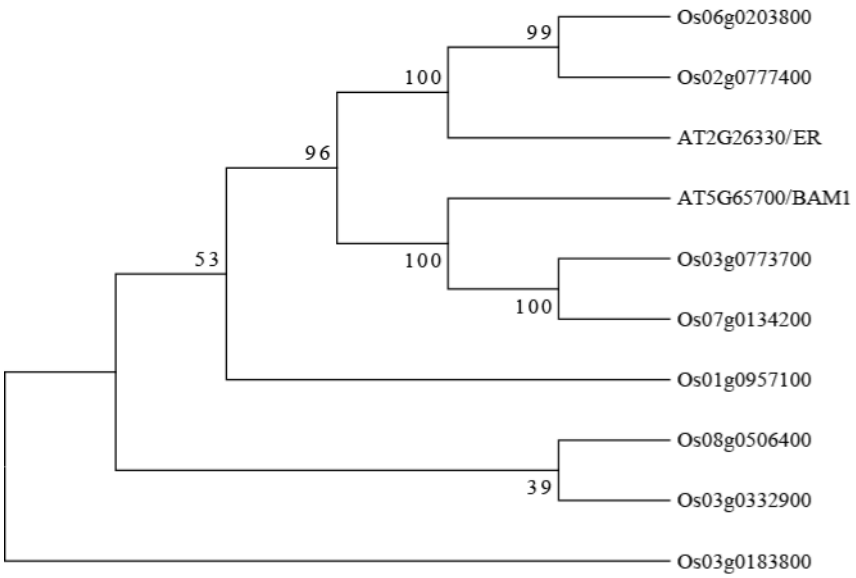


Figure 3-15 Phylogenetic tree of LRR-RLK genes in Clade 1. LRR-RLK protein sequences selected in rice and Arabidopsis are shown in Appendix IV.

3.2.8 Clade 3 includes genes expressed at the later stages of pollen development

Among the 572 genes in Clade 3 which are expressed in the late stage of anther development, 139 are DEGs and 106 of them are anther specific genes (Figure 3-16), suggesting that these genes play specific role on pollen development at the later stages. Os01g0939100 in Clade 3, is the putative ortholog of *ACA11* in Arabidopsis, is specifically in anther in the later stages. *ACA11* is type V p-type ATPase expressed in pollen, required for ion homeostasis and pollen development (Jakobsen et al., 2005), and so Os01g0939100 might be required for pollen development in rice.

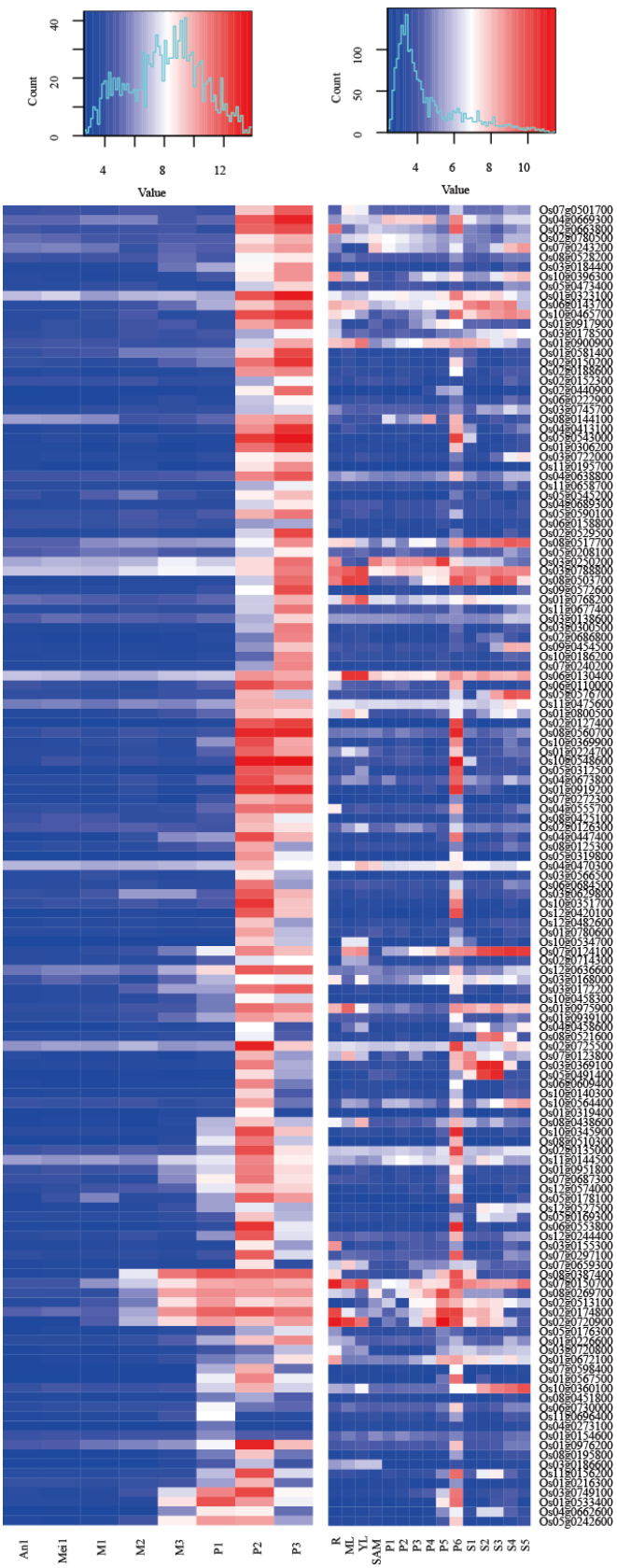


Figure 3-16 Expression patterns of Clade 3 in rice anther and other tissues. Left panel: heatmap of Clade 3 in anther (dataset GSE14304); Right panel:

heatmap of Clade 3 in other tissues. Stages in left panel: An1 = Anther Stage 1-2; Mei = Anther Stage 3-6; M1 = Anther Stage 7; M2 = Anther Stage 8a; M3 = Anther Stage 8b; P1 = Anther Stage 9 and 10; P2 = Anther Stage 11; P3 = Anther Stage 12. Tissues in right panel: R = Root; ML = Mature Leaf; YL = Young Leaf; SAM = Shoot Apical Meristem; P1 = Inflorescence upto 3 cm; P2 = Inflorescence, 3 - 5 cm; P3 = Inflorescence, 5 - 10 cm; P4 = Inflorescence, 10 - 15 cm; P5 = Inflorescence, 15 - 22 cm; S1 = Seed, 0 - 2 DAP; S2 = Seed, 3 - 4 DAP; S3 = Seed, 5 - 10 DAP; S4 = Seed , 11 - 20 DAP; S5 = Seed, 21 - 29 DAP.

3.2.9 Database for gene expression pattern query

A database containing gene expression patterns of 24,462 genes in rice anther and their co-expression network was built, which is available from a freely accessible webpage (https://www.cpib.ac.uk/anther/ riceindex.html). Figure 3-17 shows a brief functional framework of this database. Annotation, plot of expression of query gene are searchable by either RAP or MSU identifier. Figure 3-3 shows the plot of *TDR* as an example. This tool also shows whether it is in RiceAntherNet and which genes are co-expressed with query gene. When querying by cluster, genes and annotation in this cluster will be shown. Gene expression plots of all genes in the cluster can also be downloaded. This database thus provides a global gene expression profile and co-expression network in rice anther.

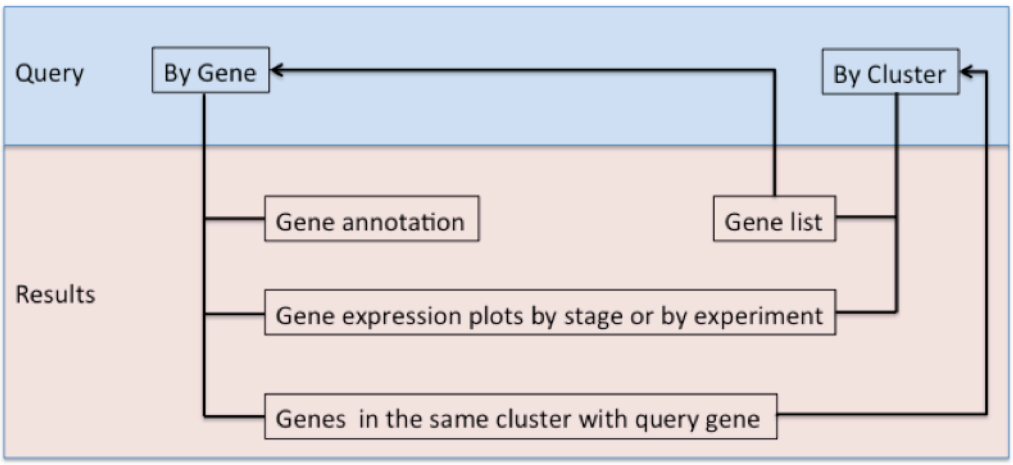


Figure 3-17 Functional framework of RiceAntherNet database.

3.3 DISCUSSION

3.3.1 RiceAntherNet is a useful tool for gene functional study during rice anther and pollen development

RiceAntherNet is a rice co-expression network which identifies clusters of genes that are co-expressed in anther tissues across a number of developmental time points, and may therefore show closely functionally-related roles. This network provides a global view of gene expression patterns throughout rice anther development, which can also be compared to that in other species, in particular Arabidopsis.

Datasets of 9 male sterile mutants were mapped on to the correlation network to identify genes with differential expression patterns associated with the various male sterile mutants as compared to wild type. RiceAntherNet is an important tool for demonstrating coordinated network regulation during anther development. Furthermore, the website is available as a resource for the rice

community, allowing immediate conformation of the gene expression profile in anthers for a gene of interest.

3.3.2 Processes of sporopollenin synthesis and transport are conserved in rice and Arabidopsis

Compared with Arabidopsis and rice, genes in RiceAntherNet Clade 1.1 also show co-expressed patterns in FlowerNet in Arabidopsis (Pearce et al., 2015). Based on previous research, 7 genes (*OsABCG15, DPW, OsLAP6, CYP703A3, OsLAP5, OsNP1, OsCP1*) out of 21 in Clade 1.1 are believed to be involved in pollen wall formation in rice. 12 genes in the RiceAntherNet have orthologs in Arabidopsis that have been reported to have a role in pollen development, suggesting a conserved function in rice. This illustrates that the gene co-expression network generated by array data is a valuable tool to identify new genes involved in related process and to predict novel gene functions. The similarity between the known gene regulation networks in Arabidopsis and rice during pollen development has been indicated previously (Wilson & Zhang, 2009). This further validates the reliability of both RiceAntherNet and the Arabidopsis FlowerNet for new gene discovery.

3.3.3 Rice has gene regulation network that is partly not conserved compared to Arabidopsis

Genes in RiceAntherNet Clade 1.2 show distinguished expression patterns in Arabidopsis and rice (Pearce et al., 2015). Genes in Clade 1.2 have both anther

specificity and stage specificity (expressed in anther stage 8 and 9) in their expression pattern in rice; while their orthologs in Arabidopsis are widely expressed. Two genes (*OsNP1 and OsCP1*) in this clade are required for tapetum degradation and male fertility (Lee et al., 2004; Liu et al., 2017), however no equivalent role during tapetum degradation process has been found for their putative orthologs in Arabidopsis. This finding suggests that gene regulation networks are not entirely conserved during tapetum PCD in rice and Arabidopsis.

3.4 THE IMPORTANCE OF THIS WORK

RiceAntherNet furthers our understanding of gene regulation during anther development, as a reverse genetic resource for novel gene identification during pollen development in rice. Compared to previous co-expression networks in rice, this work focuses on anther and pollen development, and RiceAntherNet particularly shows the gene expression pattern in anther throughout all anther stages. As opposed to the chain-like clusters commonly generated (Aya et al., 2011), the tightly connected clusters in RiceAntherNet commonly finds genes which are involved in the same process, meaning that one does not need to know the starting target genes to predict genes with related functions. Although in this analysis only novel genes involved in pollen wall formation were predicted by using mutants that are defective in pollen wall formation, RiceAntherNet is an efficient and reliable tool that can be used to predict novel

genes involved in other pathways. It would be interesting to target other example clades related to other functions by mapping DEGs in different mutants of interest. This can be readily achieved by using the resources generated to map against genes of interest from specific mutants or conditions, to identify target clusters for future analysis. This therefore broadens the application of these resources beyond pollen development and thus provides a valuable reverse genetic resource for functional genetic analysis.

4 VALIDATION OF TARGETS FROM ANTHERNET

4.1 INTRODUCTION

According to RiceAntherNet in Chapter 3, 53 genes of unknown functions were predicted involving in rice pollen development regulation (section 3.2.6). A validation of the functions of those candidates is required to prove the reliability of RiceAntherNet, as well as to deepen the understanding of rice pollen development. Forward genetics (traditional genetics) is a main way for identifying gene functions based on the characterization of mutants. In rice reproduction research, almost all gene functions revealed to date have been studied by forward genetics (section 1.1), which required a mutant with an identified phenotype as a starting point, followed by gene mapping and cloning. A large part of the gene regulation network associated with plant reproduction could however be missed as a consequence of the lack of mutants with clear phenotype (as some phenotypes are hard to observe by mass screening). The identification and functional analysis of redundant genes, where the mutation of a single locus does not lead to obvious phenotypes, would also be outside the scope of forward genetics approaches. In recent years, the CRISPR/Cas9 system has been developed for targeted gene editing by employing a Cas9 endonuclease which binds to a guide RNA complex (Zhang et al., 2014). As the CRISPR/Cas9 system is developing rapidly, it becomes more and more convenient to obtain mutations in target genes in rice

(Zhang et al., 2014). With the usage of this efficient gene-editing tool, reverse genetics has the potential to be a productive way to discovery genes with unknown functions (Shah et al., 2015). In this chapter, 9 targets predicted by RiceAntherNet were selected for gene function study. Mutaions in these genes were obtained using the CRISPR/Cas9 system (see method in section 2.5) in the 9522 genetic background for further investigation.

4.2 RESULTS

Among total 53 genes predicted by RiceAntherNet involving in rice pollen development, more attention is paid to those genes which might effect tapetum formation and degredation because normal tapetum behaviors are essential for both microspore development and pollen wall formation (section 1.1.2). As described in section 1.2, Ca^{2+} signal and lipid synthesis can regulate the tapetum PCD process. In 53 candidates, *calmodulin like 41 (CML41)* is annotated as a CML protein containing Ca^{2+} -binding sites which might related to Ca^{2+} signaling during pollen development. Meanwhile, *Dihydroflavonol-4-reductase Like1 (OSDRL1), Dihydroflavonol-4-reductase Like2 (OSDRL2), Glucose-methanolcholine oxidoreductase (GMCO)* and *POLYKETIDESYNTHASEA (OsLAP6)* are encode enzymes participate lipid synthesis pathways. In addition, two LRR family members, *LRR1* and *LRR5,* were predicted as involved in cell differentiation for tapetum iniciation because their expressed in anther as well as in other tissues with strong cell

division and differentiation, such as root, SAM and inflorescence.

The annotations and predicted functions of these 7 genes from clade1 are listed in Table 4-1. All of them showed significantly changed expression in male sterile mutants according to microarray data and have been highlighted as genes of interest in RiceAntherNet (Section 3.2.5) (Table 4-2).

Table 4-1 Annotations of candidate genes

Gene	RAPID	Annotation	Predicted Function
CML41	Os03g0769500	EF-Hand type domain containing protein	Ca^{2+} signaling
GMCO	Os02g0678300	Similar to protein HOTHEAD	Lipids Synthesis
OSDRL1	Os08g0515900	TKPR, flavonoid biosynthesis	Lipids Synthesis
OSDRL2	Os01g0127500	TKPR, flavonoid biosynthesis	Lipids Synthesis
OsLAP6	Os10g0484800	Polyketide synthase, type III domain containing protein.	Lipids Synthesis
LRR1	Os01g0170300	Serine/threonine protein kinase domain containing protein	Cell Differentiation
LRR5	Os05g0169600	Serine/threonine protein kinase domain containing protein	Cell Differentiation

Table 4-2 Fold change of candidate genes in male sterile mutants. S = Anther Stage. See full table of all DEGs in Appendix III.

Gene	*eat1-1* *S10*	*gamyb-2* *S9*	*tdr* *S8b*	*tip2* *S7*	*udt1-1* *S9*	*ptc1* *S9*	*mads3-4* *S9*	*osdex1* *S9*
CML41	-4.2	-14.6	-58.2		-2.6			-35.2
GMCO	-15.5		-194.5	-3	-73.4			
OSDRL1			-115.1	-4.3	-52.9		21.9	
OSDRL2		-12.3	-105.2		-11.5			
OSLAP6		-24.5	-266.6				6.9	
LRR1	-11.2	-57.1	-5.9		-4.3	-22.4		
LRR5	-18.7	-56.9	-12.4			-104.3	2.6	

Figure 4-1 shows the expression patterns of 7 candidate genes in rice anther as well as in other tissues. Figure 4-1A shows that all of them are highly expressed specifically in anther at stage 8 and 9; Figure 4-1B shows that a number of these genes *(CML41, GMCO, OSDRL1, OSDRL2* and *OSLAP6),* which are predicted to be associated with tapetum PCD and lipid synthesis, had tissue specific expression in the anther, while others, *LRR1* and *LRR5,* predicted as involved in cell differentiation, are expressed not only in anther but also in root, SAM and inflorescence. All the information regarding these genes, including gene annotations, expression patterns and gene performance in mutants, supports the hypothesis that these candidates might play a role in rice anther development and therefore they were selected for further study.

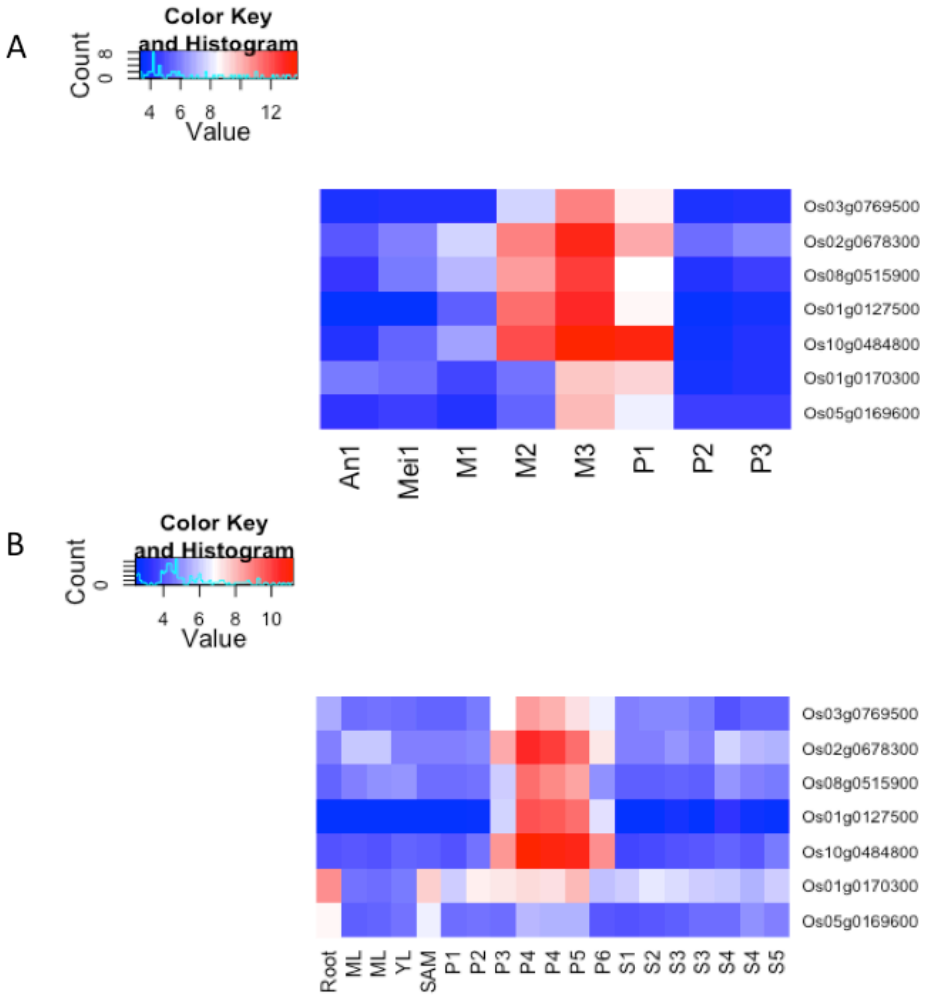


Figure 4-1 Expression patterns of 7 candidate genes in rice anther and other tissues. A: heatmap of candidate genes in anther tissues (dataset GSE14304); B: heatmap of candidate genes in other tissues. Stages in A: An1 = Anther Stage 1-2; Mei = Anther Stage 3-6; M1 = Anther Stage 7; M2 = Anther Stage 8a; M3 = Anther Stage 8b; P1 = Anther Stage 9 and 10; P2 = Anther Stage 11; P3 = Anther Stage 12. Tissues in B: R = Root; ML = Mature Leaf; YL = Young Leaf; SAM = Shoot Apical Meristem; P1 = Inflorescence up to 3 cm; P2 = Inflorescence, 3 - 5 cm; P3 = Inflorescence, 5 - 10 cm; P4 = Inflorescence, 10 - 15 cm; P5 = Inflorescence, 15 - 22 cm; S1 = Seed, 0 - 2 DAP (Days After Pollination); S2 = Seed, 3 - 4 DAP; S3 = Seed, 5 - 10 DAP; S4 = Seed, 11 - 20 DAP; S5 = Seed, 21 - 29 DAP.

In addition, the MYB transcription factor *(GAMYB Like 2) GAMYBL2* (Os03g0578900) (section 4.2.4) was selected as important gene to look at because of the high co-expression with both *LRR1* (PCC = 0.90538) and *LRR5*

(PCC = 0.92529).

In summary, seven candidate genes (*CML41* (Os03g0769500) (section 4.2.1), *GMCO* (Os02g0678300) (section 4.2.2), *OSDRL1* (Os08g0515900) (section 4.2.3), *OSDRL2* (Os01g0127500) (section 4.2.3), *OsLAP6* (Os10g0484800) (section 4.2.5), *LRR1* (Os01g0170300) (section 4.2.5) and *LRR5* (Os05g0169600) (section 4.2.5)) in RiceAntherNet clade 1 were chosen for further study of their functions during anther and pollen development in rice.

In addition, the function of *LOB DOMAIN 2 (LBD2)* (At1g06280), the ortholog of candidate gene Os01g0242400 in RiceAntherNet clade1 was studied in Arabidopsis (section 4.2.6).

4.2.1 *Calmodulin like 41 (CML41)*

CML41 is an anther-specific expressed gene which encodes a CML protein. According to RiceAntherNet, *CML41* is one of the members in Clade 1.1, which is co-expressed with lipid synthesis and transport genes *OsABCG15*, *OsLAP5*, *CYP703A3* and *DPW*. It is significantly down regulated in male sterile mutants *tdr*, *eat1-1* and *osdex1* (Table 4-2).

4.2.1.1 Molecular Identification of *CML41*

The molecular structure of *CML41* was investigated to provide a better understanding of this gene. *CML41* encodes a 185-amino acid peptide which contains 2 D-x-D-x-D-G loops as Ca^{2+} -binding sites (Rigden & Galperin, 2004) (Figure 4-2A). As *CML41* is largely down regulated in mutants of

bHLH transcription factor family, including *eat1* and *tdr*, the 3487 bp-upstream promoter sequence of *CML41* was screened for the presence of the conserved binding motif of the bHLH family, E-box (5'-CANNTG-3') (Li et al., 2015) . Figure 4-2B shows that in the 3487 bp promoter sequence of *CML41*, 28 E-boxes (only 13.6 E-boxes were expected in a random nucleotide sequence of 3487 bp) were found and 9 of them were particularly clustered in the last 900 bp before the transcriptional start site. Based on the molecular identification results and the presence of the E-box motifs, it was hypothesized that *CML41* might control Ca^{2+} homeostasis under the regulation of these bHLH family proteins.

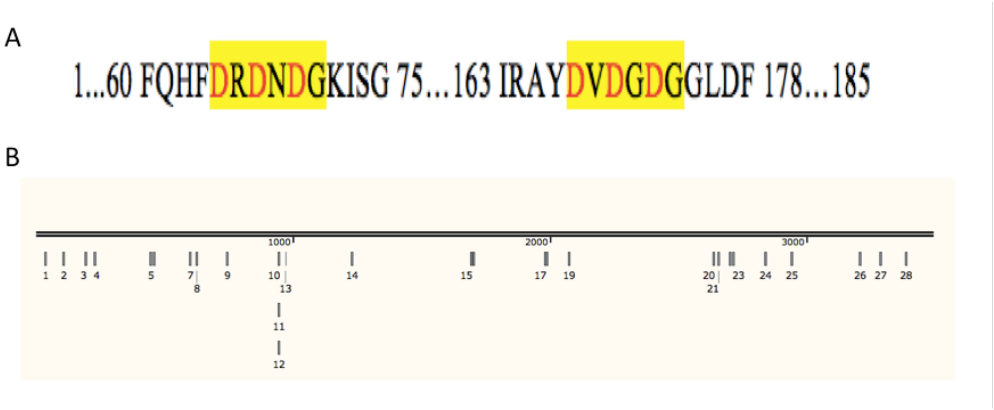


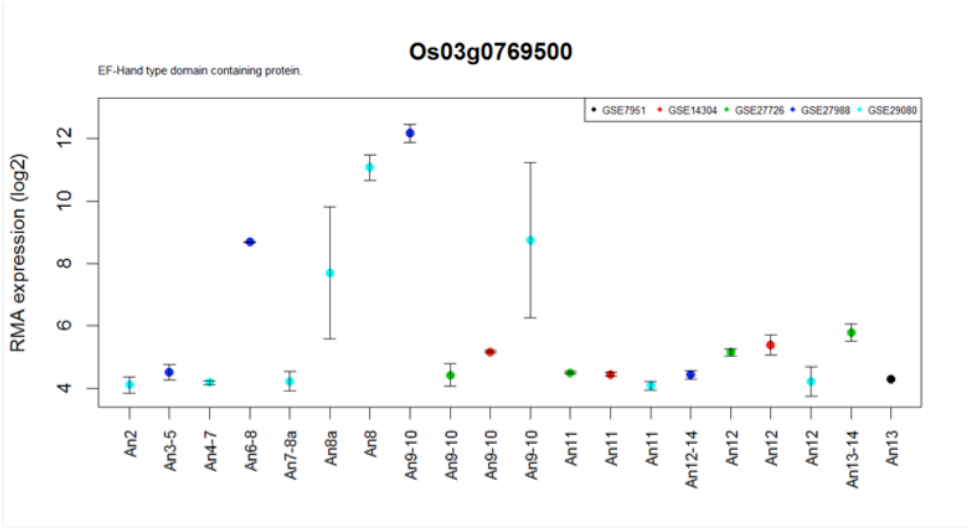
Figure 4-2 Molecular Identification of *CML41*. A: peptide structure with the 2 D-x-D-x-D-G loops; B: promoter structure with 28 E-boxes. Numbers and positions of E-boxes are labeled under the sequence map.

4.2.1.2 Expression analysis of *CML41*

An expression plot from RiceAntherNet indicated that *CML41* is expressed only in anther stage 8 and 9 (Figure 4-3A). Quantitative reverse transcriptase PCR (qRT-PCR) (Section 2.6.3) of wild type (9522) anther tissues at stage 8 to

stage 10, was performed to check the expression of *CML41* using the primers CML41_RT_F and CML41_RT_R (Appendix II and Figure 4-6). The results confirmed that *CML41* was highly expressed at anther stage 8b and decreased rapidly at stage 9 (Figure 4-3B). No expression of *CML41* was detected at stage 8a and stage 10.

A



B

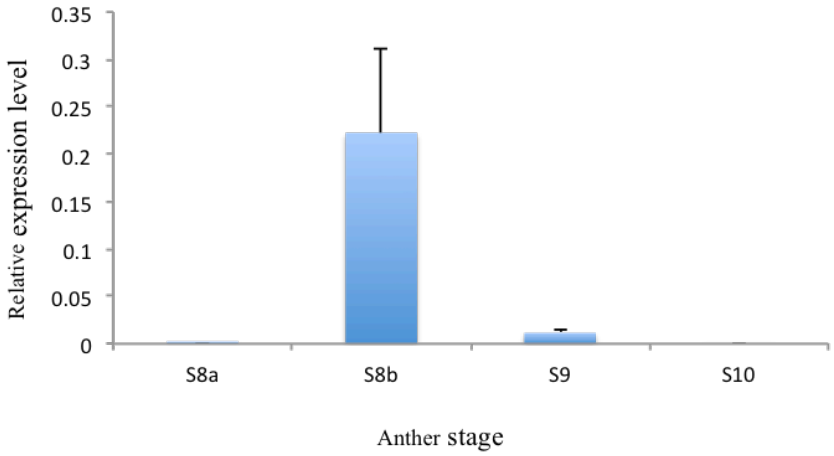


Figure 4-3 The expression of *CML41* in WT anther. A: expression plot of *CML41* (RiceAntherNet: https://www.cpib.ac.uk/anther/riceindex.html); B: qRT-PCR of *CML41*. Error bar: standard error. Data was normalized to the reference gene *OsACTIN*.

As discussed in the introduction (section 1.2.3), 33 CMLs have been identified in rice. Among them, the expression patterns of 24 CMLs were available in

RiceAntherNet (Figure 4-4). 7 CMLs (Os05g0223000, Os01g0949500, Os03g0331700, Os08g0117400, Os01g0135700, Os05g0380900 and Os01g0955100) are expressed in anther through out all anther stages (Figure 4-4A), as well as widely expressed in root, leaf, inflorescence and seed (Figure 4-4B). These 7 genes may have redundant functions in rice pollen development alongside *CML41*. However, among all of the CML genes in rice, *CML41* is the only anther-specific and anther stage 9-specific expressed gene in the rice CML family members. The very specific expression of *CML41* in rice anther suggests that it may play and important role in regulating Ca^{2+} homeostasis in tapetum cells.

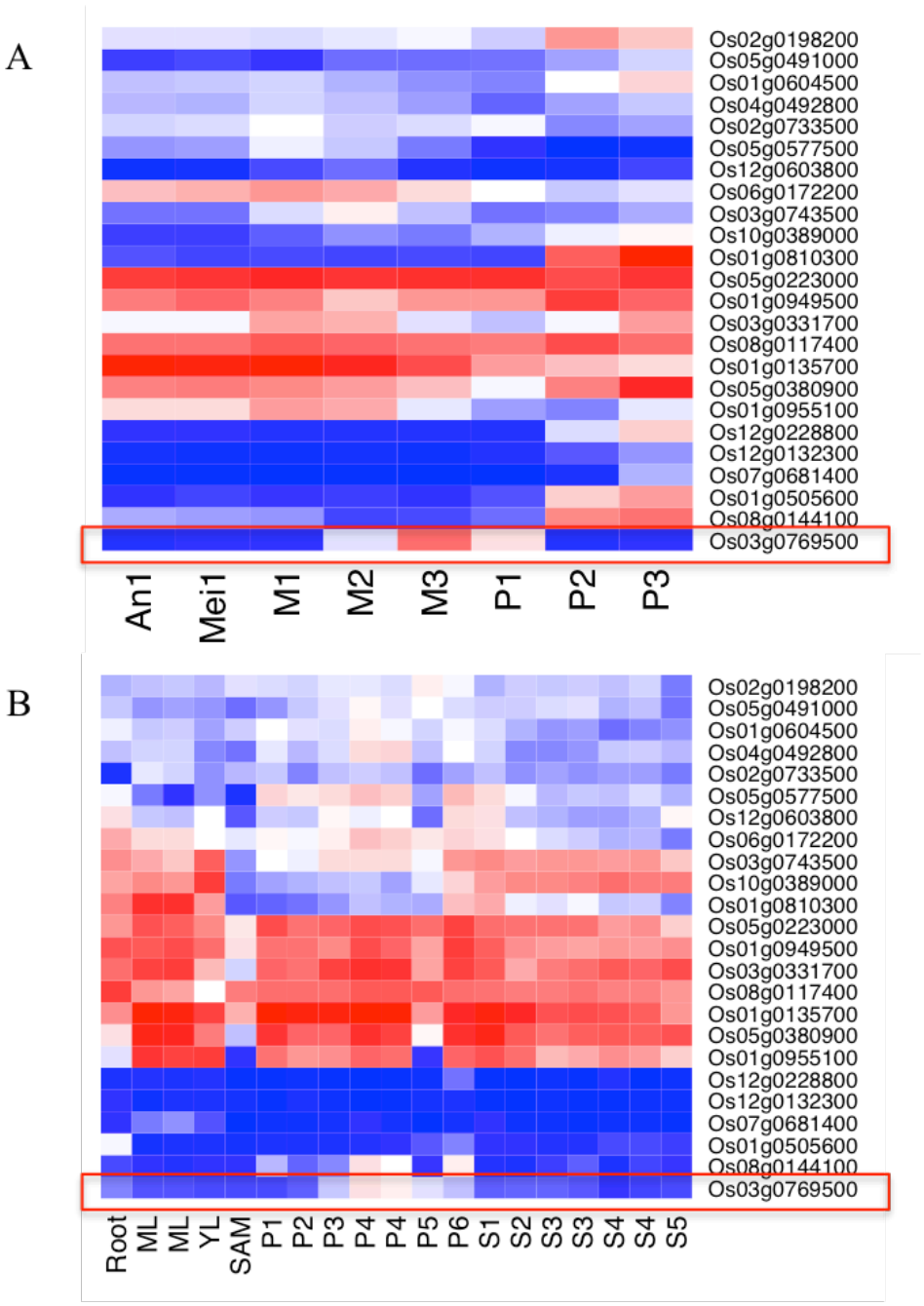


Figure 4-4 Expression patterns of 24 CML genes in rice anther and other tissues. A: heatmap of candidate genes in anther tissues (dataset GSE14304); B: heatmap of candidate genes in other tissues. Stages in A: An1 = Anther Stage 1-2; Mei = Anther Stage 3-6; M1 = Anther Stage 7; M2 = Anther Stage 8a; M3 = Anther Stage 8b; P1 = Anther Stage 9 and 10; P2 = Anther Stage 11; P3 = Anther Stage 12. Tissues in B: R = Root; ML = Mature Leaf; YL = Young Leaf; SAM = Shoot Apical Meristem; P1 = Inflorescence up to 3 cm; P2 = Inflorescence, 3 - 5 cm; P3 = Inflorescence, 5 - 10 cm; P4 = Inflorescence, 10 - 15 cm; P5 = Inflorescence, 15 - 22 cm; S1 = Seed, 0 - 2 DAP (Days After Pollination); S2 = Seed, 3 - 4 DAP; S3 = Seed, 5 - 10 DAP; S4 = Seed, 11 - 20 DAP; S5 = Seed, 21 - 29 DAP. Expression patterns of CML41 are highlighted in the red frame.

4.2.1.3 *CML41* may be regulated directly by bHLH family proteins

Previous promoter analysis of *CML41* shows that the promoter of *CML41* is enriched in the bHLH family protein binding site, the E-box. *CML41* was also significantly down regulated in the tapetum bHLH family mutants *eat1* and *tdr* (Table 4-2). Therefore, the Dual-Luciferase Reporter Assay (section 2.7) was performed to explore the regulatory relationship between *CML41* and the bHLH family proteins (Figure 4-5).

Figure 4-5A shows the schematic diagrams of the reporter plasmid and effector plasmids used in the transient expression mediated by *Agrobacterium* in tobacco leaves (section 2.4). The reporter plasmid containing the *CML41* promoter sequence (3 kb) and the firefly luciferase gene (Luc) was transiently expressed by *Agrobacterium* infiltration in tobacco leaves with the effector plasmids, with *TIP2*, *TDR* and *EAT1* genes under the control of the CaMV35S promoter. The primers for amplifying the *CML41* promoter sequence are listed in Appendix II and the effector plasmids used in this work were obtained from Zhang's lab (Fu et al., 2014). Four independent transient transformations were performed for each treatment. The data were analyzed according to method in section 2.7 The results show that the bHLH proteins EAT1 and TDR1 are positively regulating the transcription of *CML41* individually, while TIP2 is not required for the regulation (Figure 4-5B). The interaction of TIP2, TDR and EAT1 does not appear to be required for activation of luciferase in the

transient reporter system and therefore are not required for the regulation of *CML41* transcription. These results are consistent with the hypothesis that *CML41* expression is largely reduced in the *tdr* and *eat1* mutants, but have no change in the *tip2* mutant (Table 4-2). In summary, the E-box analysis, microarray data and Dual-Luciferase Reporter Assay supports the hypothesis that *CML41* is a direct downstream target of TDR and EAT1.

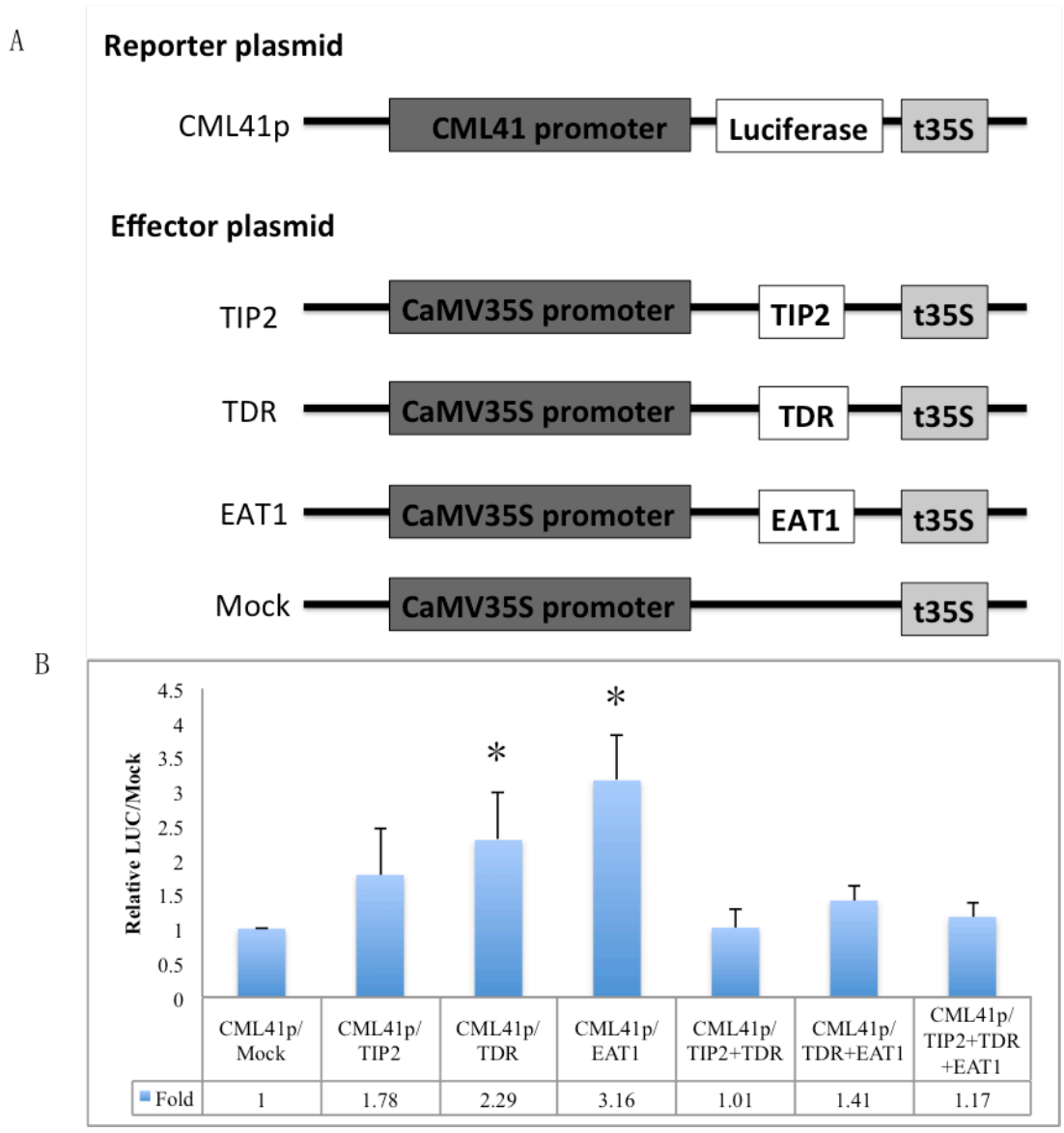


Figure 4-5 Regulation of the *CML41* Promoter by TIP2, TDR and EAT1. A: Schematic diagrams of the reporter plasmids and effector plasmids used in the transient expression in tobacco leaves; B: Transactivation of the Luc reporter gene by TIP2, TDR and EAT1. Error bar: standard deviation. Specific promoter regions of CML41 show significant activation (highlight by asterisk) by TDR and EAT1. Four repeats were performed for each treatment. Wilcoxon test was performed to show the significant of the regulation.

4.2.1.4 Mutants of *CML41* were obtained from CRISPR/Cas9 system

Two targets for *CML41* mutagenesis were designed for editing *CML41* by the CRISPR/Cas9 system (Figure 4-6). Transgenic plants were obtained according to the method in section 2.5. Mutation sites were amplified by PCR with

primers CML41_CriGT_F1 / CML41_CriGT_R1 and then sequenced. The PCR products were then sequenced to confirm the genotypes of the plants. 58 plants were genotyped and 56 of them were detected as WT (genotyping details are given in Appendix V). Two mutants of the same genotype with one base insertion at 23 bp position were obtained and named *cml41-1* and *cml41-2*. The *cml41-1* mutant (labeled in Figure 4-6) was backcrossed with wild type (9522) to obtain a stable mutant without the introduced Cas9 gene in the next generation.

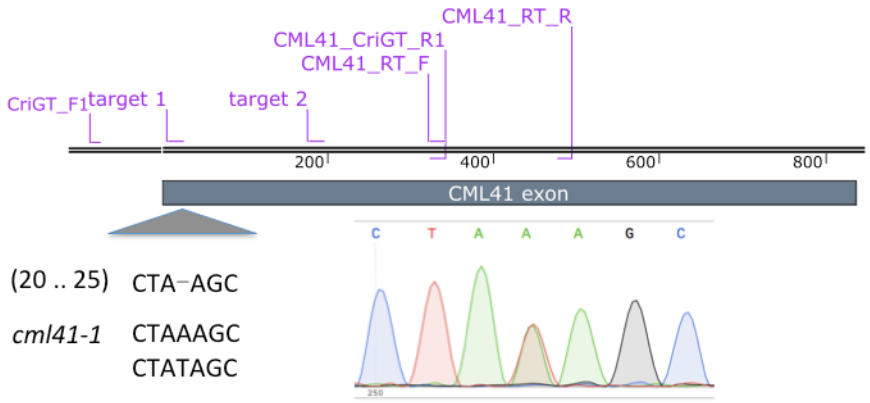


Figure 4-6 Schematic representation of *CML41* with CRISPR targets, genotyping primers and sequencing information of *cml41-1* mutant.

4.2.1.5 *cml41-1* was fertile

Anther and pollen development was analysed in the homozygous *cml41* mutant lines; iodine staining (section 2.8.1) followed by SEM (section 2.8.2) were performed to check the fertility of *cml41-1*. Figure 4-7 shows that pollen viability of *cml41-1* was normal at the mature pollen stage with about 80% pollen grains darkly stained by potassium iodide indicating a high level of viability; the pollen fertility rate in WT was about 90%.

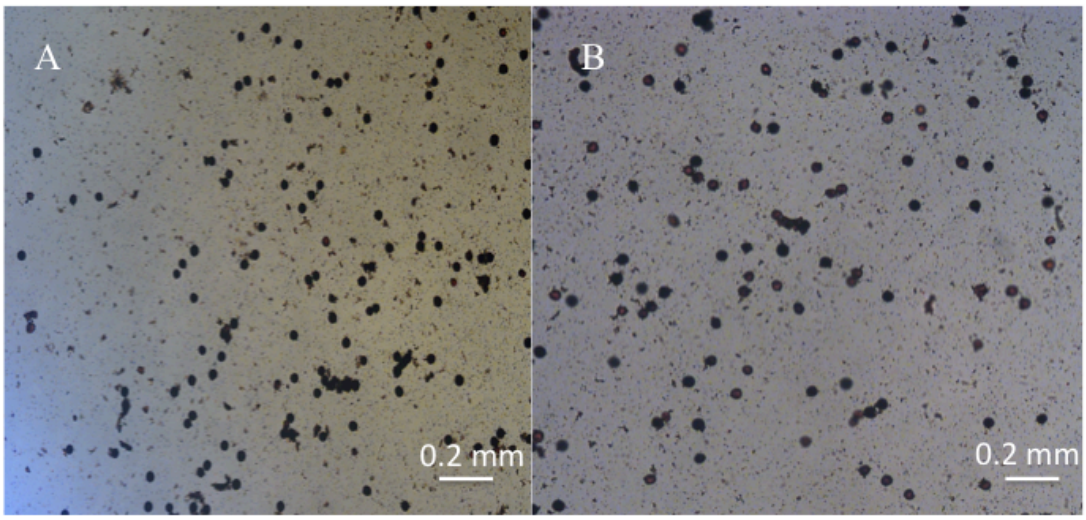


Figure 4-7 Pollen viability staining, using potassium iodine, of *cml41-1* at mature pollen stage. A: pollen from wild type (9522); B: pollen from the homozygous mutant *cml41-1*.

The morphology of anthers and pollen at anther stage 12 were observed by SEM (Figure 4-8). At stage 12, the anther of *cml41-1* was full and the pollen grains were round and of an equivalent size to the wild type (Figure 4-8A, B, G, H). The anther wall of *cml41-1* appeared similar to that of the wild type 9522 (Figure 4-8C, D); the surface of the wild type pollen was covered by evenly distributed material likely to be lipidic biopolymers (Figure 4-8I, J), however this appeared reduced in the mutant. Ubish bodies of both 9522 and *cml41-1* appeared similar, and were distributed intensively on the inner surface of the tapetum layer (Figure 4-8E, F).

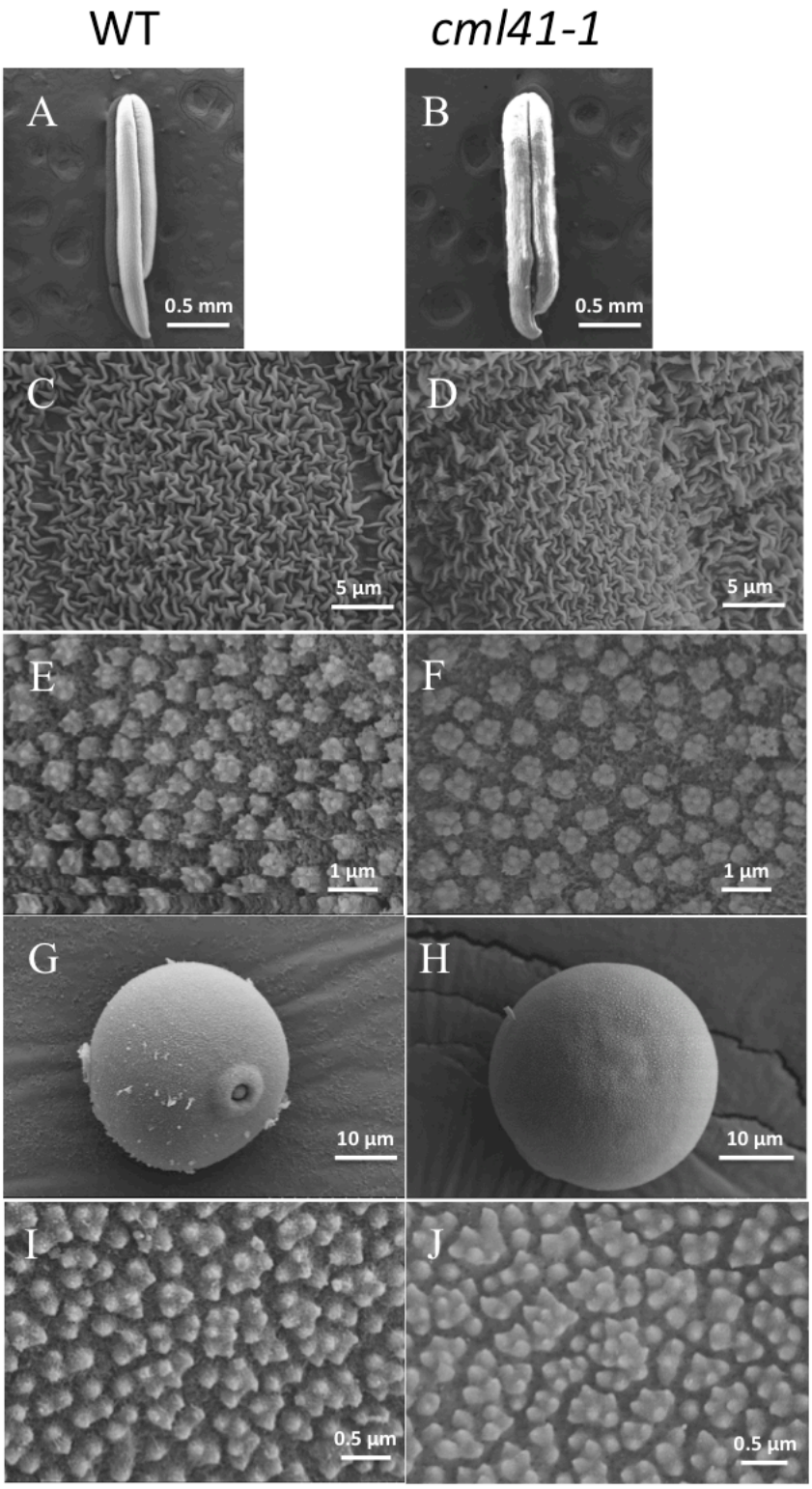


Figure 4-8 SEM pictures of anthers and pollen grains in WT and *cml41-1* at stage 12. A, B: anthers; C, D: anther wall; E, F: Ubisch bodies; G, H: pollen grains; I, J: pollen exine; left: wild type (9522); right: *cml41-1* mutant.

The phenotyping of *cml41-1* shows that although *CML41* is likely to be a direct downstream gene regulated by EAT1 and TDR, it does not effect the fertility of plant as well as the patterning of anther and pollen wall. This might be because there are 33 CMLs in rice with at least 9 of them expressed in anther at stage 8b as *CML41* (Figure 4-4A) and these CMLs might play a redundant role with *CML41* in rice. Mutation of multiple family members would therefore potentially be required to identify a clear phenotype.

4.2.2 *Glucose-methanol-choline oxidoreductase (GMCO)*

GMCO encodes a putative glucose–methanol–choline oxidoreductase in rice (Liu et al., 2017). A reported ortholog in rice, *OsNP1* is required for tapetum degradation and male fertility (Chang et al., 2016; Liu et al., 2017). *GMCO* was co-expressed with *OsNP1* in rice anther according to RiceAntherNet (section 3.2.6). Based upon the outputs from the RiceAntherNet it was predicted that *GMCO* might involved in anther and pollen wall formation by regulating lipid metabolism (Table 4-1).

4.2.2.1 Promoter of *GMCO* is enriched in bHLH family protein binding sites

Expression of *GMCO* was severely reduced in the male sterile mutants *eat1* (15.5 fold), *tdr (*194.5 fold) and *udt1-1* (73.4 fold) (Table 4-2). EAT1, TDR and UDR1 are all transcription factors in the bHLH family. Therefore, the distribution of E-box binding motifs was checked in the upstream

region/promoter of *GMCO* to evaluate the possibility of *GMCO* being directly regulated by these bHLH transcription factors. The 3000 bp-upstream promoter sequence of *GMCO* was scanned and Figure 4-9 shows that 22 E-boxes were identified as distributed in the 3000 bp promoter of *GMCO*; 12 E-boxes were identified in the 1000 up-stream sequence compared to the expected number of 3.9 in a random DNA sequence. This suggested that the expression of *GMCO* might be directly regulated by these bHLH transcriptional factors.

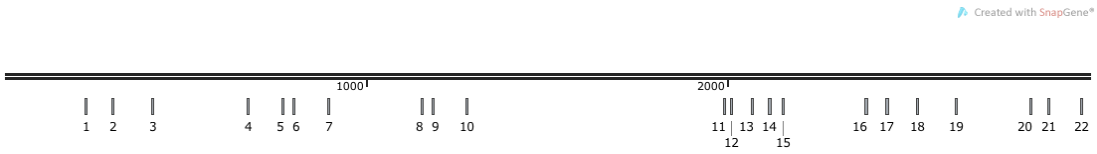


Figure 4-9 The 3000 bp Promoter/Upstream region of *GMCO*. Numbers and positions of E-boxes were labeled under the map of the promoter sequence.

4.2.2.2 Mutants of *GMCO* were obtained using the CRISPR/Cas9 system

Two targets were designed for *GMCO* for gene knocking out by CRISPR/Cas9 system (Figure 4-10). The targets were transformed into both wild type (9522) and *osnp1-4* backgrounds. Primers GMCO_CriGT_F1 / GMCO_CriGT_R1 (Figure 4-10) were used for genotyping by PCR. Ten plants were detected but unfortunately no transgenic plants in the wild type (9522) background were identified. However, one heterozygote and five mutants were identified and characterized in the *osnp1-4* background, which served to generate five double mutant alleles (mutants details in Appendix V). The *gmco-1osnp1-4* double

mutant was used for the following phenotype investigation. *gmco-1osnp1-4, gmco-2osnp1-4 and gmco-3osnp1-4* (Figure 4-10) were backcrossed with 9522 to get both single mutants of *GMCO* and stable double mutants without the Cas9 gene in the next generation.



Figure 4-10 Schematic representations of *GMCO* with CRISPR targets, genotyping primers and sequencing information of *gmco-1osnp1-4, gmco-2osnp1-4* and *gmco-3osnp1-4* mutants.

4.2.2.3 *gmco-1 osnp1-4* double mutant partially recovered the phenotype of *osnp1-4*

OsNP1 has been previously reported as involved in biosynthesis of lipid components, and mutation of *OsNP1* lead to a reduction of some cutin monomers (Liu et al., 2017). SEM was used to compare the morphology of lipidic biopolymers in the anthers of 9522 WT, *osnp1-4* mutant and the *gmco-1osnp1-4* double mutant (Figure 4-11). It was observed that, *osnp1-4* have small anthers (Figure 4-11B), collapsed pollen grains covered by big granules (Figure 4-11E), smooth epidermal surface (Figure 4-11H) and

cumulated bead- shaped Ubisch bodies (Figure 4-11K). Similar to *osnp1-4, gmco-1 osnp1-4* had small anthers and just a few shriveled pollen grains were found in anthers. However, compared with *osnp1-4,* the *gmco-1 osnp1-4* double mutant had smaller granules covering the surface of the microspores and less cumulated Ubisch bodies on the inner surface of tapetum layer. The patterns of anther cuticle were distinct between 9522, *osnp1-4* and *gmco1 osnp1-4*: typical continuous spaghetti-like cuticle surface was observed in the wild type, whereas a smooth surface without cutin was seen in *osnp1-4* and a loosely organized cuticle surface in *gmco-1osnp1-4*. It was reported that in *osnp1-4* mutant, along with the defect of cuticle development, less cutin and wax amounts were detected by gas chromatography-mass spectrometry (GC-MS) and gas chromatography-flame ionization detection (GC-FID) (Liu et al., 2017). Taken together, it may be possible that compared with *osnp1-4*, the lipidic biopolymers were reduced in *gmco-1osnp1-4,* which lead to a reduction in the accumulation of bead-shaped Ubisch bodies and smaller granules over the microspores in the double mutant; however the cuticle components were increased in *gmco-1osnp1-4* and appear to provide a partially recoved cuticle surface in the double mutant. These phenotypes suggest that *GMCO* might produce metabolites that are important for Ubisch body formation as well as pollen wall and anther wall development. The single mutant of *GMCO* is requirement to help provide deeper understanding of the function of *GMCO* and the relationship between *GMCO* and *OsNP1*.

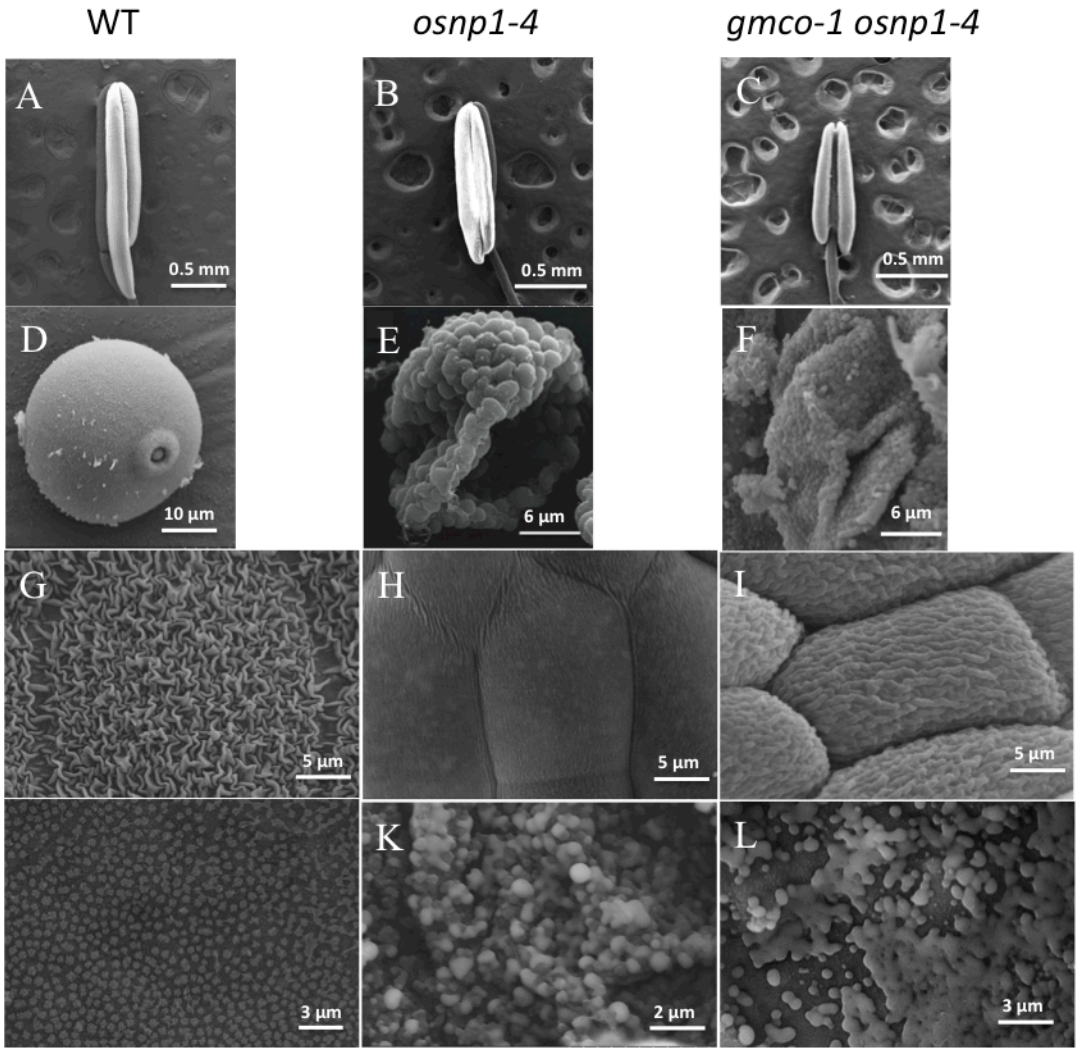


Figure 4-11 SEM analysis of WT, *osnp1-4* and *gmco-1osnp1-4* at stage 12. A, B, C: anthers; D, E, F: pollen grains; G, H, I: anther wall; J, K, L: Ubisch bodies.

4.2.3 *Dihydroflavonol-4-reductase Like1 (OSDRL1)* and *Dihydroflavonol-4-reductase Like2 (OSDRL2)*

OsDRL1 and *OsDRL2* were predicted as important genes involved in anther and pollen wall formation using RiceAntherNet (section 3.2.6.1). Both of these genes are expressed specifically in anthers at stage 8 and stage 9 (Figure 4-1). *OsDRL1* and *OsDRL2* are currently uncharacterized proteins which encode proteins similar to Dihydroflavonol-4-reductase. The orthologs of *OsDRL1*

and *OsDRL2* in Arabidopsis are *TKPR1* and *TKPR2* respectively as mentioned in section 3.2.6.1 (see Phylogenetic tree in Figure 3-13A).

4.2.3.1 Mutants of *OSDRL1* and *OSDRL2* were obtained from CRISPR/Cas9 system

Two targets were designed for both *OsDRL1* and *OsDRL2* respectively for mutagenesis using the CRISPR/Cas9 system (Figure 4-6). Genotypes of the transgenic plants were confirmed using PCR with primers OsDRL1_CriGT_F1 / OsDRL1_CriGT_R1 for OsDRL1 and OsDRL2_CriGT_F1 / OsDRL2_CriGT_R1 for OsDRL2 (Figure 4-12). The genotypes of mutants were confirmed by PCR products sequencing. 49 plants were detected for *OsDRL1*, among them seventeen were wild type, 8 were heterozygotes and 24 mutants of *OsDRL1* were obtained. 16 plants were detected for *OsDRL2*. Among them 9 were wild type, 4 were heterozygotes and 3 mutants of *OsDRL2* was obtained (details of mutants in Appendix V). Two mutants of *OsDRL1* and three mutants of O*sDRL2* were backcrossed with wild type to generate stable lines in next generation (see details in Figure 4-12).

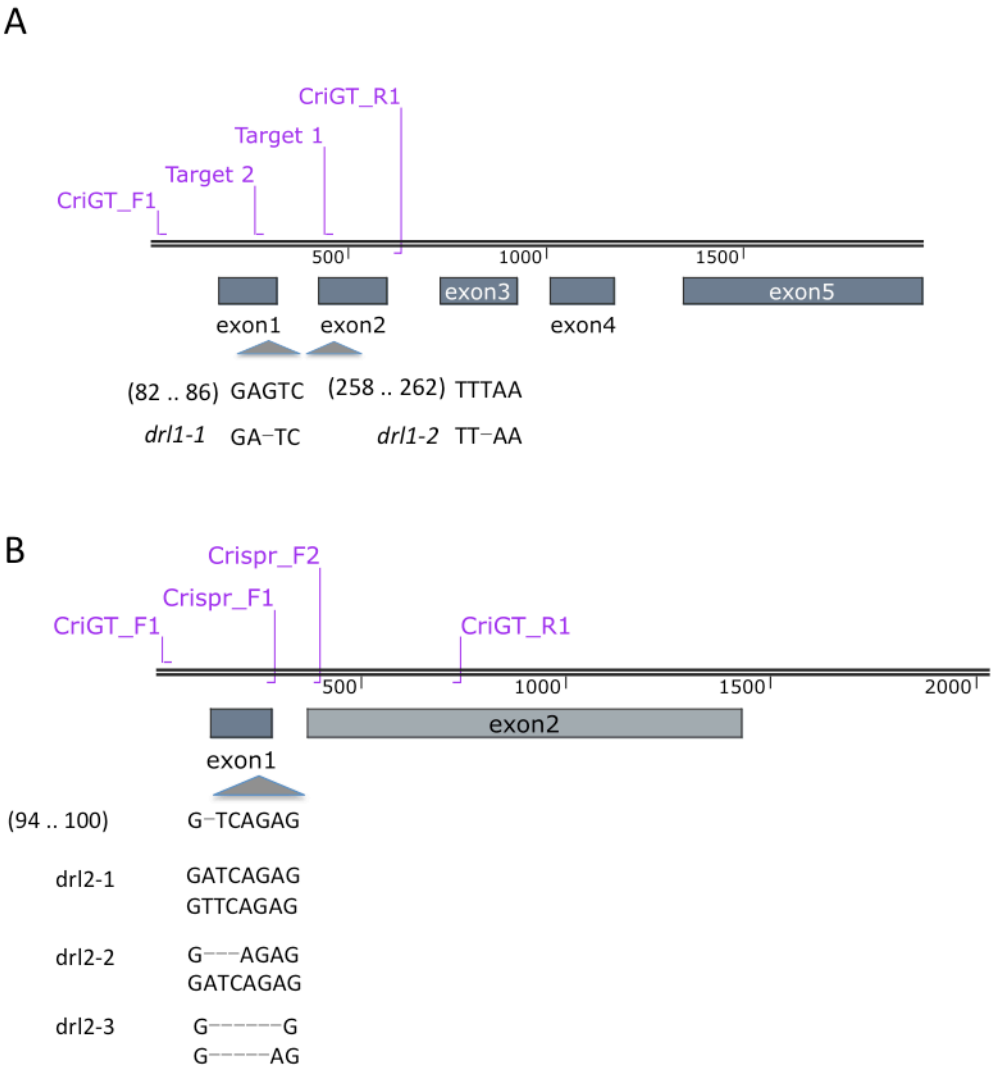


Figure 4-12 Schematic representations of *OsDRL1* and *OsDRL2*. A: *OsDRL1* shows the CRISPR targets, genotyping primers and sequencing information of drl1-1 mutant; B: *OsDRL2* indicates with CRISPR targets, genotyping primers and sequencing information of *drl2-1, drl2-2* and *drl2-3* mutants.

4.2.3.2 *osdrl1-1* and *osdrl2-1* were defective in pollen wall patterning and pollen grain release

Mutants *osdrl1-1* and *osdrl2-1* were used for further genotyping. Potassium iodine staining shows that pollen viability was normal in both *osdrl1-1* and *osdrl2-1*, however the amount of pollen was significantly reduced in *osdrl1-1* (Figure 4-13). This result suggests that *osdrl1-1* might produce less pollen, or

have defect in pollen grain releasing.

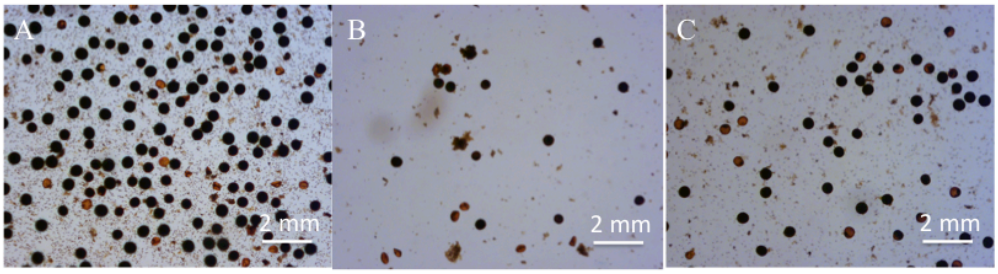


Figure 4-13 Pollen iodine viability staining of *osdrl1-1* and *osdrl2-1* at mature pollen stage. A: pollen of wild type 9522; B: pollen of *osdrl1-1*; C: pollens of *osdrl2-1*.

SEM was then performed to check the morphology of anthers and pollens at anther stage 12 in *osdrl1-1* and *osdrl2-1* (Figure 4-14). It was observed that, compared with 9522, the anthers of *osdrl1-1* and *osdrl2-1* mutants were smaller (Figure 4-14 A, D, G), however with normal shape and well patterned pollen wall.

The anthers of *osdrl1-1* and *osdrl2-1* were opened under a stereomicroscope to expose the inner surface of the tapetum layer. When WT anthers were opened, well-developed pollen grains were densely packed and observed to readily pop out of the anthers. However, when *osdrl1-1* anthers were opened, very few pollen grains came out.

After opening the anthers, the inner face was observed by SEM (Figure 4-14B, E, H). It could clearly see that the Ubisch bodies were on the tapetum layer in WT and *osdrl2-1*. However, neither the inner surface, nor Ubisch bodies, could be observed in the *osdrl1-1* mutant because pollen grains were closely

arranged and tightly attached to the tapetum layer (Figure 4-14E). What is more, it was found that among all pollen grains in *osdrl1-1*, no germ pore could be observed (Figure 4-14E). It seems that all germ pores were facing to the tapetum layer. In Arabidopsis, INP1 (INAPERTURATE POLLEN1) has been reported not only affect the specification of aperture positions but also involved in global exine patterning (Dobritsa & Coerper, 2012). The *osdrl1-1* may therefore be a good mutant for studying how pollen grains are arranged and facing into anther before releasing since this has not been reported in rice to date.

The morphology of pollen wall was also checked in *osdrl1* and *osdrl2*. Both mutants show abnormal pattering on the pollen cuticle surface. In *osdrl1-1*, pollen exine was covered by an extra layer of unknown materials (Figure 4-14F). This extra layer might be the reason why pollen grains were sticking to tapetum layer in the mutant *osdrl1-1*. In *osdrl2-1*, the pollen exine also looked more densely distributed with a cone-shape compared to wild type (Figure 4-14I).

Both *osdrl1-1* and *osdrl2-1* plants were fertile. Although the pollen of *osdrl1-1* were attached to the inner surface of anther so tightly that they could not easily be released, the pistil might still obtain a few pollen after anther dehiscence. The cross between *osdrl1-1* and *osdrl2-1* was performed to generate double mutants. However it was not successful when *osdrl1-1* as male parent because

the pollen of *osdrl1-1* could not be released by manually shaking the panicle.

These results supported the previous prediction that *OSDRL1* and *OSDRL2* play important role in pollen wall formation through the lipid synthesis pathway.

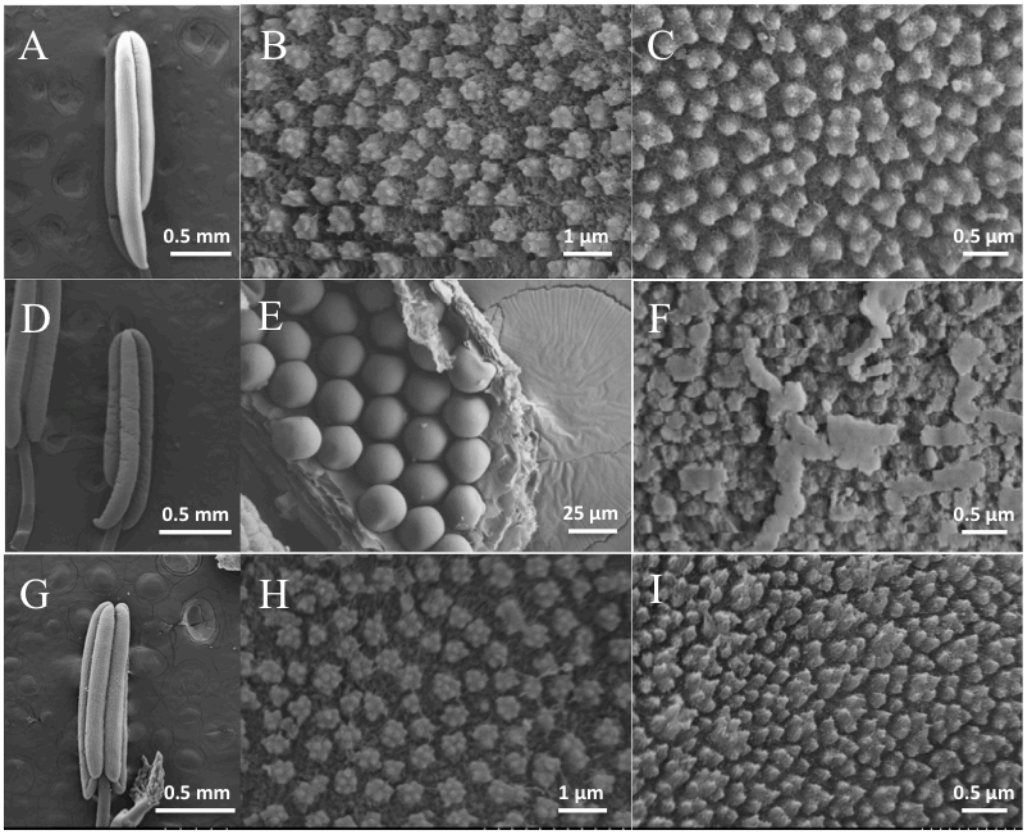


Figure 4-14 SEM pictures of anthers and pollen grains in WT, and the *osdrl1-1* and *osdrl2-1* CRISPR/CAS9 mutants at anther stage 12. A, D, G: anthers; B, H: Ubisch bodies; E: inner surface of anther wall; C, F, I: pollen exine; A, B, C: 9522 WT; D, E, F: *osdrl1-1*; G, H, I: *osdrl2-1*.

4.2.4 *GAMYB Like 2 (GAMYBL2)*

As mentioned in section 1.2.2.1, MYB transcriptional factors such as MYB33 and MYB65 in Arabidopsis as well as GAMYB in rice have been reported to regulate tapetum cell death and exine formation. GAMYBL2 is a member of

MYB family proteins which is the ortholog of GAMYB. However the expression of *GAMYBL2* is different from *GAMYB*. Previous research shows *GAMYB* is expressed from early anther stages and is highest in rice anthers at the tetrad and young microspore stages (stage 7 to stage 9) (Aya et al., 2009), while *GAMYBL2* expression is shorter and is highest in rice anthers at anther stage 8 and stage 9 (Figure 4-15A). *GAMYBL2* show very high co-expressed with candidate genes *LRR1* and *LRR5* in RiceAntherNet (Figure 4-15B) and was predicted as an important gene in pollen development.

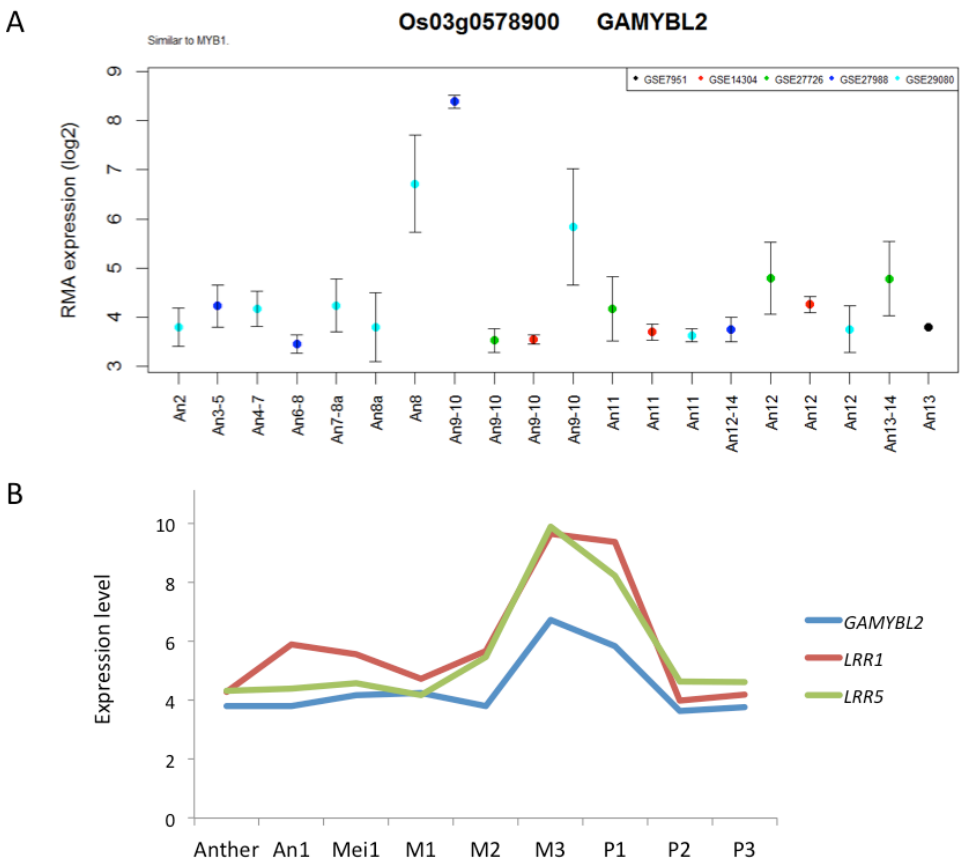


Figure 4-15 The expression of *GAMYBL2* in WT anther (RiceAntherNet: https://www.cpib.ac.uk/anther/riceindex.html). A: the expression of *GAMYBL2*; B: co-expression pattern of *GAMYBL2*, *LRR1* and *LRR5*.

4.2.4.1 Mutants of *GAMYBL2* were obtained using the CRISPR/Cas9 system

Four mutant lines of *GAMYBL2* were obtained using the CRISPR/Cas9 system after detecting fifteen plants (details in Appendix V). Primers GAMYBL2_CriGT_F1 / GAMYBL2_CriGT_R1 and GAMYBL2_CriGT_F2 / GAMYBL2_CriGT_R2 were designed for transgenic plants genotyping (Appendix II). All these four mutant lines (Figure 4-16) were backcrossed with 9522 to get stable mutants in the next generation. Mutant *gamybl2-3* line with one bp insertion at 121 bp of *GAMYBL2* was used for the phenotyping.

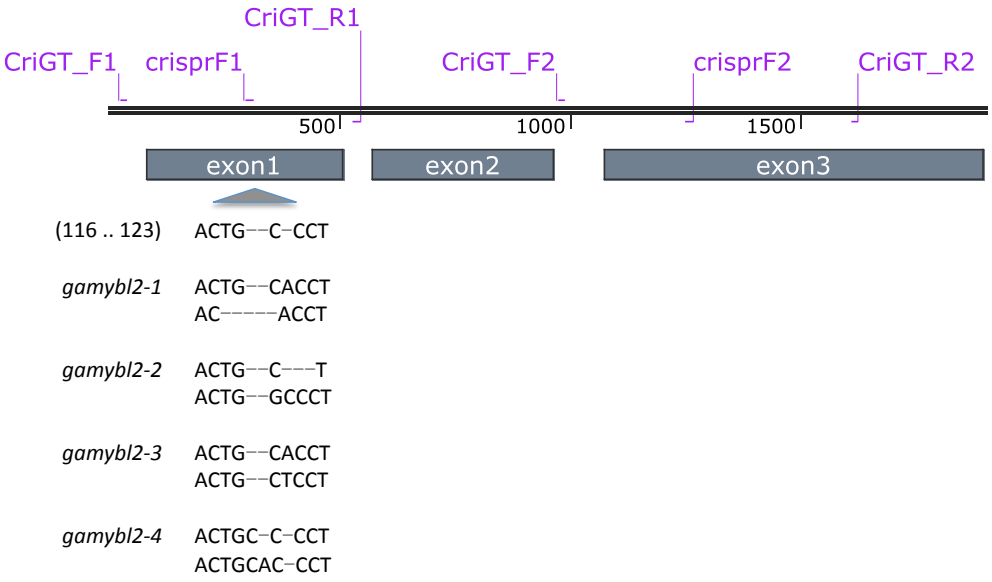


Figure 4-16 Schematic representations of *GAMYBL2* with CRISPR targets, genotyping primers and sequencing information of *gamybl2-1, gamybl2-2, gamybl2-3* and *gamybl2-4* mutants.

4.2.4.2 *GAMBYL2* was required for pollen wall formation

To further characterize the *gamybl2* mutant pollen viability and patterning of

the pollen exine, SEM was performed on WT and the *gamybl2-3* mutant. In *gamybl2-3* mutant, about half of the pollen grains were shrunken and half were normal (Figure 4-17C, E, G). No differences in Ubisch bodies and the anther wall were observed between WT and *gamybl2-3*. However, it was observed that *gamybl2-3* had diverse patterning of pollen exine (Figure 4-17D, F, H). Figure 4-17D shows the pollen exine was covered by an extra cobweb-shaped layer; Figure 4-17F shows that beads of a similar size were distributed on an uneven exine surface; Figure 4-17H shows the exine surface covered by bead-shaped materials so densely that some of them were connected to pieces. The abnormal pollen exine patterning indicates the potentially important role of *GAMYBL2* in regulating pollen exine formation.

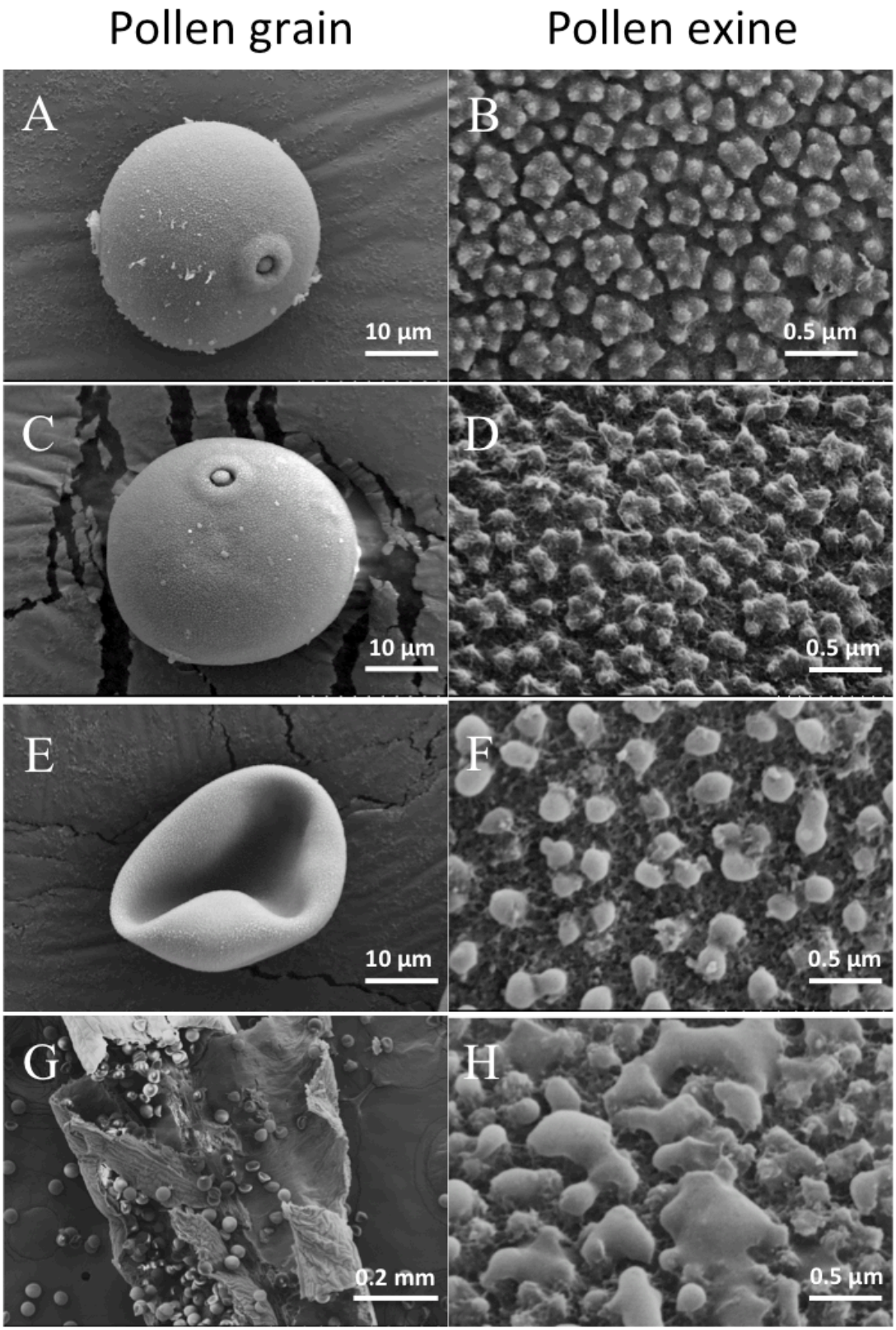


Figure 4-17 SEM pictures of pollen grains in WT and the *gamybl2-3*. A: WT pollen grain; B, E, G: *gamybl2-3* pollen grain; B: WT pollen exine; D, F, H: *gamybl2-3* pollen exine.

4.2.5 Other candidate genes that were investigated for their roles in pollen development

Other candidates including *OsLAP6*, *LRR1* and *LRR5* were identified from RiceAntherNet as potentially involved in rice pollen development and worth further investigation. *OsLAP6* encodes for a polyketide synthase involved in fatty acid modifications in rice. *OsLAP6* loss-of-function mutant was previously identified from a male sterile mutant library in Prof Dabing Zhang's laboratory (SJTU; unpublished data). Loss of function of *OsLAP5*, the ortholog of *OsLAP6*, has been published leading to a partial male sterile phenotype (Wang et al., 2013). To understand the relationship between *OsLAP5* and *OsLAP6* in regulating anther development, *OsLAP6* mutagenesis by CRISPR/Cas9 system was performed both in 9522 and in the *oslap5* mutant (Figure 4-18A). Two Leucine-rich receptor-like (LRR) protein kinases *LRR1* and *LRR5* that are predicted as involved in tapetum cell differentiation were also knocked out by the CRISPR/Cas9 system (Figure 4-18B, C). Mutants for *OsLAP6*, *LRR1* and *LRR5* were generated (genotypes in Appendix V). Mutants including *oslap6-1, lrr1-1, lrr1-2, lrr1-3, lrr1-4* and *lrr5-1* were backcrossed with 9522 and used in analysis of rice pollen development (Figure 4-18).

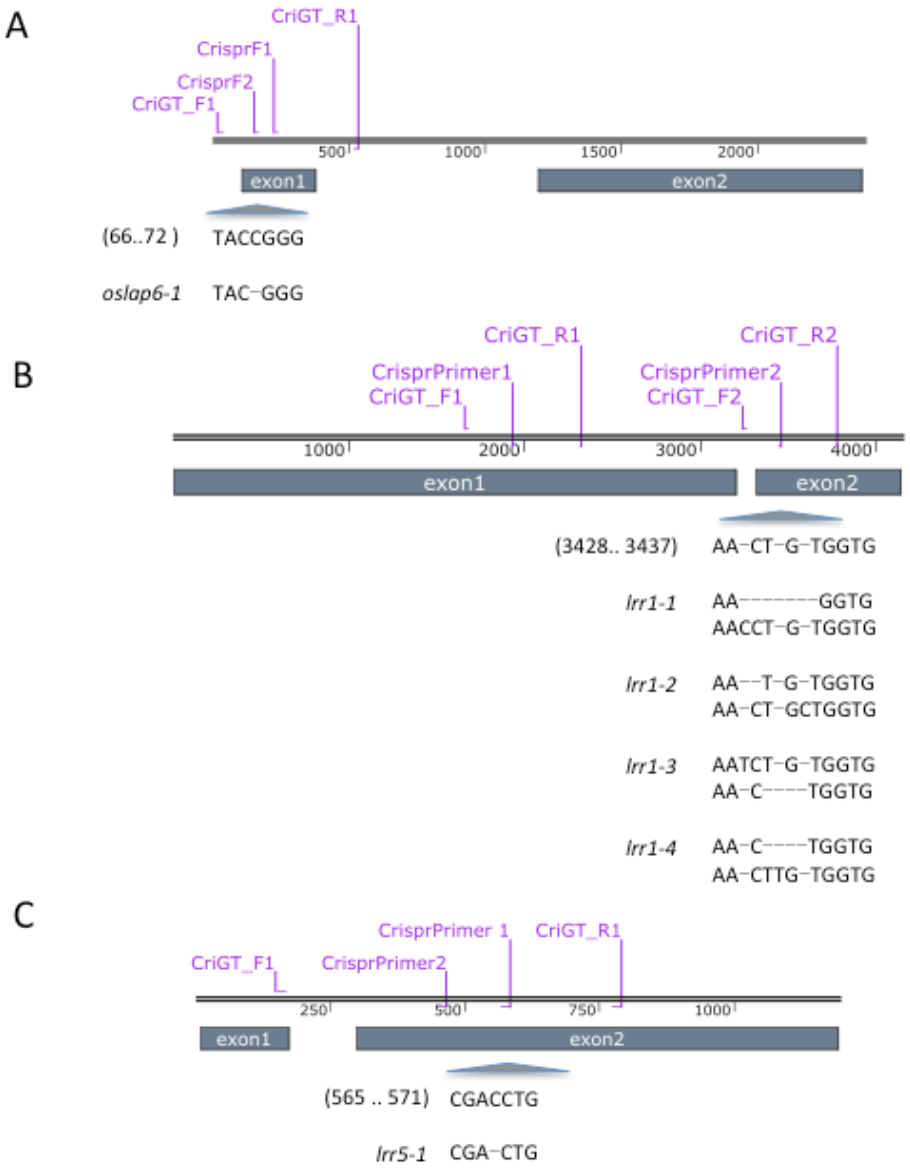


Figure 4-18 Schematic representations of *OsLAP6, LRR1* and *LRR5* with CRISPR targets, genotyping primers and sequencing information of mutants. A: Schematic representations of Os*LAP6*; B: Schematic representations of *LRR1*; C: Schematic representations of *LRR5*.

4.2.6 *LOB DOMAIN 2 (LBD2)* in Arabidopsis

Os01g0242400 in RiceAntherNet clade 1, *OsLBD2*, is one member of the LATERAL ORGAN BOUNDARIES (LOB) DOMAIN (LBD) gene family, which plays a role in anther development, downstream of *MYB80* in rice (Zhang et al., 2010b). The putative ortholog of *LBD2* in Arabidopsis has a

similar expression pattern as *OsLBD2* in rice (Figure 4-19). The expression pattern of *LBD2* in Arabidopsis shows that *LBD2* is highly expressed during anther stage 8 and 9. It also shows significantly down regulated in the *ms1* mutant (Figure 4-19B). An Arabidopsis T-DNA insertion line of *LBD2* was obtained and showed a partially male sterile phenotype.

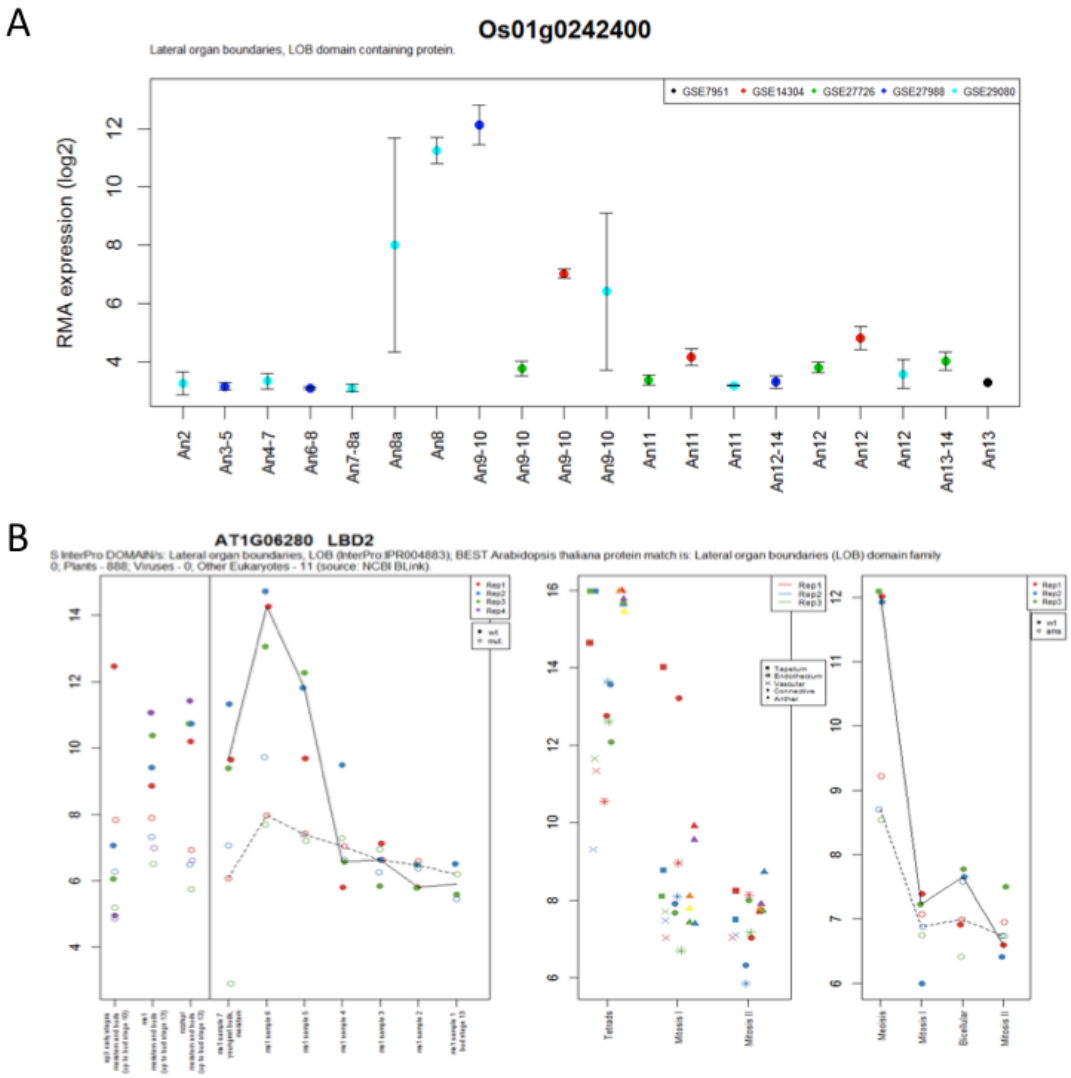


Figure 4-19 The expression of *LBD2* in rice and Arabidopsis. A: expression of *OsLBD2* in rice; B: expression of *LBD2* in Arabidopsis based on Arabidopsis FlowerNet (Pearce et al., 2015).

4.2.6.1 *LBD2* T-DNA insertion line

The SALK_149880, T-DNA insertion mutant of *LBD2* was obtained (Figure 4-20) and named *lbd2-1*. A 125bp T-DNA insertion was detected before the transcriptional start site (TSS) of *LBD2* by PCR using three primers system (http://signal.salk.edu/tdnaprimers.2.html) provided by Salk Institute Genomic Analysis Laboratory (see primers in Appendix II).

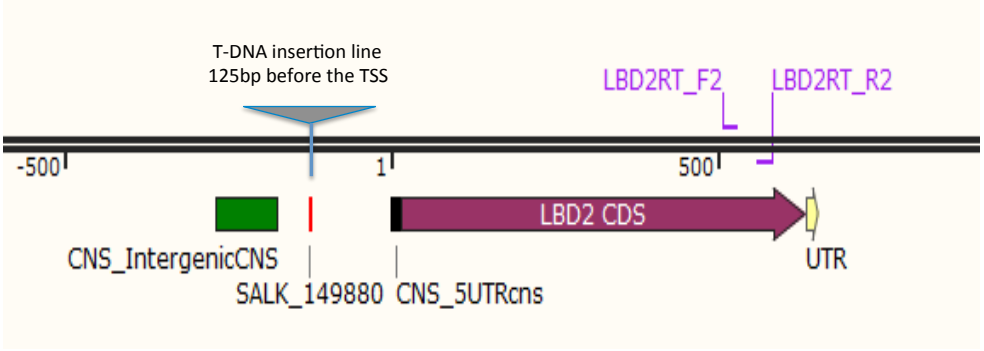


Figure 4-20 Schematic representations of *LBD2* with qRT-PCR primers and T-DNA insertion line information.

qRT-PCR was performed with primers LBD2_RT_F2 / LBD2_RT_R2 to confirm the expression of *LBD2* in *lbd2-1* (see primers in Figure 4-20). It was found that *lbd2-1* was an overexpression line with *LBD2* more than four fold higher expression than in WT Arabidopsis (Figure 4-21).

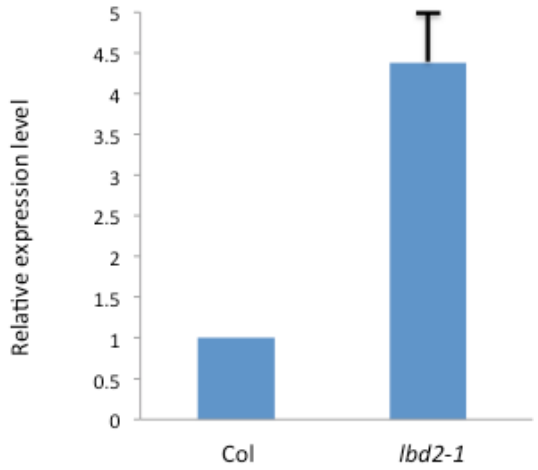


Figure 4-21 qRT-PCR of *LBD2* in leaves of WT and *lbd2-1*. Error bar: standard error. Data was normalized to *LBD2* in WT leaves.

4.2.6.2 *lbd2-1* over expression line of *LBD2* showed a partially male sterile phenotype

Phenotypes of inflorescences and flowers of WT and *lbd2-1* were examined (Figure 4-22). Figure 4-22B showed that *lbd2-1* was partially sterile, with three kinds of siliques: siliques with no seeds (arrow 1 in Figure 4-22B), siliques with a few seeds (arrow 2 in Figure 4-22B) and siliques full of seeds as in WT (arrow 3 in Figure 4-22B). Figure 4-22D illustrates that *lbd2-1* had shorter filaments but pistils of normal size, which meant that the anther could not reach to the stigmatic surface. Furthermore, the petals of *lbd2-1* were shorter and meant that the stigma was outside the petals (Figure 4-22F) while in the WT the petals totally encased the reproductive organs (Figure 4-22E). The shorter filaments and petals probably both contributed to the failure of pollination and lead to the partial sterile phenotype of *lbd2-1*. From the phenotypes of *lbd2-1*, it could be concluded that *LBD2* might regulated the

fertility of Arabidopsis by controlling the length of filaments and the size of petals.

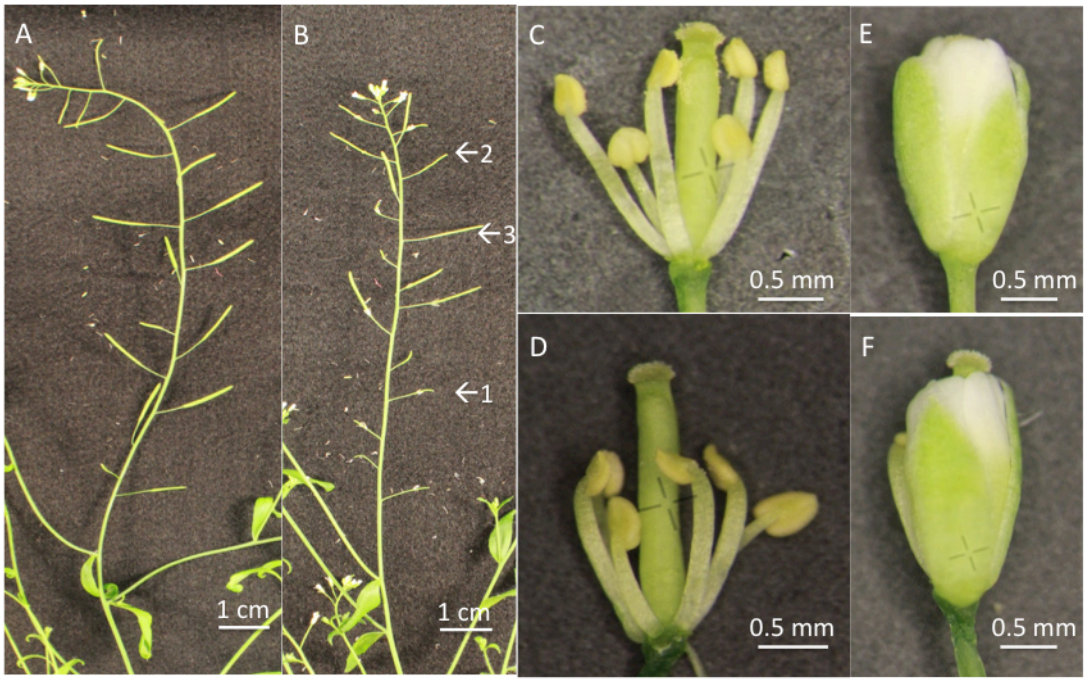


Figure 4-22 Phenotypes of *lbd2-1*. A: the inflorescence of WT; B: the inflorescence of *lbd2-1* (arrows point to different phenotypes of siliques); C: reproductive organs of WT; D: reproductive organs of *lbd2-1*; E: flower of WT; F: flower of *lbd2-1*.

4.3 DISCUSSION

In this chapter, 9 candidate genes (*CML41, OsDRL1, OsDRL2, GMCO, GAMYBL2, LBD2, OsLAP6, LRR1* and *LRR5*) predicted by reverse genetic methods (Chapter 3) to have a role in anther and pollen development were selected to verify their gene functions in rice and Arabidopsis. *CML41* was shown to be a downstream target of bHLH family protein EAT1 and TDR in rice; *GMCO, OsDRL1, OsDRL2* and *GAMYBL2* were found be required for pollen wall development and cuticle surface patterning in rice; over expression of *LBD2* in Arabidopsis revealed that *LBD2* had a function in controlling the

development of stamens and petals; other candidate genes including *OsLAP6*, *LRR1* and *LRR5* were knocked out in 9522 WT for gene functional analysis.

4.3.1 *CML41* plays as downstream of EAT1 and TDR in regulating Ca^{2+} homeostasis during rice anther development

Ca^{2+} homeostasis has been reported as a signal of tapetum PCD (section 1.2.2.1). *CML41* contains EF-hand motifs that bind Ca^{2+} to regulate Ca^{2+} homeostasis in rice anthers. It has been confirmed by qRT-PCR that *CML41* is specifically expressed in anther at stage 8 when tapetum PCD starts (section 4.2.1.2). Previous research revealed that the transcription factors TIP2, TDR and EAT1 could work together to regulate downstream genes (section 1.3). However, no cooperation of these three bHLH proteins were found in regulating the expression of *CML41* (Figure 4-5). Results of a Dual-Luciferase Reporter Assay showed that the bHLH family proteins EAT1 and TDR could bind to the promoter of *CML41* and positively regulate the expression of *CML41* individually when transiently expressed in tobacco, but TIP2 showed no role in regulating *CML41* (section 4.2.1.3). No male sterile phenotype was observed in *cml41-1* mutant indicates the redundant functions on regulating Ca^{2+} homeostasis in rice, suggesting possible redundancy with other family members.

4.3.2 Lipid synthesis genes are involved in pollen wall formation by affecting lipid metabolism

Lipid metabolism is important for normal tapetum PCD and pollen wall formation (section 1.2.5). 3 lipid synthesis genes (*GMCO*, *OSDRL1* and *OSDRL2*) were predicted and studied in this research.

4.3.2.1 Glucose-methanol-choline (GMC) oxidoreductase in rice pollen development

GMCO and *OSNP1* both encode Glucose-methanol-choline oxidoreductases in rice (section 4.2.2). Their ortholog in Arabidopsis, At1g12570 was reported to be involved in cutin biosynthesis (Kannangara et al., 2007). *OsNP1* has been found essential for anther cuticle and pollen exine formation and knock out of *OsNP1* lead to abortion of pollen development and failure of anther cuticle formation (Liu et al., 2017). The double mutant *gmco-1osnp1-4* partially recovered the cuticle formation of anther wall (section 4.2.2.3), which indicates that *GMCO* loss-of-function could accumulate cuticle materials in the *osnp1-4* mutant. It is possible that mutation of both *GMCO* and *OsNP1* may lead to a breakdown of the GMC oxido-reduction pathway but promotion of the accumulation of products in other pathways for cutin components. This hypothesis could be verified by further metabolites analysis of *gmco-1* and *osnp1-4*.

4.3.2.2 DFR/TKPR (DIHYDROFLAVONOL 4-REDUCTASE) family in rice and Arabidopsis pollen development

In Arabidopsis, *TKPR1* and *TKPR2* encode DIHYDROFLAVONOL 4-REDUCTASE family protein. In rice, *OsDRL1* and *OsDRL2* are orthologs of *TKPR1*; *OsDLR2* is the ortholog of *TKPR2* (See Phylogenetic tree in Figure 3-13A). *TKPR1*, *TKPR2* (Grienenberger et al., 2010) in Arabidopsis and *OsDFR2* (Shirasawa et al., 2013) in rice were reported to be involved in male fertility. In this research, the other two DFR family members in rice, *OsDRL1* and *OsDRL2* were predicted to be important for pollen exine formation by reverse genetic method. The phenotypes of all DFRs in rice and Arabidopsis are shown in Table 4-3. All DFRs studied so far in rice and Arabidopsis have been associated with pollen exine formation (Table 4-3). Mutation of any of them leads to an abnormal pollen exine patterning. In the DFR family, genes in TKPR1 clade (*TKPR1*, *OsDFR2* and *OsDRL1*) in phylogenetic tree (Figure 3-13A) are more important for pollen exine formation. The loss-of-function of any gene in TKPR1 clade showed more serious phenotypes in pollen exine development and lead to sterile or partial sterility. Genes in TKPR2 clade (*TKPR2* and *OsDRL2*) in phylogenetic tree (Figure 3-13A) however might slightly affect the composition of pollen exine with loss-of-function lines showing subtle defects in exine formation which do not effect for pollen viability and plant fertility. According to the phenotypes observed, this DFR pathway seems conserved in Arabidopsis and rice.

Table 4-3 Phenotypes of *DIHYDROFLAVONOL 4-REDUCTASE* (*DFR)* genes in rice and Arabidopsis

Gene	Species	Fertility	Phenotype of mutants
TKPR1	Arabidopsis	Sterile	No reticulate exine pattern of pollen exine (Grienenberger et al., 2010)
OsDFR2	Rice	Sterile	No pollen (Shirasawa et al., 2013)
OsDRL1	Rice	Partial Sterile	Extra layer over pollen exine which make pollen grains cannot release (section 4.2.3.2)
TKPR2	Arabidopsis	Fertile	A few spherical protrusions inside and outside pollen exine tecta (Grienenberger et al., 2010)
OsDRL2	Rice	Fertile	More dense pattering of pollen exine (section 4.2.3.2)

4.3.3 GA regulated MYB family members in rice pollen development

It has been reported that MYB transcriptional factor *GAMYB* in rice and the homologs *MYB33* and *MYB65* in Arabidopsis could be regulated by GA to play roles in anther development (section 1.2.2.1). *GAMYBL2* is an ortholog of *GAMYB* which was coexpressed with *LRR1* and *LRR5* (Figure 4-15B). In *gamybl2-3* mutant, a partial male sterile phenotype with various patternings of pollen exine was observed (section 4.2.4.2). These findings suggest that *GAMYBL2* plays a crucial role in regulating pollen exine formation, however further analysis is needed to confirm this role.

4.3.4 Other genes in RiceAntherNet would be good candidates for further study

In RiceAntherNet (chapter 3), genes in clade1 were predicted relating to pollen wall formation. To date, the validation of *CML41, GMCO, OsDRL1* and

OsDRL2 suggests their functions in rice pollen development which is consist with the predictions of RiceAntherNet. *LBD2* has been studied in Arabidopsis with a role in stamen and petal development. Mutants of *LAP6*, *LRR1* and *LRR5* in rice have been obtained and provide good materials for further gene functional study. In addition, other genes predicted in Clade 1 as being required for anther and pollen development would be good candidates to investigate.

5 *TIP2* REGULATION NETWORK

5.1 INTRODUCTION

TIP2 encodes a bHLH transcription factor which regulates the differentiation of middle layer and tapetum layer, as well as the tapetum cell death (Fu et al., 2014). Previous research shows that in early anther development stages, *TIP2* controls the tapetum cell death by regulating the expression of two bHLH genes *TDR* and *EAT1*, while GA also play as a signal for tapetum degradation through *GAMYB* (section 1.2.2.1). In addition, lipid metabolism and transport genes which are involved in the tapetum late function and are involed with pollen wall formation, including *CYP703A3*, *CYP704B2* and *C6*, were found significantly down regulated in the *tip2* mutant (section 5.2.1). However, whether *TIP2* regulating tapetal cell death is dependent on GA regulatory pathway on tapetal cell death remain elusive. The molecular mechanism and downstream target genes of *TIP2* regulating the lipid metabolism is unclear as well. To deepen the understanding of *TIP2* regulation network, microarray data of *tip2* anthers (unpublished data from Prof. Dabing Zhang's lab, SJTU) were used for mining downstream genes of *TIP2*.

5.2 RESULTS

Differentially Expressed Gens (DEGs) in *tip2* were found based on microarray data of *tip2* anthers at anther stage 6 and stage 7 (section 5.2.1). After

bioinformatics analysis including expression pattern analysis, GO enrichment analysis and E-box analysis, 6 genes encode lipid transport protein family proteins were recognized as possible candidate downstream target genes of *TIP2* (section 5.2.2).

5.2.1 Bioinformatics investigation for prediction of downstream gene targets of TIP2

To investigate the downstream genes of *TIP2*, 12 Affymetrix arrays (including 9522 anther at stage 6, 9522 anther at stage 7, *tip2* anther at stage 6, *tip2* anther at stage 7; three biological repeats for each sample) were used to calculate the DEGs. Genes with relative expression level over 8 (log_2 scale) were kept and genes with 2-fold change between *tip2* and WT were recognized as DEGs. In total, 280 genes were found to be significantly down regulated in *tip2* at anther stage 6 and the number was 62 in *tip2* at anther stage 7 (Appendix II).

All the DEGs were then clustered by hierarchical clustering (section 2.2.4) based upon gene expression pattern in WT (Figure 5-1A). Because *TIP2* is highly expressed at anther stage 7 to stage 9, genes which are also highly and specifically expressed in anthers at the same stages were identified as they might be the downstream target genes of *TIP2*. Therefore, the clade with 31 genes co-expressed during anther stage 7 to 9 (see gene details in Appendix III) highlighted in Figure 5-1A was recognized as candidates for the following GO enrichment analysis (section 2.2.6).

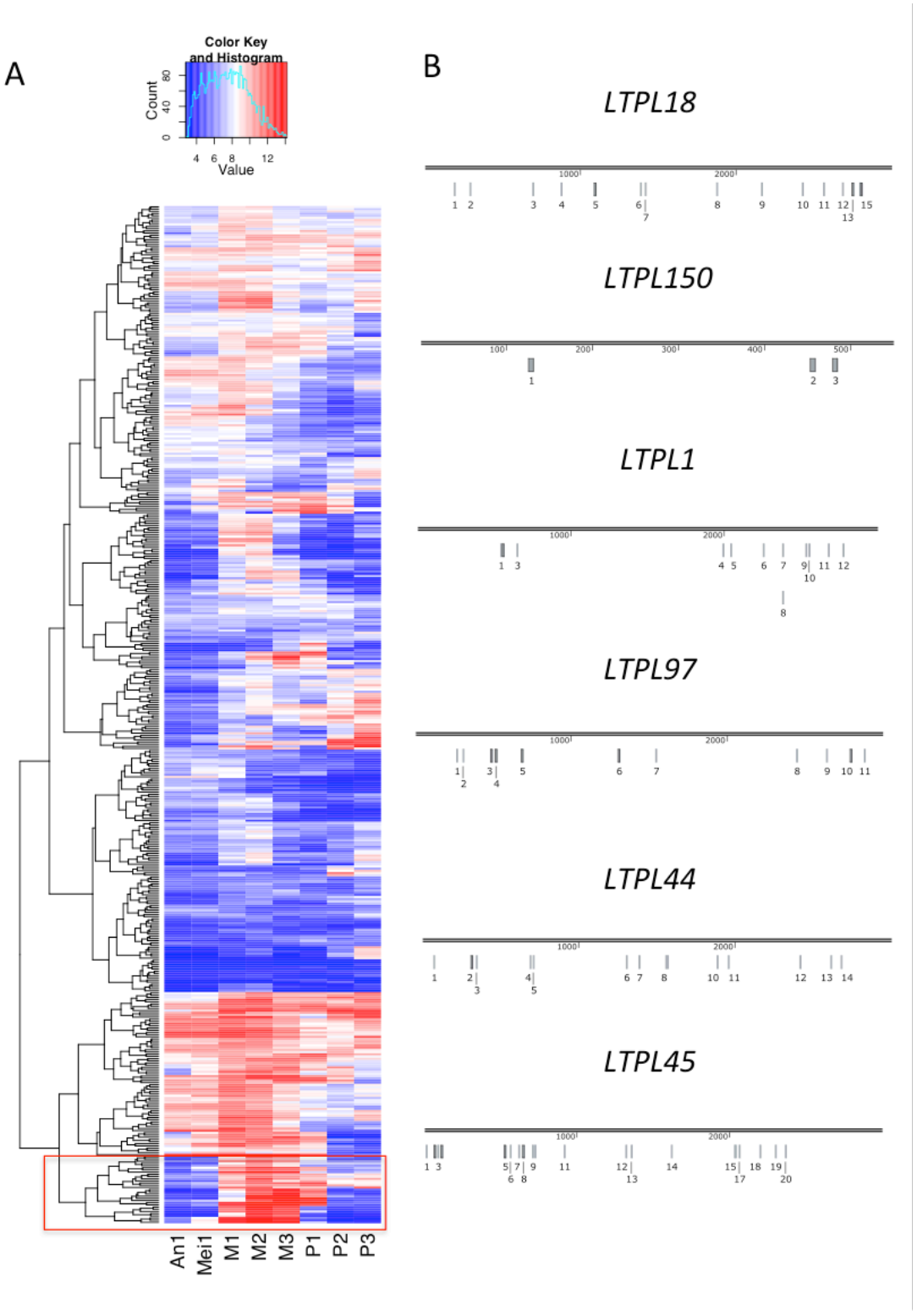


Figure 5-1 Gene expression analysis and E-box analysis of DEGs of *tip2*. A: DEGs expression pattern in WT. The clade highlighted by red frame was selected for a follow-up investigation. B: E-box analysis of LTPs.

Table 5-1 shows the result of GO enrichment for these 31 downstream

candidates of *TIP2*. In addition to GO terms shown in three example clades of RiceAntherNet (Table 3-2 in section 3.2.5), including pollen exine formation, spropollenin biosynthesis and pollen wall assembly，it is interesting to see that GO term of lipid metabolism, such as lipid localization and lipid transport were be enriched in TIP2 downstream candidates.

Table 5-1 GO enrichment analysis for 31 putative downstream genes of *TIP2*

GO.ID	Term	p-value
GO:0010876	Lipid localization	5.20E-06
GO:0010584	Pollen exine formation	1.60E-05
GO:0080110	Sporopollenin biosynthetic process	1.60E-05
GO:0010208	Pollen wall assembly	1.90E-05
GO:0010927	Cellular component assembly involved in morphogenesis	1.90E-05
GO:0006869	Lipid transport	2.40E-05

Upon close observation of those candidates, it was found that, among 31 candidates, 8 of them were related to lipid metabolism pathways (Table 5-2), including 2 reported lipid synthesis proteins CYP703A3 (Yang et al., 2014) and CYP704B2 (Li et al., 2010), as well as 6 Lipid Transfer Protein precursors (LTP) LTPL18, LTPL44, LTPL45, LTPL150, LTPL97 and LTPL1. All of them were largely down-regulated in *tip2* (Table 5-2). This indicates that *TIP2* could regulate rice anther development by effecting the expression of lipid transfer related genes. No GA related MYB family members were found in this candidate clade, as well as no GA synthesis genes show expression changes in the *tip2* mutant. This suggests that *TIP2* does not the regulate tapetum PCD through the GA pathway.

Table 5-2 Candidate genes in lipid metabolism pathway that show altered expression in the *tip2* mutant. Fold changes of gene expressions between tip2 mutant and WT at anther stage 6 and anther stage 7 are listed. Value =Log2 (Fold Change).

RAP ID	Gene	*tip2* S6	*tip2* S7
Os01g0219500	LTPL18		-5.26
Os08g0546300	LTPL44		-3.40
Os09g0525500	LTPL45	-2.83	-6.44
Os01g0691300	LTPL150	-3.14	-6.79
Os08g0545800	LTPL97	-2.39	-6.60
Os03g0357500	LTPL1	-1.44	-5.16
Os08g0131100	CYP703A3		-0.93
Os03g0168600	CYP704B2		-0.78

5.2.2 LTPs were regulated by TIP2, TDR and GAMYB

As mentioned before (section 5.1) TIP2 could interact with other bHLH proteins, for example TDR *and* EAT1*,* to regulate tapetum differentiation and cell death. GAMYB has similar function as TIP2 in regulating tapetum degradation (section 5.1). However, it is still unclear whether TIP2 regulates Lipid Transfer Protein genes through the typical TIP2-mediated transcription cascade (section 1.3) and whether GAMYB and TIP2 could both regulate lipids metabolism pathway. To answer these questions, the expression patterns of these 8 lipid metabolism related genes (section 5.2.1) were checked in *tdr*, *eat-1* and *gamyb-2* (data from Rice AntherNet, fold change details in Appendix II). Figure 5-2 shows that all LTPs which are putative downstream of *TIP2* were also significantly down-regulated in mutants of *TDR* and *GAMYB*, but no expression changes were seen in *eat1-1*. These data indicate

that TIP2 might interact with TDR to affect lipid transfer while EAT1 is not involved in this process. In addition, the overlap of LTPs which show altered expression in both *TIP2* and *GAMYB* suggest that *TIP2* and *GAMYB* might cooperate to regulate lipid metabolism in rice anthers.

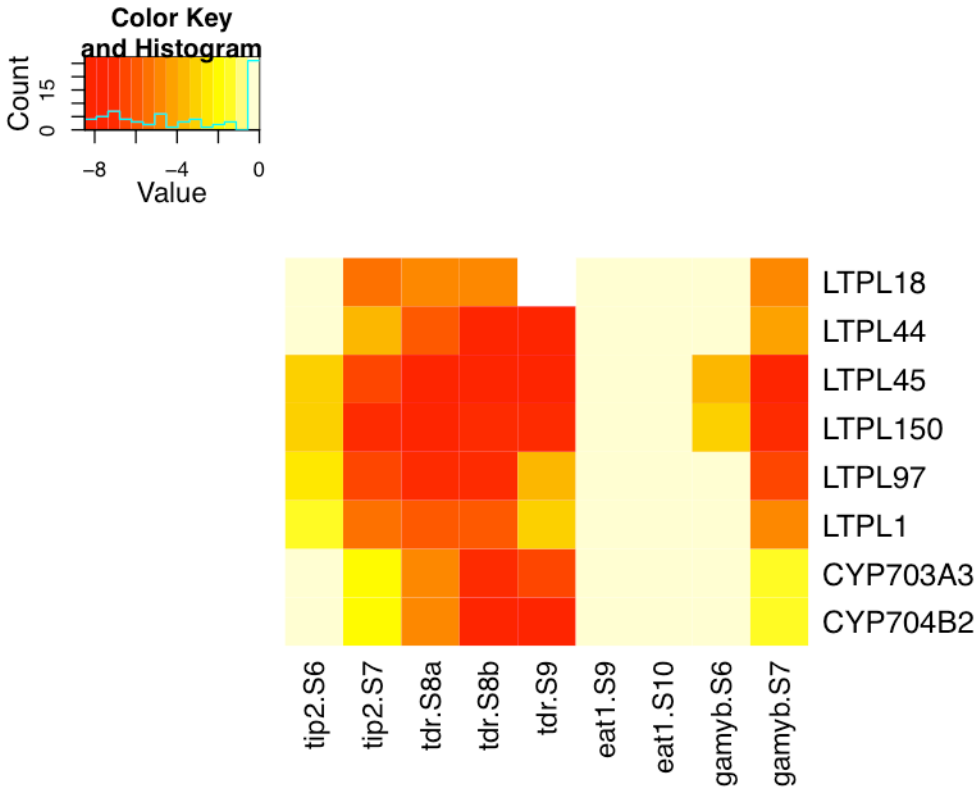


Figure 5-2 Heatmap of fold changes of LTPs in *tip2, eat1-1* and *gamyb-2* mutants. S= Anther stage. Value =Log2 (Fold Change).

5.2.3 LTPs were indirect downstream genes of TIP2

To further confirm that whether TIP2 and TDR could regulate those LTPs directly during rice anther development, E-box analysis was performed on the 3K promoters of 6 LTPs, including *LTPL18, LTPL44, LTPL45, LTPL150, LTPL97* and *LTPL1* (Figure 5-1B). However only *LTP1* and *LTP18* were found to have E-boxes accumulated near the transcriptional start site (Figure

5-1B).

Therefore, the Dual-Luciferase Reporter Assay was performed on the 3k promoter regions of *LTPL1* and *LTPL18* to explore the regulatory relationship between LTPs and the TIP2 protein (Figure 5-3). The reporter plasmid with 3k promoter of *LTPL1* and *LTPL18*, as well as the effector plasmids with cDNA of *TIP2* and *TDR* used in the transient expression mediated by *Agrobacterium* in tobacco leaves were shown in Figure 5-3A. The primers for amplifying the promoter sequence of *LTPL1* and *LTPL18* are listed in Appendix II. The effector plasmids were obtained from the Zhang lab, SJTU (Fu et al., 2014). Four independent transient transformations were performed for each treatment, and data of each treatment were normalized to mock (empty pGreen0229 vector) inoculations (Figure 5-3B, C).

No significant induction of expression, as detected by the luciferase assay in transient assays, was observed for the *LTPL1* and *LTPL18* promoters under regulatory control by the TIP2 and TDR proteins (Figure 5-3B, C). The results suggest that the bHLH proteins TIP2 and TDR do not directly regulate the expression of *LTPL1* or *LTP18* by binding to their promoter regions. The combination of TIP2 and TDR also do not show regulation of *LTPL1* or *LTP18* expression (Figure 5-3B, C).

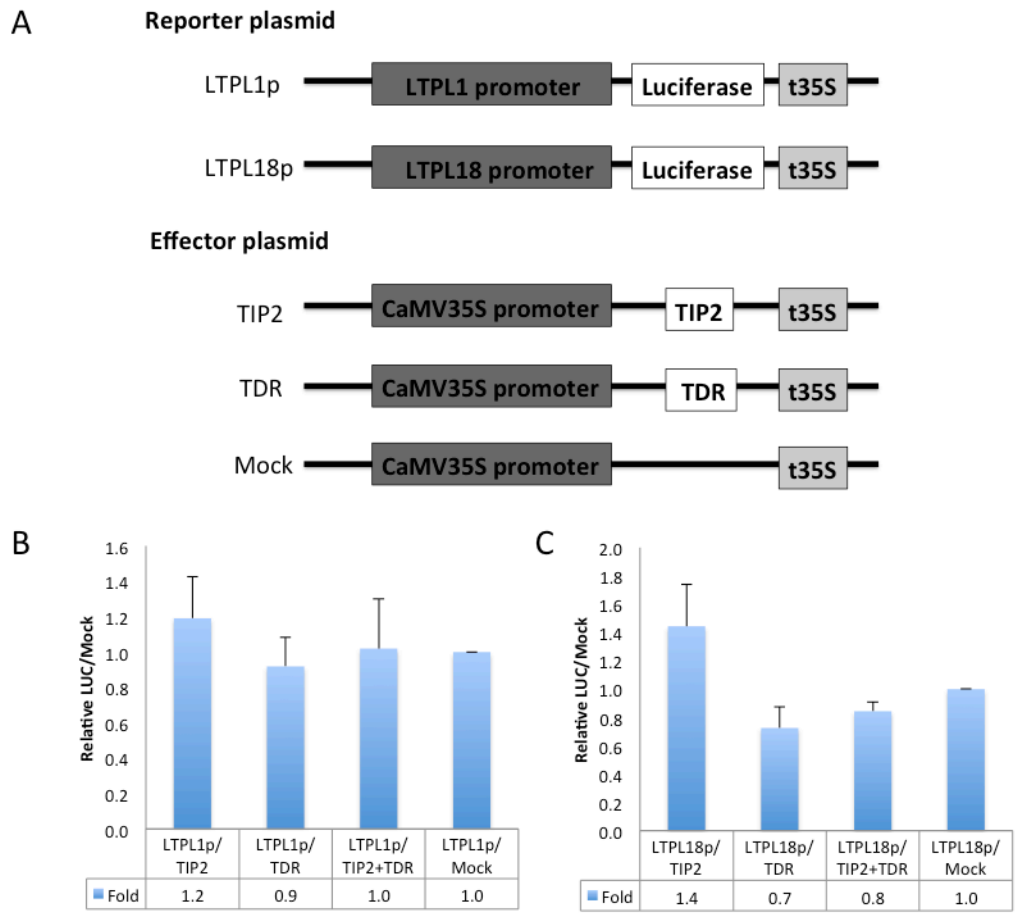


Figure 5-3 Regulation of the *LTPL1 and LTPL2* Promoters by TIP2. A: Schematic diagrams of the reporter plasmids and effector plasmids used in the transient expression in tobacco leaves; B: Transactivation of the Luc reporter gene with *LTPL1* promoter by TIP2 and TDR; C: Transactivation of the Luc reporter gene with *LTPL18* promoter by TIP2 and TDR. Error bar: standard deviation.

5.3 DISCUSSION

In this study, by comparing microarray data of the wild type 9522 and *tip2* mutant at anther stage 6 and stage 7, it is indicates that *TIP2* not only could regulate genes involved in pollen exine formation and sporopollenin biosynthetic process but also the lipid transfer process (Table 5-1).

In plants, Lipid Transfer Proteins (LTPs) are synthesized as precursors with a N-terminal peptide and they are assumed to facilitate the transfer of fatty acids

phospholipids and steroids between membranes (LIN et al., 2017). The expression of 6 unknown LTPs regulated by TIP2 were also found significantly changed in *tdr* and *gamyb-2*. According to results from the Dual-Luciferase Reporter Assay, this regulation by TIP2 is not by direct regulation and there are likely to be other regulators through which TIP2 and TDR regulate lipid metabolism.

GAMYB is required for expression of *CYP703A3* and other genes involved in the synthesis of sporopollenin and also some GA-up regulated genes to effect cell death in tapetum (section 1.2.2.1). Mutantion of *GAMYB* shows defects in PCD of tapetum cells as well as formation of exine and Ubish bodies, which is similar to the *tip2* mutant (section 1.2.2.1). However *GAMYB* is regulated by GA while *TIP2* show no upstream or downstream relationship with GA since no GA-related genes were found in the DEGs in the *tip2* mutant (section 5.2.1). It is evident that *GAMYB* and *TIP2* share a large number of downstream LTPs, which indicates their cooperation in an unknown way to regulate lipid metabolism in anther.

6 GENERAL DISCUSSION

Rice is an important food crop for the world's food security. Hybrid rice contributes a lot to the food supply all over the world because of its high yield and high quality (Chang et al., 2016). However, studies on rice reproductive development, especially gene regulatory relationships during anther development are needed to address the challenges of future rice breeding.

In this work, a rice global co-expression network (RiceAntherNet) was generated from public microarray data for studying gene regulations during rice anther and pollen development (chapter 3). Not only were gene expression patterns of annotated 24,462 genes in the anther provided, the expression profiles of these genes in nine male sterile mutants, as well as the expression relationships between paired genes were also provided for future scientific research references. In addition, genes taking part in the early differentiation stage, the tapetum PCD stage and later pollen wall formation stage were predicted to show interconnected relationships in the rice anther regulatory networks. Gene functions of nine predicted genes were investigated by molecular genetic approaches to deepen understanding on rice anther development (chapter 4).

Microarray dataset of *tip2* mutant were further analyzed to mine the downstream genes regulated by *TIP2* or *TIP2*-mediated transcription cascade (chapter 5). The expression of six unknown LTPs involved in lipid transfer

were found to be dependent upon TIP2.

Based upon the outcomes of this work, the combination of gene function predictions by reverse genetic approaches followed by gene function characterization by molecular genetic approaches is a reliable and efficient way for gene function studies in rice anther development.

6.1 GENERATION OF GENE EXPRESSION DATA FROM PUBLIC MICROARRAY DATASETS FOR REFERENCE

As biological technology develops, large amount of plant biology data have been produced in recent decades (Popescu et al., 2016). The increase of biological data opens the door for improving the understanding of plant biology by data gathering, analysis and sharing. Large numbers of public gene expression datasets of rice facilitate the rice gene expression study in anther (section 1.4.2). In this work microarray datasets of wild type rice anther which were publicly available on the Gene Expression Omnibus (GEO) (Barrett et al., 2013) were used for gene expression study in rice.

6.1.1 Gene expression profiles in rice anther were generated

57 Affymetrix arrays from anther tissues of wild type rice and 45 Affymetrix arrays on other tissues of wild type rice were downloaded from GEO (array information in Appendix III). 57 anther-specific arrays were used to give rice anther gene expression profiles during anther development (Appendix III). 45 arrays from other tissues were also analyzed to provide a rice global gene

expression profile in other rice tissues (Appendix III). Expression patterns of 24,462 genes both in different stages of anther and other tissues were available from the website (https://www.cpib.ac.uk/anther/riceindex.html) and can be seen in the attached CD. This expression data includes the gene expression pattern, but also the gene annotation and putative ortholog in Arabidopsis, which make it a good source for rice gene expression profile screening.

6.1.2 Gene fold changes in male sterile mutants

Gene expression patterns in nine male sterile mutants including *msp1-4, ostdl1a, gamyb-2, tip2, udt1-1, tdr, eat1-1, ptc1* and *mads3-4* were analysed using microarray datasets from these mutants (Table 3-1). The fold change of all DEGs which significantly changed in male sterile mutants are listed in Appendix III. This is a convenient source for querying genes of interest in these nine mutants and gives valuable information for anther development research. DEGs followed by GO enrichment analysis are a good way for studying downstream genes and corresponding gene functions (section 3.2.5).

6.1.3 Gene co-expression network during pollen and anther development were generated

Gene co-expression analysis has been used widely for gene function prediction. Genes with similar expression pattern possibly have the same function or involved in the same biological pathway. In this work, a gene co-expression network (RiceAntherNet) was generated including 3422 well-connected

clusters with 9358 nodes (genes) and 75853 edges (co-relations). This is available on website (https://www.cpib.ac.uk/anther/riceindex.html) for querying the correlations between of genes of interest.

6.2 REVERSE GENETIC APPROACHES FOR GENE FUNCTION PREDICTION

Rice anther development includes lots of important biological process and complex gene regulation networks with a lot of regulators involved in this process (section 1.1). To date, a large part of this regulation network remains unknown and almost all genes have been reported in this process were found by forward genetic approaches. In this work, reverse genetics show the power of such approaches to speed up novel gene function study.

6.2.1 Forward genetics in plant biology

Although forward genetics is widely used for identifying gene functions, it has limitations. First of all, the analysis of forward genetic approaches based on the characterization of mutants, requires the identification of clear phenotypes. The mutants with clear phenotypes are easily selected from mutant libraries as research materials, while mutants without obvious phenotypes might be ignored. Secondly, gene mapping and positional cloning are essential for the identification of genes from a mutant, but gene identification is not always successful due to technical limitations. Thirdly, the single mutation of genes that have redundant functions in plant development do not always show phenotypes and multiple gene family knockouts are needed. In these cases

above, the functional analysis of those genes would be impossible to investigate by forward genetics.

6.2.2 Reverse genetics in plant biology

Recently, the development of next generation sequencing has made it possible to study gene sequences genome-wide (Shah et al., 2015). In addition, the age of biological big data opens the door for data mining by a series of bioinformatics methods (Popescu et al., 2016). These make reverse genetics a powerful approach for novel gene function prediction and characterization. Recently the development of CRISPR/Cas9 system for targeted gene editing in rice (Zhang et al., 2014) is providing a great benefit for the application of reverse genetics in plants. Novel genes predicted by reverse genetic approaches can be easily mutated by the CRISPR/Cas9 system enabling further mutant phenotype characterization and gene function investigation.

In this work, gene expression pattern analysis, gene co-expression analysis, GO enrichment analysis, Phylogenetic tree, binding motif analysis and paper mining were all used as reverse genetic approaches for gene function prediction.

6.2.3 RiceAntherNet for prediction of genes involved in anther and pollen development

A rice gene co-expression network for anther development (RiceAntherNet) was developed for prediction of gene regulatory relationships during pollen

development (section 3.2.4). Three clades showing differential expression patterns were constructed by mapping the DEGs in male sterile mutants to RiceantherNet (Figure 3-10). Genes in Clade 1 was identified associated with pollen wall formation in late anther development stage (section 3.2.6); genes in Clade 2 may have a function during cell differentiation in early anther development stages (section 3.2.7); genes in clade 3 were predicted to be involved in the tapetum PCD process (section 3.2.8). Nine candidate genes predicted by RiceAntherNet were chose for further study which later proved the validation of RiceAntherNet.

6.2.4 Exploration of downstream genes of TIP2 by using microarray data of *tip2*

TIP2 is the only reported gene to date, which not only controls differentiation of the middle layer and tapetum layer, but also controls the tapetum PCD process (Fu et al., 2014; Ko et al., 2014). In this study, microarray data from the wild type 9522 and *tip2* mutant at anther stage 6 and stage 7 were compared to find DEGs which might be the downstream targets of TIP2. The expression patterns of these DEGs were filtered and 31 DEGs expressed at anther stage 7 to stage 9 identified, that were predicted as downstream targets of TIP2 (chapter 5). GO enrichment analysis of these 31 candidates indicated that TIP2 may not only regulate genes involved in pollen exine formation and sporopollenin biosynthetic process, but also the lipid transfer process, since

two Cytochrome P450 family member (CYP703A3 (Yang et al., 2014)and *CYP704B2* (Li et al., 2010)) and six Lipid Transfer Protein precursors (*LTPL18, LTPL44, LTPL45, LTPL150, LTPL97* and *LTPL1*) were identified as potential targets (Table 5-1). CYP703A3 and CYP704B2 have been shown to be essential for pollen wall formation (Li et al., 2010; Yang et al., 2014) and the expression of *CYP704A3* is activates by GAMYB (Aya et al., 2009). LTPs have been shown to be involved in various biological processes, such as cutin biosynthesis, defense response, plant signaling and seed maturation (LIN et al., 2017). Motif analysis and Dual-luciferase reporter analysis suggests that TIP2 and TDR are not directly regulating these targets (Figure 5-3). The additional microarray analysis of *tdr* and *gamyb-2* (Figure 5-2) shows all LTPs were also down regulated in *tdr* and *gamyb-2* as they were in *tip2*. The overlap of downstream genes in these three mutants suggests that the bHLH family proteins TIP2 and TDR may be working together in the tapetum to regulate the lipid transfer pathway and that GAMYB may interact with TIP2, in an as yet uncharacterized way, to regulate lipid metabolism in the anther.

The mutation of these candidate LTPs by CRISPR/Cas9 system is on going (primers shown in Appendix II). The further phenotyping of these mutant lines will give more evidence for their function in sporopollenin transfer and pollen wall formation.

6.3 GENE FUNCTION INVESTIGATION FOR NOVEL GENES PREDICTED FROM THE CORRELATION NETWORK

Nine genes (*CML41, OsDRL1, OsDRL2, GMCO, GAMYBL2, LBD2, OsLAP6, LRR1* and *LRR5*) were predicted as candidates that might be involved in rice pollen wall formation. Mutations of these genes were manipulated by CRISPR-Cas9 editing and mutants were obtained for all of these candidates. To investigate the gene functions of these candidates on pollen wall formation, phenotyping of mutants of five candidates (*CML41, OsDRL1, OsDRL2, GMCO* and *GAMYBL2*) were performed by SEM to check the anther wall formation and pollen wall patterning.

6.3.1 The function of DFR/TKPR (DIHYDROFLAVONOL 4-REDUCTASE) family in pollen development is conserved in rice and Arabidopsis

OsDRL1 and *OsDRL2* are predicted to function in pollen wall formation by affecting lipid synthesis (section 4.1). Phylogenetic analysis (Figure 3-13A) shows that *TKPR1* and *TKPR2* in Arabidopsis are putative orthologs of *OsDRL1* and *OsDRL2*. TKPR1 and TKPR2 function as oxidoreductases involved in sporopollenin formation, which are able to reduce the carbonyl of the tetraketide α-pyrone to a secondary alcohol (Grienenberger et al., 2010). Mutations of *TKPR1* and *TKPR2* result in defect in pollen exine patterning in Arabidopsis (Grienenberger et al., 2010). Previously, only *OsDFR2* was found as an ortholog of *TKPR1* in rice, and no pollen was observed in the mutant of

OsDFR2 (Shirasawa et al., 2013). However by reverse genetic approaches, *OsDRL1* and *OsDRL2* were identified in this work as additional orthologs of the Arabidopsis *TKPR1* and *TKPR2* genes. The mutants of *OsDRL1* and *OsDRL2* both show pollen exine patterning defects indicating the conserved function of the DFR/TKPR family in pollen exine formation (Figure 4-14, Table 4-3). However the phenotype of mutants of *TKPR2* clade (*TKPR2* in Arabidopsis and *OsDRL2* in rice) in the phylogenetic tree (Figure 3-14A) show more subtle phenotypes in pollen exine compared with the TKPR1 clade. This finding suggests that the functions of TKPR1 clade and TKPR2 clade as reductases are not fully redundant *in vivo*. Nevertheless despite exine formation defects, both mutants of *OsDRL1* and *OsDRL2* were fertile, which means they would not have been identified by forward genetic approaches. The prediction of their roles by reverse genetic approaches deepens the understanding of how *DFR/TKPR* family may work together in sporopollenin biosynthesis as partial redundant genes in regulating pollen wall formation.

Further work could be conducted to determine the enzymatic activity of OsDRL1 and OsDRL2 proteins by expression in *E.coli*, which would further confirm the function of OsDRL1 and OsDRL2. The double mutant of *drl1-1drl2-1* will be obtained in the next generation by cross *drl1-1* and *drl2-1* mutants. Comparison of the exine formation pattering among the *drl1-1, drl2-1* single and *drl1-1drl2-1* double mutants will provide more clues as to the function of the DFR family in sporopollenin synthesis in rice.

6.3.2 *GMCO (Glucose-methanol-choline oxidoreductase)* may be involved in cutin component biosynthesis

GMCO is a functionally uncharacterized gene predicted by RiceantherNet as involved in lipid synthesis (section 4.1). It was reported that GMC oxidoreductase superfamily catalyse the oxidation of oxidate alcohol moiety to aldehyde using the cofactor flavin adenine dinucleotide (Wongnate & Chaiyen, 2013). One GMC oxidoreductase *OsNP1* gene has been reported essential for anther cuticle and pollen exine formation, with an absences of cuticle observed in *osnp1* mutant (Liu et al., 2017). According to Liu et al., the measurement of cuticular wax in *osnp1-4* shows that cuticular wax, particularly hydroxylated components for aliphatic cutin monomers were reduced in *osnp1-4*, which is consist with the lack of cuticle phenotype in *osnp1-4* (Liu et al., 2017). A double mutant *gmco-1osnp1-4* was obtained in this work and SEM showed that mutation of *GMCO* in *osnp1-4* partially recovered the cuticle formation (Figure 4-14).

As reported previously, GMC enzymes have differences in the mechanism of catalyzing substrate oxidation, despite having similar protein structures (Wongnate & Chaiyen, 2013). The partial recovery of cuticle formation in the *gmco-1osnp1-4* double mutant may be because when there is a loss of function of both *GMCO* and *OsNP1*, other cutin components may be directed towards pollen exine production and accumulation for cutin patterning.

The single mutant line of *GMCO* will be obtained in the next generation by backcrossing *gmco-1osnp1-4* with wild type. Measuring cuticular wax and cutin constituents in anthers of *gmco-1, osnp1-4, gmco-1osnp1-4* and wild type will help to understand the enzyme activity of *GMCO* and *OsNP1* as well as the cutin components in rice anthers. The finding that *GMCO* is associated with pollen exine formation further validates the advantages of using reverse genetic approaches for gene functional analysis.

6.3.3 *GAMYBL2* may play important role in pollen exine formation

A MYB transcription factor, *GAMYBL2,* was predicted as an important gene in pollen (section 4.1). Mutation of *GAMYBL2* showed partial male sterility and abnormal pollen exine patterning (Figure 4-17), which suggest it plays a crucial role in pollen development. According to DEGs in RiceAntherNet (.cys file in attached CD), No significant expression changes of *GAMYBL2* were observed in *gamyb-2, udt1-1 and tdr* mutants, but a 224-fold down regulation was seen in the *ptc1* mutant and a 14.7-fold down regulation in the *eat1* mutant (Li et al., 2011a). As described in section 3.2.5, mutants of *EAT1* and *PTC1* have very distinct morphological changes compared with other male sterile mutants. The *udt1-1, gamyb-2* and *tdr* mutants have enlarged tapetum cells during the later stages of anther development (Aya et al., 2009; Hu et al., 2011; Li et al., 2006b); while *ptc1* and *eat1-1* mutants show darkly stained materials, with a high density of organelles in the tapetum in later stages (Li et al., 2011a;

Niu et al., 2013). The expression changes of *GAMYBL2* in these persistent tapetum mutants (*eat1-1* and *ptc1*) and tapetum enlarged mutants (*udt1-1, gamyb-2* and *tdr*) indicates that tapetum cell structure are still under control of unknown factors even if the tapetum cannot undergo normal PCD process. It is thought that the mutant of *GAMYBL2* may have a similar tapetum form to *ptc1* and *eat1-1*. To further investigate the tapetum behavior in the *GAMYBL2* mutant during rice anther development, TEM materials of anther at stage 7 to stage 12 have been collected for future TEM analysis.

6.3.4 Study of *CML41 (calmodulin like 41)* may help with understanding Ca^{2+} signalling during tapetum PCD

CML41 encodes a CML protein which regulates Ca^{2+} homeostasis in plants (Mohanta et al., 2017). In this work, *CML41* was found to be significantly down regulated in the *eat1, tdr* and *osdex1* mutants according to the microarray analysis (section 4.1). Results of the Dual-Luciferase Reporter Assay show that EAT1 and TDR could positively regulate the *CML41* promoter individually (Figure 4-5).

Previous work reported a CaM protein, OsDEX1, that controlled tapetum PCD, which might be via regulating Ca^{2+} homeostasis in tapetal cells (Yu et al., 2016). However, the signaling pathway and molecular mechanism behind this regulation currently remains elusive. It has been reported in animals, that the Ca^{2+} content of the Endoplasmic Reticulum (ER) act as cell death signals to

affect the apoptosis process (Giorgi et al., 2008). It would therefore be interesting to investigate the concentration of Ca^{2+} in tapetum cells at different stages to understand signaling and regulating mechanism of tapetum PCD. Further study on *CML41* would potentially help to establish the dynamics of Ca^{2+} homeostasis in tapetum cells. Mutants of *OsDEX1* and *CML41* would potentially be good material to reveal the regulatory network of *OsDEX1*, as well as find out the function and molecular mechanism of Ca^{2+} signaling during tapetum PCD. SEM shows no phenotype in both anthers and pollen of the *cml41-1* mutant which indicates potential redundancy of *CML41* with other family members. Multiple mutations of other putative redundant genes should be performed for further functional study of CMLs family in rice.

6.3.5 *LBD2* in Arabidopsis affect stamen and petal development

OsLBD2 has a similar expression pattern in rice to that of its Arabidopsis ortholog, *LBD2* (Figure 4-19). Over expression of *LBD2* in Arabidopsis caused a partially male sterile phenotype with shorter filaments and shorter petals (Figure 4-22). It is possible that *LBD2* controls the length of filaments and the size of petals and thus lead to partial sterility in Arabidopsis. It would be interesting to knock out *LBD2* both in Arabidopsis and in rice to learn more about *LBD2* function and whether the role of *LBD2* is conserved in both species.

6.3.6 Other candidates would be good targets for deepening understanding of gene regulation network during pollen development in rice

Mutants of other candidates (section 4.1) from RiceAntherNet, including *LAP6*, *LRR1* and *LRR5* were obtained in the Wild type (9522) background (details in Appendix V). Future work is needed in the genotyping and phenotyping of these mutants by SEM and TEM to understand their function in rice pollen development. Previous functional validation of candidate genes indicates the high efficiency of functional studies by reverse genetic approaches. Therefore other genes predicted in RiceAntherNet also provide good candidate for further investigation in the future.

6.4 CONCLUSIONS

This work is an example of using reverse genetic approaches to explore novel, uncharacterized genes involved in rice anther development. The ongoing increase of biological data and development of CRISPR/Cas9 system in rice provides powerful tools to capitalize upon the data generated for target identification from the RiceAntherNet correlation network.

By using public microarray data, a website (https://www.cpib.ac.uk/anther/riceindex.html) which provide global expression pattern in rice anther of 24, 462 genes has been provided for user's queries. Gene expression changes in 9 rice male sterile mutants (*msp1-4,*

ostdl1a, gamyb-2, tip2, udt1-1, tdr, eat1-1, ptc1 and *mads3-4*) are also available. These reference files are valuable for rice anther development research by gene expression pattern screening.

A co-expression network was generated to predict genes involved in pollen wall formation. This offers a much quicker approach than forward genetics to explore the functions of candidate genes in pollen and anther development. Among the candidates predicted as genes required for rice anther pollen development, *OsDRL1*, *OsDRL2*, *GMCO* and *GAMYBL2* have been characterized as essential for anther and pollen exine formation. The investigation of other candidates such as *OsLAP6*, *LRR1* and *LRR5* are on going. These findings indicate the successful application of co-expression network analysis to predict genes involved in rice pollen development process.

In addition, a study on *TIP2* downstream pathways by *tip2* microarray datasets indicates that TIP2 may interact EAT1, as well as cooperate with GAMYB to regulate lipid metabolism during anther development.

The using of reverse genetic approaches to explore genes required for rice anther development furthers our understanding of gene regulation during this process. Reverse genetic approaches combined with genotyping of mutants provide an efficient and reliable way for gene regulation study in the future.

REFERENCES

Achard, P., Herr, A., Baulcombe, D. C., & Harberd, N. P. 2004. Modulation of floral development by a gibberellin-regulated microRNA. ***Development***, 131(14): 3357-3365.

Andeme Ondzighi, C., Christopher, D. A., Cho, E. J., Chang, S. C., & Staehelin, L. A. 2008. Arabidopsis protein disulfide isomerase-5 inhibits cysteine proteases during trafficking to vacuoles before programmed cell death of the endothelium in developing seeds. ***Plant Cell***, 20(8): 2205-2220.

Aoki, K., Ogata, Y., & Shibata, D. 2007. Approaches for extracting practical information from gene co-expression networks in plant biology. ***Plant Cell Physiol***, 48(3): 381-390.

Ariizumi, T., Kawanabe, T., Hatakeyama, K., Sato, S., Kato, T., Tabata, S., & Toriyama, K. 2008. Ultrastructural characterization of exine development of the transient defective exine 1 mutant suggests the existence of a factor involved in constructing reticulate exine architecture from sporopollenin aggregates. ***Plant Cell Physiol***, 49(1): 58-67.

Ariizumi, T., & Toriyama, K. 2011. Genetic regulation of sporopollenin synthesis and pollen exine development. ***Annu Rev Plant Biol***, 62: 437-460.

Avanci, N. C., Luche, D. D., Goldman, G. H., & Goldman, M. H. 2010. Jasmonates are phytohormones with multiple functions, including plant defense and reproduction. ***Genet Mol Res***, 9(1): 484-505.

Awad, S., & Chen, J. 2014. Inferring transcription factor collaborations in gene regulatory networks. ***BMC Systems Biology***, 8(Suppl 1): S1.

Aya, K., Suzuki, G., Suwabe, K., Hobo, T., Takahashi, H., Shiono, K., Yano, K., Tsutsumi, N., Nakazono, M., Nagamura, Y., Matsuoka, M., & Watanabe, M. 2011. Comprehensive network analysis of anther-expressed genes in rice by the combination of 33 laser microdissection and 143 spatiotemporal microarrays. ***PLoS One***, 6(10):

e26162.

Aya, K., Ueguchi-Tanaka, M., Kondo, M., Hamada, K., Yano, K., Nishimura, M., & Matsuoka, M. 2009. Gibberellin modulates anther development in rice via the transcriptional regulation of GAMYB. ***Plant Cell***, 21(5): 1453-1472.

Bar-Joseph, Z., Gerber, G. K., Lee, T. I., Rinaldi, N. J., Yoo, J. Y., Robert, F., Gordon, D. B., Fraenkel, E., Jaakkola, T. S., Young, R. A., & Gifford, D. K. 2003. Computational discovery of gene modules and regulatory networks. ***Nat Biotechnol***, 21(11): 1337-1342.

Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., Marshall, K. A., Phillippy, K. H., Sherman, P. M., Holko, M., Yefanov, A., Lee, H., Zhang, N., Robertson, C. L., Serova, N., Davis, S., & Soboleva, A. 2013. NCBI GEO: archive for functional genomics data sets--update. ***Nucleic Acids Res***, 41(Database issue): D991-995.

Bassel, G. W., Lan, H., Glaab, E., Gibbs, D. J., Gerjets, T., Krasnogor, N., Bonner, A. J., Holdsworth, M. J., & Provart, N. J. 2011. Genome-wide network model capturing seed germination reveals coordinated regulation of plant cellular phase transitions. ***Proc Natl Acad Sci U S A***, 108(23): 9709-9714.

Bergmann, D. C., & Sack, F. D. 2007. Stomatal development. ***Annu Rev Plant Biol***, 58: 163-181.

Bush, S. M., & Krysan, P. J. 2007. Mutational evidence that the Arabidopsis MAP kinase MPK6 is involved in anther, inflorescence, and embryo development. ***J Exp Bot***, 58(8): 2181-2191.

Cao, P., Jung, K. H., Choi, D., Hwang, D., Zhu, J., & Ronald, P. C. 2012. The Rice Oligonucleotide Array Database: an atlas of rice gene expression. ***Rice (N Y)***, 5(1): 17.

Chang, Z., Chen, Z., Wang, N., Xie, G., Lu, J., Yan, W., Zhou, J., Tang, X., & Deng, X. W. 2016. Construction of a male sterility system for hybrid rice breeding and seed production using a nuclear male sterility gene. ***Proc Natl Acad Sci U S A***, 113(49): 14145-14150.

Chen, W., Yu, X. H., Zhang, K., Shi, J., De Oliveira, S., Schreiber, L.,

Shanklin, J., & Zhang, D. 2011. Male Sterile2 encodes a plastid-localized fatty acyl carrier protein reductase required for pollen exine development in Arabidopsis. ***Plant Physiol***, 157(2): 842-853.

Choi, H., Jin, J. Y., Choi, S., Hwang, J. U., Kim, Y. Y., Suh, M. C., & Lee, Y. 2011. An ABCG/WBC-type ABC transporter is essential for transport of sporopollenin precursors for exine formation in developing pollen. ***Plant J***, 65(2): 181-193.

Dai, M., Wang, P., Boyd, A. D., Kostov, G., Athey, B., Jones, E. G., Bunney, W. E., Myers, R. M., Speed, T. P., Akil, H., Watson, S. J., & Meng, F. 2005. Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. ***Nucleic Acids Res***, 33(20): e175.

de Azevedo Souza, C., Kim, S. S., Koch, S., Kienow, L., Schneider, K., McKim, S. M., Haughn, G. W., Kombrink, E., & Douglas, C. J. 2009. A novel fatty Acyl-CoA Synthetase is required for pollen development and sporopollenin biosynthesis in Arabidopsis. ***Plant Cell***, 21(2): 507-525.

Dekkers, B. J., Pearce, S., van Bolderen-Veldkamp, R. P., Marshall, A., Widera, P., Gilbert, J., Drost, H. G., Bassel, G. W., Muller, K., King, J. R., Wood, A. T., Grosse, I., Quint, M., Krasnogor, N., Leubner-Metzger, G., Holdsworth, M. J., & Bentsink, L. 2013. Transcriptional dynamics of two seed compartments with opposing roles in Arabidopsis seed germination. ***Plant Physiol***, 163(1): 205-215.

Dobritsa, A. A., & Coerper, D. 2012. The novel plant protein INAPERTURATE POLLEN1 marks distinct cellular domains and controls formation of apertures in the Arabidopsis pollen exine. ***Plant Cell***, 24(11): 4452-4464.

Dobritsa, A. A., Nishikawa, S., Preuss, D., Urbanczyk-Wochniak, E., Sumner, L. W., Hammond, A., Carlson, A. L., & Swanson, R. J. 2009a. LAP3, a novel plant protein required for pollen development, is essential for proper exine formation. ***Sex Plant Reprod***, 22(3): 167-177.

Dobritsa, A. A., Shrestha, J., Morant, M., Pinot, F., Matsuno, M., Swanson, R., Moller, B. L., & Preuss, D. 2009b. CYP704B1 is a long-chain fatty acid omega-hydroxylase essential for sporopollenin synthesis in pollen

of Arabidopsis. ***Plant Physiol***, 151(2): 574-589.

Edel, K. H., Marchadier, E., Brownlee, C., Kudla, J., & Hetherington, A. M. 2017. The Evolution of Calcium-Based Signalling in Plants. ***Curr Biol***, 27(13): R667-R679.

Endo, S., Shinohara, H., Matsubayashi, Y., & Fukuda, H. 2013. A novel pollen-pistil interaction conferring high-temperature tolerance during reproduction via CLE45 signaling. ***Curr Biol***, 23(17): 1670-1676.

Feller, A., Machemer, K., Braun, E. L., & Grotewold, E. 2011. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. ***Plant J***, 66(1): 94-116.

Fu, Z., Yu, J., Cheng, X., Zong, X., Xu, J., Chen, M., Li, Z., Zhang, D., & Liang, W. 2014. The Rice Basic Helix-Loop-Helix Transcription Factor TDR INTERACTING PROTEIN2 Is a Central Switch in Early Anther Development. ***Plant Cell***, 26(4): 1512-1524.

Gautier, L., Cope, L., Bolstad, B. M., & Irizarry, R. A. 2004. affy--analysis of Affymetrix GeneChip data at the probe level. ***Bioinformatics***, 20(3): 307-315.

Gill, S. S., & Tuteja, N. 2010. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. ***Plant Physiol Biochem***, 48(12): 909-930.

Giorgi, C., Romagnoli, A., Pinton, P., & Rizzuto, R. 2008. Ca2+ signaling, mitochondria and cell death. ***Curr Mol Med***, 8(2): 119-130.

Goetz, M., Godt, D. E., Guivarc'h, A., Kahmann, U., Chriqui, D., & Roitsch, T. 2001. Induction of male sterility in plants by metabolic engineering of the carbohydrate supply. ***Proc Natl Acad Sci U S A***, 98(11): 6522-6527.

Grienenberger, E., Kim, S. S., Lallemand, B., Geoffroy, P., Heintz, D., Souza Cde, A., Heitz, T., Douglas, C. J., & Legrand, M. 2010. Analysis of TETRAKETIDE alpha-PYRONE REDUCTASE function in Arabidopsis thaliana reveals a previously unknown, but conserved, biochemical pathway in sporopollenin monomer biosynthesis. ***Plant Cell***, 22(12): 4067-4083.

Guo, L., Gao, Z., & Qian, Q. 2014. Application of resequencing to rice

genomics, functional genomics and evolutionary analysis. ***Rice (N Y)***, 7(1): 4.

Hamada, K., Hongo, K., Suwabe, K., Shimizu, A., Nagayama, T., Abe, R., Kikuchi, S., Yamamoto, N., Fujii, T., Yokoyama, K., Tsuchida, H., Sano, K., Mochizuki, T., Oki, N., Horiuchi, Y., Fujita, M., Watanabe, M., Matsuoka, M., Kurata, N., & Yano, K. 2011. OryzaExpress: an integrated database of gene expression networks and omics annotations in rice. ***Plant Cell Physiol***, 52(2): 220-229.

He, Y., Fukushige, H., Hildebrand, D. F., & Gan, S. 2002. Evidence supporting a role of jasmonic acid in Arabidopsis leaf senescence. ***Plant Physiol***, 128(3): 876-884.

Hedden, P., & Phillips, A. L. 2000. Gibberellin metabolism: new insights revealed by the genes. ***Trends Plant Sci***, 5(12): 523-530.

Hellens, R. P., Allan, A. C., Friel, E. N., Bolitho, K., Grafton, K., Templeton, M. D., Karunairetnam, S., Gleave, A. P., & Laing, W. A. 2005. Transient expression vectors for functional genomics, quantification of promoter activity and RNA silencing in plants. ***Plant Methods***, 1: 13.

Hirano, K., Aya, K., Hobo, T., Sakakibara, H., Kojima, M., Shim, R. A., Hasegawa, Y., Ueguchi-Tanaka, M., & Matsuoka, M. 2008. Comprehensive transcriptome analysis of phytohormone biosynthesis and signaling genes in microspore/pollen and tapetum of rice. ***Plant Cell Physiol***, 49(10): 1429-1450.

Hong, L., Tang, D., Zhu, K., Wang, K., Li, M., & Cheng, Z. 2012. Somatic and reproductive cell development in rice anther is regulated by a putative glutaredoxin. ***Plant Cell***, 24(2): 577-588.

Hord, C. L., Chen, C., Deyoung, B. J., Clark, S. E., & Ma, H. 2006. The BAM1/BAM2 receptor-like kinases are important regulators of Arabidopsis early anther development. ***Plant Cell***, 18(7): 1667-1680.

Hord, C. L., Sun, Y. J., Pillitteri, L. J., Torii, K. U., Wang, H., Zhang, S., & Ma, H. 2008. Regulation of Arabidopsis early anther development by the mitogen-activated protein kinases, MPK3 and MPK6, and the ERECTA and related receptor-like kinases. ***Mol Plant***, 1(4): 645-658.

Hu, L., Liang, W., Yin, C., Cui, X., Zong, J., Wang, X., Hu, J., & Zhang, D.

2011. Rice MADS3 regulates ROS homeostasis during late anther development. ***Plant Cell***, 23(2): 515-533.

Hu, L., Tan, H., Liang, W., & Zhang, D. 2010. The Post-meiotic Deficicent Anther1 (PDA1) gene is required for post-meiotic anther development in rice. ***J Genet Genomics***, 37(1): 37-46.

Huang, J.-Z., E, Z.-G., Zhang, H.-L., & Shu, Q.-Y. 2014. Workable male sterility systems for hybrid rice: Genetics, biochemistry, molecular biology, and utilization. ***Rice (N Y)***, 7: 13.

Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U., & Speed, T. P. 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. ***Biostatistics***, 4(2): 249-264.

Itoh, J., Nonomura, K., Ikeda, K., Yamaki, S., Inukai, Y., Yamagishi, H., Kitano, H., & Nagato, Y. 2005. Rice plant development: from zygote to spikelet. ***Plant Cell Physiol***, 46(1): 23-47.

Jakobsen, M. K., Poulsen, L. R., Schulz, A., Fleurat-Lessard, P., Møller, A., Husted, S., Schiøtt, M., Amtmann, A., & Palmgren, M. G. 2005. Pollen development and fertilization in Arabidopsis is dependent on the MALE GAMETOGENESIS IMPAIRED ANTHERS gene encoding a Type V P-type ATPase. ***Genes & Development***, 19: 2757-2769.

Jepson, C., Karppinen, K., Daku, R. M., Sterenberg, B. T., & Suh, D. Y. 2014. Hypericum perforatum hydroxyalkylpyrone synthase involved in sporopollenin biosynthesis--phylogeny, site-directed mutagenesis, and expression in nonanther tissues. ***FEBS J***, 281(17): 3855-3868.

Jiang, J., Zhang, Z., & Cao, J. 2013. Pollen wall development: the associated enzymes and metabolic pathways. ***Plant Biol (Stuttg)***, 15(2): 249-263.

Jung, K. H., Han, M. J., Lee, Y. S., Kim, Y. W., Hwang, I., Kim, M. J., Kim, Y. K., Nahm, B. H., & An, G. 2005. Rice Undeveloped Tapetum1 is a major regulator of early tapetum development. ***Plant Cell***, 17(10): 2705-2722.

Kafieh, R., & Mehridehnavi, A. 2013. A comprehensive comparison of different clustering methods for reliability analysis of microarray data. ***J Med Signals Sens***, 3(1): 22-30.

Kannangara, R., Branigan, C., Liu, Y., Penfield, T., Rao, V., Mouille, G., Hofte, H., Pauly, M., Riechmann, J. L., & Broun, P. 2007. The transcription factor WIN1/SHN1 regulates Cutin biosynthesis in Arabidopsis thaliana. ***Plant Cell***, 19(4): 1278-1294.

Kim, O. K., Jung, J. H., & Park, C. M. 2010a. An Arabidopsis F-box protein regulates tapetum degeneration and pollen maturation during anther development. ***Planta***, 232(2): 353-366.

Kim, S. S., Grienenberger, E., Lallemand, B., Colpitts, C. C., Kim, S. Y., Souza Cde, A., Geoffroy, P., Heintz, D., Krahn, D., Kaiser, M., Kombrink, E., Heitz, T., Suh, D. Y., Legrand, M., & Douglas, C. J. 2010b. LAP6/POLYKETIDE SYNTHASE A and LAP5/POLYKETIDE SYNTHASE B encode hydroxyalkyl alpha-pyrone synthases required for pollen development and sporopollenin biosynthesis in Arabidopsis thaliana. ***Plant Cell***, 22(12): 4045-4066.

Ko, S. S., Li, M. J., Sun-Ben Ku, M., Ho, Y. C., Lin, Y. J., Chuang, M. H., Hsing, H. X., Lien, Y. C., Yang, H. T., Chang, H. C., & Chan, M. T. 2014. The bHLH142 Transcription Factor Coordinates with TDR1 to Modulate the Expression of EAT1 and Regulate Pollen Development in Rice. ***Plant Cell***, 26(6): 2486-2504.

Kolesnikov, N., Hastings, E., Keays, M., Melnichuk, O., Tang, Y. A., Williams, E., Dylag, M., Kurbatova, N., Brandizi, M., Burdett, T., Megy, K., Pilicheva, E., Rustici, G., Tikhonov, A., Parkinson, H., Petryszak, R., Sarkans, U., & Brazma, A. 2015. ArrayExpress update--simplifying data submissions. ***Nucleic Acids Res***, 43(Database issue): D1113-1116.

Kumar, S., Stecher, G., & Tamura, K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. ***Mol Biol Evol***, 33(7): 1870-1874.

Kuriyama, H., & Fukuda, H. 2002. Developmental programmed cell death in plants. ***Curr Opin Plant Biol***, 5(6): 568-573.

Lampard, G. R., & Bergmann, D. C. 2007. A Shout-Out to Stomatal Development: How the bHLH Proteins SPEECHLESS, MUTE and

FAMA Regulate Cell Division and Cell Fate. ***Plant Signal Behav***, 2(4): 290-292.

Lee, S., Jung, K. H., An, G., & Chung, Y. Y. 2004. Isolation and characterization of a rice cysteine protease gene, OsCP1, using T-DNA gene-trap system. ***Plant Mol Biol***, 54: 11.

Li, H., Pinot, F., Sauveplane, V., Werck-Reichhart, D., Diehl, P., Schreiber, L., Franke, R., Zhang, P., Chen, L., Gao, Y., Liang, W., & Zhang, D. 2010. Cytochrome P450 family member CYP704B2 catalyzes the {omega}-hydroxylation of fatty acids and is required for anther cutin biosynthesis and pollen exine formation in rice. ***Plant Cell***, 22(1): 173-190.

Li, H., Yuan, Z., Vizcay-Barrena, G., Yang, C., Liang, W., Zong, J., Wilson, Z. A., & Zhang, D. 2011a. PERSISTENT TAPETAL CELL1 encodes a PHD-finger protein that is required for tapetal cell death and pollen development in rice. ***Plant Physiol***, 156(2): 615-630.

Li, L., Li, Y., Song, S., Deng, H., Li, N., Fu, X., Chen, G., & Yuan, L. 2015. An anther development F-box (ADF) protein regulated by tapetum degeneration retardation (TDR) controls rice anther development. ***Planta***, 241(1): 157-166.

Li, N., Zhang, D. S., Liu, H. S., Yin, C. S., Li, X. X., Liang, W. Q., Yuan, Z., Xu, B., Chu, H. W., Wang, J., Wen, T. Q., Huang, H., Luo, D., Ma, H., & Zhang, D. B. 2006a. The rice tapetum degeneration retardation gene is required for tapetum degradation and anther development. ***Plant Cell***, 18(11): 2999-3014.

Li, X., Duan, X., Jiang, H., Sun, Y., Tang, Y., Yuan, Z., Guo, J., Liang, W., Chen, L., Yin, J., Ma, H., Wang, J., & Zhang, D. 2006b. Genome-wide analysis of basic/helix-loop-helix transcription factor family in rice and Arabidopsis. ***Plant Physiol***, 141(4): 1167-1184.

Li, X., Gao, X., Wei, Y., Deng, L., Ouyang, Y., Chen, G., Li, X., Zhang, Q., & Wu, C. 2011b. Rice APOPTOSIS INHIBITOR5 coupled with two DEAD-box adenosine 5'-triphosphate-dependent RNA helicases regulates tapetum degeneration. ***Plant Cell***, 23(4): 1416-1434.

Li, Y., Li, D., Guo, Z., Shi, Q., Xiong, S., Zhang, C., Zhu, J., & Yang, Z. 2016.

OsACOS12, an orthologue of Arabidopsis acyl-CoA synthetase5, plays an important role in pollen exine formation and anther development in rice. ***BMC Plant Biol***, 16(1): 256.

Li-Beisson, Y., Shorrosh, B., Beisson, F., Andersson, M. X., Arondel, V., Bates, P. D., Baud, S., Bird, D., Debono, A., Durrett, T. P., Franke, R. B., Graham, I. A., Katayama, K., Kelly, A. A., Larson, T., Markham, J. E., Miquel, M., Molina, I., Nishida, I., Rowland, O., Samuels, L., Schmid, K. M., Wada, H., Welti, R., Xu, C., Zallot, R., & Ohlrogge, J. 2010. Acyl-lipid metabolism. ***Arabidopsis Book***, 8: e0133.

Li-Beisson, Y., Shorrosh, B., Beisson, F., Andersson, M. X., Arondel, V., Bates, P. D., Baud, S., Bird, D., Debono, A., Durrett, T. P., Franke, R. B., Graham, I. A., Katayama, K., Kelly, A. A., Larson, T., Markham, J. E., Miquel, M., Molina, I., Nishida, I., Rowland, O., Samuels, L., Schmid, K. M., Wada, H., Welti, R., Xu, C., Zallot, R., & Ohlrogge, J. 2013. Acyl-lipid metabolism. ***Arabidopsis Book***, 11: e0161.

LIN, K. C., WU, T. M., CHANDRIKA, N. N. P., CHOU, S. J., & HONG, C. Y. 2017. Molecular characterization and subcellular localization of salt-inducible lipid transfer proteins in rice. ***Biologia Plantarum***, 61(3): 501-510.

Liu, Z., Lin, S., Shi, J., Yu, J., Zhu, L., Yang, X., Zhang, D., & Liang, W. 2017. Rice No Pollen 1 (NP1) is required for anther cuticle formation and pollen exine patterning. ***Plant J***, 91(2): 263-277.

Londo, J. P., Chiang, Y. C., Hung, K. H., Chiang, T. Y., & Schaal, B. A. 2006. Phylogeography of Asian wild rice, Oryza rufipogon, reveals multiple independent domestications of cultivated rice, Oryza sativa. ***Proc Natl Acad Sci U S A***, 103(25): 9578-9583.

Luo, Q., Li, Y., Shen, Y., & Cheng, Z. 2014. Ten years of gene discovery for meiotic event control in rice. ***J Genet Genomics***, 41(3): 125-137.

Ma, H. 2005. Molecular genetic analyses of microsporogenesis and microgametogenesis in flowering plants. ***Annu Rev Plant Biol***, 56: 393-434.

Men, X., Shi, J., Liang, W., Zhang, Q., Lian, G., Quan, S., Zhu, L., Luo, Z., Chen, M., & Zhang, D. 2017. Glycerol-3-Phosphate Acyltransferase 3

(OsGPAT3) is required for anther development and male fertility in rice. ***J Exp Bot***, 68(3): 513-526.

Mohanta, T. K., Kumar, P., & Bae, H. 2017. Genomics and evolutionary aspect of calcium signaling event in calmodulin and calmodulin-like proteins in plants. ***BMC Plant Biol***, 17(1): 38.

Morant, M., Jorgensen, K., Schaller, H., Pinot, F., Moller, B. L., Werck-Reichhart, D., & Bak, S. 2007. CYP703 is an ancient cytochrome P450 in land plants catalyzing in-chain hydroxylation of lauric acid to provide building blocks for sporopollenin synthesis in pollen. ***Plant Cell***, 19(5): 1473-1487.

Morris, J. H., Apeltsin, L., Newman, A. M., Baumbach, J., Wittkop, T., Su, G., Bader, G. D., & Ferrin, T. E. 2011. clusterMaker: a multi-algorithm clustering plugin for Cytoscape. ***BMC Bioinformatics***, 12: 436.

Mou, Z., He, Y., Dai, Y., Liu, X., & Li, J. 2000. Deficiency in fatty acid synthase leads to premature cell death and dramatic alterations in plant morphology. ***Plant Cell***, 12(3): 405-418.

Murre, C., McCaw, P. S., & Baltimore, D. 1989. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. ***Cell***, 56(5): 777-783.

Niu, N., Liang, W., Yang, X., Jin, W., Wilson, Z. A., Hu, J., & Zhang, D. 2013. EAT1 promotes tapetal cell death by regulating aspartic proteases during male reproductive development in rice. ***Nat Commun***, 4: 1445.

Nonomura, K., Miyoshi, K., Eiguchi, M., Suzuki, T., Miyao, A., Hirochika, H., & Kurata, N. 2003. The MSP1 gene is necessary to restrict the number of cells entering into male and female sporogenesis and to initiate anther wall formation in rice. ***Plant Cell***, 15(8): 1728-1739.

Panikashvili, D., Shi, J. X., Bocobza, S., Franke, R. B., Schreiber, L., & Aharoni, A. 2010. The Arabidopsis DSO/ABCG11 transporter affects cutin metabolism in reproductive organs and suberin in roots. ***Mol Plant***, 3(3): 563-575.

Parish, R. W., & Li, S. F. 2010. Death of a tapetum: A programme of developmental altruism. ***Plant Science***, 178(2): 73-89.

Paxson-Sowders, D. M., Dodrill, C. H., Owen, H. A., & Makaroff, C. A. 2001.

DEX1, a Novel Plant Protein, Is Required for Exine Pattern Formation during Pollen Development in Arabidopsis. ***Plant Physiology***, 127(4): 1739-1749.

Pearce, S., Ferguson, A., King, J., & Wilson, Z. A. 2015. FlowerNet: a gene expression correlation network for anther and pollen development. ***Plant Physiol***, 167(4): 1717-1730.

Phan, H. A., Li, S. F., & Parish, R. W. 2012. MYB80, a regulator of tapetal and pollen development, is functionally conserved in crops. ***Plant Mol Biol***, 78(1-2): 171-183.

Popescu, G. V., Noutsos, C., & Popescu, S. C. 2016. Big Data in Plant Science: Resources and Data Mining Tools for Plant Genomics and Proteomics. ***Methods Mol Biol***, 1415: 533-547.

Pruitt, K. D., Tatusova, T., & Maglott, D. R. 2005. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. ***Nucleic Acids Res***, 33(Database issue): D501-504.

Qin, M., Tian, T., Xia, S., Wang, Z., Song, L., Yi, B., Wen, J., Shen, J., Ma, C., Fu, T., & Tu, J. 2016. Heterodimer Formation of BnPKSA or BnPKSB with BnACOS5 Constitutes a Multienzyme Complex in Tapetal Cells and is Involved in Male Reproductive Development in Brassica napus. ***Plant Cell Physiol***, 57(8): 1643-1656.

Qin, P., Tu, B., Wang, Y., Deng, L., Quilichini, T. D., Li, T., Wang, H., Ma, B., & Li, S. 2013. ABCG15 encodes an ABC transporter protein, and is essential for post-meiotic anther and pollen exine development in rice. ***Plant Cell Physiol***, 54(1): 138-154.

Quilichini, T. D., Douglas, C. J., & Samuels, A. L. 2014. New views of tapetum ultrastructure and pollen exine development in Arabidopsis thaliana. ***Ann Bot***, 114(6): 1189-1201.

Ranjan, R., Khurana, R., Malik, N., Badoni, S., Parida, S. K., Kapoor, S., & Tyagi, A. K. 2017. bHLH142 regulates various metabolic pathway-related genes to affect pollen development and anther dehiscence in rice. ***Sci Rep***, 7: 43397.

Rhee, S. Y., Beavis, W., Berardini, T. Z., Chen, G., Dixon, D., Doyle, A.,

Garcia-Hernandez, M., Huala, E., Lander, G., Montoya, M., Miller, N., Mueller, L. A., Mundodi, S., Reiser, L., Tacklind, J., Weems, D. C., Wu, Y., Xu, I., Yoo, D., Yoon, J., & Zhang, P. 2003a. The Arabidopsis Information Resource (TAIR): a model organism database providing a centralized, curated gateway to Arabidopsis biology, research materials and community. ***Nucleic Acids Res***, 31(1): 224-228.

Rhee, S. Y., Osborne, E., Poindexter, P. D., & Somerville, C. R. 2003b. Microspore separation in the quartet 3 mutants of Arabidopsis is impaired by a defect in a developmentally regulated polygalacturonase required for pollen mother cell wall degradation. ***Plant Physiol***, 133(3): 1170-1180.

Rigden, D. J., & Galperin, M. Y. 2004. The DxDxDG motif for calcium binding: multiple structural contexts and implications for evolution. ***J Mol Biol***, 343(4): 971-984.

Sakai, H., Lee, S. S., Tanaka, T., Numa, H., Kim, J., Kawahara, Y., Wakimoto, H., Yang, C. C., Iwamoto, M., Abe, T., Yamada, Y., Muto, A., Inokuchi, H., Ikemura, T., Matsumoto, T., Sasaki, T., & Itoh, T. 2013. Rice Annotation Project Database (RAP-DB): an integrative and interactive database for rice genomics. ***Plant Cell Physiol***, 54(2): e6.

Sakamoto, T., Miura, K., Itoh, H., Tatsumi, T., Ueguchi-Tanaka, M., Ishiyama, K., Kobayashi, M., Agrawal, G. K., Takeda, S., Abe, K., Miyao, A., Hirochika, H., Kitano, H., Ashikari, M., & Matsuoka, M. 2004. An overview of gibberellin metabolism enzyme genes and their related mutants in rice. ***Plant Physiol***, 134(4): 1642-1653.

Sanders, P. M., Bui, A. Q., Weterings, K., McIntire, K. N., Hsu, Y.-C., Lee, P. Y., Truong, M. T., Beals, T. P., & Goldberg, R. B. 1999. Anther developmental defects in Arabidopsis thaliana male-sterile mutants. ***Sex Plant Reprod***, 11: 297-322.

Sasaki, T. 1997. [Rice genome research as a model for plant genome analysis]. ***Tanpakushitsu Kakusan Koso***, 42(17 Suppl): 2953-2961.

Sato, Y., Namiki, N., Takehisa, H., Kamatsuki, K., Minami, H., Ikawa, H., Ohyanagi, H., Sugimoto, K., Itoh, J., Antonio, B. A., & Nagamura, Y. 2013. RiceFREND: a platform for retrieving coexpressed gene

networks in rice. ***Nucleic Acids Res***, 41(Database issue): D1214-1221.

Scott, R. J., Spielman, M., & Dickinson, H. G. 2004. Stamen structure and function. ***Plant Cell***, 16 Suppl: S46-60.

Shah, A. N., Davey, C. F., Whitebirch, A. C., Miller, A. C., & Moens, C. B. 2015. Rapid reverse genetic screening using CRISPR in zebrafish. ***Nat Methods***, 12(6): 535-540.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. ***Genome Res***, 13(11): 2498-2504.

Sheldon, P. S., Kekwick, R. G., Smith, C. G., Sidebottom, C., & Slabas, A. R. 1992. 3-Oxoacyl-[ACP] reductase from oilseed rape (Brassica napus). ***Biochim Biophys Acta***, 1120(2): 151-159.

Shi, J., Cui, M., Yang, L., Kim, Y. J., & Zhang, D. 2015. Genetic and Biochemical Mechanisms of Pollen Wall Development. ***Trends Plant Sci***, 20(11): 741-753.

Shi, J., Tan, H., Yu, X. H., Liu, Y., Liang, W., Ranathunge, K., Franke, R. B., Schreiber, L., Wang, Y., Kai, G., Shanklin, J., Ma, H., & Zhang, D. 2011. Defective pollen wall is required for anther and microspore development in rice and encodes a fatty acyl carrier protein reductase. ***Plant Cell***, 23(6): 2225-2246.

Shirasawa, S., Kifuji, Y., Komiya, R., Kitashiba, H., Nishimura, M., & Nishio, T. 2013. Identification of a single nucleotide deletion causing frameshift mutation in OsDFR2A in a genic male sterile mutant of rice and its possible application to F1 hybrid breeding. ***Molecular Breeding***, 31(4): 805-814.

Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L., & Ideker, T. 2011. Cytoscape 2.8: new features for data integration and network visualization. ***Bioinformatics***, 27(3): 431-432.

Sorensen, A. M., Krober, S., Unte, U. S., Huijser, P., Dekker, K., & Saedler, H. 2003. The Arabidopsis ABORTED MICROSPORES (AMS) gene encodes a MYC class transcription factor. ***Plant J***, 33(2): 413-423.

Takada, K., Ishimaru, K., Kamada, H., & Ezura, H. 2006. Anther-specific

expression of mutated melon ethylene receptor gene Cm-ERS1/H70A affected tapetum degeneration and pollen grain production in transgenic tobacco plants. ***Plant Cell Rep***, 25(9): 936-941.

Tzeng, J. D., Hsu, S. W., Chung, M. C., Yeh, F. L., Yang, C. Y., Liu, M. C., Hsu, Y. F., & Wang, C. S. 2009. Expression and regulation of two novel anther-specific genes in Lilium longiflorum. ***J Plant Physiol***, 166(4): 417-427.

Usadel, B., Obayashi, T., Mutwil, M., Giorgi, F. M., Bassel, G. W., Tanimoto, M., Chow, A., Steinhauser, D., Persson, S., & Provart, N. J. 2009. Co-expression tools for plant biology: opportunities for hypothesis generation and caveats. ***Plant Cell Environ***, 32(12): 1633-1651.

van Doorn, W. G., & Woltering, E. J. 2005. Many ways to exit? Cell death categories in plants. ***Trends Plant Sci***, 10(3): 117-122.

Wang, Y., Lin, Y. C., So, J., Du, Y., & Lo, C. 2013. Conserved metabolic steps for sporopollenin precursor formation in tobacco and rice. ***Physiol Plant***, 149(1): 13-24.

Wang, Y., Wang, Y. F., & Zhang, D. B. 2006. [Identification of the rice (Oryza sativa L.) mutant msp1-4 and expression analysis of its UDT1 and GAMYB genes]. ***Zhi Wu Sheng Li Yu Fen Zi Sheng Wu Xue Xue Bao***, 32(5): 527-534.

Wang, Z., Liang, Y., Li, C., Xu, Y., Lan, L., Zhao, D., Chen, C., Xu, Z., Xue, Y., & Chong, K. 2005. Microarray analysis of gene expression involved in anther development in rice (Oryza sativa L.). ***Plant Mol Biol***, 58(5): 721-737.

Wilson, Z. A., & Zhang, D. B. 2009. From Arabidopsis to rice: pathways in pollen development. ***J Exp Bot***, 60(5): 1479-1492.

Wittkop, T., Emig, D., Lange, S., Rahmann, S., Albrecht, M., Morris, J. H., Bocker, S., Stoye, J., & Baumbach, J. 2010. Partitioning biological data with transitivity clustering. ***Nat Methods***, 7(6): 419-420.

Wongnate, T., & Chaiyen, P. 2013. The substrate oxidation mechanism of pyranose 2-oxidase and other related enzymes in the glucose-methanol-choline superfamily. ***FEBS J***, 280(13): 3009-3027.

Xu, D., Shi, J., Rautengarten, C., Yang, L., Qian, X., Uzair, M., Zhu, L., Luo,

Q., An, G., Wassmann, F., Schreiber, L., Heazlewood, J. L., Scheller, H. V., Hu, J., Zhang, D., & Liang, W. 2017. Defective Pollen Wall 2 (DPW2) Encodes an Acyl Transferase Required for Rice Pollen Development. ***Plant Physiol***, 173(1): 240-255.

Xu, J., Ding, Z., Vizcay-Barrena, G., Shi, J., Liang, W., Yuan, Z., Werck-Reichhart, D., Schreiber, L., Wilson, Z. A., & Zhang, D. 2014. ABORTED MICROSPORES Acts as a Master Regulator of Pollen Wall Formation in Arabidopsis. ***Plant Cell***, 26(4): 1544-1556.

Xu, J., Yang, C., Yuan, Z., Zhang, D., Gondwe, M. Y., Ding, Z., Liang, W., Zhang, D., & Wilson, Z. A. 2010. The ABORTED MICROSPORES Regulatory Network Is Required for Postmeiotic Male Reproductive Development in Arabidopsis thaliana. ***The Plant Cell***, 22(1): 91-107.

Yang, L., Qian, X., Chen, M., Fei, Q., Meyers, B. C., Liang, W., & Zhang, D. 2016. Regulatory Role of a Receptor-Like Kinase in Specifying Anther Cell Identity. ***Plant Physiol***, 171(3): 2085-2100.

Yang, X., Liang, W., Chen, M., Zhang, D., Zhao, X., & Shi, J. 2017. Rice fatty acyl-CoA synthetase OsACOS12 is required for tapetum programmed cell death and male fertility. ***Planta***, 246(1): 105-122.

Yang, X., Wu, D., Shi, J., He, Y., Pinot, F., Grausem, B., Yin, C., Zhu, L., Chen, M., Luo, Z., Liang, W., & Zhang, D. 2014. Rice CYP703A3, a cytochrome P450 hydroxylase, is essential for development of anther cuticle and pollen exine. ***J Integr Plant Biol***, 56(10): 979-994.

Yeung, K. Y. M., M. ; Bumgarner, R. E. 2004. From co-expression to co-regulation: how many microarray experiments do we need? ***Genome Biology***, 5(7): R48.

Yi, J., Moon, S., Lee, Y. S., Zhu, L., Liang, W., Zhang, D., Jung, K. H., & An, G. 2016. Defective Tapetum Cell Death 1 (DTC1) Regulates ROS Levels by Binding to Metallothionein during Tapetum Degeneration. ***Plant Physiol***, 170(3): 1611-1623.

Yu, J., Meng, Z., Liang, W., Behera, S., Kudla, J., Tucker, M. R., Luo, Z., Chen, M., Xu, D., Zhao, G., Wang, J., Zhang, S., Kim, Y. J., & Zhang, D. 2016. A Rice Ca2+ Binding Protein Is Required for Tapetum Function and Pollen Formation. ***Plant Physiol***, 172(3): 1772-1786.

Zhang, C., Guinel, F. C., & Moffatt, B. A. 2002. A comparative ultrastructural study of pollen development in Arabidopsis thaliana ecotype Columbia and male-sterile mutant apt1-3. ***Protoplasma***, 219(1-2): 59-71.

Zhang, D., & Liang, W. 2016. Improving food security: Using male fertility for hybrid seed breeding. ***Science***, sponsored collection: 45.

Zhang, D., Liang, W., Yin, C., Zong, J., Gu, F., & Zhang, D. 2010a. OsC6, encoding a lipid transfer protein, is required for postmeiotic anther development in rice. ***Plant Physiol***, 154(1): 149-162.

Zhang, D., Luo, X., & Zhu, L. 2011. Cytological analysis and genetic control of rice anther development. ***J Genet Genomics***, 38(9): 379-390.

Zhang, D., & Yang, L. 2014. Specification of tapetum and microsporocyte cells within the anther. ***Curr Opin Plant Biol***, 17: 49-55.

Zhang, D. S., Liang, W. Q., Yuan, Z., Li, N., Shi, J., Wang, J., Liu, Y. M., Yu, W. J., & Zhang, D. B. 2008. Tapetum degeneration retardation is critical for aliphatic metabolism and gene regulation during rice pollen development. ***Mol Plant***, 1(4): 599-610.

Zhang, H., Zhang, J., Wei, P., Zhang, B., Gou, F., Feng, Z., Mao, Y., Yang, L., Zhang, H., Xu, N., & Zhu, J. K. 2014. The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. ***Plant Biotechnol J***, 12(6): 797-807.

Zhang, L., Mao, D., Xing, F., Bai, X., Zhao, H., Yao, W., Li, G., Xie, W., & Xing, Y. 2015. Loss of function of OsMADS3 via the insertion of a novel retrotransposon leads to recessive male sterility in rice (Oryza sativa). ***Plant Sci***, 238: 188-197.

Zhang, S., Fang, Z. J., Zhu, J., Gao, J. F., & Yang, Z. N. 2010b. OsMYB103 is required for rice anther development by regulating tapetum development and exine formation. ***Chinese Science Bulletin***, 55(29): 3288-3297.

Zhao, G., Shi, J., Liang, W., Xue, F., Luo, Q., Zhu, L., Qu, G., Chen, M., Schreiber, L., & Zhang, D. 2015. Two ATP Binding Cassette G Transporters, Rice ATP Binding Cassette G26 and ATP Binding Cassette G15, Collaboratively Regulate Rice Male Reproduction. ***Plant Physiol***, 169(3): 2064-2079.

Zhao, X., de Palma, J., Oane, R., Gamuyao, R., Luo, M., Chaudhury, A., Herve, P., Xue, Q., & Bennett, J. 2008. OsTDL1A binds to the LRR domain of rice receptor kinase MSP1, and is required to limit sporocyte numbers. ***Plant J***, 54(3): 375-387.

Zheng, Z., Xia, Q., Dauk, M., Shen, W., Selvaraj, G., & Zou, J. 2003. Arabidopsis AtGPAT1, a member of the membrane-bound glycerol-3-phosphate acyltransferase gene family, is essential for tapetum differentiation and male fertility. ***Plant Cell***, 15(8): 1872-1887.

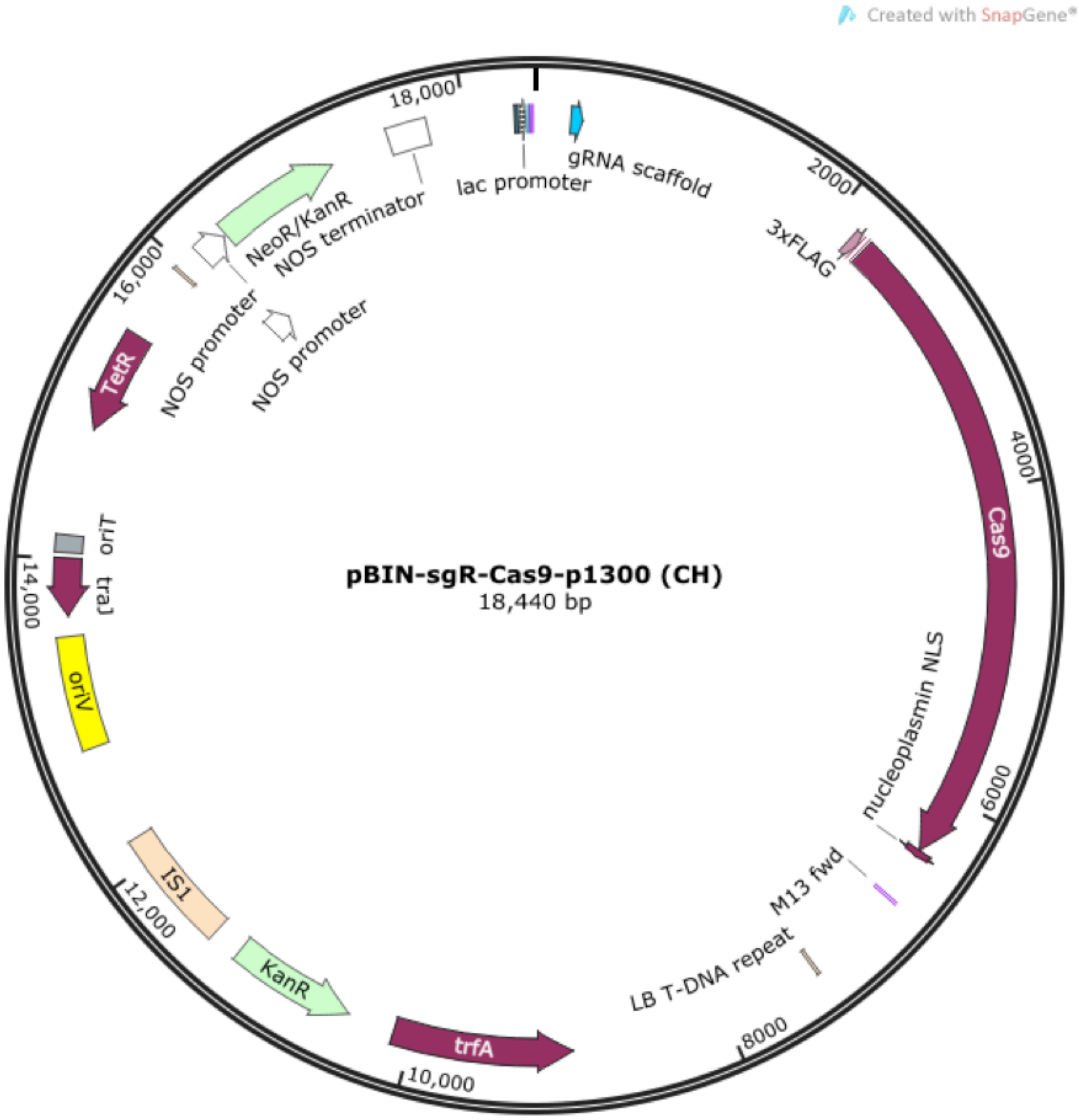
Zhu, E., You, C., Wang, S., Cui, J., Niu, B., Wang, Y., Qi, J., Ma, H., & Chang, F. 2015. The DYT1-interacting proteins bHLH010, bHLH089 and bHLH091 are redundantly required for Arabidopsis anther development and transcriptome. ***Plant J***, 83(6): 976-990.

Zinkl, G. M., Zwiebel, B. I., Grier, D. G., & Preuss, D. 1999. Pollen-stigma adhesion in Arabidopsis: a species-specific interaction mediated by lipophilic molecules in the pollen exine. ***Development***, 126(23): 5431-5440.

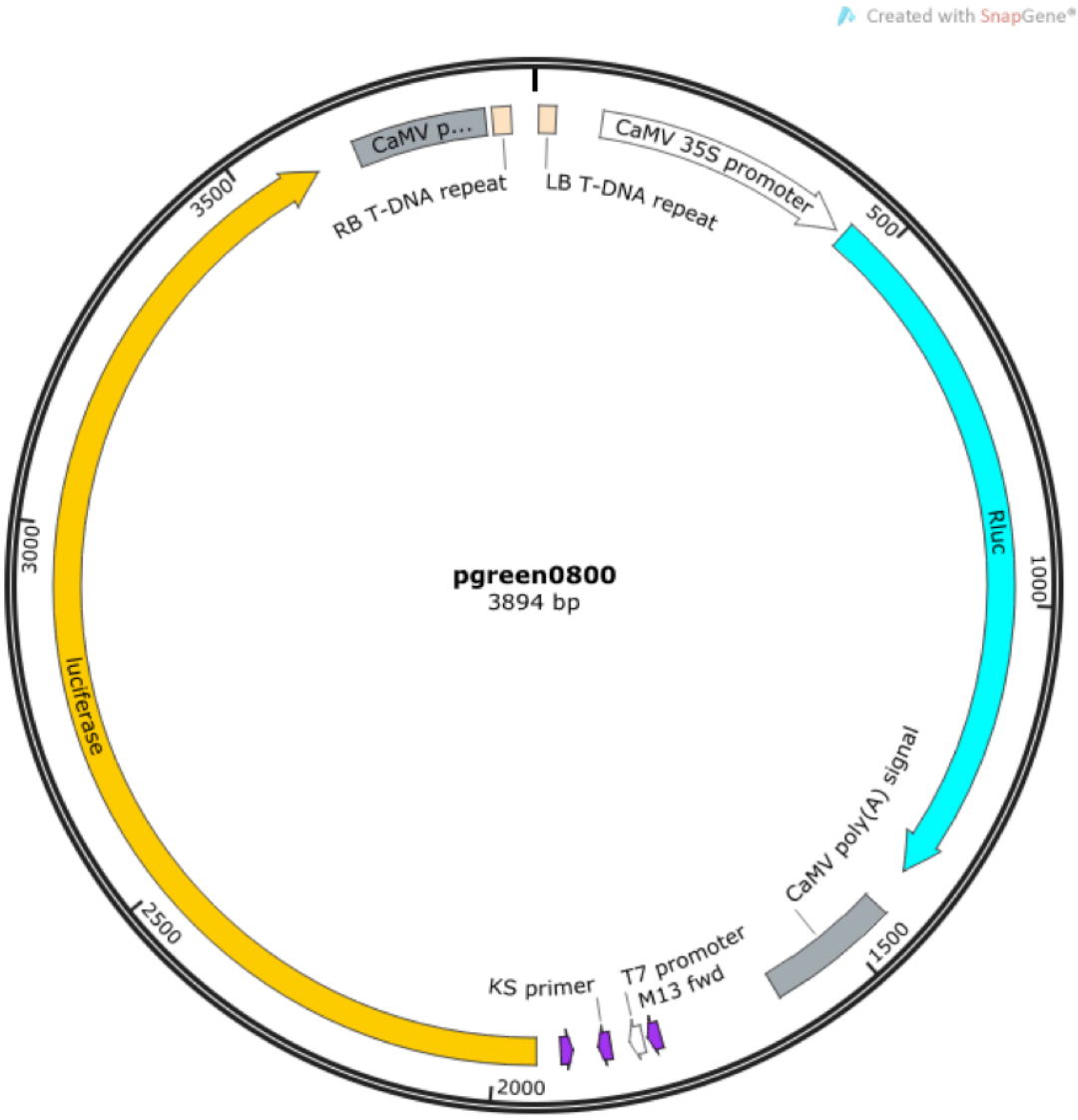
APPENDIX

Appendix I. VECTORS

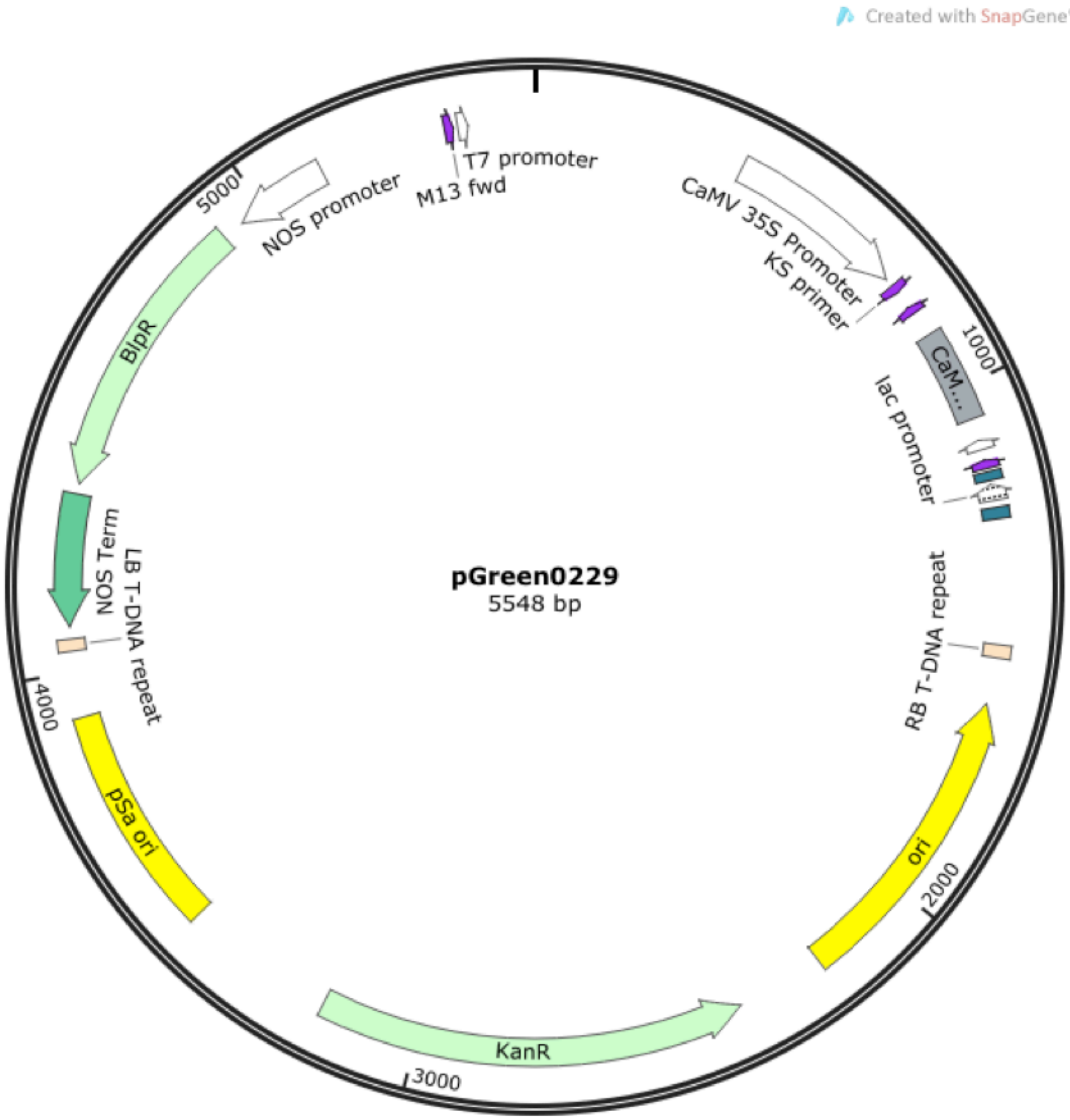
Map of pBin-sgR-Cas9-p1300 (CH)



Map of pGreen0800



Map of pGreen0229



Map of pSoup


Created with SnapGene®
pSoup
9275 bp
TetR
traJ
oriT
oriV
lac promoter
MCS
tet promoter
TcR
trfA
RepA
1000
2000
3000
4000
5000
6000
7000
8000
9000

Appendix II. Primers

CRISPR Primers		
Name	Sequence (5'…3')	RAP_ID
LAP6_Cri_F1	TGGCGCAGGAGCCAATGCTCCAGG	Os10g0484800
LAP6_Cri_R1	AAACCCTGGAGCATTGGCTCCTGC	Os10g0484800
LAP6_Cri_F2	TGGCGGAGGCGAAAGGATACCGGG	Os10g0484800
LAP6_Cri_R2	AAACCCCGGTATCCTTTCGCCTCC	Os10g0484800
GAMYBL2_Cri_F1	TGGC GGACAGCCATGTACTGCCCT	Os06g0604000
GAMYBL2_Cri_R1	AAACAGGGCAGTACATGGCTGTCC	Os06g0604000
GAMYBL2_Cri_F2	TGGCGGAGAGGACATTGCTCCCCG	Os06g0604000
GAMYBL2_Cri_R2	AAACCGGGGAGCAATGTCCTCTCC	Os06g0604000
LRR5_Cri_F1	TGGCGTCGGTCGATCCTCAGGTCG	Os05g0169600
LRR5_Cri_R1	AAACCGACCTGAGGATCGACCGAC	Os05g0169600
LRR5_Cri_F2	TGGCGGTCTCGCTTCCGGCAAAGG	Os05g0169600
LRR5_Cri_R2	AAACCCTTTGCCGGAAGCGAGACC	Os05g0169600
LRR1_Cri_F1	TGGCGTGCCCTGGAAAAGTCCCGG	Os01g0170300
LRR1_Cri_R1	AAACCCGGGACTTTTCCAGGGCAC	Os01g0170300
LRR1_Cri_F2	TGGCGTACCCATTGCACCACAGTT	Os01g0170300
LRR1_Cri_R2	AAACAACTGTGGTGCAATGGGTAC	Os01g0170300
CML41_Cri_F1	TGGCGAACGCGAGTGTTCCTAAGC	Os03g0769500
CML41_Cri_R1	AAACGCTTAGGAACACTCGCGTTC	Os03g0769500
CML41_Cri_F2	TGGCGGATCTTCCAGCACTTCGAC	Os03g0769500
CML41_Cri_R2	AAACGTCGAAGTGCTGGAAGATCC	Os03g0769500
DRL1_Cri_F1	TGGCGGGCACCTGTGGAATTTAAC	Os08g0515900
DRL1_Cri_R1	AAACGTTAAATTCCACAGGTGCCC	Os08g0515900
DRL1_Cri_F2	TGGCGGAAGCGACTTCTTGAGTCA	Os08g0515900
DRL1_Cri_R2	AAACTGACTCAAGAAGTCGCTTCC	Os08g0515900
DRL2_Cri_F1	TGGCGTTACCTGGGTCTCTGACGG	Os01g0127500
DRL2_Cri_R1	AAACCCGTCAGAGACCCAGGTAAC	Os01g0127500
DRL2_Cri_F2	TGGCGTCCCATAGGAACCCCACCT	Os01g0127500
DRL2_Cri_R2	AAACAGGTGGGGTTCCTATGGGAC	Os01g0127500

GMCO_Cri_F1	TGGCGTACGTGAGGGCCCTCGGGT	Os02g0678300
GMCO_Cri_R1	AAACACCCGAGGGCCCTCACGTAC	Os02g0678300
GMCO_Cri_F2	TGGCGCCAAACAGAAACAGGGTCG	Os02g0678300
GMCO_Cri_R2	AAACCGACCCTGTTTCTGTTTGGC	Os02g0678300
LTPL1_Cri_F1	TGGCGCTCGTCGGGGGGACTAGGC	Os03g0357500
LTPL1_Cri_R1	AAACGCCTAGTCCCCCCGACGAGC	Os03g0357500
LTPL1_Cri_F2	TGGCGTTGATTACCGAGCTGAGTT	Os03g0357500
LTPL1_Cri_R2	AAACAACTCAGCTCGGTAATCAAC	Os03g0357500
LTPL150_Cri_F1	TGGCGAAGACCCTCTTGGTGCGAT	Os01g0691300
LTPL150_Cri_R1	AAACATCGCACCAAGAGGGTCTTC	Os01g0691300
LTPL150_Cri_F2	TGGCGTCGCACGAAGCCTTCTTCG	Os01g0691300
LTPL150_Cri_R2	AAACCGAAGAAGGCTTCGTGCGAC	Os01g0691300
LTPL18_Cri_F1	TGGCGTCACCGTATAGTAGAGATC	Os01g0219500
LTPL18_Cri_R1	AAACGATCTCTACTATACGGTGAC	Os01g0219500
LTPL18_Cri_F2	TGGCGCACAGATCGATGGCCTCTC	Os01g0219500
LTPL18_Cri_R2	AAACGAGAGGCCATCGATCTGTGC	Os01g0219500

Genotyping Primers		
Name	Sequence	RAP_ID
LAP6_CriGT_F1	CGTCCCCTGCAATCCTGAAT	Os10g0484800
LAP6_CriGT_R1	CACACAGTGATGCAGCCAAG	Os10g0484800
LRR5_CriGT_F1	CGGATCCTACGGCTACATCG	Os05g0169600
LRR5_CriGT_R1	AGGAACCGGAATTAGCACGG	Os05g0169600
LRR1_CriGT_F1	CTGCTCATCGACAACACCCT	Os01g0170300
LRR1_CriGT_R1	GTACGAGATGTTGAGCGCGA	Os01g0170300
LRR1_CriGT_F2	TGGTTACAAGGAGCACCGTC	Os01g0170300
LRR1_CriGT_R2	CGGTTGTTGAGTAGGCCAGT	Os01g0170300
GAMYBL2_CriGT_F1	TCCGTGCTTACTGGTGTGAC	Os03g0733600
GAMYBL2_CriGT_R1	CCGGAGAAAGAATCGGTGGT	Os03g0733600
GAMYBL2_CriGT_F2	AAAGCCCCAAATCTCCGGTT	Os03g0733600
GAMYBL2_CriGT_R2	GCAATTGGCTTGAAGGGAGC	Os03g0733600
CML41_CriGT_F1	GATGGCACTGTCTGTTGGGA	Os03g0769500
CML41_CriGT_R1	TCCATCAGGGCAACGAATCC	Os03g0769500
DRL1_CriGT_F1	CACAGGCCGTTGTGAAATCT	Os08g0515900
DRL1_CriGT_R1	TGCAGGGTGAAAGTGGTCTT	Os08g0515900
DRL2_CriGT_F1	CGGCCAACGGAGTTTCTTTC	Os01g0127500
DRL2_CriGT_R1	CTCCAGTGCGACTCGTTCAG	Os01g0127500
GMCO_CriGT_F1	ACACGACGCACATTCCTGAT	Os02g0678300
GMCO_CriGT_R1	GTCAAAGATGGAGCCACCGA	Os02g0678300
LBD2_LP	ATTACCATGCATGGATCTTGC	AT1g06280
LBD2_RP	AGAGAGCTTGGTACGGTCCTC	AT1g06280
LTPL1_CriGT_F1	GCCTTCTGCATCAGCCAGTT	Os03g0357500
LTPL1_CriGT_R1	AGCTCGGTAATCAACCACGA	Os03g0357500
LTPL1_CriGT_F2	CAGGAGGTGTCCAAGGACTG	Os03g0357500
LTPL1_CriGT_R2	GTCTCCTACTCTCGGTTCGC	Os03g0357500
LTPL150_CriGT_F1	GCAAGAGATCCATCAAGCCG	Os01g0691300
LTPL150_CriGT_R1	GCAGCAAGGTTGGAATGGTT	Os01g0691300
LTPL18_CriGT_F1	GCAGGGTTACAGAGGAACTGA	Os01g0219500
LTPL18_CriGT_R1	TCTGGGTCATCGGCTACAAG	Os01g0219500

LTPL18_CriGT_F2	AGATAGTTAACGCGGCCCAC	Os01g0219500
LTPL18_CriGT_R2	CTGGTTGCAGTCCATGTTGG	Os01g0219500

qRT-PCR Primers		
Name	Sequence	RAP_ID
OsActin_F	CCTTCAACACCCCTGCTATG	Os03g0718100
OsActin_R	CAATGCCAGGGAACATAGTG	Os03g0718100
CML41_RT_F	GGATTCGTTGCCCTGATGGA	Os03g0769500
CML41_RT_R	CTGATCATGGCCTCGCAGT	Os03g0769500
LBD2_RT_F	AACCAGCCTCTCATTGGTTG	AT1G06280
LBD2_RT_R	CACTGCATAACATTGTTGCC	AT1G06280

Primers for Dual-Luciferase Reporter Assay		
Name	Sequence	RAP_ID
CMLPro_0800F	TATAGGGCGAATTGGGTACCTCGCTAGCTTAGCTTCTCC	Os03g0769500
CMLPro_0800R	TAGAACTAGTGGATCCTCAAGAAACCCAAGAAAATG	Os03g0769500
LTPL1_0800F	TATAGGGCGAATTGGGTACCAAGATCGGCAAGCATGCAGC	Os03g0357500
LTPL1_0800R	TAGAACTAGTGGATCCGGCCTCACTCTCTGGCTAAG	Os03g0357500
LTPL18_0800F	TATAGGGCGAATTGGGTACCGCCATGGCGATGGCGAGCGA	Os01g0219500
LTPL18_0800R	TAGAACTAGTGGATCCATGATTCGGTATTTGATTTG	Os01g0219500

Appendix III. Microarray datasets

5 wild type datasets (57 arrays) from GEO used for generating RiceAntherNet					
Array	Experiment	Sample ID	Sample name	Anther Stage	Stage Name
1	GSE7951	GSM195227	Anther	An12	MP (Mature Pollen)
2	GSE14304	GSM351427	An1.1	An2	PM (Premeiosis)
3	GSE14304	GSM351428	An1.2	An2	PM (Premeiosis)
4	GSE14304	GSM351429	An1.3	An2	PM (Premeiosis)
5	GSE14304	GSM351430	Mei1.1	An4-7	M (Meiosis)
6	GSE14304	GSM351431	Mei1.2	An4-7	M (Meiosis)
7	GSE14304	GSM351432	Mei1.3	An4-7	M (Meiosis)
8	GSE14304	GSM351433	M1.1	An7-8a	M (Meiosis)
9	GSE14304	GSM351434	M1.2	An7-8a	M (Meiosis)
10	GSE14304	GSM351435	M1.3	An7-8a	M (Meiosis)
11	GSE14304	GSM351436	M1.4	An7-8a	M (Meiosis)
12	GSE14304	GSM351437	M2.1	An8a	M (Meiosis)
13	GSE14304	GSM351438	M2.2	An8a	M (Meiosis)
14	GSE14304	GSM351439	M2.3	An8a	M (Meiosis)
15	GSE14304	GSM351440	M3.1	An8	M (Meiosis)
16	GSE14304	GSM351441	M3.2	An8	M (Meiosis)
17	GSE14304	GSM351442	M3.3	An8	M (Meiosis)
18	GSE14304	GSM351443	M3.4	An8	M (Meiosis)
19	GSE14304	GSM351444	P1.1	An9-10	UM (Uninucleate Microspore)
20	GSE14304	GSM351445	P1.2	An9-10	UM (Uninucleate Microspore)
21	GSE14304	GSM351446	P1.3	An9-10	UM (Uninucleate Microspore)
22	GSE14304	GSM351447	P2.1	An11	BP (Bicellula Pollen)
23	GSE14304	GSM351448	P2.2	An11	BP (Bicellula Pollen)
24	GSE14304	GSM351449	P2.3	An11	BP (Bicellula Pollen)
25	GSE14304	GSM351450	P3.1	An12	TP (Tricellular Pollen)
26	GSE14304	GSM351451	P3.2	An12	TP (Tricellular Pollen)
27	GSE14304	GSM351452	P3.3	An12	TP (Tricellular Pollen)

28	GSE27726	GSM686458	PreMeio.1	An3-5	PM (Premeiosis)
29	GSE27726	GSM686459	PreMeio.2	An3-5	PM (Premeiosis)
30	GSE27726	GSM686460	PreMeio.3	An3-5	PM (Premeiosis)
31	GSE27726	GSM686461	Meio.1	An6-8	M (Meiosis)
32	GSE27726	GSM686462	Meio.2	An6-8	M (Meiosis)
33	GSE27726	GSM686463	Meio.3	An6-8	M (Meiosis)
34	GSE27726	GSM686464	AntherSCP.1	An9-10	UM (Uninucleate Microspore)
35	GSE27726	GSM686465	AntherSCP.2	An9-10	UM (Uninucleate Microspore)
36	GSE27726	GSM686466	AntherSCP.3	An9-10	UM (Uninucleate Microspore)
37	GSE27726	GSM686467	AntherTCP.1	An12-14	TP (Tricellular Pollen)
38	GSE27726	GSM686468	AntherTCP.2	An12-14	TP (Tricellular Pollen)
39	GSE27726	GSM686469	AntherTCP.3	An12-14	TP (Tricellular Pollen)
40	GSE27988	GSM692536	UNM.1	An9-10	UM (Uninucleate Microspore)
41	GSE27988	GSM692537	UNM.2	An9-10	UM (Uninucleate Microspore)
42	GSE27988	GSM692538	UNM.3	An9-10	UM (Uninucleate Microspore)
43	GSE27988	GSM692539	BCP.1	An11	BP (Bicellula Pollen)
44	GSE27988	GSM692540	BCP.2	An11	BP (Bicellula Pollen)
45	GSE27988	GSM692541	BCP.3	An11	BP (Bicellula Pollen)
46	GSE27988	GSM692542	TCP.1	An12	TP (Tricellular Pollen)
47	GSE27988	GSM692543	TCP.2	An12	TP (Tricellular Pollen)
48	GSE27988	GSM692544	TCP.3	An12	TP (Tricellular Pollen)
49	GSE27988	GSM692545	MPG.1	An13-14	MP (Mature Pollen)
50	GSE27988	GSM692546	MPG.2	An13-14	MP (Mature Pollen)
51	GSE27988	GSM692547	MPG.3	An13-14	MP (Mature Pollen)
52	GSE29080	GSM720598	PollenPa.1	An9-10	UM (Uninucleate Microspore)
53	GSE29080	GSM720599	PollenPa.2	An9-10	UM (Uninucleate Microspore)
54	GSE29080	GSM720600	PollenPb.1	An11	BP (Bicellula Pollen)
55	GSE29080	GSM720601	PollenPb.2	An11	BP (Bicellula Pollen)
56	GSE29080	GSM720602	PollenPc.1	An12	TP (Tricellular Pollen)
57	GSE29080	GSM720603	PollenPc.2	An12	TP (Tricellular Pollen)

1 wild type dataset (45 arrays) from GEO used for global expression pattern analysis				
Array	Experiment	Sample	Description	Sample name
1	GSE6893	GSM159177	Root, 1	R
2	GSE6893	GSM159178	Root, 2	R
3	GSE6893	GSM159179	Root, 3	R
4	GSE6893	GSM159180	Mature leaf, 1	ML
5	GSE6893	GSM159181	Mature leaf, 2	ML
6	GSE6893	GSM159182	Mature leaf, 3	ML
7	GSE6893	GSM159183	Young Leaf, 1	YL
8	GSE6893	GSM159184	Young Leaf, 2	YL
9	GSE6893	GSM159185	Young Leaf, 3	YL
10	GSE6893	GSM159186	Shoot apical meristem (SAM, upto 0.5 mm), 1	SAM
11	GSE6893	GSM159187	Shoot apical meristem (SAM, upto 0.5 mm), 2	SAM
12	GSE6893	GSM159188	Shoot apical meristem (SAM, upto 0.5 mm), 3	SAM
13	GSE6893	GSM159189	Inflorescence (P1, upto 3 cm), 1	P1
14	GSE6893	GSM159190	Inflorescence (P1, upto 3 cm), 2	P2
15	GSE6893	GSM159191	Inflorescence (P1, upto 3 cm), 3	P3
16	GSE6893	GSM159192	Inflorescence (P2, 3 - 5 cm), 1	P2
17	GSE6893	GSM159193	Inflorescence (P2, 3 - 5 cm), 2	P2
18	GSE6893	GSM159194	Inflorescence (P2, 3 - 5 cm), 3	P2
19	GSE6893	GSM159195	Inflorescence (P3, 5 - 10 cm),1	P3
20	GSE6893	GSM159196	Inflorescence (P3, 5 - 10 cm), 2	P3
21	GSE6893	GSM159197	Inflorescence (P3, 5 - 10 cm), 3	P3
22	GSE6893	GSM159198	Inflorescence (P4, 10 - 15 cm), 1	P4
23	GSE6893	GSM159199	Inflorescence (P4, 10 - 15 cm), 2	P4
24	GSE6893	GSM159200	Inflorescence (P4, 10 - 15 cm), 3	P4
25	GSE6893	GSM159201	Inflorescence (P5, 15 - 22 cm), 1	P5
26	GSE6893	GSM159202	Inflorescence (P5, 15 - 22 cm), 2	P5
27	GSE6893	GSM159203	Inflorescence (P5, 15 - 22 cm), 3	P5
28	GSE6893	GSM159204	Inflorescence (P6, 22 - 30 cm), 1	P6
29	GSE6893	GSM159205	Inflorescence (P6, 22 - 30 cm), 2	P6

30	GSE6893	GSM159206	Inflorescence (P6, 22 - 30 cm), 3	P6
31	GSE6893	GSM159207	Seed (S1, 0 - 2 dap), 1	S1
32	GSE6893	GSM159208	Seed (S1, 0 - 2 dap), 2	S1
33	GSE6893	GSM159209	Seed (S1, 0 - 2 dap), 3	S1
34	GSE6893	GSM159210	Seed (S2, 3 - 4 dap), 1	S2
35	GSE6893	GSM159211	Seed (S2, 3 - 4 dap), 2	S2
36	GSE6893	GSM159212	Seed (S2, 3 - 4 dap), 3	S2
37	GSE6893	GSM159213	Seed (S3, 5 - 10 dap), 1	S3
38	GSE6893	GSM159214	Seed (S3, 5 - 10 dap), 2	S3
39	GSE6893	GSM159215	Seed (S3, 5 - 10 dap), 3	S3
40	GSE6893	GSM159216	Seed (S4, 11 - 20 dap), 1	S4
41	GSE6893	GSM159217	Seed (S4, 11 - 20 dap), 2	S4
42	GSE6893	GSM159218	Seed (S4, 11 - 20 dap), 3	S4
43	GSE6893	GSM159219	Seed (S5, 21 - 29 dap), 1	S5
44	GSE6893	GSM159220	Seed (S5, 21 - 29 dap), 2	S5
45	GSE6893	GSM159221	Seed (S5, 21 - 29 dap), 3	S5

Fold changes (Log2) of DEGs in Clade 1, Clade 2 and Clade 3																			
RAP ID	Cluster	Clade	*eat1-1* *S9*	*eat1-1* *S10*	*gamyb* *S9*	*msp1-4* *S1*	*msp1-4* *S3*	*msp1-4* *S5*	*ostdl1a* *S1*	*ostdl1a* *S3*	*ostdl1a* *S5*	*tdr* *S8a*	*tdr* *S8b*	*tdr* *S9*	*tip2* *S6*	*tip2* *S7*	*udt1* *S9*	*ptc1* *S9*	*mads3* *S9*
Os03g0564600	60	1.1	-1.7	-5.5	-3.8								-3.1					-3.5	1.9
Os11g0215400	60	1.1	-1.7	-7.8									-6.9	-6.7			-2.5		
Os08g0298600	60	1.1	3.2		-3.2								-4.3	-3.5					
Os01g0772100	60	1.1		-6.3	-6.6								-4.6	-5.6				-9.8	1.6
Os07g0585250	60	1.1		-5.6									-4.9	-4.6					
Os09g0401000	60	1.1		-5.5	-4.9								-5.0	-5.1				-8.9	1.4
Os03g0212000	60	1.1		-5.2	-6.8								-3.8	-2.2			-2.9		
Os12g0131900	60	1.1		-4.9	-4.9								-1.8					-6.1	
Os05g0169600	60	1.1		-4.2	-5.8								-3.6	-3.5				-6.7	1.4
Os04g0448500	60	1.1		-3.1	-3.9								-5.0	-2.7			-4.4		4.5
Os03g0261100	60	1.1		-3.1	-4.2								-5.0	-4.7				-2.3	
Os10g0337400	60	1.1			-3.3	1.5							-2.5						
Os07g0556800	88	1.1	-3.4	-3.6	-4.6							-4.1	-7.8	-7.2				-5.7	1.4
Os02g0110000	88	1.1	1.4		-6.6							-5.3	-7.7	-6.7			-3.8	-8.4	
Os01g0242400	88	1.1		-4.7	-4.2							-1.8	-5.6	-3.0					
Os03g0850900	88	1.1			-0.6								-1.2					2.3	1.5
Os04g0613200	88	1.1			-3.1							-1.7	-4.8	-2.8			-5.0	-3.3	1.2

Os06g0228600	88	1.1			-1.7							-1.0	-4.1						
Os10g0524500	88	1.1			0.3							-4.2	-6.5	-4.2					6.2
Os02g0503500	88	1.1			-2.1								-2.7						2.5
Os08g0360700	88	1.1			0.0								-5.8	-3.9			-4.8		
Os04g0165000	88	1.1											-6.4	-2.7			-6.8	1.3	
Os09g0273800	152	1.1	1.4										-2.3						
Os08g0375700	152	1.1	1.1	-1.7									-3.7					-5.4	
Os01g0641200	152	1.1		-1.5									-3.5						
Os04g0670500	152	1.1		-1.5									-3.0					-4.9	
Os02g0219000	152	1.1		-1.5									-2.3						
Os04g0267000	152	1.1		-1.4									-3.0					-5.7	1.0
Os03g0783100	152	1.1							-1.2				-1.9						
Os01g0112400	225	1.1		-8.8	-5.2								-7.7	-8.2			-1.5	-9.8	1.7
Os01g0825000	225	1.1		-4.3	-2.5									-1.4				-4.1	
Os12g0242700	225	1.1		-4.0	-5.9								-5.4	-6.0			-3.2	-6.2	
Os04g0543700	225	1.1		-3.7	-6.1								-6.9	-6.9			-5.0	-9.3	1.9
Os01g0170300	225	1.1		-3.5	-5.8								-2.6	-3.8			-2.1	-4.5	
Os02g0653200	225	1.1			-2.1								-1.1						
Os02g0790600	385	1.1	-1.2	-3.2	-2.7								-3.8	-2.7					
Os06g0611400	385	1.1		-6.4	-3.1								-2.4						4.0

Os03g0145900	385	1.1		-5.2	-2.3								-2.6	-2.0				-3.5	
Os03g0701400	385	1.1		-1.7	-2.7								-2.9						
Os03g0167600	116	1.2		-2.1	-3.4								-5.7	-6.3			-3.5	-3.0	2.2
Os10g0576900	116	1.2			-7.8								-4.4	-3.9					1.4
Os07g0655800	116	1.2			-3.7							-4.8	-8.2	-8.2			-3.5		
Os06g0607700	116	1.2			-4.9							-2.9	-6.3	-5.6			-4.5		2.9
Os10g0484800	116	1.2			-4.6							-5.7	-8.1	-8.1					2.8
Os04g0310700	116	1.2										-4.7	-7.6	-7.0					
Os01g0252250	116	1.2										-3.3	-6.9	-7.1					
Os06g0604000	116	1.2																	1.6
Os02g0678300	121	1.2		-4.0									-7.6	-6.7		-1.6	-6.2		
Os03g0769500	121	1.2		-2.1	-3.9								-5.9	-5.3			-1.4		
Os01g0121900	121	1.2		-1.7									-3.4	-2.5					
Os08g0131100	121	1.2			-4.0							-4.6	-7.1	-6.4		-1.9			3.3
Os02g0596200	121	1.2			-2.4							-3.5	-6.2	-3.2					5.5
Os01g0127500	121	1.2			-3.6							-3.5	-6.7	-4.9			-3.5		
Os08g0515900	121	1.2										-4.6	-6.8	-5.3	-1.4	-2.1	-5.7		4.5
Os07g0411300	121	1.2										-3.1	-5.8	-3.5			-3.8	2.8	5.5
Os04g0404400	281	1.2		-2.6	-4.0							-2.6	-5.6	-5.9		-3.6	-3.5		2.6
Os03g0263600	281	1.2			-1.6								-5.6	-2.6					2.8

Os02g0101900	281	1.2			-3.7							-2.5	-4.7	-4.6		-2.3	-2.8	-4.3	3.0
Os03g0687700	281	1.2										-7.0	-7.8	-5.0		-2.1	-3.4		4.5
Os03g0781600	281	1.2										-5.6	-7.2	-5.4			-2.4		
Os01g0213800	8	2	-1.1										-1.5						
Os07g0556000	8	2		-1.0															
Os02g0600200	8	2				-1.4	-1.3												
Os09g0489500	8	2						-2.9			-3.4								
Os08g0302000	8	2						-1.6			-1.6	-2.9	-4.5		-4.8	-6.0			2.8
Os03g0161100	8	2										-2.1	-1.2						
Os01g0631700	8	2										-1.3							
Os01g0125600	8	2										-1.0							
Os08g0506400	8	2											-1.7						1.0
Os09g0363900	8	2											-1.7						2.0
Os03g0332900	8	2											-1.6						
Os04g0685200	8	2											-1.5						1.4
Os02g0794500	8	2											-1.4						1.6
Os05g0582800	8	2											-1.4						
Os01g0957100	8	2											-1.3						2.0
Os04g0631200	8	2											-1.2						1.3
Os03g0758500	8	2											-1.1						1.2

Os05g0394200	8	2											-1.0						
Os03g0183800	8	2												1.5					
Os02g0757700	8	2												1.0					
Os01g0130400	8	2															-3.1		
Os06g0645500	8	2																	1.0
Os07g0501800	8	2																	1.0
Os02g0777400	8	2																	1.0
Os09g0456100	8	2																	1.0
Os02g0730400	8	2																	1.1
Os03g0756300	8	2																	1.2
Os03g0703100	8	2																	1.7
Os06g0643500	8	2																	1.9
Os02g0629800	8	2																	2.4
Os03g0239000	28	2	-1.5									-1.4							1.5
Os02g0226600	28	2	-1.4	-1.3	2.6									2.1					
Os03g0221500	28	2		-1.2															
Os05g0467700	28	2																	-1.0
Os06g0265400	28	2																	1.3
Os10g0347000	28	2																	1.5
Os04g0553500	28	2																	1.6

Os03g0815100	31	2	2.1			-2.9	-2.0							1.3	-1.8				
Os04g0244800	31	2	1.7			-2.7	-1.6		-2.3	-1.7			1.0			1.4			1.1
Os04g0591100	31	2	1.5			-2.6	-1.8		-1.5	-1.1				1.3					
Os01g0256500	31	2	1.3			-1.8	-1.7			-1.7									
Os04g0584600	31	2	1.1			-1.3			-1.3										
Os04g0683900	31	2				-1.8			-1.5										
Os07g0568700	31	2				-1.2													
Os11g0207200	31	2				-1.1													
Os03g0244200	31	2												1.3					
Os06g0708000	31	2												1.2					
Os02g0693000	31	2												1.1					
Os06g0697000	31	2													-1.3				
Os02g0549700	31	2															1.1		
Os05g0517400	32	2					-1.7	-4.6		-1.6	-6.2	-2.8	-4.0	-6.7	-1.4	-3.1	-4.9		2.8
Os08g0562500	32	2											-1.8						1.6
Os03g0254000	32	2											-1.5						
Os08g0524400	32	2											-1.4						1.0
Os01g0756600	32	2											-1.3						1.8
Os01g0514700	32	2											-1.1						
Os04g0418800	32	2												1.2					1.4

Os01g0323000	32	2															-1.3		
Os04g0540600	32	2																	1.1
Os03g0703000	32	2																	1.2
Os04g0233400	32	2																	1.5
Os07g0145400	42	2	-1.0										-1.3						1.5
Os07g0591700	42	2					-1.2			-1.0									
Os01g0863500	42	2										-1.2						-2.1	
Os04g0507700	42	2											-1.9						3.7
Os11g0523700	42	2											-1.5						1.4
Os04g0396800	42	2											-1.4						
Os03g0683800	42	2											-1.2						2.9
Os09g0423500	42	2																	1.0
Os05g0507000	42	2																	1.1
Os05g0419800	42	2																	1.4
Os06g0320500	56	2	-1.9										-1.3						
Os04g0414700	56	2	-1.9					-1.5			-3.1					1.1			
Os03g0592500	56	2	-1.7					-1.9			-3.4					1.0			
Os03g0778100	56	2	-1.0								-1.0		-1.2						
Os07g0148900	56	2				-1.2		-1.1			-1.7					1.1	-1.5		
Os01g0501800	56	2						-1.0			-1.4								

Os04g0457000	56	2									-1.5		-1.2						
Os02g0731600	56	2									-1.3								
Os07g0435300	56	2									-1.1								
Os03g0343900	56	2											-1.7						
Os10g0496900	56	2											-1.2						
Os02g0691500	149	2											-1.3						1.4
Os04g0532400	149	2															1.8		
Os12g0599800	149	2																	1.1
Os08g0566100	149	2																	1.1
Os06g0638500	149	2																	1.3
Os07g0501700	1	3	1.9	4.0															-1.3
Os04g0669300	1	3	1.0																
Os02g0663800	1	3		-1.7															
Os02g0780500	1	3		-1.0															
Os07g0243200	1	3		-1.0															
Os08g0528200	1	3		1.4	-1.2														
Os03g0184400	1	3			-5.2									-2.0					-1.5
Os10g0396300	1	3				1.1							1.6	1.1					
Os05g0473400	1	3					1.8												
Os01g0323100	1	3										1.3	1.9						

Os06g0143700	1	3										1.3	1.4					-1.6	
Os10g0465700	1	3											1.5	2.8					-2.8
Os01g0917900	1	3											1.3						
Os03g0178500	1	3											1.3						
Os01g0900900	1	3											1.0	1.7					
Os01g0581400	1	3															-1.2		
Os02g0150200	1	3															-4.1		
Os02g0188600	1	3															-2.3		-1.4
Os02g0152300	1	3																2.2	
Os02g0440900	1	3																-2.6	
Os06g0222900	1	3																-3.0	
Os03g0745700	1	3																	-3.1
Os08g0144100	1	3																	-2.8
Os04g0413100	1	3																	-2.7
Os05g0543000	1	3																	-1.8
Os01g0306200	1	3																	-1.6
Os03g0722000	1	3																	-1.5
Os11g0195700	1	3																	-1.2
Os04g0638800	1	3																	-1.2
Os11g0658700	1	3																	-1.2

Os05g0545200	1	3																	-1.1
Os04g0689300	1	3																	1.1
Os05g0590100	1	3																	1.3
Os06g0158800	1	3																	1.5
Os02g0529500	2	3		1.3															
Os08g0517700	2	3		1.3									1.3	1.3					
Os05g0208100	2	3			1.3														
Os03g0250200	2	3										-1.8							1.2
Os03g0788800	2	3										1.1							
Os08g0503700	2	3											2.0						
Os09g0572600	2	3											1.3				-2.2		-1.1
Os01g0768200	2	3											1.1						
Os11g0677400	2	3												-2.1					
Os03g0138600	2	3															-3.1		-1.3
Os03g0300500	2	3															-3.1		
Os02g0686800	2	3																	-3.7
Os09g0454500	2	3																	-2.1
Os10g0186200	2	3																	1.1
Os07g0240200	2	3																	2.9
Os06g0130400	5	3		-1.2															

Os06g0110000	5	3			-2.6	1.3													-1.5
Os05g0576700	5	3			-0.9														
Os11g0475600	5	3			-2.5														-1.2
Os01g0800500	5	3											2.4						-1.1
Os02g0127400	5	3																-1.1	
Os08g0560700	5	3																	-3.8
Os10g0369900	5	3																	-3.2
Os01g0224700	5	3																	-2.9
Os10g0548600	5	3																	-2.7
Os05g0312500	5	3																	-2.0
Os04g0673800	5	3																	-1.9
Os01g0919200	5	3																	-1.8
Os07g0272300	5	3																	-1.8
Os04g0555700	5	3																	-1.6
Os08g0425100	5	3																	-1.4
Os02g0126300	5	3																	-1.3
Os04g0447400	17	3		-2.4	-4.7													-2.7	-1.9
Os08g0125300	17	3			-2.6														
Os05g0319800	17	3			-4.2														-4.6
Os04g0470300	17	3										-1.3							

Os03g0566500	17	3															-1.8		-1.3
Os06g0684500	17	3															-2.4		
Os03g0629800	17	3																3.3	
Os10g0351700	17	3																	-3.9
Os12g0420100	17	3																	-2.8
Os12g0482600	17	3																	-2.5
Os01g0780600	17	3																	-2.2
Os10g0534700	17	3																	-1.3
Os07g0124100	37	3	-1.3			1.0								-1.3				-1.9	
Os02g0714300	37	3										1.5	1.3					-1.9	
Os12g0636600	37	3												-3.2					
Os03g0168000	37	3												-1.2				-2.1	
Os03g0172200	37	3																-5.6	-2.2
Os10g0458300	37	3																-2.9	-1.1
Os01g0975900	37	3																	-2.1
Os01g0939100	37	3																	-1.2
Os04g0458600	48	3			-3.2														-1.6
Os08g0521600	48	3										3.2	1.3						-1.4
Os02g0725500	48	3																	-6.1
Os07g0123800	48	3																	-3.5

Os03g0369100	48	3																	-2.2
Os05g0491400	48	3																	-1.9
Os06g0609400	48	3																	-1.8
Os10g0140300	48	3																	-1.5
Os10g0564400	48	3																	-1.2
Os01g0319400	48	3																	-1.1
Os08g0438600	54	3			-3.7													-2.0	-1.2
Os10g0345900	54	3			-4.6									-2.1					-3.9
Os08g0510300	54	3			-4.4									-1.6					-2.1
Os02g0135000	54	3			-2.2														-3.7
Os11g0144500	54	3											-1.4						
Os01g0951800	54	3												-2.1					-2.5
Os07g0687300	54	3												-1.9					
Os12g0574000	54	3												-1.4					-1.9
Os05g0178100	54	3																	-3.3
Os12g0527500	54	3																	-1.3
Os05g0169300	113	3			-6.5														-3.2
Os06g0553800	113	3											1.3						-5.3
Os12g0244400	113	3																	-2.3
Os03g0155300	113	3																	-2.2

Os07g0297100	113	3																	-2.1
Os07g0659300	113	3																	-1.5
Os08g0387400	190	3	1.3																-3.1
Os07g0150700	190	3	1.3											-2.2				-1.5	-1.1
Os08g0269700	190	3			-1.3							-1.2	-1.7				2.0		
Os02g0513100	190	3			-4.7	2.1			1.7				-2.6	-5.8				-3.5	
Os02g0174800	190	3										1.1					2.5		-1.9
Os02g0720900	190	3												1.5					
Os05g0176300	205	3			-3.2														
Os01g0226600	205	3			-1.3														-3.5
Os03g0720800	205	3			-2.6														
Os01g0672100	205	3											1.4						-1.0
Os07g0598400	259	3			-4.7														-3.4
Os01g0567500	259	3											-1.7	-3.0			-1.9		
Os10g0360100	259	3															-2.5		-1.5
Os08g0451800	259	3																	-1.4
Os06g0730000	274	3			-3.7									-1.4					-2.4
Os11g0696400	274	3												-2.7					-2.3
Os04g0273100	274	3																	-1.2
Os01g0154600	274	3																	-1.1

Os01g0976200	306	3			-3.0														-3.3
Os08g0195800	306	3			-2.9													-2.7	-2.5
Os03g0186600	306	3																-2.5	-1.6
Os11g0156200	306	3																	-2.2
Os01g0216300	306	3																	-1.1
Os03g0749100	379	3			-4.2									-4.3					-2.1
Os01g0533400	379	3			-5.6								-3.1	-5.3					-1.7
Os04g0662600	379	3			-5.9									-2.6					-1.1
Os05g0242600	379	3											-3.4	-4.5			-3.3		-1.9

31 putative downstream genes of *TIP2*				
RAP_ID	Gene name	Annotation	Cluster in RiceAntherNet	Ortholog in Arabidopsis
Os01g0219500	*LTPL18*	LTP family protein precursor		AT4G33355
Os08g0208200		expressed protein		AT4G30030
Os01g0860500	*CHIA*	glycosyl hydrolase, putative		AT5G24090
Os09g0400300	*CAD9*	dehydrogenase, putative		AT4G39330
Os08g0453700		ferric reductase, putative		AT1G19230
Os01g0891000	*HEXO3*	beta-hexosaminidase precursor, putative		AT1G65590
Os08g0207800		expressed protein		AT4G30030
Os04g0677600		CRP2 - Cysteine-rich family protein precursor		
Os02g0611800	*HCT*	transferase family protein, putative	184	AT5G48930
Os01g0363300	*UPF0497*	membrane associated DUF588 domain containing protein, putative	347	AT4G15630
Os04g0652700	*ENDO 2*	nuclease PA3, putative		AT1G68290
Os05g0209600		GDSL-like lipase/acylhydrolase, putative	629	AT5G45910
Os03g0709300	*ARPN*	plastocyanin-like domain containing protein, putative		AT2G02850
Os03g0597600		transposon protein, putative, unclassified		AT5G08100
Os03g0180900	*JAZ1*	ZIM domain containing protein, putative		AT1G19180
Os02g0101900		POEI25 - Pollen Ole e I allergen and extensin family protein precursor	281	AT2G16630
Os04g0404400		desiccation-related protein PCC13-62 precursor, putative	281	AT1G47980

Os08g0131100	*CYP703A2*	cytochrome P450, putative	121	AT1G01280
Os03g0687700		expressed protein	281	
Os08g0546300	*LTPL44*	LTP family protein precursor		AT5G62080
Os03g0168600	*CYP704B1*	cytochrome P450, putative	428	AT1G69500
Os09g0525500	*LTPL45*	LTP family protein precursor	453	AT5G52160
Os01g0691300	*LTPL150*	LTP family protein precursor	453	AT3G18280
Os08g0545800	*LTPL97*	LTP family protein precursor	330	
Os05g0498200		glycine-rich cell wall protein, putative	330	
Os03g0357500	*LTPL1*	LTP family protein precursor	453	AT3G63095
Os01g0579000		expressed protein	330	
Os04g0529700		glycosyltransferase sugar-binding region containing DXD motif, putative	330	AT3G09020
Os03g0245300	*POEI16*	POEI16 - Pollen Ole e I allergen and extensin family protein precursor	403	AT4G38770
Os11g0703000		thaumatin family domain containing protein	403	AT2G28790
Os04g0528200		CRP1 - Cysteine-rich family protein precursor	403	

APPENDIX IV. PROTEIN SEQUENCES

Protein sequences used for generating phylogenetic trees		
TKPR genes		
Gene	Species	Protein sequence
At1g68540 /TKPR2	Arabidopsis	MSEYLVTGGTGFIASYIIKSLLELGHTVRTTVRNPRDEEKVGFLWEFQGAKQRLKILQADLTVEGSFDEAV NGVDGVFHTASPVLVPQDHNIQETLVDPIIKGTTNVMSSCAKSKATLKRIVLTSSCSSIRYRFDATEASPLN ESHWSDPEYCKRFNLWYGYAKTLGEREAWRIAEEKGLDLVVVNPSFVVGPLLGPKPTSTLLMILAIAKGL AGEYPNFTVGFVHIDDVVAAHVLAMEEPKASGRIICSSSVAHWSEIIELMRNKYPNYPFENKCSNKEGDN SPHSMDTRKIHELGFGSFKSLPEMFDDCIISFQKKGLL
At4g35420 /TKPR1	Arabidopsis	MDQAKGKVCVTGASGFLASWLVKRLLLEGYEVIGTVRDPGNEKKLAHLWKLEGAKERLRLVKADLME EGSFDNAIMGCQGVFHTASPVLKPTSNPEEEILRPAIEGTLNVLRSCRKNPSLKRVVLTSSSSTVRIRDDFD PKIPLDESIWTSVELCKRFQVWYALSKTLAEQAAWKFSEENGIDLVTVLPSFLVGPSLPPDLCSTASDVLG LLKGETEKFQWHGQMGYVHIDDVARTHIVVFEHEAAQGRYICSSNVISLEELVSFLSARYPSLPIPKRFEK LNRLHYDFDTSKIQSLGLKFKSLEEMFDDCIASLVEQGYLSTVLP
Os09g0493500 /OsDFR2	Rice	MLSRILHGYGGHGGRGFEQTYRCYSAAAFNKPQLEGGDKVIMPASALHRLASLHIDYPMLFELSHHGDA AAHRVTHCGVLEFVADEGTVIMPRWMMRGMRLDDGGLVVVRSASLPKGSYAKLQPHTGDFLDTANPK AVLEKTLRSFTCLTTGDTIMVAYNNKEFLIDIVETKPASAVCIIETDCEVDFAPPLDYKEPEKVQQKPSVPS SKAASEDQDQIKDEPEFRAFTGSGNRLDGKASKPLAAGISSNPAAASSAISDSNKKVNQETAASGVSNSTR QKKGKLVFGSNKSSSSSKEPEKAPPVKVDELAKKEEPKFQAFSGTSYSLKRNRDKVSHLWRLPSAKERLQ

		LEEMLVPAINGTLNVLKSCKKNPFLKRVVLTSSSSTVRIRDESKHPEISLDETIWSSVALCEKLQLWYALA KISAEKAAWEFAKENNIDLVTVLPSFVIGPSLSHELSVTASDILGLLQGDTDRFISYGRMGYVHIDDVASC HILVYEAPQATGRYLCNSVVLDNNELVALLAKQFPIFPIPRSLRNPYEK
Os01g0127500	Rice	MPEYCVTGGTGFIASHLIRALLAASHTVRATVRDPEDEAKVGFLWELDGASERLQLVKADLMVEGSFDD AVRGVDGVFHAASPVVVVGNSSSNNGKPNDDDDEEEVQQRLVEPIVRGASNVLRSCARASPRPRRVVFT SSCSCVRYGAGAAAALNESHWSDAAYCAAHGLWYAYAKTLAEREAWRLAKERGLDMVAVNPSFVVG PILSQAPTSTALIVLALLRGELPRYPNTTVGFVHVDDAVLAHVVAMEDARASGRLICSCHVAHWSEIVGS LRERYPGYPIPAECGSHKGDDRAHKMDTAKIRALGFPPFLSVQQMFDDCIKSFQDKGLLPPHA
At5G42800 /DFR	Arabidopsis	MVSQKETVCVTGASGFIGSWLVMRLLERGYFVRATVRDPGNLKKVQHLLDLPNAKTLLTLWKADLSEE GSYDDAINGCDGVFHVATPMDFESKDPENEVIKPTVNGMLGIMKACVKAKTVRRFVFTSSAGTVNVEEH QKNVYDENDWSDLEFIMSKKMTGWMYFVSKTLAEKAAWDFAEEKGLDFISIIPTLVVGPFITTSMPPSLIT ALSPITRNEAHYSIIRQGQYVHLDDLCNAHIFLYEQAAAKGRYICSSHDATILTISKFLRPKYPEYNVPSTFE GVDENLKSIEFSSKKLTDMGFNFKYSLEEMFIESIETCRQKGFLPVSLSYQSISEIKTKNENIDVKTGDGLTD GMKPCNKTETGITGERTDAPMLAQQMCA
Os08g0515900	Rice	MENTTKGKVCVTGASGYVASWLVKRLLESGYHVLGTVRDPGNHKKVGHLWNLTGAKERLELVRADLL EEGSFDDAVMACEGVFHTASPVITETDSSKAAVLDSAINGTLNVLRSCKKNPSLKRVVLTSSSSTVRLKDE ADLPPNVLLDETSWSSMEFCESLQIWYAIAKTLAEKAAWEFAKENGIDLVAVLPTFVVGPNLSHELSPTT TDVLGLFQGETTKFTMYGRMGYVHIDDVASCHILLYETPRAAGRYICNSAVLDVNELVTLLARRFPSYPI PKRSRTPAKWNFFFSAD*

LRR-RLK genes		
Gene	Species	Protein sequence
AT2G26330 /ER	Arabidopsis	MALFRDIVLLGFLFCLSLVATVTSEEGATLLEIKKSFKDVNNVLYDWTTSPSSDYCVWRGVSCENVTFNV VALNLSDLNLDGEISPAIGDLKSLLSIDLRGNRLSGQIPDEIGDCSSLQNLDLSFNELSGDIPFSISKLKQLEQ LILKNNQLIGPIPSTLSQIPNLKILDLAQNKLSGEIPRLIYWNEVLQYLGLRGNNLVGNISPDLCQLTGLWYF DVRNNSLTGSIPETIGNCTAFQVLDLSYNQLTGEIPFDIGFLQVATLSLQGNQLSGKIPSVIGLMQALAVLD LSGNLLSGSIPPILGNLTFTEKLYLHSNKLTGSIPPELGNMSKLHYLELNDNHLTGHIPPELGKLTDLFDLN VANNDLEGPIPDHLSSCTNLNSLNVHGNKFSGTIPRAFQKLESMTYLNLSSNNIKGPIPVELSRIGNLDTLD LSNNKINGIIPSSLGDLEHLLKMNLSRNHITGVVPGDFGNLRSIMEIDLSNNDISGPIPEELNQLQNIILLRLE NNNLTGNVGSLANCLSLTVLNVSHNNLVGDIPKNNNFSRFSPDSFIGNPGLCGSWLNSPCHDSRRTVRVSI SRAAILGIAIGGLVILLMVLIAACRPHNPPPFLDGSLDKPVTYSTPKLVILHMNMALHVYEDIMRMTENLS EKYIIGHGASSTVYKCVLKNCKPVAIKRLYSHNPQSMKQFETELEMLSSIKHRNLVSLQAYSLSHLGSLLF YDYLENGSLWDLLHGPTKKKTLDWDTRLKIAYGAAQGLAYLHHDCSPRIIHRDVKSSNILLDKDLEARL TDFGIAKSLCVSKSHTSTYVMGTIGYIDPEYARTSRLTEKSDVYSYGIVLLELLTRRKAVDDESNLHHLIM SKTGNNEVMEMADPDITSTCKDLGVVKKVFQLALLCTKRQPNDRPTMHQVTRVLGSFMLSEQPPAATD TSATLAGSCYVDEYANLKTPHSVNCSSMSASDAQLFLRFGQVISQNSE
Os06g0203800	Rice	MTPAPAAASYRALVALLLVAVAVADDGSTLLEIKKSFRNVDNVLYDWAGGDYCSWRGVLCDNVTFAV AALNLSGLNLGGEISPAVGRLKGIVSIDLKSNGLSGQIPDEIGDCSSLKTLDLSFNSLDGDIPFSVSKLKHIES LILKNNQLIGVIPSTLSQLPNLKILDLAQNKLSGEIPRLIYWNEVLQYLGLRGNNLEGSISPDICQLTGLWYF DVKNNSLTGPIPETIGNCTSFQVLDLSYNKLSGSIPFNIGFLQVATLSLQGNMFTGPIPSVIGLMQALAVLD LSYNQLSGPIPSILGNLTYTEKLYMQGNKLTGPIPPELGNMSTLHYLELNDNQLSGFIPPEFGKLTGLFDLN

		LANNNFEGPIPDNISSCVNLNSFNAYGNRLNGTIPPSLHKLESMTYLNLSSNFLSGSIPIELSRINNLDTLDLS CNMITGPIPSTIGSLEHLLRLNLSNNGLVGFIPAEIGNLRSIMEIDMSNNHLGGLIPQELGMLQNLMLLNLK NNNITGDVSSLMNCFSLNILNVSYNNLAGVVPTDNNFSRFSPDSFLGNPGLCGYWLGSSCRSSGHQQKPLI SKAAILGIAVGGLVILLMILVAVCRPHSPPVFKDVSVSKPVSNVPPKLVILHMNLSLLVYEDIMTMTENLS EKYIIGYGASSTVYKCVSKNRKPVAVKKLYAHYPQSFKEFETELETVGSIKHRNLVSLQGYSLSPVGNLLF YDYMENGSLWDVLHEGPTKKKKLDWETRLRIALGAAQGLAYLHHDCSPRIIHRDVKSKNILLDKDYEA HLTDFGIAKSLCVSKTHTSTYVMGTIGYIDPEYARTSRLNEKSDVYSYGIVLLELLTGKKPVDNECNLHHL ILSKTANNAVMETVDPDIADTCKDLGEVKKVFQLALLCTKRQPSDRPTMHEVVRVLDCLVRPDPPPKSA QQLAMPQRPAVPSYINEYVSLRGTSVLSCANSSCTSDAELFLKFGEVISQNTE
Os02g0777400	Rice	MILAAACWPHWAQVPKDVSLCKPDIHALPSSNVPPKLVILHMNMAFLVYEDIMRMTENLSEKYIIGYGA SSTVYKCVLKNCKPVAIKKLYAHYPQSLKEFETELETVGSIKHRNLVSLQGYSLSPAGNLLFYDYLENGS LWDVLHAGSSKKQKLDWEARLRIALGAAQGLAYLHHDCNPRIIHRDVKSKNILLDKDYEAHLADFGIAK SLCTSKTHTSTYVMGTIGYIDPEYACTSRLNEKSDVYSYGIVLLELLTGKKPVDNECNLHHLILSKAADNT VMEMVDPDIADTCKDLGEVKKVFQLALLCSKRQPSDRPTMHEVVRVLDCLVYPDPPSKPALPPALPQSS TVPSYVNEYVSLRGGSTLSCENSSSASDAELFLKFGEVISQNTE
AT5G65700 /BAM1	Arabidopsis	MKLFLLLLFLLHISHTFTASRPISEFRALLSLKTSLTGAGDDKNSPLSSWKVSTSFCTWIGVTCDVSRRHVT SLDLSGLNLSGTLSPDVSHLRLLQNLSLAENLISGPIPPEISSLSGLRHLNLSNNVFNGSFPDEISSGLVNLRV LDVYNNNLTGDLPVSVTNLTQLRHLHLGGNYFAGKIPPSYGSWPVIEYLAVSGNELVGKIPPEIGNLTTLR ELYIGYYNAFEDGLPPEIGNLSELVRFDGANCGLTGEIPPEIGKLQKLDTLFLQVNVFSGPLTWELGTLSSL KSMDLSNNMFTGEIPASFAELKNLTLLNLFRNKLHGEIPEFIGDLPELEVLQLWENNFTGSIPQKLGENGK LNLVDLSSNKLTGTLPPNMCSGNKLETLITLGNFLFGSIPDSLGKCESLTRIRMGENFLNGSIPKGLFGLPK

		LTQVELQDNYLSGELPVAGGVSVNLGQISLSNNQLSGPLPPAIGNFTGVQKLLLDGNKFQGPIPSEVGKLQ QLSKIDFSHNLFSGRIAPEISRCKLLTFVDLSRNELSGEIPNEITAMKILNYLNLSRNHLVGSIPGSISSMQSL TSLDFSYNNLSGLVPGTGQFSYFNYTSFLGNPDLCGPYLGPCKDGVAKGGHQSHSKGPLSASMKLLLVL GLLVCSIAFAVVAIIKARSLKKASESRAWRLTAFQRLDFTCDDVLDSLKEDNIIGKGGAGIVYKGVMPNG DLVAVKRLAAMSRGSSHDHGFNAEIQTLGRIRHRHIVRLLGFCSNHETNLLVYEYMPNGSLGEVLHGKK GGHLHWDTRYKIALEAAKGLCYLHHDCSPLIVHRDVKSNNILLDSNFEAHVADFGLAKFLQDSGTSECM SAIAGSYGYIAPEYAYTLKVDEKSDVYSFGVVLLELVTGRKPVGEFGDGVDIVQWVRKMTDSNKDSVLK VLDPRLSSIPIHEVTHVFYVAMLCVEEQAVERPTMREVVQILTEIPKLPPSKDQPMTESAPESELSPKSGVQ SPPDLLNL
Os03g0773700	Rice	MHLRLLLLLVGVAAAAADADADADADALLAAKAAMSDPTGALASWGGNGTRTNTTAAAAAHCAWAGV TCSSRGAVVGLDVSGLNLSGALPAELTGLRGLMRLSVGANAFSGPIPASLGRLQFLTYLNLSNNAFNGSF PAALARLRGLRVLDLYNNNLTSPLPMEVVQMPLLRHLHLGGNFFSGEIPPEYGRWGRMQYLAVSGNELS GKIPPELGNLTSLRELYIGYYNSYSGGLPPELGNLTELVRLDAANCGLSGEIPPELGKLQNLDTLFLQVNSL AGGIPSELGYLKSLSSLDLSNNVLTGEIPASFSELKNLTLLNLFRNKLRGDIPDFVGDLPSLEVLQLWENNF TGGVPRRLGRNGRLQLLDLSSNRLTGTLPPELCAGGKMHTLIALGNFLFGAIPDSLGECKSLSRVRLGENY LNGSIPKGLFELPKLTQVELQDNLLTGNFPAVSGAAAPNLGEISLSNNQLTGALPASIGNFSGVQKLLLDR NSFSGVVPPEIGRLQKLSKADLSSNALEGGVPPEIGKCRLLTYLDLSRNNISGKIPPAISGMRILNYLNLSRN HLDGEIPPSIATMQSLTAVDFSYNNLSGLVPGTGQFSYFNATSFVGNPGLCGPYLGPCRPGVAGTDHGGH GHGGLSNGVKLLIVLGLLACSIAFAVGAILKARSLKKASEARVWKLTAFQRLDFTCDDVLDCLKEENVIG KGGAGIVYKGAMPNGDHVAVKRLPAMGRGSSHDHGFSAEIQTLGRIRHRHIVRLLGFCSNNETNLLVYE YMPNGSLGELLHGKKGGHLHWDTRYKIAIEAAKGLCYLHHDCSPLILHRDVKSNNILLDSDFEAHVADF

		GLAKFLQDTGASECMSAIAGSYGYIAPEYAYTLKVDEKSDVYSFGVVLLELVTGRKPVGEFGDGVDIVQ WVRMMTDSNKEQVMKVLDPRLSTVPLHEVMHVFYVALLCIEEQSVQRPTMREVVQILSELPKLAPRQG EVLSHAVDGFASNPPAPVPSGSAEALTGDAKDQQQQQTNSESTTPPDLISI
Os07g0134200	Rice	MRLHYHHLAVVLLAAVAAAATAAAGGEADALLAVKAALDDPTGALASWTTNTTSSPCAWSGVACNA RGAVVGLDVSGRNLTGGLPGAALSGLQHLARLDLAANALSGPIPAALSRLAPFLTHLNLSNNGLNGTFPP QLSRLRALRVLDLYNNNLTGALPLEVVSMAQLRHLHLGGNFFSGGIPPEYGRWGRLQYLAVSGNELSGK IPPELGNLTSLRELYIGYFNSYSGGIPPELGNMTDLVRLDAANCGLSGEIPPELGNLANLDTLFLQVNGLAG GIPRELGKLASLSSLDLSNNALAGEIPATFADLKNLTLLNLFRNKLRGDIPEFVGDLPSLEVLQLWENNFTG GIPRRLGRNGRFQLLDLSSNRLTGTLPPDLCAGGKLETLIALGNSLFGAIPASLGKCTSLTRVRLGDNYLN GSIPEGLFELPNLTQVELQDNLISGGFPAVSGTGAPNLGQISLSNNQLTGALPAFIGSFSGVQKLLLDQNAF TGEIPPEIGRLQQLSKADLSGNSFDGGVPPEIGKCRLLTYLDLSRNNLSGEIPPAISGMRILNYLNLSRNQLD GEIPATIAAMQSLTAVDFSYNNLSGLVPATGQFSYFNATSFVGNPGLCGPYLGPCHPGAPGTDHGGRSHG GLSNSFKLLIVLGLLALSIAFAAMAILKARSLKKASEARAWKLTAFQRLEFTCDDVLDSLKEENIIGKGGA GTVYKGTMPDGEHVAVKRLPAMSRGSSHDHGFSAEIQTLGRIRHRYIVRLLGFCSNNETNLLVYEYMPN GSLGELLHGKKGGHLHWDTRYKVAVEAAKGLCYLHHDCSPPILHRDVKSNNILLDSDFEAHVADFGLA KFLQDSGTSECMSAIAGSYGYIAPEYAYTLKVDEKSDVYSFGVVLLELITGKKPVGEFGDGVDIVQWVKT MTDSNKEHVIKILDPRLSTVPVHEVMHVFYVALLCVEEQSVQRPTMREVVQILSELPKPTSKQGEEPPSGE GAVFDLVVPAESAEANEAKEQQQQQLNSPSSPPPDLISI
Os01g0957100	Rice	MATPIALLLFVLVVAAAADSTMPMPVNEEVLGLVVFKSALSDPSGALATWTESDATPCGWAHVECDPA TSRVLRLALDGLGLSGRMPRGLDRLAALQSLSVARNNLSGELPPGLSLLASLRSIDLSYNAFSGPLPGDVP LLASLRYLDLTGNAFSGPLPATFPATVRFLMLSGNQFSGPLPQGLSKSSFLLHLNLSGNQLSGSPDFAGAL

		WPLSRLRALDLSRNQFSGTVTTGIANLHNLKTIDLSGNRFFGAVPSDIGLCPHLSTVDISSNAFDGQLPDSI AHLGSLVYFAASGNRFSGDVPAWLGDLAALQHLDFSDNALTGRLPDSLGKLKDLRYLSMSENQLSGAIP DAMSGCTKLAELHLRANNLSGSIPDALFDVGLETLDMSSNALSGVLPSGSTKLAETLQWLDLSVNQITGG IPAEMALFMNLRYLNLSRNDLRTQLPPELGLLRNLTVLDLRSSGLYGTMPSDLCEAGSLAVLQLDGNSLA GPIPDNIGNCSSLYLLSLGHNSLTGPIPVGMSELKKLEILRLEYNNLSGEIPQQLGGIESLLAVNVSHNRLVG RLPASGVFQSLDASALEGNLGICSPLVTQPCRMNVAKPLVLDPNEYPHGGDGDNNLETSGRGPASPRKRR FLSVSAMVAICAAVFIILGVIVITLLNMSARRRAGDGGTTTPEKELESIVSSSTKSSKLATGKMVTFGPGNS LRSEDFVGGADALLSKATEIGRGVFGTVYRASVGEGRVVAIKKLATASIVESRDDFDREVRILGKARHPN LLPLKGYYWTPQLQLLITDYAPHGSLEARLHGNGDGAFPPLTWAERFRIVAGTARGLAHLHQSFRPPMIH YNVKPSNILLDEQCNPMVGDFGLARLLPKLDKHVMSSRFQGGMGYVAPELACQSLRINEKCDIYGFGVLI LELVTGRRAVEYGDDDVVILIDQVRVLLDHGGGSNVLECVDPSIGEFPEEEVLPVLKLGMVCTSQIPSNRP SMAEVVQILQVIKAPVAASSARIEAF
Os08g0506400	Rice	MAAAARAAAVAVVVVVVVLGVEVAAALNTDGLALLALKFAVSEDPNGALSTWRDADNDPCGWSGVT CVDGGGGRVAGVELANFSLAGYLPSELSLLSELVTLSLPYNQLAGQIPVAITALQKLAALDLAHNLLSGQ VPAGIGRLVSLSRLDLSSNQLNGSLPPAIAGLPRLSGVLNLSYNHFTGGIPPEFGGIPVAVSLDLRGNDLAG EIPQVGSLVNQGPTAFDDNPRLCGFPLKVECAGEKEDPRIPEANGGMNPGAAAAVGRPPRRRSSPTVPVL AAIVVVAIVAGVILQWQCRRRCAAATARDEEKESAKDKSGAVTLAGSEERRSGGEEGEVFVAVDDGFG MELEELLRASAFVVGKSRGGIVYRVVPGHGPAVAVRRLSEPDDGDGGSDSGWRRRRAFETEAAAIGRAR HPNVARLRAYYYAPDEKLLIYDYLSNGSLHSALHGGPTASPTPLPWSMRLSIVQGAARGLAYLHECSPRR YVHGCIKSSKILLDDELRAHVSGFGLARLVAGGAHKAAAAQSKKLGGAACALRGGGGALAYVAPELRT PGGAAAAATQKGDVFALGVVLLEAVTGREPTEGEGGLELEAWVRRAFKEERPLSEVVDPTLLGEVHAK

		KQVLAVFHVALGCTEPDAELRPRMRAVAESLDRINA
Os03g0332900	Rice	MGYKIFEGSLVLLSLLLFTSLAVADIASEKQALLAFASAVYRGNKLNWDVNISLCSWHGVTCSPDRSRIS ALRVPAAGLIGAIPPNTLGRLVSLQVLSLRSNRLIGSIPSDITSLPSLQSIFLQDNELSGDLPSFFSPTLNTIDL SYNSFAGQIPASLQNLTQLSTLNLSKNSLSGPIPDLKLPSLRQLNLSNNELNGSIPPFLQIFSNSSFLGNPGLC GPPLAECSLPSPTSSPESSLPPPSALPHRGKKVGTGSIIAAAVGGFAVFLLAAAIFVVCFSKRKEKKDDGLD NNGKGTDNARIEKRKEQVSSGVQMAEKNKLVFLDGCSYNFDLEDLLRASAEVLGKGSYGTAYKAILED GTIVVVKRLKDVVAGKKEFEQQMEQIGRVGKHANLVPLRAYYYSKDEKLVVYEYVATGSFSAMLHGIK GIAEKTPLDWNTRMKIILGTARGIAHIHAEGGSKLAHGNIKATNVLLDQDHNPYVSDYGLSALMSFPISTS RVVVGYRAPETFESRKFTHKSDVYSFGVLLMEMLTGKAPLQSQGQDDVVDLPRWVHSVVREEWTAEVF DVELMKYLNIEDELVQMLQLAMACTSRSPERRPTMAEVIRMIEELRQSASESRDSSNENARESNPPSA
Os03g0183800	Rice	MGRADAPWLPLLLLCSSCCFCIWPQKQILVAADTDPNDVTVLNTLFTSLNSPGQLKGWQASGGDPCGQS WQGITCSGSSVTAIKLPSLGLSGNLAYNMNTMGSLIEIDMSQNNLGGGQQIQYNLPTNKLERLNLAGNQF TGNLPYSIFSMSNLKYLNLNHNQLQGNITDVFSSLYSLTTLDLSFNSLAGDLPQGFTSLSSLKKLYLQNNQ FTGYINVLANLPLDDLNVANNHFTGWIPSQLKKINNLQTDGNSWSNGPAPPPPPYSAPPPPNRPNSPGQNN GGSSSGGSSGIGGGGVAGIIISLLVVGAVVAFFVIRRRKRRAALEEHFEQHQPFTSFPSNEVKDMKPIEEST TIDVESLPSPASFSLKPPPKIERHKSFDDDDLSNKPVLKKTNVAPIKATVYSVADLQMATESFSMDNLVGE GTFGRVYRAQFTGGKVLAVKKLDSTVMPFHSSDDFAELVSDISKLHHPNLNELVGYCMEHGQHLLVYD FHRNGSLHDLLHLSDEYSKPLSWNSRVKIALGSARALEYLHEICSPSIIHKNFKSSNLLLDSEFNPHLSDAG LASFISDAEFQAAQQSAGCTAPEVDMTGQYTLKSDVYSFGVVMLELLTGRRPFDSTRPRSEQSLVRWATP QLHDIDALDRMVDPALKGLYPAKSLSRFADVLALCVQPEPEFRPPMSEVVQALVRLVQRANMTKRMLD GDTSSRRTDDQEQDFI

Appendix V. GENOTYPES OF MUTANTS

Gene	RAP ID	Genotype	Mutant	Background
CML41	Os03g0769500	Mutant：insertion of A at 23bp; insertion of T at 23bp	*cml41-1*	9522
CML41	Os03g0769500	Mutant：insertion of A at 23bp; insertion of T at 23bp	*cml41-2*	9522
DRL1	Os08g0515900	mutant：deletion of G at 84bp	*drl1-1*	9522
DRL1	Os08g0515900	mutant: deletion of 84bp;deletion of 84-87bp	*drl1-10*	9522
DRL1	Os08g0515900	mutant: deletion of 84bp;deletion of 84-88bp	*drl1-11*	9522
DRL1	Os08g0515900	mutant：deletion of 83-84bp; insertion of G at 85bp	*drl1-12*	9522
DRL1	Os08g0515900	mutant: deletion of 83-84bp;deletion of G at 84bp	*drl1-13*	9522
DRL1	Os08g0515900	mutant: deletion of 85-98bp;deletion of 85-105bp	*drl1-14*	9522
DRL1	Os08g0515901	mutant: deletion of 257-260bp; insertion of A at 257bp	*drl1-15*	9522
DRL1	Os08g0515903	mutant: deletion of 257-260bp; insertion of A at 257bp	*drl1-16*	9522
DRL1	Os08g0515904	mutant: insertion of A at 257bp	*drl1-17*	9522
DRL1	Os08g0515900	mutant: deletion of 82-83bp;deletion of G at 84bp	*drl1-18*	9522
DRL1	Os08g0515907	mutant: deletion of 258bp; insertion of A at 259bp	*drl1-19*	9522
DRL1	Os08g0515906	mutant：deletion of T at 260bp	*drl1-2*	9522
DRL1	Os08g0515908	mutant: insertion of A at 261bp; insertion of T at 261bp	*drl1-20*	9522
DRL1	Os08g0515909	mutant: insertion of A at 257bp	*drl1-21*	9522

DRL1	Os08g0515911	mutant：deletion of 259-260bp;insertion of T at 260bp	*drl1-22*	9522
DRL1	Os08g0515913	mutant：deletion of 259-260bp	*drl1-23*	9522
DRL1	Os08g0515914	mutant: insertion of A at 261bp; insertion of T at 261bp	*drl1-24*	9522
DRL1	Os08g0515900	mutant: deletion of 81-83bp;insertion of G at 84bp	*drl1-3*	9522
DRL1	Os08g0515900	mutant: deletion of 83-84bp;deletion of G at 84bp	*drl1-4*	9522
DRL1	Os08g0515900	mutant：deletion of G at 84bp;insertion of A at 85bp	*drl1-5*	9522
DRL1	Os08g0515900	mutant: deletion of 85-86bp; deletion of 81-85bp	*drl1-6*	9522
DRL1	Os08g0515900	mutant：deletion of G at 84bp	*drl1-7*	9522
DRL1	Os08g0515900	mutant: deletion of 84bp;deletion of 84-93bp	*drl1-8*	9522
DRL1	Os08g0515900	mutant：deletion of G at 84bp; insertion of A at 85bp	*drl1-9*	9522
DRL1	Os08g0515900	Heterozygote		9522
DRL1	Os08g0515900	Heterozygote		9522
DRL1	Os08g0515900	Heterozygote		9522
DRL1	Os08g0515902	Heterozygote		9522
DRL1	Os08g0515905	Heterozygote		9522
DRL1	Os08g0515910	Heterozygote		9522
DRL1	Os08g0515912	Heterozygote		9522
DRL1	Os08g0515915	Heterozygote		9522
DRL2	Os01g0127500	mutant: insertion of A at 95bp; insertion of T at 95bp	*drl2-1*	9522

DRL2	Os01g0127500	mutant: insertion of A at 95bp; deletion of 95-96bp	*drl2-2*	9522
DRL2	Os01g0127500	mutant: insertion of 95-99bp; deletion of 95-98bp	*drl2-3*	9522
DRL2	Os01g0127500	Heterozygote		9522
DRL2	Os01g0127500	Heterozygote		9522
DRL2	Os01g0127500	Heterozygote		9522
DRL2	Os01g0127500	Heterozygote		9522
GAMYBL2	Os03g0578900	mutant: insertion of A at 121bp; substitution of 118-120bp with A	*gamybl2-1*	9522
GAMYBL2	Os03g0578900	mutant: insertion of G at 120bp; deletion of 121-122bp	*gamybl2-2*	9522
GAMYBL2	Os03g0578900	mutant: insertion of A at 121bp; insertion of T at 121bp	*gamybl2-3*	9522
GAMYBL2	Os03g0578900	mutant: insertion of C at 120bp; insertion of CA at 120bp	*gamybl2-4*	9522
GMCO	Os02g0678300	mutant：insertion of C at 771bp	*gmco-1osnp1-4*	*osnp1-4*
GMCO	Os02g0678300	mutant：insertion of T at 771bp	*gmco-2osnp1-4*	*osnp1-4*
GMCO	Os02g0678300	mutant：insertion of A at 771bp;insertion of T at 771bp	*gmco-3osnp1-4*	*osnp1-4*
GMCO	Os02g0678300	mutant：insertion of C at 771bp	*gmco-4osnp1-4*	*osnp1-4*
GMCO	Os02g0678300	mutant：insertion of C at 771bp	*gmco-5osnp1-4*	*osnp1-4*
GMCO	Os02g0678300	Heterozygote		*osnp1-4*
LAP6	Os10g0484800	mutant:deletion of 69bp	*lap6-1*	9522
LAP6	Os10g0484800	mutant:deletion of 68-81bp:deletion of 66-76bp	*lap6-10*	9522
LAP6	Os10g0484800	mutant:deletion of 69bp	*lap6-2*	9522

LAP6	Os10g0484800	mutant:deletion of 68-81bp:deletion of 66-76bp	*lap6-3*	9522
LAP6	Os10g0484800	mutant:deletion of 69bp	*lap6-3*	9522
LAP6	Os10g0484800	mutant:deletion of 68-81bp:deletion of 66-76bp	*lap6-4*	9522
LAP6	Os10g0484800	mutant:deletion of 68-81bp:deletion of 66-76bp	*lap6-5*	9522
LAP6	Os10g0484800	mutant:deletion of 68-81bp:deletion of 66-76bp	*lap6-6*	9522
LAP6	Os10g0484800	mutant:deletion of 68-81bp:deletion of 66-76bp	*lap6-7*	9522
LAP6	Os10g0484800	mutant:deletion of 68-81bp:deletion of 66-76bp	*lap6-8*	9522
LAP6	Os10g0484800	mutant:deletion of 68-81bp:deletion of 66-76bp	*lap6-9*	9522
LRR1	Os01g0170300	mutant: deletion of 3430-3433bp; insertion of C at 3430bp	*lrr1-1*	9522
LRR1	Os01g0170300	mutant: deletion of 3430bp; insertion of C at 3431bp	*lrr1-2*	9522
LRR1	Os01g0170300	mutant: deletion of 3432-3433bp; insertion of T at 3433bp	*lrr1-3*	9522
LRR1	Os01g0170300	mutant: deletion of 3431-3432bp; insertion of T at 3430bp	*lrr1-4*	9522
LRR1	Os01g0170300	Heterozygote		9522
LRR1	Os01g0170300	Heterozygote		9522
LRR1	Os01g0170300	Heterozygote		9522
LRR1	Os01g0170300	Heterozygote		9522
LRR1	Os01g0170300	Heterozygote		9522
LRR1	Os01g0170300	Heterozygote		9522
LRR1	Os01g0170300	Heterozygote		9522

LRR5	Os05g0169600	mutant: deletion of 568bp	*lrr5-1*	9522
LRR5	Os05g0169600	mutant: deletion of 568bp	*lrr5-2*	9522
LRR5	Os05g0169600	Heterozygote		9522