the plant journal

SEB Society for Experimental Biology

The Plant Journal (2017) **92,** 1076–1091 doi: 10.1111/tpj.13744

RiceAntherNet: a gene co-expression network for identifying anther and pollen development genes

Hong Lin[1,2], Jing Yu[1], Simon P. Pearce[3,4], Dabing Zhang[1,*] and Zoe A. Wilson[2,5,*]

[1]*School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China,*
[2]*Division of Plant & Crop Sciences, School of Biosciences, University of Nottingham, Loughborough, UK,*
[3]*School of Mathematics, University of Manchester, Manchester, UK,*
[4]*Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK, and*
[5]*Centre for Plant Integrative Biology, School of Biosciences, University of Nottingham, Nottingham, UK*

Received 18 August 2017; revised 29 September 2017; accepted 2 October 2017; published online 14 October 2017.
*For correspondence (e-mails zoe.wilson@nottingham.ac.uk; zhangdb@sjtu.edu.cn).
H.L. and J.Y. are joint first authors.

SUMMARY

In plants, normal anther and pollen development involves many important biological events and complex molecular regulatory coordination. Understanding gene regulatory relationships during male reproductive development is essential for fundamental biology and crop breeding. In this work, we developed a rice gene co-expression network for anther development (RiceAntherNet) that allows prediction of gene regulatory relationships during pollen development. RiceAntherNet was generated from 57 rice anther tissue microarrays across all developmental stages. The microarray datasets from nine rice male sterile mutants, including *msp1-4*, *ostdl1a*, *gamyb-2*, *tip2*, *udt1-1*, *tdr*, *eat1-1*, *ptc1* and *mads3-4*, were used to explore and test the network. Among the changed genes, three clades showing differential expression patterns were constructed to identify genes associated with pollen formation. Many of these have known roles in pollen development, for example, seven genes in Clade 1 (*OsABCG15*, *OsLAP5*, *OsLAP6*, *DPW*, *CYP703A3*, *OsNP1* and *OsCP1*) are involved in rice pollen wall formation. Furthermore, Clade 1 contained 12 genes whose predicted orthologs in Arabidopsis have been reported as key during pollen development and may play similar roles in rice. Genes in Clade 2 are expressed earlier than Clade 1 (anther stages 2–9), while genes in Clade 3 are expressed later (stages 10–12). RiceAntherNet serves as a valuable tool for identifying novel genes during plant anther and pollen development. A website is provided (https://www.cpib.ac.uk/anther/riceindex.html) to present the expression profiles for gene characterization. This will assist in determining the key relationships between genes, thus enabling characterization of critical genes associated with anther and pollen regulatory networks.

Keywords: anther, pollen, rice, correlation network, expression network, reproduction.

INTRODUCTION

Anther development and pollen formation are of vital importance throughout the life cycle of flowering plants. Development starts from a single layer of cells, which undergo a series of cell divisions and differentiation to form the innermost meiocytes 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 and Yang, 2014). After initiation of anthers, meiocytes undergo meiosis to form tetrads surrounded by callose; meanwhile, the innermost sporophytic cell layer, the tapetum, secretes enzymes for callose degradation, pollen wall materials and initiates programmed cell death (PCD). After the haploid microspores separate and are released into the locule, the tapetum synthesizes and transports sporopollenin onto the surface of microspores before degenerating. Anther development can be divided into 14 stages through morphology analysis (Figure S1; Wilson and Zhang, 2009). During the cell division and differentiation stages (stages 1–7), anther primordium differentiates into three meristematic cell layers (L1, L2 and L3). L1 forms the epidermis, L2 generates archesporial cells that form microspore mother cells and primary parietal cells, which further differentiate into the three inner anther wall layers (Ma, 2005; Zhang and Yang, 2014).

A high level of conservation of the regulatory gene networks involved in pollen development has been seen between rice and Arabidopsis. In rice, *MSP1* (*MULTIPLE SPOROCYTE 1*; Nonomura *et al.*, 2003; Zhao *et al.*, 2008) and *OsTDL1A* (*TPD1-like 1A*; Zhao *et al.*, 2008; Yang *et al.*, 2016), and the Arabidopsis orthologs: *EMS1* (*EXCESS MICROSPOROCYTES1*; Jia and Zhao, 2008), *TPD1* (*TAPETUM DETERMINANT1*; Huang *et al.*, 2016) and *TIP2/bHLH* (*TDR INTERACTING PROTEIN2*; Fu *et al.*, 2014) are required for anther cell differentiation. *UDT1* (*UNDEVELOPED TAPETUM1*; Jung *et al.*, 2005), *TDR* (*TAPETUM DEGENERATION RETARDATION*; Li *et al.*, 2006), *GAMYB* (Aya *et al.*, 2009), *PTC1* (*PERSISTENT TAPETAL CELL1*; Li *et al.*, 2011), *EAT1/DTD* (*ETERNAL TAPETUM1*; Niu *et al.*, 2013), *OsDEX1* (*DEFECTIVE IN EXINE PATTERN FORMATION*; Yu *et al.*, 2016) and *OsMADS3* (Hu *et al.*, 2011) are required for tapetum function and pollen wall formation in rice; orthologous genes are found in Arabidopsis [*DYT1* (*DYSFUNCTIONAL TAPETUM1*; Feng *et al.*, 2012), *AMS* (*ABORTED MICROSPORES*; Xu *et al.*, 2014), *MYB33* and *MYB65* (Millar and Gubler, 2005), *MS1* (*MALE STERILITY1*; Wilson *et al.*, 2001) and *DEX1* (Paxson-Sowders *et al.*, 2001), respectively].

During the pollen wall formation stages, sporopollenin is synthesized in tapetal cells with associated high levels of gene expression observed. Lipid synthesis genes *CYP704B2*, *CYP703A3* (Yang *et al.*, 2014) and *DPW* (*DEFECTIVE POLLEN WALL*; Xu *et al.*, 2016) in rice, and the corresponding orthologs in Arabidopsis [*CYP704B1* (Dobritsa *et al.*, 2009a,b), *CYP703A2* (Morant *et al.*, 2007) and *MS2* (*MALE STERILE2*; Chen *et al.*, 2011)] are required for the pollen wall formation. In addition, a 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 (Qin *et al.*, 2013; Jepson *et al.*, 2014; Li *et al.*, 2016; Yang *et al.*, 2017). Sporopollenin transport-related genes are also required; *OsABCG15* (*ATP Binding Cassette G Transporters*; Qin *et al.*, 2013) and *OsABCG26* (Zhao *et al.*, 2015), whose orthologs are *ABCG11* (Panikashvili *et al.*, 2010) and *WBC27/ABCG26* (Choi *et al.*, 2011) in Arabidopsis, as well as *OsC4* and *OsC6* (Li and Zhang, 2010; Zhang *et al.*, 2010) in rice are also required for pollen wall development.

Although many genes required for viable pollen formation have been identified, the process of pollen development remains elusive. It has been reported using a 10-k rice cDNA microarray that over 2000 genes (approximately 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, 2005). 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 and Mehridehnavi, 2012); 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 genes of unknown function (Usadel *et al.*, 2009).

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 hypothetic 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 normalizes 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 that can have very

different expression patterns (Wittkop *et al.*, 2010). Here we use co-expression analysis, followed by tight clustering, to generate an untargeted genome-wide gene co-expression network focusing on rice anther development. Genes with similar expression patterns and correlated functions, especially during anther and pollen development, were connected, and putative key genes involved in pollen wall formation were revealed by this approach. Our web-based resource capitalizes upon these approaches 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.

RESULTS AND DISCUSSION

Generation of gene expression profiles from rice anthers

To globally investigate the key genes critical for rice anther development and pollen formation throughout anther development, 57 Affymetrix arrays generated specifically from anther tissue were selected from GEO (Data S1). These samples were derived from wild-type rice anthers from five experiments that covered anther development stages 2–14 (Ma, 2005; Zhang and Yang, 2014), which present an overview of genes expressed during anther development in rice.

These microarrays were renormalized to make the datasets comparable (see Experimental Procedures and Figure S2). Figure S2d shows a histogram of the normalized data, showing a transition between noise and signal with a peak about 8 ($\log_2$ scale); genes above this level are above average expressed in these samples.

Plots showing how the expression levels varied for all 24 462 genes were made, grouped by either dataset or stage. These are available from a freely accessible webpage searchable by either RAP or MSU identifier (https://www.cpib.ac.uk/anther/riceindex.html). This tool also shows which genes are co-expressed, and whether they are in clusters. Figure S3 shows the plot of *TDR* as an example.

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

A correlation network was developed (see Experimental Procedures), which we have called RiceAntherNet, with 411 060 edges between 9358 genes, based upon wild-type expression data. To generate the edges of the network, Pearson's correlation coefficient (PCC) was calculated between each pair of genes, with a cut-off of 0.9 used to denote two genes as being co-expressed (see Experimental Procedures). The full network is available as a Cytoscape (Shannon *et al.*, 2003; Smoot *et al.*, 2011) file (Data S2). Figure 1 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.

Nine datasets of rice male sterile mutants (detailed in Data S1), 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. Differentially expressed genes (DEGs; those with a fold change ≥ 2) in these mutant datasets mapped onto the network (Figure 2). Mutants that show defects in anther cell division and differentiation (*msp1-4, ostdl1a and tip2*; Nonomura *et al.*, 2003; Zhao *et al.*, 2008; Hong *et al.*, 2012; Fu *et al.*, 2014) had DEGs clustered in region 2 (Figure 2a); mutants with enlarged tapetal cells during the later stages of anther development

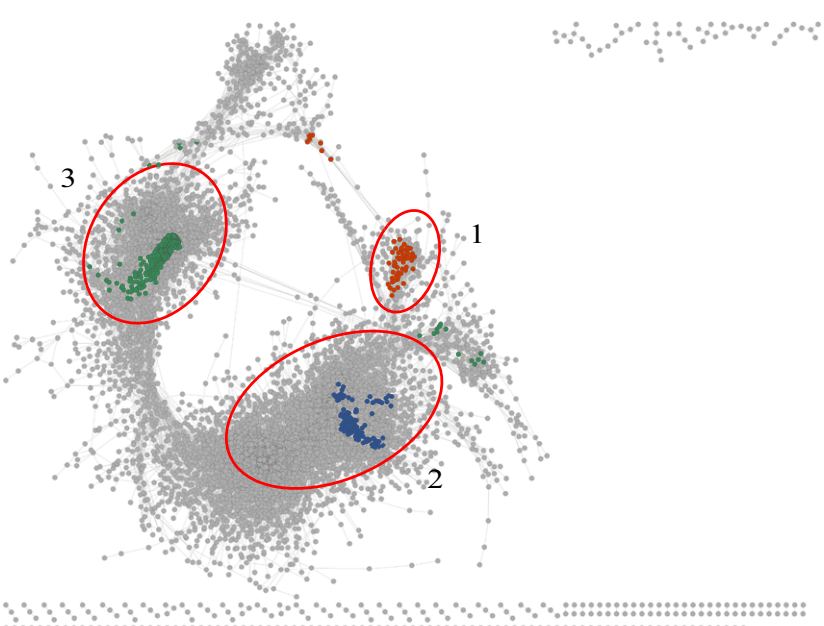

Figure 1. Global co-expression network plot by organic layout in Cytoscape.
The three circled regions refer to the three principal interaction regions. Clade 1 (red) is in region 1; Clade 2 (blue) is in region 2; Clade 3 (green) is in region 3.

(*udt1-1*, *gamyb-2* and *tdr*; Li *et al.*, 2006; Aya *et al.*, 2009; Hu *et al.*, 2011) had a high density of DEGs in region 1 (Figure 2b); 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.*, 2011; Niu *et al.*, 2013) had DEGs distributed throughout all three regions (Figure 2c). 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 upregulated in region 2 (Figure 2d).

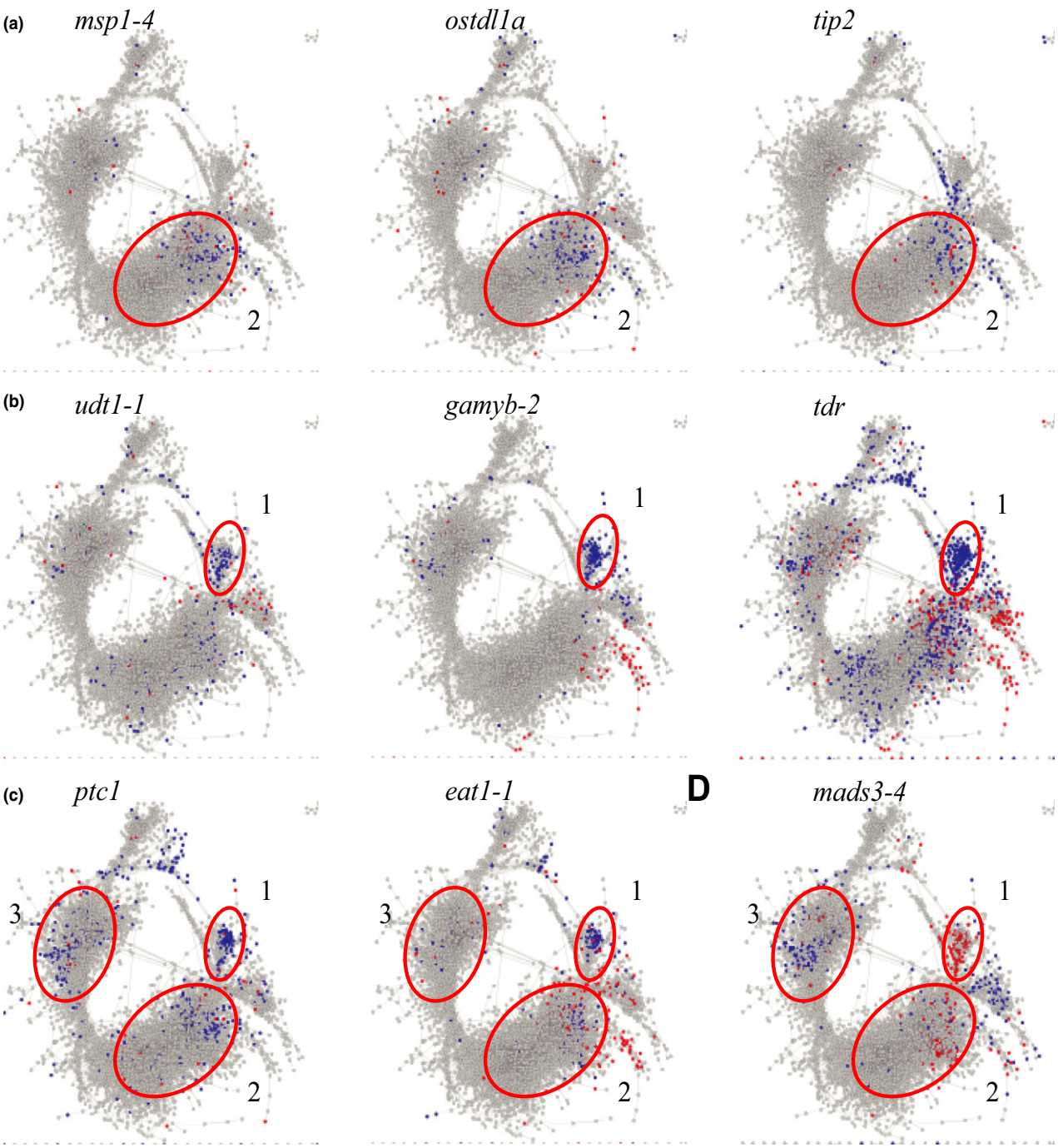

Figure 2. Differentially expressed genes (DEGs) of male sterile mutants in the global RiceAntherNet co-expression network.
(a) Distribution of the DEGs of *msp1-4*, *ostdl1a* and *tip2* mutants.
(b) Distribution of the DEGs of *udt1-1*, *gamyb-2* and *tdr* mutants.
(c) Distribution of the DEGs of *ptc1* and *eat1-1* mutants.
(d) Distribution of the DEGs of *mads3-4* mutants. Upregulated genes are shown in red; downregulated genes are shown in blue.

To further identify genes with similar expression patterns from the global network, the TransClust algorithm (Wittkop *et al.*, 2010) was used to identify clusters of highly intra-connected genes, giving 545 clusters with at least 4 nodes. TransClust produces strongly connected clusters with almost all the possible edges present, Data S3 details the allocation of genes to clusters. Genes in the same cluster had highly correlated expression patterns over whole rice anther developmental stages, and many of them were annotated with the same Gene Ontology (GO) term. For example, Cluster 116 contains eight genes (Figure 3a), all of which have very similar expression levels over the 57 arrays (Figure 3b). Gene expression patterns of published genes within these clusters were compared to validate the clustering and the relative expression patterns were consistent with published expression data. GO analysis using

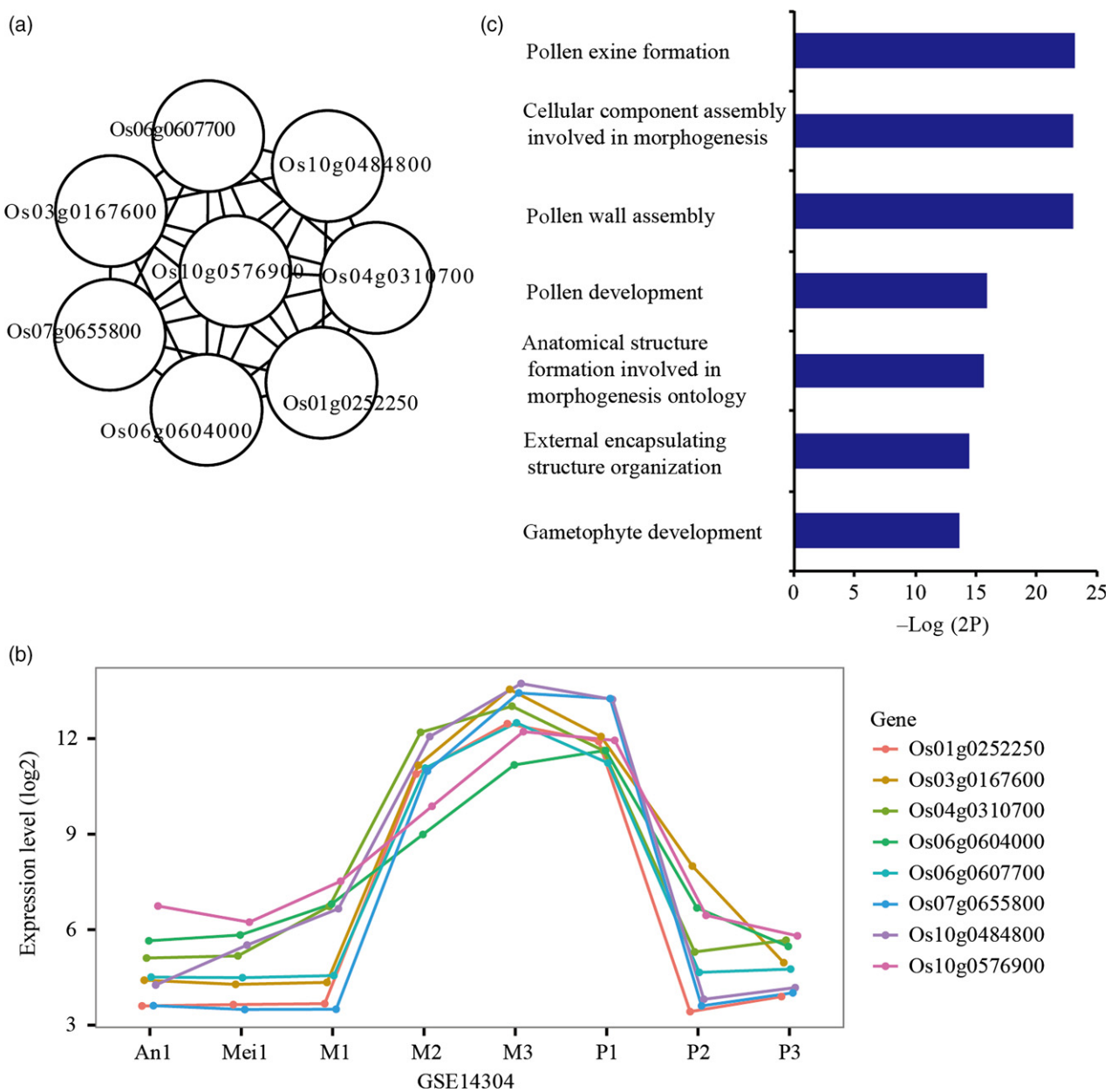

Figure 3. Analysis of gene Cluster 116.
(a) Eight genes are interconnected in Cluster 116.
(b) Gene expression pattern of Cluster 116 in GSE14304. An1, anther stage 1–6; Mei1, anther stage 7; M1, anther stage 8a; M2, anther stage 8b; M3, anther stage 9; P1, anther stage 10; P2, anther stage 11; P3, anther stage 12.
(c) Gene Ontology (GO) terms enriched in Cluster 116. -Log(2P) = -log2(*P*-value).

the orthologous Arabidopsis genes shows that in addition to showing conserved expression patterns, this cluster is functionally enriched for involvement in pollen exine formation (Figure 3c).

When DEGs of nine male sterile mutants were mapped to the clusters, a number of clusters were identified as extremely rich in DEGs from the different mutants, while most had no DEGs. These enriched clusters are therefore likely to be important in relation to anther development, and we therefore selected 29 of them for further analysis (Figure 4). Other clusters that were not enriched for genes associated with anther development often contained genes involved in particular pathways or responses, for example, Cluster 3 is highly enriched for ribosomal proteins.

Using hierarchical clustering (see Experimental procedures), we grouped the 29 selected clusters 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 5). These three clades are referred to as interaction regions 1, 2 and 3, respectively (Figure 1). Most genes in Clade 1 are significantly downregulated 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 upregulated in Clade 1 and Clade 2, and downregulated in Clade 3, which is consistent with its distinctive role (Figure 4). 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 suggests the divided functions of different clades and clusters (Data S4). Clade 1 shows GO enrichment of pollen exine formation, Clade 2 has overrepresented GO terms in organ development, and Clade 3 has overrepresented GO terms in developmental growth involved in morphogenesis.

Clade 1 includes genes for pollen wall formation

There are 60 genes in Clade 1, these are predominantly expressed during anther stages 8 and 9. 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 that 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 (Data S5; Figure 6). The high percentage of DEGs in Clade 1 indicates the important role of these genes in anther and pollen development.

Genes in Clade 1.1 are specifically expressed at stage 8 and stage 9

There are three co-expression clusters identified in Clade 1.1 (Figure 5), which include Clusters 116, 121 and 281, which contain eight, eight and five genes, respectively

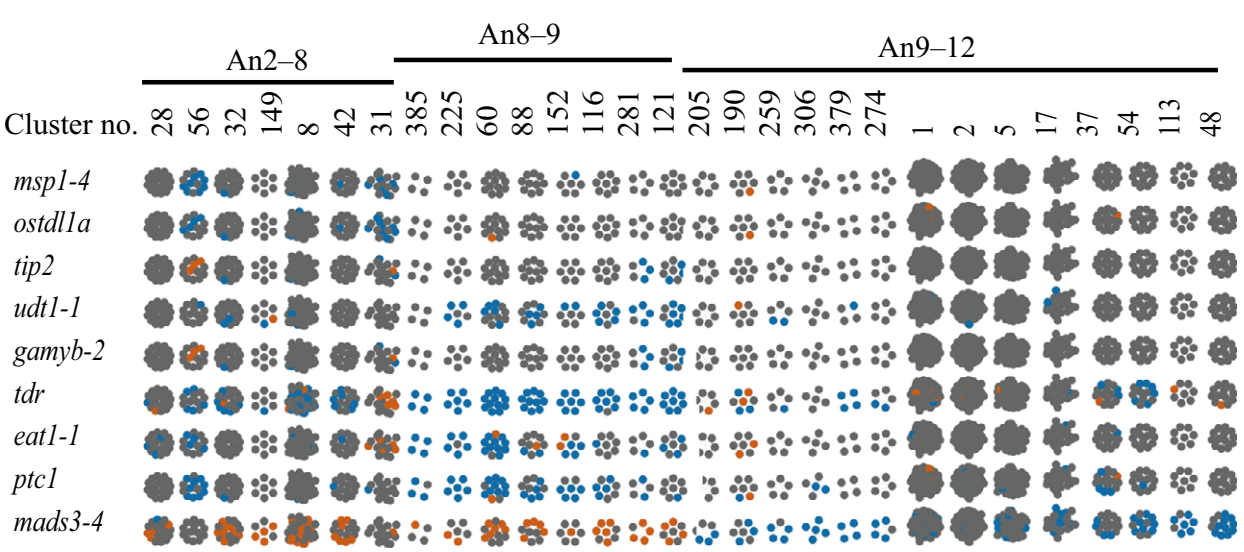

Figure 4. Example clusters from the RiceAntherNet.
Differentially expressed genes (DEGs) from male sterile mutants datasets were mapped onto RiceAntherNet. Downregulated genes are marked in blue, and upregulated genes are marked in orange.

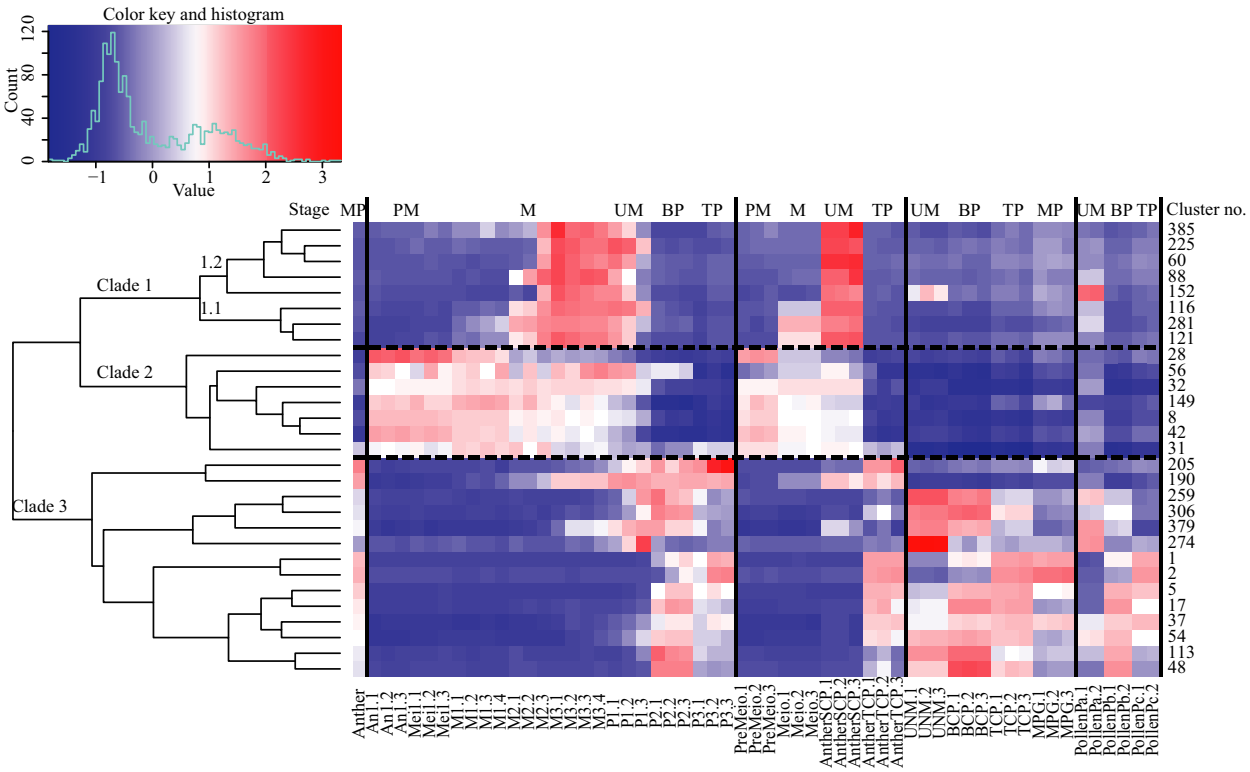

Figure 5. Hierarchical clustering of example RiceAntherNet clusters.
Clusters that are expressed at different anther develop stages separate into three Clades. Clade 2 genes are expressed throughout An2–8; Clade 1 are expressed during An8 and An9; Clade 3 expressed after An9. PM (premeiosis), anther stage 1–6; M (meiosis), anther stage 7 and 8; UM (ininucleate 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.

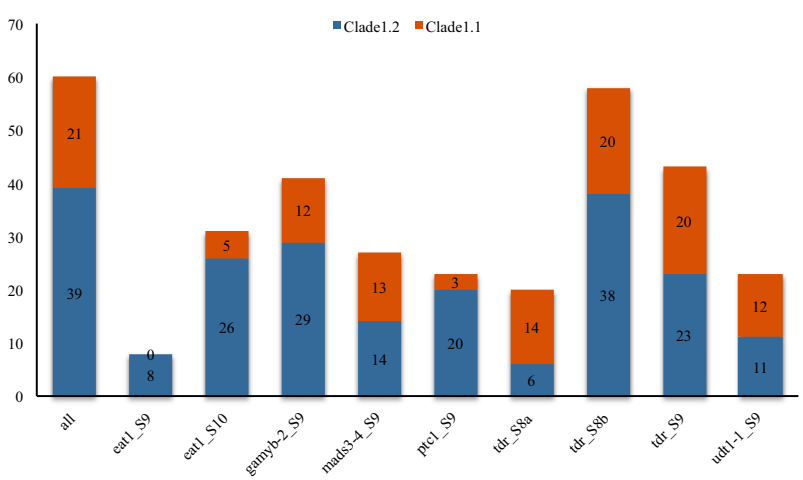

Figure 6. Number of differentially expressed genes (DEGs) in Clade 1.
The first bar shows the size of Clade 1. The remaining bars are DEGs in various mutants: *eat1-1* mutant at anther stage 9 and 10; *gamyb-2* mutant at anther stage 9; *mads3-4* mutant at anther stage 9; *ptc1* mutant at anther stage 9; *tdr* mutant at anther stage 8a, stage 8b and stage 9; *udt1-1* mutant at anther stage 9, respectively. Clade 1.1 is in orange; Clade 1.2 is in blue. [Colour figure can be viewed at wileyonlinelibrary.com]

(Table 1). Genes in Clade 1.1 are specifically expressed at stage 8 and stage 9 (Figure 5). 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.

Among the 21 genes within the clade, there are a number of lipid synthesis genes, such as Os03 g0167600 (*DPW*), Os08 g0131100 (*CYP703A3*) that are important for sporopollenin synthesis. Rice *DPW* and its putative ortholog in Arabidopsis *MS2* (*MALE STERILE2*) 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

Table 1 Genes in Clade 1.1

RAP ID	Cluster	Arabidopsis orthologs	Reported function/annotation
Os06 g0607700/OsABCG15	116	AT3G13220/WBC27	**Involved in sporopollenin transport in rice and Arabidopsis**
Os10 g0484800/OsLAP6	116	AT1G02050/LAP6	**Involved in fatty acid modifications in rice and Arabidopsis**
Os06 g0604000	116	AT1G15360/SHN1	**Involved in lipid synthesis in Arabidopsis**
Os03 g0167600/DPW	116	AT3G11980/MS2	**Involved in fatty acid modifications in rice and Arabidopsis**
Os10 g0576900	116	AT2G23910	NAD(P)-binding domain containing protein
Os07 g0655800	116	AT3G63095	Similar to aspartic acid-rich protein
Os04 g0310700	116	NA	Putative uncharacterized protein
Os01 g0252250	116	NA	Putative uncharacterized protein
Os03 g0769500	121	AT3G50770/CML41	EF-hand type domain containing protein
Os02 g0596200	121	AT1G13130	Glycoside hydrolase, family 5 protein
Os07 g0411300/OsLAP5	121	AT4G34850/LAP5	**Involved in sporopollenin biosynthesis in rice and Arabidopsis**
Os08 g0131100/CYP703A3	121	AT1G01280/CYP703A2	**Involved in sporopollenin biosynthesis in rice and Arabidopsis**
Os08 g0515900	121	AT4G35420/DRL1	**Involved in flavonoid biosynthesis in Arabidopsis**
Os02 g0678300	121	AT1G12570	**Downstream of SHN1 (SHINE1) in Arabidopsis**
Os01 g0127500	121	AT1G68540/TKPR2	**Involved in flavonoid biosynthesis in Arabidopsis**
Os01 g0121900	121	AT5G37870	Seven-in-absentia protein, sina domain containing protein
Os03 g0263600	281	AT3G59530/LAP3	**Involved in flavonoid biosynthesis in Arabidopsis**
Os02 g0101900	281	AT2G16630	Conserved hypothetical protein
Os04 g0404400	281	AT1G47980	Similar to H0502B11.4 protein
Os03 g0687700	281	NA	Ribosome-inactivating protein domain containing protein
Os03 g0781600	281	NA	Conserved hypothetical protein

Genes with known functions related to pollen development are in bold.

hydroxylation in rice, while the homolog *CYP703A2* has a similar conserved function in Arabidopsis (Morant *et al.*, 2007; de Azevedo Souza *et al.*, 2009; Dobritsa *et al.*, 2009a, b; Li *et al.*, 2010; Yang *et al.*, 2014). Loss of function mutants of *DPW* and *CYP703A3* are defective in pollen wall formation (Yang *et al.*, 2014; Xu *et al.*, 2016).

In Arabidopsis, the medium- to long-chain fatty acids that are required for sporopollenin synthesis and pollen wall formation 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 and PKSB into tri- and tetra-ketide α-pyrone (Kim *et al.*, 2010; Wang *et al.*, 2013) and reduced by *DRL1* (*Dihydroflavonol 4-reductase-like1*)/*TKPR1* and *TKPR2* (Grienenberger *et al.*, 2010). These genes [*ACOS5* (*CoA Synthetase5*; Qin *et al.*, 2016), *LAP6/PKSA* (*LAP6/POLYKETIDESYNTHASEA*; Kim *et al.*, 2010), *LAP5/PKSB* (*LAP5/POLYKETIDESYNTHASEB*; Kim *et al.*, 2010), *DRL1/TKPR1* and *TKPR2* (Grienenberger *et al.*, 2010)] are all co-expressed in Arabidopsis. *Os10 g0484800* (*OsLAP6*), ortholog of *PKSA/LAP6*, and Os07 g0411300 (*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, *OsLAP5* and *OsLAP6* show conserved biochemical activity in sporopollenin synthesis (Kim *et al.*, 2010; Wang *et al.*, 2013). 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 os*acos12* mutant shows defects in pollen development (Li *et al.*, 2016; Yang *et al.*, 2017). For the TKPRs in Clade 1.1, Os08 g0515900 is an ortholog of *TKPR1* and Os01 g0127500 is the ortholog in *TKPR2* (Figure S4a). 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 S4b). For the conservation of the co-expression pattern as well as the biochemical and biological function of genes in this pathway, we speculate that Os01 g0127500 and Os08 g0131100 are likely to have key roles in pollen development. These findings illustrate the importance of RiceAntherNet for identifying novel genes required for pollen development.

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,b). Correspondingly, the ortholog of *LAP3* (*LESS ADHESIVE POLLEN3*) in rice, Os03 g0263600, is also clustered in Clade 1.1 in RiceAntherNet.

The ortholog of Os06 g0604000, *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 Os02 g0678300 (At1 g12570) is a WIN1/SHN1 inducible gene that 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 (Libeisson *et al.*, 2011). Both Os06 g0604000 and Os02 g0678300 are clustered in Clade 1.1, further validating the robust correlation of functionally related genes in RiceAntherNet.

In summary, for all 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 (Os06 g0604000, Os02 g0678300, Os08 g0515900, Os01 g0127500 and Os03 g0263600) 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 we hypothesize that they are also likely to be involved in pollen wall formation, as they are co-expressed with 10 key genes involved in pollen exine formation. These genes are therefore good targets for future study.

Clade 1.2 genes show lower level of conservation to Arabidopsis

The genes clustering in Clade 1.2 are distinct from those in Clade 1.1 and 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 2). Unlike Clade 1.1 and its associated cluster in Arabidopsis FlowerNet (Cluster 37), which have conserved expression patterns and functions between them, genes in Clade 1.2 did not show a 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 7), while their orthologs in Arabidopsis are widely expressed (eFP Browser; Winter *et al.*, 2007; http://bar.utoronto.ca/efp/). These findings suggest that although pollen regulatory networks are similar between monocots and dicots, there are still significant differences between them. It may be that pollen development involves more integrative coordination and is a more complicated process in rice, as compared with Arabidopsis.

Os10 g0524500, *NO POLLEN 1* (*OsNP1*) 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, 2003). This finding suggests a distinct process for pollen development in rice in which RiceAntherNet can help to enable new gene identification. Accordingly, Os12 g0242700, 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 (Mou *et al.*, 2000; Libeisson *et al.*, 2011). The Arabidopsis *KAR* gene, At1 g24360, 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). Os12 g0242700 is specifically expressed in rice anthers, it therefore may act as a fatty acid synthesis gene required for anther development.

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 (Ondzighi *et al.*, 2008); however, no role has been reported during pollen development. These findings further support the idea that rice has distinct and possibly more complicated aspects of its pollen development compared with Arabidopsis. In addition to cysteine proteases, aspartyl proteases are of vital importance for tapetum cell death (Niu *et al.*, 2013). Os04 g0448500 is a eukaryotic aspartyl protease family protein, specifically expressed during tapetum PCD stage, which is co-expressed with *OsCP1*. However, the homologous gene to Os04 g0448500 in Arabidopsis (AT2G03200) is not expressed in anther. Here we have found genes specifically expressed during rice anther development, which are not conserved in Arabidopsis, orthologs of these genes show no evidence that they are required for pollen development. These findings suggest that although there is conservation in many regulatory aspects of pollen development in rice and Arabidopsis, 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.

Table 2 Genes in Clade 1.2

RAP ID	Cluster	Arabidopsis orthologs	Reported function/annotation
Os04 g0448500	60	AT2G03200	Peptidase A1 domain containing protein
Os11 g0215400	60	NA	Peptidase aspartic, catalytic domain containing protein
Os05 g0169600	60	AT5G56040/SKM2	**Regulates fertility in Arabidopsis**
Os10 g0337400	60	AT3G47570	Similar to Taxa-1
Os01 g0772100	60	AT3G03760/LBD20	Lateral organ boundaries, LOB domain containing protein
Os09 g0401000	60	AT5G16600/MYB43	Cell differentiation and regulation of secondary cell wall biogenesis
Os03 g0564600	60	AT3G22060	Protein of unknown function DUF26 domain containing protein
Os12 g0131900	60	AT1G09890	Similar to LG27/30-like
Os03 g0212000	60	NA	Hypothetical conserved gene
Os03 g0261100	60	NA	Phospholipase A2 family protein
Os07 g0585250	60	NA	Putative uncharacterized protein
Os08 g0298600	60	NA	Putative uncharacterized protein
Os06 g0228600	88	AT1G10010/AAP8	Similar to amino acid carrier
Os02 g0110000	88	AT5G45950	Lipase, GDSL domain containing protein
Os10 g0524500/OsNP1	88	AT1G12570	**Required for tapetum degradation in rice**
Os01 g0242400	88	AT1G06280/LBD2	LOB domain containing protein
Os02 g0503500	88	AT1G24360	Similar to 3-oxoacyl-[acyl-carrier-protein] reductase
Os04 g0613200	88	AT4G20050/QRT3	**Degradation of the pollen mother cell wall in Arabidopsis**
Os03 g0850900	88	AT2G02850/ARPN	**Involved in anther development and pollination in Arabidopsis**
Os04 g0165000	88	NA	Similar to thaumatin-like protein 1a
Os07 g0556800	88	NA	Ribosome-inactivating protein family protein
Os08 g0360700	88	AT3G63095	Conserved hypothetical protein
Os04 g0670500/OsCP1	152	AT5G43060	**Downstream of TDR in rice**
Os02 g0219000	152	AT1G27760/SAT32	Involved in salt tolerance in Arabidopsis
Os09 g0273800	152	AT1G75950/SKP1	Similar to Fimbriata-associated protein
Os08 g0375700	152	AT1G20140/SK4	BTB/POZ fold domain containing protein
Os03 g0783100	152	NA	Conserved hypothetical protein
Os04 g0267000	152	NA	Putative uncharacterized protein
Os01 g0641200	152	NA	Ubiquitin-like protein-NEDD8-like protein RUB3
Os02 g0653200	225	AT3G60270	Cupredoxin domain containing protein
Os01 g0170300	225	AT5G56040/SKM2	**Regulates fertility in Arabidopsis**
Os01 g0825000	225	NA	Similar to LOB domain protein 12
Os12 g0242700	225	AT1G24360	Similar to 3-oxoacyl-[acyl-carrier-protein] reductase 1
Os01 g0112400	225	AT2G23910	Major intrinsic protein family protein
Os04 g0543700	225	AT3G63095	Similar to serine proteinase
Os06 g0611400	385	AT4G18180	Pectin lyase fold/virulence factor domain containing protein
Os03 g0145900	385	AT5G22920	Similar to CHY zinc finger family protein, expressed
Os03 g0701400	385	AT2G13600	Similar to pentatricopeptide repeat protein PPR868-14
Os02 g0790600	385	AT2G18650/MEE16	Involved in embryo development in Arabidopsis

Genes with known functions related to pollen development are in bold.

Os05 g0169600, Os01 g0170300 are leucine-rich receptor-like (LRR) protein kinases downregulated in *eat1-1*, *gamyb-2*, and *tdr* and *ptc1* mutants. They share the putative ortholog AT5G56040 in Arabidopsis that encodes a LRR 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 signalling pathway to regulate fertility in Arabidopsis (Endo *et al.*, 2013). Therefore, Os05 g0169600, Os01 g0170300 might be required for fertility in rice.

The ortholog of Os04 g0613200, *QRT3* (QUARTET 3), is a polygalacturonase, specifically expressed in the tapetum when microspores separate from their meiotic siblings in Arabidopsis (Rhee *et al.*, 2003). Mutant *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.*, 2003; Figure 7). *QRT3* is largely reduced in *dyt1-3*, a callose degradation defective mutant (Zhu *et al.*, 2015). Compared with Arabidopsis, Os04 g0613200 shows a similar expression pattern in rice, which is highly reduced in those mutants in which callose is not degraded 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 (Data S5). These findings suggest that Os04 g0613200 might be a regulator for callose degradation in rice.

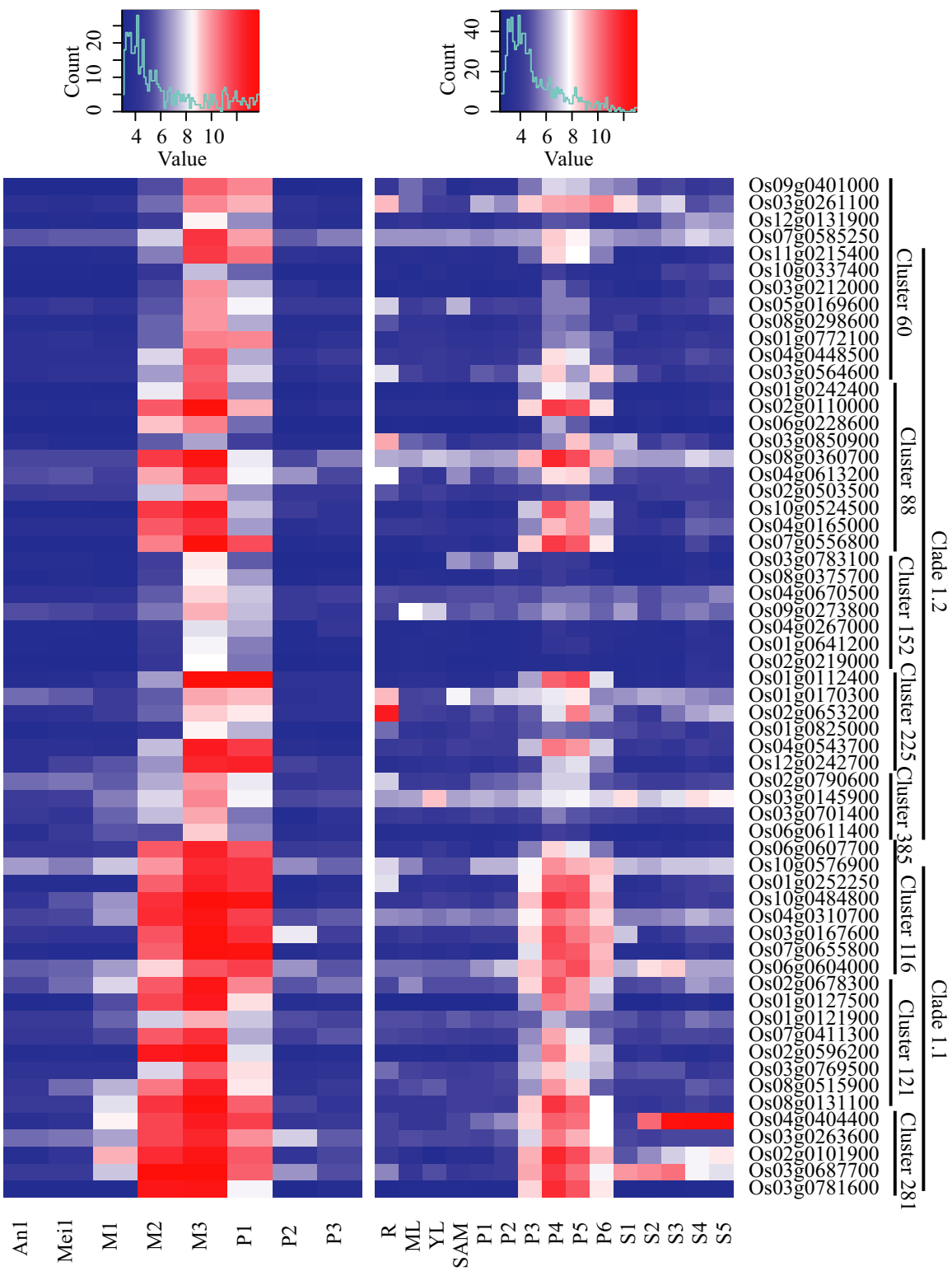

Figure 7. Expression patterns of Clade 1 genes in rice anther and other tissues.
Left panel: heatmap of Clade 1 in anther tissues (dataset GSE14304); right panel: heatmap of Clade 1 in other tissues. Stages in left panel: An1, anther stage 1–2; Mei, anther stage 3–6; M1, anther stage 7; M2, anther stage 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 days after pollination (DAP); S2, seed, 3–4 DAP; S3, seed, 5–10 DAP; S4, seed, 11–20 DAP; S5, seed, 21–29 DAP.

Of the 39 genes in Clade 1.2, 37 are currently uncharacterized, 33 of these show different expression patterns between rice and Arabidopsis, and have 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.

Clade 2 and Clade 3 genes show different expression patterns from Clade 1 genes

Genes in Clade 2 are expressed earlier than Clade 1, from anther stage 2 to stage 9, while genes in Clade 3 are expressed later than Clade 1, from stage 10 to stage 12. Compared with Clade 1, Clades 2 and 3 do not have high percentages 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 S5). Surprisingly, Cluster 8 in Clade 2 contains eight LRR-RLK (leucine-rich repeat receptor-like kinase) genes (Figure S6 shows the phylogenetic analysis of these genes). Of these eight genes, Os03 g0773700 and Os07 g0134200 are orthologs of *BAM1*, which is required for anther cell differentiation (Hord *et al.*, 2006); while another early anther development involved gene, *ER* (*ERECTA*; Hord *et al.*, 2008) has two orthologs in rice, Os06 g0203800 and Os02 g0777400, 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 four LRR-RLKs have not been reported either in rice or in Arabidopsis, but we 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.

Among the 572 genes in Clade 3 that are expressed in late anther development, 139 are DEGs and 106 of them are anther-specific genes (Figure S7), suggesting that these genes play specific roles in pollen development at these later stages. Os01 g0939100 in Clade 3 is the putative ortholog of *ACA11* in Arabidopsis, which is specifically expressed in the anther in the later stages. *ACA11* is a Type V P-type ATPase expressed in pollen, which is required for ion homeostasis and pollen development (Jakobsen *et al.*, 2005); therefore, Os01 g0939100 may function in an orthologous manner and might be required for pollen development in rice.

CONCLUSIONS

RiceAntherNet is a rice co-expression network that 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 with that in other species, in particular Arabidopsis.

Datasets of nine 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 with wild-type. Based on previous research, seven genes (*OsABCG15*, *DPW*, *OsLAP6*, *CYP703A3*, *OsLAP5*, *OsNP1*, *OsCP1*) out of 21 in Clade 1.1 have characterized roles in pollen wall formation in rice. Twelve genes in the RiceAntherNet have orthologs in Arabidopsis that have been reported to act in pollen development, suggesting a conserved function in rice. Using this tool, we have identified numerous uncharacterized co-expressed genes that are likely to be critical components of pollen regulatory networks. This illustrates that the gene co-expression network generated by array data is a valuable tool to identify new genes involved in the related process and to predict novel gene functions.

Compared with Arabidopsis and rice, genes in RiceAntherNet Clade 1.1 also show co-expressed patterns in FlowerNet in Arabidopsis (Pearce *et al.*, 2015). The similarity between the known gene regulation networks in Arabidopsis and rice during pollen development has been indicated previously (Wilson and Zhang, 2009). This further validates the reliability of both RiceAntherNet and the Arabidopsis FlowerNet for new gene discovery. 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 confirmation of the gene expression profile in anthers for a gene of interest.

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 with previous co-expression networks in rice, our focus is on anther and pollen development, and RiceAntherNet highlights gene expression patterns throughout all anther developmental stages. As opposed to the chain-like clusters commonly generated, the tightly connected clusters in RiceAntherNet find genes that are frequently involved in the same process, meaning that starting target genes are not needed to predict new genes. Although here we only predict novel genes involved in pollen wall formation by using mutants that are defective in pollen wall formation, RiceAntherNet is an efficient and reliable method 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 (Data

S2) 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.

EXPERIMENTAL PROCEDURES

Datasets

We used 57 Affymetrix arrays of anther samples of wild-type rice (from different cultivars) to generate a correlation network. These samples originated from five publicly available experiments (Li *et al.*, 2007; Fujita *et al.*, 2010; Wei *et al.*, 2010; Deveshwar *et al.*, 2011; Peng *et al.*, 2012), including samples from anther stages 2–14. A further 45 Affymetrix microarrays of samples from other tissues were used to characterize the global expression patterns in rice. The raw.cel data files of Affymetrix GeneChip Rice Genome Array were downloaded from NCBI's GEO (Barrett *et al.*, 2013); details of all arrays are listed in Data S1.

Data processing

Quality control was performed to check the signal intensity and variance of each microarray (Figure S2). RMA normalization (Irizarry, 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.

The set of genes was filtered by removing those that have no samples with expression level above 8 (on the log2 scale), leaving 14 468 genes that have at least one higher expressed value. We then used PCC to calculate the correlation between all remaining genes, giving a correlation matrix, the histogram of which gives a smooth curve (Figure S8). A semi-arbitrary cut-off was then applied to this matrix to give a set of edges to analyse; in line with previous studies that 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 cut-off of 0.9, leading to 411 060 edges between 9358 genes. Changing this cut-off would lead to more or less edges, but the general structure of the network would stay the same. This choice of a cut-off of 0.9 was confirmed by analysing known co-expressed genes, such as CYP704B2 and CYP703A3 (PCC = 0.96), ABCG15 and DPW (PCC = 0.94; Li and Zhang, 2010; Yang *et al.*, 2014).

Network construction

Those 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 (Figure 1). 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. When used on this network, TransClust produces 545 clusters that contain at least four genes; these clusters contain 4580 genes and 73 556 intra-cluster edges, 99.74% of the possible number (Data S2).

Clades analysis

To aggregate the clusters with similar expression trend together, the expression data for each gene were standardized, and an average expression pattern calculated for each cluster. These average expression patterns were clustered using complete-linkage hierarchical clustering to give clades of clusters with similar expression trends (Figure 5).

Annotation and ortholog analysis

Rice (IRGSP-1.0_representative_2016-08-05) and Arabidopsis annotation files (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, 2003), respectively. Arabidopsis orthologs were found by Local BLAST using BLAST+ v.2.2.23 with e-value 0.0000001 (otherwise default values). Protein sequences of rice and Arabidopsis were generated from GFF files downloaded from TIGR (http://rice.plantbiology.msu.edu/pub/data/Eukaryotic_Projects/o_sativa/annotation_dbs/pseudomolecules/) and TAIR (http://www.arabidopsis.org/download/index.jsp), respectively.

Phylogenetic analysis

Protein sequences (see Data S6) were used to construct a neighbour-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 cut-off = 30%). Phylogenetic tree was construct by neighbour-joining (bootstrap replications = 1000).

ACKNOWLEDGEMENTS

SPP is supported through an Early Career Fellowship from the Leverhulme Trust. This work was supported by the Biotechnology and Biological Sciences Research Council [grant number BB/J001295/1; LH was supported by a Vice Chancellors PhD Scholarship of Research Excellence from University of Nottingham and Shanghai Jiao Tong University.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

ZAW, DZ: designed the research; SPP, HL: bioinformatic analysis; HL, JY, SPP: performed the research; ZAW, HL, JY, SPP: wrote the manuscript; ZAW, HL, JY, SPP, DZ: edited the manuscript; SPP, HL: website creation.

FUNDING INFORMATION

SPP is supported through an Early Career Fellowship from the Leverhulme Trust. This work was supported by the

Biotechnology and Biological Sciences Research Council [grant number BB/J001295/1; LH was supported by a Vice Chancellors PhD Scholarship of Research Excellence from University of Nottingham and Shanghai Jiao Tong University.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Diagrams of biological events and major regulators during anther dsevelopment.

Figure S2. Boxplots and density plots of log-intensity distribution for 57 arrays before and after normalization.

Figure S3. Gene expression profile of *TDR* on the webtool (www.cpib.ac.uk/anther/riceindex.html).

Figure S4. Analysis of TKPR genes in Clade 1.

Figure S5. Expression patterns of Clade 2 in rice anther and other tissues.

Figure S6. Phylogenetic tree of LRR-RLK genes in Clade 1.

Figure S7. Expression patterns of Clade 3 in rice anther and other tissues.

Figure S8. Distribution of edge correlations.

Data S1. Wild-type and mutants microarray datasets.

Data S2. RiceAntherNet Cytoscape file.

Data S3. Annotation of genes in RiceAntherNet.

Data S4. GO enrichment analysis for three clades and main clusters.

Data S5. Fold changes of DEGs in Clade 1, Clade 2 and Clade 3.

Data S6. Protein sequences used for generating phylogenetic trees.

REFERENCES

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

Aya, K., Suzuki, G., Suwabe, K. *et al.* (2011) Comprehensive network analysis of anther-expressed genes in rice by the combination of 33 laser microdissection and 143 spatiotemporal microarrays. *PLoS ONE*, **6**, e26162.

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

Barrett, T., Wilhite, S.E., Ledoux, P. *et al.* (2013) NCBI GEO: archive for functional genomics data sets-update. *Nucleic Acids Res.* **41**, D991–D995.

Bassel, G.W., Glaab, E., Marquez, J., Holdsworth, M.J. and Bacardit, J. (2011) Functional network construction in Arabidopsis using rule-based machine learning on large-scale data sets. *Plant Cell*, **23**, 3101–3116.

Cao, P., Jung, K.H., Choi, D., Hwang, D., Zhu, J. and Ronald, P.C. (2012) The rice oligonucleotide array database: an atlas of rice gene expression. *Rice (N Y)*, **5**, 17.

Chang, Z., Chen, Z., Wang, N., Xie, G., Lu, J., Yan, W., Zhou, J., Tang, X. and 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. USA*, **113**, 14145–14150.

Chen, W., Yu, X.H., Zhang, K., Shi, J., Oliveira, S.D., Schreiber, L., Shanklin, J. and Zhang, D. (2011) Male Sterile2 encodes a plastid-localized fatty acyl carrier protein reductase required for pollen exine development in Arabidopsis. *Plant Physiol.* **157**, 842–853.

Choi, H., Jin, J.Y., Choi, S., Hwang, J.U., Kim, Y.Y., Suh, M.C. and 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**, 181–193.

Dai, M., Wang, P., Boyd, A.D. *et al.* (2005) Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res.* **33**, e175.

Dekkers, B.J., Pearce, S., van Bolderen-Veldkamp, R.P. *et al.* (2013) Transcriptional dynamics of two seed compartments with opposing roles in Arabidopsis seed germination. *Plant Physiol.* **163**, 205–215.

Deveshwar, P., Bovill, W.D., Sharma, R., Able, J.A. and Kapoor, S. (2011) Analysis of anther transcriptomes to identify genes contributing to meiosis and male gametophyte development in rice. *BMC Plant Biol.* **11**, 78.

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

Dobritsa, A.A., Shrestha, J., Morant, M., Pinot, F., Matsuno, M., Swanson, R., Møller, B.L. and Preuss, D. (2009b) CYP704B1 is a long-chain fatty acid omega-hydroxylase essential for sporopollenin synthesis in pollen of Arabidopsis. *Plant Physiol.* **151**, 574–589.

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

Feng, B., Lu, D., Ma, X., Peng, Y., Sun, Y., Ning, G. and Ma, H. (2012) Regulation of the Arabidopsis anther transcriptome by DYT1 for pollen development. *Plant J.* **72**, 612–624.

Fu, Z., Yu, J., Cheng, X., Zong, X., Xu, J., Chen, M., Li, Z., Zhang, D. and 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**, 1512–1524.

Fujita, M., Horiuchi, Y., Ueda, Y. *et al.* (2010) Rice expression atlas in reproductive development. *Plant Cell Physiol.* **51**, 2060–2081.

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

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

Hamada, K., Hongo, K., Suwabe, K. *et al.* (2011) OryzaExpress: an integrated database of gene expression networks and omics annotations in rice. *Plant Cell Physiol.* **52**, 220–229.

Hong, L., Tang, D., Shen, Y., Hu, Q., Wang, K., Li, M., Lu, T. and Cheng, Z. (2012) MIL2 (MICROSPORELESS2) regulates early cell differentiation in the rice anther. *New Phytol.* **196**, 402–413.

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

Hord, C.L., Sun, Y.J., Pillitteri, L.J., Torii, K.U., Wang, H., Zhang, S. and 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**, 645–658.

Hu, L., Liang, W., Yin, C., Cui, X., Zong, J., Wang, X., Hu, J. and Zhang, D. (2011) Rice MADS3 regulates ROS homeostasis during late anther development. *Plant Cell*, **23**, 515–533.

Huang, J., Zhang, T., Linstroth, L., Tillman, Z., Otegui, M.S., Owen, H.A. and Zhao, D. (2016) Control of anther cell differentiation by the small protein ligand TPD1 and its receptor EMS1 in Arabidopsis. *PLOS Gen.* **12**, e1006147.

Irizarry, R.A. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*, **4**, 249–264.

Jakobsen, M.K., Poulsen, L.R., Schulz, A., Fleuratlessard, P., Møller, A., Husted, S., Schiøtt, M., Amtmann, A. and 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 Dev.* **19**, 2757–2769.

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

Jia, G. and Zhao, D. (2008) Signaling of cell fate determination by the TPD1 small protein and EMS1 receptor kinase. *Proc. Natl Acad. Sci. USA*, **105**, 2220–2225.

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

Kafieh, R. and Mehridehnavi, A. (2012) A comprehensive comparison of different clustering methods for reliability analysis of microarray data. *J. Med. Signals Sensors*, **3**, 22–30.

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

Kim, S.S., Grienenberger, E., Lallemand, B. *et al.* (2010) 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**, 4045–4066.

Kolesnikov, N., Hastings, E., Keays, M. *et al.* (2015) ArrayExpress update – simplifying data submissions. *Nucleic Acids Res.* **43**, D1113–D1116.

Kumar, S., Stecher, G. and Tamura, K. (2016) MEGA7: molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874.

Lee, S., Jung, K.H., An, G. and Chung, Y.Y. (2004) Isolation and characterization of a rice cysteine protease gene, OsCP1, using T-DNA gene-trap system. *Plant Mol. Biol.*, **54**, 755–765.

Li, H. and Zhang, D.B. (2010) Biosynthesis of anther cuticle and pollen exine in rice. *Plant Signal. Behav.* **5**, 1121–1123.

Li, N., Zhang, D.S., Liu, H.S. *et al.* (2006) The rice tapetum degeneration retardation gene is required for tapetum degradation and anther development. *Plant Cell*, **18**, 2999–3014.

Li, M., Xu, W., Yang, W., Kong, Z. and Xue, Y. (2007) Genome-wide gene expression profiling reveals conserved and novel molecular functions of the stigma in rice. *Plant Physiol.* **144**, 1797–1812.

Li, H., Pinot, F., Sauveplane, V., Werckreichhart, D., Diehl, P., Schreiber, L., Franke, R., Zhang, P., Chen, L. and Gao, Y. (2010) Cytochrome P450 family member CYP704B2 catalyzes the hydroxylation of fatty acids and is required for anther cutin biosynthesis and pollen exine formation in rice. *Plant Cell*, **22**, 173–190.

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

Li, Y., Li, D., Guo, Z., Shi, Q., Xiong, S., Zhang, C., Zhu, J. and 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**, 256.

Libeisson, Y., Shorrosh, B., Beisson, F., Andersson, M.X., Arondel, V., Bates, P.D., Baud, S., Bird, D., Debono, A. and Durrett, T.P. (2011) Acyl-lipid metabolism. *Arabidopsis Book*, **11**, e0133.

Liu, Z., Lin, S., Shi, J., Yu, J., Zhu, L., Yang, X., Zhang, D. and Liang, W. (2017) Rice No Pollen 1 (NP1) is required for anther cuticle formation and pollen exine patterning. *Plant J.* **91**, 263–277. https://doi.org/10.1111/tpj.13561.

Ma, H. (2005) Molecular genetic analyses of microsporogenesis and microgametogenesis in flowering plants. *Ann. Rev. Plant Biol.*, **56**, 393–434.

Millar, A.A. and Gubler, F. (2005) The Arabidopsis GAMYB-like genes, MYB33 and MYB65, are microRNA-regulated genes that redundantly facilitate anther development. *Plant Cell*, **17**, 705–721.

Morant, M., Jørgensen, K., Schaller, H., Pinot, F., Møller, B., Werck-Reichhart, D. and 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**, 1473–1487.

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

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

Niu, N., Liang, W., Yang, X., Jin, W., Wilson, Z.A., Hu, J. and Zhang, D. (2013) EAT1 promotes tapetal cell death by regulating aspartic proteases during male reproductive development in rice. *Nat. Comms.*, **4**, 66–78.

Nonomura, K.I., Miyoshi, K., Eiguchi, M., Suzuki, T., Miyao, A., Hirochika, H. and 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**, 1728–1739.

Ondzighi, C.A., Christopher, D.A., Cho, E.J., Chang, S.C. and 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**, 2205–2220.

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

Paxson-Sowders, D.M., Dodrill, C.H., Owen, H.A. and Makaroff, C.A. (2001) DEX1, a novel plant protein, is required for exine pattern formation during pollen development in Arabidopsis. *Plant Physiol.* **127**, 1739–1749.

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

Peng, H., Chun, J., Ai, T.B., Tong, Y.A., Zhang, R., Zhao, M.M., Chen, F. and Wang, S.H. (2012) MicroRNA profiles and their control of male gametophyte development in rice. *Plant Mol. Biol.* **80**, 85–102.

Pruitt, K.D., Tatusova, T. and Maglott, D.R. (2005) NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* **33**, D501–D504.

Qin, P., Tu, B., Wang, Y., Deng, L., Quilichini, T.D., Li, T., Wang, H., Ma, B. and 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**, 138–154.

Qin, M., Tian, T., Xia, S., Wang, Z., Song, L., Yi, B., Wen, J., Shen, J., Ma, C. and Fu, T. (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**, pcw092.

Rhee, S.Y. (2003) 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**, 224–228.

Rhee, S.Y., Osborne, E., Poindexter, P.D. and Somerville, C.R. (2003) 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**, 1170–1180.

Sakai, H., Lee, S.S., Tanaka, T. *et al.* (2013) Rice Annotation Project Database (RAP-DB): an integrative and interactive database for rice genomics. *Plant Cell Physiol.* **54**, e6.

Sato, Y., Namiki, N., Takehisa, H. *et al.* (2013) RiceFREND: a platform for retrieving coexpressed gene networks in rice. *Nucleic Acids Res.* **41**, D1214–D1221.

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

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

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

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

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

Wang, Z. (2005) Microarray analysis of gene expression involved in anther development in rice (Oryza sativa L.). *Plant Mol. Biol.* **58**, 721–737.

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

Wei, L.Q., Xu, W.Y., Deng, Z.Y., Su, Z., Xue, Y. and Wang, T. (2010) Genome-scale analysis and comparison of gene expression profiles in developing and germinated pollen in Oryza sativa. *BMC Genom.* **11**, 338.

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

Wilson, Z.A., Morroll, S.J., Swarup, R. and Tighe, P.J. (2001) The Arabidopsis MALE STERILITY1 (MS1) gene is a transcriptional regulator of male gametogenesis, with homology to the PHD-finger family of transcription factors. *Plant J.* **28**, 27–39.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G.V. and Provart, N.J. (2007) An "electronic fluorescent pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS ONE*, **2**, e718. https://doi.org/10.1371/journal.pone.0000718.

Wittkop, T., Emig, D., Lange, S., Rahmann, S., Albrecht, M., Morris, J.H., Böcker, S., Stoye, J. and Baumbach, J. (2010) Partitioning biological data with transitivity clustering. *Nat. Meth.* **7**, 419–420.

Xu, J., Ding, Z., Vizcay-Barrena, G., Shi, J., Liang, W., Yuan, Z., Werck-Reichhart, D., Schreiber, L., Wilson, Z.A. and Zhang, D. (2014) ABORTED MICROSPORES acts as a master regulator of pollen wall formation in Arabidopsis. *Plant Cell*, **26**, 1544–1556.

Xu, D., Shi, J., Rautengarten, C., Yang, L., Qian, X., Uzair, M., Zhu, L., Luo, Q., An, G. and Waßmann, F. (2016) Defective Pollen Wall 2 (DPW2) encodes an acyl transferase required for rice pollen development. *Plant Physiol.* **173**, 240.

Yang, X., Di, W., Shi, J., Yi, H., Pinot, F., Grausem, B., Yin, C., Lu, Z., Chen, M. and Luo, Z. (2014) Rice CYP703A3, a cytochrome P450 hydroxylase, is essential for development of anther cuticle and pollen exine. *J. Integr. Plant Biol.* **56**, 979–994.

Yang, L., Qian, X., Chen, M., Fei, Q., Meyers, B.C., Liang, W. and Zhang, D. (2016) Regulatory role of a receptor-like kinase in specifying anther cell identity. *Plant Physiol.* **171**, 2085–2100.

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

Yu, J., Meng, Z., Liang, W., Kudla, J., Tucker, M.R., Luo, Z., Chen, M., Xu, D., Zhao, G. and Wang, J. (2016) A rice Ca^{2+} binding protein is required for tapetum function and pollen formation. *Plant Physiol.* **172**, 01261–02016.

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

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

Zhang, D., Luo, X. and Zhu, L. (2011) Cytological analysis and genetic control of rice anther development. *J. Genet. Genom.* **38**, 379–390.

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

Zhao, G., Shi, J., Liang, W., Xue, F., Luo, Q., Zhu, L., Qu, G., Chen, M., Schreiber, L. and Zhang, D.B. (2015) Two ATP binding cassette G (ABCG) transporters, OsABCG26 and OsABCG15, collaboratively regulate rice male reproduction. *Plant Physiol.* **169**, 2064–2079.

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