Genome-wide identification of *DUF506* gene family in *Oryza sativa* and expression profiling under abiotic stresses

Wei Dong[1,*], Jian Tu[1,*], Wei Deng[1], Jianhua Zhang[1], Yuran Xu[1], Anyu Gu[1], Hua An[1], Kui Fan[2], Rui Wang[2], Jianping Zhang[2], Limei Kui[1] and Xiaolin Li[1]

[1] Yunnan Academy of Agricultural Sciences, Food Crops Research Institute, Kunming, China
[2] Yunnan Grain Industry Group Co., Ltd, Kunming, China
[*] These authors contributed equally to this work.

ABSTRACT

The domain of unknown function 560 (DUF560), also known as the PDDEXK_6 family, is a ubiquitous plant protein that has been confirmed to play critical roles in Arabidopsis root development as well as ABA and abiotic responses. However, genome-wide identification and expression pattern analysis in rice (*Oryza sativa*) still need to be improved. Based on the phylogenetic relationship, 10 *OsDUF506* genes were identified and classified into four subfamilies. Segmental duplication was essential to the expansion of *OsDUF506s*, which were subjected to purifying selective pressure. Except for *OsDUF50609* and *OsDUF50610*, the *OsDUF506s* shared colinear gene pairs with five monocot species, showing that they were conserved in evolution. Furthermore, the conserved domains, gene structures, SNPs distribution, and targeting miRNAs were systematically investigated. Massive cis-regulatory elements were discovered in promoter regions, implying that *OsDUF506s* may be important in hormone regulation and abiotic stress response. Therefore, we analyzed plant hormone-induced transcriptome data and performed qRT-PCR on eight *OsDUF506s* under drought, cold, and phosphorus-deficient stresses. The results revealed that most *OsDUF506s* respond to ABA and JA treatment, as well as drought and cold conditions. In conclusion, our findings provided insights into the evolution and function of *OsDUF506s*, which could benefit crop breeding in the future.

Subjects Agricultural Science, Bioinformatics, Genetics, Genomics, Plant Science
Keywords DUF506, Gene family, Identification, Abiotic stress, Expression profile

INTRODUCTION

Domains of unknown functions (DUFs) are batches of gene families with conserved domains but unknown functions that are common in eukaryotes (*Bateman, Coggill & Finn, 2010*). The Pfam database contains 4,716 DUF families (https://www.ebi.ac.uk/interpro/entry/pfam/). Although the majority of DUF families remain unknown, some have been studied. In *Oryza sativa*, *OsDUF1618* (*Wang et al., 2014*), *OsDUF221* (*Ganie, Pani & Mondal, 2017*), *OsDUF1110* (*Harada et al., 2016*), *OsDUF810* (*Li et al., 2018*), OsDUF668 (*Zhong et al., 2019a*), *OsDUF231* (*Zhong, Cui & Ye, 2019b*), *OsDUS936* (*Li et al., 2017*) have been characterized. Previous studies showed that DUF genes were engaged in different

Submitted 13 July 2023
Accepted 3 September 2023
Published 29 September 2023

Corresponding authors
Limei Kui, klm@yaas.org.cn
Xiaolin Li, xiaolinli@163.com

Academic editor
Elsayed Mansour

Additional Information and Declarations can be found on page 21

DOI **10.7717/peerj.16168**

biological functions in *Oryza sativa*. For instance, *SWOLLEN TAPETUM AND STERILITY 1* (*STS1*), which contained a DUF726 domain, interacted with *Polyketide Synthase 2* (*OsPKS2*) and *Acyl-CoA Synthetase 12* (*OsACOS12*) to contribute to sporopollenin biosynthesis (*Yuan et al., 2022*). Another DUF726 protein, encoded by *Leaked and Delayed Degraded Tapetum 1* (*OsLDDT1*), was involved in fatty acid synthesis and anther epidermis formation (*Sun et al., 2023*). *ROLLED and ERECT LEAF 2* (REL2), which contained DUF630 and DUF632 conserved domains, was involved in regulating leaf morphology, its functional loss resulted in rolling leaves (*Yang et al., 2016*). DUF genes have also been implicated in various biotic and abiotic responses. For example, the *Oryza sativa Stress Responsive DUF740 Protein* (*OsSRDP*) gene belonged to the DUF740 family, and its overexpressed transgenic plants, driven by promoter AtRd29A, revealed increased resistance to drought, salinity, and cold stresses as well as rice blast fungus (*Jayaraman et al., 2022*). The *DUF966-stress repressive gene 2* (*OsDSR2*) gene was involved in the negative regulation of salt and drought stress responses (*Luo et al., 2014*).

The DUF506 family, also called the PDDEXK_6 family, is a group of plant proteins that are distant homologs of the PD-(D/E)XK nuclease superfamily. The nuclear structure is retained as α-β-β-α-β and includes the typical PDDEXK motifs II and III in modified forms as xDxxx motif located in the second core beta-strand, where x is any hydrophobic residue, and a (D/E)X(D/N/S/C/G) pattern. The missing positively charged residue of motif III may be replaced by a conserved arginine in motif IV, which is located in the proceeding alpha-helix (*Knizewski et al., 2007*). So far, DUF506 proteins in *Oryza sativa* have not been characterized systemically and functionally. Previous research merely identified that the expressions of 13 *AtDUF506s* were ubiquitous in organs and associated with abiotic stresses and ABA response (*Ying, 2021*). The comparative microarray data showed that *AT2G20670* was inhibited by *B. cinereal*, heat, salinity, and osmotic stress (*Sham et al., 2019*). Recent studies revealed that *REPRESSOR OF EXCESSIVE ROOT HAIR ELONGATION 1* (*AtRXR1*) gene encoded AT3G25240 protein, which was strongly induced by phosphorus limitation and suppressed root hairs (RHs) extension by interacting with RabD2c GTPase. Moreover, its function under phosphorus limitation was conserved in both monocot and dicot, as proven by the analogous function of *Brachypodium distachyon DUF506* (*Ying et al., 2022*). *AtRXR3* (*AT1G62420*), another P-inducible *AtDUF506* gene, inhibited RHs elongation by a distinct mechanism. *AtRXR3* was transactivated by *ROOT HAIR DEFECTIVE6-LIKE4* (*RSL4*) and interacted with cytosolic calmodulins to repress RHs elongation (*Ying & Scheible, 2022*). Current studies on DUF506s in Arabidopsis suggested that DUF506 family members were significant in plant growth and abiotic resistance, but DUF506s functions in *Oryza sativa* have rarely been investigated. *LOC_Os01g68650*, the closest homologous gene of *AtRXR1*, was upregulated under drought stress, and its expression in *drought tolerance 11 (OsDT11*) overexpression line was higher than in the wild line, indicating that this gene may participate in drought response and be enhanced by *OsDT11* (*Zhao et al., 2020*). *LOC_Os01g54340* was identified as a nitrogen-sensitive gene that was rapidly repressed by nitrogen starvation (*Hsieh et al., 2018*). Until now, the *OsDUF506* family has not been genome-wide identified and the functions are still unknown.

In this study, we identified all *OsDUF506* family members in *Oryza sativa* by bioinformatic methods. The phylogeny, conserved motifs, cis-acting regulatory elements, distribution of non-synonymous SNPs, target miRNA, synteny, and tissue expression specificity were analyzed. The expression pattern of *OsDUF506s* under plant hormones treatments and drought, cold, and phosphorus-deficient stresses was investigated using transcriptome and qRT-PCR. Our study provided a more comprehensive identification and classification of *OsDUF506s*, broadened our recognition of the functions under abiotic stresses, and laid the groundwork for molecular breeding in *Oryza sativa*.

MATERIALS AND METHODS

Identification of *DUF506* members in 10 plant species

All the genome databases were downloaded from the EnsemblPlants database (http://plants.ensembl.org). Thirteen Arabidopsis DUF506 protein sequences were obtained from UniProtKB/Swiss-Prot (SwissProt) database (https://www.uniprot.org) and used as queries to perform protein blast search in 10 plant species (*Oryza sativa*, *Arabidopsis thaliana*, *Glycine max*, *Solanum tuberosum*, *Gossypium raimondii*, *Zea mays*, *Hordeum vulgare*, *Triticum aestivum*, *Sorghum bicolor* and *Ananas comosus*) by using TBtools (*Chen et al., 2020*). Meantime, the typical domain of DUF506 (PF04720, PDDEXK_6) was downloaded from the PFAM database (http://pfam.xfam.org) and was used to search for DUF506 with the HMMER tool (https://www.ebi.ac.uk/Tools/hmmer/search/hmmscan). All the candidates were merged and edited to eliminate redundancies. The candidates were then submitted to NCBI-CDD (https://www.ncbi.nlm.nih.gov/cdd/) to test for the existence of the complete DUF506 conserved domain and those with an incomplete N end or C end were eliminated. The molecular weight (Mw), isoelectric point (pI), instability index, aliphatic index, and grand average of hydropathicity (GRAVY) of DUF506 members were predicted with ExPASy (http://web.expasy.org/protparam). The subcellular localizations were predicted with WoLF PSORT (https://wolfpsort.hgc.jp).

Phylogenetic relationship, structure, and conserved motifs analysis of *OsDUF506* members

A total of 130 *DUF506s* was used for aligning with the MUSCLE method (default parameters) and construction of an ML phylogenetic tree (default parameters) using MEGA 11 software by setting bootstrap to 1,000 and JTT+G model (*Tamura, Stecher & Kumar, 2021*). The result was displayed by ChiPlot (https://www.chiplot.online). The conserved motifs were predicted with MEME (https://meme-suite.org/meme/doc/meme.html) and visualized by TBtools (*Chen et al., 2020*).

Prediction of CREs of *OsDUF506* members

The 2,000 bp upstream sequences of promoters were used to search for CREs with PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html) and visualized by TBtools (*Chen et al., 2020*).

Analysis of SNPs of *OsDUF506* members

The *OsDUF506* sequences were used to query the SNP-Seek database for nsSNPs against Nipponbare reference (https://snp-seek.irri.org). SNP-index was used to assess the subpopulation specificity of nsSNPs. The SNP-index was calculated as SNP_{GJ}-index = $N_{ref}/N_{all} \times 100\%$, SNP_H-index = $N_h/N_{all} \times 100\%$, SNP_{XI}-index = 1-SNP_{GJ}-index-SNP_H-index, N_{ref} represented the number of varieties sharing the same allele with reference, N_h represented the number of varieties with heterozygous allele, N_{all} represented the total number of varieties with determined alleles at the SNP locus.

Syntenic analysis of *OsDUF506* members

The duplication events and syntenic relationship of *DUF506s* between *Oryza sativa* and other plants were obtained by MCScanX (*Wang et al., 2012*). The results were visualized by TBtools (*Chen et al., 2020*). The nonsynonymous (Ka) and synonymous (Ks) calculations were performed by the simple Ka/Ks calculator kit of TBtools (*Chen et al., 2020*).

Predict analysis of miRNAs interacting with *OsDUF506* members

The miRNAs targeting *OsDUF506s* were predicted by psRNATarget (https://www.zhaolab.org/psRNATarget/analysis?function=2) and visualized by ChiPlot (https://www.chiplot.online). The expressions of the predicted miRNAs were obtained from the PmiRExAt database (http://pmirexat.nabi.res.in/searchdb.html) and the normalized TPM values (log2) were used for constructing a heatmap, the scale method of normalized was used to intuitively reflect the expressing differences between tissues and treatments by TBtools (*Chen et al., 2020*).

Expression analysis of *OsDUF506* members in different tissues and induced by plant hormones

The expression data in various tissues and induced by 50 μM abscisic acid (ABA), 10 μM gibberellin 3 (GA_3), 10 μM auxin (IAA), 100 μM jasmonic acid (JA), 1 μM brassinolide (BL), and 1μM trans-Zeatin (tZ) were obtained from RiceXPro database (https://ricexpro.dna.affrc.go.jp/quick-guide.html), the normalized signal intensity values (log_2) were used for constructing heatmap, and the scale method of normalized was used to intuitively reflect the expressing changes of particular genes at different treating time points by TBtools (*Chen et al., 2020*).

Plant growth conditions and abiotic stresses treatments

Japonica rice variety Yunkegeng 5 was used in the expression analysis. The plants were cultivated in a climate chamber (160 μmol^{-2} s^{-1} light intensity, 14 h-light and 10 h-dark a cycle, 28 °C, 45% RH) for 14 days in Yoshida rice nutrient solution (NS1040; Coolaber Technology Co., Ltd., Beijing, China). For drought stress, plants were transferred to a nutrient solution of 20% (w/v) polyethylene glycol (PEG-6000) for 3 h. For cold stress, plants were transferred to the climate chamber under 4 °C treatment for 3 h.

For phosphorus deficiency stress, plants were transferred to phosphorus-deficient Yoshida nutrient solution (NSP1040-P; Coolaber Technology Co., Ltd., Beijing, China) for 7 days.

Three biological replicates of the treated and control plants were harvested and stored at −80 °C.

Expression analysis of *OsDUF506* members by qRT-PCR

Primer design and specificity check was performed by Primer-BLAST of NCBI (https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi). The total RNA was extracted by TaKaRa MiniBEST Plant RNA Extraction Kit, cDNA was synthesized using Vazyme HiSript III 1st Strand cDNA Synthesis Kit (+gDNA wiper). The qRT-PCR was accomplished by Vazyme ChamQ SYBR Color qPCR Master Mix (Without ROX) in LightCycler96 system under the PCR condition of 95 °C for 60 s, 45 cycles of 95 °C for 10 s, 54 to 60 °C for 20 s and 72 °C for 20 s. The relative expressions data were calculated by $2^{-\Delta\Delta CT}$ method, and the reference used in this study was *OsActin*. The primer sequences were listed in Table S9. The significant differences level was analyzed by unpaired t-test with GraphPad Prism 8.0 software.

RESULTS

Identification and phylogenetic relationship of *DUF506* members in *Oryza sativa* and nine additional plant species

Using the previous method, we identified 10 *DUF506*s in *Oryza sativa* and named them *OsDUF50601* to *OsDUF50610*. Moreover, 13, 25, 13, 20, 8, 4, 22, 8, 7 *DUF506s* were identified in *Arabidopsis thaliana*, *Glycine max*, *Solanum tuberosum*, *Gossypium raimondii*, *Zea mays*, *Hordeum vulgare*, *Triticum aestivum*, *Sorghum bicolor* and *Ananas comosus*, respectively (Table S1). A total of 130 *DUF506s* were used to construct an ML phylogenetic tree and divided into four subfamilies based on the evolutionary distance referred to previous literature (*Ying, 2021*). The IIIb subfamily contained the most *DUF506s*, while the IIIa subfamily had the fewest. Three members belonged to subfamily I (*OsDUF50602*, *OsDUF50609*, and *OsDUF50610*), two belonged to subfamily II (*OsDUF50601* and *OsDUF50606*), only *OsDUF50603* belonged to subfamily IIIa, while the rest four members composed the largest subfamily IIIb (Fig. 1).

Characterization and conserved motifs of Os*DUF506* members

The 10 *OsDUF506s* were located on chromosome 1, 3, 5, 7, 10, 11. The summary of characteristics was shown in Table 1. Their protein lengths ranged from 269 to 507. The molecular weight, theoretical isoelectric points, and aliphatic index were predicted between 29,861.74 to 54,428.55 Da, 5.58 to 9.12, and 62.39 to 86.85, respectively. The instability index of proteins exceeded 40, and their GRAVY values were negative, suggesting that they were unstable hydrophilic proteins. The subcellular localization of subfamily II members (*OsDUF50601* and *OsDUF50606*) was predicted to be in the nucleus, while the rest were predicted to be in the chloroplast. *OsDUF506s* possessed 1 to 3 exons, and the exon/intron structure of genes within a subfamily was comparable (Fig. 2A). The *OsDUF506s* from subfamily I had only one CDS region, while those from subfamilies II and IIIa had three CDS regions. The number of CDS regions led to the division of

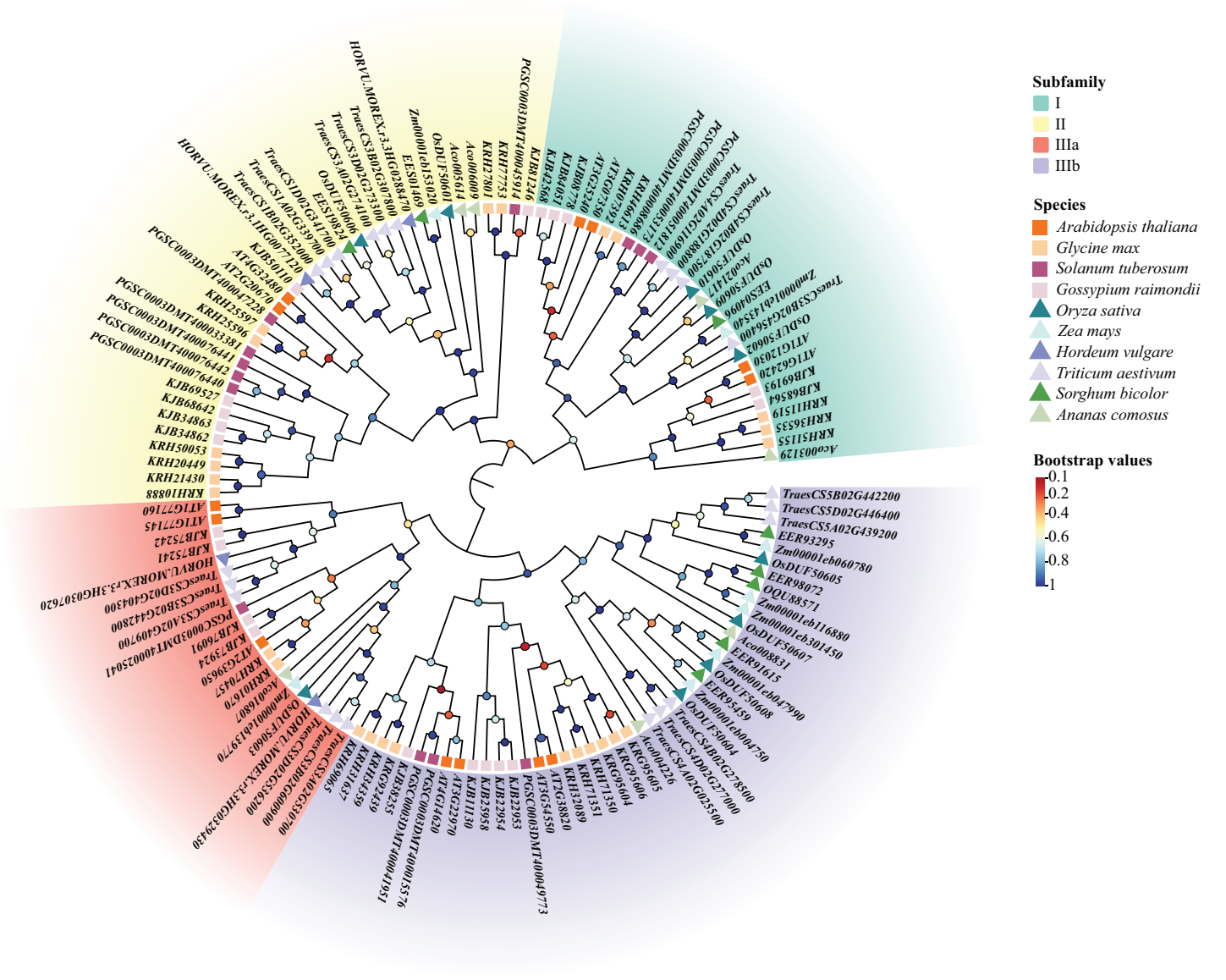

Figure 1 Phylogenetic tree of *DUF506* members identified in ten plant species. The members were divided into four subfamilies shown in different colors. Circles in different colors represent different bootstrap values. Bootstrap values were calculated from 1,000 replications and only the values over 0.7 bootstrapping were considered significant. Full-size DOI: 10.7717/peerj.16168/fig-1

subfamily IIIb into two branches. The result suggested that members of different subfamilies might have distinct functions.

Fifteen conserved protein motifs of *OsUDF506s* were identified (Fig. 2B). All *Oryza sativa* and Arabidopsis *DUF506* members shared motifs 1, 2, 3, and 5. In subfamily I, the most conserved motif was 4 and 11. Ultimately, *OsDUF50602* and *OsDUF50610* shared the same motifs with *AT1G62420* and *AT3G25240*, indicating they may have the same biological function. The subfamily II members shared motifs 8 and 11. The motif construction of *OsDUF50602* from subfamily IIIa was identical to that of *AT2G39650*. All subfamily IIIb members shared motifs 4, 6, and 7. Three-quarters possessed motif 9,

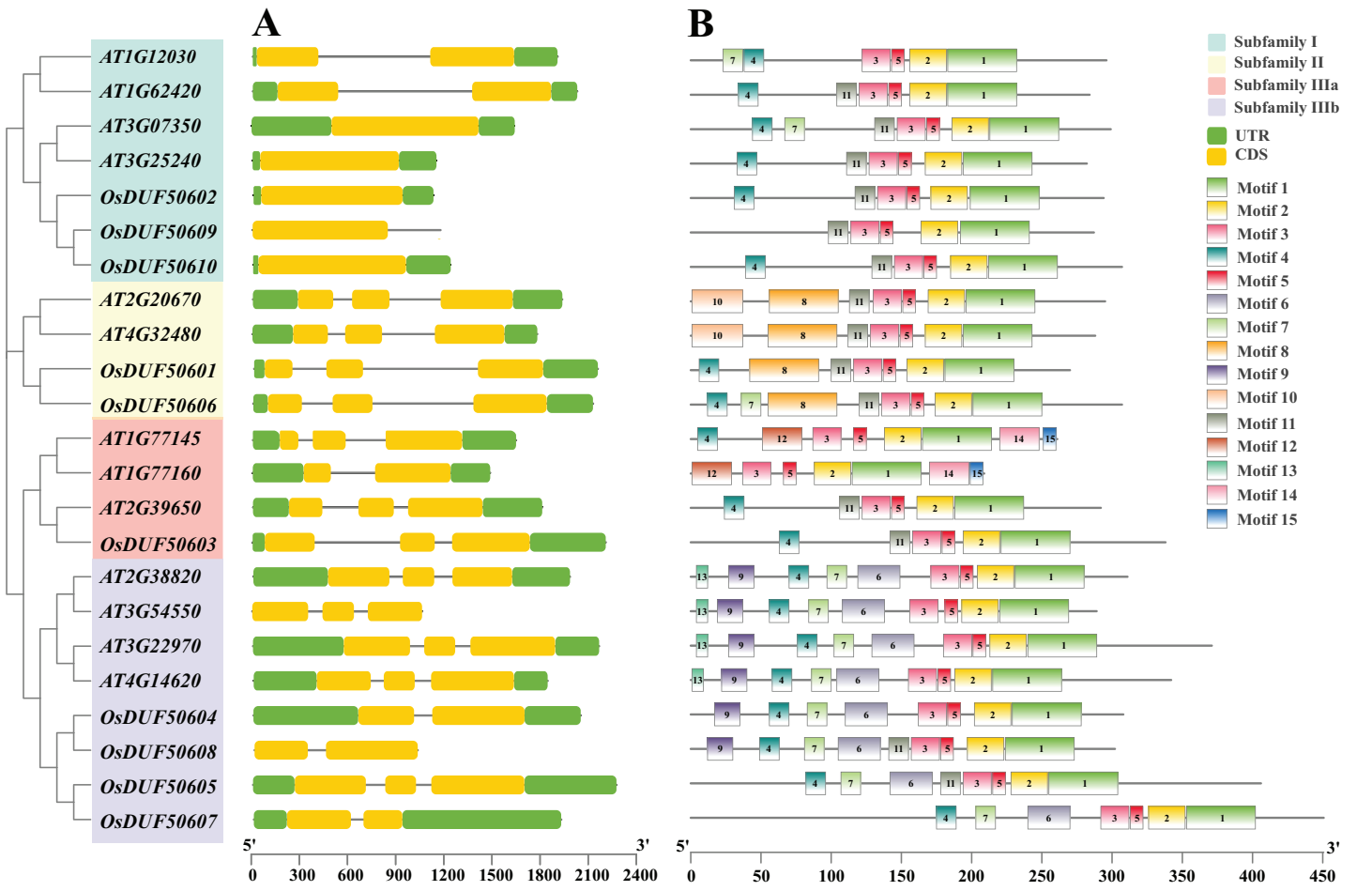

Figure 2 Gene structure, and conserved motifs of *DUF506* in *Oryza sativa* and *Arabidopsis thaliana*. Four background colors indicate four subfamilies. (A) Exon/intron structures. Green bars, yellow bars, and lines indicate UTRs, exons, and introns, respectively. (B) Distributions of conserved motifs. Motifs are shown by 15 different color bars. Full-size DOI: 10.7717/peerj.16168/fig-2

but motif 13 was exclusive to genes from Arabidopsis. The result showed that *DUF506s* belonging to the same subfamily shared similar motif characteristics, indicating that they have a similar function.

Analysis of cis-acting regulatory elements (CREs) of *OsDUF506* members

To analyze the prospective function of *OsDUF506s*, we searched the promoter regions (2,000 bp upstream of the start codon) and predicted 321 potential CREs (Fig. 3). *OsDUF50604* possessed the most CREs, while *OsDUF50603* and *OsDUF50609* possessed the fewest (Table S2). Nineteen types of CREs were identified and grouped into three functional categories: hormone response, abiotic stress response, and plant growth and metabolism. Hormone response-related CREs included MeJA-response elements (TGACG-motif and CGTCA-motif), abscisic acid response elements (ABRE), auxin response elements (TGA and AuxRR-core), salicylic acid response elements (TCA) and gibberellin response elements (GARE-motif and P-box). Each member of *OsDUF506s*

Table 1 Characteristics of *DUF506* genes in *Oryza sativa*.

Gene ID	Subfamily	Locus ID	Size (aa)	Protein MW (Da)	pI	Instability index	Aliphatic index	GRAVY	Chromosome	Subcellular localization
OsDUF50601	II	LOC_Os01g54340	269	29,861.74	7.61	47.02	65.76	−0.460	1	Nucleus
OsDUF50602	I	LOC_Os01g68650	293	30,429.34	8.97	52.76	75.22	−0.134	1	Chloroplast
OsDUF50603	IIIa	LOC_Os01g74250	337	37,616.96	7.25	60.92	86.85	−0.325	1	Chloroplast
OsDUF50604	IIIb	LOC_Os03g06680	307	33,011.21	6.68	45.12	80.59	−0.416	3	Chloroplast
OsDUF50605	IIIb	LOC_Os03g58230	405	43,737.47	9.12	50.54	68.52	−0.489	3	Chloroplast
OsDUF50606	II	LOC_Os05g44300	306	32,169.30	6.76	56.42	62.39	−0.267	5	Nucleus
OsDUF50607	IIIb	LOC_Os07g08390	507	54,428.55	9.07	59.08	66.07	−0.481	7	Chloroplast
OsDUF50608	IIIb	LOC_Os10g28210	301	32,041.06	6.52	50.73	79.60	−0.347	10	Chloroplast
OsDUF50609	I	LOC_Os11g25020	286	30,211.19	9.03	42.62	76.71	−0.108	11	Chloroplast
OsDUF50610	I	LOC_Os11g25040	306	32,464.54	5.58	50.32	82.06	−0.176	11	Chloroplast

contained at least two types of hormone response elements, MeJA-response elements and abscisic acid response elements. Regarding abiotic stress response, light response elements, anaerobic induction elements, and anoxic specific inducible elements were the most prevalent. In subfamily II and IIIb, there were low-temperature response (LTR) elements. Drought inducible elements (MSB) existed in *OsDUF50601*, *OsDUF50602*, *OsDUF50607*, and *OsDUF50609*, indicating that they may be associated with drought stress. Fewer CREs were related to plant growth and metabolism. CAT-boxes relating to meristem expression were found in half of *OsDUF506s*, while Motif I in *OsDUF50601*, *OsDUF50605*, and *OsDUF50606* was involved in root specificity. Only a few *OsDUF506* members possessed CREs associated with circadian control, cell cycle regulation, zein metabolism regulation, endosperm expression, and flavonoid biosynthetic genes regulation. These results suggested that *OsDUF506s* might be essential in hormone regulation and abiotic stress response.

Analysis of SNPs in *OsDUF506* members

According to the Rice SNP-Seek Database, 78 SNPs were identified in *OsDUF506s*, of which 30 were non-synonymous SNP (nsSNP) (Table 2). *OsDUF50609* and *OsDUF50610* possessed 13 and six nsSNPs, respectively, while the other members possessed between one to three nsSNP, indicating they were relatively conserved. To explore the subpopulation-specific variants, 30 nsSNP genotyping data of 2,644 *Oryza sativa* varieties from nine subpopulations (five *Xian/indicia (XI)* subpopulations and four *Geng/japonica (GJ)* subpopulations) were analyzed. SNP_{GJ}-index represented the proportion of varieties that shared the same allele as the reference Nipponbare. Three nsSNPs (OsDUF50609.1, OsDUF50609.7, and OsDUF50610.6) from subfamily I and one nsSNP (OsDUF50607.1) from subfamily IIIb were specific between Indica and Japonica varieties, because the SNP_{GJ}-index values in Indica subpopulations were less than 5%, while the values in the Japonica subpopulations were greater than 85% (Fig. 4 and Table S3).

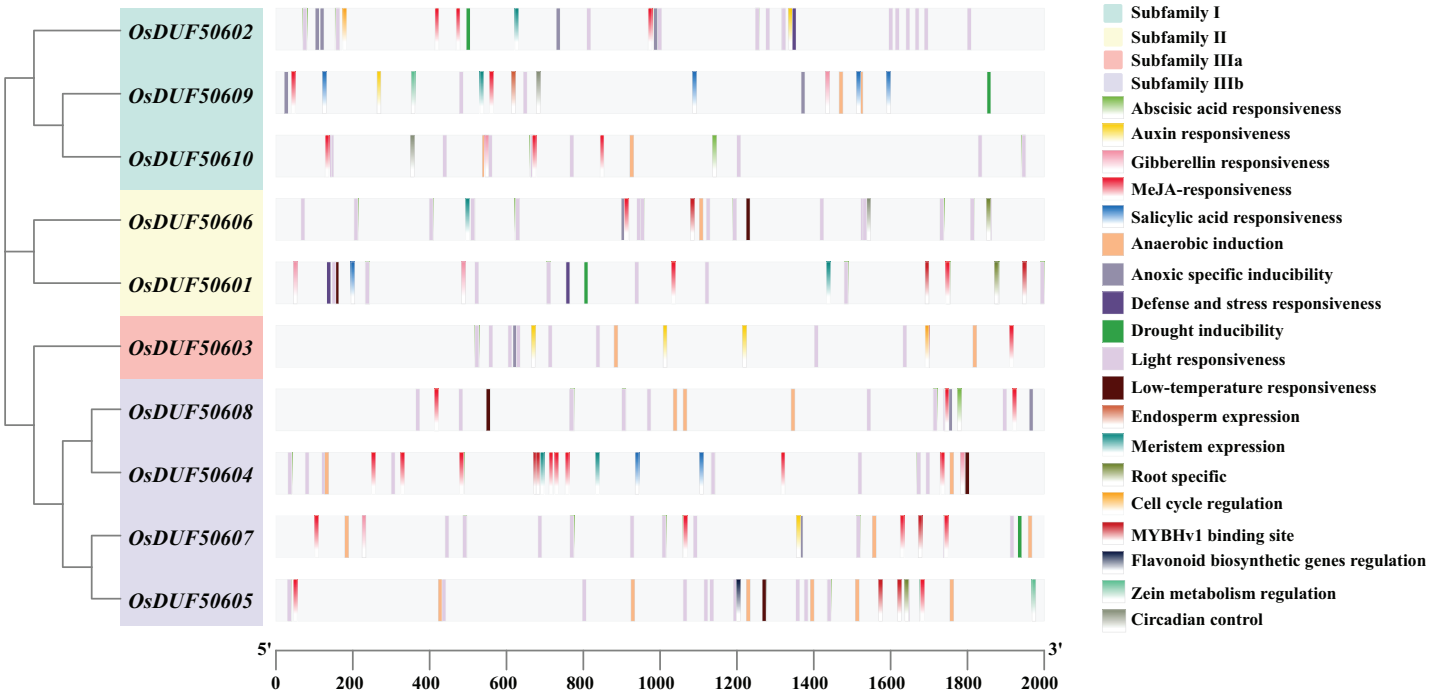

Figure 3 Distribution of CREs in *OsDUF506s* promoters. Predicted CREs with different functions are displayed with boxes in different colors. The light gray bars represent the 2,000 bp sequence upstream of *OsDUF506s*. Full-size DOI: 10.7717/peerj.16168/fig-3

Synteny analysis of *OsDUF506* members

Gene duplications are significant for gene family evolution. The synteny analysis result showed three segmental duplications (*OsDUF50601-OsDUF50606*, *OsDUF50604-OsDUF50608*, and *OsDUF50605-OsDUF50607*) and one tandem duplication (*OsDUF50609-OsDUF50610*) in *Oryza sativa* (Fig. 5), indicating that segmental duplication was the core expansion dynamic of *OsDUF506s* evolution. All duplications had Ka/Ks ratios below 0.5 (Table S4), demonstrating that *OsDUF506s* have undergone purifying selective pressure during evolution. In addition, these duplicated events were speculated to occurred at least 20.48 million years ago (Table S4).

Besides, duplicated events of *DUF506s* in *Oryza sativa*, the dicot model plant (*Arabidopsis thaliana*), and other monocotyledonous species (*Zea mays*, *Hordeum vulgare*, *Triticum aestivum*, *Sorghum bicolor*, and *Ananas comosus*) were also identified. *OsDUF50604*, *OsDUF50605*, *OsDUF50607*, and *OsDUF50608* in *Oryza sativa* and *AT3G22970* and *AT4G14620* in *Arabidopsis* from subfamily IIIb formed six homologous gene pairs and showed multiple collinearities (Fig. 5). In five monocotyledonous species, other than *OsDUF50609* and *OsDUF50610* on chromosome 11, the other *OsDUF506s* all had homologous genes. A total of 10 colinear gene pairs were found between *Oryza sativa* and *Hordeum vulgare*, 11 pairs with *Zea mays* and *Sorghum bicolor*, respectively, 12 pairs with *Ananas comosus*, and 30 pairs with *Triticum aestivum* (Fig. 6 and Table S5). The results suggested that these *DUF506s* might be derived from the same ancestral type

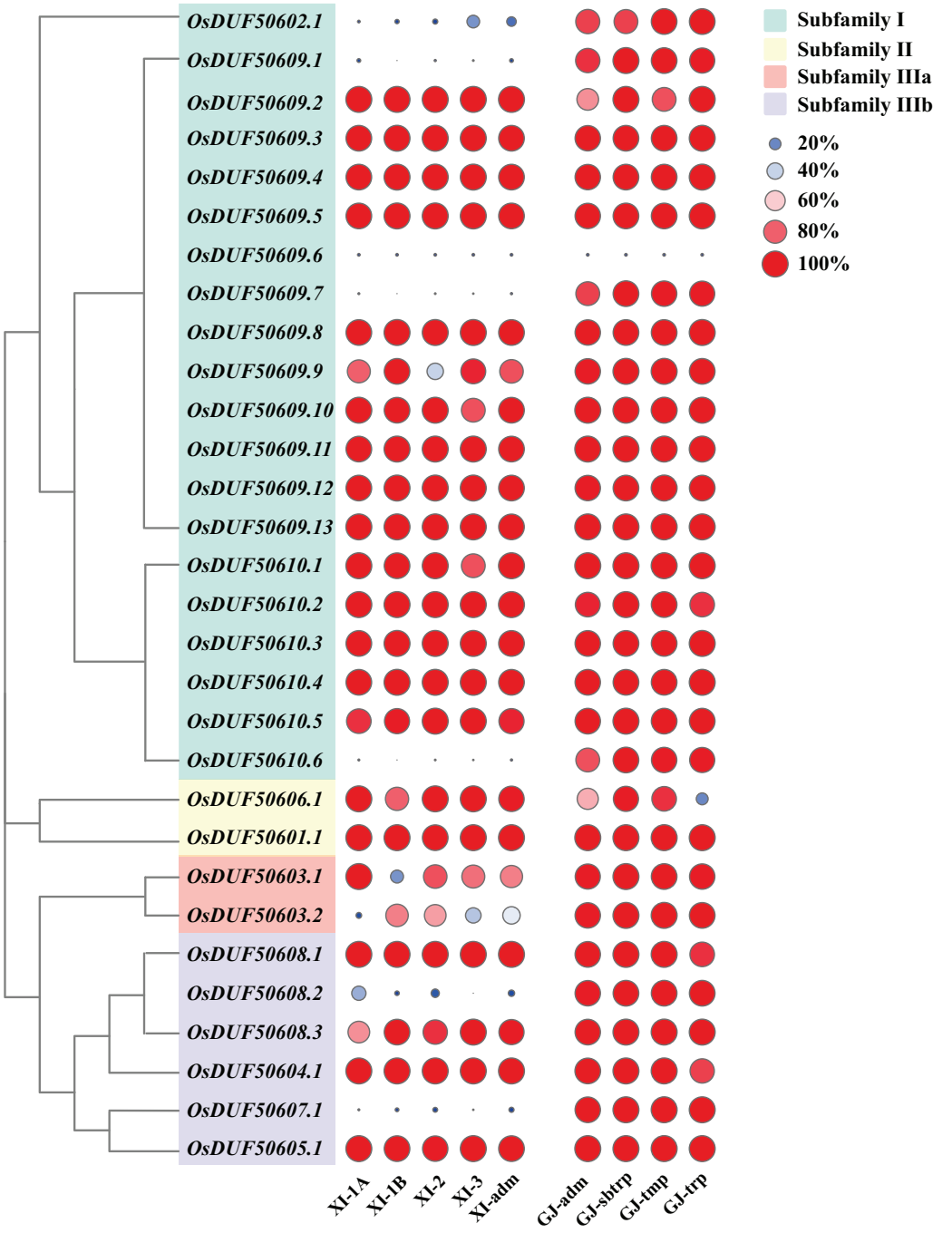

Figure 4 **SNP_{GJ}-index values of *OsDUF506s* nsSNPs in nine subpopulations.** Circles in different colors and sizes represent different SNP_{GJ}-index values. Full-size DOI: 10.7717/peerj.16168/fig-4

and function similarly. The *OsDUF506s* from subfamily IIIb were involved in gene pairs with *Arabidopsis thaliana and* those five monocots, respectively, indicating that they might be critical to the evolution of the DUF506 family.

Table 2 Non-synonymous SNP distribution of *OsDUF506s* in 3,024 *Oryza sativa* varieties.

Subfamily	Gene ID	No. total SNP	nsSNP ID	Ref allele	Alt allele	Position
I	*OSDUF50602*	2	OSDUF50602.1	C	T	Chr1:39867256
	OsDUF50609	23	OsDUF50609.1	G	T	Chr11:14248387
			OsDUF50609.2	C	A	Chr11:14248411
			OsDUF50609.3	C	T	Chr11:14248432
			OsDUF50609.4	A	G	Chr11:14248446
			OsDUF50609.5	C	T	Chr11:14248498
			OsDUF50609.6	A	G	Chr11:14248539
			OsDUF50609.7	C	G	Chr11:14248616
			OsDUF50609.8	C	G	Chr11:14248636
			OsDUF50609.9	C	T	Chr11:14248665
			OsDUF50609.10	G	T	Chr11:14248737
			OsDUF50609.11	C	T	Chr11:14248861
			OsDUF50609.12	A	C	Chr11:14248932
			OsDUF50609.13	T	C	Chr11:14249448
	OsDUF50610	9	OsDUF50610.1	G	T	Chr11:14267840
			OsDUF50610.2	C	T	Chr11:14267960
			OsDUF50610.3	G	A	Chr11:14268083
			OsDUF50610.4	C	G	Chr11:14268298
			OsDUF50610.5	C	T	Chr11:14268593
			OsDUF50610.6	T	C	Chr11:14268614
II	*OsDUF50606*	3	OsDUF50606.1	G	A	Chr5:25778321
	OSDUF50601	9	OSDUF50601.1	A	G	Chr1:31272839
IIIa	*OSDUF50603*	12	OSDUF50603.1	G	A	Chr1:43017098
			OSDUF50603.2	C	A	Chr1:43018592
IIIb	*OsDUF50608*	7	OsDUF50608.1	G	C	Chr10:14658953
			OsDUF50608.2	G	A	Chr10:14658996
			OsDUF50608.3	C	G	Chr10:14659216
	OsDUF50604	3	OsDUF50604.1	T	C	Chr3:3381185
	OsDUF50607	3	OsDUF50607.1	G	C	Chr7:4305870
	OsDUF50605	7	OsDUF50605.1	T	A	Chr3:33171316

Predicted miRNAs analysis

Sixty-nine predicted miRNAs that target *OsDUF506s* and might be involved in expression regulations were identified (Fig. 7A). All *OsDUF506s* were targeted by multiple miRNAs, suggesting these genes were strictly regulated by the combination of multiple miRNAs. Multiple *OsDUF506s* were targeted by *osa-miRNA2927*, *osa-miRNA5075*, *osa-miRNA1848*, *osa-miRNA2925*, and *osa-miRNA5809*, indicating that these miRNAs were essential to *OsDUF506s*. The lengths of matured miRNAs ranged between 19 and 24 nucleotides (Table S6). Most of the miRNAs inhibited *OsDUF506* expressions by cleavage, only *osa-miR2880*, *osa-miR5340*, *osa-miR2926*, *osa-miR156j-3p*, *osa-miR1875*, *osa-miR444b.1*, *osa-miR444c.1*, and *osa-miR5832* inhibited *OsDUF506* expressions by translation repression.

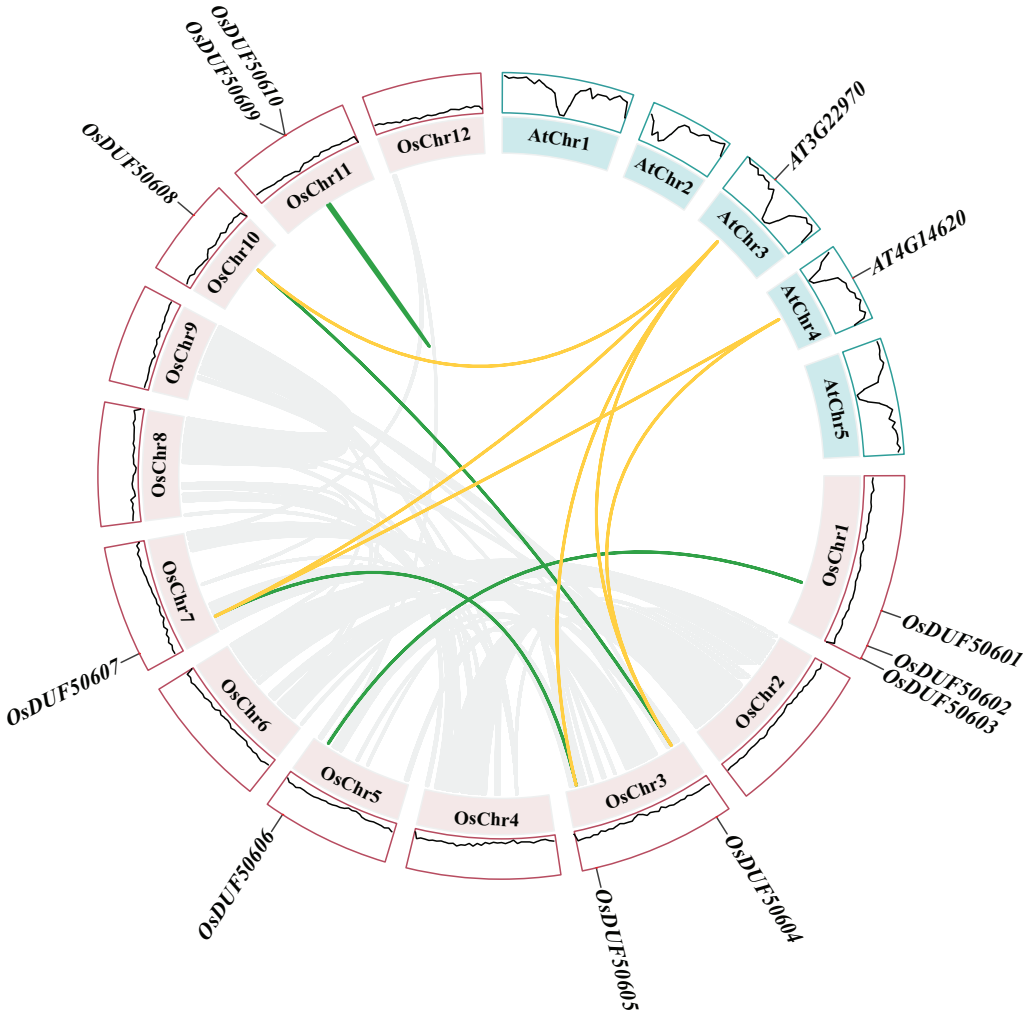

Figure 5 **Synteny analysis of *DUF506*s in *Oryza sativa* and *Arabidopsis*.** Pink and green bars represent the chromosomes of *Oryza sativa* and *Arabidopsis* respectively. The black lines in the colored box show the gene density of the chromosomes. The green lines suggest duplicated gene pairs in *Oryza sativa*, the yellow lines indicate the collinearity between *Oryza sativa* and *Arabidopsis thaliana*.
Full-size DOI: 10.7717/peerj.16168/fig-5

The expression analysis revealed that the majority of miRNAs were expressed at a modest level in *Oryza sativa* tissues (Fig. 7B). The osa-miR1874-3p, which targeted *OsDUF50605* specifically, showed the highest expression level in embryo. The osa-miR444b.1 targeting *OsDUF50606* was highly expressed in all tissues except anther, indicating that *OsDUF50606* expression was limited in these tissues, which may inhibit the function of *OsDUF50606* in plant growth. The expression levels of miRNAs under drought, salt, and cold stress showed no significant changes, suggesting they did not participate in abiotic stress responses.

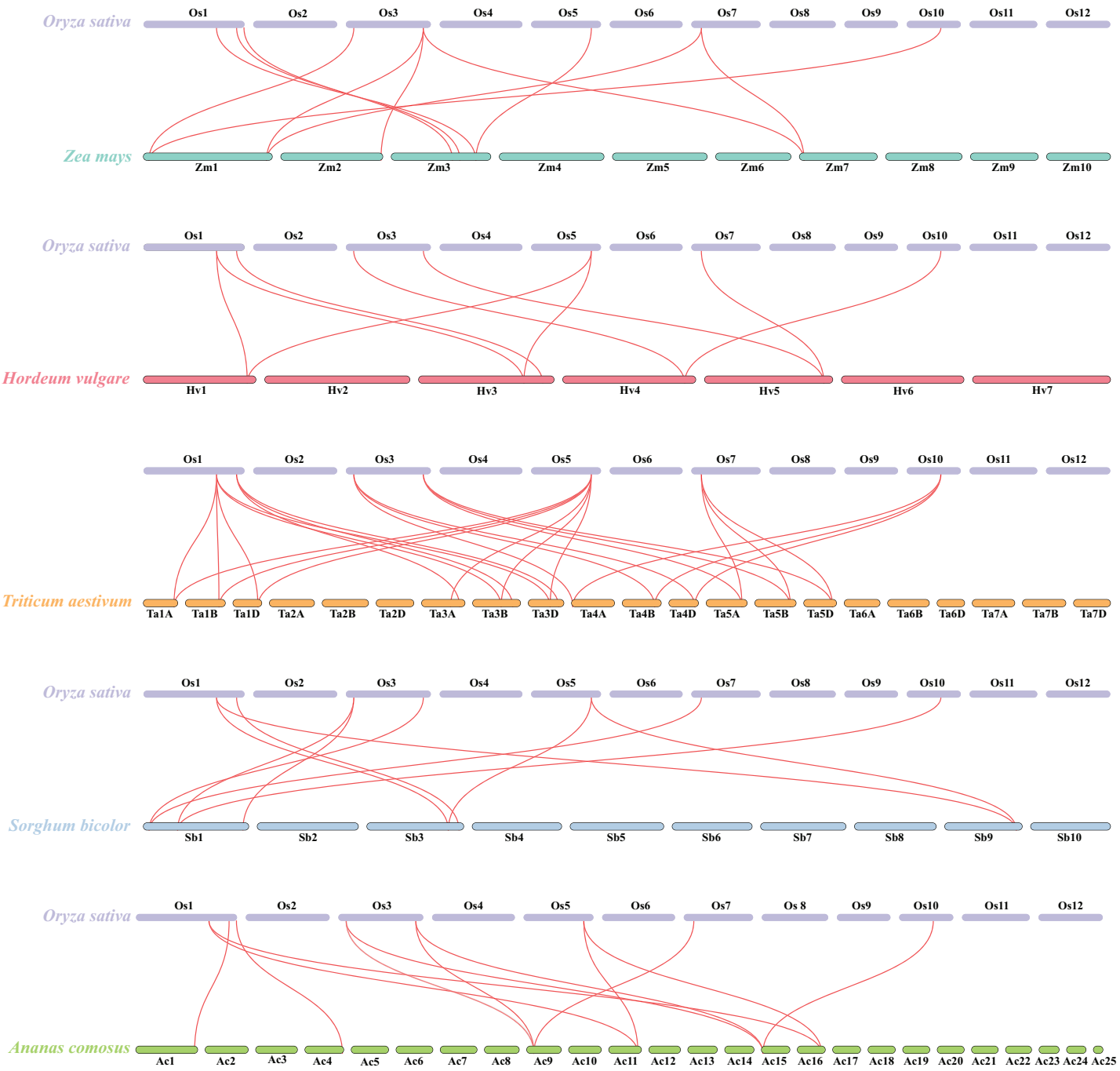

Figure 6 Synteny analysis of *DUF506s* between *Oryza sativa* and five monocot species. Red lines suggest syntenic gene pairs.
Full-size DOI: 10.7717/peerj.16168/fig-6

Expression analysis of *OsDUF506* members in different tissues and induced by plant hormones

To investigate the expression specificities of *OsDUF506s*, the expression values in the leaf blade, leaf sheath, root, stem, inflorescence, anther, pistil, lemma, palea, ovary, embryo, and endosperm were analyzed (Fig. 8; Table S7). *OsDUF50602* from subfamily I expressed

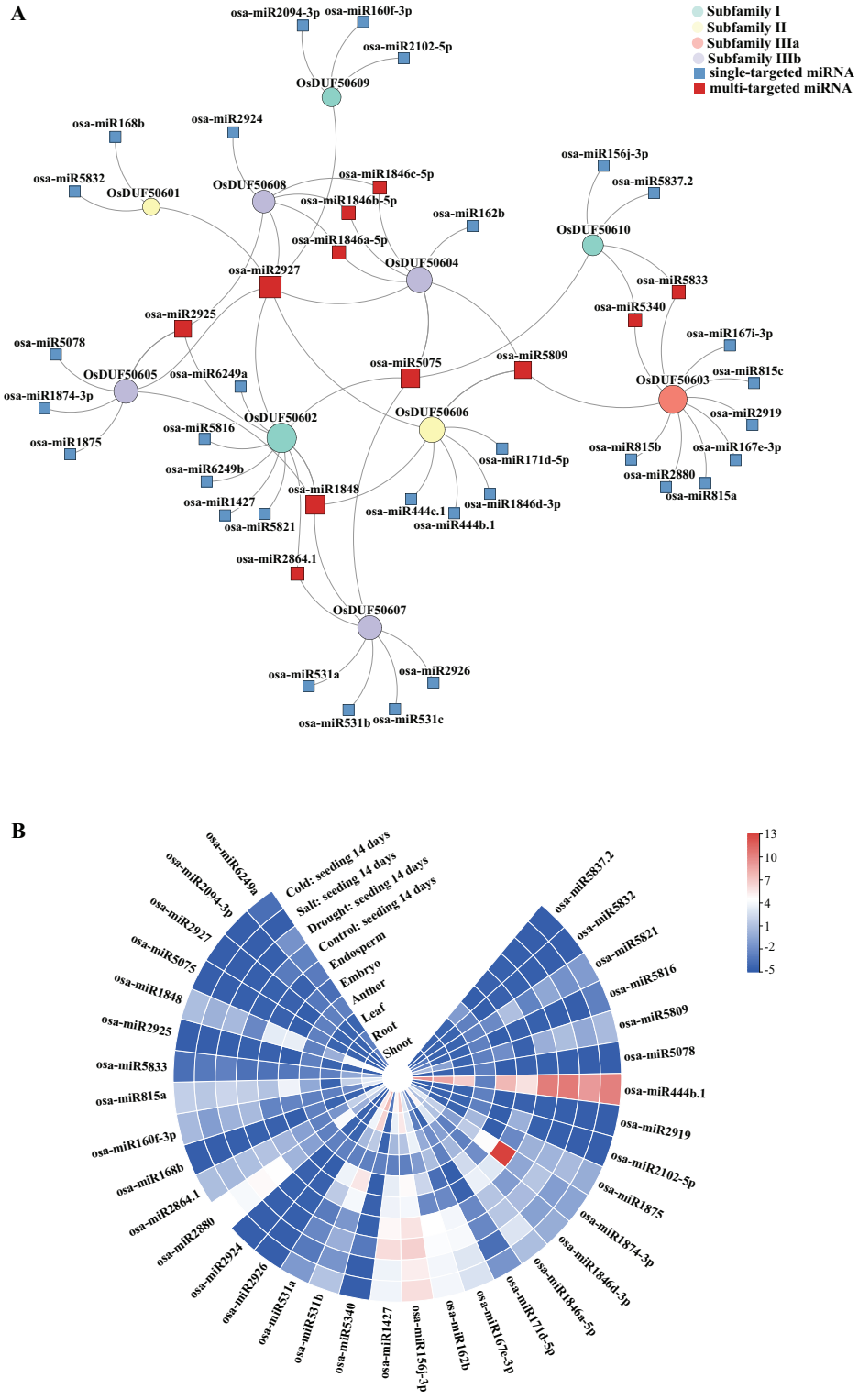

Figure 7 Analysis of predicted miRNA targeting *OsDUF506s*. (A) Identified miRNAs targeting *OsDUF506s*. Circles represent *OsDUF506s*, squares represent the related miRNAs. (B) Expressions of predicted miRNAs in different tissues and under abiotic stresses. The heatmap demonstrates the expression level, the color gradient from blue to red presents increasing expression values.

Full-size DOI: 10.7717/peerj.16168/fig-7

in all tissues, with the highest expressions in the embryo and endosperm 7 days after flowering, whereas *OsDUF506010* was expressed at an exceedingly low level in all tissues. The expression modes of the segment duplication gene pair *OsDUF50601-OsDUF50606* from subfamily II were wholly dissimilar, suggesting that they might be involved in functional redundancy. Compared to the genes from subfamily IIIa, the expression levels of four genes from subfamily IIIb were higher in most tissues. The segment duplication gene pair *OsDUF50604-OsDUF50608* was highly expressed in anther and leaf sheath, respectively. In contrast, another segment pair, *OsDUF50605-OsDUF50607*, was highly expressed in the leaf blade and leaf sheath, revealing that they might be involved in functional differentiation.

The CREs analysis of *OsDUF506s* suggested that they were widely involved in hormonal regulation, thus we analyzed their expression profiles in root and shoot treated with six plant hormones, including abscisic acid (ABA), gibberellic acid (GA_3), indole-3-acetic acid (IAA), jasmonic acid (JA), brassinolide (BL), and trans-zeatin (tZ). *OsDUF506s* showed distinct regulating modes (Fig. 9 and Table S8). Significant upregulation of *OsDUF50602* was induced by ABA and JA, whereas opposite regulation was induced by tZ in root and shoot. Subfamily II members, *OsDUF50601* and *OsDUF50606* exhibited identical regulation under ABA treatment. In the root, except *OsDUF50605* was upregulated by ABA induction, the other members from subfamily IIIb were downregulated by ABA, IAA, and JA. Under BL treatment, *OsDUF50604* and *OsDUF50608* were upregulated, whereas *OsDUF50605* and *OsDUF50607* were downregulated in root. Interestingly, *OsDUF50607* was significantly upregulated by ABA induction in shoot, while the opposite was observed in root. *OsDUF50605* also showed opposite effects on shoot and root regulations by ABA induction.

Expression analysis of *OsDUF506* members under drought stress, cold stress, and phosphorus-deficient stress

To explore the responses of *OsDUF506s* under abiotic stresses, the relative expressions of eight *OsDUF506s* under drought, cold, and phosphorus-deficient stresses were analyzed by qRT-PCR (Fig. 10). Under drought condition, the expressions of *OsDUF50601*, *OsDUF50603*, *OsDUF50604*, *OsDUF50607*, and *OsDUF50608* were significantly downregulated, whereas only *OsDUF50602* was significantly upregulated with a more than 2-fold increase. By contrast, *OsDUF506s* were more sensitive to cold stress. Under cold treatment, six *OsDUF506s* were upregulated with the exception of *OsDUF50606* and *OsDUF50607*, *OsDUF50601*, and *OsDUF50504*, which showed a 4-fold increase in gene expression. Under phosphorus-deficient condition within a week, *OsDUF50601*, *OsDUF50604*, and *OsDUF50607* were significantly downregulated, whereas only *OsDUF50605* was significantly upregulated but by less than 2-fold. The result demonstrated the majority of *OsDUF506s* were induced by drought and cold stresses, suggesting these genes might be implicated in these stress responses.

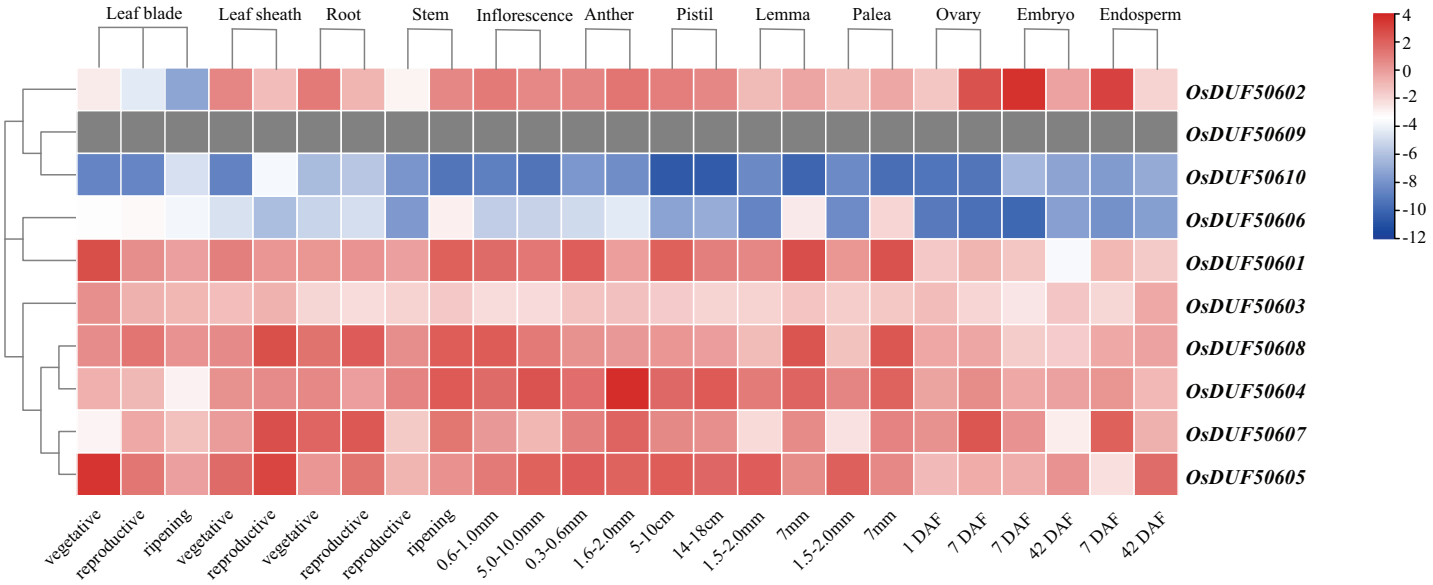

Figure 8 Expressions of *OsDUF506* members in different tissues. The heatmap demonstrates the expression level, the color gradient from blue to red presents increasing expression values. Grey presents missing data. Full-size DOI: 10.7717/peerj.16168/fig-8

DISCUSSION

DUFs are a specific type of genes with conserved domains but unknown functions. DUF506 family belongs to the PD-(D/E)XK nuclease superfamily, which consists of numerous enzymes involved in significant cellular processes (*Knizewski et al., 2007*). Most families are certified to be distinct restriction endonucleases with functions including repair of damaged DNA, resolution of holliday junctions, and excessive cleaving in DNA recombination (*Bujnicki, 2003*). The family retains the characteristic motif II of PD-(D/E)XK, but lacks a functionally-characterized domain (*Knizewski et al., 2007*). Until now, *DUF506* family research has only performed in *Arabidopsis* (*Ying, 2021*). In this study, we identified 10 *OsDUF506s* with an intact DUF506 domain in *Oryza sativa* and categorized them into four subfamilies based on the previous study in *Arabidopsis* (Table 1). Additionally, 120 *DUF506* genes from five other monocots and four dicots were also identified and their phylogenetic relationship with *OsDUF506s* was analyzed. In contrast to *DUF1618s*, which only existed in monocot, *DUF506s* were present in both monocots and dicots (Fig. 1), indicating that *DUF506* was an ancient gene family that originated prior to the dicotyledon-monocotyledon divergence (*Wang et al., 2014*). *DUF506* members from monocots or dicots preferred to congregate on the same branch (Fig. 1), indicating a substantial divergence in *DUF506*s between monocots and dicots. Moreover, genome size did not correlate with the number of *DUF506* genes. In dicots, for example, *Arabidopsis thaliana* and *Solanum tuberosum* possessed the same quantity of *DUF506* genes, but their genome sizes differed significantly.

Gene duplication is widespread in plants and contributes to genome expansion and the evolution of new functions. The majority of gene duplications consist of whole-genome duplications and tandem duplications (*Panchy, Lehti-Shiu & Shiu, 2016*). However, in

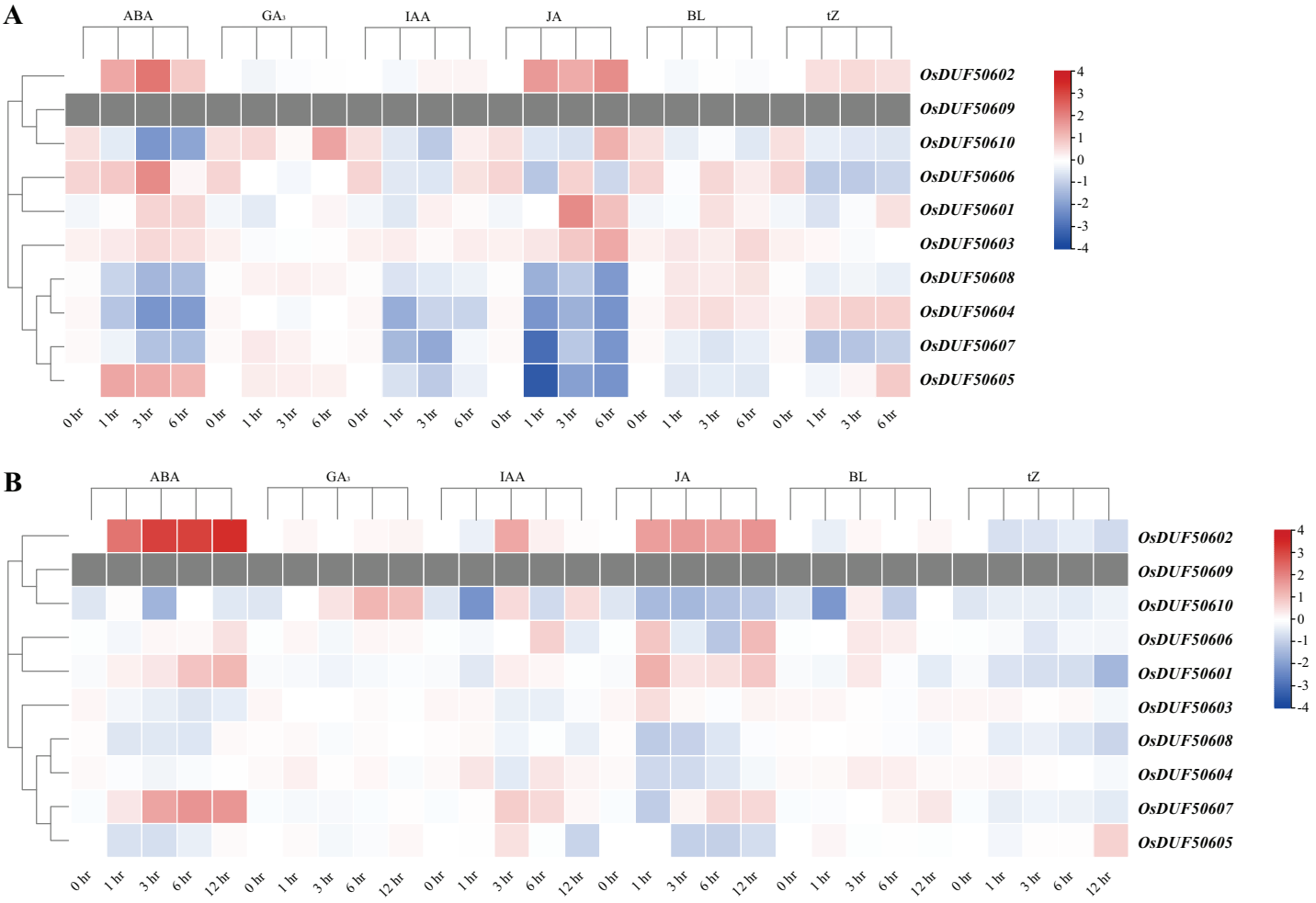

Figure 9 Expressions of *OsDUF506s* induced by plant hormones in root (A) and shoot (B). The heatmap demonstrates the expression levels, the color gradient from blue to red presents increasing expression values. Grey presents missing data. Full-size DOI: 10.7717/peerj.16168/fig-9

OsDUF506 duplication events, segmental duplication accounted for 75%, while tandem duplication only accounted for 25% (Table S4), suggesting that segmental duplication was critical for expanding *OsDUF506* family members. The dynamic processes between gene duplications and gene losses contribute to genomes differences (*Holland et al., 2017*). Research on *Oryza sativa* duplications revealed that 85% of duplicates underwent loss, subfunctionalization, or neofunctionalization during 50-70 million years of evolution (*Throude et al., 2009*). *OsDUF50602* and *OsDUF50603* did not form duplicate gene pairs (Fig. 5), indicating they could undergo gene losses. The mechanisms of duplicate gene loss include the deletion of duplicate sequence and pseudogenization, the latter of which refers to gene silence and genetic redundancy (*Ho-Huu et al., 2012*; *Lynch & Conery, 2000*; *Thibaud-Nissen, Ouyang & Buell, 2009*). The duplicated gene with low expression may experience pseudogenization, and pseudogenes typically originate from tandem duplications (*Yang et al., 2011*). This viewpoint is supported by the tandem duplicate pair *OsDUF50609-OsDUF50610*, which originated 58.69 million years ago. According to the

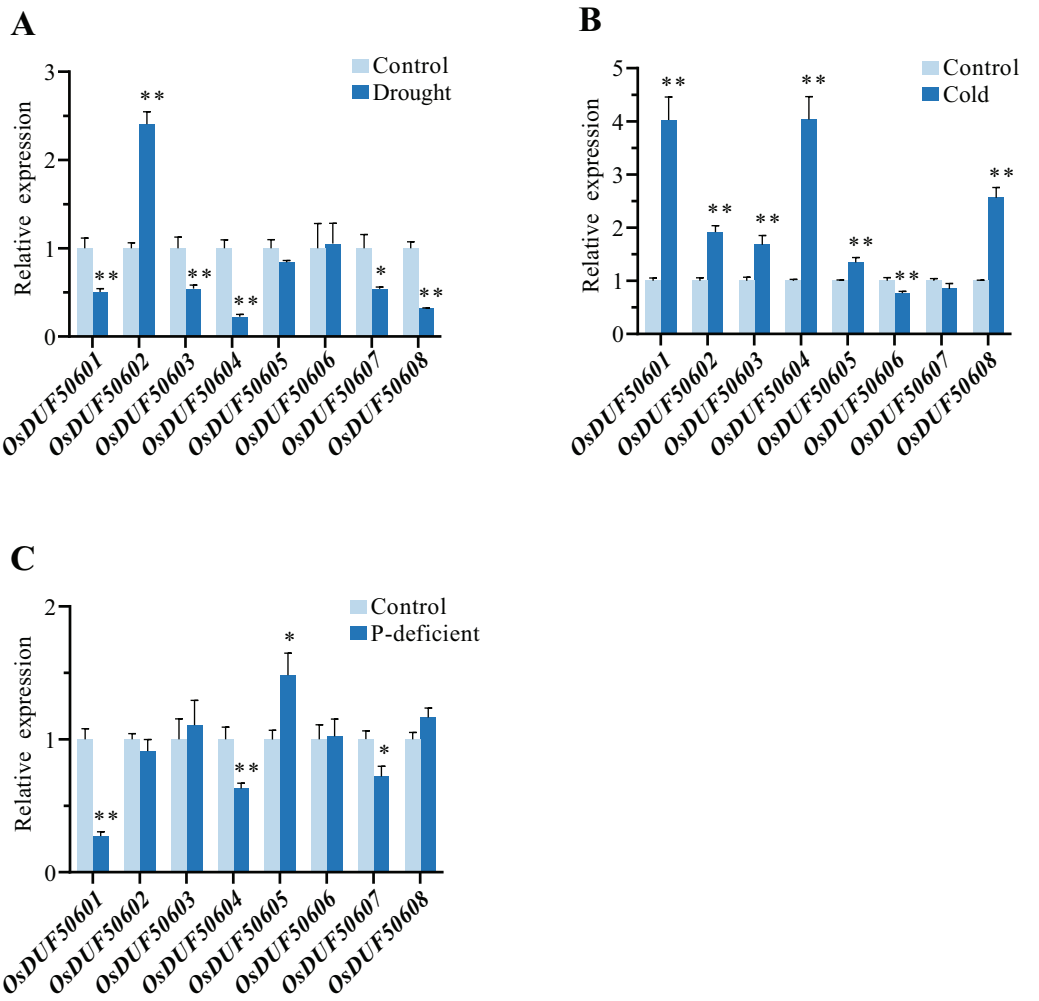

Figure 10 Relative gene expressions of *OsDUF506s* under drought (A), cold (B) and phosphorus-deficient (C) stresses. Data represent the mean ± SE of three biological and technical replicates. The significant differences level was analyzed by unpaired t-test ($^{*}p < 0.05$, $^{**}p < 0.01$).
Full-size DOI: 10.7717/peerj.16168/fig-10

gene expression profiles of the RiceXPro database (Fig. 8) and Rice Genome Annotation Project database (http://rice.uga.edu/expression.shtml), they were barely expressed in any tissue. Besides, compared to other members, *OsDUF50609* and *OsDUF50610* contained the most nsSNP mutations (Table 2), and excessive nonsynonymous nucleotide mutants are a characteristic of pseudogenes (*Balakirev & Ayala, 2003*). DUF family appears to be abundant in pseudogenes, such as DUF1311, DUF 1124, and DUF 3054 (*Thibaud-Nissen, Ouyang & Buell, 2009*). On the other side, *OsDUF50609* and *OsDUF50610* exhibited a general loss bias in Gramineae species, including *Zea mays*, *Hordeum vulgare*, *Triticum aestivum*, and *Sorghum bicolor* (Fig. 6). Except for the presumed pseudogenes, the other *OsDUF506*s all have colinear gene pairs with the other five monocots (Fig. 6), revealing that they were conserved in the evolution and expansion of the *DUF506* family in plants and they might have originated from the same ancestor. Previous studies suggested that more than half of duplicated genes have diverged in gene expression in both *Arabidopsis*

and *Oryza sativa* (*Blanc & Wolfe, 2004*; *Yim, Lee & Jang, 2009*). Two duplicated gene pairs from subfamily IIIb (*OsDUF50604-OsDUF50608* and *OsDUF50605-OsDUF50607*) revealed different tissue specificity (Fig. 8), and duplicates respond conversely under phosphorus-deficient conditions (Fig. 10), suggesting that they might be neofunctionalized.

MicroRNAs (miRNAs) are fundamental noncoding riboregulators for gene expression. In plants, miRNA silences genes by guiding RNA cleavage or translation inhibition (*Song et al., 2019*). They cooperate closely with target genes and transcription factors to regulate plant growth and resistance; it could be an effective strategy for precisely improving *Oryza sativa* varieties by derepressing specific genes using CRISPR/Cas9 (*Lin et al., 2021*; *Nadarajah & Kumar, 2019*). We predicted the miRNAs targeting *OsDUF506s* and analyzed their expression profiles in tissues and under stresses of drought, cold, and salt (Fig. 7). The osa-miR444b.1, which specifically targeted *OsDUF50606*, was the only one highly expressed in tissues excluding the anthers and under the four stresses. It was functionally unknown. One copy of the segmental duplicates from subfamily II (*OsDUF50606*) expressed extremely low in all tissues, it showed no significant changes in expression level under those abiotic stresses. In contrast, the other copy (*OsDUF50601*) was constitutively expressed and was strongly induced by drought, cold, and phosphorus-deficient stresses (Figs. 8 and 10). Further verification is required to ascertain if the expression difference originates from the translation inhibition by osa-miR444b.1.

The results of expression induced by plant hormones showed that *OsDUF506s* were more sensitive to ABA and JA treatments than IAA and GA_3 treatments, which was consistent with the presence of hormone-responsive CREs observed in their promoters (Figs. 3 and 9). Although the promoter region of some *OsDUF506s* contained a few GA_3 and IAA CREs, they show no significant and specific response trend under corresponding hormone treatment.

The CREs analysis also showed that *OsDUF50601*, *OsDUF50602*, *OsDUF50607*, and *OsDUF50609* each contained one MBS, which was a MYB transcription factor binding site responsive to drought. However, only *OsDUF50602* was upregulated under drought treatment (Figs. 3 and 10), indicating that their promotors might recruit different MYBs to active *OsDUF50602* expression and inhibit *OsDUF50601* and *OsDUF50607* expressions to regulate drought tolerance. Perhaps this is due to the vital function of hormones under the drought condition. ABA is an important plant hormone regulating water status and stomatal movement. When plants suffer from a drought environment, they synthesize ABA. Increasing ABA could induce plants to close their stomatal and retain water (*Lim et al., 2015*). Seven copies of ABRE(ABA-responsive element)existed in the promotor region of *OsDUF50602*. Its expression significantly increased in both shoot and root under ABA treatment, suggesting that it might also be involved in the ABA-related signaling pathway of drought responses (Fig. 9). Moreover, a previous study revealed that the expression of *OsDUF50602* was higher in *OsDT11* OE lines than in the wild line under drought treatment, indicating that the drought response of *OsDUF50602* might be enhanced in the particular genetic background (*Zhao et al., 2020*). ABRE has been proved to be the most conserved drought-inducible promoter in Arabidopsis, *Oryza sativa*, and

soybean, indicating that the transcriptional regulation of drought-inducible genes like *OsDUF50602* is similar across these species (*Maruyama et al., 2012*). Therefore, we analyzed the expressions of its homologous genes under drought stress in the eplant database (http://bar.utoronto.ca/). In *Triticum aestivum*, TraesCS3A02G420900 gene expression increased 3.34-fold under drought stress. On the other hand, JA and its derivatives, which occur at low levels under normal condition, accumulate to high levels and are transmitted over long distances under abiotic stress (*Wang et al., 2021a*). The promoter of *OsDUF50602* contained six copies of the CGTCA/TGACG-motif (JA-responsive element). With JA treatment, the expression of *OsDUF50602* was rapidly and significantly upregulated in both the shoot and root (Fig. 9). However, further verification is required to determine whether these two hormones collaboratively induce the drought responses of *OsDUF50602*.

The *OsDUF506s* were more sensitive to cold stress than to drought stress. Under cold stress, except for the slight downregulation of *OsDUF50606* and *OsDUF50607*, the other *OsDUF506s* were significantly upregulated, and three genes containing LTR elements (*OsDUF5060601*, *OsDUF5060604*, and *OsDUF50608*) showed the highest level (Fig. 10). *DUF506s* from other species also revealed an active response to cold stress. For instance, *At1g62420*, *At3g25240*, *Bradi2g58590*, and *Bradi2g62310* were strongly induced by cold stress in *Arabidopsis* and *Brachypodium* (*Ying, 2021*). The above results demonstrated that *DUF506* genes play a crucial role in cold response with different mechanisms.

Recent studies showed five *AtDUF506* genes belonging to subfamilies I and II (*At1g62420, At3g07350, At3g25240, At2g20670*, and *At4g32480*) were strongly upregulated by P-limitation (*Ying et al., 2022*; *Ying & Scheible, 2022*). However, although the *OsDUF506s*, which belong to subfamilies I and II, shared a highly similar exon/intron structure and conserved motifs (Fig. 3), only *OsDUF50601* was significantly downregulated by P-limitation, indicating that monocot and dicot species differed in phosphorus responding signal pathway.

Rice (*Oryza sativa*) is an essential staple food for approximately half of the world's population, and stable rice production is crucial for food security, especially in Asia (*Zhang, 2007*). However, extreme weather and climate change, such as drought, flooding, salinity, low temperature, high temperature, and mineral deficiency, seriously affect crop productivity and sustainability. Among the abiotic stress, drought is the most destructive threat. Half of the world's arable land will suffer from drought in the next three decades (*Singhal et al., 2016*). Hence, it is urgent to breed rice varieties with superior drought resistance. Drought resistance is a complex agronomic trait regulated by multiple genes. Exploring drought resistance genes through linkage analysis or GWAS is inefficient and genes with minor effects are hard to clone (*Sun et al., 2022*). With the development of bioinformatics and gene editing technology, the reverse genetics method can be combined to identify and characterize abiotic resistance-related genes such as *OsDUF506s*. Abiotic resistance improvements usually require increased expression of related genes, including *DROUGHT1* (*DROT1*), *ONAC066*, *OsMADS23*, *DROUGHT-INDUCED BRANCHED-CHAIN AMINO ACID AMINOTRANSFERASE* (*OsDIAT*), *CHILLING-TOLERANCE DIVERGENCE 1* (*COLD1*), and *Cyclic Nucleotide-gated Channels 9*

(*OsCNGC9*) (*Sun et al., 2022*; *Yuan et al., 2007*; *Li et al., 2021*; *Shim et al., 2023*; *Ma et al., 2015*; *Wang et al., 2021b*). A donor-DNA-free CRISPR/Cas-based approach to knock-up genes in rice could be helpful for rice breeding practices and simplify regulatory approval of the edited plants (*Lu et al., 2021*).

CONCLUSIONS

In this study, 10 *OsDUF506* family members in *Oryza sativa* were identified and classified into four subfamilies. We analyzed the phylogenetic relationship, gene structures, conserved motif, CREs, SNP distribution, and targeting miRNA, which filled the gap of *DUF506* family in rice. The results of public expression profiles and RT-qPCR data in tissues and under plant hormones and abiotic stresses demonstrated that *OsDUF506s* were actively involved in ABA and JA response and had different expression patterns under drought and cold, which laid the foundation for further functional analysis of *OsDUF506* family.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was supported by the Yunnan Province Major Science and Technology Project (No. 202102AE090016, No. 202302AE090001, No. 202302AE090003), and the Yunnan Province Scientific and Technological Talent and Platform Projects (No. 202105AE160009, No. 202305AF150195, Title. Colored Rice Varieties Breeding and Green Technologies Development Project). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
Yunnan Province Major Science and Technology: 202102AE090016, 202302AE090001, 202302AE090003.
Yunnan Province Scientific and Technological Talent and Platform: 202105AE160009, 202305AF150195.

Competing Interests

Kui Fan, Rui Wang & Jianping Zhang are employed by Yunnan Grain Industry Group Co., Ltd.

Author Contributions

- Wei Dong conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, and approved the final draft.
- Jian Tu performed the experiments, analyzed the data, prepared figures and/or tables, and approved the final draft.
- Wei Deng performed the experiments, prepared figures and/or tables, and approved the final draft.

- Jianhua Zhang performed the experiments, prepared figures and/or tables, and approved the final draft.
- Yuran Xu performed the experiments, prepared figures and/or tables, and approved the final draft.
- Anyu Gu analyzed the data, prepared figures and/or tables, and approved the final draft.
- Hua An analyzed the data, prepared figures and/or tables, and approved the final draft.
- Kui Fan analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Rui Wang performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Jianping Zhang analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Limei Kui conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Xiaolin Li conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

The raw data is available in the Supplemental File.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.16168#supplemental-information.

REFERENCES

Balakirev ES, Ayala FJ. 2003. Pseudogenes: are they "junk" or functional DNA? *Annual Review of Genetics* **37(1)**:123–151 DOI 10.1146/annurev.genet.37.040103.103949.

Bateman A, Coggill P, Finn RD. 2010. DUFs: families in search of function. *Acta Crystallographica Section F Structural Biology and Crystallization Communications* **66(10)**:1148–1152 DOI 10.1107/S1744309110001685.

Blanc G, Wolfe KH. 2004. Functional divergence of duplicated genes formed by polyploidy during Arabidopsis evolution. *Plant Cell* **16(7)**:1679–1691 DOI 10.1105/tpc.021410.

Bujnicki JM. 2003. Crystallographic and bioinformatic studies on restriction endonucleases: inference of evolutionary relationships in the "midnight zone" of homology. *Current Protein & Peptide Science* **4(5)**:327–337 DOI 10.2174/1389203033487072.

Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. 2020. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Molecular Plant* **13(8)**:1194–1202 DOI 10.1016/j.molp.2020.06.009.

Ganie SA, Pani DR, Mondal TK. 2017. Genome-wide analysis of DUF221 domain-containing gene family in Oryza species and identification of its salinity stress-responsive members in rice. *PLOS ONE* **12(8)**:e01824691371 DOI 10.1371/journal.pone.0182469.

Harada K, Yamashita E, Inoue K, Yamaguchi K, Fujiwara T, Nakagawa A, Kawasaki T, Kojima C. 2016. Plant-specific DUF1110 protein from Oryza sativa: expression, purification

and crystallization. *Acta Crystallographica Section F Structural Biology and Crystallization Communications* **72(6)**:480–484 DOI 10.1107/S2053230X16007573.

Ho-Huu J, Ronfort J, De Mita S, Bataillon T, Hochu I, Weber A, Chantret N. 2012. Contrasted patterns of selective pressure in three recent paralogous gene pairs in the Medicago genus (L.). *BMC Evolutionary Biology* **12**:195 DOI 10.1186/1471-2148-12-195.

Holland PW, Marletaz F, Maeso I, Dunwell TL, Paps J. 2017. New genes from old: asymmetric divergence of gene duplicates and the evolution of development. *Philosophical Transactions of the Royal Socoety Biolological Sciences* **372(1713)**:20150480 DOI 10.1098/rstb.2015.0480.

Hsieh PH, Kan CC, Wu HY, Yang HC, Hsieh MH. 2018. Early molecular events associated with nitrogen deficiency in rice seedling roots. *Scientific Report* **8(1)**:12207 DOI 10.1038/s41598-018-30632-1.

Jayaraman K, Sevanthi AM, Raman KV, Jiwani G, Solanke AU, Mandal PK, Mohapatra T. 2022. Overexpression of a DUF740 family gene (LOC_Os04g59420) imparts enhanced climate resilience through multiple stress tolerance in rice. *Frontiers Plant Science* **13**:947312 DOI 10.3389/fpls.2022.947312.

Knizewski L, Kinch LN, Grishin NV, Rychlewski L, Ginalski K. 2007. Realm of PD-(D/E)XK nuclease superfamily revisited: detection of novel families with modified transitive meta profile searches. *BMC Structural Biology* **7(1)**:40 DOI 10.1186/1472-6807-7-40.

Li LH, Lv MM, Li X, Ye TZ, He X, Rong SH, Dong YL, Guan Y, Gao XL, Zhu JQ, Xu ZJ. 2018. The rice OsDUF810 family: OsDUF810.7 may be involved in the tolerance to salt and drought. *Molecular Biology (Mosk)* **52(4)**:567–575 DOI 10.1134/S002689331804012X.

Li XX, Yu B, Wu Q, Min Q, Zeng RF, Xie ZZ, Huang JL. 2021. OsMADS23 phosphorylated by SAPK9 confers drought and salt tolerance by regulating ABA biosynthesis in rice. *PLoS Genetics* **17(8)**:e1009699 DOI 10.1371/journal.pgen.1009699.

Li L, Ye T, Guan Y, Lv M, Xie C, Xu J, Gao X, Zhu J, Cai L, Xu Z. 2017. Genome-wide identification and analyses of the rice OsDUF936 family. *Biotechnology & Biotechnological Equipment* **32(2)**:309–315 DOI 10.1080/13102818.2017.1413421.

Lim CW, Baek W, Jung J, Kim JH, Lee SC. 2015. Function of ABA in stomatal defense against biotic and drought stresses. *International Journal of Molecular Sciences* **16(7)**:15251–15270 DOI 10.3390/ijms160715251.

Lin Y, Zhu Y, Cui Y, Chen R, Chen Z, Li G, Fan M, Chen J, Li Y, Guo X, Zheng X, Chen L, Wang F. 2021. Derepression of specific miRNA-target genes in rice using CRISPR/Cas9. *Journal of Experimental Botany* **72(20)**:7067–7077 DOI 10.1093/jxb/erab336.

Lu Y, Wang J, Chen B, Mo S, Lian L, Luo Y, Ding D, Ding Y, Cao Q, Li Y, Li Y, Liu G, Hou Q, Cheng T, Wei J, Zhang Y, Chen G, Song C, Hu Q, Sun S, Fan G, Wang Y, Liu Z, Song B, Zhu JK, Li H, Jiang L. 2021. A donor-DNA-free CRISPR/Cas-based approach to gene knock-up in rice. *Nature Plants* **7(11)**:1445–1452 DOI 10.1038/s41477-021-01019-4.

Luo C, Guo C, Wang W, Wang L, Chen L. 2014. Overexpression of a new stress-repressive gene OsDSR2 encoding a protein with a DUF966 domain increases salt and simulated drought stress sensitivities and reduces ABA sensitivity in rice. *Plant Cell Reports* **33(2)**:323–336 DOI 10.1007/s00299-013-1532-0.

Lynch M, Conery JS. 2000. The evolutionary fate and consequences of duplicate genes. *Science* **290(5494)**:1151–1155 DOI 10.1126/science.290.5494.1151.

Ma Y, Dai X, Xu Y, Luo W, Zheng X, Zeng D, Pan Y, Lin X, Liu H, Zhang D, Xiao J, Guo X, Xu S, Niu Y, Jin J, Zhang H, Xu X, Li L, Wang W, Qian Q, Ge S, Chong K. 2015. COLD1 confers chilling tolerance in rice. *Cell* **160(6)**:1209–1221 DOI 10.1016/j.cell.2015.01.046.

Maruyama K, Todaka D, Mizoi J, Yoshida T, Kidokoro S, Matsukura S, Takasaki H, Sakurai T, Yamamoto YY, Yoshiwara K, Kojima M, Sakakibara H, Shinozaki K, Yamaguchi-Shinozaki K. 2012. Identification of cis-acting promoter elements in cold-and dehydration-induced transcriptional pathways in Arabidopsis, rice, and soybean. *DNA Research* **19(1)**:37–49 DOI 10.1093/dnares/dsr040.

Nadarajah K, Kumar IS. 2019. Drought response in rice: the miRNA story. *Internatiocal Journal of Molecular Sciences* **20(15)**:3766 DOI 10.3390/ijms20153766.

Panchy N, Lehti-Shiu M, Shiu SH. 2016. Evolution of gene duplication in plants. *Plant Physiology* **171(4)**:2294–2316 DOI 10.1104/pp.16.00523.

Sham A, Al-Ashram H, Whitley K, Iratni R, El-Tarabily KA, AbuQamar SF. 2019. Metatranscriptomic analysis of multiple environmental stresses identifies RAP2.4 gene associated with Arabidopsis immunity to botrytis cinerea. *Scientific Reports* **9(1)**:17010 DOI 10.1038/s41598-019-53694-1.

Shim JS, Jeong HI, Bang SW, Jung SE, Kim G, Kim YS, Redillas MCFR, Oh SJ, Seo JS, Kim JK. 2023. DROUGHT-INDUCED BRANCHED-CHAIN AMINO ACID AMINOTRANSFERASE enhances drought tolerance in rice. *Plant Physiology* **191(2)**:1435–1447 DOI 10.1093/plphys/kiac560.

Singhal P, Jan AT, Azam M, Haq QMR. 2016. Plant abiotic stress: a prospective strategy of exploiting promoters as alternative to overcome the escalating burden. *Frontiers Life Sciences* **9(1)**:52–63 DOI 10.1080/21553769.2015.1077478.

Song X, Li Y, Cao X, Qi Y. 2019. MicroRNAs and their regulatory roles in plant-environment interactions. *Annual Reviewes Plant Biology* **70(1)**:489–525 DOI 10.1146/annurev-arplant-050718-100334.

Sun Z, Liu K, Chen C, Chen D, Peng Z, Zhou R, Liu L, He D, Duan W, Chen H, Huang C, Ruan Z, Zhang Y, Cao L, Zhan X, Cheng S, Sun L. 2023. OsLDDT1, encoding a transmembrane structural DUF726 family protein, is essential for tapetum degradation and pollen formation in rice. *Plant Sciences* **329(49)**:111596 DOI 10.1016/j.plantsci.2023.111596.

Sun XM, Xiong HY, Jiang CH, Zhang DM, Yang ZL, Huang YP, Zhu WB, Ma SS, Duan JZ, Wang X, Liu W, Guo HF, Li GL, Qi JW, Liang CB, Zhang ZY, Li JJ, Zhang HL, Han LJ, Zhou YH, Peng YL, Li ZC. 2022. Natural variation of DROT1 confers drought adaptation in upland rice. *Nature Communication* **13(1)**:4265 DOI 10.1038/s41467-022-31844-w.

Tamura K, Stecher G, Kumar S. 2021. MEGA11: molecular evolutionary genetics analysis version 11. *Molecular Biology Evolution* **38(7)**:3022–3027 DOI 10.1093/molbev/msab120.

Thibaud-Nissen F, Ouyang S, Buell CR. 2009. Identification and characterization of pseudogenes in the rice gene complement. *BMC Genomics* **10(1)**:317 DOI 10.1186/1471-2164-10-317.

Throude M, Bolot S, Bosio M, Pont C, Sarda X, Quraishi UM, Bourgis F, Lessard P, Rogowsky P, Ghesquiere A, Murigneux A, Charmet G, Perez P, Salse J. 2009. Structure and expression analysis of rice paleo duplications. *Nucleic Acids Research* **37(4)**:1248–1259 DOI 10.1093/nar/gkn1048.

Wang Y, Mostafa S, Zeng W, Jin B. 2021a. Function and mechanism of jasmonic acid in plant responses to abiotic and biotic stresses. *International Journal of Molecular Sciences* **22(16)**:8568 DOI 10.3390/ijms22168568.

Wang J, Ren Y, Liu X, Luo S, Zhang X, Liu X, Lin Q, Zhu S, Wan H, Yang Y, Zhang Y, Lei B, Zhou C, Pan T, Wang Y, Wu M, Jing R, Xu Y, Han M, Wu F, Lei C, Guo X, Cheng Z, Zheng X, Wang Y, Zhao Z, Jiang L, Zhang X, Wang YF, Wang H, Wan J. 2021b. Transcriptional activation and phosphorylation of OsCNGC9 confer enhanced chilling tolerance in rice. *Molecular Plant* **14(2)**:315–329 DOI 10.1016/j.molp.2020.11.022.

Wang L, Shen R, Chen LT, Liu YG. 2014. Characterization of a novel DUF1618 gene family in rice. *Journal of Integrative Plant Biology* **56(2)**:151–158 DOI 10.1111/jipb.12130.

Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H, Kissinger JC, Paterson AH. 2012. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Research* **40(7)**:e49 DOI 10.1093/nar/gkr1293.

Yang SQ, Li WQ, Miao H, Gan PF, Qiao L, Chang YL, Shi CH, Chen KM. 2016. REL2, a gene encoding an unknown function protein which contains DUF630 and DUF632 domains controls leaf rolling in rice. *Rice (N Y)* **9(1)**:37 DOI 10.1186/s12284-016-0105-6.

Yang L, Takuno S, Waters ER, Gaut BS. 2011. Lowly expressed genes in Arabidopsis thaliana bear the signature of possible pseudogenization by promoter degradation. *Molecular Biology and Evolution* **28(3)**:1193–1203 DOI 10.1093/molbev/msq298.

Yim WC, Lee BM, Jang CS. 2009. Expression diversity and evolutionary dynamics of rice duplicate genes. *Molecular Genetics and Genomics* **281(5)**:483–493 DOI 10.1007/s00438-009-0425-y.

Ying S. 2021. Genome-wide identification and transcriptional analysis of Arabidopsis DUF506 gene family. *International Journal of Molecular Sciences* **22(21)**:111442 DOI 10.3390/ijms222111442.

Ying S, Blancaflor EB, Liao F, Scheible WR. 2022. A phosphorus-limitation induced, functionally conserved DUF506 protein is a repressor of root hair elongation in plants. *New Phytologist* **233(3)**:1153–1171 DOI 10.1111/nph.17862.

Ying S, Scheible WR. 2022. A novel calmodulin-interacting domain of unknown function 506 protein represses root hair elongation in Arabidopsis. *Plant Cell Environment* **45(6)**:1796–1812 DOI 10.1111/pce.14316.

Yuan X, Wang H, Cai J, Bi Y, Li D, Song F. 2007. Rice NAC transcription factor ONAC066 functions as a positive regulator of drought and oxidative stress response. *BMC Plant Biology* **19(1)**:278 DOI 10.1186/s12870-019-1883-y.

Yuan G, Zou T, He Z, Xiao Q, Li G, Liu S, Xiong P, Chen H, Peng K, Zhang X, Luo T, Zhou D, Yang S, Zhou F, Zhang K, Zheng K, Han Y, Zhu J, Liang Y, Deng Q, Wang S, Sun C, Yu X, Liu H, Wang L, Li P, Li S. 2022. SWOLLEN TAPETUM AND STERILITY 1 is required for tapetum degeneration and pollen wall formation in rice. *Plant Physiology* **190(1)**:352–370 DOI 10.1093/plphys/kiac307.

Zhang QF. 2007. Strategies for developing green super rice. *Proceedings of National Academy of Sciences of USA* **104(42)**:16402–16409 DOI 10.1073/pnas.0708013104.

Zhao M, Ju Y, Zhao B, Li X, Dai L, Qu J, Chu Z, Ding X. 2020. Transcriptional profiling analysis of OsDT11-mediated ABA-dependent signal pathway for drought tolerance in rice. *Plant Biotechnology Reports* **14(5)**:613–626 DOI 10.1007/s11816-020-00637-2.

Zhong R, Cui D, Ye ZH. 2019b. Evolutionary origin of O-acetyltransferases responsible for glucomannan acetylation in land plants. *New Phytologist* **224(1)**:466–479 DOI 10.1111/nph.15988.

Zhong H, Zhang H, Guo R, Wang Q, Huang X, Liao J, Li Y, Huang Y, Wang Z. 2019a. Characterization and functional divergence of a novel DUF668 gene family in rice based on comprehensive expression patterns. *Genes (Basel)* **10(12)**:980 DOI 10.3390/genes10120980.