Genome-wide identification of *R2R3-MYB* transcription factor subfamily genes involved in salt stress in rice (*Oryza sativa* L.)

Hao-Cheng Zhang[1†], Yuan-Hang Gong[1†], Tao Tao[1†], Shuai Lu[1], Wen-Yu Zhou[1], Han Xia[1], Xin-Yi Zhang[1], Qing-Qing Yang[1], Ming-Qiu Zhang[1], Lian-Min Hong[1], Qian-Qian Guo[1], Xin-Zhe Ren[1], Zhi-Di Yang[1], Xiu-Ling Cai[1,2], De-Yong Ren[4], Ji-Ping Gao[1,2*], Su-Kui Jin[1,2,3*] and Yu-Jia Leng[1,2*]

Abstract

Background R2R3-MYB transcription factors belong to one of the largest gene subfamilies in plants, and they are involved in diverse biological processes. However, the role of *R2R3-MYB* transcription factor subfamily genes in the response of rice (*Oryza sativa* L.) to salt stress has been rarely reported.

Results In this study, we performed a genome-wide characterization and expression identification of rice *R2R3-MYB* transcription factor subfamily genes. We identified a total of 117 *R2R3-MY*B genes in rice and characterized their gene structure, chromosomal location, and *cis*-regulatory elements. According to the phylogenetic relationships and amino acid sequence homologies, the *R2R3-MYB* genes were divided into four groups. qRT-PCR of the *R2R3-MYB* genes showed that the expression levels of 10 genes significantly increased after 3 days of 0.8% NaCl treatment. We selected a high expression gene *OsMYB2-115* for further analysis. *OsMYB2-115* was highly expressed in the roots, stem, leaf, and leaf sheath. OsMYB2-115 was found to be localized in the nucleus, and the yeast hybrid assay showed that OsMYB2-115 has transcriptional activation activity.

Conclusion This result provides important information for the functional analyses of rice *R2R3-MYB* transcription factor subfamily genes related to the salt stress response and reveals that *OsMYB2-115* may be an important gene associated with salt tolerance in rice.

Keywords Gene expression, R2R3-MYB, Rice (*Oryza sativa* L.), Salt stress, Transcription factor

†Hao-Cheng Zhang, Yuan-Hang Gong and Tao Tao contributed equally to this work.

*Correspondence:
Ji-Ping Gao
jpgao@yzu.edu.cn
Su-Kui Jin
skjin@yzu.edu.cn
Yu-Jia Leng
yujialeng@yzu.edu.cn
1Jiangsu Key Laboratory of Crop Genomics and Molecular Breeding/
Zhongshan Biological Breeding Laboratory/Key Laboratory of Plant

Functional Genomics of the Ministry of Education, Agricultural College of Yangzhou University, Yangzhou 225009, China
2Jiangsu Co-Innovation Center for Modern Production Technology of Grain Crops/Jiangsu Key Laboratory of Crop Genetics and Physiology, Yangzhou University, Yangzhou 225009, China
3Joint International Research Laboratory of Agriculture and Agri-Product Safety, the Ministry of Education of China, Yangzhou University, Yangzhou 225009, Jiangsu, China
4State Key Laboratory of Rice Biology and Breeding, National Rice Research Institute, Hangzhou 310006, China

Background

Soil salinization is an important factor that affects the growth and development of plants, especially salt-sensitive crops. Currently, more than 6% of the world's arable land is affected by salt damage [1]. Rice is the staple food for more than half of the world's population, and it is sensitive to salt. With continuous urbanization and population growth and a decrease in effective cultivated land area, development and utilization of saline-alkali land is the future development trend for agriculture. Cloning salt-tolerant genes and cultivating salt-tolerant rice varieties are important ways to quickly and effectively resist salt damage and increase rice planting area and yield.

Plants respond and adapt to salt stress through a series of molecular, cellular, physiological and biochemical reaction, which involve in the regionalization of Na^+ in the cytoplasm, the selective absorption, accumulation, or excretion of ions, the induction of stress tolerance gene expression, etc. [2]. Transcription factors play a central roles in signal transduction networks which modulate the expression of a number of stress responsive genes by binding with *cis*-elements in the promoter [3]. The myeloblastosis (MYB) transcription factor family is large, functionally diverse, and represented in all eukaryotes [4]. Understanding the regulatory mechanism of MYB transcription factors in response to salt stress is of great significance in improving plant processes and enhancing agricultural production. In plants, MYB proteins are divided into four subfamilies on the basis of the number of approximately 52 amino acid sequence repeats (R) of a highly conserved DNA-binding domain at the N-terminus: 1R-MYBs/MYB-related, 2R/R2R3-MYBs, 3R/R1R2R3-MYBs, and 4R-MYBs [5]. Each repeat folds into three α-helices, and the second and third helices with three regularly spaced tryptophan (or hydrophobic) residues form a helix–turn–helix structure with a hydrophobic core [6, 7]. In contrast, a high variation in the C-terminus (activation domain) of MYB transcription factors leads to its broad regulatory role [8, 9]. The 4R-MYB subfamily is the smallest in plants. For example, only one or two 4R-MYB genes have been identified in *Arabidopsis* and rice [10]. There is limited information on the functions of these proteins in plants. R1R2R3-MYB is an evolutionarily conserved subfamily in plants, and it has been found in most eukaryotic genomes with roles in cell cycle regulation [11]. MYB-related proteins contain a single or partial MYB repeat, and they belong to the second largest subfamily of MYB transcription factors in plants [4, 5]. The first reported MYB-related transcription factor gene, *MybSt1*, was identified in potato [12]. Subsequently, many MYB-related genes were isolated and found to play an important role in cellular and organ morphogenesis, secondary metabolism, circadian rhythm, etc. [13, 14]. The most common MYB subfamily is R2R3-MYBs, and it is thought to have derived from an R1R2R3-MYB gene ancestor and lost R1 repeats during evolution [15]. Many plant species encode more than 100 R2R3-MYB transcription factors and exhibit diverse functions during expansion [16]. The first plant MYB gene identified was *COLORED1* (*C1*, R2R3-MYB) from *Zea mays* [17]. Recently, numerous R2R3-MYB genes have been identified and found to play key regulatory roles in metabolism [18–23], plant growth and development [24–27], hormone signal transduction [28–31], and responses to biotic and abiotic stresses [32–35].

In rice, several *R2R3-MYB* transcription factor genes involved in salt tolerance have been isolated. *OsMYB2* encodes a stress-responsive R2R3-MYB transcription factor that plays a positive regulatory role under salt stress conditions [36]; Overexpression of *OsMYB6* results in increased proline content and catalase and superoxide dismutase activities and decreased relative electrolyte leakage and malondialdehyde content, which improves the tolerance of rice to salt stress [37]. Similarly, overexpression of *OsMYB91* enhances tolerance to salt stress, with a significant increase in proline content and capacity to scavenge reactive oxygen species as well as increased induction of *OsP5CS1* and *LOC_Os03g44130* under salt stress [38]. Although the rice genome was sequenced and published years ago, the systematic analysis of *R2R3-MYB* transcription factor genes involved in salt stress has not yet been conducted. In this study, we identified 117 *R2R3-MYB* transcription factor subfamily genes in rice and analyzed their basic information, chromosomal distribution, gene structures, phylogenic relationships, *cis*-regulatory elements, and gene expression patterns under salt stress. Additionally, a gene highly expressed under salt stress was selected for further analysis. Our results provide important information for further functional characterization of *R2R3-MYB* transcription factor subfamily genes involved in salt stress.

Results and discussion

Identification and characterization of *R2R3-MYB* genes in rice

To identify *R2R3-MYB* genes in rice, we searched and integrated data from the Plant Transcription Factor Database and a previously published paper [39]. Further screening was conducted using the Pfam database to confirm the presence of the conserved MYB domain (PF00249), and a total of 117 putative *R2R3-MYB* genes in rice were obtained and named *OsMYB2-1–OsMYB2-117* (Supplementary Table 1). The number of *R2R3-MYB* genes identified in this study was more than Kang et al. (2022) reported [40], which provided a more updated and comprehensive depiction of R2R3-MYB subfamily members in rice. In addition, the number of *R2R3-MYB* genes in rice was less than the number of

Zhang *et al. BMC Genomics*   (2024) 25:797

Page 3 of 12

genes identified in *Arabidopsis* (126) [9], *Z. mays* (157) [41], *M. truncatula* (155) [42], and *S. lycopersicum* (127) [43] and more than the number of genes identified in *P. patens* (62) [4] and *C. reinhardtii* (10) [41]. This indicated that *R2R3-MYB* genes in different plants have expanded to different degrees during evolution.

The nucleotide lengths, protein lengths, molecular weights, and isoelectric points of rice and *Arabidopsis* R2R3-MYB transcription factors all vary greatly. The nucleotide lengths of 117 rice and 125 *Arabidopsis R2R3-MYB* genes ranged from 402 bp to 4743 bp and 417 bp to 2532 bp, with an average of 1022.1 bp and 952.8 bp, respectively. The corresponding protein lengths ranged from 134 to 1581 and 139 to 844, with an average of 340.7 and 317.6, respectively. The molecular weights ranged from 14780.4 Da to 170922.0 Da and 16006.6 Da to 95765.7 Da, with an average of 36896.5 and 36040.0, respectively. The isoelectric points of these proteins ranged from 4.4 to 10.9 and 4.7 to 10.9, with an average of 7.0 and 7.1, respectively (Table 1, Supplementary Tables 1 and 2). Subcellular localization prediction results showed that the 117 rice and 125 *Arabidopsis* R2R3-MYB proteins were all located in the nucleus and corresponded to their transcription factor family functions (Supplementary Tables 1 and 2).

Chromosomal distribution and gene structure of *R2R3-OsMYB*

To better understand the genomic distribution of *R2R3-MYB* genes, the localization of these genes in rice chromosomes were marked. As shown in Fig. 1, the 117 *R2R3-OsMYB* genes were unevenly distributed on 12 chromosomes. The maximum number of *R2R3-OsMYB* genes was found on chromosome 1 (22; 18.8%), followed by chromosome 2 (13; 11.1%), whereas the least number of *R2R3-OsMYB* genes was detected on chromosome 10 (2; 1.7%). Twelve *R2R3-OsMYB* genes were found on chromosomes 3, 4 and 5, respectively. Similarly, eight *R2R3-OsMYB* genes were detected on chromosomes 6 and 7, respectively. In addition, 10, 6, 5, and 7 *R2R3-OsMYB* genes were anchored on chromosomes 8, 9, 11, and 12, respectively (Fig. 1). Other than that, *R2R3-OsMYB* genes located on chromosome 2, 4, 5, and 8 accumulate at the lower end of the arms in comparison to other chromosomes.

To analyze the conserved motifs, the distributions of R2R3-OsMYB protein motifs were checked using MEME. A total of eight conserved motifs were identified (Supplementary Fig. 1). Most R2R3-OsMYB proteins contained similar types of conserved motifs (Fig. 2A). Motif 3 was the most conserved, followed by motif 2; Motif 8 was the least conserved. Notably, *R2R3-OsMYB* genes within the same phylogenetic cluster showed similar conserved motifs.

Table 1 Nucleic acid and protein sequence features of R2R3-MYB transcription factor subfamily genes in rice and *Arabidopsis*

Name	No. of genes	Nucleotide length			Protein length			Molecular weight			Isoelectric point		
		Max.	Mix.	Avg.	Max.	Mix.	Avg.	Max.	Mix.	Avg.	Max.	Mix.	Avg.
R2R3-OsMYB	117	4743	402	1022.1	1581	134	340.7	170922.0	14780.4	36896.5	10.9	4.4	7.0
R2R3-AtMYB	125	2532	417	952.8	844	139	317.6	95765.7	16006.6	36040.0	10.9	4.7	7.1

Zhang *et al. BMC Genomics*   (2024) 25:797
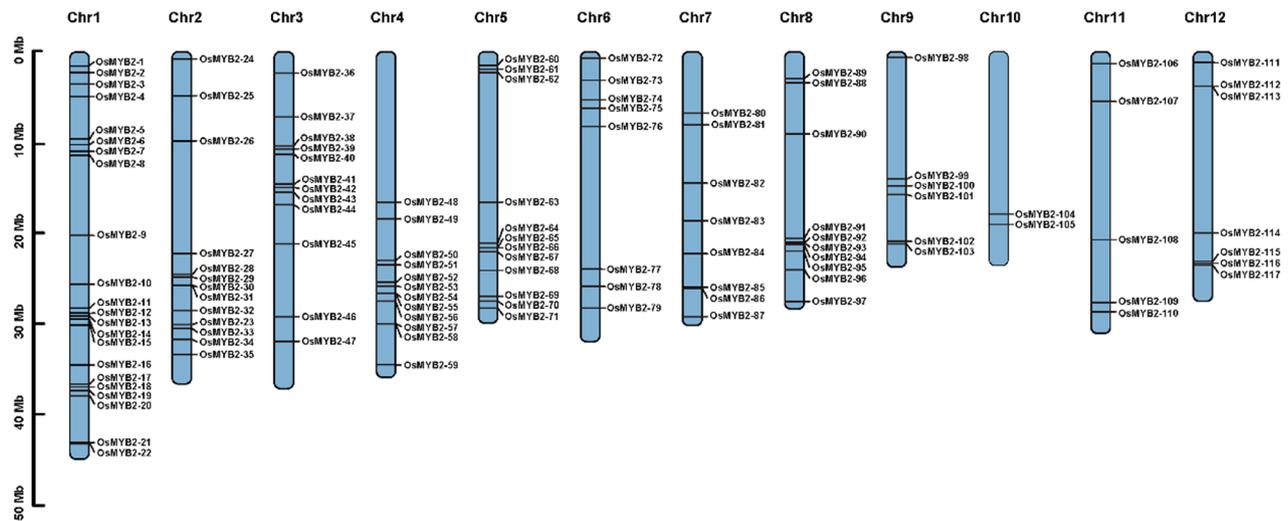
Page 4 of 12

Fig. 1 Distribution of *R2R3-MYB* transcription factor subfamily genes on rice chromosomes. Chromosome numbers are showed at the top of each chromosome. The name of each *R2R3-OsMYB* genes are shown on the right side of the chromosome. The scale represents 50 Mb chromosomal distance. The bars indicate the positions of the *R2R3-OsMYB* genes on the chromosomes

Gene structural diversity may be related to gene family evolution [44]. To evaluate the structural diversity of *R2R3-OsMYB* genes, the exon–intron arrangements were analyzed to gain insights into the gene structural organization. The number of exons among *R2R3-OsMYB* genes varied largely, ranging from 1 to 12, and 41.9% of the *R2R3-OsMYB* genes had three exons (Fig. 2B). *OsMYB2-85* had the highest exon number (12), whereas *OsMYB2-7*, *OsMYB2-9*, *OsMYB2-11*, *OsMYB2-21*, *OsMYB2-22*, *OsMYB2-24*, *OsMYB2-25*, *OsMYB2-30*, *OsMYB2-34*, *OsMYB2-44*, *OsMYB2-49*, *OsMYB2-53*, *OsMYB2-61*, *OsMYB2-63*, *OsMYB2-78*, *OsMYB2-81*, *OsMYB2-82*, *OsMYB2-88*, and *OsMYB2-95* harbored only one exon each. In general, members with close relationships in the phylogenetic tree tend to have similar exon-intron structures [44]. Interestingly, genes in the same phylogenetic cluster showed different gene structures in our study.

Phylogenetic analysis of R2R3-OsMYB proteins
To understand the phylogenetic relationships of the R2R3-OsMYB proteins, a neighbor-joining phylogenetic tree was constructed using the aligned amino acid sequences from rice (117) and *Arabidopsis* (125) in MEGA X [45]. The R2R3-OsMYB proteins were clustered into four groups (Fig. 3). Group 4 contained 100 members, which was the highest among the four groups. Group 3 was the second highest, with 15 members. Group 1 and Group 2 contained only one member each (OsMYB2-91 and OsMYB2-113, respectively). In contrast, *Arabidopsis R2R3-MYB* genes were divided into two groups, with 4 and 121 genes (Fig. 3). The result indicated that the R2R3-MYB subfamily was not very conserved during evolution.

Analysis of *cis*-regulatory elements in *R2R3-MYB* promoter regions
Cis-regulatory elements (CREs) interact with *trans*-regulators to regulating gene expression [46]. To analyze the transcriptional regulation of *R2R3-OsMYB* genes, 2000-bp sequences upstream of the start codon of *R2R3-OsMYB* genes were regarded as promoter sequences and submitted to the PlantCARE database to search for *cis*-regulatory elements. A total of 2916 putative *cis*-regulatory elements were predicted, including 1269 growth and biological process responsive *cis*-regulatory elements, 499 stress responsive *cis*-regulatory elements, 1065 hormone responsive *cis*-regulatory elements, and 83 metabolism responsive *cis*-regulatory elements (Supplementary Fig. 2). A total of 40 growth and biological process responsive *cis*-regulatory element types were predicted; Box 4, GT1-motif, and CAT-box were the top three identified *cis*-regulatory elements (254, 142, and 114, respectively), followed by hormone responsive and stress responsive *cis*-regulatory elements (13 and 6, respectively). With respect to metabolism responsive *cis*-regulatory elements, only one type (O2-site) of *cis*-regulatory element was identified. The number of *cis*-regulatory elements for each *R2R3-OsMYB* gene was analyzed (Fig. 4), and it ranged from 12 to 58, with a minimum of 12 for *OsMYB2-105* and a maximum of 58 for *OsMYB2-78*. In addition, the 117 *R2R3-OsMYB* genes contained at least one growth and biological process responsive *cis*-regulatory element and hormone responsive *cis*-regulatory element, and 2 and 58 genes did not contain stress responsive *cis*-regulatory elements and metabolism responsive *cis*- regulatory elements, respectively. These results suggest that the *R2R3-OsMYB* genes

Fig. 2 Conserved motifs and gene structure analyses of rice *R2R3-MYB* genes. **A**. The conserved motif distribution of rice R2R3-MYB proteins; Different conserved motifs were represented by seven different colors. **B**. Exon-intron structures of rice *R2R3-MYB* genes. The green boxes, yellow boxes and black lines indicate untranslated regions, exons, and introns, respectively

are essential regulators of growth and biological processes and hormones.

Expression patterns of *R2R3-MYB* genes under salt treatment

Salt stress, apart from drought and high temperature stresses, is a major environmental factor that affects plant

growth and productivity. Osmotic and ion stresses are two main methods of salt damage to plants. Both stress signals are transmitted to a secondary messenger through membrane-bound cytosolic sensors, which then activate the protein phosphorylation cascade that helps plants to adapt to salt stress [47]. In *Arabidopsis*, R2R3-MYB transcription factors are mainly involved in the regulation of
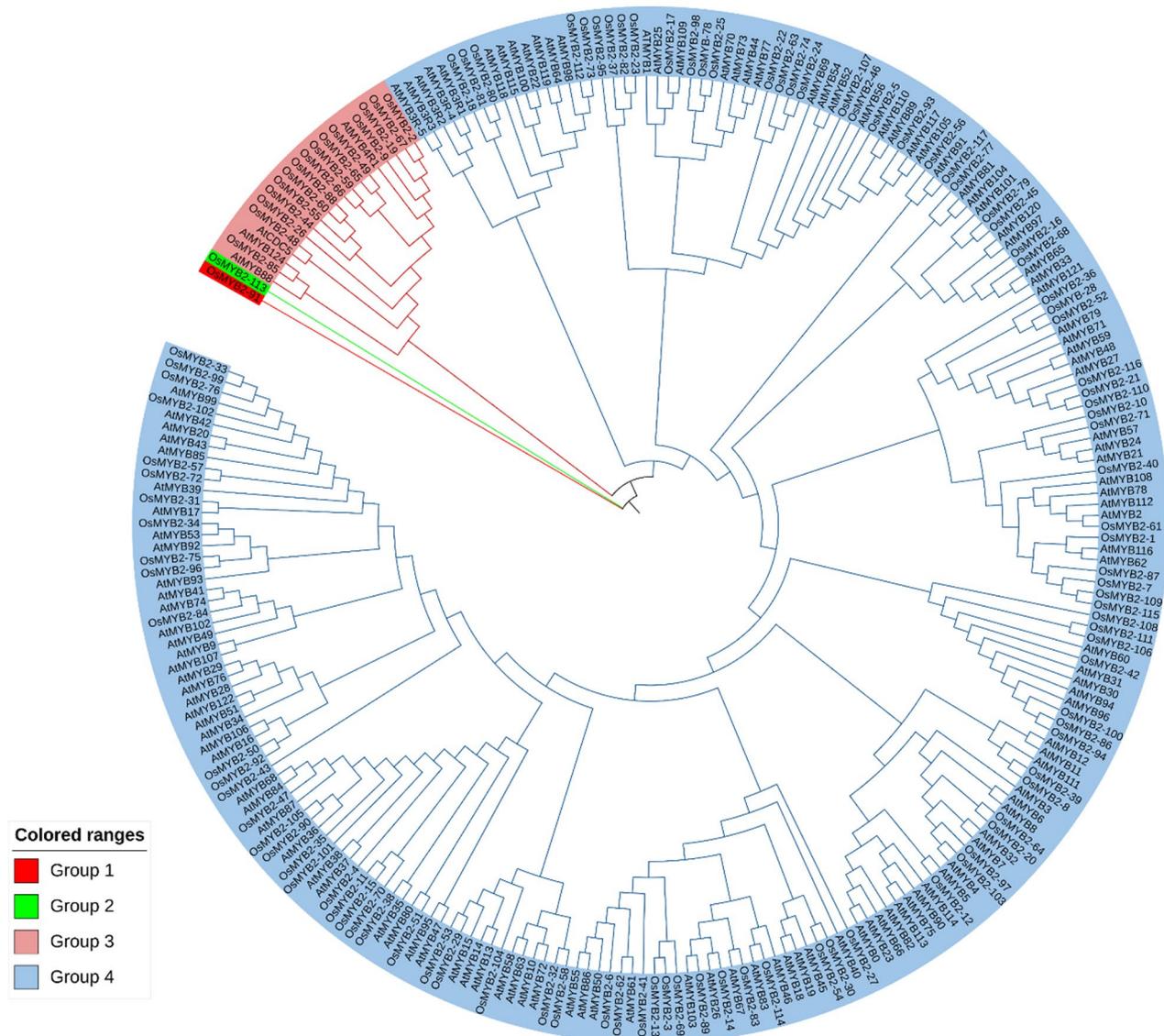
Fig. 3 Phylogenetic analysis of R2R3-MYB proteins in rice and *Arabidopsis*. The neighbor-joining method were used to constructed phylogenetic tree of rice (117) and *Arabidopsis* (125) MYB proteins by Mega X software with a bootstrap. The MYB proteins were classified into four subfamilies: Group 1, 2, 3 and 4 depicted by red, green, pink, and blue, respectively

downstream target genes for salt tolerance [48]. In rice, five *R2R3-MYB* genes, *OsMYB2* [36], *OsMYB6* [37], *OsMYB48-1* [49], *OsMYB91* [38], and *OsMPS* [50], have been shown to play an important role in the response to salt stress. However, there is limited information on the regulatory functions of other *R2R3-MYB* genes in rice under salt stress conditions. In this study, rice *R2R3-MYB* subfamily genes involved in the salt stress response were systematically identified.

Because of differences in gene expression patterns after abiotic stress treatments, stress-responsive genes have been extensively studied. Zhang et al. (2022) identified three upregulated genes and seven downregulated genes from 12 rice MDH family members under salt

treatment [51]. Gene expression analysis of 26 *BpGRASs* showed that six *BpGRASs* were significantly upregulated under salt treatment [52]. Wang et al. (2020) analyzed the expression of *RsCPA* family genes in radish under salt treatment, and the expression of 9 of 60 *RsCPA* genes increased significantly [53]. In this study, 117 *R2R3-MYB* genes were identified in rice, and the expression levels of 10 genes were four times higher than the control after 3 days of 0.8% NaCl treatment (Fig. 5A). Interestingly, the expression levels of the five reported rice *R2R3-MYB* genes (*OsMYB2*, *OsMYB6*, *OsMYB48-1*, *OsMYB91*, and *OsMPS*) involved in the response to salt stress showed no significant differences, except for *OsMYB48-1*. A possible reason is that the previous studies all focused on salt
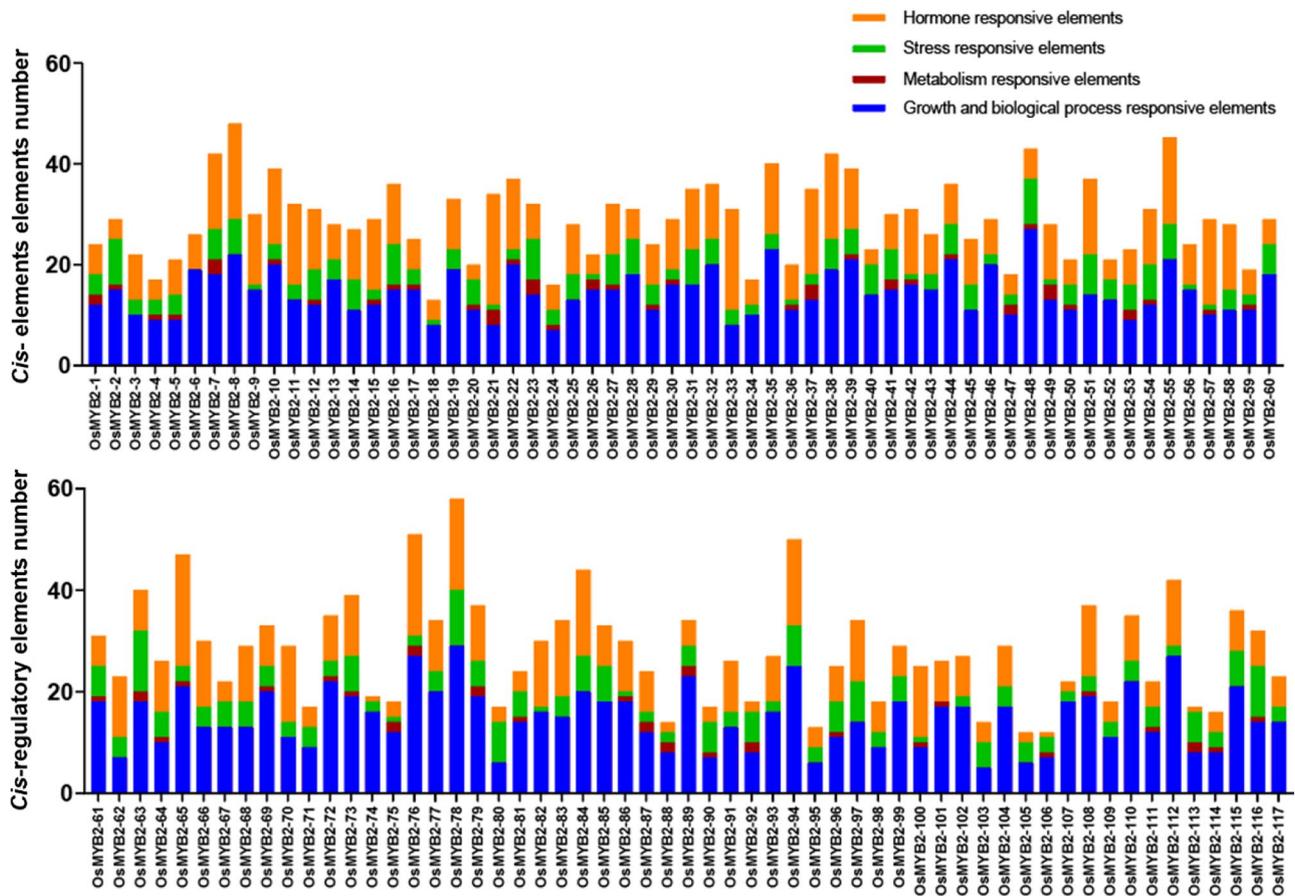
Fig. 4 The number of putative regulatory *cis*-regulatory elements in the promoters of *R2R3-MYB* genes. The predicted *cis*-regulatory elements in the promoters of *R2R3-MYB* gene were classified into four broad categories, including hormone responsive elements (orange), stress responsive elements (green), metabolism responsive elements (brown), and growth and biological process responsive elements (blue)
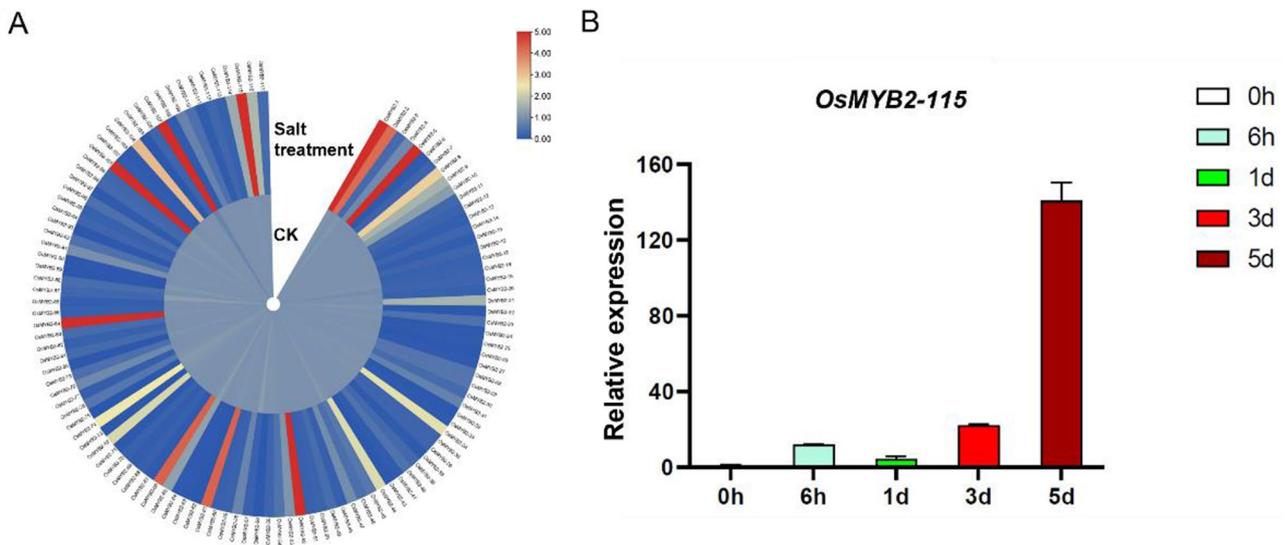

Fig. 5 Expression patterns of *R2R3-MYB* subfamily genes after NaCl treatment. **A**. The expression pattern of *R2R3-MYB* subfamily genes after 0.8% NaCl treatment for 3 days in rice; **B**. The expression patterns of *OsMYB2-115* genes after 0.8% NaCl treatment for 0 h, 6 h, 1 d, 3 d and 5 d in rice. Data are means ± SD of three biological replicates

treatment for less than 24 h, whereas salt treatment was conducted for 3 days in this study, indicating that these genes have different regulatory mechanisms for salt stress response.

The expression differences of *R2R3-MYB* genes under salt treatment may be closely related to *cis*-regulatory elements of the promoter region. ABA-response elements (ABREs) have been functionally identified in the promoters of various genes, and they are involved in regulating the expression of ABA-related genes and improving the adaptability of plants to environmental stress [54]. For example, bZIP transcription factors can bind to ABRE *cis*-regulatory elements and regulate the expression of downstream target genes and improve the salt tolerance of rice [55, 56]. In this study, ABREs (hormone-responsive elements) accounted for 40% of the *R2R3-MYB* promoter *cis*-regulatory elements (Fig. 4; Supplementary Fig. 2), which may play important roles in regulating the response of *R2R3-MYB* genes to abiotic stress.

Expression patterns, subcellular localization, and transcriptional activity analysis

We randomly selected a gene (*OsMYB2-115*) to analyze its expression levels at 0 h, 6 h, 1 day, 3 days and 5 days after salt treatment. The results showed that the expression levels of *OsMYB2-115* increased with the duration of salt treatment (Fig. 5B). To analyze the spatial and developmental expressions of *OsMYB2-115*, we isolated RNA from different tissues, namely, root and leaf at the seedling stage and culm, leaf, sheath, internode, panicle, and spikelet at the heading stage, and determined the transcript levels of *OsMYB2-115* by using qRT-PCR. The results showed that *OsMYB2-115* transcripts were detected in all examined tissues (Fig. 6A), and the expression levels were high in the spikelet, leaf, sheath, and internode at the heading stage and leaf at the seedling stage. However, lower expression levels were found in the young root, culm, and panicle.

To examine the subcellular localization of OsMYB2-115, the full-length CDS of *OsMYB2-115* was cloned upstream of the green fluorescent protein (GFP) gene under the control of the cauliflower mosaic virus 35S promoter to construct N-terminus OsMYB2-115-GFP fusion proteins. We transiently expressed single GFP and OsMYB2-115-GFP fusion proteins in rice protoplasts; GFP alone was found throughout the cell, whereas the OsMYB2-115-GFP fusion proteins were specifically localized in the nucleus (Fig. 6B).

We analyzed the transcriptional activation activity of OsMYB2-115 by performing the yeast one-hybrid assay. The coding regions of OsMYB2-115 were fused to the DNA-binding domain of yeast GAL4. Empty pGBKT7,
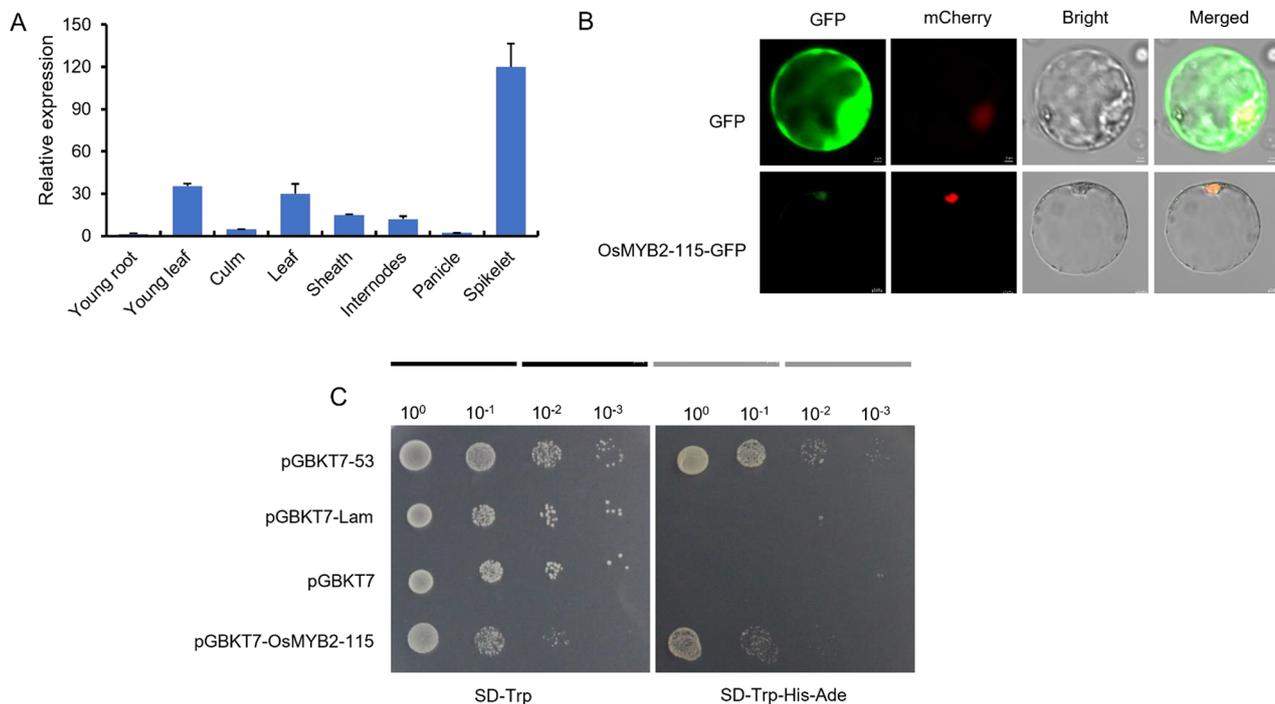

Fig. 6 Expression pattern, subcellular localization, and transcriptional activity analyses. **A**. *OsMYB2-115* expression in different tissues was detected by qRT-PCR. Data are means ± SD of three biological replicates; **B**. Subcellular localization of the OsMYB2-115 protein. GFP alone, mCherry and OsMYB2-115-GFP recombinant vector were transformed into rice protoplasts. The corresponding GFP alone was used as a positive control. mCherry fused with the nuclear localization signal as a nuclear localization control; **C**. Transcriptional activity using yeast-one-hbrid assay. pGBKT7, pGBKT7-53 and pGBKT7-Lam as negative, negative and positive controls, respectively

pGBKT7-Lam, and pGBKT7-53 were regarded as negative, negative, and positive controls, respectively. The vectors OsMYB2-115-BD, pGBKT7, pGBKT7-Lam, and pGBKT7-53 were expressed in yeast cells. The yeast cells containing OsMYB2-115-BD and pGBKT7-53 were able to grow on the SD/-Trp-His-Ade medium, whereas the yeast cells containing pGBKT7 and pGBKT7-Lam could not grow (Fig. 6C); This suggested that OsMYB2-115-BD has transcriptional activation activity.

Materials and methods

Screening and identification of *R2R3-OsMYB* genes in rice

The whole genome and protein sequences of *R2R3-MYB* transcription factor subfamily genes in rice (*Oryza sativa* L.) were obtained from the Rice Genome Annotation Project (http://rice.uga.edu/index.shtml) [57], Rice Annotation Project Database (RAP-DB, https://rapdb.dna.affrc.go.jp/index.html) [58], and Plant Transcription Factor Database (http://planttfdb.gao-lab.org/) [59]. Protein sequences of the *R2R3-MYB* transcription factor subfamily genes in *Arabidopsis* were obtained from the Arabidopsis Information Resource (https://www.arabidopsis.org/). All proteins were confirmed to contain two MYB DNA-binding domains (Pfam number: PF00249) by using the Pfam database (http://pfam-legacy.xfam.org/) [60].

The nucleotide length, predicted protein length, predicted molecular weight, predicted isoelectric point, and related annotation of *R2R3-OsMYB* and *R2R3-AtMYB* were obtained from the Rice Genome Annotation Project (http://rice.uga.edu/index.shtml) [57] and Arabidopsis Information Resource (https://www.arabidopsis.org/) [61], respectively. The subcellular localizations of R2R3-OsMYB and R2R3-AtMYB were identified using Cell-PLoc 2.0 (http://www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc-2/) [62]. The names of the rice *R2R3-OsMYB* genes were according to their chromosome locations, and the names of *Arabidopsis R2R3-AtMYB* genes were according to Stracke et al. (2001) [9].

Chromosomal distribution and structural analysis of *R2R3-OsMYB* genes

The 117 *R2R3-OsMYB* genes were located in the corresponding chromosomes of rice, according to the chromosomal location information extracted from RAP with TBtools software [63]. Gene Structure Display Server (GSDS v.2.0) was used to analyze intron and exon structures of the *R2R3-OsMYB* genes [64]. The conserved motifs of the 117 R2R3-OsMYB proteins were identified using Multiple Em for Motif Elicitation (MEME; https://meme-suite.org/meme/tools/meme) motif search tool with default settings [65].

Phylogenetic analysis of *R2R3-OsMYB* genes

The R2R3-OsMYB protein sequences of rice (117) and *Arabidopsis* (125) were aligned using ClustalW algorithm of MEGA X software [45]. A neighbor-joining phylogenetic tree was constructed for rice and *Arabidopsis thaliana* on the basis of the amino acid sequence alignment results, with 1000 bootstrap replications. Then, a Newick file was uploaded on iTOL (https://itol.embl.de) to visualize the circle tree [66].

Identification of putative *cis*-regulatory elements in the promoters

To determine *cis*-regulatory elements of the predicted promoters, 2000-bp sequences upstream from the start codon (ATG) of *R2R3-OsMYB* genes were obtained from the Rice Genome Annotation Project Database (http://rice.uga.edu/index.shtml). Putative *cis*-regulatory elements in these sequences were used to search the PlantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) [67]. Promoter structure and motif were visualized with TBtools software [63].

Plant materials and salt treatments

The *japonica* rice variety Nipponbare used in this study were preserved by our laboratory. The planting and harvesting of seeds were completed at the experimental farms of Yangzhou University (119∘42′ E, 32∘39′ N). Seeds were soaked in water in the dark at 30 °C for 2 days and then germinated in water at 37 °C for 1 day. The germinated seeds were transferred to 96-well PCR plates with the bottoms cut out and placed in trays with 5 L ddH$_2$O for 5 days. The uniform seedlings were transferred to foam boards and grown hydroponically in Yoshida solution [68]; HBD-500DH plant growth chamber (ChangzhouHaibo, China) was used. The parameters were set as light for 16 h at 30 °C and dark for 8 h at 28 °C during the experiments. Ten-day-old seedlings were treated with 0.8% NaCl solution. Seedlings in Yoshida solution were used as the control.

RNA extraction and qRT-PCR analyses of *R2R3-OsMYB* genes

To analyze the expression of *R2R3-OsMYB* genes under salt stress, leaf samples of the control and salt treatment groups were collected at 0 h, 6 h, 1 day, 3 days, and 5 days. The expression patterns of *OsMYB2-115* in different organs and growth stages, namely, the root, seedling, stem, leaf, leaf sheath, node, spikelet, and panicle, were analyzed. Total RNA was extracted using the RNAprep Pure Plant kit (TIANGEN, Germany), and reverse transcription into cDNA was conducted with the PrimeScript™ RT reagent kit with gDNA Eraser (TaKaRa, Japan), according to the manufacturer's instructions. The cDNA (1 μl) was used for real-time RT-PCR with

the TB Green® Premix EX Taq™ kit (TaKaRa, Japan) and gene-specific primers (Supplementary Table 3) in the CFX Connect™ Real-Time System (Bio-Rad Laboratories, USA). The relative expression levels of *R2R3-OsMYB* genes were calculated according to the comparative Ct method by using the $2^{-\Delta\Delta Ct}$ formula [69].

Subcellular localization of OsMYB2-115

The coding region of *OsMYB2-115* was amplified from NPB leaf cDNA, and the fragment was fused into the PA7-GFP vector to generate the OsMYB2-115-GFP recombinant vector by using the NovoRec® plus One step PCR Cloning Kit (Novoprotein, China). Then, plasmids of GFP, mCherry-NLS (mCherry fused with the nuclear localization signal), and OsMYB2-115-GFP were transformed into rice protoplasts. The protoplasts were incubated for 16 h at 28 °C, and green fluorescent signals were observed using a Zeiss LSM 880NLO confocal microscope (Carl Zeiss, Germany). The primer sequences are listed in Supplementary Table 4.

Transcriptional activity analysis

Transcription activation was analyzed using the Matchmaker® Gold Yeast Two-Hybrid System (Clontech, Japan). The coding sequence of *OsMYB2-115* was amplified from NPB leaf cDNA and inserted into the vector pGBKT7 to fuse with the GAL4 DNA-binding domain by using the NovoRec® plus One step PCR Cloning Kit (Novoprotein, China). pGBKT7-53 and pGBKT7-Lam were used as positive and negative controls, respectively. All vectors were transformed into the yeast strain Y2H Gold, and the strains were grown on SD/-Trp and SD/-Trp-His-Ade culture media.

Conclusions

In this study, we identified a total of 117 *R2R3-MY*B transcription factor subfamily genes in rice. The phylogenetic analysis indicated that the rice *R2R3-MYB* genes were divided into four groups, and the genes exhibited an exon–intron structure and motif composition similar to those of other species. The promoter *cis*-regulatory element analysis indicated that *R2R3-MYB* genes are essential regulators of growth, biological processes, and hormones. Ten *R2R3-MYB* genes showed significant upregulation after 0.8% NaCl treatment for 3 days. The expression levels of *OsMYB2-115* increased with salt treatment duration, and OsMYB2-115 was found to be localized in the nucleus and displayed transcriptional activation activity.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-024-10693-5.

Supplementary Material 1

Supplementary Material 2

Data availability
The whole genome sequence information for R2R3-MYB transcription factor subfamily genes in rice (Oryza sativa L.) was obtained from the Rice Annotation Project (https://rapdb.dna.affrc.go.jp/index.html), Rice Genome Annotation Project (http://rice.uga.edu/index.shtml), and Plant Transcription Factor Database (http://planttfdb.gao-lab.org/). Protein sequences of the R2R3-MYB transcription factor subfamily genes in Arabidopsis were obtained from the Arabidopsis Information Resource (https://www.arabidopsis.org/). The datasets supporting the conclusions of this study are included in the article and in additional files.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

References
1. Munns R, Tester M. Mechanisms of salinity tolerance. Annu Rev Plant Biol. 2008;59:651–81.
2. Ma L, Liu X, Lv W, Yang Y. Molecular mechanisms of plant responses to salt stress. Front. Plant Sci. 2022;13:934877.
3. Jisha V, Dampanaboina L, Vadassery J, Mithöfer A, Kappara S, Ramanan R. Overexpression of an AP2/ERF type transcription factor *OsEREBP1* confers biotic and abiotic stress tolerance in rice. PLoS ONE. 2015;10:e0127831.
4. Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L. MYB transcription factors in *Arabidopsis*. Trends Plant Sci. 2010;15(10):573–81.
5. Yang XY, Guo T, Li J, Chen Z, Guo B, An XM. Genome-wide analysis of the MYB-related transcription factor family and associated responses to abiotic stressors in *Populus*. Int J Biol Macromol. 2021;191:359–76.
6. Ogata K, Hojo H, Aimoto S, Nakai T, Nakamura H, Sarai A, Ishii S, Nishimura Y. Solution structure of a DNA-binding unit of myb: a helix-turn-helix-related motif with conserved tryptophans forming a hydrophobic core. Proc Natl Acad Sci U S A. 1992;89(14):6428–32.
7. Ogata K, Kanei-Ishii C, Sasaki M, Hatanaka H, Nagadoi A, Enari M, Nakamura H, Nishimura Y, Ishii S, Sarai A. The cavity in the hydrophobic core of myb DNA binding domain is reserved for DNA recognition and *trans*-activation. Nat Struct Biol. 1996;3(2):178–87.

8. Jin H, Martin C. Multifunctionality and diversity within the plant MYB-gene family. Plant Mol Biol. 1999;41(5):577–85.

9. Stracke R, Werber M, Weisshaar B. The R2R3–MYB gene family in *Arabidopsis thaliana*. Curr Opin Plant Bio. 2001;4(5):447–56.

10. Katiyar A, Smita S, Lenka SK, Rajwanshi R, Chinnusamy V, Bansal KC. Genome-wide classification and expression analysis of *MYB* transcription factor families in rice and Arabidopsis. BMC Genom. 2012;13:544.

11. Cao YP, Li K, Li YL, Zhao XP, Wang LH. MYB transcription factors as regulators of secondary metabolism in plants. Biology. 2020;9(3):61.

12. Baranowskij N, Frohberg C, Prat S, Willmitzer L. A novel DNA binding protein with homology to myb oncoproteins containing only one repeat can function as a transcriptional activator. EMBO J. 1994;13(22):5383–92.

13. Elena B, Annamaria G, Eleonora C. Plant MYB transcription factors: their role in drought response mechanisms. Int J Mol Sci. 2015;16:15811–51.

14. Roy S. Function of MYB domain transcription factors in abiotic stress and epigenetic control of stress response in plant genome. Plant Signal Behav. 2016;11(1):e1117723.

15. Rosinski JA, Atchley WR. Molecular evolution of the Myb family of transcription factors: evidence for polyphyletic origin. J Mol Evol. 1998;46(1):74–83.

16. Millard PS, Kragelund BB, Burow M. R2R3 MYB transcription factors–functions outside the dna-binding domain. Trends Plant Sci. 2019;24(10):934–46.

17. Paz-Ares J, Ghosal D, Wienand U, Peterson PA, Saedler H. The regulatory c1 locus of Zea mays encodes a protein with homology to myb proto-oncogene products and with structural similarities to transcriptional activators. EMBO J. 1987;6(12):3553–8.

18. Guo YH, Guo ZY, Zhong J, Liang YL, Feng Y, Zhang P, Zhang QX, Sun M. Positive regulatory role of R2R3 MYBs in terpene biosynthesis in *Lilium* 'Siberia'. Hortic. Plant J. 2023;9(5):1024e1038.

19. He G, Zhang R, Jiang S, Wang H, Ming F. The MYB transcription factor RcMYB1 plays a central role in rose anthocyanin biosynthesis. Hortic Res. 2023;10(6):uhad080.

20. Jiang LY, Yue ML, Liu YQ, Zhang NT, Lin YX, Zhang YT, Wang Y, Li MY, Luo Y, Zhang Y, et al. A novel R2R3-MYB transcription factor FaMYB5 positively regulates anthocyanin and proanthocyanidin biosynthesis in cultivated strawberries (*Fragaria × ananassa*). Plant Biotechnol J. 2023;21(6):1140–58.

21. Jiao TM, Huang YP, Wu YL, Jiang T, Li TT, Liu YZ, Liu YC, Han YY, Liu YJ, Jiang XL, et al. Functional diversity of subgroup 5 R2R3-MYBs promoting proanthocyanidins biosynthesis and their key residues and motifs in tea plant. Hortic Res. 2023;10(8):uhad135.

22. Li JY, Liu SQ, Chen PF, Cai JR, Tang S, Yang W, Cao FR, Zheng P, Sun BM. Systematic analysis of the R2R3-MYB family in *camellia sinensis*: evidence for galloylated catechins biosynthesis regulation. Front Plant Sci. 2022;12:782220.

23. Tang N, Cao ZY, Yang C, Ran DS, Wu PY, Gao HM, He N, Liu GH, Chen ZX. A R2R3-MYB transcriptional activator LmMYB15 regulates chlorogenic acid biosynthesis and phenylpropanoid metabolism in *Lonicera macranthoides*. Plant Sci. 2021;308:110924.

24. Millar AA, Gubler F. The Arabidopsis GAMYB-like genes, MYB33 and MYB65, are microRNA-regulated genes that redundantly facilitate anther development. Plant Cell. 2005;17(3):705–21.

25. Tominaga R, Iwata M, Okada K, Wada T. Functional analysis of the epidermal–specific MYB genes *CAPRICE* and *WEREWOLF* in *Arabidopsis*. Plant Cell. 2007;19(7):2264–2277.

26. Zhang ZB, Zhu J, Gao JF, Wang C, Li H, Zhang HQ, Zhang S, Wang DM, Wang QX, Huang H, et al. Transcription factor *AtMYB103* is required for anther development by regulating tapetum development, callose dissolution and exine formation in Arabidopsis. Plant J. 2007;52(3):528–38.

27. Cheng H, Song SS, Xiao LT, Soo HM, Cheng ZW, Xie DX, Peng JR. Gibberellin acts through jasmonate to control the expression of *MYB21*, *MYB24*, and *MYB57* to promote stamen filament growth in *Arabidopsis*. PLoS Genet. 2009;5(3): e1000440.

28. Ding ZH, Li SM, An XL, Liu X, Qin HJ, Wang DW. Transgenic expression of *MYB15* confers enhanced sensitivity to abscisic acid and improved drought tolerance in *Arabidopsis thaliana*. J Genet Genom. 2009;36(1):17–29.

29. Seo PJ, Xiang F, Qiao M, Park JY, Lee YN, Kim SG, Lee YH, Park WJ, Park CM. The MYB96 transcription factor mediates abscisic acid signaling during drought stress response in Arabidopsis. Plant Physiol. 2009;151(1):275–89.

30. Park MY, Kang JY, Kim SY. Overexpression of AtMYB52 confers ABA hypersensitivity and drought tolerance. Mol Cells. 2011;31(5):447–54.

31. Lee HG, Seo PJ. The MYB96-HHP module integrates cold and abscisic acid signaling to activate the CBF-COR pathway in Arabidopsis. Plant J. 2015;82(6):962–77.

32. Cominelli E, Sala T, Calvi D, Gusmaroli G, Tonelli C. Over-expression of the Arabidopsis *AtMYB41* gene alters cell expansion and leaf surface permeability. Plant J. 2008;53(1):53–64.

33. Xu R, Wang Y, Zheng H, Lu W, Wu C, Huang J, Yan K, Yang GD, Zheng CC. Salt-induced transcription factor *MYB74* is regulated by the RNA-directed DNA methylation pathway in *Arabidopsis*. J Exp Bot. 2015;66(19):5997–6008.

34. Sun Y, Zhao J, Li X, Li Y. E2 conjugases UBC1 and UBC2 regulate MYB42-mediated SOS pathway in response to salt stress in *Arabidopsis*. New Phytol. 2020;227(2):455–72.

35. Zhang P, Wang RL, Yang XP, Ju Q, Li WQ, Lü SY, Tran LSP, Xu J. The R2R3-MYB transcription factor AtMYB49 modulates salt tolerance in *Arabidopsis* by modulating the cuticle formation and antioxidant defence. Plant Cell Environ. 2020;43(8):1925–43.

36. Yang A, Dai XY, Zhang WH. A R2R3-type MYB gene, *OsMYB2*, is involved in salt, cold, and dehydration tolerance in rice. J Exp Bot. 2012;63(7):2541–56.

37. Tang YH, Bao XX, Zhi YL, Wu Q, Guo YR, Yin XH, Zeng LQ, Li J, Zhang J, He WL, et al. Overexpression of a MYB family gene, *OsMYB6*, increases drought and salinity stress tolerance in transgenic rice. Front Plant Sci. 2019;10:168.

38. Zhu N, Cheng SF, Liu XY, Du H, Dai MQ, Zhou DX, Yang WJ, Zhao Y. The R2R3-type MYB gene *OsMYB91* has a function in coordinating plant growth and salt stress tolerance in rice. Plant Sci. 2015;236:146–56.

39. Smita S, Katiyar A, Chinnusamy V, Pandey DM, Bansal KC. Transcriptional regulatory network analysis of MYB transcription factor family genes in rice. Front Plant Sci. 2015;6:1157.

40. Kang L, Teng Y, Cen Q, Fang Y, Tian Q, Zhang X, Wang H, Zhang X, Xue D. Genome-wide identification of R2R3-MYB transcription factor and expression analysis under abiotic stress in rice. Plants. 2022;11:1928.

41. Du H, Liang Z, Zhao S, Nan MG, Tran LSP, Lu K, Huang YB, Li JN. The evolutionary history of R2R3-MYB proteins across 50 eukaryotes: new insights into subfamily classification and expansion. Sci Rep. 2015;5:11037.

42. Zheng XW, Yi DX, Shao LH, Li C. *In silico* genome-wide identification, phylogeny and expression analysis of the *R2R3-MYB* gene family in *Medicago truncatula*. J Integr Agric. 2017;16(7):1576–91.

43. Li Z, Peng RH, Tian YS, Han HJ, Xu J, Yao QH. Genome-wide identification and analysis of the MYB transcription factor superfamily in Solanum lycopersicum. Plant Cell Physiol. 2016;57(8):1657–77.

44. Hou QC, Wang LL, Qi YC, Yan TW, Zhang F, Zhao W, Wan XY. A systematic analysis of the subtilase gene family and expression and subcellular localization investigation of anther-specific members in maize. Plant Physiol Biochem. 2023;203:108041.

45. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018;35(6):1547–9.

46. Bilas R, Szafran K, Hnatuszko-Konka K, Kononowicz AK. Cis-regulatory elements used to control gene expression in plants.PLANT CELL TISS. ORG. 2016;127:269–87.

47. Ponce KS, Meng LJ, Guo LB, Leng YJ, Ye GY. Advances in sensing, response and regulation mechanism of salt tolerance in rice. Int J Mol Sci. 2021;22(5):2254.

48. Wang XP, Niu YL, Zheng Y. Multiple functions of MYB transcription factors in abiotic stress responses. Int J Mol Sci. 2021;22(11):6125.

49. Xiong HY, Li JJ, Liu PL, Duan JZ, Zhao Y, Guo X, Li Y, Zhang HL, Ali J, Li ZC. Over-expression of *OsMYB48-1*, a novel MYB-related transcription factor, enhances drought and salinity tolerance in rice. PLoS ONE. 2014;9(3):e92913.

50. Schmidt R, Schippers JHM, Mieulet D, Obata T, Fernie AR, Guiderdoni E, Mueller-Roeber B. MULTIPASS, a rice R2R3-type MYB transcription factor, regulates adaptive growth by integrating multiple hormonal pathways. Plant J. 2013;76(2):258–73.

51. Zhang YH, Wan YL, Sun XM, Yuan J, Zhao ZQ, Gao J, Wen XR, Tang FS, Kang MT, Abliz B, et al. Genome-wide identification of MDH family genes and their association with salt tolerance in rice. Plants. 2022;11(11):1498.

52. He ZH, Tian ZZ, Zhang Q, Wang ZB, Huang RK, Xu X, Wang YC, Ji XY. Genome-wide identification, expression and salt stress tolerance analysis of the GRAS transcription factor family in *Betula platyphylla*. Front Plant Sci. 2022;13:1022076.

53. Wang Y, Ying JL, Zhang Y, Xu L, Zhang WT, Ni M, Zhu YL, Liu LW. Genome-wide identification and functional characterization of the cation proton antiporter (CPA) family related to salt stress response in radish (*Raphanus sativus* L). Int J Mol Sci. 2020;21(21):8262.

54. Hobo T, Asada M, Kowyama Y, Hattori T. ACGT-containing abscisic acid response element (ABRE) and coupling element 3 (CE3) are functionally equivalent. Plant J. 1999;19(6):679–89.

55. Ito Y, Katsura K, Maruyama K, Taji T, Kobayashi M, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. Functional analysis of rice DREB1/CBF-type transcription factors involved in cold-responsive gene expression in transgenic rice. Plant Cell Physiol. 2006;47(1):141–53.

56. Jin XF, Xiong AS, Peng RH, Liu JG, Gao F, Chen JM, Yao QH. OsAREB1, an ABRE-binding protein responding to ABA and glucose, has multiple functions in *Arabidopsis*. BMB Rep. 2010;43(1):34–9.

57. Ouyang S, Zhu W, Hamilton J, Lin H, Campbell M, Childs K, Thibaud-Nissen F, Malek RL, Lee YD, Zheng L, et al. The TIGR Rice Genome Annotation Resource: improvements and new features. Nucleic Acids Res. 2007;35:D846–851.

58. Sakai H, Lee SS, Tanaka T, Numa H, Kim J, Kawahara Y, Wakimoto H, Yang CC, Iwamoto M, Abe T, et al. Rice Annotation Project Database (RAP-DB): an integrative and interactive database for rice genomics. Plant Cell Physiol. 2013;54(2):e6.

59. Tian F, Yang DC, Meng YQ, Jin JP, Gao G. PlantRegMap: charting functional regulatory maps in plants. Nucleic Acids Res. 2019;48(D1):D1104–13.

60. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer EL, Sonnhammer ELL, Tosatto SCE, Paladin L, Raj S, et al. Pfam: the protein families database in 2021. Nucleic Acids Res. 2021;49(D1):D412–9.

61. Poole RL. The TAIR database. In: Edwards, D, editors Plant Bioinformatics. Methods in Molecular Biology. Humana Press, 2005;406:179–212.

62. Chou KC, Shen HB. Plant-mPLoc: a top-down strategy to augment the power for predicting plant protein subcellular localization. PLoS ONE. 2010;5(6):e11335.

63. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. TBtools: an integrative toolkit developed for interactive analyses of big biological data. Mol Plant. 2020;13(8):1194–202.

64. Hu B, Jin JP, Guo AY, Zhang H, Luo JC, Ga G. GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics. 2015;31(8):1296–7.

65. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren JY, Li WW, Noble WS. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 2009;37:W202–208.

66. Letunic I, Bork P. Interactive tree of Life v2: online annotation and display of phylogenetic trees made easy. Nucleic Acids Res. 2011;39:W475–8.

67. Stephane R, Patrice D, Marc VM, Pierre R. PlantCARE, a plant cis-acting regulatory element database. Nucleic Acids Res. 1999;27(1):295–6.

68. Yoshida S. Routine procedure for growing rice plants in culture solution. In: Yoshida S, Forno DA, Cock JH, editors. Laboratory manual for physiological studies of rice. Los Baños: International Rice Research Institute; 1976. pp. 61–6.

69. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods. 2001;25(4):402–8.

Publisher's Note