Research Article

Alfin-like (AL) transcription factor family in *Oryza sativa* L.: Genome-wide analysis and expression profiling under different stresses

Jeba Faizah Rahman [a], Hammadul Hoque [a], Abdullah -Al- Jubayer [b], Nurnabi Azad Jewel [a], Md. Nazmul Hasan [a], Aniqua Tasnim Chowdhury [a], Shamsul H. Prodhan [a,*]

[a] *Department of Genetic Engineering and Biotechnology, Shahjalal University of Science and Technology, Sylhet, 3114, Bangladesh*
[b] *Department of Biotechnology and Genetic Engineering, Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Gopalganj, 8100, Bangladesh*

A B S T R A C T

Oryza sativa L. is the world's most essential and economically important food crop. Climate change and ecological imbalances make rice plants vulnerable to abiotic and biotic stresses, threatening global food security. The Alfin-like (AL) transcription factor family plays a crucial role in plant development and stress responses. This study comprehensively analyzed this gene family and their expression profiles in rice, revealing nine AL genes, classifying them into three distinct groups based on phylogenetic analysis and identifying four segmental duplication events. RNA-seq data analysis revealed high expression levels of OsALs in different tissues, growth stages, and their responsiveness to stresses. RT-qPCR data showed significant expression of OsALs in different abiotic stresses. Identification of potential *cis*-regulatory elements in promoter regions has also unveiled their involvement. Tertiary structures of the proteins were predicted. These findings would lay the groundwork for future research to reveal their molecular mechanism in stress tolerance and plant development.

1. Introduction

During their lifespan, plants encounter a range of challenging and ever-changing environmental circumstances, including drought, low temperatures, high salinity, and heat, along with being susceptible to various pathogens due to their sessile nature. The existence of various abiotic and biotic stresses represents an obstacle to the growth, maturation, and overall yield of plants [1]. To endure and enhance their tolerance to various stresses, plants have evolved intricate adaptation strategies at molecular, cellular, physiological, and biochemical levels [2–4]. One notable aspect of the plant stress response is the activation of a complex interwoven network of events involving interactions between genes, signaling molecules, and cross-talk within multiple molecular pathways [3,5–8]. Specific stresses trigger the expression of various stress-responsive genes [9], and stress-signaling pathways play a crucial role in helping plants adapt to environmental challenges by connecting the sensing mechanism with the genetic response [10–15]. Furthermore, the regulation of plant defense mechanisms is significantly influenced by different gene families that function as transcription factors (TFs). These

transcription factors are responsible for functionally regulating plants in terms of metabolism and physiology, enabling them to respond effectively to abiotic stresses and defend against invading pathogens [16–18]. Acting as molecular switches that control gene expression, they can govern critical downstream responses by regulating the expression of target genes that respond to stress [19,20]. Alongside other developmental and genetic responses [21,22], the function of these TFs stands out as a prominent and dynamic mechanism that aids higher terrestrial plants in adjusting to their surroundings [23]. To date, researchers have identified a variety of transcription factors in plants and categorized them into different families according to the conserved domains they possess [24]. Amongst them, the Alfin-like TF family holds significant importance in the realm of plants as it participates in a range of important biological processes such as root development and elongation, formation of meristem, and response to salt stress [25,26]. As a 7S storage protein, the *Alfin-like* gene was first identified in *Medicago sativa* (alfalfa) plant [27,28]. The AL TF family members share the distinctive feature of having two highly conserved regions: the DUF3594/Alfin domain at the N termini and the PHD-finger motif at the C termini. With

about 140 conserved amino acids, the Alfin domain is still poorly understood and has not been identified in prokaryotes, fungi, and animals [27]. Conversely, the C terminal PHD-finger motif, which consists of about 50 conserved amino acids, is present in both the animal and plant kingdoms and plays a pivotal role in governing epigenetic and transcriptional regulation mediated by chromatin [29]. The presence of the PHD-finger motif in the AL proteins indicates that this TF family likely serves fundamental biological roles in plants. Apart from Alfin1 protein in alfalfa, it is currently unknown how these TFs participate in biological processes. *Alfin1*'s cDNA was first isolated from salt-tolerant alfalfa cells [30], and subsequent research revealed that the Alfin1 protein interacts with specific DNA sequences [31]. In alfalfa, *Alfin1* plays a role in regulating the expression of the salt-inducible *MsPRP2* gene, and transgenic alfalfa plants that have been genetically modified to overexpress *Alfin1* demonstrate enhanced salt tolerance, indicating *Alfin1*'s role as a transcriptional regulator in plants. [32]. Furthermore, these transgenic plants exhibit a notable enhancement in root growth, regardless of whether they are exposed to typical or saline conditions, which suggests that *Alfin1* plays a role in root growth [25].

Numerous studies have elucidated the function of the *AL* gene family in plants. In the model plant *Arabidopsis thaliana*, there are seven *AL* genes. Specifically, when *AtAL7* is either overexpressed or subjected to T-DNA insertion mutants, it has been observed to have a detrimental impact on salt tolerance during the early stages of seedling development [33]. Additionally, it has been discovered that *AtAL6* plays a crucial role in regulating the elongation of root hairs in *Arabidopsis*, particularly under conditions of phosphate deficiency [26]. In *Brassica rapa*, the fifteen identified *BrAL* genes have exhibited positive responses when subjected to cold, salt, and drought stresses and, notably, following infection with the fungus *Fusarium oxysporum f.*sp. *conglutinans*, ten members of the *BrAL* gene family (e.g., *BrAL1–4, 7, 9–10, 13–15*) have shown significantly elevated levels of expression in response, indicating their involvement in the plant's response to both abiotic stresses and fungal infection [34]. Furthermore, twelve *AL* genes were found within the *Brassica oleracea* genome, with *BoAL8* and *BoAL12* displaying notable responsiveness when exposed to various abiotic and biotic stresses. Additionally, the *AL* genes in *B. oleracea* exhibited significant expression levels following inoculation with *Pectobacterium carotovorum* subsp. *Carotovorum*, highlighting their crucial roles in the plant's defense mechanisms [35]. Similarly, in *Zea mays*, a comprehensive study has unveiled the presence of eighteen *AL* genes [36]. These *ZmAL* genes exhibit notable expressions in response to various environmental challenges, including salinity, drought, and cold stress, further underscoring their significance in maize's adaptive response. In addition, the expression of *AhAL*, sourced from the stress-tolerant species *Atriplex hortensis*, has been investigated in transgenic *Arabidopsis* plants. The results indicate that *AhAL* enhances the plant's tolerance to salt and drought conditions, demonstrating its potential as a valuable genetic resource for improving crop stress resilience [37].

Rice, scientifically known as *Oryza sativa* L., holds a paramount position in global trade as one of the most secured food crops, serving as a vital dietary staple for over 50 % of the world's populace [38]. As the global population continues to grow, ensuring food security necessitates increased rice production with a dual focus on cost-efficiency and high yields. However, the advancement of rice production faces impediments stemming from an array of stress factors. The complete genome sequence of rice was published long ago [39] has paved the way for a profound understanding of individual gene families' functionalities and further their utilization to improve the stress tolerance and productivity of rice. The Alfin-like transcription family has drawn significant attention in the realm of plant research, with studies conducted in Alfalfa, *Arabidopsis*, and crop plants such as *B. rapa, B. oleracea*, and *Zea mays*. Diverse functions and profound significance attributed to this TF family in crop plants have spurred our interest in an in-depth research endeavour.

This work extensively explores the *OsAL* transcription factor family

in rice. Within the scope of the study- chromosomal mapping, duplication events, gene structure, conserved motifs, phylogenetic relationship, *cis*-regulatory elements, and miRNAs targeting the *OsAL* TF genes were investigated. The tertiary structure of the OsAL proteins was also predicted in this study. In addition, to gain insights into their functions, RNA-Seq and RT-qPCR data were used to analyze their expression profiles across different tissues, developmental stages, and under biotic and abiotic stresses. Furthermore, microarray data was used in this work to investigate their expression patterns in response to different hormonal treatments. This study enables the understanding of the functional roles associated with *OsAL* genes in rice across various physiological and environmental contexts.

2. Materials and methods

2.1. Screening of Alfin-like gene family in the rice genome and their characterization

The rice AL proteins were found by conducting a BLASTP search against the *Oryza sativa* v7.0 genome dataset in the Phytozome v13 database (https://phytozome-next.jgi.doe.gov/info/Osativa_v7_0) [40]. The *A. thaliana* AL protein sequences were used as queries. We obtained the protein sequences of seven *AtAL* genes (At5g05610, At3g11200, At3g42790, At5g26210, At5g20510, At2g02470, At1g14510) [16] from the TAIR database (https://www.arabidopsis.org/browse/genefamily/Alfinlike.jsp) [41]. To ensure thorough identification, these AtAL protein sequences were also employed in separate BLASTP searches within the Rice Genome Annotation Project (RGAP) database [39] and the Protein database of NCBI (https://www.ncbi.nlm.nih.gov/protein). All candidate members retrieved from these databases underwent further analysis in the Pfam (http://pfam.xfam.org/) [42] and NCBI conserved domain database (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) [43] to validate the presence of the characteristic conserved regions associated with AL proteins. Any sequence without the Alfin/DUF3594 domain and the PHD zinc-finger motif were eliminated. The proteins were designated as 'OsAL' with the 'Os' prefix denoting *Oryza sativa*. They were numbered sequentially based on their chromosomal positions (from top to bottom), following a previously described naming convention [34,36]. Information about the identified genes' locus IDs, strand positions, coding sequence (CDS) coordinates (5′ to 3′), gene lengths and CDS lengths were collected from Phytozome v13 (https://phytozome-next.jgi.doe.gov/) database [40]. Physicochemical properties like molecular weight and theoretical isoelectric point of the proteins were calculated using the ExPASy ProtParam (https://web.expasy.org/protparam/) [44] tool. Furthermore, to predict their subcellular localization, CELLO v.2.5 (http://cello.life.nctu.edu.tw/) [45] and WoLF PSORT (https://www.genscript.com/wolf-psort.html) [46] web tools were employed. In addition, the similarities in the protein sequences of *OsALs* were searched through a multiple sequence alignment performed in the MEGA X [47] using the ClustalW alignment function with default parameters. Then, the software GeneDoc version 2.7 [48] was used to visualize the alignment result.

2.2. Phylogenetic analysis

Using the full-length amino acid sequence of *OsALs*, the Neighbor-joining method with 1000 bootstrap replicates was used to construct a phylogenetic tree in MEGA X [47]. To examine the evolutionary relationship of AL proteins of different plant species, the complete identified AL protein sequences of *Arabidopsis thaliana* [33], *Zea mays* [36], *Brassica rapa* [34]*,* and Alfin-1 of alfalfa (*Medicago sativa*) were used. Amino acid sequences of *Zea mays*, and *Brassica rapa AL* genes were downloaded from Phytozome, and the *Alfin-1* gene of alfalfa was obtained from the NCBI database (https://www.ncbi.nlm.nih.gov/) under the accession: AAA20093.2. These sequences were used to create an unrooted phylogenetic tree in MEGA X by employing the

Neighbor-joining method and 1000 bootstrap replicates. The previously reported *Arabidopsis thaliana AtAL* classification [33] was used to categorize the *OsAL* genes into various groups. The tree was then annotated, manipulated, and displayed via the online tool iTOL (https://itol.embl.de/) [49].

2.3. Analysis of gene structure and conserved motifs of OsALs

The web tool Gene Structure Display Server-GSDS 2.0 (http://gsds.gao-lab.org/) [50] was used to analyze the intron-exon organization of the *OsALs*. To find the conserved motifs, another web tool, the Multiple Expectation Maximization for Motif Elicitation-MEME version 5.4.0 (https://meme-suite.org/meme/tools/meme) [51] was employed. All the parameters were left at their defaults, except motif number, which was set to 15.

2.4. Chromosomal mapping and gene duplication study

MapChart software [52] was employed to generate a chromosomal distribution diagram to pinpoint the location of the *OsAL* genes across the 12 rice chromosomes. The data retrieved from the Phytozome database was used for this mapping, with particular reliance on the CDS coordinate information provided in Table 1. To understand the evolutionary history of these genes, gene duplication information for rice was obtained from the Plant Genome Duplication Database (PGDD) (http://chibba.agtec.uga.edu/duplication/index/downloads) [53]. Instances, where gene pairs exhibited over 90 % sequence similarity, were considered segmental duplications, and tandem duplications were defined as two or more homologous genes within a 100 kb region on the same chromosome [54]. To ascertain the evolutionary dynamics of these duplicated genes, the synonymous substitution rate (Ks) and non-synonymous substitution rate (Ka) data were gathered from the same database, and then using these values, the evolutionary constraint, represented by the Ka/Ks ratio, was calculated. An approximate timing of the duplication event was computed following the formula: T million years (Mya) = Ks/2λ. For dicotyledonous plants, λ is defined as a fixed rate of 1.5×10^{-8} substitutions per site per year [55]. Furthermore, to provide a visual representation of the paralogous gene pairs, a circular plot was generated using the TBtools program [56].

2.5. Cis-regulatory elements CREs) analysis of OsAL genes

To find out the potential CREs within the promoter of *OsAL* genes, the PlantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) [57] was used with default parameters and a 2.0 kb genomic sequence segment, upstream of the "ATG" start codon of each gene was searched. Then, different groups were created to categorize the identified *cis*-elements. Further, a visual representation of the essential CREs that respond to biotic and abiotic stresses, cellular development, and hormonal regulation was generated via TBtools software [56].

2.6. Finding miRNAs that target OsAL genes

psRNATarget (https://www.zhaolab.org/psRNATarget/) [58] was used to identify miRNAs that target *AL* genes in rice. The analysis was conducted on the mature miRNAs of rice listed in miRbase (https://www.mirbase.org/) [59]. Further, the connections between these miRNAs and their target *OsAL* genes were graphically depicted using the Cytoscape software [60]. The PmiRExAt (http://pmirexat.nabi.res.in/index.html) [61] database was then used to explore their expression patterns to understand their functional significance.

2.7. Structure analysis of OsAL proteins

The STRIDE program (https://webclu.bio.wzw.tum.de/stride/) [62]

Table 1
Characteristics of the Alfin-like (AL) transcription factor family in *Oryza sativa* L.

Sl No	Gene name	Locus id	Transcript	Chr No	Strand	CDS coordinate (5' to 3')	Length (bp)			Exons	Protein (aa)	MW (KDa)	pI	Subcellular localization
							Gene	cDNA	CDS					
1	*OsAL1*	LOC_Os01g66420	LOC_Os01g66420.1*	1	R	38,566,147–38,570,861	4715	1298	819	5	272	29.6	5.47	Nuc[a], Cyt[b], Mit[b]
			LOC_Os01g66420.2	1	R	38,566,147–38,570,861	4715	1295	816	5	271	29.48	5.47	Nuc[a,b], Cyt[b], Mit[b]
2	*OsAL2*	LOC_Os02g35600	LOC_Os02g35600.1	2	F	21,413,884–21,417,221	3338	1238	804	5	267	29.23	6.03	Nuc[a], Cyt[b]
3	*OsAL3*	LOC_Os03g60390	LOC_Os03g60390.1	3	R	34,342,899–34,346,745	3847	1459	744	5	247	27.78	5.71	Nuc[a], Ext[a], Cyt[b]
4	*OsAL4*	LOC_Os04g36730	LOC_Os04g36730.1	4	F	22,189,014–22,191,831	2818	1289	771	5	256	28.28	6.02	Nuc[a], Ext[a], Cyt[a], Chl[b]
5	*OsAL5*	LOC_Os05g07040	LOC_Os05g07040.1	5	R	3,697,729–3,702,359	4631	1294	777	5	258	28.87	5.52	Nuc[a,b]
6	*OsAL6*	LOC_Os05g34640	LOC_Os05g34640.1	5	F	20,542,370–20,547,114	4745	1236	777	5	258	29.1	5.44	Nuc[a], Cyt[b], Mit[b]
7	*OsAL7*	LOC_Os07g12910	LOC_Os07g12910.1*	7	R	7,414,568–7,418,059	3492	1161	735	5	244	27.4	6.37	Nuc[a], Ext[a], Cyt[b]
			LOC_Os07g12910.2	7	R	7,415,667–7,418,059	2393	683	522	4	173	19.35	5.15	Ext[a], Cyt[b]
			LOC_Os07g12910.4	7	R	7,415,665–7,418,059	2395	690	375	4	124	13.86	4.77	Cyt[a,b], Mit[a]
8	*OsAL8*	LOC_Os07g41740	LOC_Os07g41740.1	7	F	25,013,506–25,017,814	4309	1302	837	5	278	30.41	5.66	Nuc[a], Mit[b]
9	*OsAL9*	LOC_Os11g14010	LOC_Os11g14010.1*	11	R	7,776,646–7,784,194	7549	1290	765	5	254	28.54	5.32	Nuc[a], Cyt[a], Mit[b]
			LOC_Os11g14010.2	11	R	7,776,646–7,783,855	7210	1136	519	4	172	19.33	6.23	Ext[a], Cyt[b]

* Primary transcripts; MW: Molecular Weight; pI: Isoelectric point; bp: base pair; aa: amino acid; kDa: kilodalton; Nuc: Nucleus; Cyt: Cytoplasm; Mit: Mitochondria; Ext: Extracellular; Chl: Chloroplast; R: reverse strand; F: forward strand.
[a] Localization prediction by CELLO v.2.5 (http://cello.life.nctu.edu.tw/).
[b] Localization prediction by WoLF PSORT (https://www.genscript.com/wolf-psort.html).

was used with default settings to predict the secondary structure of OsAL proteins. The number and percentage of turns, coils, 310 helices, extended beta sheets, and alpha helices were shown in different colors as part of the structural analysis. The transmembrane motif examination was performed in TMHMM v.2.0 web server (https://services.healthtech.dtu.dk/services/TMHMM-2.0/) [63] with default parameters.

The RoseTTAFold method was employed to create the OsAL proteins' tertiary structure in the Robetta web tool (https://robetta.bakerlab.org/) [64]. Subsequently, the generated structures underwent refinement through the GalaxyRefine server (https://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE) [65], and then Swiss-PdbViewer v4.1 (https://spdbv.unil.ch/) [66] was used to minimize energy. To ensure their quality, the predicted structures were validated using PROCHECK (https://saves.mbi.ucla.edu/) [67] and ERRAT (https://saves.mbi.ucla.edu/) [68] servers. ProSA-web server (https://prosa.services.came.sbg.ac.at/prosa.php) [69] was then employed to evaluate Z-scores and energy plots. Finally, the Biovia Discovery Studio program [70] was utilized for visualizing the three-dimensional structures.

2.8. OsALs gene expression analysis in rice using RNA-seq and microarray data

The Rice Genome Annotation Project database (http://rice.plantbiology.msu.edu/) [39] and Rice Expression Database (http://expression.ic4r.org/) [71] were used to collect the RNA-seq expression data of nine *OsAL* transcripts in different tissues during different periods. To explore their expression at various developmental phases and also under different biotic and abiotic stress conditions, GENEVESTIGATOR (https://genevestigator.com/) [72] was utilized to obtain the rice mRNA-seq data for these nine *OsAL* transcripts. Further, microarray expression data were also obtained from GENEVESTIGATOR to examine their expression in response to different hormones essential for plant growth. Finally, all the retrieved expression data were subjected to normalization and profiling in the GraphPad Prism 9.0.0 software (https://www.graphpad.com/company/) [73], allowing for a comprehensive analysis of expression across various contexts, including tissues, developmental stages, stress conditions, and hormone treatments.

2.9. Plant material, stress treatments, and total RNA isolation

BRRI dhan28, an *indica* rice variety, was used in this experiment. Bangladesh Rice Research Institute (BRRI) provided the seeds, and only the mature and healthy seeds were selected for use as explants and subsequent treatment. Following a thorough cleaning, the seeds were placed onto wet tissue paper in a petri dish to initiate germination. The germination of seedlings was followed by their transfer into a hydroponic cultivation system after 3–4 days. The culture system was controlled by 16 h of light and 8 h of darkness, at a temperature of 25 ± 3 °C and having 1500–2000 lux light intensity. 20 days old seedlings were treated with various stress conditions which include submergence (the seedlings were entirely kept under distilled water), drought (3 mg/L PEG), heat (42 °C), cold (4 °C), cadmium (100 mM Cd) and salinity (100 mM NaCl). The stress conditions were maintained for 18–20 h. As a control, untreated seedlings were employed. Following the treatments, whole seedlings were obtained and repeatedly rinsed in distilled water and 70 % alcohol before the extraction of RNA. Total RNA was then extracted using the Invitrogen™ TriZOL™ reagent (Thermo Fisher Scientific Corporation, USA). Further, the genomic DNA contamination was eliminated using DNase I from Thermo Fisher Scientific Corporation's (USA) Invitrogen™ DNA-free™ DNA Removal Kit. Subsequently, the complementary DNA (cDNA) of the mRNA from the extracted total RNA sample was synthesized by using the GoScript™ Reverse Transcription System of Promega Corporation, USA. In each of these procedures, the manufacturer's protocol was followed strictly.

2.10. RT-qPCR and expression profiling of OsALs in the context of different abiotic stresses

We examined the expression of each of the nine *OsAL* genes in response to several abiotic stress conditions using all of their transcripts through RT-qPCR. The NCBI Primer-BLAST(https://www.ncbi.nlm.nih.gov/tools/primer-blast/) [74] and the OligoAnalyzer Tool (https://sg.idtdna.com/pages/tools/oligoanalyzer) of Integrated DNA Technologies, Inc were used to design the primers (S7 table). The BioRad CFX96™ Real-Time PCR Detection System was used to conduct the real-time qPCR. In this study, the Promega Corporation, USA's GoTaq® qPCR Master Mix (2X) was used. *eEF-1α*, a eukaryotic elongation factor, was employed as a reference gene [75]. We added 7.5 μL of GoTaq® qPCR Master Mix (2X), 1 μL of gene-specific primer, 2 μL of diluted cDNA sample (10 times diluted), and 3.5 μL of nuclease-free water to each 15 μL reaction mixture. The following conditions were used to carry out each reaction: first denaturation for 10 min at 95 °C, then 40 cycles of denaturation for 15 s at 95 °C, annealing for 30 s, and extension for 40 s at 72 °C. For *OsAL1.1, OsAL2, OsAL3, OsAL6, OsAL7.2* and *OsAL7.3*, the annealing temperature was 60 °C; however, for *OsAL1.2, OsAL4, OsAL5, OsAL7.1, OsAL8, OsAL9.1, OsAL9.2* and *eEF-1α*, the annealing temperature was 57 °C. Three separate experiments were performed on each sample, and after PCR amplification, melting curve analysis was conducted. To determine each gene's relative expression ratio, we employed the delta-delta Ct value approach [76]. To determine the average expression value among the different treatments, technical replication was utilized. Software used for data analysis included Microsoft Office 2019 and GraphPad Prism 9.0.0 [73]. Following a one-way ANOVA and the Bonferroni post hoc test, significant differences ($P \leq 0.05$) were identified. The significant differences were indicated by assigning varying star ratings.

3. Results

3.1. Screening Alfin-like (AL) genes in the rice genome and their characterization

Nine *Alfin-like* TF genes were found in *Oryza sativa*. Upon performing BLASTP searches using the seven *AtAL*'s protein sequences as a reference against the *Oryza sativa* v7_JGI dataset on Phytozome [40], 28 hits were found. Out of 28 transcripts examined, 13 were verified to possess both the Alfin/ DUF3594 domain and the PHD-zinc finger domain through sequence analysis performed on Pfam [42] and NCBI conserved domain database [43]. These 13 transcripts belong to the nine *AL* genes in rice which have alternative splice forms. Then, further BLASTP searches on RGAP [39] and the NCBI protein database confirmed that no putative AL protein was missed. The Phytozome annotated primary transcripts were chosen as the representative transcripts for the genes in this current study. The chromosomal positions of the genes led to their naming as *OsAL1-OsAL9*. The positions of the conserved regions within the nine OsAL proteins are detailed in the S1 table.

According to sequence analysis, the deduced proteins had lengths ranging from 244 to 278 amino acids, and the lengths of the gene and CDS varied from 2818 bp (*OsAL4*) to 7549 bp (*OsAL9*) and 735 bp (*OsAL7*) to 837 bp (*OsAL8*), respectively. These proteins exhibited varying sizes. With 278 amino acids in length and a molecular weight (MW) of 30.41 kDa, *OsAL8* encodes the largest protein in the family; in contrast, *OsAL7* is the smallest, with 244 amino acids in length and a MW of 27.40 kDa. Moreover, the proteins' predicted isoelectric points (pI) varied from 5.32 (*OsAL9*) to 6.37 (*OsAL7*). The proteins exhibited an average MW of about 27.86 kDa and a pI of about 5.73. Subcellular localization prediction revealed that the majority of the OsAL proteins would be located in the nucleus and then cytoplasm, mitochondria, and extracellular space. All of the representative *OsAL* genes are listed in Table 1, along with their physicochemical characteristics.

According to multiple sequence alignment, all *OsAL* genes have high

sequence similarity and contain two highly conserved regions: an Alfin/DUF3594 domain and a PHD-finger motif at their N- and C-termini, respectively (Fig. 1). The PHD-finger motif, which spans approximately 50 amino acids, contains the conserved Zinc-coordinating residues, namely Cys4-His-Cys3, and the Alfin domain has an approximate length of 128 amino acids.

3.2. Phylogenetic analysis

A phylogenetic tree was built to examine the evolutionary connections among the *AL* genes from various species by using the alignment of full-length amino acid sequences of AL proteins from *Oryza sativa*, *Arabidopsis thaliana* [33], *Zea mays* [36], *Brassica rapa* [34] and Alfin1 of the Alfalfa. This unrooted tree was constructed through the Neighbor-joining method in MEGA X software. The mode and gap were set to Poisson correlation and pairwise deletion, respectively and 1000 bootstrap replicates as verification parameters. The AtALs classification [33] allowed for the division of the phylogenetic tree into four groups, denoted respectively as groups I, II, III and IV (Fig. 2). The 9 OsAL proteins belonged to groups I, II and IV. Group I have OsAL3, OsAL5, and OsAL7 proteins along with 6 ZmAL, 4 BrAL, and 2 AtAL proteins. OsAL8 belongs to Group IV. There was 1 AtAL, 1 ZmAL, and 3 BrAL proteins in Group IV. OsAL1, OsAL2, OsAL4, OsAL6, OsAL9 belongs to Group II. Group II has 11 ZmAL, 5 BrAL, Alfin1, and 2 AtAL proteins. There is no OsAL protein in group III.

3.3. Gene structure and conserved motif analysis

The evolution of a gene family critically depends on intron-exon structural divergence. Gene structure analysis revealed that every member of the *OsAL* TF family shared a common pattern with four introns and five exons (Fig. 3a). The distribution of exons and introns in *Arabidopsis* [33] and *Brassica rapa* [34] *AL* TFs is in line with these findings. In addition, most of the *AL* TF genes in *Zea mays* [36] also contain four introns and five exons. This intron-exon structure indicated a clear evolutionary relationship between them.

The web-based MEME tool [51] was utilized to examine the conserved motifs of the OsAL proteins as a way to explore further the motif distribution and diversity of *AL* genes in *Oryza sativa*. In this study,

15 conserved motifs were discovered and labeled as motifs 1–15 (Fig. 3b). Because of the high similarity between OsAL peptide sequences, most motifs were common among the members. Among the 15 motifs, 1–7 appeared in all OsAL proteins. Motifs 1, 3, 4 comprise the Alfin/DUF3594 domain, and motif 2 encodes the PHD- zinc finger motif. All proteins had motif 7 and 6 at their N-terminus and C-terminus, respectively. OsAL1, OsAL8, and OsAL5 all contain motif 14. OsAL2 and OsAL3 shared motif 12. OsAL9 and OsAL5 both had motif 13. Motif 8 was in OsAL2 and OsAL4. Motif 10 was in OsAL3 and OsAL5, whereas motif 11 was found in OsAL1 and OsAL9. Motif 9 appeared in OsAL3, OsAL4, and motif 15 in OsAL8, OsAL7. The functional annotations of these motifs have yet to be investigated.

3.4. Chromosomal mapping and gene duplication analysis

All the *OsAL* genes are randomly and unequally distributed across the 7 out of 12 chromosomes of rice (Fig. 4). Chromosomes 5 and 7 contain two *AL* genes apiece, followed by chromosomes 1, 2, 3, 4, and 11 with one *AL* gene each. There was no *AL* gene in chromosomes 6, 8, 9, 10 and 12. *OsAL5, OsAL7,* and *OsAL9* are located on the p arm, while the rest are on the q arm of their respective chromosomes. *OsAL7* and *OsAL9* are located towards the centromere.

The evolution of a plant has been greatly aided by gene duplication as it results in the emergence of new functions, and thus the expansion of different gene families. The analysis of *OsALs* gene duplication aimed to uncover the evolutionary history of the gene family in rice. Four segmentally duplicated gene pairs were identified, while no tandem duplication was found between the genes (Fig. 5 and S2 Table). To assess the selection pressure, Ka/Ks values were calculated; values greater than 1 denoted positive selection, values equal to 1 denoted negative selection, and values less than 1 denoted purifying selection [77]. The duplicated *OsAL* pairs' ka/Ks values were all less than 0.3, which indicates that purifying selection has been a driving force in their evolution. In addition, it is estimated that these duplication events occurred between 23.59 and 80.05 million years ago (Mya), with the *OsAL3/OsAL7* duplication being the most recent at 23.59 Mya. The *OsAL1/OsAL2* duplication likely corresponds to the oldest event at 80.05 Mya, with other events falling in between at approximately 33.76 and 23.97 Mya.
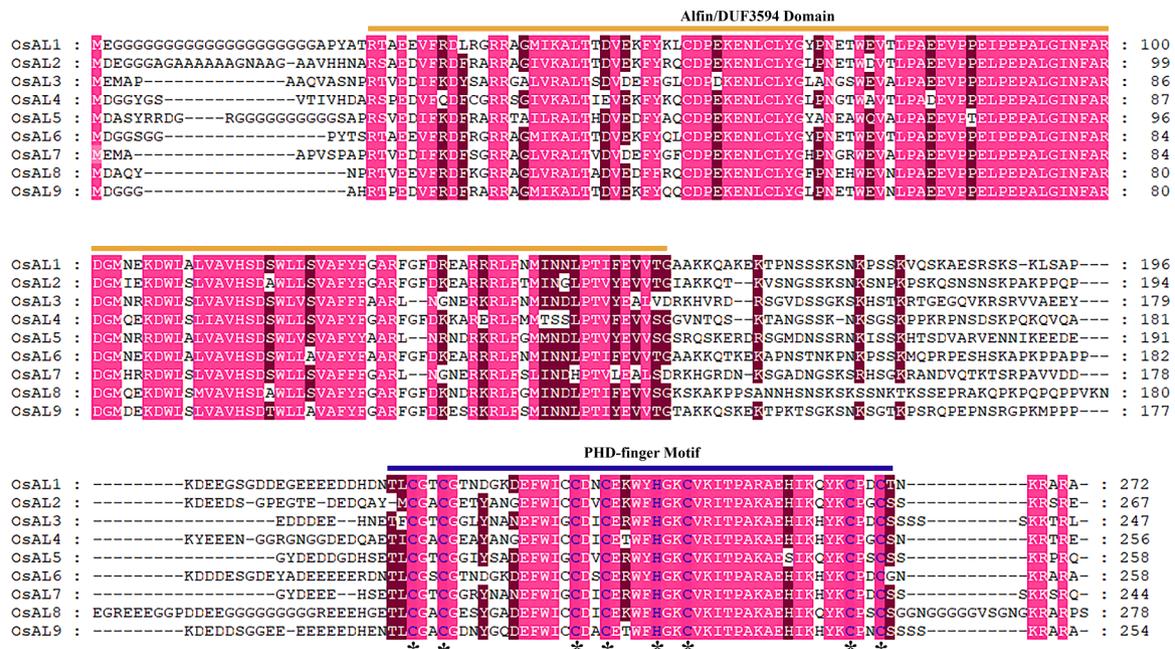

Fig. 1. Multiple sequence alignment of OsAL proteins. The Alfin/DUF3594 domain and PHD-finger motif are labeled, and the asterisks denote the conserved Cys4-His-Cys3 zinc finger motifs. Identical and similar amino acids are shaded as well.
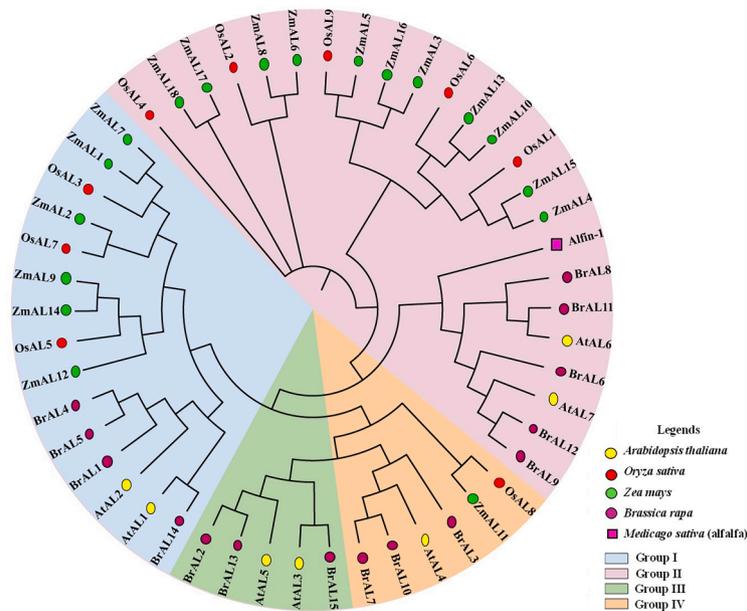
Fig. 2. Phylogenetic investigation of AL proteins in *Oryza sativa, Zea mays, Brassica rapa, Arabidopsis thaliana* and *Medicago sativa* (Alfalfa). Neighbor-joining method with 1000 bootstrap replicates was used to construct the tree in MEGA X. Based on AtAL classification, the tree was clustered into four major groups depicted in different colors.

When searching duplication events among rice *AL* genes and those of other species, it was observed that no duplication event was found in *Arabidopsis*, potato, wheat, tomato, cotton, grape, poplar, soybean, or sorghum except maize. Between rice and maize, four pairs of orthologous *AL* genes were found (Table 2). The range of Ka/Ks values for these pairs fell between 0.146 and 0.251, with a mean value of 0.210.

3.5. Cis-regulatory elements (CREs) analysis of OsALs' promoter region

CREs are pivotal in controlling gene expression. To explore the putative regulatory mechanisms of *OsALs* in response to various stimuli, the *cis*-elements in the promoter region of each *OsAL* gene up to 2.0 kb upstream from the translation start site were scanned in the PlantCARE database. This study identified a total of 70 CREs in the 5′ UTR promoter region of the *OsALs*. A list of all identified *cis*-regulatory elements, along with their corresponding function, is mentioned in the S3 table. These CREs were categorized into eight functional groups, including elements responsive to light (27 % of all identified CREs), abiotic stress (15 %), hormonal regulation (24 %), cellular development (10 %), promoter-related (3 %), biotic stress (4 %), miscellaneous functions (7 %). In addition, some CREs with unknown functions (constituting 10 % of the total identified CREs) were also found in the *OsAL* genes (S3 table). Fig. 6 depicts the abundance of hormonal regulation, cellular development, and stress -responsive CREs in the *OsALs*.

Abiotic stress-responsive elements encompass various elements such as ARE (anaerobic-response element), GC-motif (G-rich sequence motif), LTR (low-temperature response element), CCAAT-box, DRE (dehydration-response element) core, MYB, MYB-recognition site, MYC, and STRE (stress-responsive element), MBS (MYB binding site). They participate in anaerobic, anoxic, drought, salt, and cold stress responses. Additionally, biotic stress-responsive elements such as the W box (WRKY binding site), WRE3 (W box-related element 3), and WUN-motif (wound-responsive element) are primarily associated with fungal elicitors and wound responses. Furthermore, the TATA box and CAAT box are notably prevalent in all *OsAL* genes (S3 table), which are the essential core promoter elements. They mainly function in the promoter and enhancer regions.

Cis-regulatory elements that are responsive to cellular development are crucial for several processes, such as meristem expression, meristem-

specific and xylem-specific expression, circadian control, and secondary xylem development. These encompass the AAGAA-motif, CAT-box, CCGTCC-motif, AC-II, O_2-site, Circadian and CCGTCC-box elements. The light-responsive elements are the GA-motif, AE-box, Box 4, G-Box, G-box, GATA-motif, GT1-motif, MRE, Sp 1, TCT-motif, Box II, C-box, chs-Unit 1 m1, Gap-box, ATC-motif, ACE, TCCC-motif and LAMP-element (S3 table). Hormonal-regulation responsive elements comprise ABRE, TCA, ABRE3a, ABRE4, ABRE2, GARE-motif, TCA-element, TGA-element, CARE, p-box, CGTCA-motif, TGACG-motif, ERE, as-1, TATC-box, AuxRR-core. The Myb-binding site, MYB-like sequence, A-box, DRE1, and telo-box are examples of *cis*-elements with various functions. Furthermore, several unnamed *cis*-elements were found with unknown functions (S3 table). The abundance of *cis*-elements in the promoter region highly indicates that they play a role in *OsAL* gene transcriptional regulation. Additionally, an array of hormone and stress-related *cis*-elements suggests a direct correlation with the genes' activity under different stress conditions in rice. Hence, it is plausible that the *OsAL* TF genes are important in regulating rice stress responses.

3.6. Identification of miRNAs targeting the OsAL genes

MicroRNAs (miRNAs) serve as significant post-transcriptional regulators of the expression of genes in both animals and plants. Down-regulation of miRNAs implies increased expression of their target mRNA. We used the psRNATarget database [58] to identify the miRNAs that target the *AL* genes in rice. In total, 88 unique potential miRNAs of 19–24 nucleotide long targeting the *OsALs* were identified. The S4 table contains information on the identified miRNAs and their mode of inhibition.

Fig. 7 shows a schematic representation of miRNAs-*OsALs* interaction. Targeted regulatory relation between miRNAs and their *OsALs* revealed that *OsAL2* was targeted by the least number of miRNAs: osa-miR5815 and osa-miR1440b (Fig. 7). The highest number of miRNAs targeted *OsAL4*. The number of miRNAs targeting *OsAL* genes were 9 for *OsAL1*, 2 for *OsAL2*, 9 for *OsAL3*, 29 for *OsAL4*, 8 for *OsAL5*, 14 for *OsAL6*, 4 for *OsAL7*, 6 for *OsAL8*, 7 for *OsAL9* (Fig. 7). In addition, 7 miRNAs targeted more than one *OsAL* gene. These miRNAs include osamiR2101–5p, osa-miR5075, osa-miR1858a, osa-miR1858b, osa-miR2925, osa-miR5832, osamiR5814 (Fig. 7). osa-miR5832 target
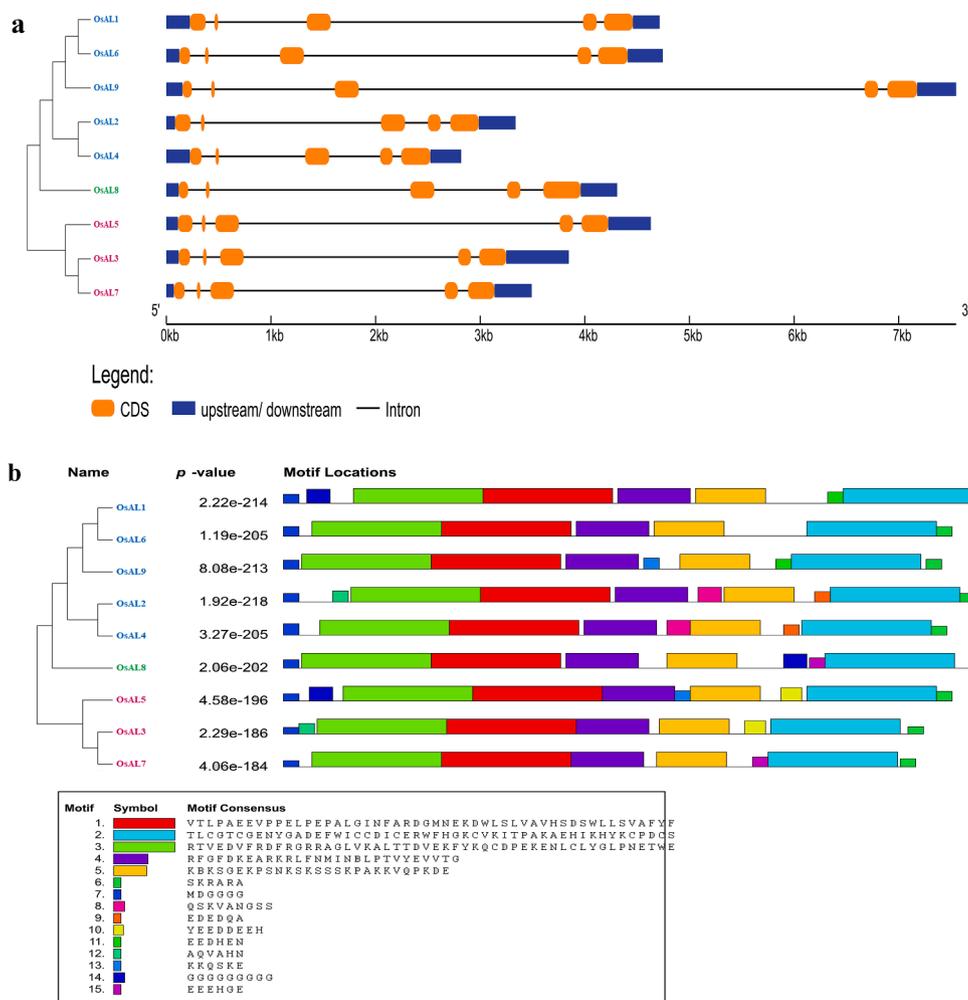
Fig. 3. Phylogeny, gene structure and conserved motifs in nine *OsALs* of rice. The phylogenetic tree generated via Neighbor-joining method shows evolutionary relationship among themselves. OsALs are labeled in pink, blue, and green colors to denote Group I, II and IV. (a) Gene structure of *OsALs*. Black lines represent introns and orange boxes represent exons. Blue boxes are used to symbolize untranslated regions (UTRs). The scale below can be used to estimate the lengths of exons and introns. (b) Conserved motifs of OsALs. The motifs are depicted by 15 different colored boxes. Protein sequence consensus for the corresponding motif is displayed in the legends below.
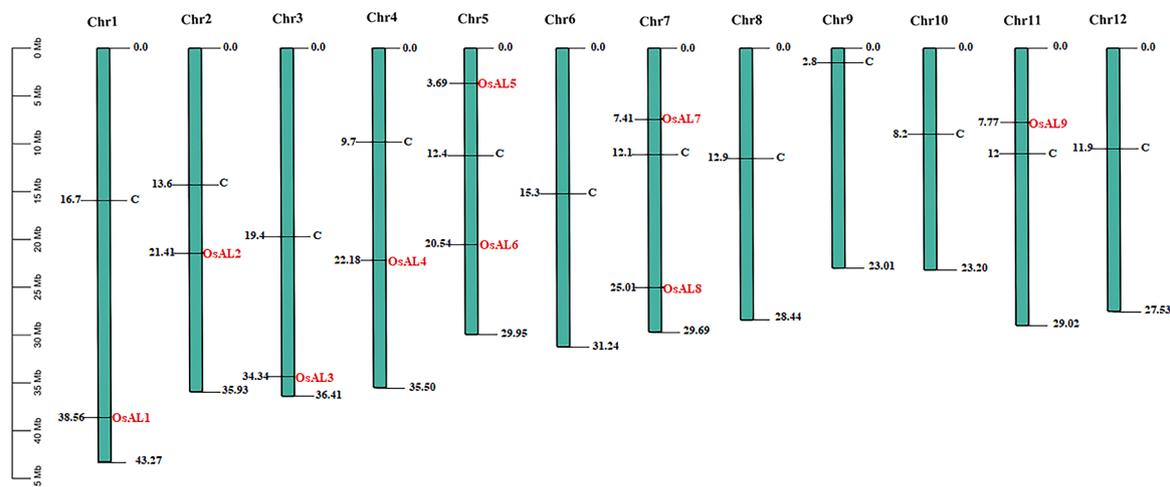

Fig. 4. Chromosomal distribution *OsAL* TF genes. The figure was generated via MapChart software. The chromosome number is indicated at the top, with the position of the centromere denoted by "C." By referencing the scale on the left, it's easy to compare the locations of each gene.
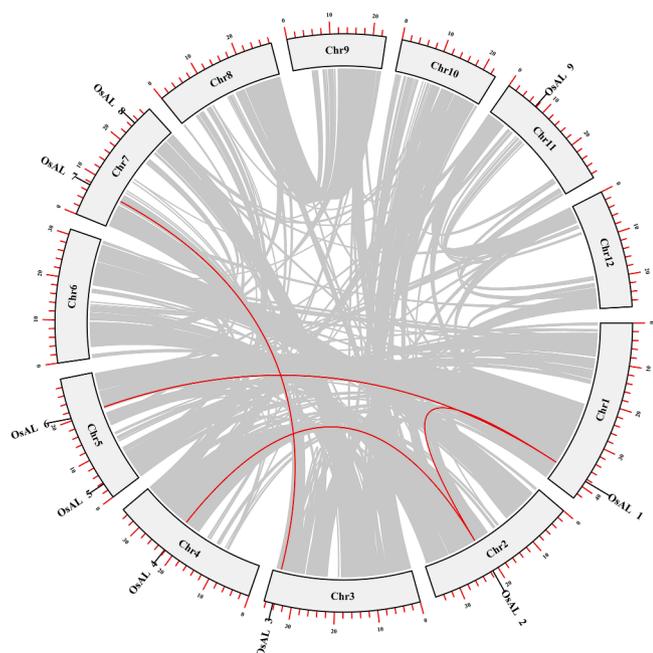
OsAL3, OsAL4 and *OsAL6*. osa-miR5075 target *OsAL1, OsAL5, OsAL4, OsAL9*. osa-miR5814 target *OsAL7* and *OsAL8*. osa-miR1858a target *OsAL1, OsAL5, OsAL6*. osa-miR2925 target *OsAL1, OsAL4, OsAL6*. osa-miR1858b target *OsAL1, OsAL5, OsAL6* (Fig. 7). Most miRNA-mediated *OsAL* repression involved mRNA cleavage, while a small fraction of targets was blocked at the translational level. Therefore, it suggests that the identified miRNAs may regulate *OsALs* expression post-transcriptionally by cleaving mRNA and inhibiting translation.

To investigate the expression patterns of the identified miRNAs, a heatmap (Fig. 8) was generated using the miRNAs' expression data from PmiRExAt [61] in different tissue types and abiotic stress conditions. For this study, miRNAs having an expected value of 5 were chosen. Among these miRNAs, osa-miR444b.2, which targets *OsAL5*, exhibited a high degree of expression in every tissue and under several stress conditions. Additionally, osamiR5797 which targets *OsAL3*, displayed prominent expression in anther, while in leaves, osa-miR812k targeting *OsAL6* exhibited significant expression. Also, significant expression of osa-miR812k was found under H_2O_2 stress response during seedling development. Furthermore, osa-miR408–5p targeting *OsAL9* was expressed in leaf, root, embryo, and endosperm, osa-miR1875 targeting *OsAL8* expressed in leaf and root, osa-miR1858a targeting *OsAL1, OsAl5, OsAL6* was expressed in endosperm. The majority of miRNAs exhibited decreased expression during various stress conditions. Downregulation of these miRNAs during stress conditions suggests the upregulated expression of *OsALs* in response to different stress environments.
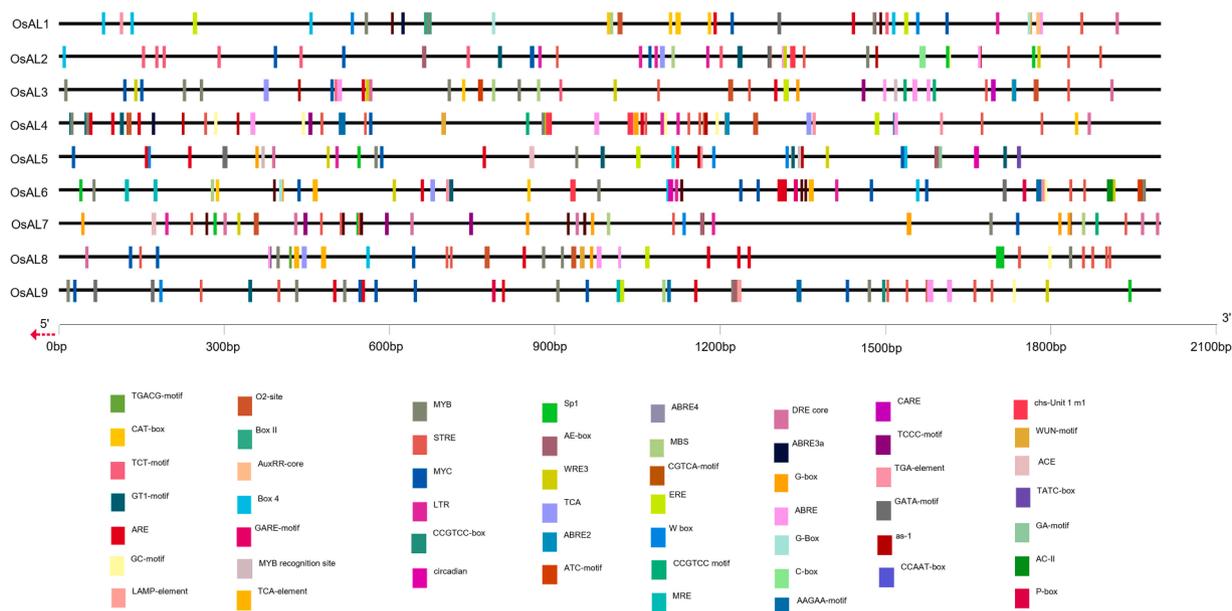

Fig. 5. Duplication events of OsALs within rice genome. The red lines indicate specific duplications of *OsALs*, while the gray lines represent duplications across the entire rice genome. TBtools was used to generate this plot.

Table 2
Duplication events between *AL* genes in rice and maize.

Duplicated gene 1	Duplicated gene 2	Ka	Ks	Ka/Ks	Duplication time (Mya)	Purifying selection	Duplicate type
OsAL3	AC225147.4_FG003	0.1016	0.6979	0.1455796	23.26	Yes	Segmental
OsAL5	GRMZM2G148810	0.2175	0.7432	0.2926534	24.77	Yes	Segmental
OsAL5	GRMZM5G893976	0.1188	0.7845	0.151434	26.15	Yes	Segmental
OsAL7	GRMZM2G047316	0.1263	0.5028	0.2511933	16.76	Yes	Segmental

Fig. 6. Schematic depiction of hormonal-regulation, cellular development and stress-responsive *cis*-regulatory elements (CREs) in *OsALs*. The scale at the bottom depicts the nucleotides upstream of the translation start site indicated by the red arrow. Distinct colors denote distinct CREs. Abiotic stress-responsive CREs: GC-motif, ARE, LTR, MBS, CCAAT-box, DRE core, MYB, MYB-recognition site, MYC, and STRE. Biotic stress-responsive CREs: W box, WRE3, and WUN-motif. Hormonal regulation CREs: ABRE, ABRE3a, ABRE4, ABRE2, GARE-motif, p-box, TCA, TCA-element, TGA-element, CARE, CGTCA-motif, TGACG-motif, ERE, as-1, TATC-box, and AuxRR-core. Cellular developmental CREs: CAT-box, O2-site, circadian, CCGTCC-motif, CCGTCC-box, AC-II, and AAGAA-motif.
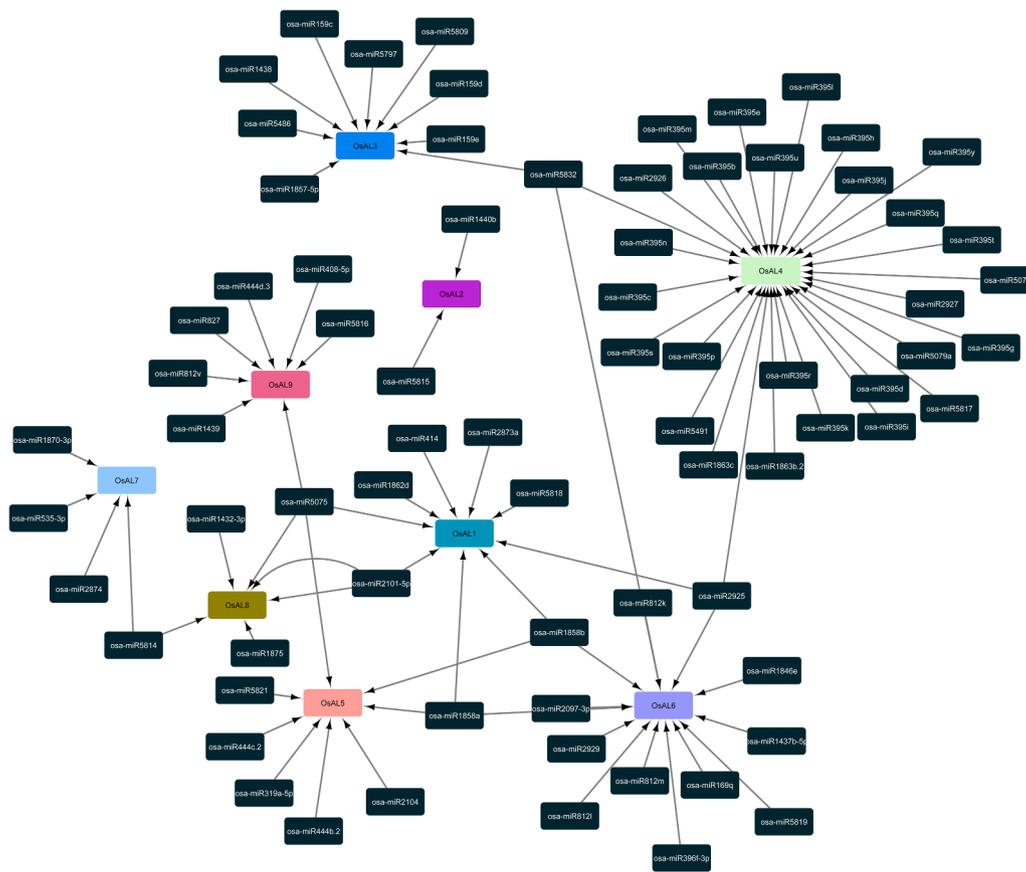
Fig. 7. Identification of putative miRNAs that target *OsALs*. Cytoscape was used to create the schematic representation of miRNAs-*OsALs* interaction. Black-colored boxes display the miRNAs, variously colored boxes correspond to distinct *OsALs*, and the arrows denote the regulatory relationship.

3.7. Analysis of protein structure

Secondary structure analysis of OsALs unveiled the positioning of various elements, including alpha-helix, beta-sheet, isolated beta bridge, turn, coil, 310-helix, and transmembrane helix (Fig. 9). The S5 Table provides the percentage of these elements and the location of the transmembrane helix. Notably, the most prevalent secondary structure is the alpha-helix, closely followed by the turn, coil, beta-sheet, and 310-helix. However, exceptions were observed in the case of OsAL1, 4, 7 and 9. In OsAL1 and 4, the coil percentage exceeded that of the turn, while in OsAL1, 7, and 9, the 310-helix percentage was greater than the beta-sheet percentage. The greater proportion of helical structures suggests the stability of OsALs [78], whereas structures like random coils play a crucial role in signaling cascades [79]. Furthermore, an analysis of membrane-spanning motifs (MSM) [64] revealed that only OsAL5 contains a single membrane-spanning motif.

3D models of the OsAL proteins were generated through the RoseTTAFold method of Robetta [64]. To enhance the precision of these models, we employed GalaxyRefine [65] to refine the resulting models and conducted energy minimization with Swiss-PdbViewer [66]. All of the OsAL proteins' modeled tertiary structures are shown in Fig. 10 and were visualized using Discovery Studio [70]. The validation process for these generated structures involved PROCHECK [67], ERRAT [68], and ProSA-web server [69] (as documented in the S6 table). The quality of the proteins was assessed through the PROCHECK's Ramachandran plot analysis. Remarkably, more than 90 % of the residues were found in the favored and additional allowed regions, with less than 1.5 % falling into the disallowed regions. This confirms the good quality of our projected models (S1 Fig). According to the analysis by ERRAT [68], the models demonstrated an overall quality factor of >87 (S2 Fig). Meanwhile, ProSA-web [69] provided the Z-score, a measure of how closely the

designed models resemble native structures (S6 Table), along with an energy plot depicting the local quality of the models. Notably, the Z-score fell within the range typically associated with native proteins, signifying the high quality of our generated structures, which also remained well within the range of X-ray crystal structures (S3 Fig). The energy plot (S4 Fig) displayed that all of the residues in the simulated structures had a lower energy value. These collective outcomes underscore the high standard of the modeled tertiary protein structures.

3.8. OsALs gene expression analysis in rice from the RNA-Seq and microarray data

Gene expression analysis examines how often or actively a protein is produced from its corresponding gene. This serves as a highly sensitive indicator of biological activity, as alterations in gene expression patterns correspond to changes in natural biological processes. Gene expression data of the 9 *OsAL* genes in various tissues were collected from the Rice Genome Annotation Project database [39] and the Rice Expression Database [71]. A heatmap was then generated using the normalized RNA-Seq FPKM expression values (Fig. 11).

The analysis revealed that *OsAL1, OsAL2, OsAL8,* and *OsAL9* had high expression in most tissues, while only *OsAL4* showed downregulation across all tissues. Genes with tissue-specific expression patterns were present in some cases. For example, *OsAL3* upregulated in anther, pistil (BF), and seed (5 DAP), while *OsAL7* upregulated in shoots, endosperm (25 DAP), seed (5 DAP), seed (10 DAP), and mature seed. In the same manner, *OsAL5* showed tissue-specific expression in anther and anther (BF). *OsAL2* showed expression in all the tissues except downregulation in roots, endosperm (25 DAP), and seed (10 DAP). *OsAL6* was highly expressed in roots, roots (7 days S), shoots, and seeds (5 DAP). *OsAL8* showed downregulation in root (14 days S), endosperm (25
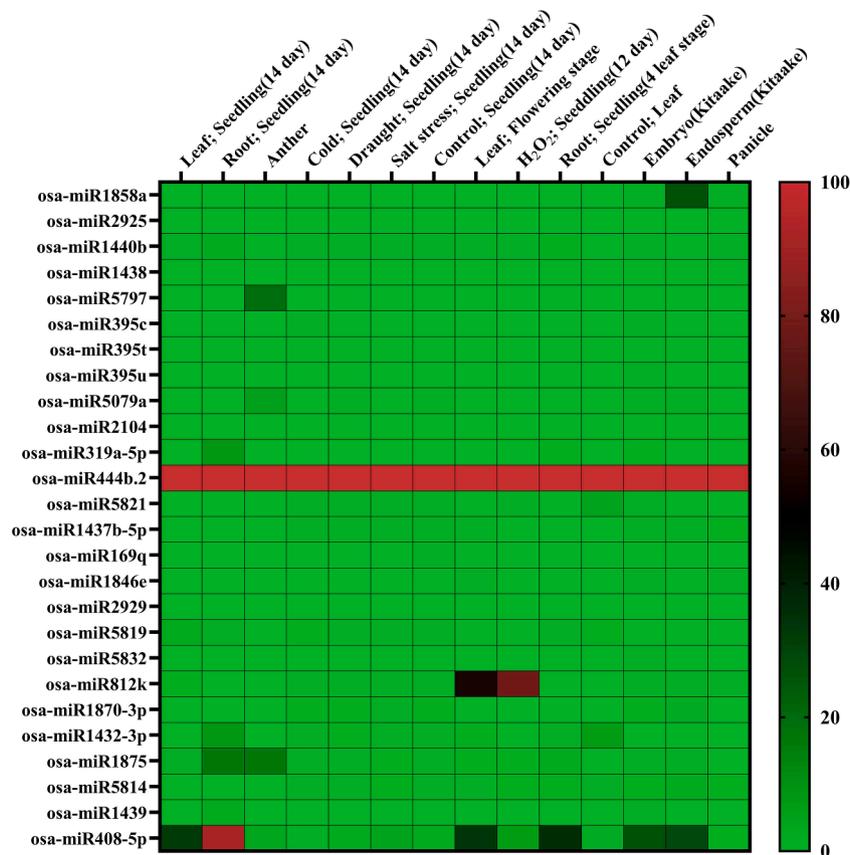
Fig. 8. The expression patterns of OsALs-targeting miRNAs in various tissues and abiotic stresses. GraphPad Prism 9.0.0 was utilized to construct the heatmap. The color scale on the right-side indicates the relative expression of each miRNA.

DAP), seed (10 DAP), and mature seed. In endosperm (25 DAP), all the genes were downregulated except *OsAL7*, and in seedling (4 LS), only *OsAL2* showed upregulation.

To assess *OsAL* expression during different developmental stages, pre-analyzed expression data from mRNA-seq expression values of ten developmental stages of rice were retrieved from GENEVESTIGATOR. Then, the GraphPad Prism 9.0.0 was used to generate a scatter plot with error bars (Fig. 12). The analysis showed that most genes were highly expressed during development, except for *OsAL4*, which had decreased expression in the early stages of development of the rice plant.

Furthermore, to gain insight into how *OsALs* respond to different types of stresses, the expression profiles of *OsALs* in response to both biotic and abiotic conditions were examined. To conduct the study, RNA-seq data retrieved from GENEVESTIGATOR was employed to create the corresponding heatmaps in GraphPad Prism 9.0.0. Figs. 13 and 14 demonstrate these heatmaps, showcasing their expression profiles under different biotic and abiotic stresses, respectively. The color scale in each figure represents Log_2 values of relative expression.

In reaction to an infection by the rice stripe virus (RSV), *OsAL4, 5, 6, 7,* and *9* exhibited an increase in their expression levels, while *OsAL1, 2,* and *3* were found to be downregulated (Fig. 13). In the case of bacterial leaf streak pathogen, *Xanthomonas oryzae* pv. *Oryzicola,* strains infecting rice, the expression of *OsAL3, 5,* and *7* were high, whereas *OsAL2* and *8* showed low expression. Furthermore, rice dwarf virus (RDV) infection led to the upregulation of *OsAL1* and *9*, while the remaining genes were found to be downregulated. It was also observed that *OsAL2, 3,* and *8* showed upregulation in response to *Rhizoctonia solani* infection that causes rice sheath blight, whereas *OsAL4, 7,* and *9* displayed downregulation. When the rice plant was exposed to rice blast fungus (*Magnaporthe oryzae*) infection, it increased *OsAL2, 3,* and *8* expression, suggesting their potential role to combat rice blast disease. These

findings addressed the involvement and significance of *OsALs* in controlling rice stress responses to different pathogens as well as biotic stresses.

The genes had distinct expression patterns with varying time duration heat treatments (Fig. 14). Except for *OsAL4*, all *OsALs* were downregulated after 30 and 60 min of heat treatment. In the case of a 120-minute heat treatment, *OsAL1, 2, 4, 6, 7,* and *8* were elevated. In addition, *OsAL2, 4* and *8* were upregulated after 165 min of heat treatment, whereas *OsAL1, 2,* and *4* were upregulated during 225 min of heat treatment. It was observed that *OsAL4* was highly expressed in heat stress of all time duration. Similar distinctive expression patterns were seen after dehydration (Fig. 14). Dehydration stress for 30 min increased *OsAL8* and *9* expression levels, whereas stress for 60 min increased *OsAL4, 8,* and *9*. In addition, *OsAL1* and *7* were upregulated in 90 min dehydration condition while a dehydration period of 135 min induced the expression of *OsAL1, 3, 4,* and *6*. During varying durations of anoxic conditions (Fig. 14), *OsAL3, 6,* and *8* were found to be upregulated. It was observed that the expression of *OsAL3* and *6* were proportional to the duration of the anoxic conditions, while *OsAL8* showed a constant upregulation. In addition, *OsAL4, 5* were expressed in 3 h of anoxic conditions, and *OsAL7* also exhibited a moderate level of expression in both 3 and 12 h of anoxic conditions. In the case of drought stress (Fig. 14), the exposure for 2 and 3 days caused upregulation of all *OsALs* except *OsAL4* and *8*. Under 8 days-long drought conditions, only *OsAL4* was upregulated, with the rest being in downregulation. Further, when exposed to 13 days-long drought stress *OsAL1, 3, 4, 5,* and *9* showed upregulation, and 20 days-long drought caused downregulation of all *OsALs* except *OsAL4*. In a drought condition of 26 days, *OsAL2* and *4* were highly expressed. In case of salt stress, 1 hour of salt treatment caused upregulation of *OsAL1, 2, 3, 4, 5, 6, 7* and *9* in the shoot, while 5 h of treatment caused upregulation of *OsAL3, 6, 7, 8,* and *9* (Fig. 14). In
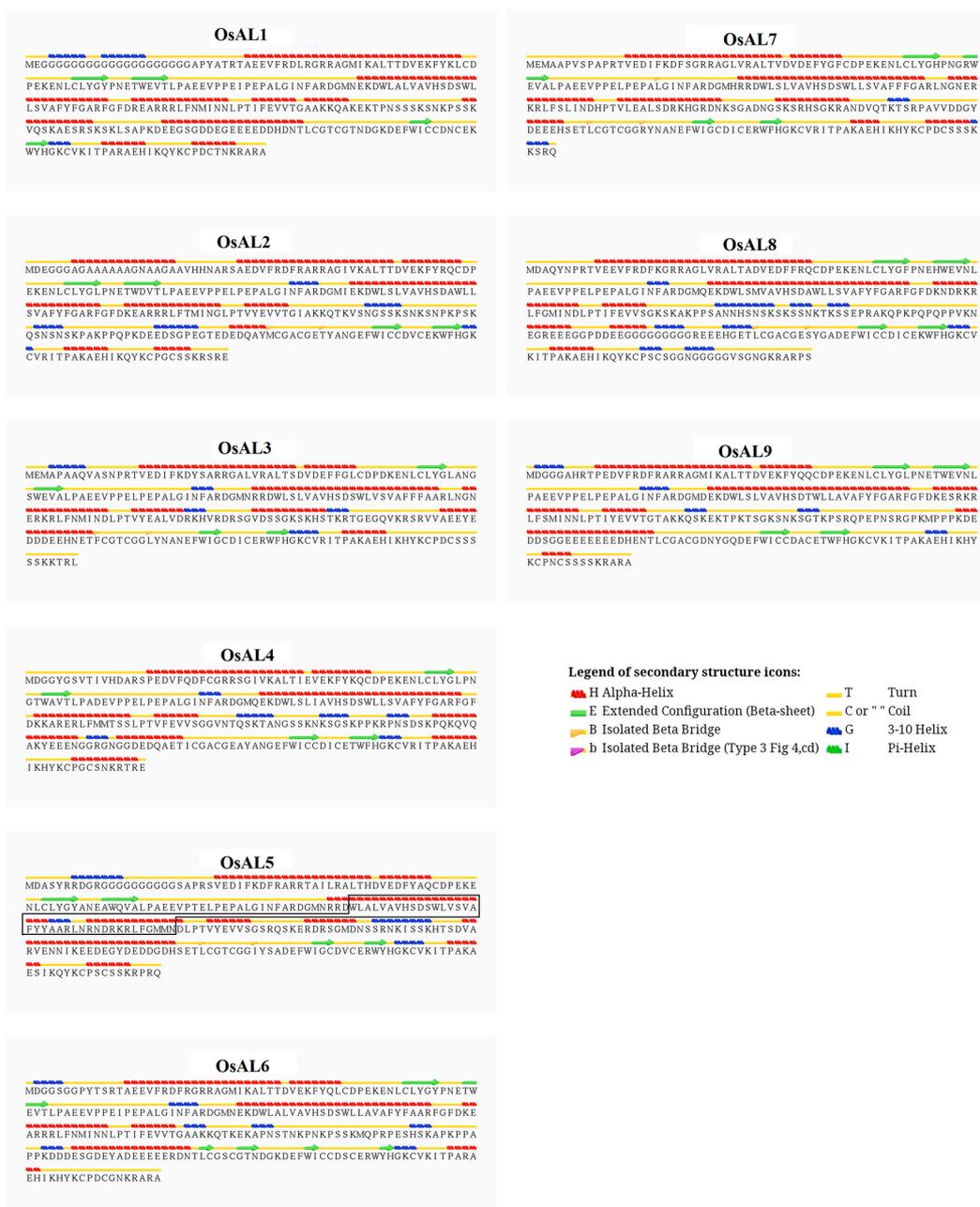
Fig. 9. Secondary structure analysis of rice OsAL proteins. STRIDE program was used to generate the structure, and TMHMM v2.0 server identified the membrane-spanning motif. The black box represents the transmembrane region. Legend of secondary structure icons are shown at the bottom-right corner.

addition, 24 h of treatment induced the expression of *OsAL1, 2, 3, 4* and *6* in shoot. It was also found that *OsAL1, 2, 3*, and *5* expression levels were high in the roots under salinity, whereas *OsAL1, 2, 3*, and *4* were elevated in the leaves. The analysis also observed that *OsAL1* and *3* were elevated, whereas others were downregulated during the submergence stress (Fig. 14). Cold treatment upregulated the expression of *OsAL1, 2, 4, 5, 6, 8* and *9*, whereas *OsAL1, 3, 6, 7, 8* and *9* were upregulated under alkali treatment (Fig. 14). The distinct expression patterns of *OsALs* under different stressful conditions revealed their role in regulating the stress response in rice.

Additionally, to explore their expression patterns in response to different hormones required for plant growth and development, such as abscisic acid, auxins, cytokinins, gibberellin, salicylic acid, ethylene, jasmonic acid, and kinetin, microarray expression data of nine *OsALs* were retrieved from GENEVESTIGATOR, and then represented in a heatmap via GraphPad Prism 9.0.0 (Fig. 15). The *OsALs* had distinct expression patterns. In response to indole-3-acetic acid (IAA), all *OsALs* exhibited elevated expression except *OsAL2*. Only *OsAL1* demonstrated upregulation under abscisic acid (ABA) treatment, while salicylic acid (SA) induced the high expression of *OsAL1* and *4*. Additionally, the direct precursor of ethylene, 1-amicocyclopropane-1-carboxylic acid (ACC), led to increased expression of *OsAL1, 4, 6*, and *7*. Only *OsAL7* was highly expressed upon exposure to jasmonic acid (JA). Notably, gibberellin (GA3) triggered the high expression of all *OsALs*. Both 1-naphthaleneacetic acid (NAA) and kinetin induced the upregulation of *OsAL3* and *OsAL5–9*. In addition, *OsAL2, 4, 6*, and *7* demonstrated elevated expression in response to 6-benzylaminopurine (BAP), while under the treatment of trans-zeatin (tZ), *OsAL3–7* exhibited significant upregulation.
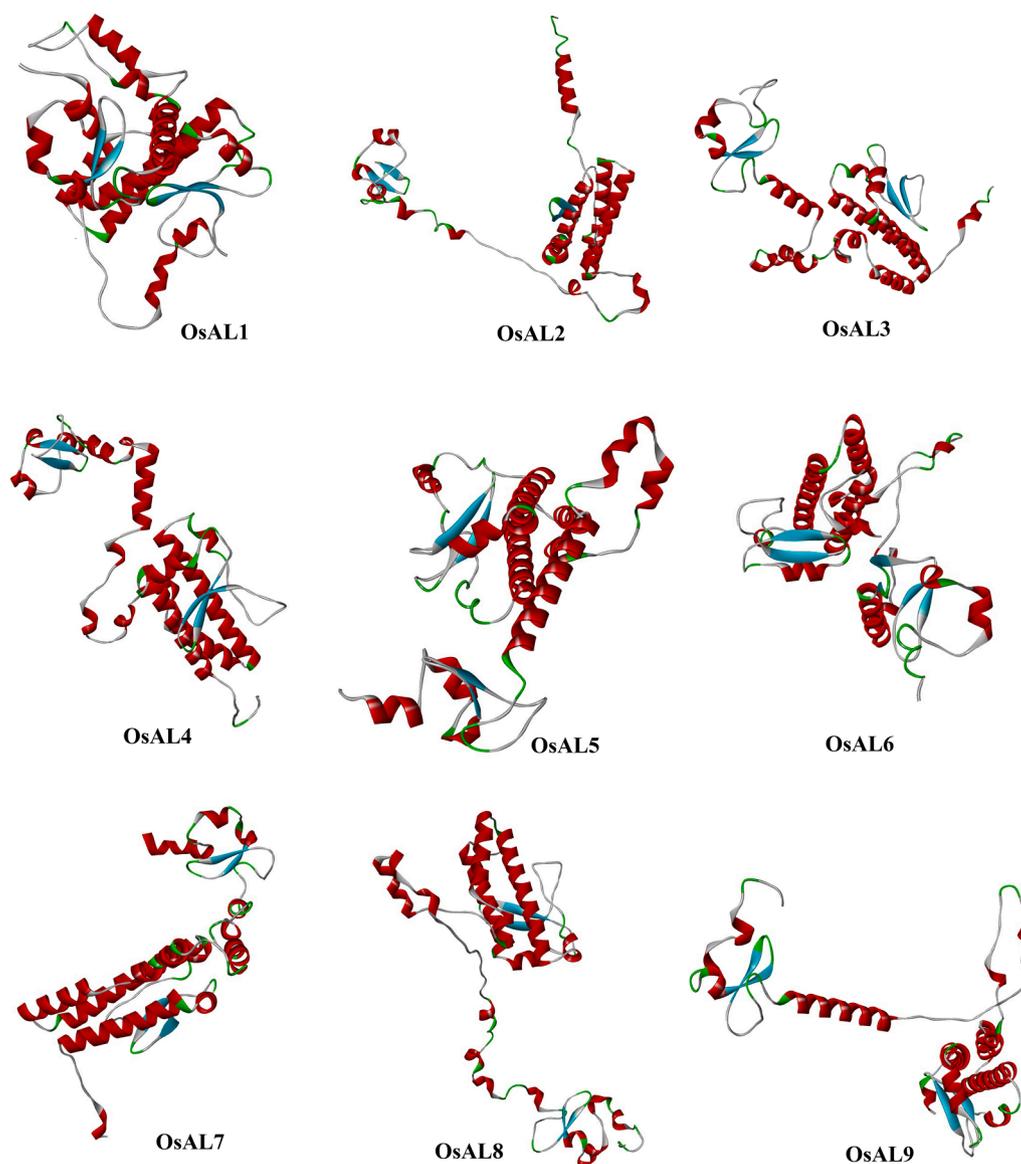
Fig. 10. Predicted 3D structures of OsAL proteins. The tertiary structures were constructed by Robetta online server, and then Biovia Discovery Studio Visualizer was used to visualize them.

3.9. OsALs gene expression analysis under multiple abiotic stresses from RT-qPCR data

The relative expression levels of *OsALs* were evaluated in rice seedling under a range of stress conditions such as heat, cold, cadmium, salt, submergence, and drought. Fig. 16 shows a heatmap of the real-time expression data of *OsALs* after being exposed to these stresses for 18 to 20 h.

As shown in Fig. 16, both transcripts of *OsAL1, OsAL1.1* and *OsAL1.2* exhibited significant higher expression under cold, submergence and drought stress. However, under salt stress, *OsAL1.2* showed a significant expression than *OsAL1.1*, and none of the transcripts had significant expression in response to heat stress. The RNA-seq data (Fig. 14) also demonstrated the increased expression of *OsAL1* under conditions of heat, drought, salt, submergence, cold, dehydration, and alkali stress. *OsAL2* showed significant upregulation in all the stress conditions except cadmium and heat, and the expression level was more than ten-fold in response to cold and submergence stress. The expression pattern of *OsAL2* obtained from the RNA-Seq data (Fig. 14) was also consistent with this RT-qPCR result, except for its expression under submergence

stress. *OsAL3* exhibited substantially elevated expression levels in cold, submerged, and drought environments, and the RNA-Seq data (Fig. 14) also showed consistent results. However, when exposed to salt stress, the gene significantly downregulated in RT-qPCR analysis, even though the RNA-Seq data showed an increase (Fig. 14). No significant results were observed for *OsAL4* and *OsAL5* expression under any of these stress conditions, and the expression patterns from the RNA-Seq data (Fig. 14) didn't show consistent results. *OsAL6* showed a nearly three-fold significantly higher expression under the conditions of cold and heat stress, and the RNA-Seq data (Fig. 14) also supported this finding. Although there was a higher expression under salt, submergence, and drought conditions, it was not significant. The RT-qPCR data analysis found that the transcripts of *OsAL7* showed different expression patterns under stress conditions. The primary transcript, *OsAL7.1,* exhibited substantial upregulation under cold and drought stress with no significant downregulation under other stresses. The other transcripts, *OsAL7.2* and *OsAL7.3,* showed significant downregulation in all of these stressful situations. However, the RNA-Seq data of *OsAL7* (Fig. 14) didn't show consistent results except for heat, drought, and salt stress for the primary transcript *OsAL7.1. OsAL8* showed significantly higher
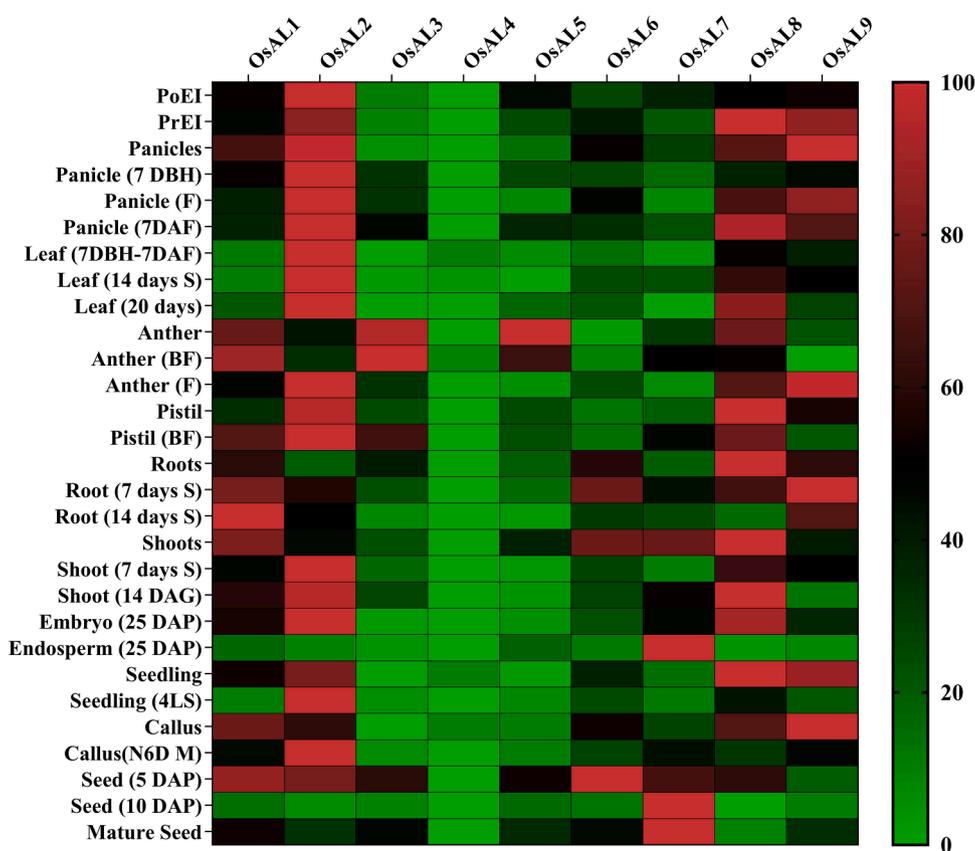
Fig. 11. Expression profiling of *OsALs* of rice in different tissues. RNA-seq data was retrieved from the Rice Expression Database and Rice Genome Annotation Project database. GraphPad Prism 9.0.0 was used to generate the heatmap. The color bar at right-side of the heatmap represents relative expression values, where red denotes a high level of expression, black a moderate level, and green a low level. Abbreviations: PoEI- post emergence inflorescence; PrEI- pre-emergence inflorescence; 7 DBH- 7 days before heading; F- flowering; 7 DAF- 7 days after flowering; DAG-days after germination; S- seedling; BF- before flowering; DAP- days after pollination; 4 LS- 4 leaf seedling.
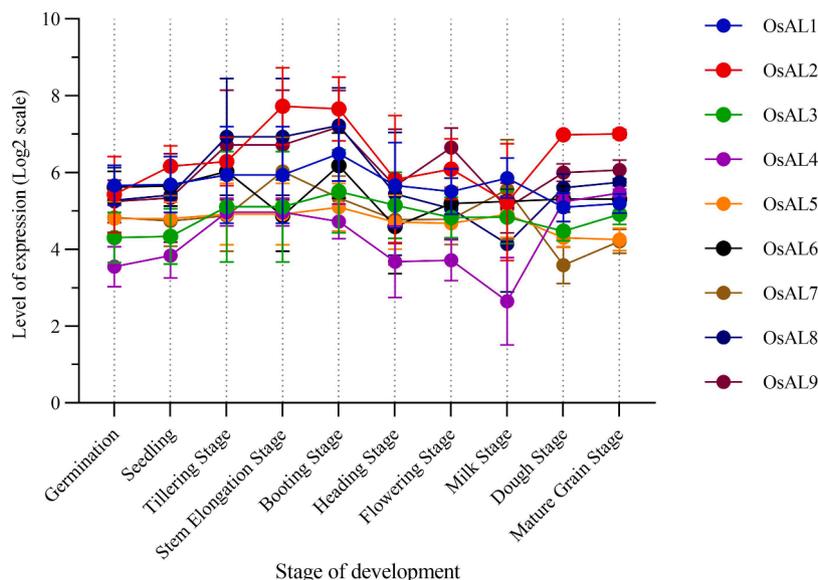

Fig. 12. A scatter plot illustrating the expression pattern of *OsALs* across ten developmental stages of rice plant. Pre-analyzed RNA-Seq data from GENEVESTIGATOR was used to generate the plot with error bar in GraphPad Prism 9.0.0. Different colored circles represent different *OsAL* genes. Ten developmental stages are also mentioned at the bottom. The scale on the left-side represents the level of expression in Log2 scale.

expression under all the stress conditions except cadmium and heat, and the relative expression level was four-fold under cold stress. However, under drought and submergence stress, the RNA-Seq data of Fig. 14

didn't demonstrate consistent results. Regarding the expression profile of *OsAL9*, it was shown that the primary transcript *OsAL9.1* exhibited more than five-fold significant higher expression level under cold stress.
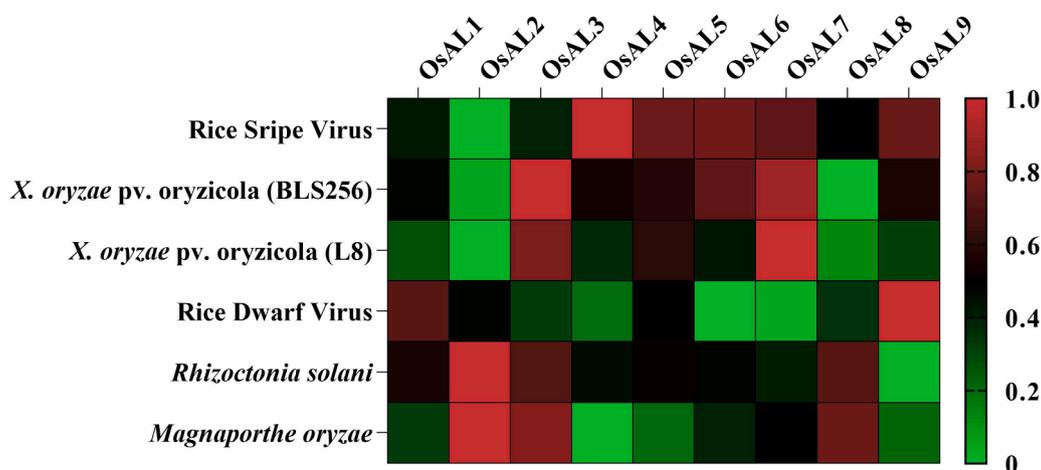
Fig. 13. Expression profiling of *OsALs* in response to different biotic stress conditions. RNA-Seq data from GENEVESTIGATOR was employed to create the heatmap in GraphPad Prism 9.0.0 software. The right-side scale of the heatmap represents relative expression values. High expression level is shown in red, low expression level in green, and moderate expression level in black.

Additionally, it was substantially upregulated in response to salt and drought stress. This finding was also supported by the RNA-Seq data (Fig. 14). Furthermore, a significant higher expression was observed in *OsAL9.2* under cold and submergence stress, and more than five-fold expression level was found in response to cold stress. The expression profiling under these stress environments using RT-qPCR data revealed that none of the *OsAL* genes showed significant upregulation or down-regulation in response to cadmium. Under cold stress, all the genes except *OsAL4, 5, 7.2*, and *7.3* showed significantly higher expression. Additionally, in response to salt, *OsAL1.2, 2, 8*, and *9.1* showed significant upregulation. Under heat stress, only *OsAL6* showed a significant higher expression level. It was also observed that under submergence and drought conditions, most of the genes were significantly expressed. This finding suggests the potential involvement of *OsALs* family in controlling rice's abiotic stress response.

4. Discussion

The Alfin-like transcription factor (TF) family exhibits apparent involvement in various physiological processes in plants. Extensive research has focused on understanding its features and functions in *Arabidopsis* [26,33] and crop plants maize and *Brassica* [34–36]. Rice is a primary staple meal for more than half of the world's population, with Asia being the primary contributor to global rice production [80]. However, rice production is facing challenges due to different types of biotic and abiotic stresses. The rising occurrence of these stresses emphasizes the necessity of developing rice varieties that can withstand and tolerate stresses. To address this, the complete and accurate annotation of genes and gene families, along with their functional exploration, especially their role in stress responses, is a crucial starting point to potentially modify and manipulate them for developing stress-tolerant rice varieties through genetic engineering to increase the annual yield worldwide.

As the *AL* gene family plays a role in plant development and stress responses, it is important to extensively investigate this gene family in rice. Therefore, this study was designed to conduct a comparative genome-wide analysis of the *AL* gene family in rice and to explore its potential involvement in responding to different stresses.

A total of nine *Alfin-like* TF genes in rice were found in this study, whereas ten *OsAL* TF genes had been reported previously [81]. All the genes possess both the Alfin domain and the PHD-zinc finger motif at their respective N and C termini [33,34] (Table S1). These genes vary in sequence length, molecular weight, and pI value, indicating that the gene family is diverse and the proteins were found to be localized in the

nucleus, cytoplasm, mitochondria, and extracellular space (Table 1). However, there are differences in the localization of the proteins predicted by the two tools. The majority of the proteins, according to CELLO's findings, are found in the nucleus, but WoLF PSORT suggested that they may also be found in the cytoplasm and mitochondria. Finding their precise location will require more investigation. Additionally, a multiple sequence alignment of the AL proteins revealed that the members have high sequence similarity (Fig. 1).

The phylogenetic tree showed the evolutionary relationship of *AL* TF genes among *Oryza sativa, Arabidopsis thaliana, Zea mays, Brassica rapa*, and Alfin1 gene of alfalfa (Fig. 2). Based on the AtALs classification [33] used in the earlier research on maize and *Brassica* [34–36], we found that the phylogenetic tree could be split into four different groups, namely I, II, III, and IV. The 9 OsAL protein belonged to groups I, II and IV. No new groups were identified for any of the OsALs. The *OsAL* gene family was discovered to share an equal number of groups with the *ZmAL* gene family but not with the *BrAL* and *AtAL* gene families. This finding suggests that the *OsAL* gene family is a family of genes that originated after the division of monocotyledon and dicotyledon plants [82].

Intron-exon structural divergence in plant genes plays a vital role in evolution [83]. All members of the OsAL TF family were found to have four introns and five exons by gene structure analysis (Fig. 3). The absence of intronless genes suggests that they are highly expressed and have not recently evolved [84,85].

Conserved motifs are crucial for the function and specificity of proteins. Protein interactions can be better understood by identifying common structural patterns that are shared by many different proteins [86]. Conserved motif analysis found that because of the high sequence similarity among OsAL proteins, the motifs were common among the members across the three phylogenetic groups (Fig. 3). The first seven of the fifteen motifs were found in all OsAL proteins. All the proteins had motif 7 and 6 at their N-terminus and C-terminus, respectively. Motif 14 was found in OsAL1, OsAL8, and OsAL5. OsAL2 and OsAL3 shared motif 12, and OsAL9 and OsAL5 both had motif 13. Motif 8 appeared in both OsAL2 and OsAL4. Motif 10 was discovered in OsAL3 and OsAL5, while motif 11 was discovered in OsAL1 and OsAL9. Motif 9 appeared in OsAL3 and OsAL4, while motif 15 appeared in OsAL8 and OsAL7. The functional annotation revealed that motif 1, 3, and 4 comprise the Alfin/DUF3594 domain, whereas motif 2 encodes the PHD- zinc finger motif. The functional annotations of other motifs have yet to be investigated.

The position of a gene on a chromosome impacts the evolution and expression of a trait in plants [87]. The chromosomal distribution study
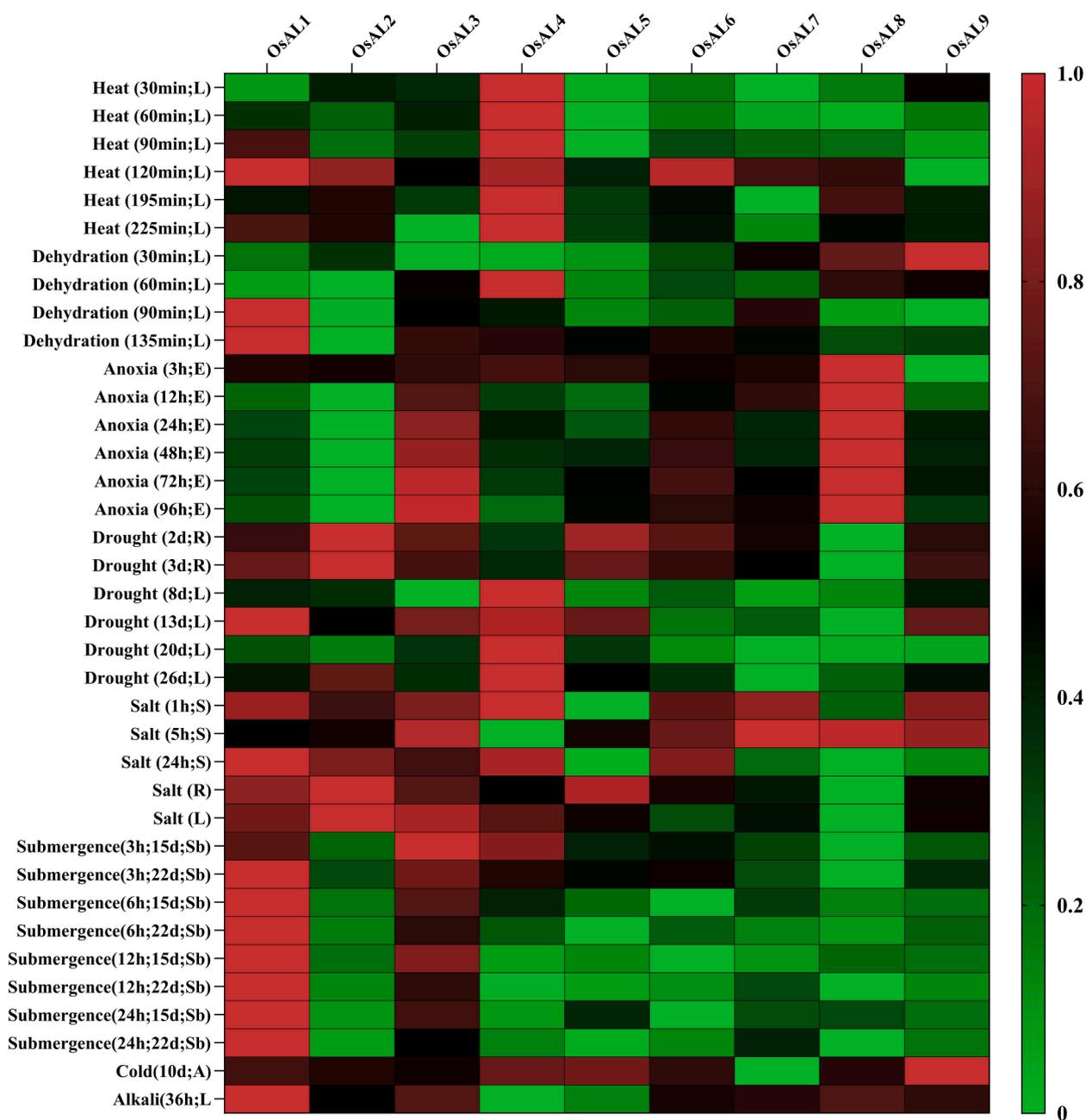
Fig. 14. Expression profiling of *OsALs* under different abiotic stresses. RNA-Seq data retrieved from GENEVESTIGATOR was used to create the heatmap in GraphPad Prism 9.0.0. Relative expression values are denoted by the right-side scale of the heatmap and the colors red, black, and green represent different expression levels: high expression, moderate, and low. Abbreviations: h-hour; d- day; L- leaf; S- shoot; R- root; E- embryo; A- anther; Sb- shoot base.

revealed that *OsAL* genes are randomly and unequally distributed across seven of rice's twelve chromosomes (Fig. 4). There is no *AL* gene in chromosomes 6, 8, 9, 10, and 12. During evolution, gene duplication is important for the expansion of gene families and the emergence of novel functions [88]. Four instances of gene duplication were found in the genome by the gene duplication study (Fig. 5). The occurrences of gene duplication on different chromosomes suggest that segmental duplication was the primary driver of diversification. Ka/Ks values < 1 in the duplicated *OsAL* pairs indicated that purifying selection had played a role in their evolution [77] (Table S2).

Cis-elements and transcription factors (TFs) work together in a gene's promoter region to form the transcription initiation complex and activate the RNA polymerase, which initiates the transcription of a gene [89]. Within the promoter region of the *OsAL* TF gene family, a total of 70 *cis*-regulatory elements were found and further categorized into eight

functional groups (Table S3). Various types of hormones, cellular-development, and stress-related *cis*-elements suggest that this gene family may be directly correlated with rice growth, development, and stress responses (Fig. 6). miRNA-mediated gene regulation plays a crucial role in modulating plant response to environmental stimuli [90]. This study identified a total of 88 distinct candidate miRNAs that could target rice *OsAL* gene family members (Table S4). Seven miRNAs have several *OsAL* genes as their targets, and the rest were specific to only one gene (Fig. 7). The targeted regulatory relationship between miRNAs and their *OsALs* revealed that *OsAL2* was targeted by the fewest miRNAs, while *OsAL4* was targeted by the most miRNAs. It implies that miRNAs could have a potential role in the expression of *OsAL2* and *OsAL4*. Additionally, osa-miR5075 targeted four different genes, suggesting that this miRNA may play a significant role in the expression of *OsALs*. The study also revealed that most of *OsAL* repression mediated by miRNA
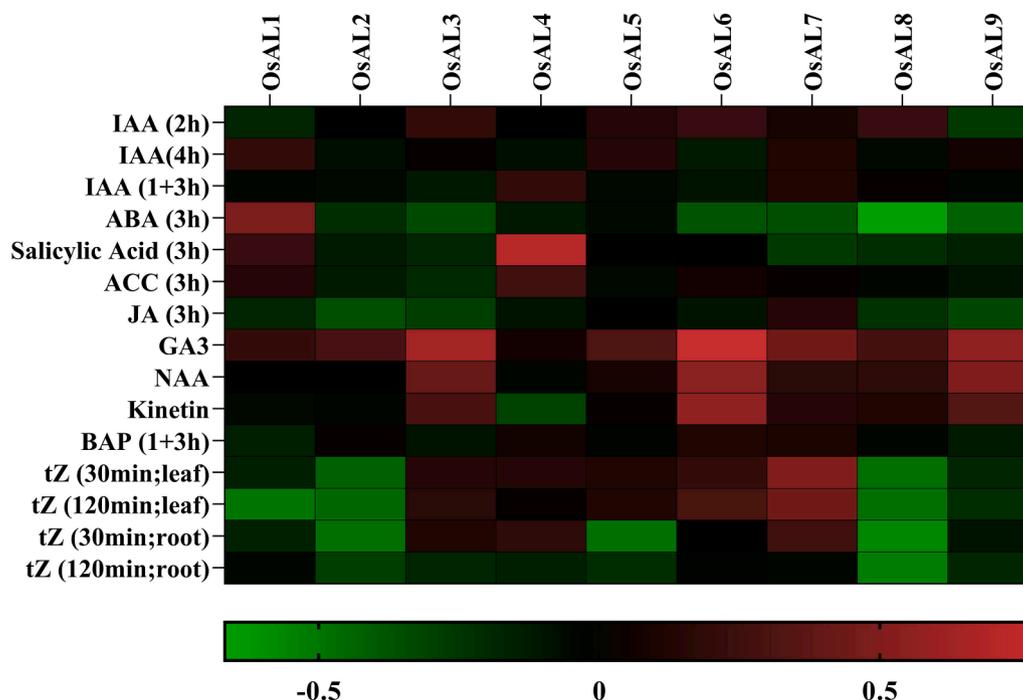
Fig. 15. Expression profiling of *OsALs* in response to various hormonal treatments. Microarray expression data obtained from GENEVESTIGATOR was utilized to construct the heatmap in GraphPad Prism 9.0.0. The scale at the bottom of the heatmap represents Log2 values of relative expression. A high, moderate, and low level of expression are indicated by the colors-red, black, and green, respectively. Abbreviations: IAA- indole-3-aceric acid; ABA- abscisic acid; SA- salicylic acid; ACC- 1-amicocyclopropane-1-carboxylic acid; JA- jasmonic acid; GA3- gibberellin; NAA- 1-naphthaleneacetic acid; BAP- 6-benzylaminopurine; tZ- trans-zeatin.

involves cleavage of mRNA, with a small fraction of targets being inhibited at the translational level (Table S4). It indicates that the identified miRNAs may exert post-transcriptional regulation on *OsAL* expression through mRNA cleavage and translation inhibition.

The miRNA expression analysis observed that most of them have reduced expressions in many tissues and stresses, except for osamiR444b.2, osamiR5797, osa-miR812k, osa-miR408–5p, osa-miR1875, and osa-miR1858a (Fig. 8). The miRNAs mostly displayed reduced expression in response to various stress conditions; this implies that the regulation of *OsAL* genes by miRNAs may play an important role in plants' ability to respond to environmental obstacles and growth conditions.

Secondary structure analysis revealed that all of the proteins had similar proportions of alpha-helix, beta-sheet, isolated beta bridge, turn, coil, and 310-helix, with some exceptions existing in OsAL1, 4, 6, 7, 8, and 9 (Fig. 9 and Table S5). No isolated beta bridge was found in OsAL1, 4, 6, and 9. OsAL7 and OsAL8 had higher beta-sheet, and 310-helix percentages than the other proteins, respectively. The alpha-helix's dominance over other structures, which suggests that the protein structure is stable, was also demonstrated in the study [78,91]. Only OsAL5 was found to have a single membrane-spanning motif (MSM) (Fig. 9 and Table S5).

3D structures of the OsAL proteins were also predicted in this study (Fig. 10). Several tools confirmed the high quality of all the predicted structures found in the tertiary structure analysis (Fig S1, S2, S3, S4 and Table S6). These structures might be used in further research to more precisely analyze the proteins.

To understand how *OsALs* function throughout the growth and development of the plant, RNA-seq data was used to examine *OsAL* expression patterns in different kinds of tissues in rice (Fig. 11). *OsALs* were found to be differentially expressed by the tissue-specific expression analysis.

Higher expression levels of *OsAL1, OsAL2, OsAL8*, and *OsAL9* in most of the tissues revealed the involvement of these proteins in the development and growth of rice. In addition, *OsAL3* had abundant expression

in anther, pistil, and seed, *OsAL7* in shoots, endosperm, and seed, *OsAL5* in anther, and *OsAL6* in roots, shoots, and seed. Only *OsAL4* showed downregulation in every tissue. The high expression of *OsALs* in anther, pistil, embryo, and seed suggests their potential involvement in pollen germination, fertilization, and reproduction processes, whereas their abundant expressions in panicle, inflorescence, and leaves suggest their respective roles in flower and leaf development. In addition, the high expression levels of these genes in both root and shoot imply their substantial contribution to the development of these plant organs. Furthermore, the way they express themselves during the seedling stage suggests that they may be involved in the early stages of plant growth. Alfalfa [25] and *Arabidopsis* [26] have both previously reported on the function of *ALs* in root growth. In addition, previous studies in *Brassica rapa* [34], *Brassica oleracea* [35], and *Zea mays* [36] have all revealed evidence of the involvement of *ALs* in a variety of physiological processes in plants, such as flower development, root and shoot growth, and leaf development.

Furthermore, the expression of *OsALs* in the ten developmental stages of rice was examined using pre-analyzed expression data from mRNA-seq expression values retrieved from GENEVESTIGATOR. All *OsALs* were found to be substantially expressed during the developmental stages, except for *OsAL4*, which displayed a low expression pattern during the early developmental stages of the rice plant (Fig. 12). The findings also suggests that the *AL* genes play a role in rice development and growth.

In addition, to understand the responsiveness of *OsALs* to different plant hormones, their expression patterns under various hormonal treatments were investigated via microarray data (Fig. 15). The expression analysis revealed that in response to indole-3-acetic acid (IAA), an auxin class of hormones, all *OsALs* exhibited elevated expression except *OsAL2*. Only *OsAL1* demonstrated upregulation under abscisic acid (ABA) hormone treatment, while salicylic acid (SA) hormone induced the elevated expression of *OsAL1* and *4*. Additionally, the direct precursor of ethylene hormone, 1-amicocyclopropane-1-carboxylic acid (ACC), led to the increased expression of *OsAL1, 4, 6*, and *7*.
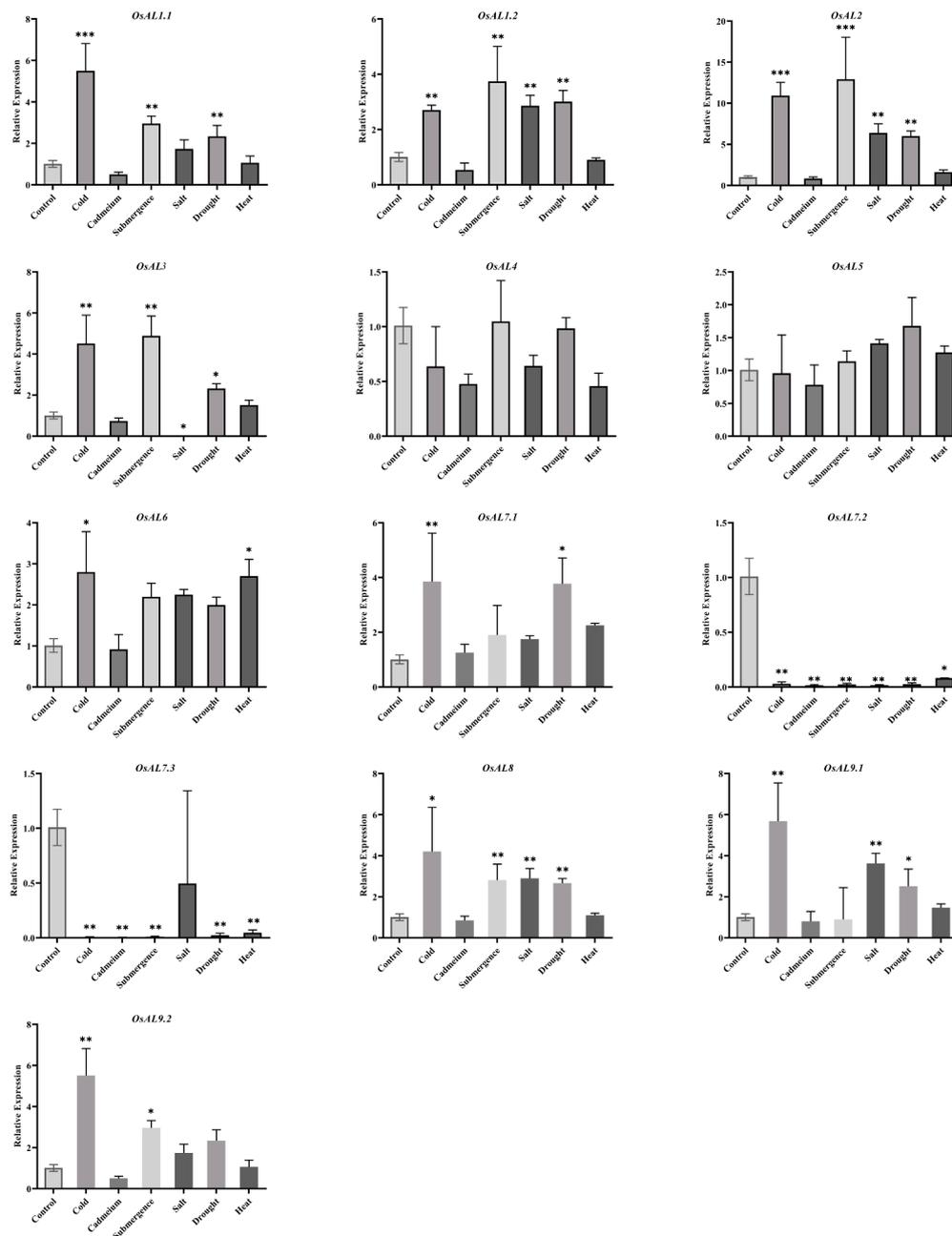
Fig. 16. Analysis of *OsALs*' relative expression in response to different abiotic stress conditions by RT-qPCR. The relative expression value of each OsAL in the leaves of rice plant is shown on the Y-axis, and the X-axis shows the stress conditions. Mean value of expression at each stress treatment was determined via technical replication. Data were analyzed in GraphPad Prism 9.0.0 and Microsoft Office 2019. One-way ANOVA and Bonferroni post-hoc test were performed to determine the significant difference ($P \leq 0.05$). To show the significant differences, different means were labeled with different numbers of stars (*).

Only *OsAL7* was highly expressed upon exposure to jasmonic acid (JA), a jasmonate class of lipid-based hormone. Notably, the gibberellin (GA3) hormone triggered the high expression of all *OsALs*. Both the synthetic auxin, 1-naphthaleneacetic acid (NAA), and a cytokinin-like synthetic hormone, kinetin induced the upregulation of *OsAL3* and *OsAL5–9*. In addition, *OsAL2, 4, 6*, and *7* demonstrated elevated expression in response to synthetic cytokinin 6-benzylaminopurine (BAP), while under the treatment of another cytokinin derived hormone trans-zeatin (tZ*), OsAL3–7* exhibited significant upregulation. Therefore, this study also demonstrated the responsiveness of *OsALs* to different types of plant hormones which are important for the growth and development of rice.

Additionally, to investigate the involvement of *OsALs* in different stress conditions, expression levels were examined using RNA-seq data. *OsALs* activity upon exposure to bacteria, fungi, and viruses was

investigated. The study explored how *OsALs* responded to different biotic factors including bacterial, fungal, and viral infections (Fig. 13). Our research revealed that when exposed to rice stripe virus (RSV), *OsAL4–7* and *9* exhibited increased expression, while rice dwarf virus (RDV) infection led to upregulation of *OsAL1* and *9*. Moreover, *OsAL3, 5*, and *7* showed higher activity when confronted with Xanthomonas oryzae pv. Oryzicola, a leaf-streak disease causing bacterium.

The study also demonstrated that both rice sheath blight-causing *Rhizoctonia solani* and rice blast-causing *Magnaporthe oryzae* fungal infection elevated the expression of *OsAL2, 3*, and *8*. The precise role of *Arabidopsis* AL proteins under biotic stresses remains undetermined. The current finding is consistent with prior studies that have demonstrated the role of *Brassica* AL proteins in providing resistance to fungal and bacterial infections [34,35]. Furthermore, the biotic stress-responsive

cis-elements W-box, WRE3, and WUN motifs were found in all the *OsALs* promoter regions. Our investigation of the function of these genes in providing defenses under such stressful circumstances is also supported by the existence of these motifs.

To ascertain the role of *OsALs* under various abiotic stresses, we examined the information taken after analyzing the RNA-seq data (Fig. 14), and compared it with the findings of the RT-qPCR analysis (Fig. 16). Despite a few inconsistencies, the expression level of the majority of the genes remained consistent in both instances of the RNA-Seq and RT-qPCR analysis. In both cases, *OsAL1–3* and *OsAL6–9* genes were found to be crucial for dealing with cold stress, pointing to their potential role in this situation. In addition, *OsAL1, 2, 8,* and *9* exhibited a higher level of expression when the plant was exposed to salt stress, suggesting their function in helping the plant survive under such conditions. Under submergence, *OsAL1* and *3* exhibited upregulation, which raises the possibility that they are involved in this stress response. *OsAL6* had a significantly higher expression level under heat stress in both cases, implying its role in heat stress. The upregulated expression of *OsAL1–3, 7,* and *9* under drought conditions suggests that these genes may aid plants in surviving in this environment. Positive responses of *Alfin-like* genes under cold, salt, and drought stresses have been reported in *Brassica rapa* [34], *Brassica oleracea* [35], and *Zea mays* [36]. These plants also showed higher expression levels for the majority of their *AL* genes. In addition, transgenic alfalfa with overexpressed *Alfin1* exhibited enhanced salinity tolerance in a study that previously documented the role of the *Alfin1* gene in alfalfa under salt stress [25,32]. Furthermore, transgenic *Arabidopsis* that expressed the *AhAL1* gene of *Atriplex hortensis* was shown to have improved tolerance to salt and drought stress, and this finding suggested *AhAL1* as a novel candidate gene for enhancing crop plants' resistance to abiotic stress [37]. The notion concerning the involvement and function of *OsALs* in contexts of abiotic stress is also supported by the presence of multiple *cis*-elements in the promoter region which show responsiveness to different abiotic stresses. Since gene expression is among the most intricate biological processes, more research is required to comprehend the mechanism through which *OsALs* control plant's reactions to different stresses. These results imply that *OsALs* are important for responding to a broad spectrum of biotic and abiotic stresses; therefore, this gene family holds promise as a potential target for genetic engineering that can offer defense against a variety of stress conditions.

5. Conclusion

This research study involved the characterization of the rice *Alfin-like* TF gene family and a comprehensive analysis of their expression profiles under different contexts. This study identified nine *OsAL* genes, which are located on different chromosomes (1–5, 7, and 11). Through phylogenetic analysis, these genes were grouped into three categories, and it was observed that this gene family evolved through segmental duplication over time. Examining *cis*-regulatory elements indicated their role in hormone regulation, developmental stages, and stress response to a range of stimuli. Additionally, miRNA analysis aided in elucidating their function to regulate the activity of their targeted *OsALs*. A thorough analysis of the expression profiles by employing the RNA-seq data revealed the role of *OsALs* during different growth phases and stresses. In addition, their expression profiles in response to different hormones and biotic stresses were also analyzed. Furthermore, the activity of *OsALs* to a range of abiotic stresses was validated by the expression analysis performed with RT-qPCR data. The three-dimensional structures of the OsAL proteins were also predicted in this study. The discoveries above contribute to improved knowledge and a more comprehensive understanding of the roles played by *AL* genes in rice. These findings would also lay the groundwork for future research on *OsALs*, potentially leading to the genomic modifications aimed at furthering rice improvement and developing stress-tolerant varieties.

CRediT authorship contribution statement

Jeba Faizah Rahman: Conceptualization, Methodology, Data curation, Software, Investigation, Writing – original draft. **Hammadul Hoque:** Validation, Investigation. **Abdullah -Al- Jubayer:** Validation, Data curation. **Nurnabi Azad Jewel:** Validation, Methodology. **Md. Nazmul Hasan:** Conceptualization, Methodology, Software, Investigation. **Aniqua Tasnim Chowdhury:** Software. **Shamsul H. Prodhan:** Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.btre.2024.e00845.

References

[1] J.S. Boyer, Plant productivity and environment, Science 218 (4571) (1982) 443–448.
[2] J.D. Jones, J.L. Dangl, The plant immune system, Nature 444 (7117) (2006) 323–329.
[3] K. Yamaguchi-Shinozaki, K. Shinozaki, Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses, Annu. Rev. Plant Biol. 57 (2006) 781–803.
[4] K. Urano, Y. Kurihara, M. Seki, K. Shinozaki, 'Omics' analyses of regulatory networks in plant abiotic stress responses, Curr. Opin. Plant Biol. 13 (2) (2010) 132–138.
[5] H. Knight, M.R. Knight, Abiotic stress signalling pathways: specificity and cross-talk, Trends Plant Sci. 6 (6) (2001) 262–267.
[6] V. Chinnusamy, K. Schumaker, J.K. Zhu, Molecular genetic perspectives on cross-talk and specificity in abiotic stress signalling in plants, J. Exp. Bot. 55 (395) (2004) 225–236.
[7] Z. Peleg, E. Blumwald, Hormone balance and abiotic stress tolerance in crop plants, Curr. Opin. Plant Biol. 14 (3) (2011) 290–295.
[8] A.S. Duque, A.M. de Almeida, A.B. da Silva, J.M. da Silva, A.P. Farinha, D. Santos, P. Fevereiro, S. de Sousa Araújo, Abiotic stress responses in plants: unraveling the complexity of genes and networks to survive, Abiot. Stress-Plant Resp. Appl. Agric. (2013) 49–101.
[9] S. Mahajan, N. Tuteja, Cold, salinity and drought stresses: an overview, Arch. Biochem. Biophys. 444 (2) (2005) 139–158.
[10] M.F. Thomashow, Plant cold acclimation: freezing tolerance genes and regulatory mechanisms, Annu. Rev. Plant Biol. 50 (1) (1999) 571–599.
[11] K. Shinozaki, K. Yamaguchi-Shinozaki, Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways, Curr. Opin. Plant Biol. 3 (3) (2000) 217–223.
[12] L. Xiong, K.S. Schumaker, J.-K. Zhu, Cell signaling during cold, drought, and salt stress, Plant Cell 14 (suppl_1) (2002) S165–S183.
[13] J.-K. Zhu, Salt and drought stress signal transduction in plants, Annu. Rev. Plant Biol. 53 (1) (2002) 247–273.
[14] K. Shinozaki, K. Yamaguchi-Shinozaki, M. Seki, Regulatory network of gene expression in the drought and cold stress responses, Curr. Opin. Plant Biol. 6 (5) (2003) 410–417.
[15] G.-T. Huang, S.-L. Ma, L.-P. Bai, L. Zhang, H. Ma, P. Jia, J. Liu, M. Zhong, Z.-F. Guo, Signal transduction during cold, salt, and drought stresses in plants, Mol. Biol. Rep. 39 (2012) 969–987.

[16] J.L. Riechmann, J. Heard, G. Martin, L. Reuber, C.-Z. Jiang, J. Keddie, L. Adam, O. Pineda, O. Ratcliffe, R. Samaha, Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes, Science 290 (5499) (2000) 2105–2110.

[17] L.-J. Qu, Y.-X. Zhu, Transcription factor families in Arabidopsis: major progress and outstanding issues for future research, Curr. Opin. Plant Biol. 9 (5) (2006) 544–549.

[18] T.A. Meraj, J. Fu, M.A. Raza, C. Zhu, Q. Shen, D. Xu, Q. Wang, Transcriptional factors regulate plant stress responses through mediating secondary metabolism, Genes (Basel) 11 (4) (2020) 346.

[19] S.-A. Khan, M.-Z. Li, S.-M. Wang, H.-J. Yin, Revisiting the role of plant transcription factors in the battle against abiotic stress, Int. J. Mol. Sci. 19 (6) (2018) 1634.

[20] Y. Yoon, D.H. Seo, H. Shin, H.J. Kim, C.M. Kim, G. Jang, The role of stress-responsive transcription factors in modulating abiotic stress tolerance in plants, Agronomy 10 (6) (2020) 788.

[21] P.N. Benfey, D. Weigel, Transcriptional networks controlling plant development, Plant Physiol. 125 (1) (2001) 109–111.

[22] J. Schiefelbein, Cell-fate specification in the epidermis: a common patterning mechanism in the root and shoot, Curr. Opin. Plant Biol. 6 (1) (2003) 74–78.

[23] S.B. Carroll, Evolution at two levels: on genes and form, PLoS Biol. 3 (7) (2005) e245.

[24] J. Jin, F. Tian, D.-C. Yang, Y.-Q. Meng, L. Kong, J. Luo, G. Gao, PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants, Nucleic Acids Res. (2016) gkw982.

[25] I. Winicov, Alfin1 transcription factor overexpression enhances plant root growth under normal and saline conditions and improves salt tolerance in alfalfa, Planta 210 (2000) 416–422.

[26] N.N.P. Chandrika, K. Sundaravelpandian, S.M. Yu, W. Schmidt, ALFIN-LIKE 6 is involved in root hair elongation during phosphate deficiency in Arabidopsis, New Phytol. 198 (3) (2013) 709–720.

[27] J.E. Krochko, J.D. Bewley, Use of electrophoretic techniques in determining the composition of seed storage proteins in alfalfa, Electrophoresis 9 (11) (1988) 751–763.

[28] J.E. Krochko, S.K. Pramanik, J.D. Bewley, Contrasting storage protein synthesis and messenger RNA accumulation during development of zygotic and somatic embryos of alfalfa (Medicago sativa L.), Plant Physiol. 99 (1) (1992) 46–53.

[29] M. Bienz, The PHD finger, a nuclear protein-interaction domain, Trends Biochem. Sci. 31 (1) (2006) 35–40.

[30] I. Winicov, cDNA encoding putative zinc finger motifs from salt-tolerant alfalfa (Medicago sativa L.) cells, Plant Physiol. 102 (2) (1993) 681.

[31] D.R. Bastola, V.V. Pethe, I. Winicov, Alfin1, a novel zinc-finger protein in alfalfa roots that binds to promoter elements in the salt-inducible MsPRP2 gene, Plant Mol. Biol. 38 (1998) 1123–1135.

[32] I. Winicov, D.R. Bastola, Transgenic overexpression of the transcription factor Alfin1 enhances expression of the endogenous MsPRP2 gene in alfalfa and improves salinity tolerance of the plants, Plant Physiol. 120 (2) (1999) 473–480.

[33] Y. Song, J. Gao, F. Yang, C.-S. Kua, J. Liu, C.H. Cannon, Molecular evolutionary analysis of the Alfin-like protein family in Arabidopsis lyrata, Arabidopsis thaliana, and Thellungiella halophila, PLoS ONE 8 (7) (2013) e66838.

[34] M.A. Kayum, J.-I. Park, N.U. Ahmed, H.-J. Jung, G. Saha, J.-G. Kang, I.-S. Nou, Characterization and stress-induced expression analysis of Alfin-like transcription factors in Brassica rapa, Mol. Genet. Genomics 290 (2015) 1299–1311.

[35] M.A. Kayum, J.-I. Park, N.U. Ahmed, G. Saha, M.-Y. Chung, J.-G. Kang, I.-S. Nou, Alfin-like transcription factor family: characterization and expression profiling against stresses in Brassica oleracea, Acta Physiologiae Plantarum 38 (2016) 1–14.

[36] W. Zhou, J. Wu, Q. Zheng, Y. Jiang, M. Zhang, S. Zhu, Genome-wide identification and comparative analysis of Alfin-like transcription factors in maize, Genes Genomics 39 (2017) 261–275.

[37] J.-J. Tao, W. Wei, W.-J. Pan, L. Lu, Q.-T. Li, J.-B. Ma, W.-K. Zhang, B. Ma, S.-Y. Chen, J.-S. Zhang, An Alfin-like gene from Atriplex hortensis enhances salt and drought tolerance and abscisic acid response in transgenic Arabidopsis, Sci. Rep. 8 (1) (2018) 1–13.

[38] W. Liu, J. Liu, L. Triplett, J.E. Leach, G.-L. Wang, Novel insights into rice innate immunity against bacterial and fungal pathogens, Annu. Rev. Phytopathol. 52 (2014) 213–241.

[39] Y. Kawahara, M. de la Bastide, J.P. Hamilton, H. Kanamori, W.R. McCombie, S. Ouyang, D.C. Schwartz, T. Tanaka, J. Wu, S. Zhou, Improvement of the Oryza sativa Nipponbare reference genome using next generation sequence and optical map data, Rice 6 (2013) 1–10.

[40] D.M. Goodstein, S. Shu, R. Howson, R. Neupane, R.D. Hayes, J. Fazo, T. Mitros, W. Dirks, U. Hellsten, N. Putnam, Phytozome: a comparative platform for green plant genomics, Nucleic Acids Res. 40 (D1) (2012) D1178–D1186.

[41] P. Lamesch, T.Z. Berardini, D. Li, D. Swarbreck, C. Wilks, R. Sasidharan, R. Muller, K. Dreher, D.L. Alexander, M. Garcia-Hernandez, The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools, Nucleic Acids Res. 40 (D1) (2012) D1202–D1210.

[42] J. Mistry, S. Chuguransky, L. Williams, M. Qureshi, G.A. Salazar, E.L. Sonnhammer, S.C. Tosatto, L. Paladin, S. Raj, L.J. Richardson, Pfam: the protein families database in 2021, Nucleic Acids Res. 49 (D1) (2021) D412–D419.

[43] S. Lu, J. Wang, F. Chitsaz, M.K. Derbyshire, R.C. Geer, N.R. Gonzales, M. Gwadz, D. I. Hurwitz, G.H. Marchler, J.S. Song, CDD/SPARCLE: the conserved domain database in 2020, Nucleic Acids Res. 48 (D1) (2020) D265–D268.

[44] Gasteiger, E., C. Hoogland, A. Gattiker, S. e. Duvaud, M.R. Wilkins, R.D. Appel and A. Bairoch (2005). Protein identification and analysis tools on the ExPASy server, Springer.

[45] C.S. Yu, Y.C. Chen, C.H. Lu, J.K. Hwang, Prediction of protein subcellular localization, Proteins: Struct. Funct. Bioinform. 64 (3) (2006) 643–651.

[46] P. Horton, K.-J. Park, T. Obayashi, N. Fujita, H. Harada, C. Adams-Collier, K. Nakai, WoLF PSORT: protein localization predictor, Nucleic Acids Res. 35 (suppl_2) (2007) W585–W587.

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

[48] Nicholas, K.B. (1997). "Genedoc: a tool for editing and annoting multiple sequence alignments." http://wwwpscedu/biomed/genedoc.

[49] I. Letunic, P. Bork, Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation, Bioinformatics 23 (1) (2007) 127–128.

[50] B. Hu, J. Jin, A.-Y. Guo, H. Zhang, J. Luo, G. Gao, GSDS 2.0: an upgraded gene feature visualization server, Bioinformatics 31 (8) (2015) 1296–1297.

[51] Bailey, T.L. and C. Elkan (1994). "Fitting a mixture model by expectation maximization to discover motifs in bipolymers".

[52] R. Voorrips, MapChart: software for the graphical presentation of linkage maps and QTLs, J. Hered. 93 (1) (2002) 77–78.

[53] T.-H. Lee, H. Tang, X. Wang, A.H. Paterson, PGDD: a database of gene and genome duplication in plants, Nucleic Acids Res. 41 (D1) (2012) D1152–D1158.

[54] H. Kong, L.L. Landherr, M.W. Frohlich, J. Leebens-Mack, H. Ma, C.W. DePamphilis, Patterns of gene duplication in the plant SKP1 gene family in angiosperms: evidence for multiple mechanisms of rapid gene birth, Plant J. 50 (5) (2007) 873–885.

[55] M.A. Koch, B. Haubold, T. Mitchell-Olds, Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in Arabidopsis, Arabis, and related genera (Brassicaceae), Mol. Biol. Evol. 17 (10) (2000) 1483–1498.

[56] C. Chen, H. Chen, Y. Zhang, H.R. Thomas, M.H. Frank, Y. He, R. Xia, TBtools: an integrative toolkit developed for interactive analyses of big biological data, Mol. Plant 13 (8) (2020) 1194–1202.

[57] M. Lescot, P. Déhais, G. Thijs, K. Marchal, Y. Moreau, Y. Van de Peer, P. Rouzé, S. Rombauts, PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences, Nucleic Acids Res. 30 (1) (2002) 325–327.

[58] X. Dai, P.X. Zhao, psRNATarget: a plant small RNA target analysis server, Nucleic Acids Res. 39 (suppl_2) (2011) W155–W159.

[59] S. Griffiths-Jones, H.K. Saini, S. Van Dongen, A.J. Enright, miRBase: tools for microRNA genomics, Nucleic Acids Res. 36 (suppl_1) (2007) D154–D158.

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

[61] A.K.S. Gurjar, A.S. Panwar, R. Gupta, S.S. Mantri, PmiRExAt: plant miRNA expression atlas database and web applications, Database (2016). **2016**.

[62] M. Heinig, D. Frishman, STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins, Nucleic Acids Res. 32 (suppl_2) (2004) W500–W502.

[63] A. Krogh, B. Larsson, G. Von Heijne, E.L. Sonnhammer, Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes, J. Mol. Biol. 305 (3) (2001) 567–580.

[64] D.E. Kim, D. Chivian, D. Baker, Protein structure prediction and analysis using the Robetta server, Nucleic Acids Res. 32 (suppl_2) (2004) W526–W531.

[65] J. Ko, H. Park, L. Heo, C. Seok, GalaxyWEB server for protein structure prediction and refinement, Nucleic Acids Res. 40 (W1) (2012) W294–W297.

[66] N. Guex, M.C. Peitsch, SWISS-MODEL and the SWISS-Pdb Viewer: an environment for comparative protein modeling, Electrophoresis 18 (15) (1997) 2714–2723.

[67] R.A. Laskowski, J.A.C. Rullmann, M.W. MacArthur, R. Kaptein, J.M. Thornton, AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR, J. Biomol. NMR 8 (1996) 477–486.

[68] C. Colovos, T.O. Yeates, Verification of protein structures: patterns of nonbonded atomic interactions, Protein Sci. 2 (9) (1993) 1511–1519.

[69] M. Wiederstein, M.J. Sippl, ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins, Nucleic Acids Res. 35 (suppl_2) (2007) W407–W410.

[70] Biovia, D.S. (2017). Discovery studio modeling environment, release.

[71] L. Xia, D. Zou, J. Sang, X. Xu, H. Yin, M. Li, S. Wu, S. Hu, L. Hao, Z. Zhang, Rice Expression Database (RED): an integrated RNA-Seq-derived gene expression database for rice, J. Genet. Genomics 44 (5) (2017) 235–241.

[72] T. Hruz, O. Laule, G. Szabo, F. Wessendorp, S. Bleuler, L. Oertle, P. Widmayer, W. Gruissem, P. Zimmermann, Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes, Adv. Bioinformatics (2008). **2008**.

[73] H. Motulsky, Analyzing data with GraphPad prism, GraphPad Softw. Incorp. (1999).

[74] J. Ye, G. Coulouris, I. Zaretskaya, I. Cutcutache, S. Rozen, T.L. Madden, Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction, BMC Bioinformatics 13 (2012) 1–11.

[75] M. Jain, A. Nijhawan, A.K. Tyagi, J.P. Khurana, Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR, Biochem. Biophys. Res. Commun. 345 (2) (2006) 646–651.

[76] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method, Methods 25 (4) (2001) 402–408.

[77] Z. Liu, Y. Liu, J.A. Coulter, B. Shen, Y. Li, C. Li, Z. Cao, J. Zhang, The WD40 gene family in potato (Solanum Tuberosum L.): genome-wide analysis and identification of anthocyanin and drought-related WD40s, Agronomy 10 (3) (2020) 401.

[78] J. Muthukumaran, P. Manivel, M. Kannan, J. Jeyakanthan, R. Krishna, A framework for classification of antifreeze proteins in over wintering plants based on their sequence and structural features, J. Bioinform. Seq. Anal. 3 (2011) 70–88.

[79] S. Deng, J. Sun, R. Zhao, M. Ding, Y. Zhang, Y. Sun, W. Wang, Y. Tan, D. Liu, X. Ma, Populus euphratica APYRASE2 enhances cold tolerance by modulating vesicular

trafficking and extracellular ATP in Arabidopsis plants, Plant Physiol. 169 (1) (2015) 530–548.

[80] P. Schneider, F. Asch, Rice production and food security in Asian Mega deltas—A review on characteristics, vulnerabilities and agricultural adaptation options to cope with climate change, J. Agron. Crop Sci. 206 (4) (2020) 491–503.

[81] Y. Xiong, T. Liu, C. Tian, S. Sun, J. Li, M. Chen, Transcription factors in rice: a genome-wide comparative analysis between monocots and eudicots, Plant Mol. Biol. 59 (2005) 191–203.

[82] Q. Liu, C. Sun, J. Han, L. Li, K. Wang, Y. Wang, J. Chen, M. Zhao, Y. Wang, M. Zhang, Identification, characterization and functional differentiation of the NAC gene family and its roles in response to cold stress in ginseng, Panax ginseng CA Meyer, PLoS ONE 15 (6) (2020) e0234423.

[83] T.E. Koralewski, K.V. Krutovsky, Evolution of exon-intron structure and alternative splicing, PLoS ONE 6 (3) (2011) e18055.

[84] R.K. Deshmukh, H. Sonah, N.K. Singh, Intron gain, a dominant evolutionary process supporting high levels of gene expression in rice, J. Plant Biochem. Biotechnol. 25 (2016) 142–146.

[85] P. Nagar, A. Kumar, M. Jain, S. Kumari, A. Mustafiz, Genome-wide analysis and transcript profiling of PSKR gene family members in Oryza sativa, PLoS ONE 15 (7) (2020) e0236349.

[86] K.D. MacIsaac, E. Fraenkel, Practical strategies for discovering regulatory DNA sequence motifs, PLoS Comput. Biol. 2 (4) (2006) e36.

[87] B. Piechulla, N. Merforth, B. Rudolph, Identification of tomato Lhc promoter regions necessary for circadian expression, Plant Mol. Biol. 38 (1998) 655–662.

[88] Z. Huang, W. Duan, X. Song, J. Tang, P. Wu, B. Zhang, X. Hou, Retention, molecular evolution, and expression divergence of the auxin/indole acetic acid and auxin response factor gene families in Brassica rapa shed light on their evolution patterns in plants, Genome Biol. Evol. 8 (2) (2016) 302–316.

[89] K. Sutoh, D. Yamauchi, Two cis-acting elements necessary and sufficient for gibberellin-upregulated proteinase expression in rice seeds, Plant J. 34 (5) (2003) 635–645.

[90] S.K. Yadav, V.V. Santosh Kumar, R.K. Verma, P. Yadav, A. Saroha, D.P. Wankhede, B. Chaudhary, V. Chinnusamy, Genome-wide identification and characterization of ABA receptor PYL gene family in rice, BMC Genomics 21 (2020) 1–27.

[91] R. Islam, M. Rahaman, H. Hoque, N. Hasan, S.H. Prodhan, A. Ruhama, N.A. Jewel, Computational and structural based approach to identify malignant nonsynonymous single nucleotide polymorphisms associated with CDK4 gene, PLoS ONE 16 (11) (2021) e0259691.