scientific reports

Check for updates

OPEN

Genome-wide identification of CRF gene family members in four rice subspecies and expression analysis of OsCRF members in response to cold stress at seedling stage

Lei Lei[1,2,3], Guohua Ding[1,2,3], Liangzi Cao[1,2,3], Jinsong Zhou[1,2,3], Yu Luo[1], Liangming Bai[1,2], Tianshu Xia[1], Lei Chen[4], Jiangxu Wang[4], Kai Liu[4], Yang Ren[4], Yusong Miao[4], Qingjun Lei[5], Tingting Xie[1], Guang Yang[1], Wan Li[1], Xueyang Wang[4] & Shichen Sun[1,2,3]✉

Cytokinin Response Factors (CRFs) play a crucial role in plant growth and development, hormone signaling, and responses to biotic and abiotic stresses. However, there have been no reports on CRF genes in rice until now. We analyzed the CRF families in four rice subspecies: cultivated rice *Oryza sativa Japonica* Group, *Oryza sativa Indica* Group, and *Oryza sativa* (circum-Aus1 var. N22), as well as wild rice *Oryza rufipogon*. We identified 7, 6, 6, and 7 CRF in their genomes, respectively, distributed across different chromosomes. The protein motifs and gene structures of CRF in these four types of rice show high conservation. Cis-regulatory element analysis revealed that the promoter regions of the CRF contain numerous hormone and stress-related elements. The number of CRF in these four types of rice is not influenced by gene duplication. The expression pattern showed that OsCRF exhibit significant tissue-specific expression. The qRT-PCR results showed that OsCRF strongly responded to low-temperature stress and can be induced by melatonin and cytokinin to increase expression levels. In addition, the nuclear localisation of *OsCRF4/5* was confirmed to be as predicted. The results above will provide a foundation for further and deeper investigation of CRFs.

Keywords *Oryza sativa*, CRF family, Expression pattern, Low temperature stress

AP2/ERF (APETALA2/ETHYLENE RESPONSIVE FACTOR) is a widely present superfamily of transcription factors in plants. A significant structural feature of this family is the presence of an AP2 domain, which is a DNA binding domain composed of 60–70 amino acids[1]. Cytokinin Response Factors (CRFs) genes belong to the ERF-type members of the AP2/ERF transcription factor superfamily[2]. The CRF gene sequence consists of a single AP2/ERF transcription factor DNA-binding domain, a highly conserved CRF domain, and a variable C-terminal region. The latter is involved in protein-protein interactions, distinguishing CRFs from other AP2/ERF members[3,4]. The CRF genes are a branch of the two-component signaling system. In this system, cytokinins bind to self-phosphorylating histidine kinase receptors, which then transfer the phosphate group to histidine phosphotransfer proteins (HPT). The HPTs subsequently transfer the phosphate group to transcription factors, specifically the type-B response regulators (RRB)[5]. Currently, CRF genes have been found in mosses, lycophytes, ferns, and most flowering plants[4].

CRF genes are widely involved in plant growth and development, hormone signaling, and responses to biotic and abiotic stresses[6,7]. In *Arabidopsis*, approximately 25% of the F_2 progeny with knockouts of the *CRF5* and *CRF6* genes exhibit impaired embryo development[8]. Research on cork oak has found that *CRF3* plays a role in embryo development, suggesting that CRF genes may be involved in seed embryo development in flowering plants[9]. Both *AtCRF3* and *AtCRF4* can enhance *Arabidopsis* tolerance to low-temperature stress[10]. Both *AtCRF1* and *AtCRF2* are associated with salt stress. Mutants of *Arabidopsis* lacking the functions of crf1 and crf2 exhibit

[1]Institute of Crop Cultivation and Tillage, Heilongjiang Academy of Agricultural Sciences, Harbin 150028, China. [2]Heilongjiang Rice Quality Improvement and Genetic Breeding Engineering Research Center, Harbin 150086, China. [3]Northeast of National Center of Technology Innovation for Saline-Alkali Tolerant Rice, Harbin 150086, China. [4]Heilongjiang Academy of Agricultural Sciences, Harbin 150086, China. [5]Branch of Animal Husbandry and Veterinary of Heilongjiang Academy of Agricultural Sciences, Qiqihar 161005, China. ✉email: sunshichen1979@163.com

higher photosynthetic efficiency than the wild type after salt treatment[11]. In Solanum lycopersicum, four genes, *SlCRF1*, *SlCRF2*, *SlCRF3*, and *SlCRF5*, respond to various abiotic stresses, such as low temperature, drought, and high temperature. Additionally, *SlCRF5* is involved in the development of tomato roots, leaves, and flowers[12–14]. In *Arabidopsis*, plants overexpressing the *CRF2* gene exhibit significantly increased levels of gene expression associated with disease resistance and demonstrate stronger resistance to pathogens compared to wild-type *Arabidopsis*[15]. In addition, the CRF gene interacts with auxin-transporting PIN proteins and plays a more direct role in regulating auxin distribution throughout the plant[8][8]. And *CRF4* in *Arabidopsis* was shown to be the earliest nitrogen-responsive transcription factor[16].

Rice is one of the major staple food crops, with half of the world's population consuming it. Previous studies have identified and analyzed the AP2/ERF transcription factor family in rice, but research on CRF genes has not been conducted[17]. And the CRF gene response to low temperature in rice is not clear. Moreover, rice has widely different genetic backgrounds depending on the classification of subspecies, so revealing the distribution and variation of the CRF genes in different rice subspecies is important for understanding its functional evolution. Therefore, three cultivated rice varieties (*Oryza sativa Japonica* Group (*Os*), *Oryza sativa Indica* Group (*Oi*), *Oryza sativa* circum-Aus1 var. N22 (*Osn*)) and one wild rice variety (*Oryza rufipogon*, *Or*) were selected in this study for CRF gene family Bioinformatics analysis of the CRF gene family was performed in *Oryza sativa* (*circum-Aus1 var*. N22) and a wild rice variety (*Oryza rufipogon*) with the aim of exploring the characteristics of the CRF gene in rice. The main components of the analysis included the physicochemical properties of the proteins, the construction of phylogenetic trees, the conserved patterns of the genes, the gene structures, the chromosomal localisation of the genes, the duplication types of the genes, the cis-acting elements and the expression patterns. Finally, we further investigated the expression levels of OsCRF members in response to low temperature, as well as the effects of melatonin and cytokinin on gene expression levels under low temperature stress. The results of this study will help to gain insights into the characterisation of CRF genes in rice and their response to low temperature stress, and provide a reference for understanding their functions.

Results

Identification and physicochemical analysis of CRF family members in four rice species

We identified 157, 162, 157, and 171 AP2/ERF members in the genomes of *Os*, *Oi*, *Osn*, and *Or* rice species, respectively. Subsequently, we constructed neighbor-joining phylogenetic trees using the protein sequences of AP2/ERF members from these four rice species and CRF member protein sequences from 14 other species including *A. thaliana*, *G. max*, *P. tomentosa*, *T. hispida*, *M. caespitosa*, *M. mohrii*, *O. punctata*, *P. dulcis*, *T. cacao*, *S. asiatica*, *A. rufa*, *H. trionum*, *V. angularis*, *and A. annua*. This enabled the identification of potential CRF genes in the four rice species (Fig. S1). The results showed that 9, 30, 13, and 13 members potentially belong to CRF genes in the AP2/ERF members of Os, Oi, Osn, and Or rice species, respectively. Further, we conducted multiple alignments of the AP2/ERF members obtained from the four rice species with the protein sequences of *Arabidopsis* thaliana CRF members, identifying members belonging to the CRF-type transcription factors (Fig. S2). The results indicated that 7, 6, 6, and 7 AP2/ERF members were determined to be CRF genes in *Os*, *Oi*, *Osn*, and *Or* rice species, respectively. These members were renamed based on their chromosomal positions in the respective genomes, as *OsCRF1* to *OsCRF6* (a pair of homologous genes renamed as *OsCRF2a/2b*), *OiCRF1* to *OiCRF6*, *OsnCRF1* to *OsnCRF6*, and *OrCRF1* to *OrCRF7* (Table S1). Further analysis of the physicochemical properties of the CRF members in the four rice species revealed that the protein lengths ranged from 241 amino acids (aa) (*OsCRF4*) to 382 aa (*OsnCRF2*), and the protein molecular weights ranged from 25487.02 Dalton (Da) (*OsCRF4*) to 40869.68 Da (*OsnCRF2*). The isoelectric points of 20 members were less than 7, while 6 members had isoelectric points greater than 7. The instability index of the proteins ranged from 43.99 (*OiCRF1*) to 90.17 (*OsCRF5*), and the aliphatic index ranged from 56.62 (*OiCRF3*) to 75.65 (*OsnCRF6*), indicating hydrophilic proteins. Subcellular localization analysis revealed that 21 members were localized in the nucleus, 3 members in the chloroplast, and 1 member in the cytoplasm (Table S2). Due to the protein sequence of *OsCRF1* not starting with the amino acid M (non-start codon), the subcellular distribution of the protein could not be predicted.

Phylogenetic analysis

We constructed Neighbor-Joining phylogenetic trees utilizing the protein sequences of CRF members from five species: *Os*, *Oi*, *Osn*, *Or*, and *Arabidopsis thaliana* (*At*). The phylogenetic analysis revealed four distinct groups, labeled as Group I through Group IV (Fig. 1). In Group I, three CRF members were identified from each of the *Os*, *Oi*, *Osn*, and *Or* species. Group II was exclusively composed of eight AtCRF members. Group III included 2, 2, and 3 CRF members from *Os*, *Oi*, and *Or*, respectively. Finally, Group IV comprised 2, 3, and 2 CRF members from *Os*, *Osn*, and *Or*, respectively, along with 4 AtCRF members. These findings suggest a high degree of conservation among CRF members across the different rice varieties.

Protein conserved motif and gene structure analysis

Based on the protein sequences of CRF members from *Os*, *Oi*, *Osn*, and *Or* rice varieties, we constructed a Neighbor-joining phylogenetic tree. The clustering results for each member were consistent with the clustering results in Fig. 1, and they were grouped accordingly (Fig. 2A). Using the MEME tool, we annotated 10 motifs in the protein sequences of CRF members from the four rice varieties. Each motif's conserved sequence was identified using the NCBI-CDD database. It was found that Motif 1 and Motif 2 corresponded to the AP2 domain, while Motif 3 to Motif 10 were not identified as any domain (Fig. S3). Except for the protein *OiCRF2*, all other members contained Motif 2 (Fig. 2B). The CRF member proteins in Group I were divided into two categories based on motif type and location. *OiCRF2*, *OsCRF2a*, *OsCRF2b*, *OrCRF2*, and *OsnCRF2* proteins corresponded to nine types of motifs, namely Motif 10, Motif 3, Motif 6, Motif 2, Motif 1, Motif 4, Motif 9, Motif 7, and Motif 5. *OiCRF3*, *OsnCRF3*, *OrCRF3*, *OsCRF3*, *OsnCRF4*, *OiCRF5*, and *OrCRF6* proteins mainly corresponded to five

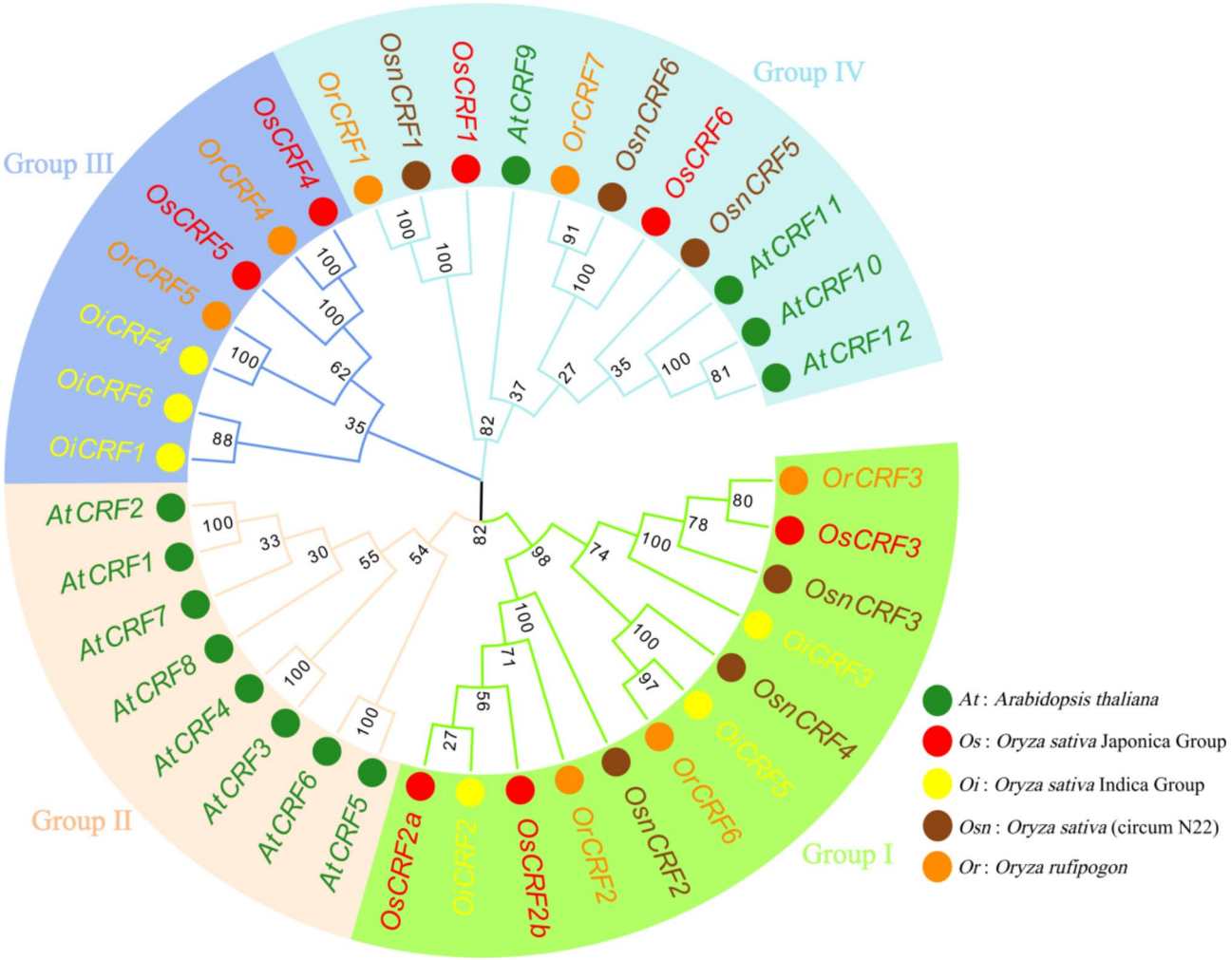

Fig. 1. Phylogenetic tree of CRF members from *Os*, *Oi*, *Osn*, *Or*, and *Arabidopsis thaliana*. Different colored regions represent different groups, and solid circles in species colors represent the corresponding species names.

types of motifs, namely Motif 3, Motif 2, Motif 1, Motif 4, and Motif 8. The seven CRF member proteins in Group II mainly contained three types of motifs, namely Motif 3, Motif 2, and Motif 1. The seven CRF member proteins in Group IV mainly contained five types of motifs, namely Motif 3, Motif 2, Motif 1, Motif 5, and Motif 7.

The gene structure of CRF members exhibits a high degree of conservation (Fig. 2C). In Group I, except for the genes *OiCRF2*, *OiCRF3*, and *OsCRF3*, which contain 2 exons and 1 intron, the remaining 9 genes have only 1 exon and 0 introns. In Group II, the genes *OiCRF1* and *OiCRF6* have 2 exons and 1 intron, while the other 5 genes have only 1 exon and 0 introns. In Group III, all seven genes have only 1 exon and 0 introns. In summary, the conservation of protein motifs and gene structure suggests a high degree of conservation of CRF genes among different rice varieties during the evolutionary process.

Gene chromosome location, collinearity, and selection pressure analysis

OsCRF members are distributed on chromosomes 1, 3, 7, and 9, with each chromosome containing 4, 1, 1, and 1 gene, respectively (Fig. 3A). OiCRF members are distributed on chromosomes 1, 5, and 6, with each chromosome containing 3, 1, and 2 genes, respectively (Fig. 3B). OsnCRF members are distributed on chromosomes 1, 6, 8, and 9, with each chromosome containing 3, 1, 1, and 1 gene, respectively (Fig. 3C). OrCRF members are distributed on chromosomes 1, 3, 5, 6, and 9, with each chromosome containing 3, 1, 1, 1, and 1 gene, respectively (Fig. 3D). The CRF members of the four rice species are all distributed in regions of relatively high gene density on the chromosomes. It is worth noting that all four rice species have a relatively high distribution on chromosome 1, each with three genes (*OsCRF2a/2b* is considered one gene), and these three genes are also relatively close in chromosomal distribution. Further analysis of chromosomal collinearity within the four rice species identified only one pair of segmental duplicated genes, *OsCRF4/OsCRF5*, in *Oryza sativa*.

We selected Os as the main reference species and identified homologous genes between Os and CRF members of *Arabidopsis thaliana*, *Oi*, *Osn*, and *Or* (Fig. 4). The results showed that Os shares one pair of homologous genes with *Arabidopsis thaliana*, which is *OsCRF3/AtCRF6*. Additionally, Os shares 2 (*OsCRF2a/OiCRF2*,

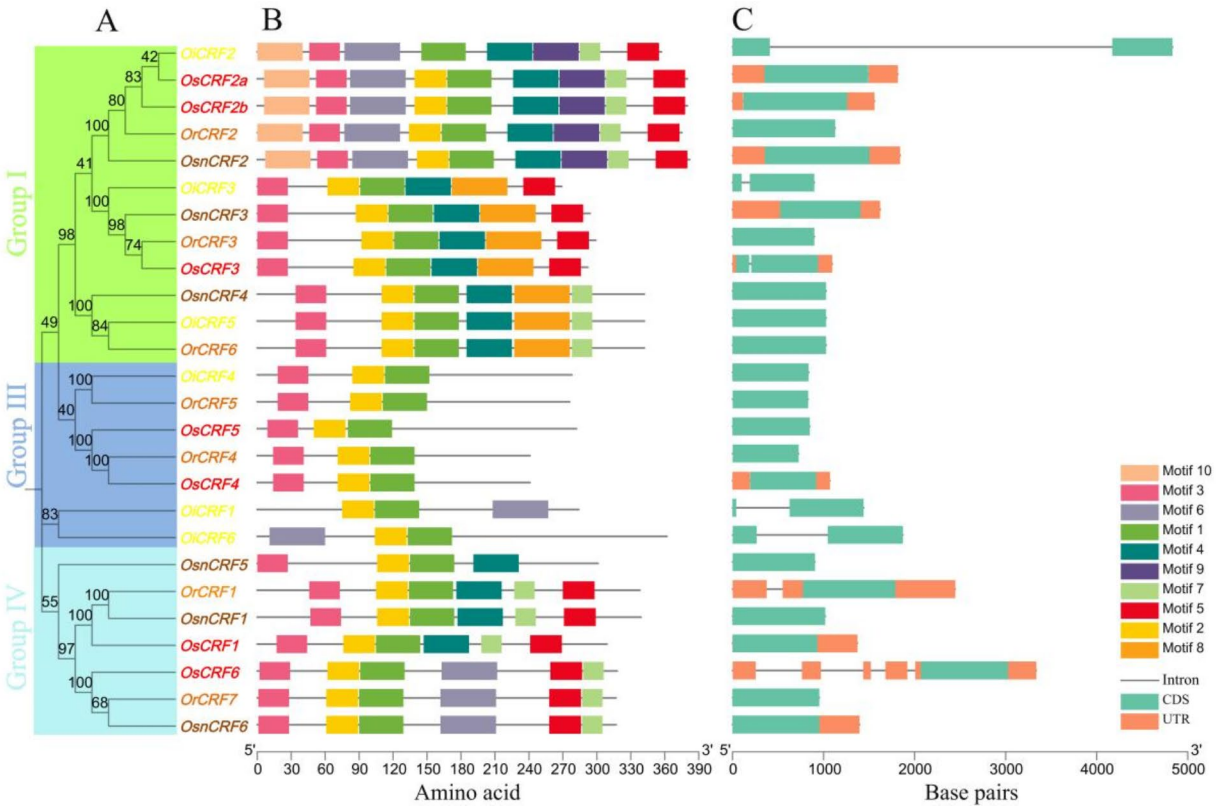

Fig. 2. Conservation motifs and gene structure of *Os*, *Oi*, *Osn*, *Or*, and CRF members. A: Neighbor-joining phylogenetic tree of *Os*, *Oi*, *Osn*, *Or*, and CRF members. B: Conserved protein motifs. C: Gene structure.

OsCRF3/OiCRF3), 4 (*OsCRF3/OsnCRF3*, *OsCRF1/OsnCRF1*, *OsCRF2a/OsnCRF2*, and *OsCRF6/OsnCRF6*), and 6 (*OsCRF1/OrCRF1*, *OsCRF2a/OrCRF2*, *OsCRF3/OrCRF3*, *OsCRF4/OrCRF4*, *OsCRF5/OrCRF4*, and *OsCRF6/OrCRF7*) pairs of homologous genes with *Oi*, *Osn*, and *Or*, respectively. It is noteworthy that *Os* shows high homology and conservation in chromosomal distribution with the other three rice species, indicating low separation during evolution. The results with *Arabidopsis thaliana* suggest a greater degree of separation between dicots and monocots in terms of CRF genes, consistent with the results shown in Fig. 1.

We conducted Ka, Ks, and Ka/Ks value calculations for segment duplication and homologous gene pairs to explore the evolutionary pressures affecting CRF genes (Table 1). The results showed that for six pairs of homologous genes, including *OsCRF3/OiCRF3*, *OsCRF1/OsnCRF1*, *OsCRF2a/OsnCRF2*, *OsCRF1/OrCRF1*, *OsCRF3/OrCRF3*, and *OsCRF4/OrCRF4*, both Ka and Ks were 0, indicating no synonymous or non-synonymous mutations occurred, resulting in a Ka/Ks value of 0. For the pair *OsCRF2a/OiCRF2*, the Ka/Ks value was greater than 1, while for the remaining six pairs of homologous genes, the Ka/Ks value was less than 1. The combination of homologous gene pairs further validates the high conservation of CRF genes among rice varieties during the evolutionary process.

Os, *Oi*, *Osn*, and *Or* rice CRF members' upstream 2000 bp promoter regions were collectively annotated with 86 cis-acting elements, of which 34 had confirmed functionalities (Table S3). The annotation revealed a plethora of basic elements such as CAAT-box and TATA-box, alongside various light-responsive elements including ACE, G-box, Box 4, and ATCT-motif. We focused on three functional cis-acting elements: Plant growth and development, Abiotic and biotic stresses, and Phytohormone responsive, for positional distribution display (Fig. 5, Table S4). Among them, the Plant growth and development-related elements comprised seven cis-acting elements: CAT-box, O2-site, NON-box, circadian, GCN4_motif, MBSI, and HD-Zip 1. Abiotic and biotic stresses-related elements included five cis-acting elements: ARE, MBS, GC-motif, LTR, and WUN-motif. Phytohormone responsive-related elements encompassed eight cis-acting elements: TGACG-motif, ABRE, CGTCA-motif, TCA-element, GARE-motif, TGA-element, P-box, and AuxRR-core. Notably, the conservation of cis-acting element types and distribution positions among different rice CRF members on the same branch of the phylogenetic tree was observed, such as *OsnCRF4*, *OiCRF5*, and *OrCRF6* genes, *OrCRF7*, and *OsnCRF7*. In summary, the results of promoter cis-acting elements indicate the extensive involvement of CRF members in rice growth and development processes.

Analysis of the expression patterns of OsCRF members

We utilized heatmaps to illustrate the FPKM values obtained from root, stem, leaf, and leaf samples under low-temperature stress (Table S5). The results (Fig. 6A) showed that the FPKM values corresponding to *OsCRF2b* in roots, stems, and leaves were all less than 1, suggesting that this gene may not be expressed in rice tissues.

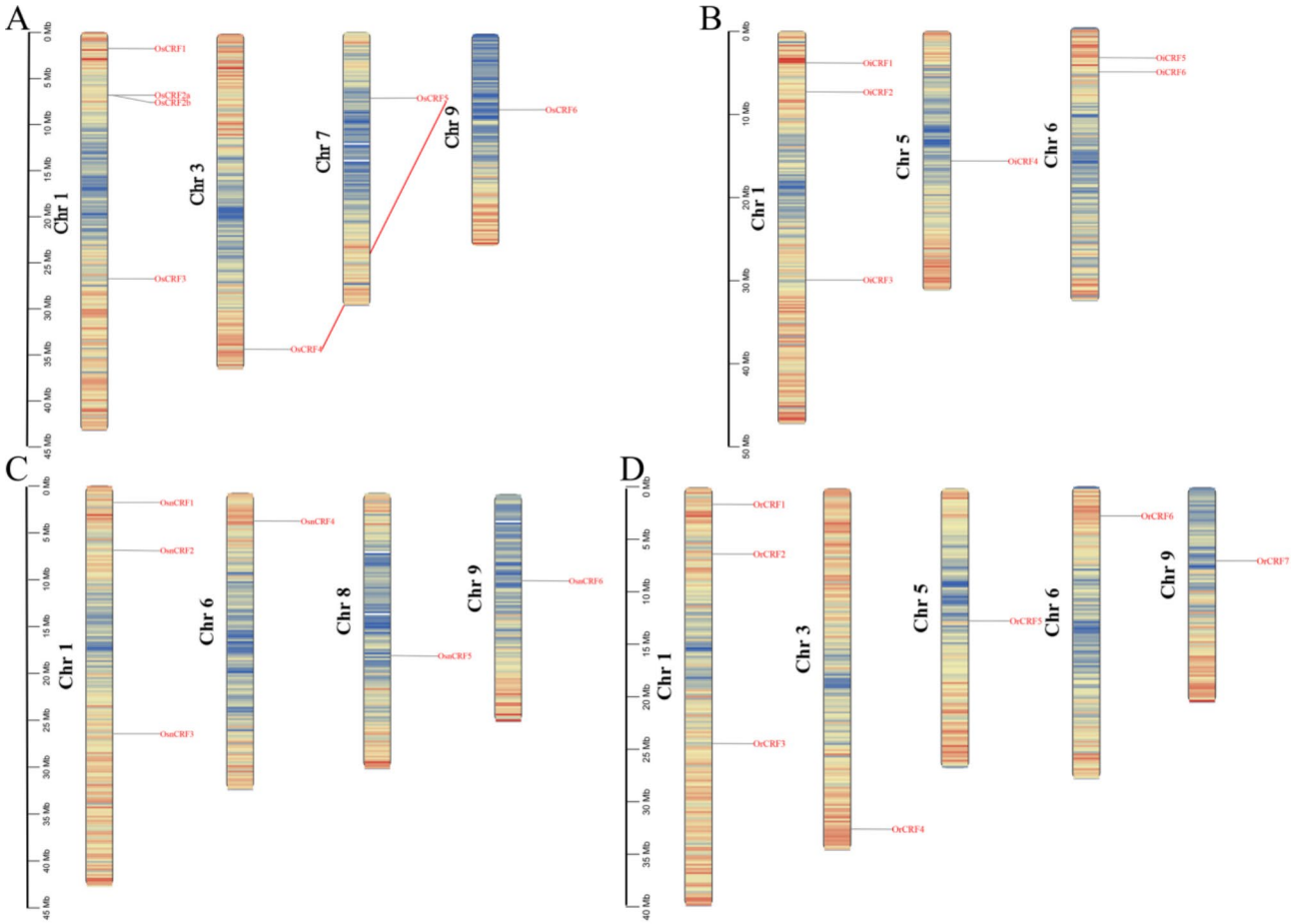

Fig. 3. Chromosomal distribution of CRF members in *Os*, *Oi*, *Osn*, and *Or*. A: *Oryza sativa Japonica* Group. B: *Oryza sativa Indica* Group. C: *Oryza sativa* (circum-Aus1 var. N22). D: *Oryza rufipogon*. The left scale represents chromosome lengths, with "Chr" indicating chromosomes. Genetic distances were set at 200 kb to calculate the gene density of each chromosome, represented by a gradient color from blue (low gene density) to red (high gene density). Blank regions indicate genetic regions lacking gene distribution information. Red lines indicate segmental duplication gene pairs.

OsCRF2a exhibited the highest FPKM value in leaves, indicating potentially high expression activity in leaves. *OsCRF3* showed FPKM values greater than 1 in stems and leaves, while *OsCRF6* showed FPKM values greater than 1 in roots and stems, and *OsCRF5* displayed FPKM values greater than 1 in all three tissues, suggesting the involvement of these three genes in the growth and development processes of different rice tissues. It is noteworthy that both *OsCRF1* and *OsCRF4* exhibited high FPKM levels in all three tissues, with particularly high expression levels in leaves, suggesting that these two genes may mainly participate in leaf growth, development, and related functions. According to the results obtained under low-temperature stress (Fig. 6B), *OsCRF1*, *OsCRF2a*, *OsCRF4*, *OsCRF5*, and *OsCRF6* all responded to low temperatures, showing a significant increase in FPKM. However, *OsCRF2b* and *OsCRF3* showed no response to low-temperature stress. This result suggests that CRF genes may be involved in rice's response to low-temperature stress.

Fluorescence quantification analysis

To investigate whether OsCRF members occur in response to low-temperature stress (4 °C), we selected leaves of Longdao 18 variety of three-leaf stage plants for quantitative fluorescence analysis (Fig. 7). The results showed that the expression levels of two genes, *OsCRF2b* and *OsCRF3*, were less than 1 in all six time-point samples. five genes, *OsCRF1*, *OsCRF2a*, *OsCRF4*, *OsCRF5*, and *OsCRF6*, all showed significantly higher expression levels with the lapse of time of the low-temperature stress treatments, which was characterized by a significant temporal gradient of expression. In addition to this, we further explored whether melatonin and cytokinin treatments would affect the expression levels of OsCRF members in rice leaves after low-temperature stress treatment. The results showed that the expression levels of two genes, *OsCRF2b* and *OsCRF3*, varied little in the six time point samples and were less than 1. The expression levels of three genes, *OsCRF4*, *OsCRF5* and *OsCRF6*, increased slightly in the six time point samples. the expression levels of *OsCRF2a* were significantly increased in all of the 48 h samples and increased slightly in all of the other time point samples. It is noteworthy that the expression level of *OsCRF1* gene was greatly elevated after melatonin treatment corresponding to the four time points of

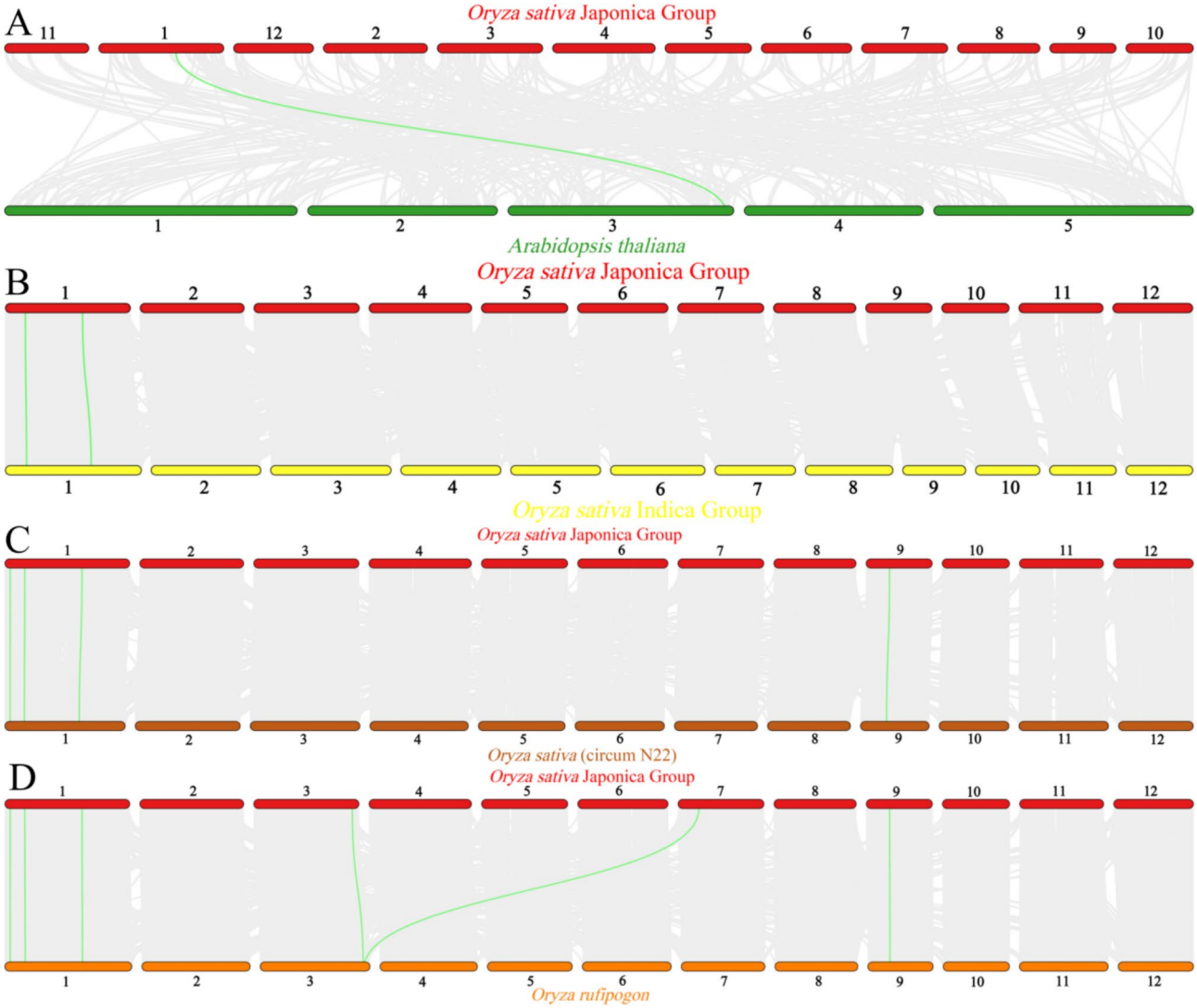

Fig. 4. Distribution of CRF orthologous gene pairs among species. Each horizontal line represents a chromosome, with numbers indicating chromosome numbers. Green lines represent CRF orthologous gene pairs.

6, 12, 24 and 48 h. In summary, the expression level of CRF genes in rice under low temperature environment occurs a strong response and will be induced by melatonin and cytokines.

Target Gene Prediction Analysis

We obtained the binding motif of the CRF4 transcription factor from the JASPAR database and identified 23,301 target genes by searching the 2000 bp upstream promoter sequences of Os genes for CRF4 binding sites. These target genes include thirteen matching sequences: ACGCCGCC, CAGCCGCC, CCGCCGCC, CCGCCGTC, GAGCCGCC, GCGCCGAC, GCGCCGCA, GCGCCGCC, GCGCCGCG, GCGCCGGC, GCGCCGTC, GGG CCGCC, and CGCCGCC. Most of the target genes contain multiple matching sequences in their promoter sequences. (Fig. S4, Table S6). The results indicate that 16,284 target genes have clear structural domains, 7,936 target genes have annotated Gene Ontology (GO) functions, and 8,632 target genes are annotated to KEGG pathways (Fig. 8, Table S7). According to the GO enrichment analysis, in biological processes, most target genes are mainly enriched in various functions such as cellular process (GO:0009987), metabolic process (GO:0008152), and developmental process (GO:0032502). In molecular function, most target genes are primarily enriched in catalytic activity (GO:0003824), binding (GO:0005488), and signal transducer activity (GO:0004871). In cellular components, most target genes are mainly enriched in functions such as cell (GO:0005623), organelle (GO:0043226), and membrane (GO:0016020). The KEGG annotation results show that most target genes are mainly enriched in metabolism-related pathways. Target genes are enriched in various pathways related to growth, development, and stress responses, such as Carbon metabolism (259 target genes, ko01200), MAPK signaling pathway (179 target genes, ko04010), and Starch and sucrose metabolism (152 target genes, ko00500), among others.

Gene pairs	Ka	Ks	Ka/Ks
OsCRF4/OsCRF5	0.447557027	0.75014049	0.596630942
OsCRF3/AtCRF6	0.175790755	3.444685171	0.051032459
OsCRF2a/OiCRF2	0.005067587	0.003631968	1.395272852
OsCRF3/OiCRF3	0	0	NaN
OsCRF3/OsnCRF3	0.006176051	0.013413174	0.460446668
OsCRF1/OsnCRF1	0	0	NaN
OsCRF2a/OsnCRF2	0	0	NaN
OsCRF6/OsnCRF6	0.00135318	0.004746851	0.28506904
OsCRF1/OrCRF1	0	0	NaN
OsCRF2a/OrCRF2	0.001203128	0.003454238	0.348305054
OsCRF3/OrCRF3	0	0	NaN
OsCRF4/OrCRF4	0	0	NaN
OsCRF5/OrCRF4	0.447557027	0.75014049	0.596630942
OsCRF6/OrCRF7	0.00135318	0.004746851	0.28506904

Table 1. CRF genes replication and homologous gene Ka/Ks values. Promoter analysis

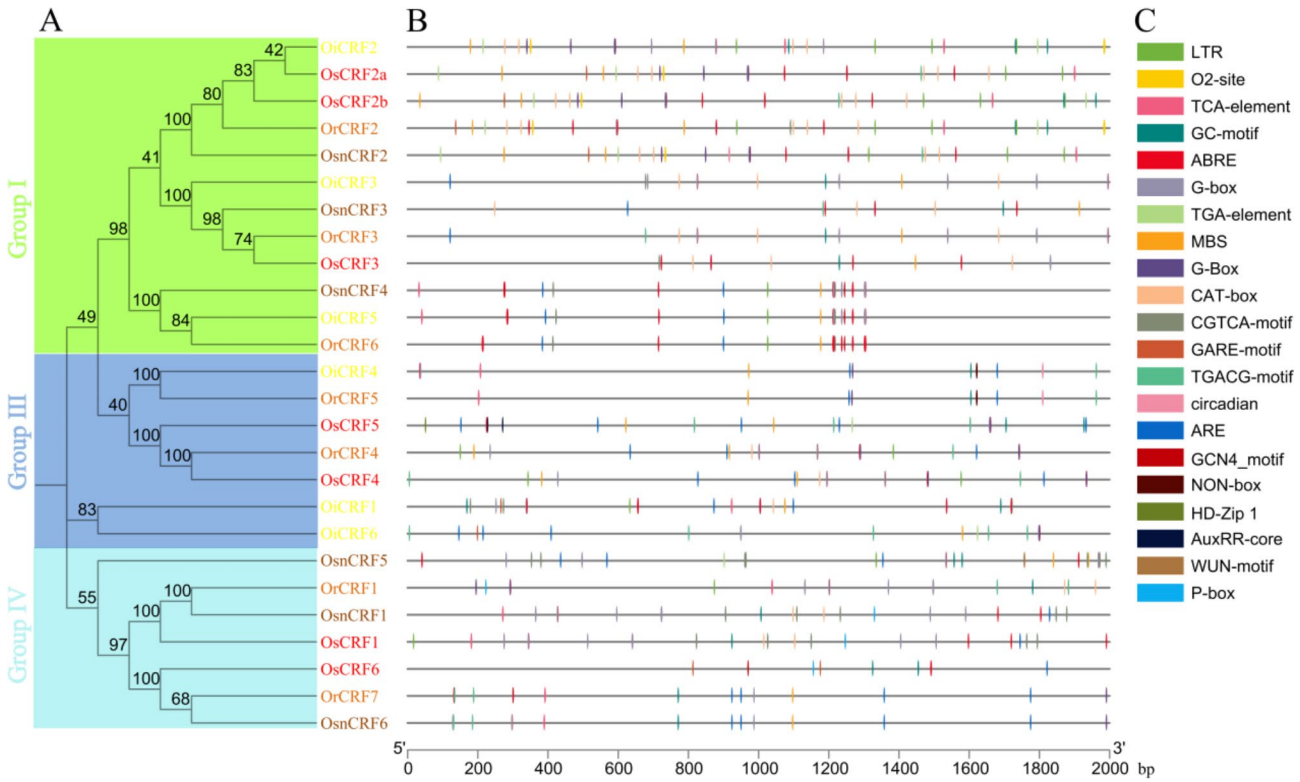

Fig. 5. Annotation of cis-acting elements in the upstream 2000 bp promoter regions of *Os*, *Oi*, *Osn*, and *Or* CRF members. A: Neighbor-joining phylogenetic tree of CRF members from *Os*, *Oi*, *Osn*, and *Or*. B: Distribution of cis-acting elements. C: Names of cis-acting elements.

Subcellular localization analysis

The subcellular localization of rice CRF proteins was predicted (Table S4), multiple CRF proteins localised to the nucleus, chloroplasts and cytoplasm. Using tobacco leaves, OsCRF4-GFP and OsCRF5-GFP were selected for transient expression assays. Plasmids of OsCRF4-GFP and OsCRF5-GFP were transiently expressed together with the nuclear markers. As shown in Fig. 9, the fluorescence signals of OsCRF4-GFP and OsCRF5-GFP co-localised extensively with the nuclear markers, indicating that OsCRF4-GFP and OsCRF5-GFP were localised at the plasma membrane and endoplasmic reticulum.

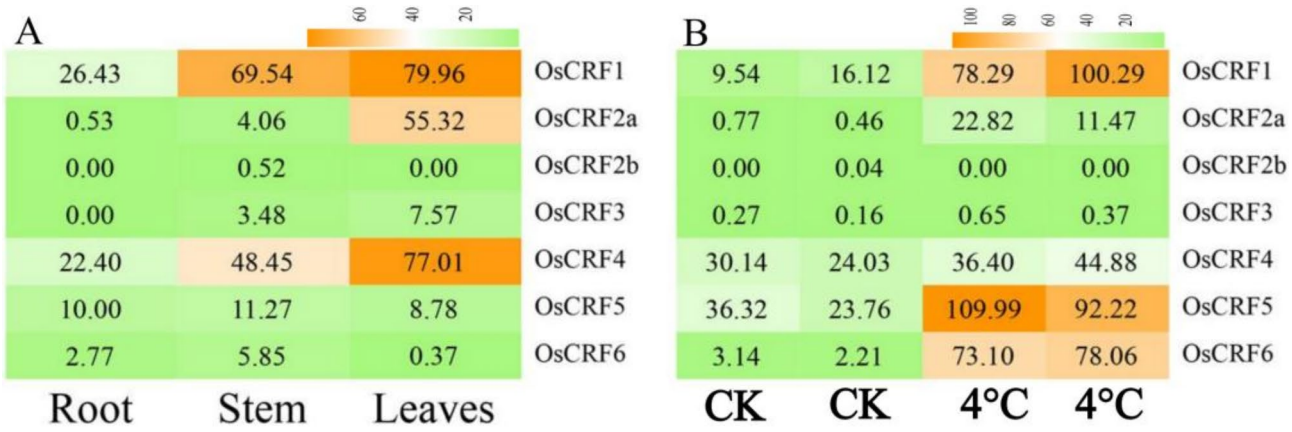

Fig. 6. Heatmaps illustrating FPKM values of OsCRF members. A: FPKM values of OsCRF members in roots, stems, and leaves. B: FPKM values of OsCRF members in leaves under normal temperature (CK) and low-temperature stress (4 °C). Green to yellow indicates an increase in FPKM values.

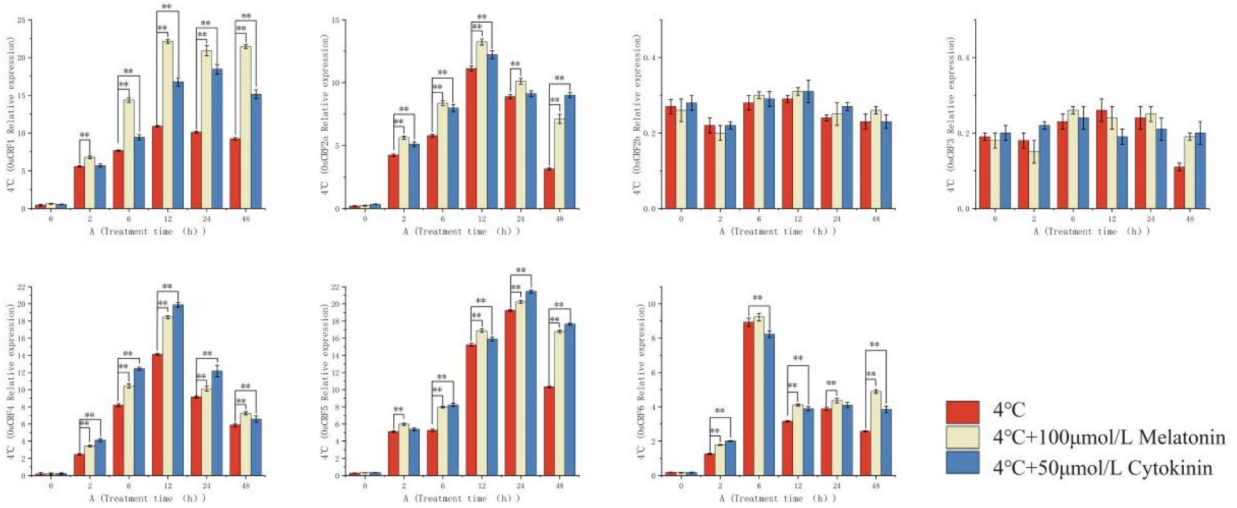

Fig. 7. Time gradient fluorescence quantification of expression levels of OsCRF members. Red color indicates normal low temperature stress (4 °C) treated samples. Yellow indicates melatonin-treated post low temperature stress (4 °C) treated samples. Blue color indicates cytokinin-treated post low-temperature stress (4 °C) treated samples.

Discussion

The AP2/ERF family, as one of the largest transcription factor families, has achieved significant research progress in various aspects such as growth and development, abiotic stress, and signal transduction. The CRF genes constitute a small fraction of the AP2/ERF transcription factor superfamily. Currently, CRF genes have been extensively studied in *Arabidopsis thaliana* and to a lesser extent in a few other plants such as *Solanum lycopersicum*[14], *Brassica rapa*[18], *Glycine max*[19], *Tamarix hispida*[20], and *Quercus suber*[9]. In this study, we conducted a preliminary screening of AP2/ERF members in four rice varieties, Os, Oi, Osn, and Or, and identified genes belonging to the CRF type transcription factors through multiple alignments with *Arabidopsis* CRF members.

We identified 7, 6, 6, and 7 CRF genes in the Os, Oi, Osn, and Or rice varieties, respectively. The number of CRF members varies across different plant species and does not correlate directly with genome size. For instance, *Arabidopsis thaliana* has 12 CRF genes, tomato has 11, and soybean has 26. Notably, there's significant diversity in the number of CRF genes between diploid and polyploid varieties of the same species; for example, *B. rapa* has 21 CRF genes, while the allotetraploid *B. napus* has 44 [21]. The physicochemical properties of CRF proteins are similar among the four rice varieties, with most CRF proteins localized in the nucleus, consistent with the subcellular localization of CRF members in *Solanum lycopersicum*, *Brassica rapa*, and *Brassica napus*[18]. Phylogenetic analysis revealed distinct clustering of CRF members from the four rice varieties with significant divergence from *Arabidopsis* CRF members, unlike the significant conservation observed between *B. napus* and

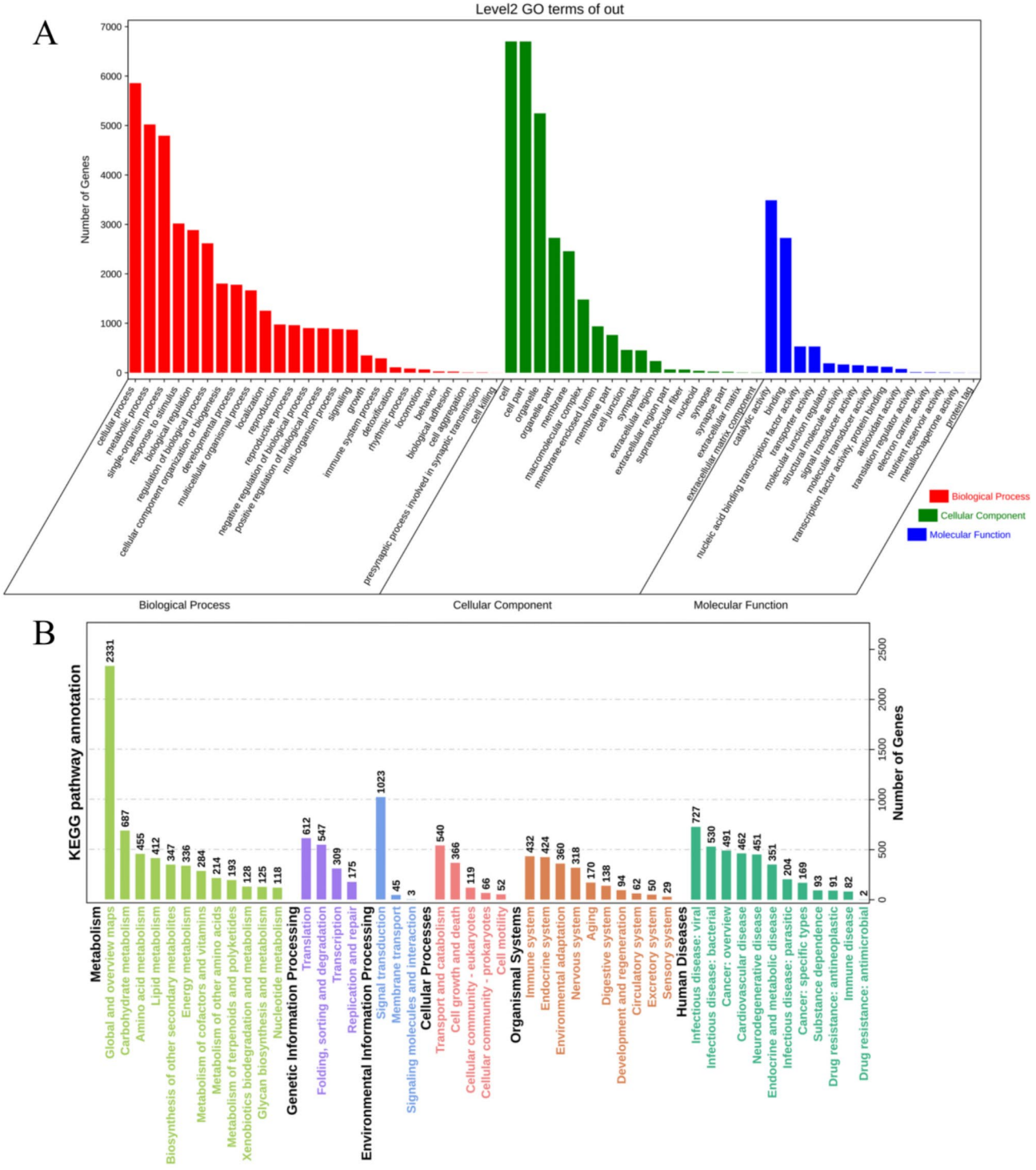

Fig. 8. Annotation of Target Genes with GO and KEGG. A: Annotation results of target genes with GO functions. B: Annotation results of target genes with KEGG pathways.

the 12 CRF genes of *Arabidopsis*. This suggests that there is a high degree of divergence in the CRF gene during the evolution of rice and *Arabidopsis*.

Gene function is closely associated with conserved motifs within protein sequences[22]. Based on the phylogenetic tree, it's evident that CRF members from Group III and Group IV in Os, Oi, Osn, and Or have higher conservation of motifs. However, for the five CRF members in Group I (*OiCRF2*, *OsCRF2a*, *OsCRF2b*, *OrCRF2*, and *OsnCRF2*), corresponding to nine conserved motifs, this result aligns with the findings for CRF3 and CRF4 members in the phylogenetic tree of *Brassica napus*, where members of these two groups corresponded to nine motifs, while the other seven CRF members corresponded to CRF6 group members. Based on this, we

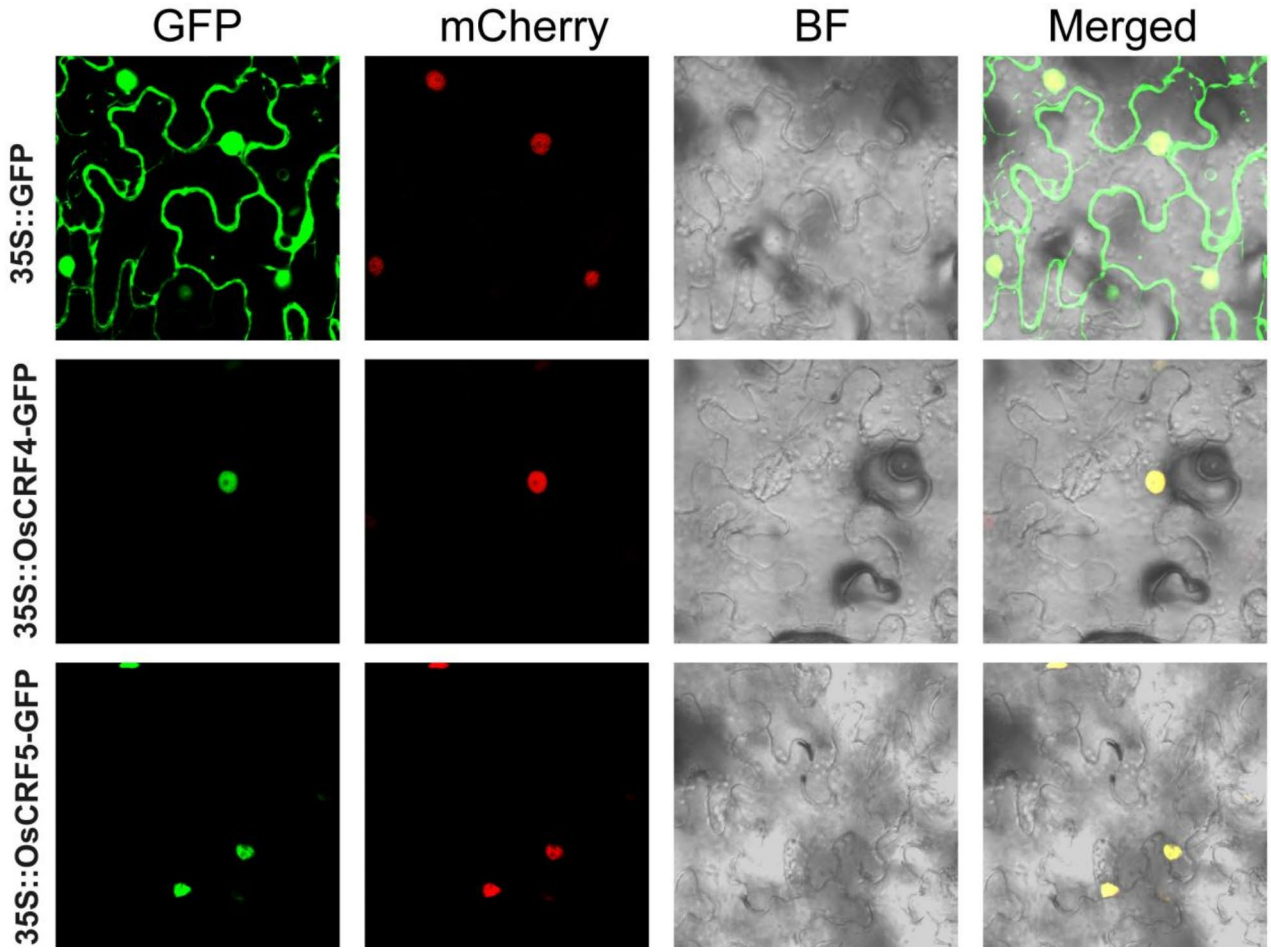

Fig. 9. Transient expression of OsCRF4-GFP and OsCRF5-GFP with nuclearbiomarker fusion proteins in tobacco epidermal cells.

infer that the degree of sequence variation in CRF genes in rice may not be as high as in dicotyledonous plants. Gene structure information can provide clues for the evolution of gene family members. The gene structure of CRF members in the four rice varieties is highly conserved, with 21 members having one exon and five members having two exons, consistent with the gene structure of corresponding CRF members in *Brassica rapa* and *Brassica napus*[21].

The chromosomal location of a gene may influence its functional expression[23]. Chromosomal localization results indicate that CRF1, CRF2, and CRF3 genes in Os, Oi, Osn, and Or rice varieties are all distributed on chromosome 1, with relatively close positions, while the remaining genes are distributed on other chromosomes. Gene duplication is a major factor in the expansion and evolution of gene members, contributing to species adaptation to environmental changes and maintaining normal life processes[24]. We identified only one pair of segmental duplicate genes, *OsCRF4/OsCRF5*, in Os, indicating that gene duplication is not the main cause of CRF member expansion in rice. Inter-species collinearity results show that Os has 2, 3, and 6 pairs of homologous genes with Oi, Osn, and Or, respectively, with 2, 3, and 3 pairs of homologous genes on chromosome 1, indicating the high conservation of CRF members on chromosome 1 during the evolution of rice. It is noteworthy that Os and Or share 6 pairs of CRF homologous genes, indicating a close phylogenetic relationship between Or, originating from Asia, and Os. Similarly, the Ka and Ks values of the 6 pairs of homologous genes are all 0, indicating that some CRF member sequences among rice varieties are highly conserved, and the Ka/Ks values suggest that CRFs in rice have undergone strong purifying selection pressure during evolution, indicating similar functions.

Cis-acting elements in the promoter region regulate gene expression[25]. Annotation of promoter sequences reveals that CRF genes in Os, Oi, Osn, and Or rice varieties are involved in various functions such as growth, stress response, and hormone regulation. Investigating gene expression during tissue development and under adverse environmental conditions is important for understanding the molecular mechanisms of biological development[26]. Soybean CRF members are expressed in multiple tissues and respond to cold stress. For example, *GmCRF15* and *GmCRF25* are significantly expressed in seeds, *GmCRF6* and *GmCRF8* are expressed significantly in root hairs, and GmCRF3 and *GmCRF15* are significantly expressed under cold stress[19]. CRF members in *B. rapa* show significant expression in nutritional assemblies and reproductive tissues. Some CRF genes in *B. napus* can enhance its tolerance to low phosphorus stress, such as the BnaCRF7. Similarly, FPKM values obtained from

transcriptome data show that OsCRF members are expressed in roots, stems, and leaves, and also participate in resisting low-temperature stress. Fluorescence quantification showed that two genes, *OsCRF2b* and *OsCRF3*, might not be expressed in rice leaves. In contrast, five genes, *OsCRF1*, *OsCRF2a*, *OsCRF4*, *OsCRF5*, and *OsCRF6*, increased their expression levels over time under low-temperature stress in rice and were affected by melatonin and cytokinin to increase their expression levels. Taken together, CRF may have a function in rice to resist low-temperature stress.

Transcription factors regulate gene expression by binding to specific sequences in the promoter regions of target genes[27]. Numerous studies have identified target genes corresponding to AP2-type transcription factors. For instance, Ag*DREB1* and *AgDREB2* may act as transcriptional activators by binding to corresponding DRE elements to enhance celery's stress resistance[28]. Sm128 and Sm152 can interact with ERF to regulate the biosynthesis of tanshinones and phenolic acids[29]. AP2/ERF can interact with WRKY, bHLH, bZIP, MYB, NAC, and C2H2 to enhance plant resistance to cold stress[30]. However, there are no reports on the regulation of target genes by CRF-type transcription factors. In this study, we obtained the binding motif of the CRF4 transcription factor through the JASPAR database and identified a total of 23,301 target genes in Os, including 13 matching sequences. This indicates that CRF can bind to multiple sequences, demonstrating a broader range of target gene binding functions. Enrichment analysis of Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways revealed that target genes possess various protein functions related to stress resistance and are distributed across pathways associated with stress defense. This suggests that CRF genes regulate multiple pathways through the modulation of downstream genes, highlighting the importance of identifying and exploring these CRF-corresponding target genes in the rice genome.

Conclusions

Seven, six, six, and seven CRF genes were identified in the genomes of the four rice subspecies, respectively. These CRF members shared consistent features. Phylogenetic analysis classified these members into three distinct groups, revealing a high degree of conservation in protein motifs and gene structures within each group. Additionally, the amplification of CRF genes was not predominantly driven by segmental or tandem repeats, and there was significant divergence between CRF members in monocotyledons and dicotyledons. Promoter element analysis indicated that CRF genes contain cis-acting elements involved in growth, development, hormone regulation, and stress response in rice. OsCRF members were expressed in roots, stems, and leaves, responding to low-temperature stress with increased expression levels induced by melatonin and cytokinin. The nuclear localization of OsCRF4/5 was confirmed as predicted. These findings provide valuable insights into the characterization of rice CRF genes and their response mechanisms to low-temperature stress.

Materials and methods

Identification of AP2/ERF members and CRFs in four rice

Genome data for *Oryza sativa Japonica* Group (GCA_001433935.1, referred to as *Os*), *Oryza sativa Indica* Group (GCA_000004655.2, referred to as *Oi*), *Oryza sativa* (circum-Aus1 var. N22) (GCA_001952365.2, referred to as *Osn*), and *Oryza rufipogon* (GCA_000817225.1, referred to as *Or*) were obtained from the Ensembl Plants database (http://plants.ensembl.org/species.html) for the identification and analysis of CRF family members[31]. Using the Pfam database (http://pfam.xfam.org/), the hidden Markov model (HMM) of the AP2 domain (PF00847) was downloaded. The HMMER tool was then employed to search and align the protein sequences of the four rice genomes against the AP2 domain hidden Markov model, retaining protein sequences with an E-value of $\leq$-5 [32,33]. As the primary focus here is to study CRF-type members, no redundancy removal of sequences was conducted, and only screening of AP2/ERF members was performed.

We obtained the protein sequences of 12 *Arabidopsis* CRF members from the UniProt database. To ensure that the selected genes from the AP2/ERF members of the four rice species belong to the CRF-type transcription factors, we further retrieved protein sequences of CRF members from thirteen additional species, including *Glycine max*, *Populus tomentosa*, *Tamarix hispida*, *Marshallia caespitosa*, *Marshallia mohrii*, *Oryza punctata*, *Prunus dulcis*, *Theobroma cacao*, *Striga asiatica*, *Actinidia rufa*, *Hibiscus trionum*, *Vigna angularis*, and *Artemisia annua*, from the NCBI database (Table S8). Using the MEGA tool, multiple sequence alignment was performed with the MUSCLE algorithm for the AP2/ERF family members of the four rice species and the CRF member protein sequences of fourteen other species. Subsequently, the neighbor-joining method was employed to construct phylogenetic trees with 1,000 bootstrap replications, using the Poisson correction model and pairwise deletion. Next, potential CRF-type transcription factors selected from the AP2/ERF members of the four rice species were subjected to MUSCLE multiple sequence alignment analysis. Based on the structural domain features of CRF-type transcription factors, non-CRF-type members were removed. Finally, the physicochemical properties of the CRF members from the four rice species were analyzed using ExPASy (http://web.expasy.org/protparam/), and protein subcellular localization analysis was conducted using WoLF PSORT II (https://www.genscript.com/wolf-psort.html?src=leftbar).[34,35].

Construction of phylogenetic trees

Using the MEGA X tool, multiple sequence alignment was performed with the MUSCLE algorithm for the CRF members of the four rice species and the AtCRF protein sequences of *Arabidopsis*. Subsequently, phylogenetic trees were constructed using the neighbor-joining method with 1,000 bootstrap replications, employing the Poisson correction model and pairwise deletion[36]. The generated phylogenetic trees were then visualized and refined using the Evolview tool[37].

Conserved protein motifs and gene structures

The protein sequences of the CRF family members from the four rice species were annotated for conserved motifs using the MEME online tool (http://meme-suite.org/tools/meme). The number of motifs was set to 10, with other parameters kept at default values[38]. The exon and intron positions and quantities of the CRF family members were determined using the General Feature Format (GFF) annotation information from the four rice genomes. Finally, the evolutionary trees, motifs, and gene structures of the CRF family members from the four rice species were clustered and visualized using the TBtools software[39].

Gene chromosomal localization, collinearity, and identification of homologous gene pairs

Using the TBtools software, the chromosomal distribution positions of CRF members in the four rice species were plotted. The MCScanX tool was employed to identify segmental duplications and tandem duplicate gene pairs among the CRF family members of the four rice species[40]. For interspecies collinearity analysis, Oryza sativa (Os) was selected as the reference, and homologous CRF genes were analyzed with *Arabidopsis*, *Oryza indica* (Oi), *Oryza sativa* N22 (Osn), and *Oryza rufipogon* (Or). Finally, the Ka (nonsynonymous substitution rate), Ks (synonymous substitution rate), and Ka/Ks values were calculated for all types of duplicated genes using the KaKs Calculator tool[41].

Annotation of Cis-acting elements

Using the TBtools software, the upstream 2000 bp promoter sequences of CRF members from the four rice species were extracted. These sequences were then annotated for cis-acting elements using the PlantCARE online tool (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/)[42]. Functional categories such as abiotic and biotic stresses, phytohormone responsiveness, and plant growth and development were selected based on previous classifications. Subsequently, the distribution of cis-acting elements in the promoter regions of CRF members from the four rice species was plotted and analyzed using the TBtools software[43].

Expression patterns of OsCRF members

We analyzed the differential tissue expression patterns of OsCRF members using root and leaf transcriptome data obtained from our sequencing efforts[44,45], as well as stem transcriptome data from the NCBI SRA database (SRR26372771)[46]. Additionally, transcriptome data of leaves subjected to low-temperature stress (SRR22266757 and SRR22266758 at 4 °C, and SRR22266759 and SRR22266760 at normothermia (CK)) were retrieved from the NCBI SRA database[47]. These data were utilized to obtain FPKM (Fragments Per Kilobase of exon model per Million mapped fragments) values, allowing for the analysis of OsCRF member expression patterns under low-temperature stress.

Fluorescence quantification

We selected three-leaf stage plants of the Japonica rice variety Longdao 18 and implemented three treatment conditions: (1) exposure to 4 °C low-temperature stress for 0, 2, 6, 12, 24, and 48 h; and (2) separate treatments involving spraying 100μM/L melatonin and 50μM/L cytokinin solutions, followed by the same low-temperature stress intervals. Each treatment was replicated three times biologically, and the leaves were ultimately harvested and stored at -80 °C for subsequent analysis.

Three independent biological replicates, each consisting of three independent plants, were utilized for qRT-PCR detection. The qRT-PCR primers for the selected CRF genes were designed using Primer Premier 5 (Table S9). The fluorescence quantification procedure followed the previously described method[48]. The obtained cycle threshold (CT) values were quantitatively analysed by the 2-△△Ct method[49].

Prediction of target genes

We obtained the transcription factor binding motif (MA0976.1) of CRF4 from the JASPAR Plantae database (https://jaspar.elixir.no/search?q=&collection=CORE&tax_group=plants)[50]. Subsequently, using TBtools, we extracted the 2000 bp promoter sequences of all Os genes and identified the genes bound by the CRF4 transcription factor motif using the Motif FIMO tool (https://meme-suite.org/meme/)[36]. Finally, target gene domain prediction was conducted based on the PFAM database, and KEGG (Kyoto Encyclopedia of Genes and Genomes) and GO (Gene Ontology) enrichment analyses of the target genes were performed using OmicShare Tools (https://www.omicshare.com/tools).

Subcellular localization of OsCRF4/5

To determine the subcellular localisation of CRF proteins, the full-length CDS of *OsCRF4/5* was ligated to CaMV35S::GFP. and the CaMV35S::OsCRF4/5-GFP vector and nuclear marker were transformed into *Agrobacterium* EHA105 and co-transformed into *N. benthamiana* leaf blades. The fluorescence signals were observed by laser scanning confocal microscopy after 24 h of induction at 27 °C under dark conditions.

Data availability

Transcriptome sequencing data were obtained from the National Center for Biotechnology Information (NCBI) under the biological project numbers SRR26372771, SRR22266757, SRR22266758, SRR22266759 and SRR22266760.

Received: 21 July 2024; Accepted: 13 November 2024
Published online: 18 November 2024

References

1. Feng, K. et al. Advances in AP2/ERF super-family transcription factors in plant. *Crit. Rev. Biotechnol.* **40**, 750–776. https://doi.org/10.1080/07388551.2020.1768509 (2020).
2. Rashotte, A. M. et al. A subset of Arabidopsis AP2 transcription factors mediates cytokinin responses in concert with a two-component pathway. *Proc. Natl. Acad. Sci. U S A.* **103**, 11081–11085. https://doi.org/10.1073/pnas.0602038103 (2006).
3. Kim, J. CYTOKININ RESPONSE FACTORs gating environmental signals and hormones. *Trends Plant. Sci.* **21**, 993–996. https://doi.org/10.1016/j.tplants.2016.10.004 (2016).
4. Rashotte, A. M. & Goertzen, L. R. The CRF domain defines cytokinin response factor proteins in plants. *BMC Plant. Biol.* **10**, 74. https://doi.org/10.1186/1471-2229-10-74 (2010).
5. Keshishian, E. A. & Rashotte, A. M. Plant cytokinin signalling. *Essays Biochem.* **58**, 13–27. https://doi.org/10.1042/bse0580013 (2015).
6. Melton, A. E., Zwack, P. J., Rashotte, A. M. & Goertzen, L. R. Identification and functional characterization of the Marshallia (Asteraceae) clade III cytokinin response factor (CRF). *Plant. Signal. Behav.* **14**, 1633886. https://doi.org/10.1080/15592324.2019.1633886 (2019).
7. Cucinotta, M. et al. Cytokinin response factors integrate auxin and cytokinin pathways for female reproductive organ development. *Development*. **143**, 4419–4424. https://doi.org/10.1242/dev.143545 (2016).
8. Šimášková, M. et al. Cytokinin response factors regulate PIN-FORMED auxin transporters. *Nat. Commun.* **6**, 8717. https://doi.org/10.1038/ncomms9717 (2015).
9. Capote, T., Usié, A., Barbosa, P., Ramos, M. & Gonalves, S. Transcriptome dynamics of cork oak (Quercus suber) somatic embryogenesis reveals active gene players in transcription regulation and phytohormone homeostasis of embryo development. *Tree Genet. Genom* **15** (2019).
10. Zwack, P. J., Compton, M. A., Adams, C. I. & Rashotte, A. M. Cytokinin response factor 4 (CRF4) is induced by cold and involved in freezing tolerance. *Plant. Cell. Rep.* **35**, 573–584. https://doi.org/10.1007/s00299-015-1904-8 (2016).
11. Hallmark, H. T. & Rashotte, A. M. Review - cytokinin response factors: responding to more than cytokinin. *Plant. Sci.* **289**, 110251. https://doi.org/10.1016/j.plantsci.2019.110251 (2019).
12. Gupta, S. & Rashotte, A. M. Expression patterns and regulation of SlCRF3 and SlCRF5 in response to cytokinin and abiotic stresses in tomato (Solanum lycopersicum). *J. Plant. Physiol.* **171**, 349–358. https://doi.org/10.1016/j.jplph.2013.09.003 (2014).
13. Shi, X., Gupta, S. & Rashotte, A. M. Characterization of two tomato AP2/ERF genes, SlCRF1 and SlCRF2 in hormone and stress responses. *Plant. Cell. Rep.* **33**, 35–45. https://doi.org/10.1007/s00299-013-1510-6 (2014).
14. Shi, X., Gupta, S. & Rashotte, A. M. Solanum lycopersicum cytokinin response factor (SlCRF) genes: characterization of CRF domain-containing ERF genes in tomato. *J. Exp. Bot.* **63**, 973–982. https://doi.org/10.1093/jxb/err325 (2012).
15. Kwon, T. Cytokinin Response factor 2 positively regulates salicylic acid-mediated plant immunity in Arabidopsis thaliana. *Plant. Biotechnol.* (2016).
16. Varala, K. et al. Temporal transcriptional logic of dynamic regulatory networks underlying nitrogen signaling and use in plants. *Proc. Natl. Acad. Sci. U S A.* **115**, 6494–6499. https://doi.org/10.1073/pnas.1721487115 (2018).
17. Rashid, M., Guangyuan, H., Guangxiao, Y., Hussain, J. & Xu, Y. AP2/ERF transcription factor in Rice: genome-wide Canvas and Syntenic relationships between monocots and eudicots. *Evolutionary Bioinf. Online.* **8**, 321–355. https://doi.org/10.4137/ebo.s9369 (2012).
18. Bolser, D. M., Staines, D. M., Perry, E. & Kersey, P. J. Ensembl Plants: Integrating Tools for Visualizing, Mining, and Analyzing Plant Genomic Data. *Methods in molecular biology (Clifton, N.J.)* 1533, 1–31, doi: (2017). https://doi.org/10.1007/978-1-4939-6658-5_1
19. Mistry, J. et al. Pfam: the protein families database in 2021. *Nucleic Acids Res.* **49**, D412–d419. https://doi.org/10.1093/nar/gkaa913 (2021).
20. Finn, R. D., Clements, J. & Eddy, S. R. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* **39**, W29–37. https://doi.org/10.1093/nar/gkr367 (2011).
21. Gasteiger, E. et al. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* **31**, 3784–3788. https://doi.org/10.1093/nar/gkg563 (2003).
22. Horton, P. et al. WoLF PSORT: protein localization predictor. *Nucleic Acids Res.* **35**, W585–587. https://doi.org/10.1093/nar/gkm259 (2007).
23. Wang, S., Zhang, H., Shi, L., Xu, F. & Ding, G. Genome-wide dissection of the CRF Gene Family in Brassica napus indicates that BnaCRF8s specifically regulate Root Architecture and phosphate homeostasis against phosphate fluctuation in plants. *Int. J. Mol. Sci.* **21** https://doi.org/10.3390/ijms21103660 (2020).
24. Zhang, H., Gao, S., Lercher, M. J., Hu, S. & Chen, W. H. EvolView, an online tool for visualizing, annotating and managing phylogenetic trees. *Nucleic Acids Res.* **40**, W569–572. https://doi.org/10.1093/nar/gks576 (2012).
25. Bailey, T. L. et al. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* **37**, W202–208. https://doi.org/10.1093/nar/gkp335 (2009).
26. Chen, C. et al. TBtools: an integrative Toolkit developed for interactive analyses of big Biological Data. *Mol. Plant.* **13**, 1194–1202. https://doi.org/10.1016/j.molp.2020.06.009 (2020).
27. Wang, Y. et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* **40**, e49. https://doi.org/10.1093/nar/gkr1293 (2012).
28. Zhang, Z. et al. KaKs_Calculator: calculating Ka and Ks through model selection and model averaging. *Genom. Proteom. Bioinform.* **4**, 259–263. https://doi.org/10.1016/s1672-0229(07)60007-2 (2006).
29. Lescot, M. et al. 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**, 325–327. https://doi.org/10.1093/nar/30.1.325 (2002).
30. Zhang, D., Yu, Z., Zeng, B. & Liu, X. Genome-wide analysis of the ABC gene family in almond and functional predictions during flower development, freezing stress, and salt stress. *BMC Plant. Biol.* **24**, 12. https://doi.org/10.1186/s12870-023-04698-7 (2024).
31. Lei, L. et al. Identification of a major QTL and Candidate Gene Analysis of Salt Tolerance at the Bud Burst Stage in Rice (Oryza sativa L.) using QTL-Seq and RNA-Seq. *Rice (New York N Y).* **13**, 55. https://doi.org/10.1186/s12284-020-00416-1 (2020).
32. Yang, L. et al. Whole-genome mining of abiotic stress gene loci in rice. *Planta*. **252**, 85. https://doi.org/10.1007/s00425-020-03488-x (2020).
33. Gokulan, C. G. et al. Multiomics-assisted characterization of rice-yellow stem Borer interaction provides genomic and mechanistic insights into stem borer resistance in rice. *Theor. Appl. Genet.* **137**, 122. https://doi.org/10.1007/s00122-024-04628-7 (2024).
34. He, Z. et al. R-loops act as regulatory switches modulating transcription of COLD-responsive genes in rice. *New. Phytol.* **241**, 267–282. https://doi.org/10.1111/nph.19315 (2024).
35. Zhang, D. et al. Genome-wide identification of members of the Skp1 family in almond (Prunus dulcis), cloning and expression characterization of PsdSSK1. *Physiol. Mol. Biol. Plants*. **29**, 35–49. https://doi.org/10.1007/s12298-023-01278-9 (2023).
36. Rao, X., Huang, X., Zhou, Z. & Lin, X. An improvement of the 2^(-delta delta CT) method for quantitative real-time polymerase chain reaction data analysis. *Biostatistics Bioinf. Biomathematics*. **3**, 71–85 (2013).
37. Castro-Mondragon, J. A. et al. JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **50** (D165-d173). https://doi.org/10.1093/nar/gkab1113 (2022).
38. Liu, Z. et al. Genome-wide identification, phylogeny, evolution and expression patterns of AP2/ERF genes and cytokinin response factors in Brassica rapa ssp. pekinensis. *PLoS One*. **8**, e83444. https://doi.org/10.1371/journal.pone.0083444 (2013).

39. Duan, X., Zhang, K., Duanmu, H. & Yu, Y. Genome-wide identification and expression characteristics of cytokinin response factors in soybean. *J. Plant. Growth Regul.* (2023).
40. Qin, L. et al. An ERF transcription factor from Tamarix Hispida, ThCRF1, can adjust osmotic potential and reactive oxygen species scavenging capability to improve salt tolerance. *Plant. Sci.* **265**, 154–166. https://doi.org/10.1016/j.plantsci.2017.10.006 (2017).
41. Kong, L. et al. Comparative analysis of cytokinin response factors in Brassica diploids and amphidiploids and insights into the evolution of Brassica species. *BMC Genom.* **19**, 728. https://doi.org/10.1186/s12864-018-5114-y (2018).
42. Hu, R. et al. Comprehensive analysis of NAC domain transcription factor gene family in Populus trichocarpa. *BMC Plant. Biol.* **10**, 145. https://doi.org/10.1186/1471-2229-10-145 (2010).
43. Spector, D. L. The dynamics of chromosome organization and gene regulation. *Annu. Rev. Biochem.* **72**, 573–608. https://doi.org/10.1146/annurev.biochem.72.121801.161724 (2003).
44. Ren, R. et al. Widespread whole genome duplications contribute to Genome Complexity and species Diversity in Angiosperms. *Mol. Plant.* **11**, 414–428. https://doi.org/10.1016/j.molp.2018.01.002 (2018).
45. Wittkopp, P. J. & Kalay, G. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat. Rev. Genet.* **13**, 59–69. https://doi.org/10.1038/nrg3095 (2011).
46. Doherty, C. J. & Kay, S. A. Circadian control of global gene expression patterns. *Annu. Rev. Genet.* **44**, 419–444. https://doi.org/10.1146/annurev-genet-102209-163432 (2010).
47. Franco-Zorrilla, J. M. et al. DNA-binding specificities of plant transcription factors and their potential to define target genes. *Proc. Natl. Acad. Sci. U S A.* **111**, 2367–2372. https://doi.org/10.1073/pnas.1316278111 (2014).
48. Li, M. Y. et al. Genomic identification of AP2/ERF transcription factors and functional characterization of two cold resistance-related AP2/ERF genes in celery (Apium graveolens L). *Planta.* **250**, 1265–1280. https://doi.org/10.1007/s00425-019-03222-2 (2019).
49. Ji, A. J. et al. Genome-wide identification of the AP2/ERF gene family involved in active Constituent Biosynthesis in. *Plant. Genome.* **9** https://doi.org/10.3835/plantgenome2015.08.0077 (2016).
50. Ritonga, F. N. et al. AP2/ERF, an important cold stress-related transcription factor family in plants: a review. *Physiol. Mol. Biol. Plants.* **27**, 1953–1968. https://doi.org/10.1007/s12298-021-01061-8 (2021).

Acknowledgements

This research was funded by the China Postdoctoral Science Foundation (2023MD734178), the Heilongjiang Province Key R&D Program(2022ZX02B04-3), the Heilongjiang Academy of Agricultural Sciences "Agricultural Science and Technology Innovation Leapfrog Project" (CX23YQ04), the China Agriculture Research System (CARS-01).

Author contributions

Conceptualization, S.S. and L.L.; methodology, L.L.; software, G.D. and Y.R; validation, L.C., J.Z. and Y.L.; formal analysis, L.L., L.B, T.X; investigation, L.C., J.W., K.L., Y.C.; data curation, Y.M., Q.L., T.X., G.Y., W.L., X.W.; writing—original draft preparation, L.L.; writing—review and editing, L.L. and S.S; funding acquisition, L.L. and S.S. All authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by the China Postdoctoral Science Foundation (2023MD734178), the Heilongjiang Province Key R&D Program(2022ZX02B04-3), the Heilongjiang Academy of Agricultural Sciences "Agricultural Science and Technology Innovation Leapfrog Project" (CX23YQ04), and the China Agriculture Research System (CARS-01).

Declarations

Competing interests

The authors declare no competing interests.

Conflict of interest

There are no conflicts of interest to declare.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-79950-7.

Correspondence and requests for materials should be addressed to S.S.