BMC Genomics

RESEARCH Open Access

Genome-wide identification, expression profile and selection analysis of the CPK gene family in *Nelumbo nucifera*

Teng Cheng[1,4†], Mengyang Li[1†], Chufeng Zhao[2], Tao Wang[2], Xingwen Zheng[2,3], Liangbo Yang[3], Ying Diao[1*], Sai Yang[5*] and Zhongli Hu[1,2]

Abstract

Background Lotus (*Nelumbo nucifera* Gaertn.) is an ancient relic plant that has applications as an aquatic flower, herbal medicine, and vegetable. It is responsive to environmental stress. Calcium functions as a ubiquitous second messenger in various signal transduction pathways in plants. Calcium-dependent protein kinases (CPKs), which are serine/threonine-protein kinases commonly found in plants, have significant impacts on plant growth, development, and resilience to adversity. However, the genes encoding calcium-dependent protein kinases (CPKs) in lotus remain unclear.

Results In this study, the *CPK* gene family was systematically and comprehensively identified and analyzed. The 27 *CPKs* of lotus were further categorized into five subfamilies based on gene structure and phylogenetic tree analysis. Segmental duplication was found to be the primary event of *CPK* gene duplication, and all identified *CPK* genes underwent purifying selection. Comparative genomics analysis between lotus and model or non-model plants revealed that a large number of ancient *CPKs* were retained in lotus. Additionally, several distinct *CPKs* with strong elimination signals were selected from different ecotypes and cultivation types. The expression of *CPKs* was tissue-specific and regulated under abiotic stress. Therefore, it is suggested that CPK may confer potential advantages in some biological adaptations of lotus during long-term survival and artificial domestication. Overall, this research not only elucidates the relationship between *CPK* gene evolution and function among species but also lays a valuable foundation for future molecular breeding research on the function of *CPK* in lotus.

Conclusion This study represents the first comprehensive investigation of lotus *CPK* genes at a genome-wide level, revealing their uneven distribution among eight chromosomes. The *NnCPKs* were categorized into five groups, and an in-depth analysis of their structure and organization was conducted. By comparing genomes, we gained a better understanding of gene functions based on their homologs. Furthermore, the expression profiles in different tissues and responses to abiotic stresses indicated that these genes may play significant roles in lotus growth and

†Teng Cheng and Mengyang Li contributed equally to this work.

*Correspondence:
Ying Diao
yingdiao@whpu.edu.cn
Sai Yang
yangsai@hust.edu.cn

Full list of author information is available at the end of the article

development. These findings provide a valuable foundation for future functional studies of lotus *CPK* genes to explore their biological effects.

Keywords Calcium-dependent protein kinases, *Nelumbo nucifera*, Genome-wide, Synteny analysis, Abiotic stress

Introduction

Plants have developed a variety of defense mechanisms to cope with different environmental stressors [1–3]. The osmotic pressure of the cell membrane changes in response to environmental challenges experienced by plants [4]. Calcium (Ca^{2+}), a widely distributed secondary messenger in eukaryotes, plays an essential role in maintaining homeostasis and facilitating signal transduction pathways [5]. Throughout plant development and in reaction to environmental stimuli, calcium signals are detected and interpreted by calcium sensors or calcium-binding protein target proteins. As the concentration of calcium fluctuates, downstream proteins undergo phosphorylation, leading to alterations in gene expression patterns [6, 7]. Calmodulin (CaM) and CaM-like proteins (CML), calcineurin B-like proteins (CBL), and Ca^{2+}-dependent protein kinases (CPK) represent the three primary types of Ca^{2+} sensors responsible for translating chemical signals into biological cellular responses [8, 9].

CPKs, also known as CDPKs, are exclusively found in oomycetes, protists, green algae, and plants, and are absent in fungi or animals [3]. These proteins are classified as novel Ca^{2+} sensor protein kinases due to their unique structures, which consist of four distinct domains: the N-terminal domain, Ser/Thr kinase domain, autoinhibitory junction domain, and calmodulin-like domain. The N-terminal domain ranges from 40 to 180 amino acids and contains myristoylation and palmitoylation sites that are involved in subcellular localization and substrate identification [10, 11]. The Ser/Thr kinase region harbors an ATP binding site and is a highly conserved domain [12, 13]. The autoinhibitory junction domain comprises 20–31 amino acids and inhibits CPK activity by binding to the catalytic domain in a pseudosubstrate manner [14]. For Ca^{2+} binding, the calmodulin-like domain features an EF-hand motif with a helix-loop-helix structure. Due to its unique structure, CPK is the only protein capable of directly sensing, responding to, and translating changes in Ca^{2+} levels, resulting in downstream protein phosphorylation without inducing structural changes in other proteins [12, 15].

After the Quaternary Ice Age, only two species in *Nelumbo* survived, including *N. nucifera* Gaertn and *N. lutea* Pers. Some taxonomic evidence suggests that *N. lutea* should be considered a subspecies of *N. nucifera*. As a relic plant that has been cultivated for over 7000 years, *N. nucifera* (commonly known as lotus) can be classified into flower lotus, seed lotus, and rhizome lotus based on its various usage purposes [16]. Lotus is a perennial aquatic eudicot, while the majority of dicotyledons are terrestrial plants. Despite this distinction, lotus retains most of the characteristics of terrestrial plants, such as leaf emergence, anthesis, and pollination. Additionally, it exhibits key features of aquatic plants with highly developed aerenchyma tissue distributed in various organs - particularly stomata formed in the rhizome - along with degenerated mechanical tissue and vascular bundles. The other aquatic plants such as *Ceratophyllum demersum* L. and *Lemna minor* L. show the significantly simplified structure in comparison to that of lotus [17]. As a relic plant, the lotus undergoes a process of adaptation to terrestrial living conditions before returning to its aquatic habitat for survival. It not only retains both terrestrial and aquatic characteristics but also undergoes an aquatic adaptation process involving homeostatic regulation over time.

CPKs constitute a multigene family that plays a crucial role in regulating osmotic pressure and facilitating environmental adaptation processes. Previous reports have identified *CPK* genes in most plant species, with *Arabidopsis thaliana* containing 34 *CPK* genes [18], *Rice* containing 29 genes [19], *Pineapple* containing 17 genes [20], *Tomato* containing 29 genes [21], and *Grape* containing 19 genes [22]. However, the status of *CPKs* in lotus remains unclear. In recent years, the high-completeness and high-quality assembly of the lotus genome has provided a foundation for identifying and isolating lotus *CPK* genes. Additionally, whole-genome resequencing data of cultivated lotus varieties has presented an opportunity to uncover population differentiation of *CPKs*. This study aims to first identify lotus *CPK* genes at a genome-wide level, followed by comprehensive biological analysis. Furthermore, we will investigate the expression profiles of *NnCPKs* in various tissues under abiotic stress conditions and assess the expression profiles of selected *NnCPKs* in different population types. Overall, this work aims to provide insight into the evolutionary history and theoretical basis for studying the biological functions of *CPKs* in lotus survival strategies and stress responses.

Results

Identification of CPK gene members in the lotus genome

Based on the Pfam IDs of two structural domains of CPK proteins, namely the protein kinase domain (PF00069) and the EF-hand domain (PF13499), separate searches were conducted in the protein files of the lotus gene database. The results of these two searches were then intersected and duplicates removed. Subsequently, online

tools such as NCBI CDD, SMART, and Pfam were utilized to identify the protein structures contained in these candidate genes, with genes lacking conserved domains being excluded. Following this filtering process, a total of 27 *CPK* genes were identified in the lotus genome and named *NnCPK1-NnCPK27* according to their chromosomal positions (Table 1). To further understand the characteristics of lotus CPK proteins, an analysis was conducted on the physicochemical properties of all *CPK* genes. It was found that all identified *NnCPK*-encoded proteins varied in length from 218 (*NnCPK18*) to 597 (*NnCPK9*) amino acid residues, with coding sequences ranging from 657 to 1794 bp and molecular weights ranging from 25.06 to 66.2 kDa.

The isoelectric point ranged from 4.85 (*NnCPK26*) to 9.67 (*NnCPK13*) as shown in Table 1. The isoelectric points of CPK protein subgroups varied from 5 to greater than 8 [23, 24]. In this study, the CPK proteins in subgroups I and II exhibited acidity with isoelectric points less than 7, while all proteins in subfamily IV had isoelectric points greater than 8, consistent with previous studies.

The N-terminal myristoylation and palmitic sites of the CPK protein are associated with membrane localization. A previous report indicated that, following mutation of the myristoylation site of CPK18 in *Rice*, the protein was no longer specifically located on the cell membrane but rather distributed throughout the entire cell [23]. The predicted proteins for ten *CPK* genes contain both N-terminal palmitoylation and myristoylation sites. Fourteen CPKs were predicted to have only a palmitoylation site, while one (NnCPK26) had only a myristoylation site. However, neither NnCPK17 nor NnCPK27 have either a palmitoylation or myristoylation site. Taken together, these results indicate that the N-terminal myristoylation and palmitoylation sites of the CPK protein can determine its membrane localization.

Classification and exon–intron organization analysis of *NnCPK* genes

Structural divergence provides valuable insights into potential evolutionary relationships within multigene families. A distinct unrooted phylogenetic tree was constructed, and the exon/intron organizations of the corresponding sequences were compared to gain a better understanding of the structural diversity of the *NnCPK* genes. The findings revealed that 27 *NnCPK* genes could be categorized into five subfamilies. The majority of *NnCPK* genes within the same group exhibited identical

Table 1 Characteristics of CPKs in Lotus

Gene name	Gene ID	CDS	Amino acid	MW(kDa)	pI	*N*-Palm	*N*-Myristoylation	Chr.	Group	No. of EF hands
NnCPK1	OF00338	1617	538	60622.9	6.70	Y	Y	1	I	4
NnCPK2	OF09444	1596	531	59923.0	6.28	Y	N	1	I	4
NnCPK3	OF22377	1620	539	61067.2	7.20	Y	Y	1	I	4
NnCPK4	OF23826	1689	562	63732.7	9.25	Y	Y	1	IV	4
NnCPK5	OF01564	1587	528	59536.5	6.32	Y	N	2	III	4
NnCPK6	OF14765	1605	534	59784.6	5.71	Y	N	2	III	4
NnCPK7	OF21460	1593	530	59456.0	5.51	Y	Y	2	I	4
NnCPK8	OF24383	1671	556	61723.9	5.34	Y	N	2	II	4
NnCPK9	OF24399	1794	597	66227.7	5.22	Y	N	2	II	4
NnCPK10	OF26227	1587	528	59499.5	6.58	Y	N	2	III	4
NnCPK11	OF17993	1665	554	63063.6	6.52	Y	Y	3	III	4
NnCPK12	OF20158	1665	554	63211.0	8.07	Y	Y	4	III	4
NnCPK13	OF05424	1674	557	62869.7	9.67	Y	Y	5	IV	4
NnCPK14	OF05972	1521	506	57377.8	6.17	Y	Y	5	I	4
NnCPK15	OF05973	1566	521	58814.5	6.52	Y	Y	5	I	4
NnCPK16	OF06101	1683	560	62521.2	5.59	Y	N	5	II	4
NnCPK17	OF22297	1479	492	55454.9	5.33	N	N	5	II	4
NnCPK18	OF08745	657	218	25060.3	5.23	Y	N	6	I	4
NnCPK19	OF08902	1692	563	62778.8	6.42	Y	N	6	II	4
NnCPK20	OF10514	1722	573	63599.7	4.96	Y	N	6	II	4
NnCPK21	OF10528	1605	534	59097.9	4.97	Y	N	6	II	4
NnCPK22	OF00371	1566	521	58642.2	6.08	Y	Y	7	I	4
NnCPK23	OF12495	1593	530	59671.7	7.26	Y	N	7	III	4
NnCPK24	OF20406	1569	522	58858.6	5.97	Y	N	7	I	4
NnCPK25	OF21092	1470	489	54347.4	5.67	Y	N	7	IV	3
NnCPK26	OF00182	1602	533	60446.6	4.85	N	Y	Contig2	III	4
NnCPK27	OF00137	954	317	36562.0	6.78	N	N	Contig404	IV	2

exon–intron structures, as well as similar intron/exon numbers and lengths.

As illustrated in Fig. 1a, the majority of *NnCPK* genes consisted of seven or eight exons and six or seven introns. Members of *CPK* subfamily I had eight exons, with the exception of *NnCPK18*, which had six exons. Most *CPK* genes in group II contained 6–8 exons. Genes in subfamily III were composed of seven or eight exons, while *NnCPK26* specifically contained nine exons. *NnCPKs* in subgroups IV and V comprised 12 and 6 exons, respectively. These findings suggest a correlation between the evolutionary relationships and exon–intron structures of these *CPK* genes. Genes with more similar sequences tended to have the same number of exons, while those with dissimilar gene structures may have diverse functions.

Conserved motif analysis of NnCPK proteins

To gain a deeper understanding of the similarity and diversity of NnCPK proteins, we conducted an analysis of the conserved motifs using the Multiple Em for Motif Elicitation (MEME) online software. As depicted in Fig. 1b and Fig. S1, a total of ten conserved motifs were identified. It was observed that most NnCPK proteins contained eight or nine motifs, with motif 8 being present in almost all CPK family members except for NnCPK18 and NnCPK1 in group I, as well as NnCPK25 and NnCPK27 in group V. Furthermore, it was noted that proteins within the same subfamily typically exhibited similar motifs. Notably, motifs 9/7 were found to be unique to all members of Group I and Group III.

Cis-regulatory factors in the promoters of *NnCPK* genes

As the binding sites of transcription factors, cis-regulatory elements play a crucial role in transcriptional regulation. In this study, we utilized the 1.5 kb upstream sequences of all 27 *NnCPK* genes to investigate stress-related regulatory elements (Fig. 2). Our analysis revealed the presence of six hormone-responsive cis-elements, including the CGTCA motif involved in the Methyl Jasmonate (MeJA) response mechanism, the AuxRR core participating in auxin response, and the TCA element and ABRE responding to salicylic acid and abscisic acid, respectively. Additionally, P-boxes and GARE motifs related to GA-responsive elements were also identified in the upstream region. Furthermore, we identified eleven cis-regulatory elements associated with growth and stress response. These include G-Box, ATCT-motif, and ACE involved in light responsiveness; ARE for anaerobic induction; CAT-box associated with meristem expression; TC-rich repeats involved in stress responsiveness; as well as MBS1 regulating flavonoid biosynthesis. Other elements related to low-temperature responsiveness (MBS), drought resistance (TATA-Box), as well as a common cis-element factor in promoter areas (CAAT-Box) were also found upstream of *NnCPK* genes (Table S1). The clustering of these cis-elements within *NnCPKs* indicates their pivotal role in controlling gene expression during various plant growth phases and in response to external stimuli.

Evolutionary and collinearity analysis of *CPKs*

Distribution and collinearity analysis of CPKs in lotus

The *CPK* genes were mapped to their respective chromosomes based on location information within the genome. Analysis revealed that out of the 27 identified *CPK* genes, *NnCPK26* and *NnCPK27* were contiguous, while the remaining 25 *NnCPKs* were irregularly distributed across seven chromosomes. Some chromosomes showed a higher concentration of *CPK* genes, while others had a lower concentration. Figure 3 illustrates that chromosome 8 was the only one without any *CPK* genes, whereas chromosome 2 contained the highest number (six) of *NnCPK* genes. Chromosomes 3 and 4 each had only one *NnCPK* gene, while chromosomes 1, 6, and 7 harbored four *CPK* genes, and chromosome 5 contained five.

To investigate the expansion mechanism and gene duplication effect of the lotus *CPK* gene family, we conducted tandem and segmental duplication analyses. In the lotus genome, we identified eleven pairs of duplicated *CPK* genes, comprising two tandemly duplicated pairs (*NnCPK10* and *NnCPK5*, *NnCPK14* and *NnCPK15*) and nine segmentally duplicated pairs (*NnCPK11-NnCPK12*, *NnCPK21-NnCPK9*, *NnCPK19-16*, *NnCPK13-4*, *NnCPK2-NnCPK24*, *NnCPK23-6*, *NnCPK1-3*, *NnCPK20-8*, and *NnCPK14-18*) (Fig. 3). To assess the selective pressure on gene-encoded proteins, we determined the Ka/Ks ratios for these duplicated gene pairs. The results showed that all duplicated gene pairs had a Ka/Ks ratio of < 1.00 indicating purifying selection (Table S2). Furthermore, the divergence time of the duplicated gene pairs ranged from 4.2 to 37.4 million years ago(MYA). These findings suggest that both tandemly and segmentally duplicated genes underwent purifying selection and that segmental duplication played a crucial role in amplifying members of the *CPK* gene family in lotus.

Collinearity analysis of lotus and other species

To investigate the origin and evolution of lotus *CPKs*, a comparative syntenic linkage map of lotus associated with *A. thaliana*, *Rice*, *Pineapple*, and *Grape* (Fig. 4) was constructed. The analysis revealed that four *NnCPKs* showed a homologous relationship with those in *Rice*, one in *A. thaliana*, seven in *Pineapple*, and seventeen in *Grape* (Table S3). Additionally, four pairs of syntenic orthologous genes (one to one), namely *NnCPK9-OsCPK9*, *NnCPK11-OsCPK8*, *NnCPK15-OsCPK12*, and *NnCPK19-OsCPK7* were identified between *Rice*

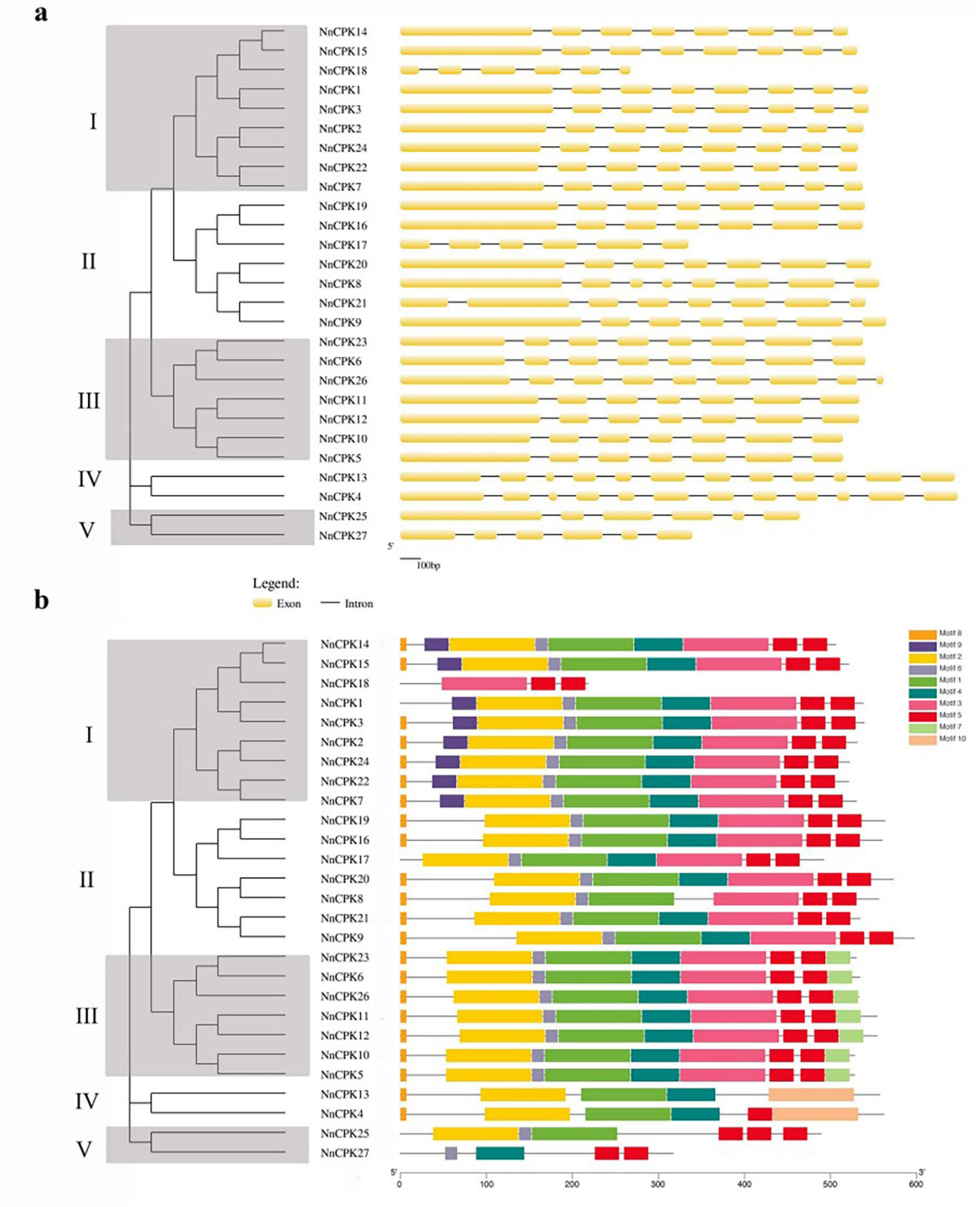

Fig. 1 Phylogenetic relationships, gene organization, and layout of conserved protein motifs in *CPK*, genes from *Nelumbo nucifera*. Different colors are used to depict cluster details. **a**. The structure of the lotus *CPK* gene exons and introns. **b**. Composition of the lotus CPK protein motif. The motifs are exhibited in distinct colored boxes, numbered 1–10. The scale at the bottom can be used to determine protein length

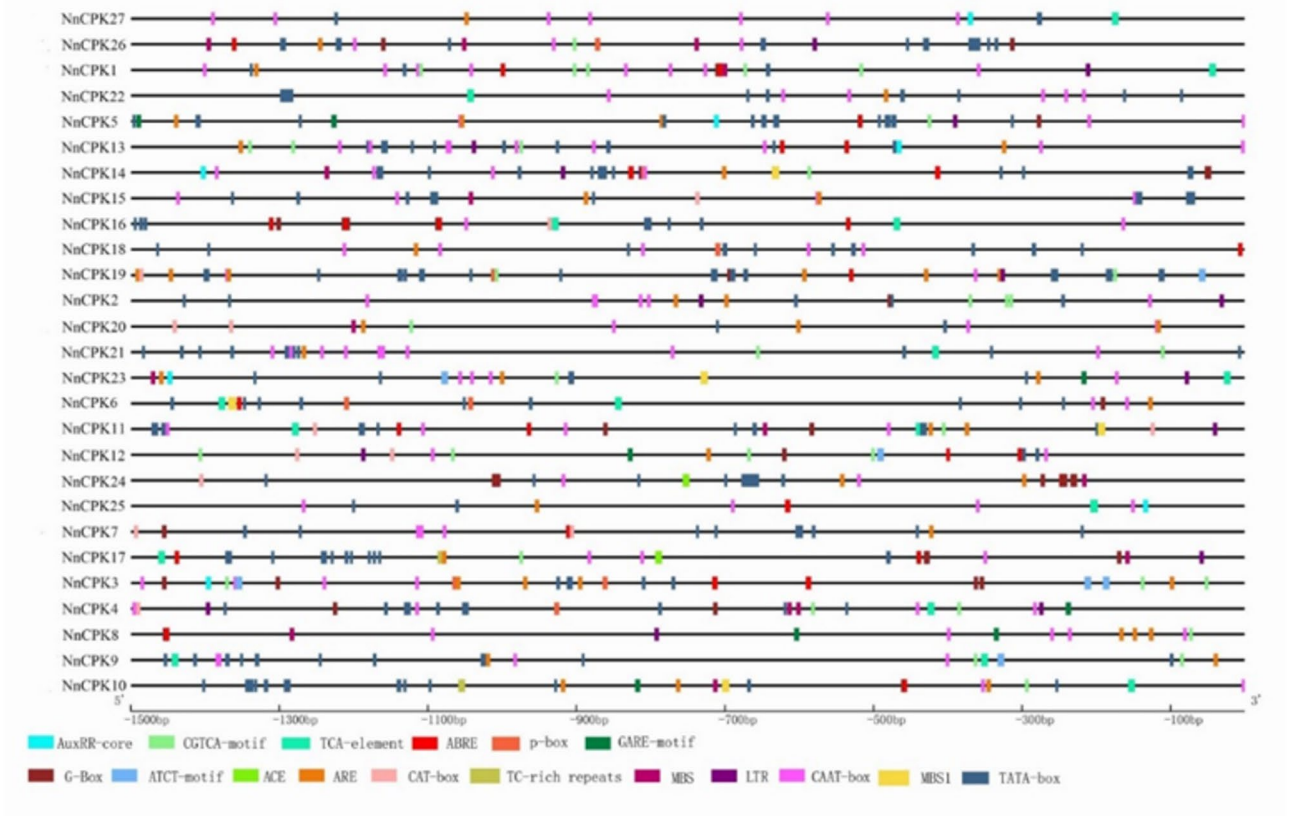

Fig. 2 Investigation of cis-acting elements in the promoter region of *CPKs* in lotus

and lotus plants. Interestingly, only one pair of genes, *NnCPK3-AtCPK23*, was found in the comparative analysis of lotus and *A. thaliana*. In the comparison map of lotus with pineapple multiple genes corresponded. There are three types of homologous gene pairs: (1) A single *CPK* lotus gene corresponding to a single *Pineapple* gene (*NnCPK16-AcoCPK10*). (2) A single lotus *CPK* gene corresponding to multiple pineapple genes including *NnCPK9-AcoCPK3/6* and *NnCPK19-AcoCPK8/10*. (3) Multiple *CPK* lotus genes corresponding to a single pineapple gene including *NnCPK1/3-AcoCPK7* and *NnCPK11/12-AcoCPK13*. Furthermore, according to the collinearity results for lotus and grape homologous *CPK* genes can be divided into two types: (1) One-to-one type including *NnCPK23-VvCDPK11*, *NnCPK1-VvCDPK12*, *NnCPK15-VvCDPK15*, and *NnCPK17-VvCDPK16*. (2) The second type included *NnCPK11/15-VvCDPK14*, *NnCPK2/24-VvCDPK1*, *NnCPK16/19-VvCDPK2*, *NnCPK7/22-VvCDPK6*, *NnCPK10/5-VvCDPK10*, *NnCPK9/20/21-VvCDPK8*, and *NnCPK9/20/21/23-VvCDPK9* had multiple lotus genes corresponding to a single grape gene. These results demonstrate that these genes might have originated from the same ancestor providing insight into the predicted functions of lotus *CPKs*.

Selection analysis of *CPK* in different population types of lotus

To investigate the role of *CPK* genes in various lotus populations, we obtained selected genes from 87 accessions through lotus resequencing data (Table S4). Six genes from the *CPK* gene family of lotus plants were chosen for analysis. Using the R language, we mapped these selected genes and identified their presence in each population type (Fig. 5). Among the genes showing strong signal elimination in flower lotus and wild lotus, *NnCPK26* was identified. Two other genes, *NnCPK4* and *NnCPK3*, were found to be present in both rhizome and wild lotus populations. Additionally, *NnCPK1* and *NnCPK6* exhibited selective scanning intensity signals in seed/wild lotus and temperate/tropical lotus populations. Lastly, *NnCPK18* was selected within the temperate lotus group of the Yangtze River-Yellow River Basin as well as the Northeast China group of lotus.

Expression patterns of *NnCPKs* in different tissues and under abiotic stress conditions

The expression patterns of all 27 *NnCPKs* were investigated using transcriptomic data from various lotus tissues, including the leaf, root, petiole, petal, seed coat, plumula, rhizome apical meristem, rhizome elongation zone, and rhizome internode (Table S5). The expression

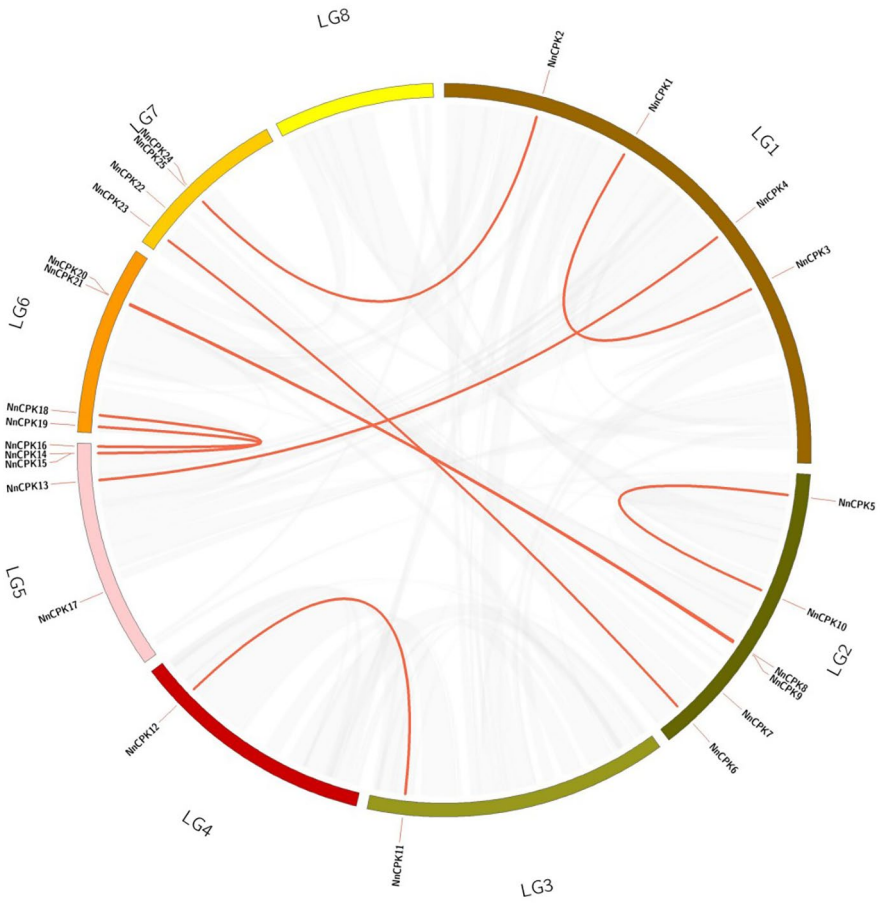

Fig. 3 A schematic diagram of the chromosomal location and interchromosomal connections of the *NnCPKs*

profiles of all 27 genes in different lotus tissues are depicted in Fig. 6 and Table S6. Certain *NnCPK* genes exhibited preferential expression across the tested tissues; for example, *NnCPK4* was expressed in the leaf and rhizome internode; *NnCPK5* was expressed in the rhizome apical meristem and elongation zone; two genes were expressed in the plumula (*NnCPK23/19*) and seed coat (*NnCPK16/6*); and three genes were expressed in the rhizome internode (*NnCPK4/16/6*). These findings suggest that *NnCPK* may have a significant impact on lotus development.

CPK genes play a crucial role in the defense mechanisms against abiotic stresses, such as salt stress and temperature stress (9,24). In order to investigate the expression patterns in response to these stresses, we conducted qRT-PCR experiments under NaCl treatment and low temperature (4 °C). The results showed that several genes were upregulated, with *NnCPK18* exhibiting a strong response to NaCl treatment (Fig. 7; Table S7). Additionally, *NnCPK21* and *NnCPK7* were found to be upregulated after 6 h of cold stress, while *NnCPK17* displayed the highest expression level at 24 h, indicating its significant role in the response to cold stress (Fig. 8; Table S8).

Discussion

Lotus CPK protein EF-hand structure

A typical CPK protein typically contains four EF-hand domains, although these domains are not fixed [18]. The number of EF-hand domains in the CPK of *A. thaliana* varies from 1 to 3, and this variation is widely observed in monocotyledons and dicotyledons [18]. *Rice*, *Zea mays*, and *Soybean* have CPK proteins with fewer than 4 EF-hand domains [25–27], while pineapple, barley, and pepper all have CPK proteins with four EF-hand domains [28, 29]. The presence of isomers in the EF-hand domain results in inconsistency in the number of EF-hand domains contained within *CPK* genes across different species. In lotus, all 25 *CPK* genes possess four EF-hand motifs in the CaM-like domain. NnCPK27 and NnCPK25 have two and three EF-hand motifs respectively; these motifs are capable of recognizing and binding calcium ions. Further research is necessary to determine whether a reduction in the number of structural domains affects

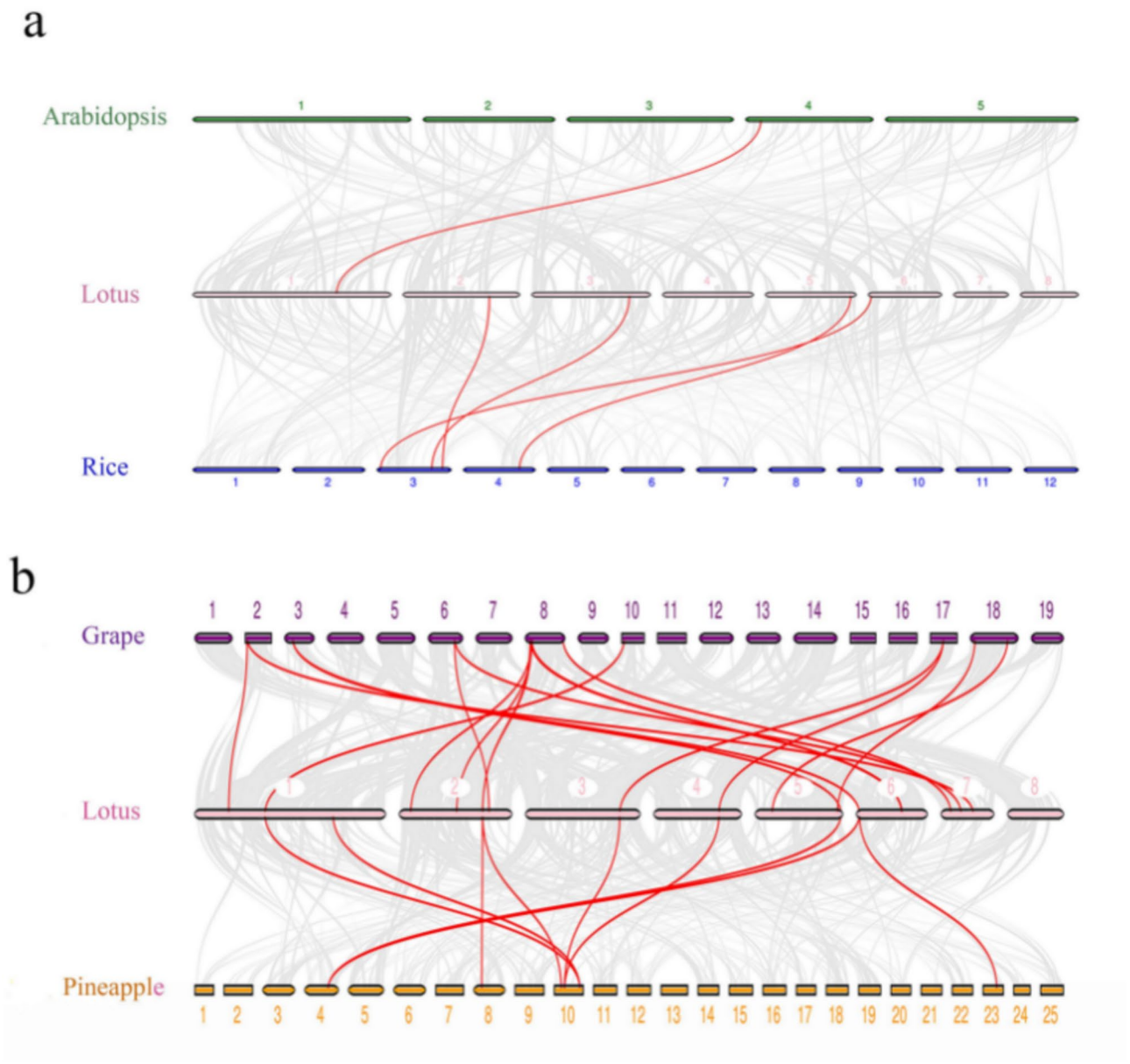

Fig. 4 Synteny analysis of *CPKs* between lotus and other plant species. **(a)** Synteny analysis of *CPKs* between lotus, *A. thaliana* and *Rice*. **(b)** Synteny analysis of *CPKs* among lotus, *Pineapple* and *Grape*. The gray lines in the background represent collinear blocks within the genomes of lotus and other plants, whereas the red lines represent syntenic *CPK* pairs

the functionality of *CPK* genes. Additionally, a study has shown that both the position of the motif formed by Aspartic acid (D) and Glutamic acid (E) within the EF-hand domain as well as the content of D/E amino acids within the motif can influence Ca^{2+} regulation by CPK [30]. Thus, systematic characterization of NnCPKs' EF-hand motifs, particularly focusing on the spatial arrangement and composition of D/E residues, would provide valuable insights into their Ca^{2+}-binding properties.

Ancient *CPK* genes in Lotus

Based on the classification results of the four *CPK* subfamilies in *A. thaliana*, *Oryza sativa*, *Pineapple*, and *Grape*, the *NnCPKs* can be categorized into five subfamilies. This classification was further supported by analyses of gene structure and conserved motifs. Previous research has shown that intron loss occurs more rapidly than intron recruitment [20]. It was observed that the number of introns in subfamily IV was highest in lotus. The *CP*K genes of subfamilies I, II, III, and V underwent intron loss during evolution. Among them, *NnCPK13* and *NnCPK4* in group IV were found to be highly conserved.

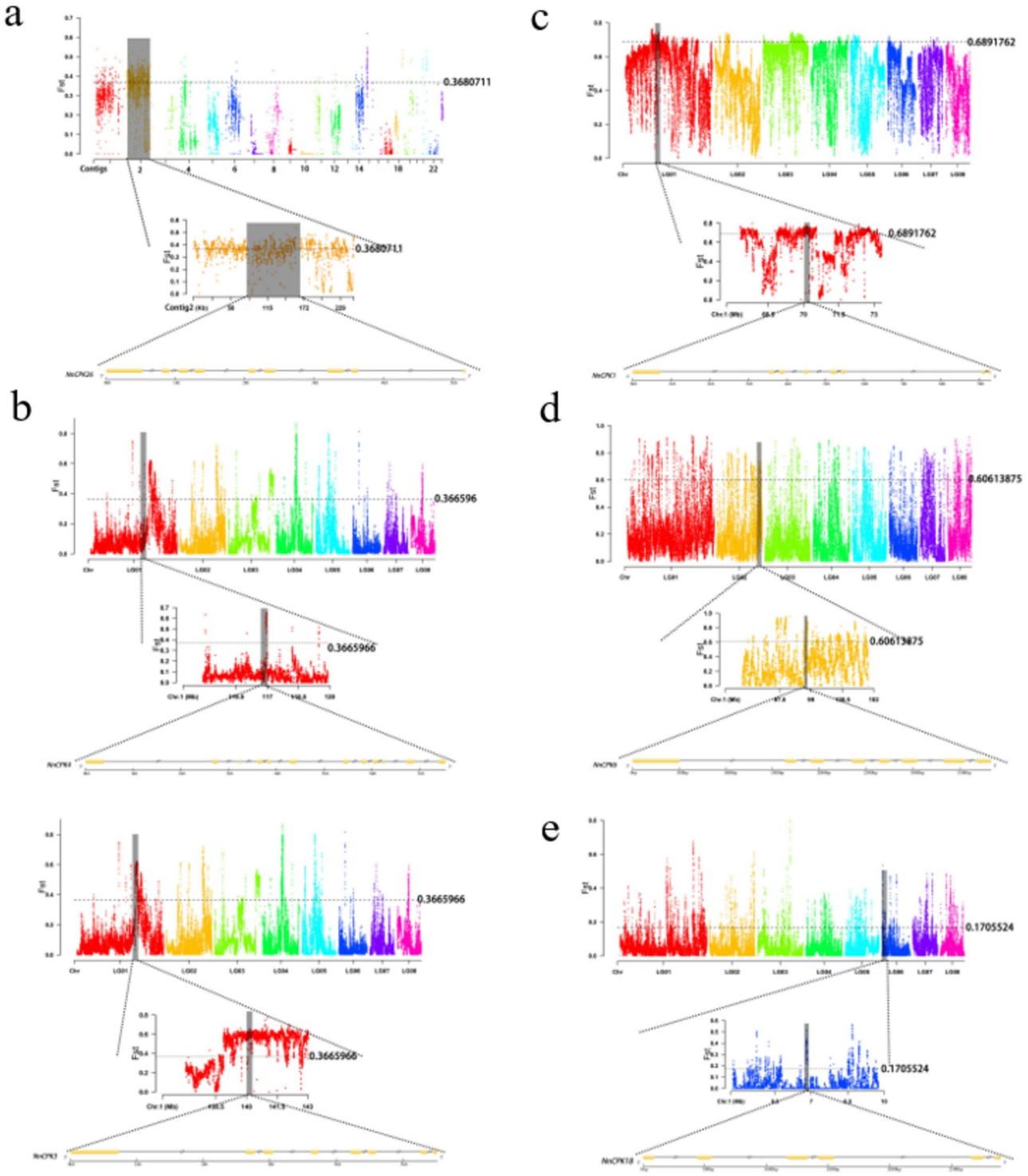

Fig. 5 Selection analysis of CPK in different population types of lotus. **(a)***NnCPK26* was selected from flower and wild lotus plants. **(b)***NnCPK3* and *NnCPK4* were selectively expressed in rhizome lotus and wild lotus. **(c)***NnCPK1* was selected from seed flowers and wild lotus plants. **(d)***NnCPK6* with signals of a selective sweep in temperate and tropical lotus. **(e)***NnCPK18* was selected from a temperate lotus plant in the Yangtze River Basin and Northeast China lotus plant

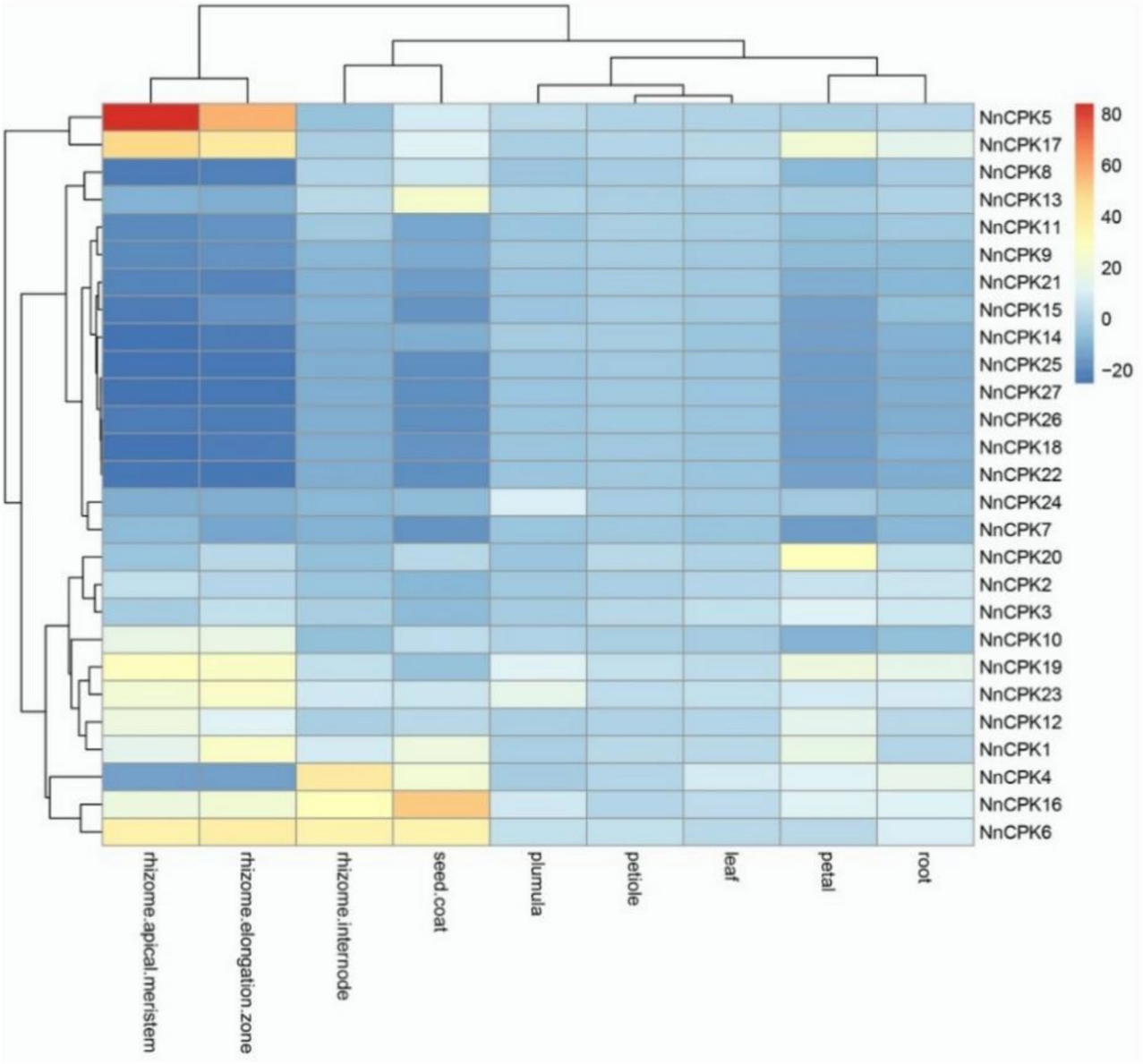

Fig. 6 Expression analysis of 27 *NnCPK* genes in different lotus tissues

Genes with a significant total intron size and/or multiple introns are likely to have greater functional significance compared to those with smaller numbers of introns [31]. Variations in the number of introns could potentially correlate with different environmental stresses [32]. Therefore, it is suggested that *NnCPK13* and *NnCPK4* may play significant roles in lotus adaptation to various environmental stresses.

Conservation of the *CPK* gene structure

Whole-genome duplication, tandem duplication, and segmental duplication play pivotal roles in genome expansion and species differentiation [33]. Our study demonstrates that there were 11 repetitive events in lotus genome, including 9 segmental duplications and 2 tandem duplications, so we indicate that segmental duplication is the primary evolutionary mechanism contributing to *NnCPKs* expansion. The Ka/Ks ratios of collinear gene pairs were counted and Ka/Ks ratios were less than 0.5, so we concluded that the gene structure of *CPK* has been highly conserved throughout evolution. Phylogenetic analysis of *CPK* genes, in conjunction with comparative analysis of similar gene structures and conserved motifs within the same subfamily, strongly supports the reliability of subfamily classification.

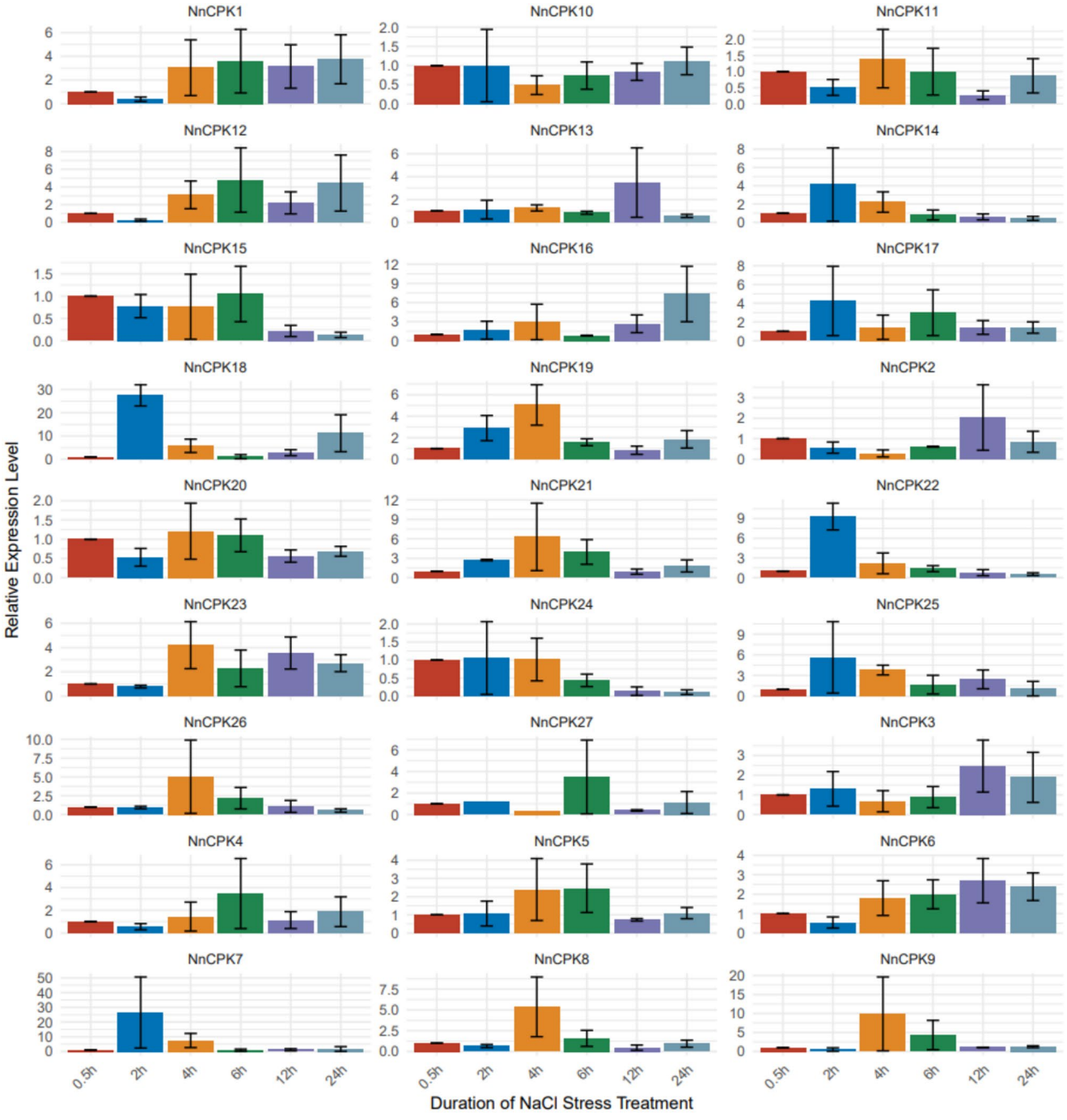

Fig. 7 Expression analysis of the *CPK* gene in lotus plants under salt stress. Error bars represent the standard deviation of gene expression data, illustrating the variability of data points relative to their mean value

Comparative analysis of different populations and species of *CPK*

Comparative genomics relies on organizing genomes into syntenic blocks with conserved characteristics across species [34]. This synteny study aims to elucidate the evolutionary and functional associations between syntenic genes of different species. Comparative genomic analysis revealed one pair of homologous genes in lotus and *A. thaliana*, *NnCPK3-AtCPK23*, and four pairs of homologous *CPK* genes in lotus and *Rice*, *NnCPK9-OsCPK9*, *NnCPK1-OsCPK8*, *NnCPK15-OsCPK12*, and *NnCPK19-OsCPK7*. The collinearity between the *CPK* genes of lotus and those of other species indicates their potential functional similarity. According to the collinearity analysis of some other dicotyledonous terrestrial plants with *Rice* and *A. thaliana*, the number of genes homologous to *A. thaliana* was significantly greater than that homologous to *Rice*. However, only one *CPK*

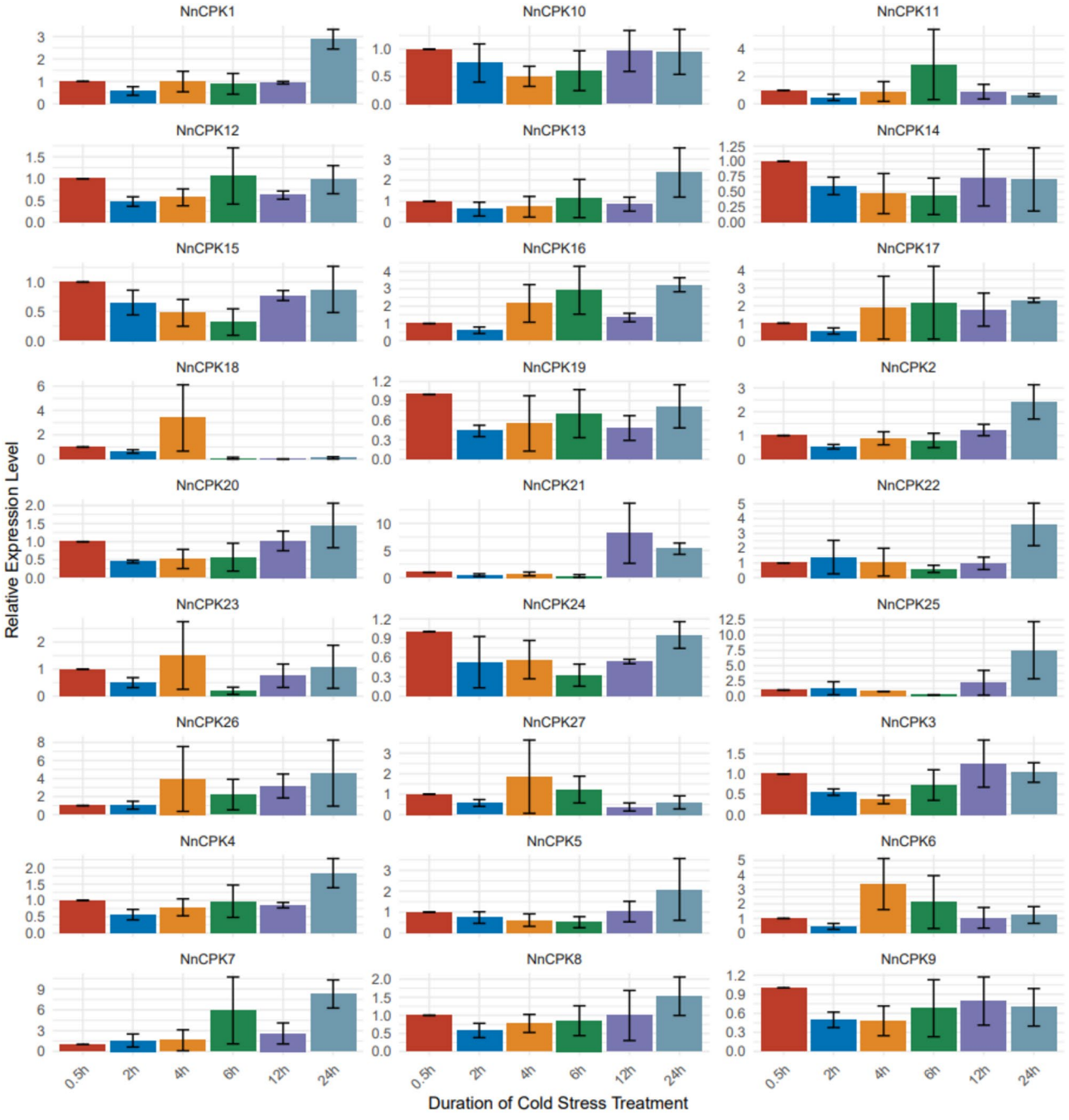

Fig. 8 Expression analysis of the *CPK* gene in lotus plants under cold stress. Error bars represent the standard deviation of gene expression data, illustrating the variability of data points relative to their mean value

gene from lotus exhibited collinearity with that from *A. thaliana*. This may be due to changes in the living environment during the process of evolution leading to loss of numerous primitive ancient *CPK* genes in *A. thaliana*. It is suggested that *A. thaliana* has evolved functional regulatory genes more suitable for growth through genome-wide duplication events as well as other genetic alterations [35].

CPK genes in *A. thaliana* and *Rice*

Some studies have investigated the mechanism of *AtCPK23*-mediated signal transduction under drought and salt stress. Mutants of *AtCPK23* exposed to abiotic drought and salt stresses exhibited enhanced tolerance and decreased stomatal aperture. It is suggested that *AtCPK21* and *AtCPK23* may have partially overlapping functions [36]. Furthermore, plants with a loss of *AtCPK21* function showed increased tolerance to

hyperosmotic stress [13]. In addition, several *CPK* genes in lotus were found to have a collinear relationship with those in *Rice*, a monocotyledonous model plant, indicating potential similar functions. *OsCPK9*, identified as a positive regulator, was shown to be involved in abiotic stress and enhances tolerance to drought stress by strengthening stomatal closure and increasing the osmotic adjustment capacity of plants [37]. Moreover, *OsCPK7*-overexpressing plants demonstrated increased resistance to salt and drought [38], while the *OsCPK12*-overexpressing line was found to be more sensitive to ABA and blast fungus [39]. The synteny analysis results revealed four pairs of genes (one-to-one) between lotus and *Rice*, suggesting potential functional similarities. Overall, these findings from synteny analysis indicate that *NnCPKs* might play a role in conferring resistance to abiotic stress and provide a foundation for further investigation into *CPK*-mediated functions under stress.

Comparative genomics analysis of ancient *CPK* genes in Lotus

To better elucidate the evolutionary relationships and functions of *NnCPKs*, we conducted a comparative analysis of the lotus genome with those of other sequenced eudicot and monocot species, including *Pineapple* and *Grape*. The lotus genome has preserved numerous ancient *CPK* genes. *Grapes*, being an ancient fruit tree domesticated in Europe for a long time, have been found as fossils in strata from the Tertiary period. The presence of a large number of syntenic blocks between grape and lotus suggests that the *CPK* gene originated before the divergence of their lineages and has been conserved throughout evolution. In contrast to *A. thaliana*, *Rice*, and *Grape* which underwent three whole-genome duplications (WGDs) (τ, σ, and ρ), *Pineapple*, distributed in tropical America, experienced only two genome doublings (τandσ) [16, 40]. Despite this difference in WGD events, *Pineapple* possesses fewer *CPK* genes but retains a significant number of ancient *CPK* genes. Our collinearity analysis also revealed multiple collinearity blocks between lotus and Pineapple. Gene combinations displayed complex functional relationships with many pairs of lotus genes associated with a single pineapple gene. However, several *CPK* genes from both species did not match any syntenic blocks due to various rounds of chromosomal restructuring and fusions as well as selective gene loss within their genomes. This has made it challenging to identify chromosomal syntenies [25, 41].

Selection analysis of the lotus *CPK* gene

In addition, based on geographical distribution and adaptability, lotus has been classified into two ecological types: tropical and temperate lotus. The temperate lotus is further divided into two populations: the Yangtze River-Yellow River Basin subgroup and the Northeast China subgroup [42]. Analysis of selected genes in different populations revealed that *NnCPK6* and *NnCPK18* were specifically chosen in tropical and temperate lotus, while *NnCPK26*, *NnCPK3*, *NnCPK4*, and *NnCPK1* were selected from flower/wild lotus, rhizome/wild lotus, and seed/wild lotus respectively. The selection signatures observed in these genes (*NnCPK6*, *NnCPK18*, *NnCPK26*, *NnCPK3*, *NnCPK4*, and *NnCPK1*) across distinct lotus populations suggest their potential involvement in ecological adaptation processes, though functional validation through comparative transcriptomics and genome-editing approaches remains essential.

Expression of Lotus *CPK* genes

Based on the expression patterns of lotus, it was observed that approximately half of the *CPK* genes demonstrated wild-type expression across various tissues. Duplicated genes exhibited varying levels of transcription in different tissues; for example, *NnCPK4* showed high expression in the rhizome internode and leaf, whereas *NnCPK13* displayed an opposite pattern of expression. The presence of cis-elements in the promoter region of *CPK* genes may account for their differential responses to stress environments and subsequent variations in expression levels. While most genes were downregulated under salt stress, *NnCPK18* showed upregulation 2 h after NaCl treatment. Additionally, *NnCPK17* exhibited upregulation with peak expression at 24 h under cold conditions. These findings suggest that lotus *CPKs* play a role in response to salt and cold stress. While our data suggest a potential role of lotus CPKs in stress responses, the specific molecular mechanisms—including signal transduction pathways, protein-protein interactions, and downstream targets—require systematic investigation through functional genomics and biochemical approaches.

Materials and methods

Identification of Lotus *CPK* gene family members

The hidden Markov model (HMM) for the protein kinase domain (PF00069) and the EF-hand domain (PF13499) was obtained from the Pfam database (https://www.ebi.ac.uk/interpro/entry/pfam/#table). The lotus genome data for analysis were downloaded from NCBI (accession number: SRR13617442). The HMM profiles of the two conserved domains were selected as the query files to perform a hemmer search of the lotus protein database with a cutoff e-value of 1.2e-28. The results of multisequence alignment by ClusterW (2.1) were compared in order to establish a species-specific HMM model for lotus, which was subsequently re-searched. All identified sequences were deduplicated, and candidate sequences with e-values less than 0.001 were chosen. All *CPK* candidate genes were validated in SMART (http://smart.e

mbl.de/), NCBI CDD (https://www.ncbi.nlm.nih.gov/cdd/), and Pfam databases to confirm their two conserved core domains. The molecular weight and theoretical isoelectric point of the CPK protein were estimated using bio-Perl and ExPASy (http://web.ExPASy.org/protparam/). Myristoylator (http://web.ExPASy.org/myristoylator) [43] and CSS-Plam program [44] were utilized to predict N-myristoylation and palmitoylation sites respectively.

Analysis of gene structure and phylogenetics

Exon-intron organization analysis and the generation of intron patterns were conducted using the Gene Structure Display Server (GSDS) bioinformatics tool, based on the default parameters (http://gsds.cbi.pku.edu.cn/). The motif-based sequence analysis tools MEME (https://meme-suite.org/meme/tools/meme) were utilized to predict conserved motifs [45]. Multiple alignments of the recognized CPK amino acid sequences were performed with ClustalW. All results are presented in the form of a phylogenetic tree, which was constructed using MEGA 7.0 software via the neighbor-joining method with a bootstrap value of 1000 replicates.

Cis-regulatory elements in the *NnCPK* gene promoter

The lotus genome sequence has provided the upstream 1.5 kb regulatory regions of *NnCPKs*. Cis-regulatory elements in the promoter regions of all *CPK* genes were identified in the PlantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) [46].

Chromosomal distribution and gene duplication

A distribution map of genes anchored on the chromosome was generated using circus software, based on location data in the lotus genome database [47]. Gene duplication events were analyzed using the Multiple Collinearity Scan toolkit software (MCScanX) [48]. The NG model in the KaKs Calculator 2.0 was utilized to estimate the values of Ka, Ks, and Ka/Ks [49]. The formula T = Ks/2r was applied to determine the value of divergence time. Here, Ks represents synonymous substitutions per site, while r denotes the rate of divergence for plant nuclear genes. The r for dicotyledonous plants was estimated to be 1.5×10^{-8}/site/year [50]. Syntenic blocks were identified and used to construct a synteny diagram connecting the lotus and *A. thaliana*, *Rice*, *Pineapple*, and *Grape* genomes.

Plant material and treatments

The harvested lotus seeds were placed in clear water to facilitate germination, with the water being changed on a daily basis. After 10–15 days of germination, buds of approximately equal height were selected and immediately transplanted into shallow water-filled plastic pots. Subsequently, the lotus plants were subjected to salt and cold stress by transferring them to a growth chamber at 4℃ containing 350 mM NaCl for a duration of 18 days. Leaves from treated plants were collected at time points of 0.5, 2, 4, 6, 12 and 24 h, rapidly submerged in liquid nitrogen and stored at -80 °C for future use.

RNA-seq expression exploration

To analyze the expression patterns of the *CPK* genes, we conducted an examination of their expression levels in various tissues. We obtained RNA-Seq data from GenBank, including plumula (SRX5056982, SRX5056981, and SRX5056980), rhizome apical meristem (SRR831190), rhizome elongation zone (SRX266486), rhizome internode (SRX268456, SRX268457, and SRX268458), seed coat (SRX4718778, SRX4718777, and SRX4718776), leaf (SRX266473, SRX266474, and SRX266475), root (SRX264995, SRX264996, and SRX265001) petiole (SRX266488, SRX266489, and SRX266490) and petal (SRX5416953, S RX3384753,and SRR8169126). The SRA formatted data was converted into FASTQ using the SRA toolkit. Subsequently, the raw reads were filtered to remove low-quality sequences with fastp (v0.20.1). An index of transcripts was constructed using Salmon (v0.9.1). Paired-end or single-end clean reads were aligned to the transcriptome data. The average value was calculated for samples with multiple SRA data points. The expression level TPM (transcripts per kilobase million) was scaled by R (version 4.0.3) to construct a heatmap.

RNA extraction and quantitative real-time PCR

A total of 30 mg of frozen plant material RNA was extracted using the RNAprep Pure Plant Plus Kit produced by Beijing Tiangen Biotechnology Co., Ltd., China. The quality of the extracted RNA was assessed through 1.2% agarose gel electrophoresis. Subsequently, the reverse transcription reaction was conducted utilizing the FastQuant RT Kit with gDNase (Tiangen Biotech Co., Ltd., Beijing, China). The relative transcript levels of genes were analyzed by real-time PCR using the StepOne Real-Time PCR System (Applied Biosystem, Foster City, USA) and ChamQ^{TM}SYBRqPCR Master Mix (High ROX Premixed) System (Vazyme Biotech Co., Ltd), with details of the primers used provided in Table S9. The qRT−PCR temperature profile consisted of an initial denaturation step at 95 °C for 30 s, followed by 40 cycles of denaturation at 95 °C for 10 s and annealing/extension at 60 °C for 30 s. Furthermore, the relative transcript levels of lotus genes were normalized to NnActin transcript levels using the $2^{-\Delta\Delta CT}$ method [51]. It is important to note that all RT−qPCR data included three technical and biological replicates. All RT−qPCR data included three technical and biological replicates.

Selection analysis of *CPK* genes in different populations

The whole-genome resequencing data of 87 lotus species (NCBI accession no. SRP095218, accession no. PRJNA343634) were compared with the reference genome of Taikonglian No. 3 lotus. Haplotype callers and genotype GVCFs from GATK (version 4.0.10.0) (https://github.com/broadinstitute/gatk) were used to screen SNPs. FST values of different populations were calculated using VCFtools (version 0.1.13), with 100,000 sliding windows and 10,000 sliding step settings. The area with selective elimination of strong signals was determined by the top 5% of the FST. Selected genes in different populations were drawn with the R language package CMplot3.6.2.

Conclusions

Overall, the first comprehensive analysis of lotus *CPK* genes at a genome-wide level has been conducted, revealing their uneven distribution among eight chromosomes. The *NnCPKs* have been categorized into five groups, and a detailed examination of their structure and organization has been carried out. By comparing genomes, we have gained a better understanding of the functions of these genes based on their homologs. Furthermore, the expression profiles in different tissues and responses to abiotic stresses indicate that these genes may play significant roles in regulating lotus growth and development. These preliminary observations could inform future functional studies exploring potential biological roles of lotus *CPK* genes.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-025-11635-5.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Supplementary Material 5

Supplementary Material 6

Supplementary Material 7

Supplementary Material 8

Supplementary Material 9

Supplementary Material 10

Acknowledgements

This work was supported by Hubei Province Funding Support for High-Quality Development of Seed Industry Project (Major Program) [2021BBA093; HBZY2023B004-1], Navy Medical University Basic Medical College "Yi Zhang" Basic Medical Talent Development and Support Program[JCYZRC-D-022] and Naval Medical University Youth Startup Fund for Basic Medical Research [2024QN003].

Author contributions

TC performed the experiments, analyzed the data and prepared the figures and tables. ML approved the final draft. CZ, TW and ML processed the data. XZ and LY assisted in completing part of the experiment. YD and ZH conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft. SY revised the manuscript and provided financial support for this research. All authors have read and agreed to the published version of the manuscript.

Funding

Hubei Province Support Fund Project for High-Quality Development of Seed Industry (Major Program) [HBZY2023B004-1]; Key R&D Program Projects of Hubei Province [2021BBA093]; Navy Medical University Basic Medical College "Yi Zhang" Basic Medical Talent Development and Support Program[JCYZRC-D-022]; Naval Medical University Youth Startup Fund for Basic Medical Research [2024QN003].

Data availability

Data is provided within the manuscript or supplementary information files.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]School of Life Science and Technology, Wuhan Polytechnic University, Wuhan 430023, P. R. China
[2]Lotus Engineering Research Center of Hubei Province, College of Life Sciences, Wuhan University, Wuhan 430072, P. R. China
[3]White Lotus Industrial Development Bureau of Guangchang County, Guangchang 344900, P. R. China
[4]Institute of Neuroscience and Key Laboratory of Molecular Neurobiology of Military of Education, Naval Medical University, Shanghai 200433, China
[5]Nursing Department, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430022, Hubei, China

Received: 27 February 2024 / Accepted: 24 April 2025
Published online: 09 May 2025

References

1. Hasanuzzaman M, Nahar K, Alam MM, Roychowdhury R, Fujita M, Physiological. Biochemical, and molecular mechanisms of heat stress tolerance in plants. Int J Mol Sci. 2013;14:9643–84. https://doi.org/10.3390/ijms14059643.
2. Lecourieux D, Ranjeva R, Pugin A. Calcium in plant Defence-Signalling pathways. New Phytol. 2006;171:249–69. https://doi.org/10.1111/j.1469-8137.2006.01777.x.
3. Valmonte GR, Arthur K, Higgins CM, MacDiarmid RM. Calcium-Dependent protein kinases in plants: evolution, expression and function. Plant Cell Physiol. 2014;55:551–69. https://doi.org/10.1093/pcp/pct200.
4. Boudsocq M, Laurière C. Osmotic signaling in plants: multiple pathways mediated by emerging kinase families. Plant Physiol. 2005;138:1185–94. https://doi.org/10.1104/pp.105.061275.
5. Edel KH, Marchadier E, Brownlee C, Kudla J, Hetherington AM. The evolution of Calcium-Based signalling in plants. Curr Biol. 2017;27:R667–79. https://doi.org/10.1016/j.cub.2017.05.020.
6. Harmon AC, Gribskov M, Harper JF. CDPKs - a kinase for every Ca2 + Signal? Trends Plant Sci. 2000;5:154–9. https://doi.org/10.1016/s1360-1385(00)01577-6.
7. Chinnusamy V, Schumaker K, Zhu J-K. Molecular genetic perspectives on Cross-Talk and specificity in abiotic stress signalling in plants. J Exp Bot. 2004;55:225–36. https://doi.org/10.1093/jxb/erh005.

8. Zhang L, Lu Y-T. Calmodulin-Binding protein kinases in plants. Trends Plant Sci. 2003;8:123–7. https://doi.org/10.1016/S1360-1385(03)00013-X.
9. Bouché N, Yellin A, Snedden WA, Fromm H. Plant-Specific Calmodulin-Binding proteins. Annu Rev Plant Biol. 2005;56:435–66. https://doi.org/10.1146/annurev.arplant.56.032604.144224.
10. Ito T, Nakata M, Fukazawa J, Ishida S, Takahashi Y. Alteration of substrate specificity: the variable N-Terminal domain of tobacco Ca(2+)-Dependent protein kinase is important for substrate recognition. Plant Cell. 2010;22:1592–604. https://doi.org/10.1105/tpc.109.073577.
11. Hamel L-P, Sheen J, Séguin A. Ancient signals: comparative genomics of green plant CDPKs. Trends Plant Sci. 2014;19:79–89. https://doi.org/10.1016/j.tplants.2013.10.009.
12. Harper JF, Harmon A, Plants. Symbiosis and parasites: A calcium signalling connection. Nat Rev Mol Cell Biol. 2005;6:555–66. https://doi.org/10.1038/nrm1679.
13. Franz S, Ehlert B, Liese A, Kurth J, Cazalé A-C, Romeis T. Calcium-Dependent protein kinase CPK21 functions in abiotic stress response in Arabidopsis Thaliana. Mol Plant. 2011;4:83–96. https://doi.org/10.1093/mp/ssq064.
14. Harmon AC, Yoo BC, McCaffery C. Pseudosubstrate Inhibition of CDPK, a protein kinase with a Calmodulin-like domain. Biochemistry. 1994;33:7278–87. https://doi.org/10.1021/bi00189a032.
15. Poovaiah BW, Du L, Wang H, Yang T. Recent advances in Calcium/Calmodulin-Mediated signaling with an emphasis on Plant-Microbe interactions. Plant Physiol. 2013;163:531–42. https://doi.org/10.1104/pp.113.220780.
16. Ming R, VanBuren R, Wai CM, Tang H, Schatz MC, Bowers JE, Lyons E, Wang M-L, Chen J, Biggers E, et al. The pineapple genome and the evolution of CAM photosynthesis. Nat Genet. 2015;47:1435–42. https://doi.org/10.1038/ng.3435.
17. Heidbüchel P, Hussner A. Fragment type and water depth determine the regeneration and colonization success of submerged aquatic macrophytes. Aquat Sci. 2019;81(6). https://doi.org/10.1007/s00027-018-0603-2.
18. Cheng S-H, Willmann MR, Chen H-C, Sheen J. Calcium signaling through protein kinases. The Arabidopsis Calcium-Dependent protein kinase gene family. Plant Physiol. 2002;129:469–85. https://doi.org/10.1104/pp.005645.
19. Ray S, Agarwal P, Arora R, Kapoor S, Tyagi AK. Expression analysis of Calcium-Dependent protein kinase gene family during reproductive development and abiotic stress conditions in rice (Oryza sativa L. Ssp. Indica). Mol Genet Genomics. 2007;278:493–505. https://doi.org/10.1007/s00438-007-0267-4.
20. Zhang M, Liu Y, He Q, Chai M, Huang Y, Chen F, Wang X, Liu Y, Cai H, Qin Y. Genome-Wide investigation of Calcium-Dependent protein kinase gene family in pineapple: evolution and expression profiles during development and stress. BMC Genomics. 2020;21. https://doi.org/10.1186/s12864-020-6501-8.
21. Hu Z, Lv X, Xia X, Zhou J, Shi K, Yu J, Zhou Y. Genome-Wide identification and expression analysis of Calcium-Dependent protein kinase in tomato. Front Plant Sci. 2016;7:469. https://doi.org/10.3389/fpls.2016.00469.
22. Zhang K, Han Y-T, Zhao F-L, Hu Y, Gao Y-R, Ma Y-F, Zheng Y, Wang Y-J, Wen Y-Q. Genome-Wide identification and expression analysis of the CDPK gene family in grape, Vitis spp. BMC Plant Biol. 2015;15:164. https://doi.org/10.1186/s12870-015-0552-z.
23. Crizel RL, Perin EC, Vighi IL, Woloski R, Seixas A, da Silva Pinto L, Rombaldi CV, Galli V. Genome-Wide identification, and characterization of the CDPK gene family reveal their involvement in abiotic stress response in Fragaria X Ananassa. Sci Rep. 2020;10:11040. https://doi.org/10.1038/s41598-020-67957-9.
24. Zhu S-Y, Yu X-C, Wang X-J, Zhao R, Li Y, Fan R-C, Shang Y, Du S-Y, Wang X-F, Wu F-Q, et al. Two Calcium-Dependent protein kinases, CPK4 and CPK11, regulate abscisic acid signal transduction in Arabidopsis. Plant Cell. 2007;19:3019–36. https://doi.org/10.1105/tpc.107.050666.
25. He D, Lei Z, Xing H, Tang B. Genome-Wide identification and analysis of the aldehyde dehydrogenase (ALDH) gene superfamily of Gossypium raimondii. Gene. 2014;549:123–33. https://doi.org/10.1016/j.gene.2014.07.054.
26. Kong X, Lv W, Jiang S, Zhang D, Cai G, Pan J, Li D. Genome-Wide identification and expression analysis of Calcium-Dependent protein kinase in maize. BMC Genomics. 2013;14:433. https://doi.org/10.1186/1471-2164-14-433.
27. Ma P, Liu J, Yang X, Ma R. Genome-Wide identification of the maize Calcium-Dependent protein kinase gene family. Appl Biochem Biotechnol. 2013;169:2111–25. https://doi.org/10.1007/s12010-013-0125-2.
28. Cai H, Cheng J, Yan Y, Xiao Z, Li J, Mou S, Qiu A, Lai Y, Guan D, He S. Genome-Wide identification and expression analysis of Calcium-Dependent protein kinase and its closely related kinase genes in Capsicum Annuum. Front Plant Sci. 2015;6:737. https://doi.org/10.3389/fpls.2015.00737.
29. Yang Y, Wang Q, Chen Q, Yin X, Qian M, Sun X, Yang Y. Genome-Wide survey indicates diverse physiological roles of the barley (Hordeum vulgare L.) Calcium-Dependent protein kinase genes. Sci Rep. 2017;7:5306. https://doi.org/10.1038/s41598-017-05646-w.
30. Mohanta TK, Mohanta N, Mohanta YK, Bae H. Genome-Wide identification of calcium dependent protein kinase gene family in plant lineage shows presence of novel D-x-D and D-E-L motifs in EF-Hand domain. Front Plant Sci. 2015;6. https://doi.org/10.3389/fpls.2015.01146.
31. Gorlova O, Fedorov A, Logothetis C, Amos C, Gorlov I. Genes with a large intronic burden show greater evolutionary conservation on the protein level. BMC Evol Biol. 2014;14. https://doi.org/10.1186/1471-2148-14-50.
32. Zuo D, Lei S, Qian F, Gu L, Wang H, Du X, Zeng T, Zhu B. Genome-Wide identification and stress response analysis of BcaCPK gene family in amphidiploid Brassica Carinata. BMC Plant Biol. 2024;24:296. https://doi.org/10.1186/s12870-024-05004-9.
33. Cannon SB, Mitra A, Baumgarten A, Young ND, May G. The roles of segmental and tandem gene duplication in the evolution of large gene families in Arabidopsis Thaliana. BMC Plant Biol. 2004;4. https://doi.org/10.1186/1471-2229-4-10.
34. Ghiurcuta CG, Moret BME. Evaluating synteny for improved comparative studies. Bioinformatics. 2014;30:i9–18. https://doi.org/10.1093/bioinformatics/btu259.
35. Veremeichik GN, Shkryl YN, Gorpenchenko TY, Silantieva SA, Avramenko TV, Brodovskaya EV, Bulgakov VP. Inactivation of the Auto-Inhibitory domain in *Arabidopsis AtCPK1* leads to increased salt, cold and heat tolerance in the *AtCPK1*-Transformed *Rubia cordifolia* L cell cultures. Plant Physiol Biochem. 2021;159:372–82. https://doi.org/10.1016/j.plaphy.2020.12.031.
36. Ma S-Y, Wu W-H. AtCPK23 functions in Arabidopsis responses to drought and salt stresses. Plant Mol Biol. 2007;65:511–8. https://doi.org/10.1007/s11103-007-9187-2.
37. Wei S, Hu W, Deng X, Zhang Y, Liu X, Zhao X, Luo Q, Jin Z, Li Y, Zhou S, et al. A rice Calcium-Dependent protein kinase OsCPK9 positively regulates drought stress tolerance and spikelet fertility. BMC Plant Biol. 2014;14. https://doi.org/10.1186/1471-2229-14-133.
38. Saijo Y, Hata S, Kyozuka J, Shimamoto K, Izui K. Over-Expression of a single Ca2+-Dependent protein kinase confers both cold and salt/drought tolerance on rice plants. Plant J. 2000;23:319–27. https://doi.org/10.1046/j.1365-313x.2000.00787.x.
39. Asano T, Hakata M, Nakamura H, Aoki N, Komatsu S, Ichikawa H, Hirochika H, Ohsugi R. Functional characterisation of OsCPK21, a Calcium-Dependent protein kinase that confers salt tolerance in rice. Plant Mol Biol. 2011;75:179–91. https://doi.org/10.1007/s11103-010-9717-1.
40. Tang H, Bowers JE, Wang X, Paterson AH. Angiosperm genome comparisons reveal early polyploidy in the monocot lineage. Proc Natl Acad Sci U S A. 2010;107:472–7. https://doi.org/10.1073/pnas.0908007107.
41. Zhang M, Liu Y, Shi H, Guo M, Chai M, He Q, Yan M, Cao D, Zhao L, Cai H, et al. Evolutionary and expression analyses of soybean basic leucine zipper transcription factor family. BMC Genomics. 2018;19:159. https://doi.org/10.1186/s12864-018-4511-6.
42. Liu Z, Zhu H, Zhou J, Jiang S, Wang Y, Kuang J, Ji Q, Peng J, Wang J, Gao L, et al. Resequencing of 296 cultivated and wild Lotus accessions unravels its evolution and breeding history. Plant J. 2020;104:1673–84. https://doi.org/10.1111/tpj.15029.
43. Bologna G, Yvon C, Duvaud S, Veuthey A-L. N-Terminal myristoylation predictions by ensembles of neural networks. Proteomics. 2004;4:1626–32. https://doi.org/10.1002/pmic.200300783.
44. Ren J, Wen L, Gao X, Jin C, Xue Y, Yao X. CSS-Palm 2.0: an updated software for palmitoylation sites prediction. Protein Eng Des Sel. 2008;21:639–44. https://doi.org/10.1093/protein/gzn039.
45. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 2009;37:W202–8. https://doi.org/10.1093/nar/gkp335.
46. Lescot M. PlantCARE, a database of plant Cis-Acting regulatory elements and a portal to tools for in Silico analysis of promoter sequences. Nucleic Acids Res. 2002;30:325–7. https://doi.org/10.1093/nar/30.1.325.
47. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. Circos: an information aesthetic for comparative genomics. Genome Res. 2009;19:1639–45. https://doi.org/10.1101/gr.092759.109.
48. Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee T, Jin H, Marler B, Guo H, et al. MCScanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40:e49. https://doi.org/10.1093/nar/gkr1293.

49. Wang D, Zhang Y, Zhang Z, Zhu J, Yu J. KaKs_Calculator 2.0: A toolkit incorporating Gamma-Series methods and sliding window strategies. Genomics Proteom Bioinf. 2010;8:77–80. https://doi.org/10.1016/S1672-0229(10)60008-3.
50. Koch MA, Haubold B, Mitchell-Olds T. Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in arabidopsis, arabis, and related genera (Brassicaceae). Mol Biol Evol. 2000;17:1483–98. https://doi.org/10.1093/oxfordjournals.molbev.a026248.
51. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using Real-Time quantitative PCR and the 2– ΔΔCT method. Methods. 2001;25:402–8. https://doi.org/10.1006/meth.2001.1262.

Publisher's note