BMC Genomics

RESEARCH

Open Access

Comprehensive analysis of the CPP gene family in Moso bamboo: insights into their role in rapid shoot growth

Jiaqi Tan[1†], Xueyun Xuan[1†], Shiying Su[1†], Yang Jiao[1], Hui Guo[1] and Zhijun Zhang[1*]

Abstract

Cysteine-rich polycomb-like proteins (CPPs), pivotal transcription factors crucial for evolution of plants from germination to maturity, and adaptation to environmental stresses, have not yet been characterized within the context of Moso bamboo. The *CPP* gene family of Moso bamboo was identified through bioinformatics, and the structural and functional attributes of the gene, including its physicochemical properties, evolutionary relationships, and gene-protein structures, were revealed. Additionally, the current study also offers valuable information on the patterns of gene expression in bamboo shoots during the period of accelerated development. The results show that the Moso bamboo genome contains 17 *CPP* members. Molecular phylogenetic relationships indicated that CPPs could be divided into three subfamilies and that *CPP* members of the same subfamily shared similar gene structures, motifs and conserved structural domains. The covariance analysis showed that the covariance between *CPP* and *Oryza sativa* was higher than that between Arabidopsis. Protein homology modeling showed that CPP proteins contain the DNA-binding domain of typical transcription factors. Transcriptomic data analysis revealed that *CPP* gene expression differs between tissues and organs. *CPP* could be regulated in response to exogenous gibberellin (GA) and naphthalene acetic acid (NAA). The qRT-PCR experiments demonstrated that *CPP* was crucial in the initial and fast expansion of bamboo shoots. Additionally, gene ontology (GO), KEGG enrichment and *CPP* regulatory network map analyses revealed multiple functional annotations of PeCPP-regulated downstream target genes. The results of this study will not only lay the foundation for further exploration of the detailed biological functions of *CPP* genes in the growth and development of Moso bamboo, but also establish the groundwork for future genetic enhancement of fast-growing forest trees.

Keywords Moso bamboo, Transcription factor, *CPP* gene family, Genome-wide identification, Expression pattern

Background

Transcriptional factors (TFs) are proteins with DNA-binding properties that regulate gene expression at the transcriptional level by interacting with specific motifs in designated promoters. Hence, transcription factors serve as crucial regulators in the processes of evolution of plants from germination to maturity, and response to abiotic stresses. In general, the DNA binding domain of each transcription factor is highly conserved. Within the PlantTFdb database, transcription factors are categorized into numerous gene families according to the sequence attributes of their DNA binding domain [1, 2]. For example, the dicotyledonous model plant *Arabidopsis thaliana* transcription factors can be divided into 62 families. This plant contains at least 1,533 genes encoding

†Jiaqi Tan, Xueyun Xuan and Shiying Su contributed equally to this work.

*Correspondence:
Zhijun Zhang
zjzhang@zafu.edu.cn
[1] Bamboo Industry Institute, Zhejiang A&F University, Lin'an, Hangzhou, Zhejiang 311300, China

transcriptional regulators, accounting for approximately 5.9% of all Arabidopsis genes [3]. The monocot model plant rice (*Oryza sativa*) transcription factors can be divided into 56 families. It contains at least 2408 genes encoding transcriptional regulators, accounting for approximately 4% to 4.8% of the total number of rice genes. Although the quantity of plant transcription factors is not dominant, they can regulate the expression levels of numerous downstream target genes. Ultimately, they have an effect on plant growth and development, as well as the plant's response to abiotic stresses. As a result, identifying and analyzing transcription factors is an important part of understanding plant gene transcriptional control mechanisms [4].

CXC structural domain-containing transcription factors, characterized by their cysteine-rich composition, are integral in regulating transcription levels and thereby play an essential part in plant tissue and organ growth, development, and response to abiotic stress. While members of the *CPP* gene family are prevalent in both plant and animal species, they have not been found in lower eukaryotes such as yeast [5]. After the discovery of the first plant CPP transcription factor, *TSO1*, in Arabidopsis in 2008, the transcription factor has been identified in 16 higher plant species [6], including rice [7], soybean [8], tomato [9], wheat [10], cucumber [11] and maize [12] through genome-wide data bioinformatics analysis. Sequence analysis of existing CPP proteins revealed that CPP proteins contain two Cysteinerich CXC conserved motifs (pfam03638), namely CXC1 and CXC2. The two CXC motifs are usually separated by a length region containing the short conserved sequence RNPXAFXPK [13, 14]. The CXC structural domain has also been named CRC, TCR and CHC [13, 15].

The documented roles of the *CPP* gene primarily pertain to its participation in plant floral organ development and the plant's response to adverse environmental conditions. For example, the Arabidopsis TOS1 (*AtCPP5*) gene not only regulates mitotic and cytoplasmic cell division during flower development, but also causes defects in flower organ development in the form of mutations [16, 17]. Specifically, soybean *CPP* exhibits heightened responsiveness to high temperature stress, thereby augmenting the stress resilience of soybean through upregulated expression of numerous *CPP* genes in the roots [8]. Additionally, *CPP1* modulates GMLBC3 gene expression in soybean by interacting with its promoter region [13]. Furthermore, previous studies have shown that the *CPP* gene family can regulate gene expression by binding DNA through its CXC domain [15]. In maize, most *ZmCPP* genes were significantly up-regulated under high and low-temperature stresses, and four *ZmCPP* genes were also up-regulated under drought stress [12].

Moso bamboo (*Phyllostachys edulis*), which belongs to the genus Gentiana in the subfamily Bamboo of the family Gramineae, is one of the most versatile herbaceous plants on the planet. Moso bamboo is one of the most rapidly growing species of bamboo, with stalk growth rates of up to 1–1.5 m/day. New bamboo can reach a height of more than 6–7 m in 2–3 months from the time the shoots break the ground [18]. Subsequently, the biomass can grow at a rate of 10–30% per year and become timber in 5–6 years. In contrast, most other woody plants have an annual biomass growth rate of only 2–5% [19] and require at least 10 years to reach maturity. Although the genetic improvement of Moso bamboo is hampered by the fact that the genetic system is not yet complete and the flowering cycle takes about 60 years, the molecular mechanism of the "explosive" rapid growth of Moso bamboo has always been a topic of great interest. Studies have shown that the rapid growth phase of shoots is closely related to cell growth processes, metabolic processes, replication and reproduction processes, stress resistance, mutual regulation of plant hormones in response to abiotic stresses. With the rapid development of molecular biology techniques and the application of genomics in bamboo plants, especially the completion of bamboo whole-genome sequencing in 2013 [20] and the update of the Moso bamboo genome data in 2018 and the generation of a large amount of transcriptome data, our understanding of the abundance, distribution and functional expansion of the bamboo gene family has been improved [21]. To date, SBP-like [22, 23], ARF [24], MADS-box [25], UBP [26], MYB [27], AP2/ERF [28], SAUR [29], AQP [30], IQD [31], HD-Zip [32], and D-J [33] gene families have been performed in Moso bamboo for genome-wide analysis [34]. However, there are no relevant literature reports on the identification, structure and function of the *CPP* gene family in Moso bamboo. Therefore, exploring CPP transcription factors involved in the proliferation, maturation, and environmental response of Moso bamboo is valuable for identifying genetic resources and the molecular breeding work of forest trees or graminaceous crops.

In this study, genome-wide identification of the *CPP* gene family in Moso bamboo was performed using a bioinformatics approach. Through systematic analysis of 17 Moso bamboo *CPP* genes and their encoded proteins, including phylogenetic relationships, expression analysis based on transcriptomic data and qRT-PCR. Combined with downstream target gene identification and expression analysis, it is proposed to reveal the role of the *CPP* gene family in the development of different tissues and organs of Moso bamboo and rapid shoot growth and development. The above results will provide a solid theoretical basis for further studies on the functions of the

CPP transcription factor gene family in Moso bamboo. Meanwhile, it will provide candidate gene resources for the genetic improvement of fast-growing forest trees.

Results

Analysis of *CPP* gene family members and physicochemical properties

The amino acid sequences of the 17 *CPP* family members were subjected to bioinformatic analysis, and the genes were assigned new names, *PeCPP01* to *PeCPP17*, depending on their chromosomal location (Table 1). The findings indicated that the *CPP* family's most substantial protein had a molecular weight of 85.24 kilodaltons (kD), whilst the smallest protein had a molecular weight of 20.20 kD. The lengths of their amino acid sequences varied from 177 to 791 residues. The isoelectric points range from 4.86 to 9.37. Out of the proteins, ten were categorized as basic due to their theoretical isoelectric points being over 7, while seven were classed as acidic with scores below 7. The aliphatic index, which quantifies protein thermostability, varied from 41.81 to 80.16, suggesting notable differences in thermal stability within this protein family. The *CPP* genes in these members have an average hydrophobic index of less than zero and are all hydrophilic proteins. The localization of CPPs in the cell was predicted based on the signal peptide of the protein sequence [35], and the results showed that all CPPs were localized in the nucleus. Only three CPPs (*PeCPP12*, *PeCPP13* and *PeCPP14*) were localized not only in the nucleus but also localized to the cell membrane, demonstrating that *CPP* acts as a transcription factor mainly in the nucleus to regulate the expression of downstream target genes, and some CPPs may also play a regulatory role in the chloroplast.

Analysis of gene family evolution

Based on the evolutionary analysis of different species, *CPP* family genes can be divided into two categories [7, 12]. The results showed that the amino acid sequences of Moso bamboo *CPP* could be divided into three subclades, namely CPP I, CPP II and CPP III (Fig. 1A). The largest number of *CPP* family members was found in CPP I, with 7. CPP II and CPP III had 4 and 6 *CPP* family members respectively. The CPPs of Arabidopsis form a distinct group, while those of rice and Moso bamboo cluster together in one family, indicating that Moso bamboo is more closely related to rice and more distantly related to Arabidopsis in the *CPP* subfamily. To further investigate the duplication events that occurred in the *PeCPP* gene family, we conducted collinearity analysis. Co-lineages within the Moso bamboo species are shown (Fig. 1B): a total of six sets of linear relationships existed between members of the Moso bamboo intraspecific family, with a total of six *PeCPP* genes exhibiting fragment duplication. A total of seven direct paralogous homologues of the 10 Moso bamboo PeCPPs were identified in the monocotyledonous rice. In contrast, only one Moso bamboo *CPP* homolog was found in the dicotyledonous

Table 1 Detailed information on17 *PeCPP*genes and their encoded proteins

Gene Name	Gene ID	A.I	GRAVY	Size (aa)	MW (kDa)	pI	I.I	Predicted Location
PeCPP01	PH02Gene12177.t1	62.82	−0.648	791	85.24	6.01	55.19	Nucleus
PeCPP02	PH02Gene43602.t1	55.86	−0.785	353	39.43	8.64	61.16	Nucleus
PeCPP03	PH02Gene44610.t2	57.17	−0.775	508	55.56	7.90	53.29	Nucleus
PeCPP04	PH02Gene08479.t1	59.23	−0.724	351	39.09	8.64	63.25	Nucleus
PeCPP05	PH02Gene46191.t1	41.81	−0.838	177	20.20	4.86	51.42	Nucleus
PeCPP06	PH02Gene08750.t2	63.09	−0.661	781	84.36	6.16	55.56	Nucleus
PeCPP07	PH02Gene01129.t1	60.77	−0.679	392	42.87	5.88	59.54	Nucleus
PeCPP08	PH02Gene01127.t1	66.26	−0.580	564	61.50	6.41	50.85	Nucleus
PeCPP09	PH02Gene09230.t1	59.62	−0.679	393	43.04	6.92	63.09	Nucleus
PeCPP10	PH02Gene09228.t1	66.85	−0.620	778	83.76	7.77	54.07	Nucleus
PeCPP11	PH02Gene49268.t1	80.16	−0.441	566	62.22	8.88	53.92	Nucleus
PeCPP12	PH02Gene42648.t1	63.86	−0.646	625	67.62	9.12	54.03	Chloroplast
PeCPP13	PH02Gene42650.t1	75.11	−0.382	593	64.29	9.12	49.87	Chloroplast
PeCPP14	PH02Gene00783.t1	75.31	−0.349	433	46.53	9.37	52.18	Chloroplast
PeCPP15	PH02Gene21071.t1	58.87	−0.801	506	55.39	8.38	62.86	Nucleus
PeCPP16	PH02Gene34347.t1	59.04	−0.781	508	55.56	8.54	64.16	Nucleus
PeCPP17	PH02Gene24371.t1	66.58	−0.624	491	54.01	6.25	54.90	Nucleus

A.I. Aliphatic index, *GRAVY* Grand average of hydropathicity score, *MW* Molecular weight, *pI* Isoelectric point, *I.I* Instability index

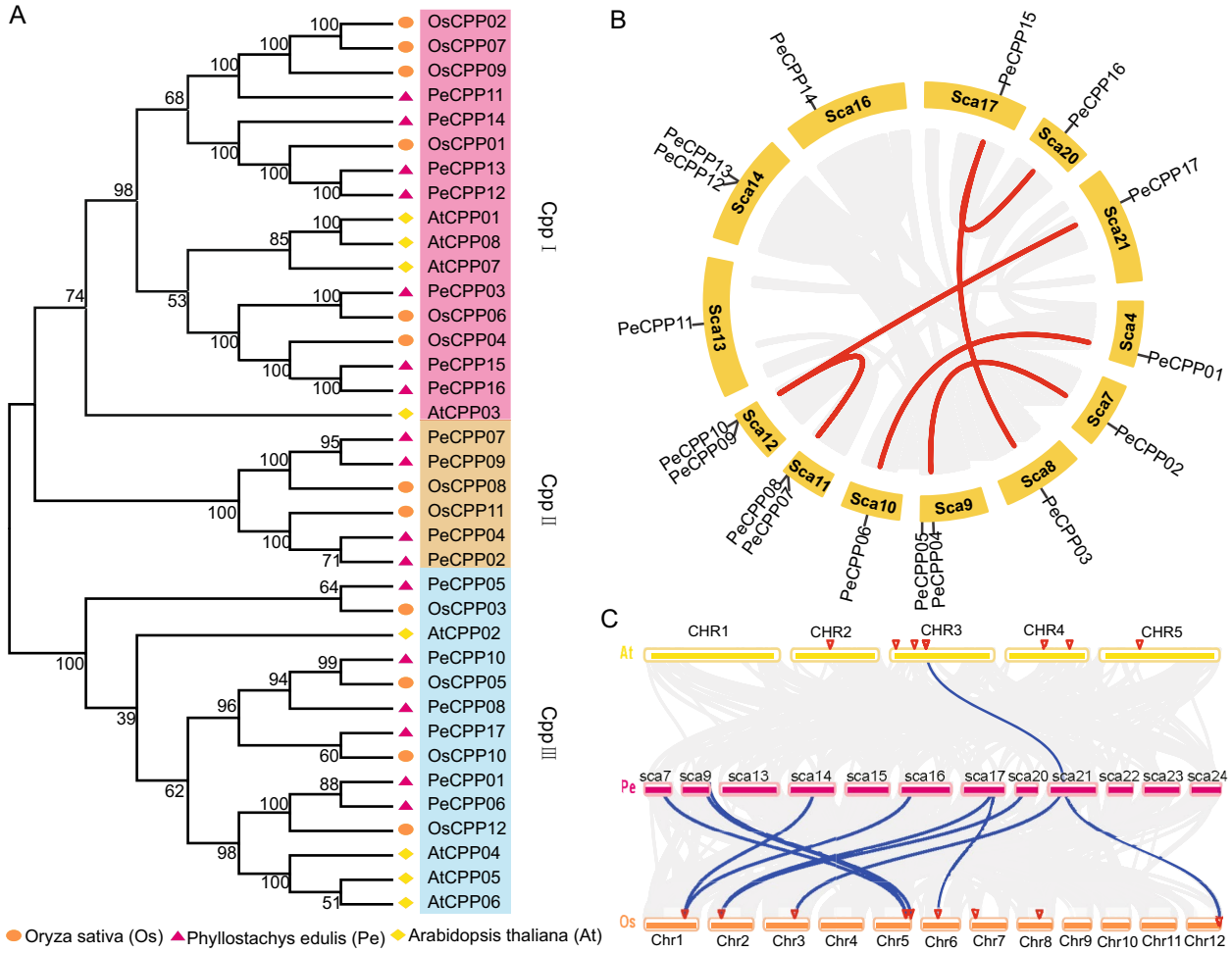

Fig. 1 **a** Phylogenetic analysis of the full-length CPP protein sequence from *Phyllostachys edulis* (Pe, bamboo), *Arabidopsis thaliana* (At, Arabidopsis), and *Oryza sativa* (Os, rice). Orange, red and yellow represent the sequences of bamboo, rice and Arabidopsis, respectively. **b** Chromosome distribution of *CPP* gene and its relationship. The gray lines indicate isomorphic blocks in the *Phyllostachys edulis* genome, and the red lines indicate repeated *PeCPP* gene pairs. **c** Simultaneous analysis of the *Phyllostachys edulis* genome of a monocotyledonous plant and a dicotyledonous plant. The gray line indicates the alignment block between the paired genomes, and the blue line indicates the integrated *CPP* gene pair

plant Arabidopsis, indicating a closer kinship between the same monocotyledonous Moso bamboo and rice (Fig. 1C).

Chromosome distribution characteristics and gene structure analysis

The chromosome scaffolds distribution of the *CPP* gene family of Moso bamboo (Fig. 2A) showed that 17 *PeCPP* genes were distributed on 13 chromosome scaffolds. Different chromosome scaffolds had different gene distribution densities. No gene underwent gene doubling (parallel duplication). Two *PeCPP* genes are present on each of scaffold11, scaffold12 and scaffold14, while all other scaffolds contain only one gene. The arrangement of exons and introns offers valuable information on the evolutionary connections across gene families [36]. The gene structure of the *CPP* family of Moso bamboo (Fig. 2B) shows that the number of introns of each Moso bamboo *CPP* gene varies from 5 to 10. *PeCPP13* and *PeCPP17* lack untranslated regions (UTRs) at both the 5' and 3' ends, while *PeCPP16* lacks a UTR at the 3' end. The 17 sequences were classified into 3 classes and the results were generally consistent with the classification of the evolutionary tree. The results showed different levels of intron variation, with the gene *PeCPP17* in CPP III having the longest intron area and the gene in CPP II having the smallest intron. The intron lengths of genes in CPP I exhibit a higher degree of concentration. Differences in the quantity of introns can lead to a wide

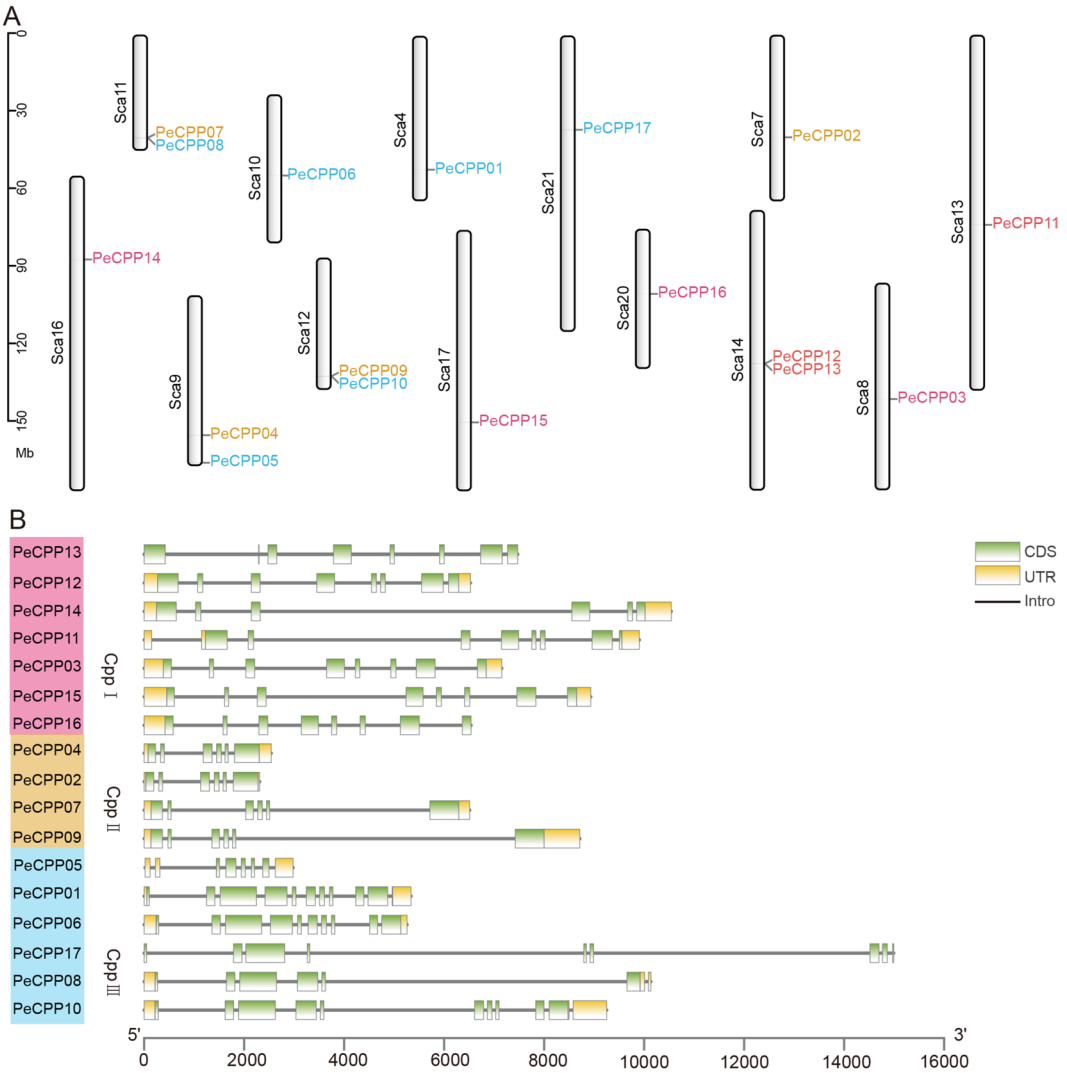

Fig. 2 Chromosomal location and gene structure of Moso bamboo *CPP* genes. **a** Chromosomal location of 17 *PeCPP* transcripts on the 13 Moso bamboo scaffold. The scale represents Mb. The chromosome numbers are indicated at the top of each bar.Pink, orange and blue indicate CPP I, CPP II and CPP III respectively. **b** Exon/intron organization of Moso bamboo *CPP* genes

range of gene structures, and the length of introns can also affect the specific functions in plants [7, 37].

Analysis of motifs and conserved structural domains

Further analysis of the conserved structural domains of the proteins encoded by *CPP* family members based on the NCBI online software CDD revealed that all *CPP* family members encoded proteins with CXC structural domains (Fig. 3A). Except for PeCPP13, PeCPP17 and PeCPP08, which have only one CXC domain, most of the family members have both CXC domains, namely CXC1 and CXC2. The CXC domain of the *CPP* gene family is highly conserved, but the distribution and number of the domains are somewhat divergent, which may lead to functional differences among the members. Members of the Moso bamboo *CPP* gene family contain two to five highly conserved motifs in number (Fig. 3B). Some motifs occur only in specific CPP proteins, suggesting that these motifs have specific functions in these genes. For example, motif3 and motif4 occur in the CPP I subfamily, and the presence of certain motifs at specific positions indicates gene-specificity. Also, some members of the *CPP* gene family suffer from motif loss. For instance, most genes in the CPP II and CPP III subfamilies contain only three motifs, while *PeCPP05*, *PeCPP08* and *PeCPP17* contain only two motifs, which may lead to functional

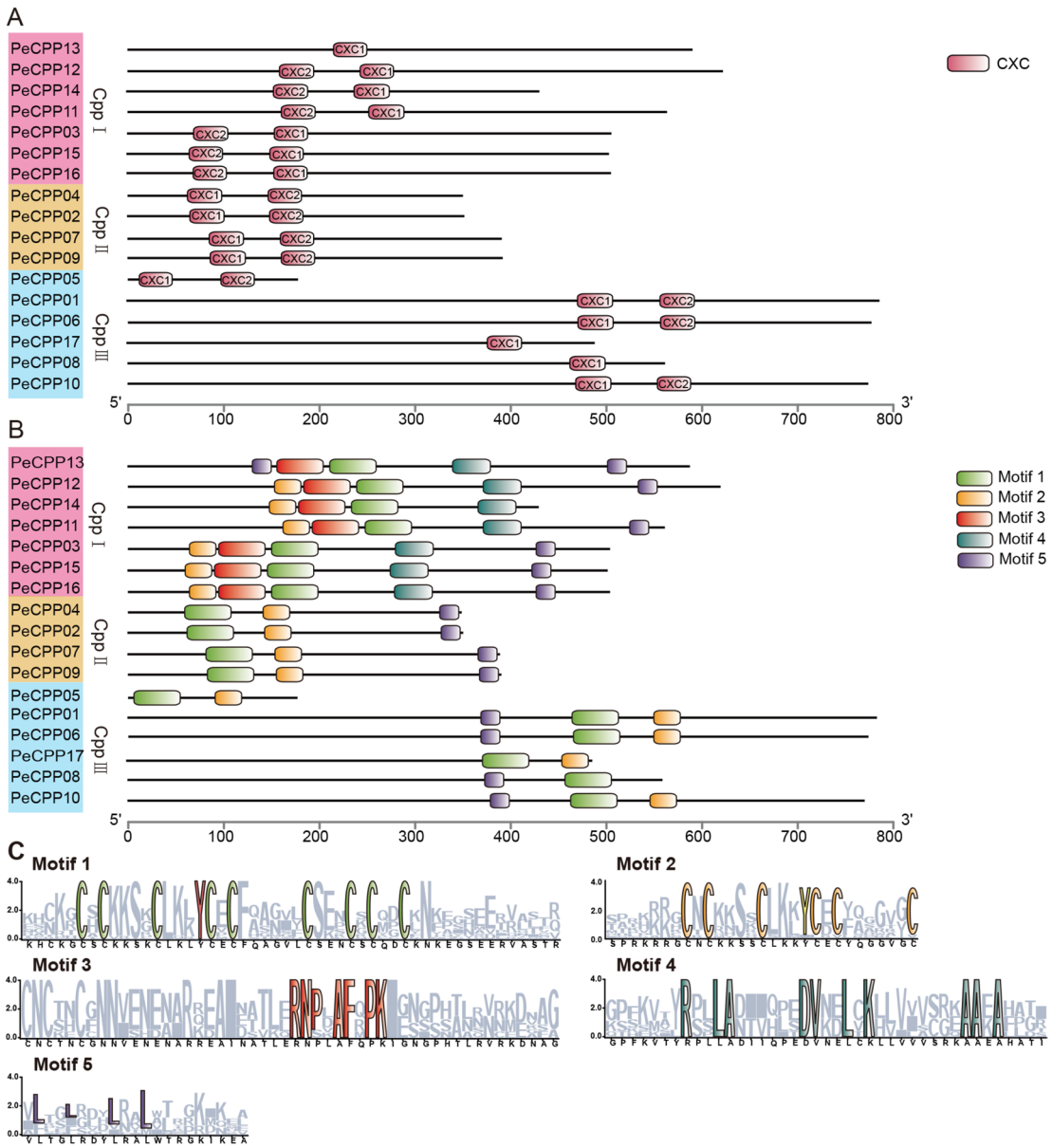

Fig. 3 Conserved domain and motif analysis. **a** Conserved domain prediction of 17 PeCPP proteins. The gray lines indicate the length of each protein sequence, and the conserved domains are indicated by colored boxes. **b** Schematic representation of the conserved motifs in Moso bamboo CPP proteins elucidated from publicly available data. Each colored box represents a motif in the protein, with the motif name indicated in the box along the bottom. **c** Sequence logos for each conserved motif are shown above

differences between genes. It was found that *PeCPP17* contains only one CXC domain, but its conserved motif contains both CXC1 and CXC2. Therefore, it is speculated that *PeCPP17* may have a gene mutation. In general, the Moso bamboo *CPP* gene family exhibits a moderate level of conservation in the CXC structural domain and motif. Nevertheless, there are variances in the composition, distribution, and quantity of the members, which could result in functional variations (Fig. 3C).

Homology modeling of *CPP* gene family proteins at three levels of structure

The highest homology, 5fd3 [38], was selected as the template for tertiary structural homology modeling of *CPP*

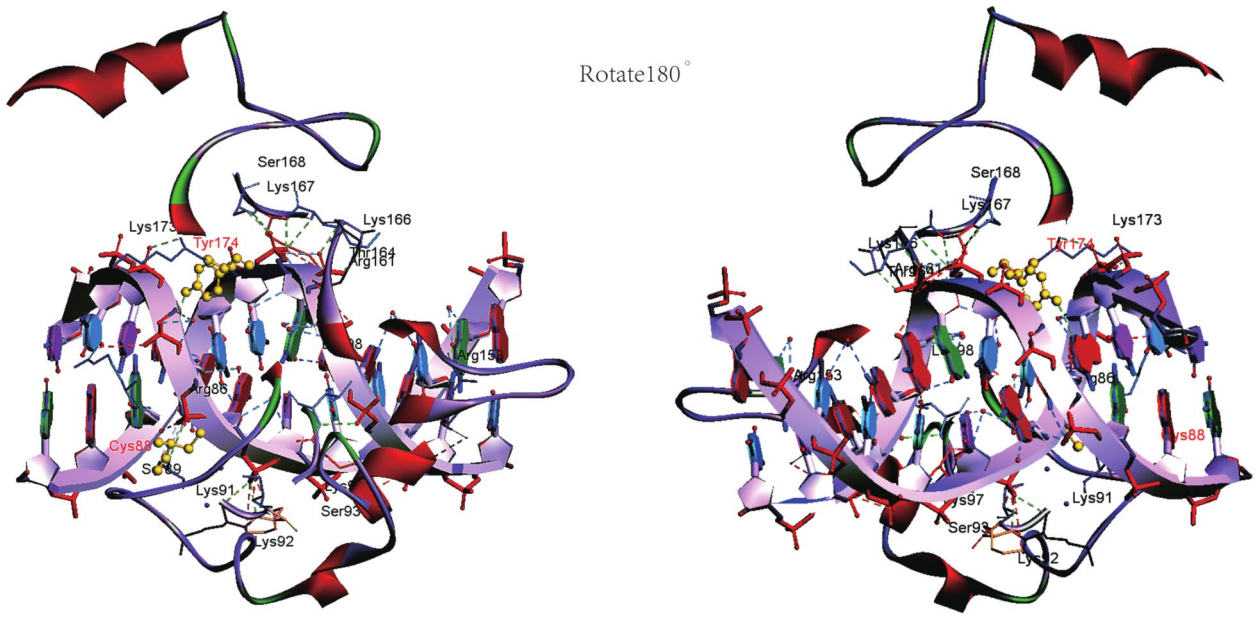

Fig. 4 The three-dimensional structure of *PeCPP07*. Helices and β-strands are colored red and green, respectively. The 4Fe-4S cluster and the side chains of key residues Cys88 and Tyr174 are shown in ball and stick representation

in Moso bamboo using the Swiss-model online server. The best conformational model was selected after SAVES measurement. The modeling results showed (Fig. 4) that the CXC sequence, a conserved Cys-rich structural domain, was located at the N terminus. There were 16 amino acid residues in the *CPP* protein, namely Lys166, 167, 173, 97, 92, 91, Ser168, 93, 89, Arg161, 153, 86, Leu98, Thr164, Tyr174 and Cys88, which could be bind to nucleic acids through non-bonding interactions such as hydrogen bonds and van der Waals forces. Among them, Cys88 is derived from the CXC structural domain. The tertiary structure of this *CPP* protein has a highly conserved DNA binding domain (DBD) and Tyr17 confers binding specificity to the *CPP* protein by interacting with DNA, a novel finding in this study that has not been found in existing studies of *CPP* genes.

Promoter characterization of the *CPP* family in Moso bamboo

Numerous members of the *CPP* family are known to be regulated by diverse abiotic stresses, underscoring their involvement in the response mechanisms to such conditions. In this study, we analyzed the nucleotide sequences located 1500 bp upstream of 17 Moso bamboo *CPP* genes to identify cis-acting regulatory elements. Promoter analysis indicated the presence of not only core promoter elements but also a multitude of other cis-acting elements involved in plant growth and development, stress response, and hormone signaling (Fig. 5A). Notably, elements related to stress response were particularly prevalent, constituting 83% of the total identified elements, thereby underlining the critical function of the *CPP* gene family in Moso bamboo's adaptation to environmental stressors (Fig. 5B-C). Focusing on growth and development, each gene within the *CPP* family exhibited CAT-box (30%) [39], which is predominantly linked with gene expression in meristematic tissues. Hormone-responsive elements detected include those responsive to methyl jasmonate, such as the CGTCA-motif (15%) and TGACG-motif (15%) [40], alongside elements like the ABRE (22%) which respond to abscisic acid, and a substantial number of elements like MYC (40%) responsive to jasmonic acid [41]. Moreover, elements such as the TATA-box and CAAT-box, ubiquitous in eukaryotic promoters, were frequently observed in the *PeCPP* promoters, emphasizing their essential role in transcriptional regulation and the necessity for the *CPP* gene family in Moso bamboo to swiftly initiate response mechanisms upon encountering environmental stress. In the domain of stress response, MYB elements, predominantly involved in mediating responses to drought, high salinity, and low temperatures, represented a significant fraction. Furthermore, the study identified an array of other stress-responsive cis-acting elements, including those associated with antioxidant responses (ARE), light-responsive elements (G-Box), and drought-responsive elements (MBS) (Fig. 5E-F). These results highlight the integral role of the *CPP* gene family in both the growth and environmental adaptability of Moso bamboo, showcasing the complexity of its regulatory mechanisms.

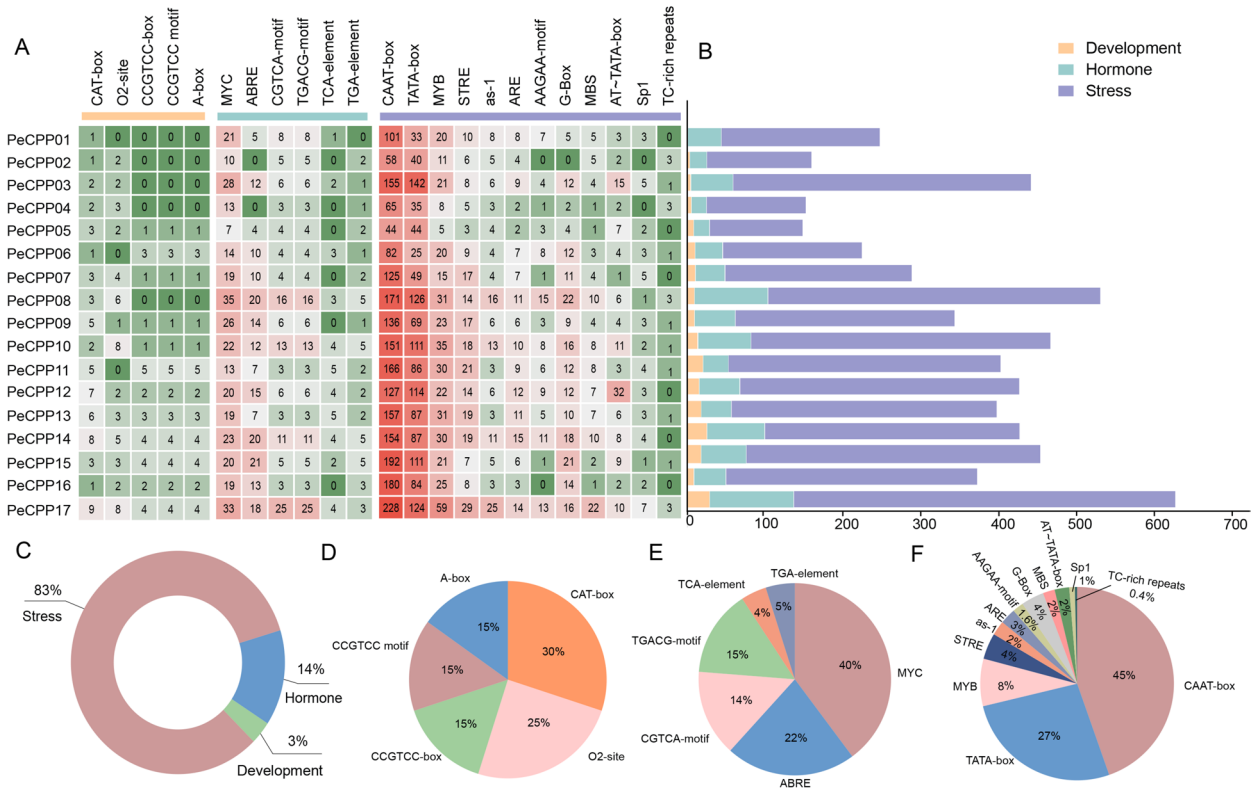

Fig. 5 Cis-elements in the *PeCPP* promoters of Moso bamboo. **a** The intensity of the red color and the numbers in the cells indicate the numbers of different cis-elements in each *PeCPP*. **b** The colored histograms indicate the number of different cis-elements in three categories. **c** Summary of all cis-elements. **d**-**f** The proportions of different cis-elements in each category: Development, Hormone, Abiotic and Biotic Stress Responsive

CPP family gene expression patterns

For investigating the expression profile characteristics of the Moso bamboo *CPP* gene family, the expression of each of the 17 *PeCPP* genes in the above samples was calculated for different growth heights of Moso bamboo shoots (0.2 m, 0.5 m, 1 m, 2 m, 3 m, 4 m, 5 m, 6 m), Moso bamboo seedlings under different hormone (GA, NAA) treatments, as well as transcriptome data from roots, leaves, flowers and whips. The expression heat map revealed that the expression of all Moso bamboo *CPP* gene family members was down-regulated after the application of GA treatment to live Moso bamboo seedlings. This indicates that all *CPP* genes may be negatively regulated by GA (Fig. 6A). After NAA treatment of live Moso bamboo seedlings as shown in (Fig. 6B): expression was up-regulated after *PeCPP06*, *PeCPP10*, *PeCPP05*, *PeCPP03* and *PeCPP16*. In contrast, the expression of *PeCPP01*, *PeCPP09* and *PeCPP04* was down-regulated after treatment with NAA, indicating that *CPP* in Moso bamboo is regulated by the NAA hormone. In addition, *CPP* was differentially expressed in different organs of the plant, as shown in (Fig. 6C): Most *CPP* was highly expressed in bamboo leaves, except for *PeCPP05*, which was expressed at a low level in leaves. *PeCPP05*, *PeCPP02*, *PeCPP09* and *PeCPP11* were highly expressed in bamboo flowers but less so in roots. *PeCPP04*, *PeCPP05*, *PeCPP08* and *PeCPP09* were less expressed in bamboo whips, while other genes were expressed at relatively high levels. This difference in expression levels suggests that the Moso bamboo *CPP* genes may be involved in the growth and development of different tissues and organs of Moso bamboo. In addition, the expression of *CPP* in different bamboo shoots at different stages of germination was analyzed separately (Fig. 6D) and validated by quantitative PCR (Fig. 7), which revealed that most *PeCPP* was at a high level in 0.2 m and 0.5 m shoots, at an intermediate level in 1 m shoots, and an increasing number of *PeCPP* showed low levels of expression after 1 m. Among them, *PeCPP03*, *PeCPP11* and *PeCPP09* showed the highest expression at 0.2 m, 2 m and 0.5 m, respectively, while *PeCPP07* and *PeCPP04* showed the lowest expression at 2 m and 3 m, respectively. As can be seen by constructing expression histograms, basically all *PeCPP* genes were highly expressed at the shoot germination stage. Overall, these PeCPPs were highly expressed at the 0.2 m (below ground) to 0.5 m (above ground) stages of shoot

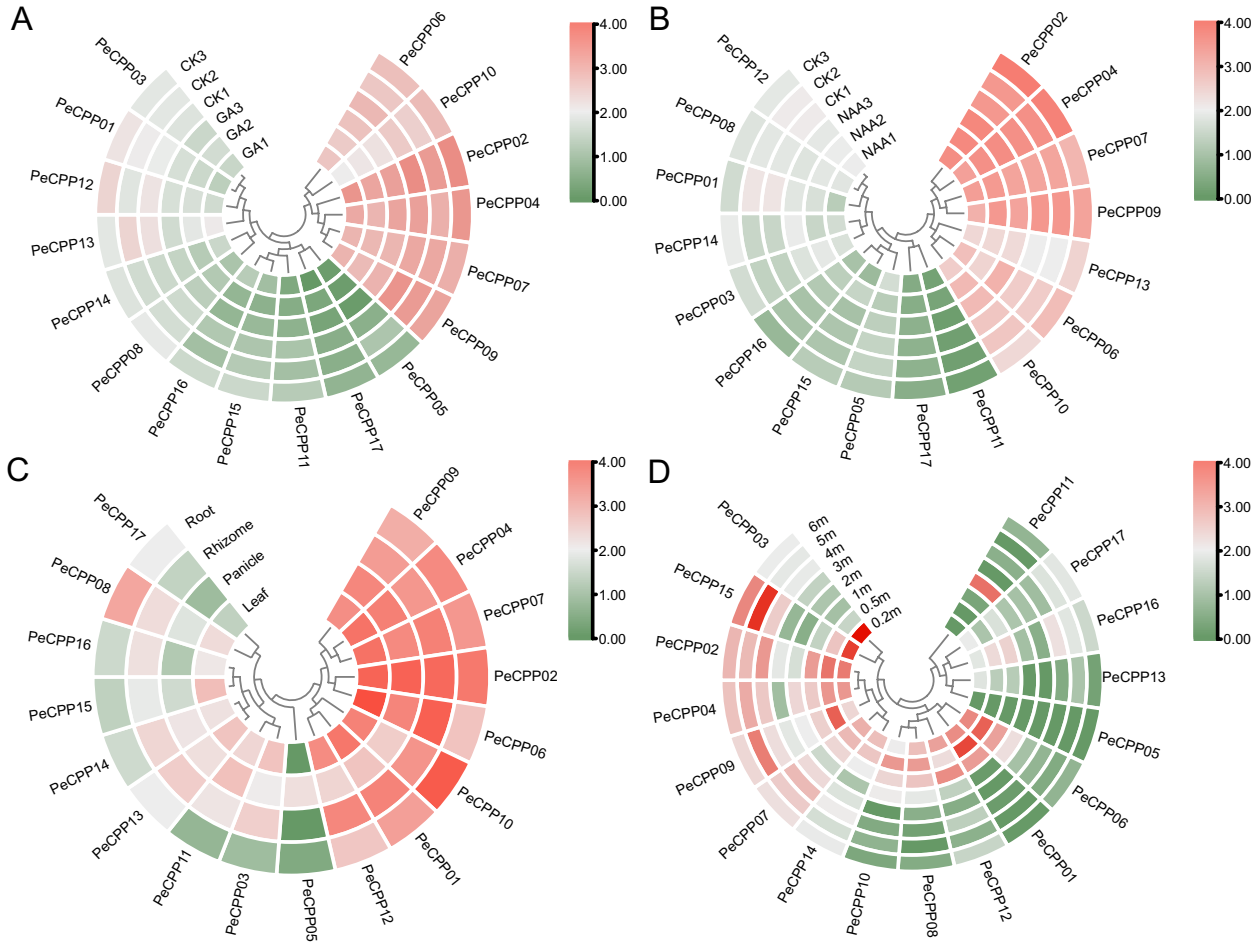

Fig. 6 Heat maps of PeCPP gene expression (log2 (TPM + 1) expression values) in multiple tissues, at different Moso bamboo heights and in response to hormone treatments. **a** GA treatment. **b** NAA treatment. **c** Roots, rhizomes, panicles, and leaves. **d** Young bamboo shoots of different heights. CK, control group. Each tissue or treatment was replicated three times. Relative expression levels are indicated by a color scale, with a change from green to red indicating low to high expression

germination, indicating that they may serve a significant role in the shoot emergence process.

Functional annotation of *CPP* target gene expression

Characteristic sequence models for *CPP* identification were downloaded through the JASPAR database. The 1-kb upstream sequences of all Moso bamboo genes were also scanned using the FIMO3 tool with a default *p*-value threshold of $1e^{-6}$, and the number of CPP-regulated target genes identified was 330. In order to conduct a more in-depth analysis of the biological roles of the target genes of Moso bamboo *CPP*, we performed GO enrichment analysis and KEGG functional annotation of 17 PeCPP protein-regulated downstream target genes. The GO enrichment analysis graph (Fig. 8A) shows that at the biological process level, *CPP* target genes are present in metabolic processes (GO:0008152) and cellular processes (GO: 0009987) and other GO enrichment terms. At the cellular component level, most of the *CPP* target genes were mainly involved in constituting cellular components (GO:0044464) and organelle components (GO:0043226), while a few were also enriched in protein-containing complexes (GO:0032991). The downstream target genes regulated by the *CPP* gene family in Moso bamboo are mainly enriched in binding (GO:0005488) and catalytic activity (GO:0003824) functions at the molecular functional level. Its binding role is a characteristic feature commonly found in transcription factors. GO enrichment analysis showed that the downstream target genes regulated by *PeCPP* transcription factors are involved in macromolecular biosynthesis, cellular biosynthetic processes, RNA biosynthesis, and the largest number of genes are associated with the regulation of cellular macromolecular biosynthetic processes (Table S1).

The KEGG Pathway classification statistics bar chart (Fig. 8B**,** Table S2) shows that the KEGG primary

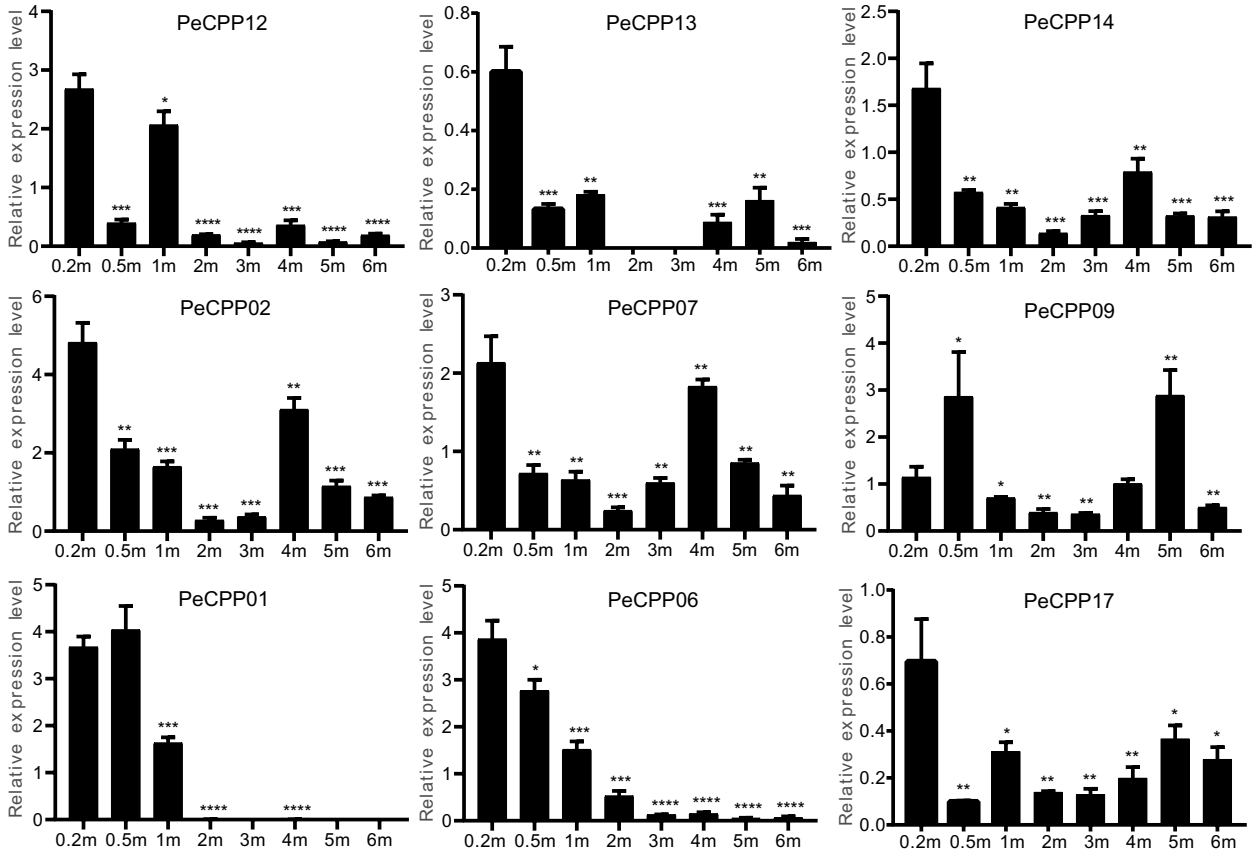

Fig. 7 RT–qPCR analysis of *PeCPP* gene in shoots at different growth heights. All experiments were performed independently at least three times, and the data are expressed as the mean ± standard deviation (SD). Asterisks indicate significant differences in transcript levels compared with those of 0.2 m shoots

metabolic pathways of *CPP* target genes are enriched in five major categories. The total number of target genes of metabolic pathways within the five major pathways was the highest, with the largest percentage of target genes involved in the secondary metabolic pathway of carbohydrates. Genetic information processing metabolic pathways contain four categories: translation, transcription, replication and repair, folding, sorting and degradation, of which the translation process involves the most *CPP* target genes, indicating that this *CPP* gene family can regulate the translation process of ribosomes. In addition *CPP* target genes are involved in metabolic pathways such as transport and catabolism.

Time-series expression analysis of target genes

To explore the regulatory effects of *CPP* genes on downstream target genes, we performed a time-series analysis of *CPP* target genes. The trend analysis graph showed (Fig. 9) that a total of 10 representative trends were summarized. Among them, the significant enrichment trend (profile 0) was negatively correlated with the change in shoot stalk height, indicating that the target genes in this profile were gradually down-regulated during shoot growth. The significant enrichment trend (profile 9) correlates positively with the change in shoot stalk height, demonstrating that the target genes in this profile are progressively up-regulated during shoot growth.

Construction of the CPP target gene regulatory networks

To comprehensively investigate the role of CPP transcription factors in the rapid development of Moso bamboo shoots, we devised a regulatory network map centered on *CPP* genes, based on temporal expression data (Fig. 10). This analysis revealed that *PeCPP01*, *PeCPP02*, *PeCPP06*, *PeCPP08*, *PeCPP10*, and *PeCPP14* exert substantial positive and negative influences on downstream target genes. Specifically, *PeCPP01*, *PeCPP06*, and *PeCPP08* were observed to suppress the expression of the *NmrA-like portein* gene family (*PH02Gene40393*), which is integral to the regulation of plant metabolism and responses to environmental stresses [42]. Furthermore, *PeCPP06* and *PeCPP08* acted as negative regulators of the Linker

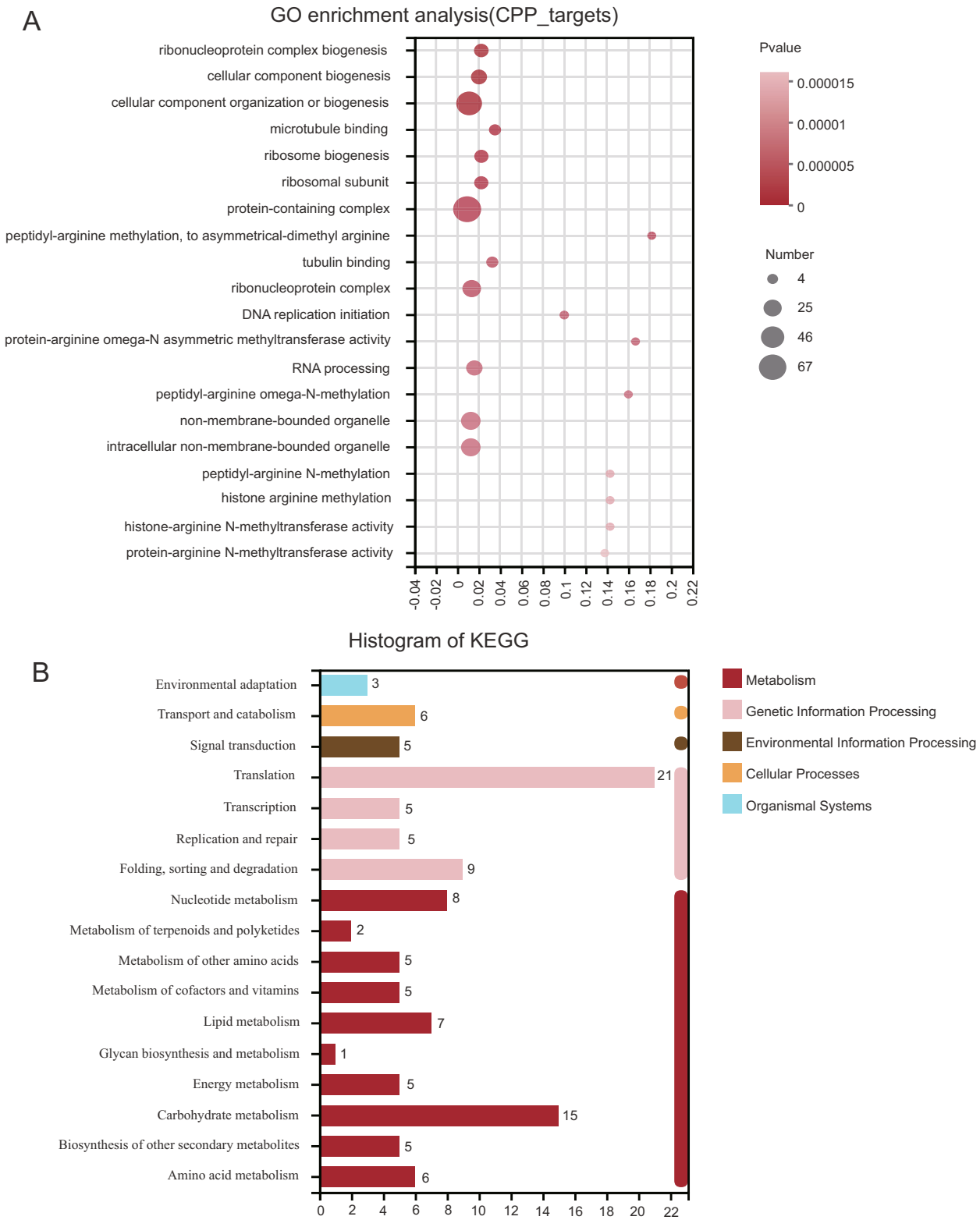

Fig. 8 GO and KEGG pathways of downstream target genes. **a** GO enrichment analysis of the *PeCPP* proteins relative to the GO database. The horizontal axis indicates the enrichment factor, and the size of the circle indicates the number of genes annotated with a given GO term. **b** Statistical map of the KEGG metabolic pathway classification of CPP-Targets. The vertical coordinate is the name of the KEGG metabolic pathway; the horizontal coordinate is the number of genes or transcripts annotated to that pathway

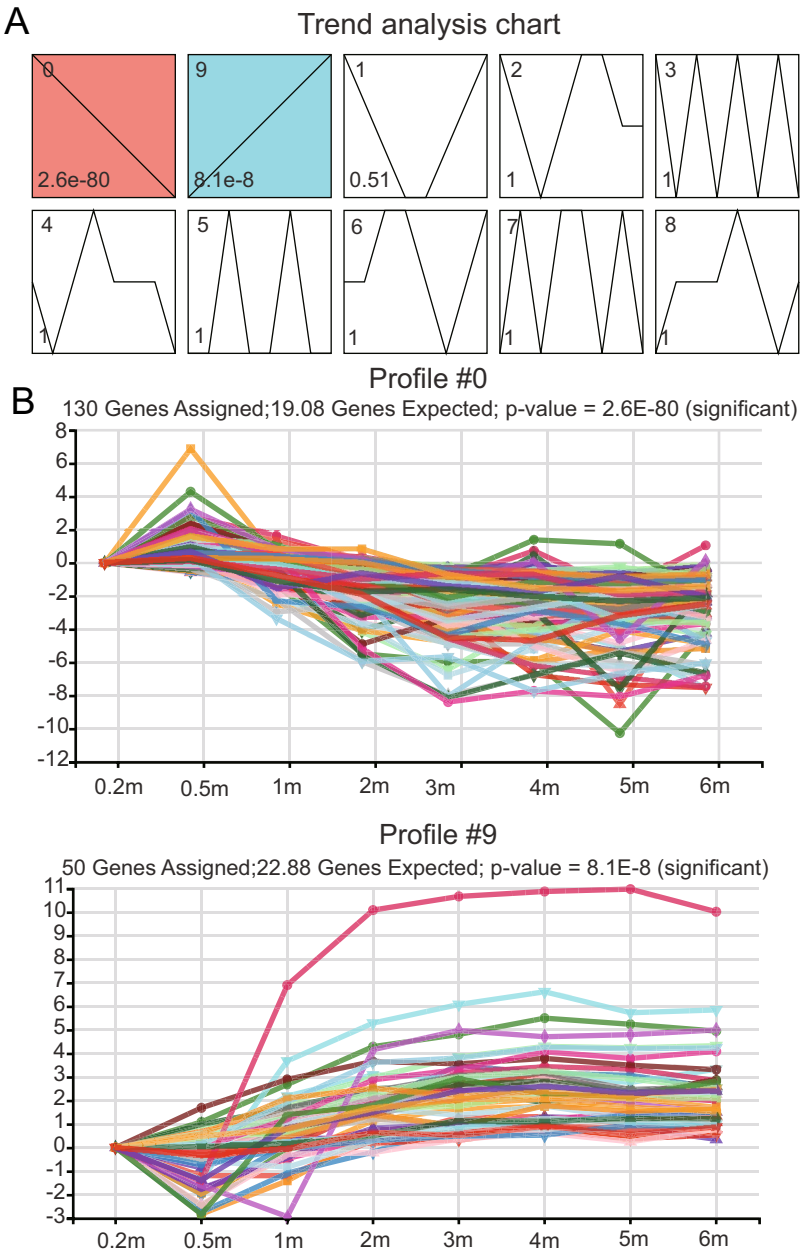

Fig. 9 STEM analysis of *CPP* target genes. **a** Each box corresponds to a type of expression profile, and only the colored profiles are significantly different. The numbers in the upper and lower left corners of each box indicate the order of the profile and the *P*-value, respectively. **b** Profile 0 (above), with a down-regulated pattern; profile 9 (below), with an up-regulated pattern

histone gene (*PH02Gene15233*), where reduced expression of this gene is unlikely to significantly affect plant growth and development under non-stressful conditions [43] (Fig. 10A). Importantly, *PeCPP06* markedly enhanced the expression of MYB [27, 44] transcription factor genes (*PH02Gene38600* and *PH02Gene47557*), which are vital for the swift growth and improved environmental resilience of Moso bamboo shoots. Additionally, *PeCPP06* also modulates the expression of cytochrome P450 genes (*PH02Gene33336* and *PH02Gene43237*) [45], which are critical for the complex biochemical regulation and defense mechanisms within the plant's metabolic network. Moreover, *PeCPP14* participates in the auxin signal transduction pathway by upregulating the AUX/IAA gene (*PH02Gene47180*), facilitating plant growth and development[24] (Fig. 10B). Collectively, these findings

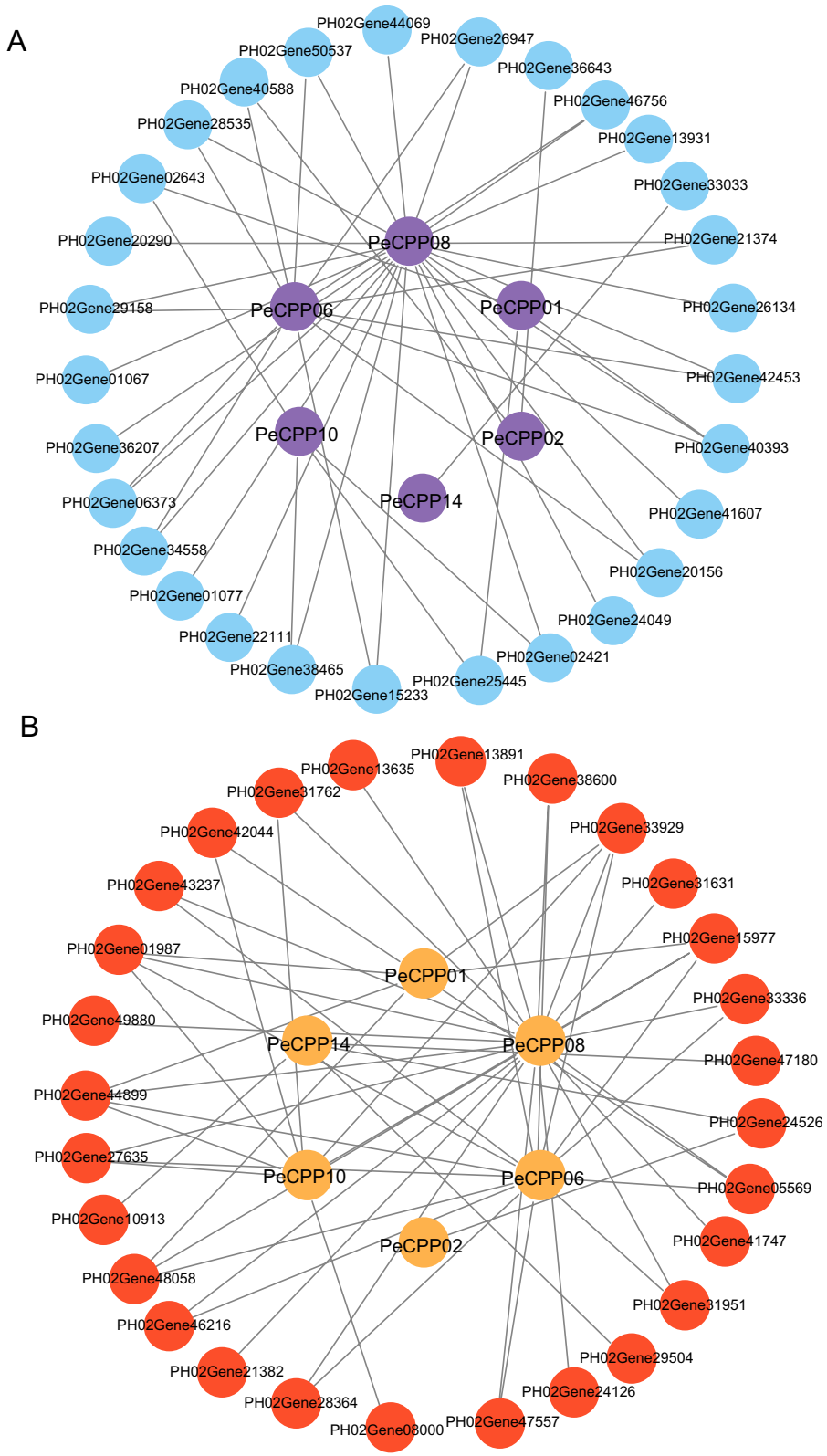

Fig. 10 *PeCPP* gene regulatory networks in Moso bamboo. **a** Predictive network of downstream target genes negatively regulated by *PeCPP* transcription factors. Purple represents *PeCPP* genes, blue represents downstream target genes. **b** Prediction network of downstream target genes positively regulated by *PeCPP* transcription factors. Yellow represents *PeCPP* genes and red represents downstream target genes

demonstrate that CPP transcription factors are pivotal in enhancing the rapid growth and environmental adaptability of Moso bamboo shoots through their regulatory impact on various downstream target genes.

Discussion

The *CPP* family is closely related to plant growth, development and abiotic stress response. It has also been found that the *CPP* family is involved in the development of floral organs and contributes significantly to the control of reproductive tissue development and cell division. Overall, there is little research on the *CPP* gene family in plants, and most studies have been limited to the identification and functional characterization of the *CPP* gene family. There remains a deficiency in the research on the *CPP* gene family in bamboo. Our study not only focuses on the basic identification, structural and functional linkage of the *CPP* gene family in Moso bamboo, but also explores in more depth the expression profile of *CPP* genes in different tissues and organs and under hormone-regulated conditions. We also characterize the role of *CPP* in the rapid development of Moso bamboo shoots by combining the expression characteristics of its regulatory target genes.

The CXC domain is a typical feature of the CPP protein family. Whether in animals or plants, the CXC domain of the CPP protein is highly conserved. But the similarity of other partial sequences is very low. The CXC domain sequence may be the binding domain of transcription factors and DNA, or it may be the binding domain of certain specific metal ions, thereby limiting the variation of its sequence. The CXC domain plays a very important role in the function of transcription factors. In this study, two CXC domains were observed in all CPP proteins except for CPP4. The conserved sequences are CXCX4CX3YCXCX6CX3CXCX2C and CXCX-4CX3YCXCX6C, namely CXC1 and CXC2. The first CXC structural domain is much more conserved than the second CXC structural domain. Additionally, there are other residues that exhibit a high degree of conservation in these sequences [11]. For example, we found that Ser, Lys, Leu, Thr, Tyr, Arg and Cys are highly conserved in the first CXC structural domain (Fig. 4). In previous studies on the *CPP* gene, only sequence information for CXC1 was found. In this study, CXC1 is separated from CXC2 in the CPP I subfamily by a variable-length region containing the short conserved sequence RNPXAFXPK [13, 46]. After structural domain analysis (Fig. 3a), it was found that CXC1 and CXC2 were not sequential. CXC1 can appear before CXC2 or after CXC2. Besides, the short conserved sequence R sequence between CXC1 and CXC2 is not present in the genes of the CPP II and CPP III subfamilies, and it is speculated that the presence or absence of the R sequence may affect the function of the gene. Interestingly, we also identified a new highly conserved Motif5 with the conserved sequence LX2LX-3LX2L, called the L sequence, but the function of this Motif has yet to be verified by further experiments.

As a transcription factor, *CPP* can regulate the expression of downstream target genes through protein-DNA interactions. A thorough assemblage of hydrogen bonds between the backbone and sidechain of the protein and the DNA phosphate backbone, as well as the insertion of a tyrosine from each subdomain into the minor groove, stabilize the protein–DNA complex [38]. In our research, we discovered that there may be a highly conserved DNA-binding domain (DBD) in the tertiary structure of PeCPP proteins, a region that relatively stabilises the protein structure by binding to nucleic acids. Additionally, Moso bamboo contains the protein LIN54 [47, 48], which also has two tandem CXC structural domains. The CXC domain in LIN54 was found to bind specific DNA sequences. LIN54 binding specificity is conferred by interactions between tyrosine and DNA [38]. The phenomenon of *CPP* with two CXC structural domains and similar Tyr-DNA interactions was also found in our research. Therefore, we hypothesized that such interactions may also exist between the tyrosine and DNA in the CPP protein of Moso bamboo, thus affecting the specificity of the CPP protein in binding nucleic acids. In this study, the Moso bamboo *CPP* tyrosine-DNA complex is formed mainly through the interaction of the Tyr174 amino acid residue with the minor groove of the DNA double strand. The discovery of this structural feature has not been documented in previous *CPP*-related studies, and it will help to guide future research into the relationship between *CPP* structure and function. However, the above structural features need to be further verified by X-RAY and other experiments.

Gene expression patterns can offer valuable insights into gene function, and *CPP* family genes have different expression patterns in plants [6]. For example, in Arabidopsis, *TSO1*/*AtCPP5* is highly expressed in flowers and reaches its highest level of expression during ovule and microspore development [5]. *CsCPP* genes are expressed in cucumber leaves, ovaries, flowers, stems, roots and deciduous stem tissues, among others. In the present study, *PeCPP* genes were found to be expressed to varying degrees in leaves, flowers, whips and roots of Moso bamboo. Some genes were only highly expressed only at specific developmental stages, while others showed little to no expression. For instance, *PeCPP05* is only highly expressed in bamboo flowers and is largely unexpressed in roots (Fig. 6c). During root development, only *PeCPP10*, *PeCPP17* and *PeCPP08* showed high levels of expression, while all other genes showed low or no

expression. We found relatively high levels of expression in bamboo leaves in general, suggesting that the *CPP* gene family plays a regulatory role in the developmental stages of Moso bamboo leaves.

The "fast growth" stage of Moso bamboo is concentrated in the short 2–3 month period after shoots. Based on transcriptomic data, many of the major transcriptional regulators required for the proliferation and maturation of bamboo sprouts, such as MYB and bZIP, have been identified [49, 50]. However, there is a gap in the understanding of the role of *CPP* transcription factors and their target genes on the rapid growth and development of Moso bamboo shoots. Expression profiling at eight-time points of a complete shoot development revealed that PeCPPs could reach their highest levels at the early stage of 0.2m bamboo shoots, and it is hypothesized that *PeCPP* is mainly involved in the early growth and development of shoots with an important role. Further GO enrichment and KEGG analysis of the identified *CPP* target genes revealed that these large numbers of downstream target genes can be involved in macromolecular biosynthesis, cellular biosynthetic processes, sugar biosynthesis and metabolism. In addition, STEM analysis of the expression trends of the target genes throughout the rapid development of the shoot revealed the types and numbers of *CPP* positively and negatively regulated target genes. The examination of the interaction between the CPP transcription factors and the downstream target genes indicates that *PeCPP01*, *PeCPP02*, *PeCPP06*, *PeCPP08*, *PeCPP10*, and *PeCPP14* have a significant impact on these genes. The reduced expression of specific gene families in plants has been noticed to have a limited effect on their growth and development. This could be because these genes are not essential under normal growth conditions or because there are other genes that can perform similar functions. During the early phases of hairy bamboo shoot growth, cellular activities such as cell division and elongation are the main focus, requiring significant DNA replication, repair, and ribosome synthesis [51]. The results indicate that the genes inhibited by *CPP* primarily include those related to the synthesis of ribosomes, the assembly of histones, and the reaction to environmental stress. This suggests an increased need for ribosomes and proteins during the initial stages of shoot development in order to support fast cellular division and growth. Once these requirements are fulfilled and the shoots increase in height, CPP's function changes to suppressing these genes. Simultaneously, as bamboo shoots rapidly grow, the *CPP* gene primarily activates genes related to physiological processes, metabolic reactions, tissue development, and environmental adaptation. This aligns with the specific requirements for the growth and development of the shoots. It is hypothesized that *CPP* can regulate the rapid growth and development process of Moso bamboo shoots through the influence of *CPP* target genes.This provides valuable clues to unravel the molecular mechanism of rapid growth of Moso bamboo.

Materials and methods

Plant material

Specimens for experimentation were obtained from Moso bamboo shoots at heights ranging from 0.2 m to 6 m. Each height was represented by three biological replicates. After being collected, the specimens were promptly immersed in liquid nitrogen and stored at −80°C to retain the integrity of the RNA. The FastPure Plant Total RNA Isolation Kit (Vazyme Biotechnology, Nanjing, China, RC401) was employed to gather RNA from each sample. The RNA was subsequently transcribed into cDNA, which was then preserved at a temperature of −20°C.

Identification and physicochemical characterisation of *CPP* family members

Complete genomic data for Moso bamboo were acquired from database (http://gigadb.org/dataset/100498). Utilizing HMMER3, found at website (http://hmmer.janelia.org/), a numerical Hidden Markov Model (Profile HMM) was developed to align with the protein database of Phyllostachys edulis, ensuring the E-value did not exceed 1×10^{-20} for significant matches [52]. The CPP structural domain (PF03638) retrieved from the Pfam database was examined and combined with the Profile HMM findings [53]. The initial screening yielded candidate gene family members of Moso bamboo. Further synthesis of the CPP structural domain of candidate family members and authentication using NCBI BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) [54], resulting in the candidate CPP transcription factor family.

The CDS, protein fasta sequences, and the gene architecture and position data of CPP family members were obtained from the Moso bamboo entire genome database using TBtools (v1.068) program [55]. CELLO (http://cello.life.nctu.edu.tw/) was employed to make predictions for subcellular localization [56]. Protein physicochemical parameters such as molecular weight (MW) and isoelectric point (pI) were predicted using the ExPASy ProtParam tool (https://web.expasy.org/protparam/) [57].

Prediction of gene structure, motifs and conserved structural domains

The Moso bamboo genome annotation files (general feature format, GFF) were analyzed and the Moso bamboo *CPP* family genes were visualized based on their gene location information. Conducted an analysis of the conserved amino acid sequences in CPP proteins utilizing the MEME online suite (http://meme-suite.org/) [58].

The NCBI conserved domain database (https://www.ncbi.nlm.nih.gov/cdd/) was assumed to anticipate the conserved domains in CPPs. Protein structure visualization was performed using DOG 2.0 (http://dog.biocuckoo.org) [59].

Construction of phylogenetic trees and colinearity analysis

Genome-wide details regarding Arabidopsis and rice originated from the TAIR10 database (http://www.arabidopsis.org/index.jsp) and the Rice Genome Annotation Project database (http://rice.plantbiology.msu.edu). A total of eight Arabidopsis CPP proteins and 12 rice CPP proteins were identified through the use of HMMER3 searches on the corresponding region protein libraries. The software ClustalX was utilized to do a multiprotein sequence alignment for Moso bamboo, rice, and Arabidopsis. The alignment data served to generate a maximum likelihood (ML) phylogenetic tree in MEGA 7.0, employing 1000 bootstrap replications to confirm the statistical results' credibility [60].

Through the BLAST module of TBtools (v1.068) software, the sequence alignment of all proteins in the Moso bamboo genome, the pairwise alignment of Moso bamboo and rice, and the genomic protein sequences of Moso bamboo and Arabidopsis thaliana were carried out. Localization of the CPP family on chromosomes and the relationship between different species were analyzed and visualized using MC ScanX, Circos and Multiple Synteny Plot [61].

Analysis of cis-acting elements

The PlantCARE tool was utilized to identify cis-acting regulatory elements within the 1,500 base pair promoter regions situated upstream of the location where transcription begins for each gene (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/). Scientific visualisation of the genomic location of promoters with the help of TBtools software [62].

Homology modeling of protein tertiary structures

The homology template of the protein sequence of CPP was retrieved using the PDB database (http://www.rcsb.org/). Swiss Model (https://www.swissmodel.expasy.org/) was used for homology modeling to obtain the tertiary structure model of CPP gene family proteins. The model was also evaluated using SAVES (https://saves.mbi.ucla.edu/Jobs/683046/pc/saves.sum).

Gene expression analysis

Thirty-one transcript data were downloaded from the EBI and NCBI databases for two hormone treatments including GA and NAA (5 μM), different plant organs of Moso bamboo (roots, rhizomes, panicles, and leaves), and shoot germination stages (0.2 m, 0.5 m, 1 m, 2 m, 3 m, 4 m, 5 m, 6 m). The expression of *PeCPP* genes and related transcription factors was determined for different hormone treatments and developmental stages. The TPM data for each gene were logarithmically transformed using a base 2 logarithm, and TBtools software was used to create heatmaps of gene expression [63].

Identification of *CPP* target genes

To attempt to figure out a collection of genes that may be controlled by CPPs, we extracted promoter sequences across all genes in Moso bamboo, specifically 1000 base pairs in length, using TBtools (v1.0697) [55]. Based on a model of cis-acting elements known to be recognized by CPPs, in combination with the JASPA_CORE database (http://jaspar.genereg.net) [64] and the Motif FIMO program in MEME (http://meme-suite.org/) [58].

GO and KEGG enrichment analysis

Gene Ontology (http://github.com/tanghaibao/GOatools) [65] and KOBAS (http://kobas.cbi.pku.edu.cn/download.php) were used for downstream target genes, GO and KEGG PATHWAY analysis. That Bonferroni multiplex test correction was employed to mitigate the occurrence of false positives [66]. GO functions were deemed substantially enriched if their Bonferroni-corrected *P*-value was < 0.00001. KEGG PATHWAY enrichment analysis of *CPP* target genes was carried out using the R script, calculated in the same way as GO functional enrichment analysis. When the adjusted *P*-value (P-adjust) was less than 0.05, the KEGG PATHWAY function was considered significantly enrichedz.

STEM analysis

Transcriptomic data from various developmental phases of tender shoots were employed. The expression profiles of the *CPP* target genes were investigated using the non-parametric clustering feature of the STEM software [67]. The algorithm proposed by Ernst and Bar-Joseph (2006) was used to calculate the statistical significance of the number of CPPs per profile compared to the expected number.

Construction of *CPP* gene regulatory network

From the steam analysis, we identified two distinct groups of downstream target genes: those that were up-regulated and those that were down-regulated, along with the associated CPP transcription factors. These groups were then correlated, and both target genes and CPP transcription factors were selected based on stringent criteria of an absolute correlation coefficient > 0.9 and *P* value ≤ 0.001. Utilizing these selection parameters,

gene regulatory networks were constructed with the aid of Cytoscape software [68].

Quantitative PCR experiments

Based on the selected gene sequences, specific qRT-PCR primers were designed using Beacon Designer 7.0 (Table S3) [69]. The qRT-PCR analysis was conducted using the CFX-96 Real-Time System (Bio-Rad, USA), with four technical duplicates established for each sample [33]. The *PeActin* was used as the endogenous control for normalising gene expression across different samples [70, 71].Through the utilisation of the $2^{(-\Delta\Delta Ct)}$ approach, the quantification of the relative gene expression levels in several samples was accomplished. Data analysis and visualisation were performed using GraphPad Prism 7 software [72].

Conclusions

In summary, we systematically identified and analyzed 17 *CPP* genes of Moso bamboo on a genome-wide scale. An investigation was conducted on the protein properties, chromosomal location, evolutionary connections, gene structure, conserved motifs, and expression patterns of the subject. Compared with the existing studies on *CPP* genes, this study further investigated the structural characteristics and functional properties of *CPP* from two perspectives: the tertiary structure of CPP proteins and the target genes. Furthermore, transcriptomic and qRT-PCR experiments showed that the *PeCPP* gene family is involved in the development and hormone response of many tissues and organs, and can be involved in the rapid growth and development of shoots through the regulation of downstream target genes.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-024-11084-6.

Supplementary Material 1.

Supplementary Material 2.

Supplementary Material 3.

Acknowledgements

We are appreciative of the lab members' aid and insightful conversations. We are grateful to the International Centre for Bamboo and Rattan for publishing the reference sequence Phyllostachys_edulis_V2.0 at BambooGDB.

Authors' contributions

JT performed data collection and processing, participated in study design and interpretation, performed experiments, and wrote the manuscript. HG and YJ participated in some of the experiments. XX and SS assisted in the interpretation of the results. ZZ is responsible for the completeness of the data and accuracy of the data analysis. The manuscript was reviewed and edited by all authors.

Funding

This study was funded by Zhejiang Provincial Natural Science Foundation of China (grant no. LTGN23C160002).

Data availability

All data generated or analysed in this study have been included in the relevant published articles and their supplements. Gene sequence information (GFF) for the Moso bamboo (Phyllostachys edulis) used in this study can be downloaded from website (http://gigadb.org/dataset/100498). Related protein family information is available on the Pfam website (http://pfam.xfam.org/). Sequence data for analysing the interspecific covariate homology of mosaic bamboo, rice and Arabidopsis thaliana can be obtained from the rice database (http://rice.plantbiology.msu.edu) and the TAIR database (https://www.arabidopsis.org/), respectively. In addition, the transcriptome data analysed in this study have been uploaded to the EMBL database (https://www.ebi.ac.uk/) (accession numbers: ERR105067, ERR105069, ERR105073, and ERR105075) and are also available in the NCBI GEO database (https://www.ncbi. nlm.nih.gov/) (Accession No. GSM2810849: SRR6171235- SRR6171258).

Declarations

Ethics approval and consent to participate

The plant materials used in this study were Tissue samples of bamboo rhizomes, inflorescences, young leaves, roots, and shoots of different heights were obtained from bamboo plants growing in Guilin, Guangxi. This study did not require ethical approval or consent, as it did not involve any endangered or protected species.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 1 June 2024 Accepted: 25 November 2024
Published online: 03 December 2024

References

1. Tatusov RL, Koonin EV, Lipman DJ. A Genomic Perspective on Protein Families. Science (New York, NY). 1997;278(5338):631–7.
2. Thornton JW, DeSalle R. Gene family evolution and homology: genomics meets phylogenetics. Annu Rev Genomics Hum Genet. 2000;1:41–73.
3. Riechmann JL, Heard J, Martin G, Reuber L, Jiang C, Keddie J, Adam L, Pineda O, Ratcliffe OJ, Samaha RR, et al. Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. Science (New York, NY). 2000;290(5499):2105–10.
4. Mitsuda N, Ohme-Takagi M. Functional analysis of transcription factors in Arabidopsis. Plant Cell Physiol. 2009;50(7):1232–48.
5. Andersen SU, Algreen-Petersen RG, Hoedl M, Jurkiewicz A, Cvitanich C, Braunschweig U, Schauser L, Oh SA, Twell D, Jensen E. The conserved cysteine-rich domain of a tesmin/TSO1-like protein binds zinc in vitro and TSO1 is required for both male and female fertility in Arabidopsis thaliana. J Exp Bot. 2007;58(13):3657–70.
6. Lu T, Dou Y, Zhang C. Fuzzy clustering of CPP family in plants with evolution and interaction analyses. BMC Bioinformatics. 2013, 14 Suppl 13(Suppl 13):S10.
7. Yang Z, Gu S, Wang X, Li W, Tang Z, Xu C. Molecular evolution of the CPP-like gene family in plants: insights from comparative genomics of Arabidopsis and rice. J Mol Evol. 2008;67(3):266–77.
8. Zhang L, Zhao HK, Wang YM, Yuan CP, Zhang YY, Li HY, Yan XF, Li QY, Dong YS. Genome-wide identification and expression analysis of the CPP-like gene family in soybean. Genet Mol Res. 2015;14(1):1260–8.
9. Sun Y, Jia X, Chen D, Fu Q, Chen J, Yang W, Yang H, Xu X. Genome-wide identification and expression analysis of cysteine-rich polycomb-like protein (CPP) gene family in tomato. Int J Mol Sci. 2023;24(6):5762.
10. Ullah U, Buttar ZA, Shalmani A, Muhammad I, Ud-Din A, Ali H. Genome-wide identification and expression analysis of CPP-like gene family in

Triticum aestivum L. under different hormone and stress conditions. Open life Sci. 2022;17(1):544–562.
11. Zhou Y, Hu L, Ye S, Jiang L, Liu S. Genome-wide identification and characterization of cysteine-rich polycomb-like protein (CPP) family genes in cucumber (Cucumis sativus) and their roles in stress responses. Biologia. 2018;73(4):425–35.
12. Song XY, Zhang YY, Wu FC, Zhang L. Genome-wide analysis of the maize (Zea may L.) CPP-like gene family and expression profiling under abiotic stress. Genet Mol Res. 2016;15(3):gmr.15038023.
13. Cvitanich C, Pallisgaard N, Nielsen KA, Hansen AC, Larsen K, Pihakaski-Maunsbach K, Marcker KA, Jensen EO. CPP1, a DNA-binding protein involved in the expression of a soybean leghemoglobin c3 gene. Proc Natl Acad Sci USA. 2000;97(14):8163–8.
14. Song JY, Leung T, Ehler LK, Wang C, Liu Z. Regulation of meristem organization and cell division by TSO1, an Arabidopsis gene with cysteine-rich repeats. Development (Cambridge, England). 2000;127(10):2207–17.
15. Hauser BA, He JQ, Park SO, Gasser CS. TSO1 is a novel protein that modulates cytokinesis and cell expansion in Arabidopsis. Development (Cambridge, England). 2000;127(10):2219–26.
16. Hauser BA, Villanueva JM, Gasser CS. Arabidopsis TSO1 regulates directional processes in cells during floral organogenesis. Genetics. 1998;150(1):411–23.
17. Liu Z, Running MP, Meyerowitz EM. TSO1 functions in cell division during Arabidopsis flower development. Development (Cambridge, England). 1997;124(3):665–72.
18. Wei Q, Jiao C, Ding Y, Gao S, Guo L, Chen M, Hu P, Xia S, Ren G, Fei Z. Cellular and molecular characterizations of a slow-growth variant provide insights into the fast growth of bamboo. Tree Physiol. 2018;38(4):641–54.
19. Atanda J. Environmental impacts of bamboo as a substitute constructional material in Nigeria. Case Studies in Construction Materials. 2015;3:33–9.
20. Peng Z, Lu Y, Li L, Zhao Q, Feng Q, Gao Z, Lu H, Hu T, Yao N, Liu K, et al. The draft genome of the fast-growing non-timber forest species moso bamboo (Phyllostachys heterocycla). Nat Genet. 2013;45(4):456–461, 461e451–452.
21. Zhao H, Gao Z, Wang L, Wang J, Wang S, Fei B, Chen C, Shi C, Liu X, Zhang H, et al. Chromosome-level reference genome and alternative splicing atlas of moso bamboo (Phyllostachys edulis). GigaSci. 2018;7(10):giy115.
22. Andrade LM, Peixoto-Junior RF, Ribeiro RV, Nóbile PM, Brito MS, Marchiori PER, Carlin SD, Martins APB, Goldman MHS, Llerena JPP, et al. Biomass Accumulation and Cell Wall Structure of Rice Plants Overexpressing a Dirigent-Jacalin of Sugarcane (ShDJ) Under Varying Conditions of Water Availability. Front Plant Sci. 2019;10:65.
23. Pan F, Wang Y, Liu H, Wu M, Chu W, Chen D, Xiang Y. Genome-wide identification and expression analysis of SBP-like transcription factor genes in Moso Bamboo (Phyllostachys edulis). BMC Genomics. 2017;18(1):486.
24. Li F, Wu M, Liu H, Gao Y, Xiang Y. Systematic identification and expression pattern analysis of the Aux/IAA and ARF gene families in moso bamboo (Phyllostachys edulis). Plant physiology and biochemistry : PPB. 2018;130:431–44.
25. Zhang Y, Tang D, Lin X, Ding M, Tong Z. Genome-wide identification of MADS-box family genes in moso bamboo (Phyllostachys edulis) and a functional analysis of PeMADS5 in flowering. BMC Plant Biol. 2018;18(1):176.
26. Wu R, Shi Y, Zhang Q, Zheng W, Chen S, Du L, Lu C. Genome-wide identification and characterization of the UBP gene family in moso bamboo (Phyllostachys edulis). Int J Mol Sci. 2019;20(17):4309.
27. Yang K, Li Y, Wang S, Xu X, Sun H, Zhao H, Li X, Gao Z. Genome-wide identification and expression analysis of the MYB transcription factor in moso bamboo (Phyllostachys edulis). PeerJ. 2019;6:e6242.
28. Wu HL, Lv H, Li L, Liu J, Mu SH, Li XP, Gao J. Genome-Wide Analysis of the AP2/ERF Transcription Factors Family and the Expression Patterns of DREB Genes in Moso Bamboo (Phyllostachys edulis). PLoS ONE. 2015;10(5):e0126657.
29. Bai Q, Hou D, Li L, Cheng Z, Ge W, Liu J, Li X, Mu S, Gao J. Genome-wide analysis and expression characteristics of small auxin-up RNA (SAUR) genes in moso bamboo (Phyllostachys edulis). Genome. 2017;60(4):325–36.
30. Sun H, Li L, Lou Y, Zhao H, Gao Z. Genome-wide identification and characterization of aquaporin gene family in moso bamboo (Phyllostachys edulis). Mol Biol Rep. 2016;43(5):437–50.
31. Wu M, Li Y, Chen D, Liu H, Zhu D, Xiang Y. Genome-wide identification and expression analysis of the IQD gene family in moso bamboo (Phyllostachys edulis). Sci Rep. 2016;6:24520.
32. Chen D, Chen Z, Wu M, Wang Y, Wang Y, Yan H, Xiang Y. Genome-wide identification and expression analysis of the HD-Zip gene family in moso bamboo (Phyllostachys edulis). J Plant Growth Regul. 2017;36:323–37.
33. Ma R, Huang B, Chen J, Huang Z, Yu P, Ruan S, Zhang Z. Genome-wide identification and expression analysis of dirigent-jacalin genes from plant chimeric lectins in Moso bamboo (Phyllostachys edulis). PLoS ONE. 2021;16(3):e0248318.
34. Wang T, Wang H, Cai D, Gao Y, Zhang H, Wang Y, Lin C, Ma L, Gu L. Comprehensive profiling of rhizome-associated alternative splicing and alternative polyadenylation in moso bamboo (Phyllostachys edulis). Plant J. 2017;91(4):684–99.
35. Chou KC, Shen HB. Plant-mPLoc: a top-down strategy to augment the power for predicting plant protein subcellular localization. PLoS ONE. 2010;5(6):e11335.
36. Rogozin IB, Sverdlov AV, Babenko VN, Koonin EV. Analysis of evolution of exon-intron structure of eukaryotic genes. Brief Bioinform. 2005;6(2):118–34.
37. Guo M, Yin YX, Ji JJ, Ma BP, Lu MH, Gong ZH. Cloning and expression analysis of heat-shock transcription factor gene CaHsfA2 from pepper (Capsicum annuum L.). Genet Mol Res. 2014;13:1865–1875.
38. Marceau AH, Felthousen JG, Goetsch PD, Iness AN, Lee HW, Tripathi SM, Strome S, Litovchick L, Rubin SM. Structural basis for LIN54 recognition of CHR elements in cell cycle-regulated promoters. Nat Commun. 2016;7(1):1–11.
39. Ibraheem O, Botha CE, Bradley G. In silico analysis of cis-acting regulatory elements in 5′ regulatory regions of sucrose transporter gene families in rice (Oryza sativa Japonica) and Arabidopsis thaliana. Comput Biol Chem. 2010;34(5–6):268–83.
40. Yu TF, Zhao WY, Fu JD, Liu YW, Chen M, Zhou YB, Ma YZ, Xu ZS, Xi YJ. Genome-wide analysis of CDPK family in foxtail millet and determination of SiCDPK24 functions in drought stress. Front Plant Sci. 2018;9:651.
41. Mou S, Liu Z, Guan D, Qiu A, Lai Y, He S. Functional analysis and expressional characterization of rice ankyrin repeat-containing protein, OsPI-ANK1, in basal defense against Magnaporthe oryzae attack. PLoS ONE. 2013;8(3):e59699.
42. Mao T, Geng Z, Zhang Y, Xue W, Ma L, Yang J, Jin Y, Wang S, Zhuo S, Zhang Y, et al. Genome-wide characterization of NmrA-like proteins and the regulatory function of soybean GmNmrA6 in response to salt and oxidative stresses. Environ Exp Bot. 2023;213:105447.
43. Sheikh AH, Nawaz K, Tabassum N, Almeida-Trapp M, Mariappan KG, Alhoraibi H, Rayapuram N, Aranda M, Groth M, Hirt H. Linker histone H1 modulates defense priming and immunity in plants. Nucleic Acids Res. 2023;51(9):4252–65.
44. Biswas D, Gain H, Mandal A. MYB transcription factor: A new weapon for biotic stress tolerance in plants. Plant Stress. 2023;10:100252.
45. Xia Y, Yang J, Ma L, Yan S, Pang Y. Genome-wide identification and analyses of drought/salt-responsive cytochrome P450 genes in Medicago truncatula. Int J MolSci. 2021;22(18):9957.
46. Sutou S, Miwa K, Matsuura T, Kawasaki Y, Ohinata Y, Mitsui Y. Native tesmin is a 60-kilodalton protein that undergoes dynamic changes in its localization during spermatogenesis in mice. Biol Reprod. 2003;68(5):1861–9.
47. Korenjak M, Taylor-Harding B, Binné UK, Satterlee JS, Stevaux O, Aasland R, White-Cooper H, Dyson N, Brehm A. Native E2F/RBF complexes contain Myb-interacting proteins and repress transcription of developmentally controlled E2F target genes. Cell. 2004;119(2):181–93.
48. Lewis PW, Beall EL, Fleischer TC, Georlette D, Link AJ, Botchan MR. Identification of a Drosophila Myb-E2F2/RBF transcriptional repressor complex. Genes Dev. 2004;18(23):2929–40.
49. Peng Z, Zhang C, Zhang Y, Hu T, Mu S, Li X, Gao J. Transcriptome sequencing and analysis of the fast growing shoots of moso bamboo (Phyllostachys edulis). PLoS ONE. 2013;8(11):e78944.
50. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 2013;14(4):R36.
51. Tao G-Y, Ramakrishnan M, Vinod KK, Yrjälä K, Satheesh V, Cho J, Fu Y, Zhou M. Multi-omics analysis of cellular pathways involved in different rapid growth stages of moso bamboo. Tree Physiol. 2020;40(11):1487–508.

52. Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity searching. Nucleic Acids Res. 2011;39(Web Server issue):W29–37.
53. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, et al. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. 2016;44(D1):D279-285.
54. Jin J, Tian F, Yang DC, Meng YQ, Kong L, Luo J, Gao G. PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. Nucleic Acids Res. 2017;45(D1):D1040–5.
55. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. Mol Plant. 2020;13(8):1194–202.
56. Huang B, Xu C. Identification and characterization of proteins associated with plant tolerance to heat stress. J Integr Plant Biol. 2008;50(10):1230–7.
57. Gasteiger E, Hoogland C, Gattiker A, Duvaud Se, Wilkins MR, Appel RD, Bairoch A. Protein identification and analysis tools on the ExPASy server; Humana Press: The Proteomics Protocols Handbook. 2005.
58. 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(Web Server issue):W202–208.
59. Ren J, Wen L, Gao X, Jin C, Xue Y, Yao X. DOG 1.0: illustrator of protein domain structures. Cell Res. 2009;19(2):271–3.
60. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 2016;33(7):1870–4.
61. Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee TH, 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(7):e49.
62. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S. 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(1):325–7.
63. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 2012;7(3):562–78.
64. Khan A, Fornes O, Stigliani A, Gheorghe M, Castro-Mondragon JA, van der Lee R, Bessy A, Chèneby J, Kulkarni SR, Tan G, et al. JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. Nucleic Acids Res. 2018;46(D1):D260-d266.
65. Klopfenstein DV, Zhang L, Pedersen BS, Ramírez F, Warwick Vesztrocy A, Naldi A, Mungall CJ, Yunes JM, Botvinnik O, Weigel M, et al. GOATOOLS: A Python library for Gene Ontology analyses. Sci Rep. 2018;8(1):10872.
66. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J Roy Stat Soc: Ser B (Methodol). 2018;57(1):289–300.
67. Ernst J, Bar-Joseph Z. STEM: a tool for the analysis of short time series gene expression data. BMC Bioinformatics. 2006;7:191.
68. Jiao Y, Tan J, Guo H, Huang B, Ying Y, Ramakrishnan M, Zhang Z. Genome-wide analysis of the KNOX gene family in Moso bamboo: insights into their role in promoting the rapid shoot growth. BMC Plant Biol. 2024;24(1):213.
69. Thornton B, Basu C. Real-time PCR (qPCR) primer design using free online software. Biochemistry and molecular biology education : a bimonthly publication of the International Union of Biochemistry and Molecular Biology. 2011;39(2):145–54.
70. Fan C, Ma J, Guo Q, Li X, Wang H, Lu M. Selection of reference genes for quantitative real-time PCR in bamboo (Phyllostachys edulis). PLoS ONE. 2013;8(2):e56573.
71. Guo H, Tan J, Jiao Y, Huang B, Ma R, Ramakrishnan M, Qi G, Zhang Z. Genome-wide identification and expression analysis of the HAK/KUP/KT gene family in Moso bamboo. Front Plant Sci. 2024;15:1331710.
72. Mitteer DR, Greer BD, Fisher WW, Cohrs VL. Teaching behavior technicians to create publication-quality, single-case design graphs in graphpad prism 7. J Appl Behav Anal. 2018;51(4):998–1010.

Publisher's Note