BMC Genomics

RESEARCH Open Access

Genome-wide identification and characterization of the *AP2/ERF* gene family in *Cinnamomum camphora*

Zerui Yang[1,2*], Xinnan Yang[1,2] and Xiasheng Zheng[3*]

Abstract

Background The camphor tree (*Cinnamomum camphora*) is a valuable source of essential oils rich in monoterpenes and sesquiterpenes, which exhibit antibacterial, antioxidant, and insecticidal properties. The AP2/ERF transcription factor family is known for its roles in regulating secondary metabolite biosynthesis and stress responses in plants. However, the functions and regulatory mechanisms of AP2/ERF transcription factors in camphor tree remain largely unexplored.

Results A total of 154 unique *AP2/ERF* genes were identified in the camphor tree genome and classified into four subfamilies (RAV, AP2, ERF, and DREB), among which the ERF subfamily comprises 75 members and represents the largest. The physicochemical properties of the proteins were diverse, with predicted nuclear localization. Phylogenetic analysis revealed distinct clustering patterns consistent with their subfamily classification. Gene structure analysis showed variation in exon-intron organization among subfamilies. A total of 5,910 cis-acting elements were identified in the promoters, with MeJA-responsive and stress responsiveness elements being most prevalent. Expression profile analysis indicated tissue-specific expression patterns, with most *CcAP2/ERF* genes highly expressed in the roots. Correlation analysis revealed that 44 *CcAP2/ERF* genes were strongly correlated with 6 terpene synthase (TPS) genes involved in monoterpene biosynthesis, suggesting regulatory functions. Strikingly, *CcERF_104* was identified as a high-potential hub regulator, co-expressed two with borneol diphosphate synthase (BPPS) genes (*CcTPS26* and *CcTPS49*). Moreover, a suite of regulators (including *CcERF_27*, *CcERF_42*, *CcERF_83*, *CcERF_87*, *CcAP2_12*, and *CcAP2_13*) showed strong positive correlation with *CcTPS72*, the most catalytically efficient BPPS synthase in *C. camphora*, highlighting sophisticated transcriptional control over monoterpene biosynthesis.

Conclusion This study provides a comprehensive analysis of the AP2/ERF transcription factor family in *C. camphora* and elucidates their potential roles in regulating terpene biosynthesis. The findings lay the foundation for future functional characterization of key *CcAP2/ERF* genes using genetic and biochemical approaches, with the goal of enhancing terpenoid content in *C. camphora* leaves through genetic improvement technologies.

*Correspondence:
Zerui Yang
yzr_gzucm1991@163.com
Xiasheng Zheng
xszheng@gzucm.edu.cn

[1]Key Specialty of Clinical Pharmacy, The First Affiliated Hospital of Guangdong Pharmaceutical University, Guangzhou, Guangdong, China
[2]School of Traditional Chinese Medicine, Guangdong Pharmaceutical University, Guangzhou, Guangdong, China
[3]School of Pharmaceutical Sciences, Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, China

Introduction

Cinnamomum camphora, commonly known as the camphor tree, is a valuable broad-leaf evergreen species within the Lauraceae family. It has been cultivated in China for over 2000 years and is highly regarded for its high-quality timber and aromatic properties [1]. Essential oil derived from the leaves and twigs of camphor tree is rich in monoterpenes, sesquiterpenes, and other aromatic compounds, which exhibits a broad spectrum of pharmacological activities, such as antibacterial, antioxidant, and insecticidal effects [2].

The camphor tree can be classified into five chemotypes based on the predominant components in its essential oil extracted from its leaves. These components include linalool, D-borneol, camphor, cineole, and nerolidol [3, 4]. These terpenoids are highly valued for their industrial and pharmaceutical applications. For instance, D-borneol is a well-known traditional Chinese medicine widely used for treating cardiovascular diseases such as stroke, coronary heart disease, and angina pectoris [5]. It has been included in multiple editions of the Chinese Pharmacopoeia and is in high demand as a key ingredient in many traditional Chinese herbal formulations [6–9]. Additionally, D-borneol is used to alleviate pain associated with wounds, injuries, burns, and cuts. To enhance the yield and quality of essential oils, which hold substantial commercial value, a comprehensive understanding of the metabolic pathways and regulatory mechanisms of monoterpenoids is crucial. Currently, the understanding of the regulatory mechanisms of terpene-related genes in *C. camphora* is still limited. Previous studies have primarily focused on the structural genes involved in the terpenoid metabolic pathway [1, 10]. To gain a deep insight into the molecular mechanisms underlying the biosynthesis of these aromatic compounds, it is essential to investigate the transcription factors that regulate this process, which will not only help elucidate the underlying mechanisms but also provide a theoretical basis for the cultivation of high-quality *C. camphora*.

The APETALA2/ethylene-responsive factor (AP2/ERF) transcription factor superfamily is extensively present in plants and has vital functions in various aspects of plant growth, development, and stress responses [11, 12]. Moreover, these transcription factors are key regulators of the biosynthesis of secondary metabolites, which are crucial for plant defense and adaptation mechanisms. Based on differences in the number and sequence of AP2 structural domains, the AP2/ERF family can be divided into five subfamilies: AP2, DREB, ERF, RAV, and Soloist [13, 14]. Previous research has extensively explored the AP2/ERF family in a variety of plant species, such as *Arabidopsis thaliana* [15], *Coptis Chinensis* [16], *Zanthoxylum armatum* [17], and *Salvia miltiorrhiza* [18]. Moreover, researchers have identified AP2/ERF genes that regulate terpenoid biosynthesis in various plant species. For example, CitERF71 in sweet orange (*Citrus sinensis* Osbeck) activates the expression of the terpene synthase gene *CiTPS16*, thereby promoting the release of geraniol [19]. LcERF19, an AP2/ERF transcription factor from *Litsea cubeba*, positively regulates geranial and neral biosynthesis [20]. In maize, EreB58 directly binds to the GCC-box in the promoter of *TPS10*, promoting its expression and thus regulating the biosynthesis of sesquiterpenes in leaves [21]. These studies demonstrate the regulatory roles of AP2/ERF genes in terpenoid biosynthesis across different plants. However, the AP2/ERF transcription factors in *C. camphora*. remain relatively under-investigated.

In this study, we employed a comprehensive approach involving genomic and transcriptome data analysis of *C. camphora*, combined with classification, phylogenetic analysis, and examination of expression patterns to identify *CcAP2/ERF* transcription factors. Additionally, terpene-related transcription factors were pinpointed through correlation analysis. Our findings offer novel insights into the role of AP2/ERF transcription factors in regulating terpene biosynthesis in *C. camphora*.

Materials & methods

Plant materials

C. camphora of the borneol chemotype was procured from Ji'an Yu Feng Natural Species Co., Ltd (Ji'an, Jiangxi, China) and was subsequently maintained under cultivation at The First Affiliated Hospital of Guangdong Pharmaceutical University. The leaf, stem, and root of *C. camphora* were sampled from plants of 2-year-old. A portion of these samples was immediately frozen in liquid nitrogen at −80 °C for subsequent RNA extraction. Additionally, all experimental procedures were carried out in triplicate for biological replication in the study.

Identification and physicochemical properties analysis of CcAP2/ERF genes

The AP2/ERF domain (PF00847) hidden Markov model (HMM) profile was obtained from the HMMER 3.0 database (http://hmmer.org/). This profile was subsequent ly employed to perform a genome-wide search against the *C. camphora* genome assembly available in the Global Pharmacopoeia Genome Database (http://www.gpgenome.com/species/206). The AP2/ERF candidate sequences were validated using the Conserved Domain Database (CDD) tool available on the NCBI website (https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi/) and the SMART database (https://smart.embl.de/). The gene structures of certain *CcAP2/ERF* family members were manually corrected using the IGV-GSAman software (https://gitee.com/CJchen/IGV-sRNA), based on next-generation transcriptome data (project number:

PRJNA1189175). Meanwhile, we used the Gene Structure Error Detect plugin in TBtools [22] to examine each gene, ensuring that every gene had a quality score greater than 0.3, which indicates no issues with the gene structure (Supplementary Table S1).

The physicochemical properties of the confirmed CcAP2/ERF proteins, including aliphatic index, grand average of hydrophobicity (GRAVY), molecular weight, protein size, and theoretical isoelectric point, were predicted using the ProtParam tool on the ExPASy website (https://web.expasy.org/protparam/) [23]. Additionally, subcellular localization patterns of these proteins were predicted using DeepLocPro (https://services.healthtech.dtu.dk/services/DeepLocPro-1.0/) [24].

Phylogenetic analysis and chromosomal locations of CcAP2/ERF genes

A neighbor-joining (NJ) phylogenetic tree was generated using MEGA11 software [25] and subsequently visualized with Evolview v3 [26]. To evaluate the robustness of the tree topology, bootstrap analysis was conducted with 1000 replicates. The positions of all genes were obtained from the *C. camphora* GFF3 annotation file and then visualized across the 12 chromosomes using TBtools software [22].

Analysis of gene structure, protein conserved motifs, and domains of CcAP2/ERFs

To characterize the conserved domains of CcAP2/ERF proteins, we employed the CD-Search tool from the NCBI Conserved Domain Database (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). The conserved motifs within these proteins were further analyzed using the MEME program available online (https://meme-suite.org/meme/tools/meme) [27]. Finally, the identified conserved domains, motifs and gene structures were visualized using TBtools software [22].

Gene duplication and genome synteny analysis

To assess interspecies replication events in *C. camphora*, the Multiple Collinear Scanning Tool (MCScanX) was employed [28]. Genomic data for *Arabidopsis thaliana*, *Camphora kanahirae*, *Solanum lycopersicum* and *Oryza sativa* were obtained from NCBI (https://www.ncbi.nlm.nih.gov/). Subsequently, the collinear gene relationships between *C. camphora* and these four species were analyzed using the Multiple Synteny Plot function in TBtools and visualized with TBtools [22].

Analysis of cis-acting elements in the promoters of CcAP2/ERF genes

To identify potential cis-acting regulatory elements, the 2.0 kb promoter sequence upstream of the transcriptional start site of the *CcAP2/ERF* genes was submitted to PlantCARE, an online bioinformatics tool accessible at http://bioinformatics.psb.ugent.be/webtools/plantcare/html/ [29].

Expression profle analysis of CcAP2/ERF superfamily genes

In our previous investigation, RNA-Seq was conducted on three distinct tissues (stem, root, and leaf) of *C. camphora* [30]. These data were utilized to examine the expression profiles of *CcAP2/ERF* transcription factors. The expression levels of *CcAP2/ERF* genes were quantified in Fragments Per Kilobase of transcript per Million mapped reads (FPKM). To visualize the transcriptomic FPKM values, Tbtools software was employed, and gene expression heatmaps were generated using log2(FPKM) transformed data.

Correlation analysis and co-expression network construction

A correlation analysis was conducted to examine the relationship between the quantitative expression values of nine TPS genes (*CcTPS23*, *CcTPS26*, *CcTPS32*, *CcTPS44*, *CcTPS49*, *CcTPS54*, *CcTPS59*, *CcTPS72* and *CcTPS74*) and *CcAP2/ERF* genes in different tissues. This was achieved by calculating Pearson's correlation coefficients (PCC) with the aid of the Hmisc package in the R programming environment. Ultimately, a network map linking *CcAP2/ERF* genes and *CcTPS* genes was generated based on the significant Spearman correlation results (with correlation coefficients exceeding 0.9 and *P-values* below 0.05), utilizing Cytoscape [31].

Quantitative real-time PCR analysis (qRT-PCR)

To ascertain the precision and dependability of the RNA-sequencing data, a selection of 9 *CcAP2/ERF* genes associated with the *CcTPSs* were chosen for confirmation through quantitative real-time PCR (qRT-PCR). Total RNA from the leaves, stems and roots of *C. camphora* were isolated using the SteadyPure Universal RNA Extraction Kit (AG21017, Accurate Biology, Hunan, China). Following this, 1 μg of the isolated RNA was converted into cDNA in a 20 μL reaction using the Evo M-MLV Reverse Transcription Mix Kit with gDNA Removal for qPCR, Version 2 (Accurate Biology, Hunan, China). The synthesized cDNA was subjected to qPCR analysis using the SYBR Green Pro Taq HS qPCR Kit (Accurate Biology, Hunan, China) on a Bio-Rad CFX96 Touch™ Real-Time PCR Detection System (Bio-Rad Laboratories, USA). Ccactin served as the internal control, and gene expression was quantified via the $2^{\wedge\text{-}\Delta\Delta Ct}$ method. The primers specific to the genes of interest are detailed in the Supplementary Table S2.

Results

Identification and analysis of AP2/ERF superfamily genes in the Camphor tree genome

Utilizing both Hidden Markov Model (HMM) and BLAST search techniques, we identified all potential members of the *AP2/ERF* gene family in the camphor tree genome. After eliminating duplicates, 154 unique *AP2/ERF* genes were retained. These candidates were then subjected to detailed analysis regarding their amino acid sequence lengths, isoelectric points (pI), molecular weights (MW), and predicted subcellular locations (Table S3). The lengths of the encoded proteins varied from a minimum of 126 amino acids for CcERF_18 to a maximum of 722 amino acids for CcAP2_21, with an overall average length of 282.5 amino acids. The molecular weights spanned from a low of 14,184.71 Da for CcERF_18 to a high of 78,405.1 Da for CcAP2_21, with an average molecular weight of 31,215.87 Da. The calculated isoelectric points for these proteins ranged between 4.58 for CcERF_43 and 10.05 for CcERF_63, averaging at 6.45. All of the AP2/ERF proteins were predicted to be nuclear-localized.

Classification and phylogenetic analysis of CcAP2/ERF protein

Each subfamily of the AP2/ERF family is characterized by distinct structural domain profiles. Following the classification approach used for *Arabidopsis*, we leveraged both the structural domain attributes and the sequence homologies to systematically classify the AP2/ERF proteins in camphor tree. According to the *A. thaliana* classification, all the CcAP2/ERF genes were grouped into four subfamilies, which were RAV, AP2, ERF and DREB. Among them, the ERF was the largest group, comprising 75 members and could be divided into seven distinct subgroups (designated as ERF-V to ERF-X). Groups ERF-I to ERF-IV belong to the DREB family, with 8, 12, 29, and 5 members, respectively. Finally, the AP2 subfamily comprised 21 genes while the RAV subfamily comprised only four candidate genes (Fig. 1, Table S4). These *CcAP2/ERF* genes were systematically renamed based on the classification of the AP2/ERF superfamily and the chromosomal positions of their respective genes (Table S3).

Gene structure, motif composition, conserved domains analysis of the CcAP2/ERF genes

To gain deeper insights into the features of *CcAP2/ERF* transcription factors, the gene structures were depicted using TBtools software. It was observed that the number of exons in *AP2/ERF* genes showed considerable variation, spanning from a single exon to as many as ten. Within the ERF family, the majority of *CcERF* genes lack introns within their open reading frames. As illustrated in Fig. 2, only nineteen *CcERF* genes possess at least one intron(s), predominantly found in Groups V, VII, and X, where the intron positions are conserved across each group. Meanwhile, AP2 family genes include at least six exons. Conversely, RAV family genes are typically intronless. These findings further confirm the accuracy of the phylogenetic trees.

In this research, the MEME software was utilized to identify a total of ten distinct motifs within the CcAP2/ERF protein family. These motifs were designated as Motif 1 through Motif 10, and motifs 1 was present in all CcAP2/ERF proteins (Fig. 3). Furthermore, a consistent pattern in the distribution of conserved motifs among CcAP2/ERF proteins within the same group suggests that these proteins may share similar functions within their respective groups.

The conserved domains of the CcAP2/ERFs were identified through the NCBI online resource, which revealed that the predominant domain types present are the AP2 domain and the B3 domain (Fig. 4). Each protein possessed at least one highly conserved AP2 domain. Specifically, members of the AP2 subfamily featured either two AP2 domains or a combination of one AP2 domain and one domain from the broader AP2 superfamily. In contrast, RAV subfamily members were characterized by the presence of one AP2 domain and one B3 domain, which corroborated the classification of the family members.

Chromosomal localization and synteny analysis of the CcAP2/ERF gene

The chromosomal positioning of genes is frequently a critical determinant of their functional roles. In the camphor tree genome, each chromosome harbors a significant number of *AP2/ERF* genes, with over 5 on each (Fig. 5). The 154 identified *CcAP2/ERF* genes are unevenly spread across 12 chromosomes. The chromosomes with the highest counts of *AP2/ERF* family members are LG02, LG08, and LG07, housing 26, 23, and 17 members, respectively. Conversely, LG10 and LG12 has the smallest number of family members, with only 5, respectively.

Collinearity analysis is a valuable tool for uncovering functional genes and tracing the evolutionary trajectories of gene families [32]. To identify homologous genes within the *CcAP2/ERF* family that may share similar functions, intraspecific synteny analysis was performed. A total 53 pairs of synteny relationships among the *CcAP2/ERF* genes were found (Table S5). Specifically, 11 pairs of AP2 genes exhibited synteny, including *CcAP2_1/CcAP2_21*, *CcAP2_10/CcAP2_12*, *CcAP2_13/CcAP2_14*, *CcAP2_13/CcAP2_20*, *CcAP2_2/CcAP2_10*, *CcAP2_2/CcAP2_12*, *CcAP2_3/CcAP2_4*, *CcAP2_5/CcAP2_18*, *CcAP2_6/CcAP2_17*, *CcAP2_7/CcAP2_20*, *CcAP2_9/CcAP2_11*. Additionally, one pairs of RAV genes were found to be syntenic: *CcRAV_2/CcRAV_3*.

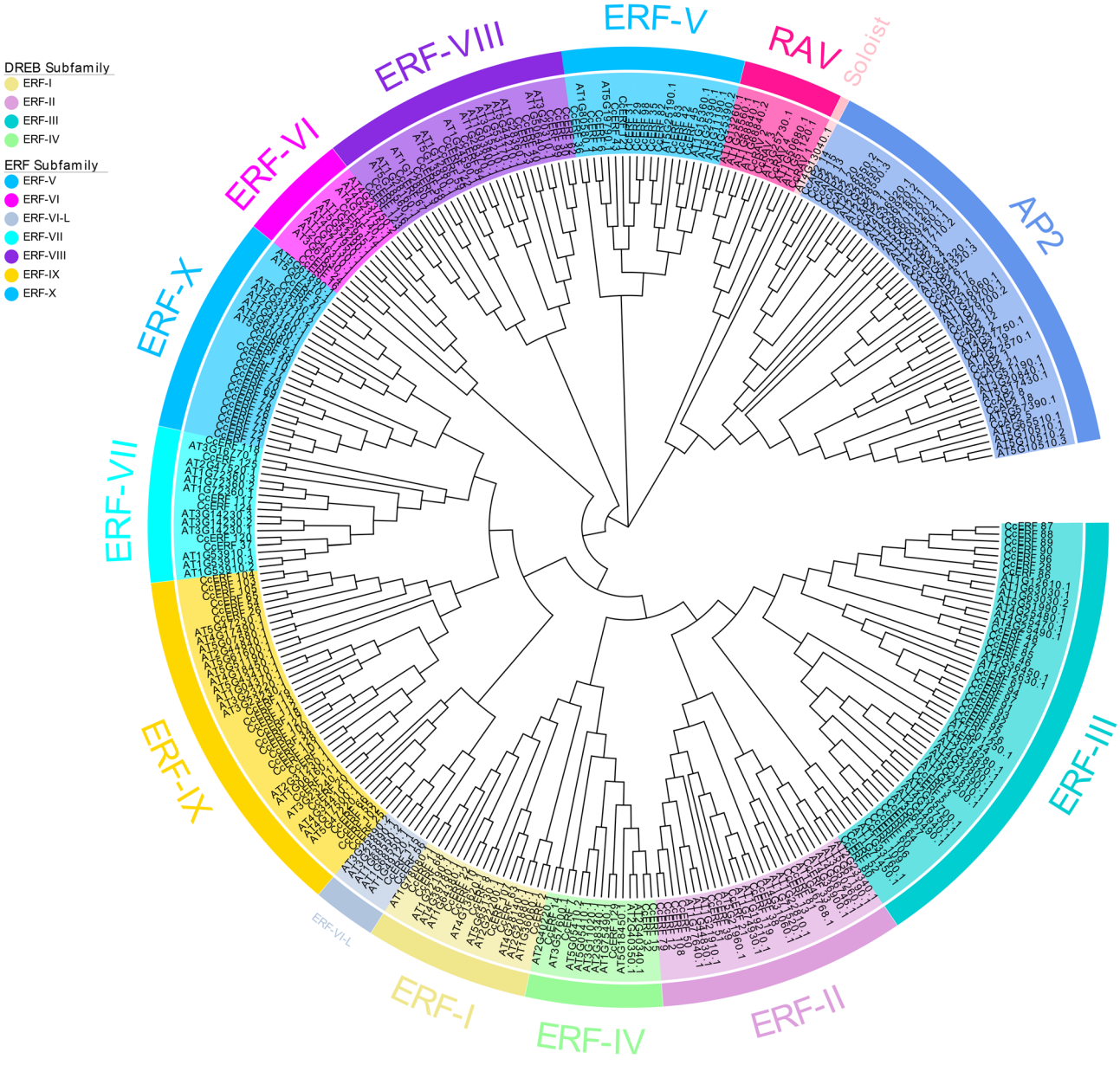

Fig. 1 Phylogenetic relationships among AP2/ERFs proteins from the *C. camphora* and the *A. thaliana* genome were analyzed using the neighbor-joining method. Each subgroup is represented by a distinct color with the name labeled in the outer circle

The remainder all belong to the ERF subfamily. The presence of synteny predominantly within the same subfamily suggests a higher likelihood of shared biological functions among genes within the same subfamily.

To explore the evolutionary connections of *CcAP2/ERF* across species and to assess the conservation and divergence of homologous genes, we generated collinearity diagrams comparing *C.camphor* with *A. thaliana*, *Camphora kanahirae*, *Solanum lycopersicum* and *Oryza sativa* (Fig. 6). We identified 120, 133, 134, and 224 pairs of orthologous genes between *C.camphor* and *A. thaliana*, *O. sativa*, *S. lycopersicum* and *C. kanahirae*, respectively. Based on the highest number of syntenic genes, the closest relationship between *C.camphor* and *C. kanahirae* has been confirmed, which aligns with the hypothesis that closely related species tend to share similar chemical compounds and genetic backgrounds. There were 29 *AP2/ERF* genes from *C.camphor* that had no collinear gene pairs with the *AP2/ERF* genes from *C. kanahirae*, these differences may have been affected by the third internal replication event of Laurales.

The cis-regulatory element analysis in the promoters of CcAP2/ERFs

Cis-regulatory elements in promoters (CREs) play a vital role in modulating both the expression patterns and

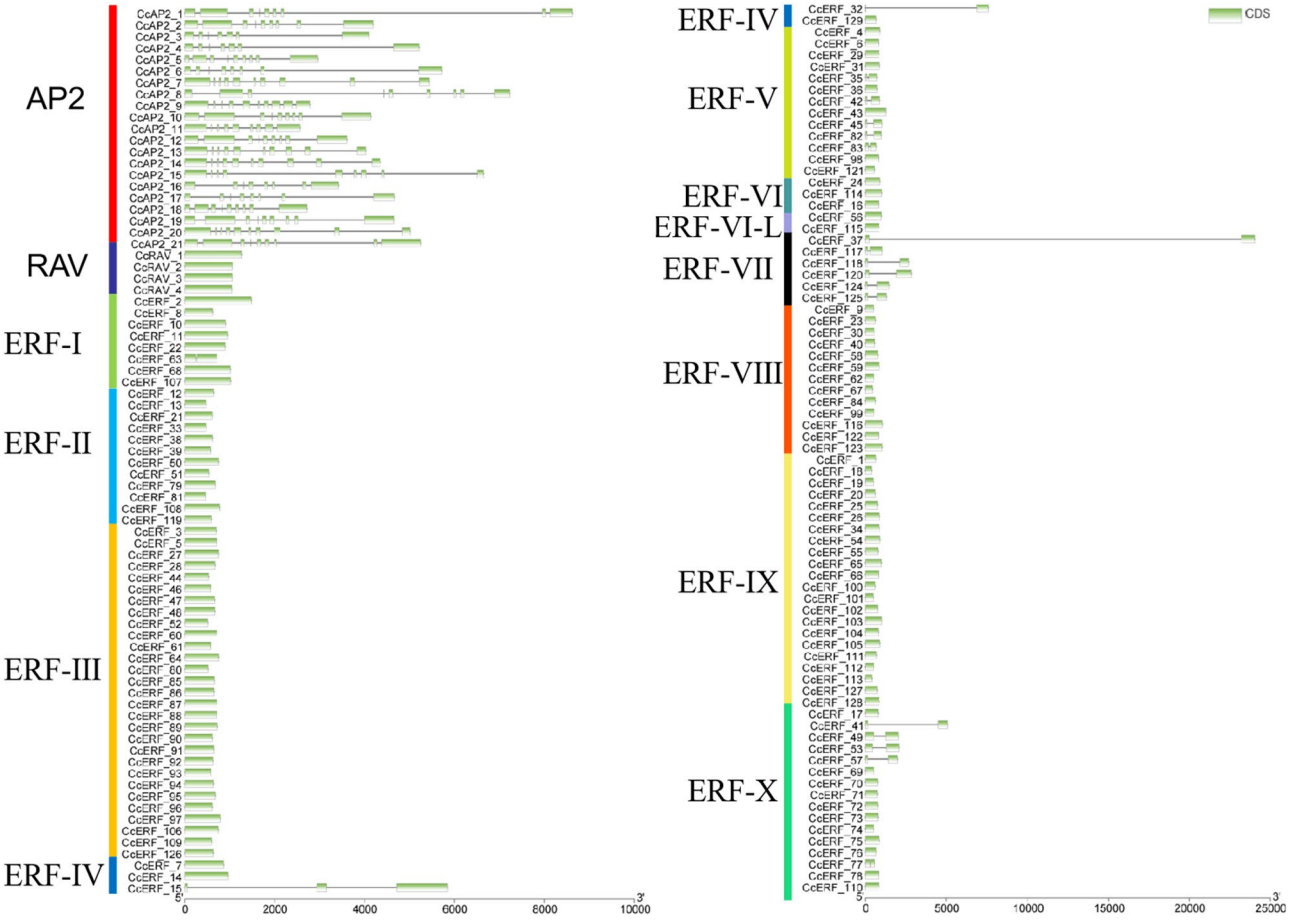

Fig. 2 Gene structure of *CcAP2/ERFs* family members

functions of genes. To identify these CREs, the 2000 bp upstream sequences of all 154 *CcAP2/ERF* genes were analyzed. A total of 5,991 CREs were identified in the promoters of *CcAP2/ERF* genes and were categorized into three main groups: stress responsiveness elements, plant growth and development, and phytohormone responsiveness (Fig. 7, Table S6). Among these groups, the stress responsiveness elements constituted the largest category, with 3,036 elements. Within this group, three elements, including a stress response element (STRE), a light-responsive elements (G-box), an anoxia response element (ARE)-were particularly prominent, contributing 486 (16.00%), 452 (14.89%), and 331 (10.90%) elements, respectively. In addition, the phytohormone responsiveness elements constituted the second largest group, comprising a total of 2,195 elements. Among these, MeJA-responsive elements, including the MYC element, CGTCA-motif and TGACG-motif, contributed 1,199 elements (54.62%), while the abscisic acid-responsive element (ABRE) accounted for 410 elements (18.67%). The smallest group consisted of plant growth and development elements, totaling 760 elements, among which AAGAA-motif were the most abundant, accounting for 187 elements (24.60%).

Transcriptome-based and real-time PCR investigation on the expression patterns of CcAP2/ERFs in various tissues

Expression profiles can provide clues to the functional differences among all members of a gene family. To compare the expression profiles of *AP2/ERF* genes across various tissues in *C.camphor*, the expression patterns of 154 *CcAP2/ERF* genes were examined using transcriptomic data from different organs of *C.camphor*, including roots, stems, and leaves. In this research, the FPKM value was computed for each *CcAP2/ERF* genes to reflect its expression level in various organs. The findings revealed that the majority of *CcAP2/ERF* genes exhibited higher expression in roots compared to other tissues. A notable trend in the expression of certain genes across different organs was observed (Fig. 8, Table S7). Out of all the genes, 145 were expressed in at least one tissue, while 9 genes, including *CcERF_51*, *CcERF_21*, *CcERF_122*, *CcERF_121*, *CcERF_77*, *CcERF_76*, *CcERF_71*, *CcERF_69* and *CcERF_59*, showed no expression (FPKM = 0).

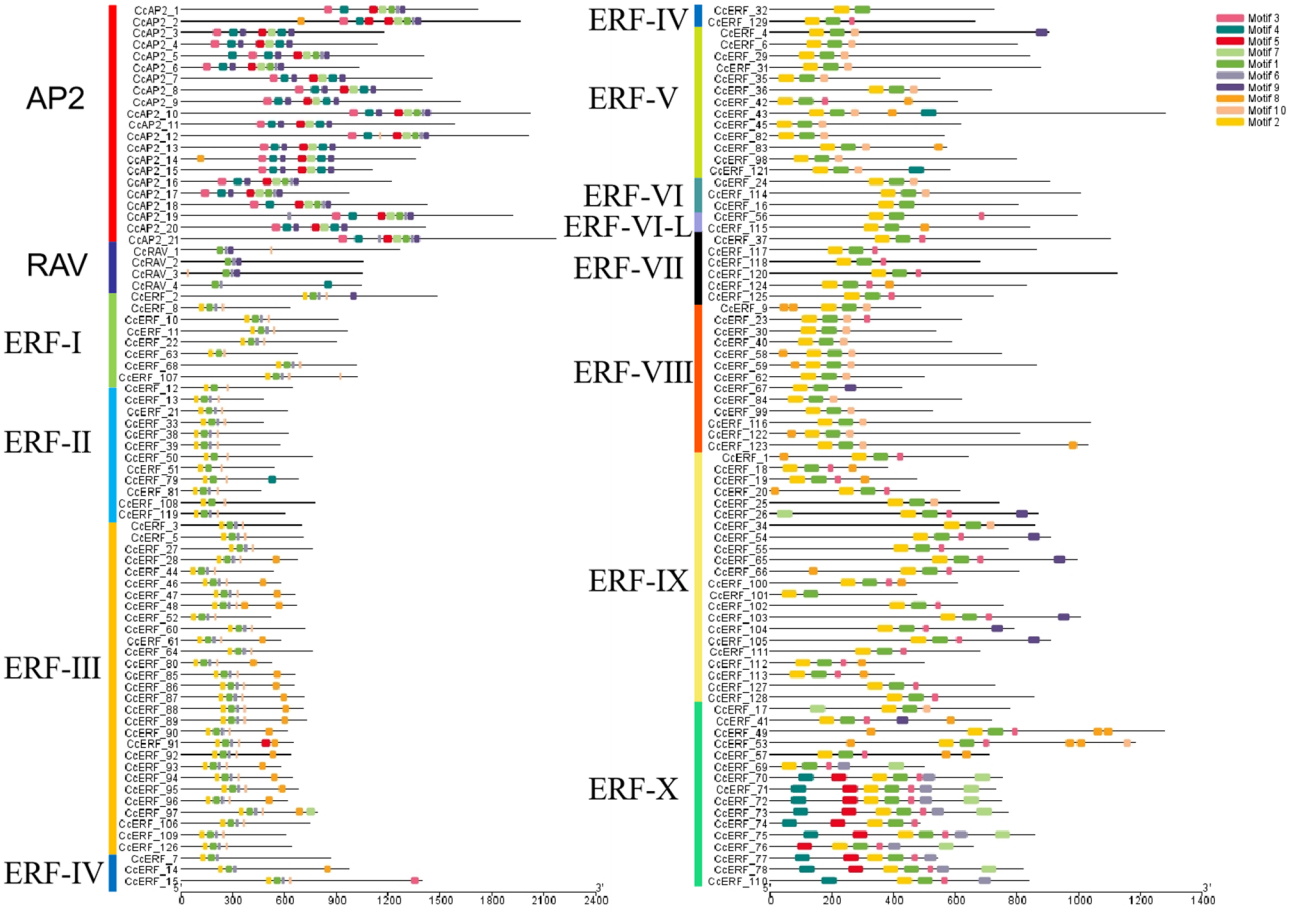

Fig. 3 Conserved motif of the AP2/ERF family in the *C. camphora* genome

In our previous studies, we have identified nine TPS genes (*CcTPS23*, *CcTPS26*, *CcTPS32*, *CcTPS44*, *CcTPS49*, *CcTPS54*, *CcTPS59*, *CcTPS72* and *CcTPS74*) related to the biosynthesis of most of the monoterpenes in the camphor tree. We subsequently computed the correlation coefficients (R-values) between the expression patterns of *CcAP2/ERF* genes and nine *CcTPS* genes in different tissues(Table S8, Fig. 9). We found that a total of 46 *CcAP2/ERFs* exhibited strong correlations (|R|>0.9, $P < 0.05$) with the six TPS genes (*CcTPS23*, *CcTPS26*, *CcTPS44*, *CcTPS49*, *CcTPS72* and *CcTPS74*), among which 32 were positively correlated, 14 were negatively correlated. Notably, borneol is one of the most valuable natural products in camphor trees, and its biosynthesis is primarily catalyzed by borneol diphosphate synthase (BPPS) [1]. Strikingly, we identified *CcERF_104* as a high-potential hub regulator, which was co-expressed with two key BPPS genes (*CcTPS26* and *CcTPS49*). Moreover, a suite of regulators (including *CcERF_27*, *CcERF_42*, *CcERF_83*, *CcERF_87*, *CcAP2_12*, and *CcAP2_13*) showed a strong positive correlation with *CcTPS72*, the most catalytically efficient BPPS synthase in *C. camphora*. Furthermore, we found that 2 *CcAP2/ERF* genes were positively correlated with *CcTPS74*, 4 *CcAP2/ERF* genes with *CcTPS44*, 1 *CcAP2/ERF* gene with *CcTPS23*. To further validate the accuracy of the transcriptome profiles, we conducted qRT-PCR analysis to quantify the expression of 9 *CcAP2/ERFs* in different tissues of *C.camphor*. The results showed that the qPCR results were consistent with the transcriptome data (Fig. 10).

Discussion

Therapeutically significant compounds in pharmaceuticals predominantly originate from natural secondary metabolites within medicinal plants [33, 34]. In *C. camphora*, terpenoids represent principal bioactive constituents, exhibiting high abundance in leaf tissues [35]. Consequently, enhancing terpenoid content in *C. camphora* leaves through genetic engineering could substantially increase the plant's commercial value by optimizing utilization of this biomass resource. The AP2/ERF transcription factor family, which is extensively present in plants, plays crucial roles in regulating secondary metabolite biosynthesis, including terpenoids [36]. However, the specific functions and regulatory mechanisms of AP2/ERF transcription factors in *C. camphora* remain

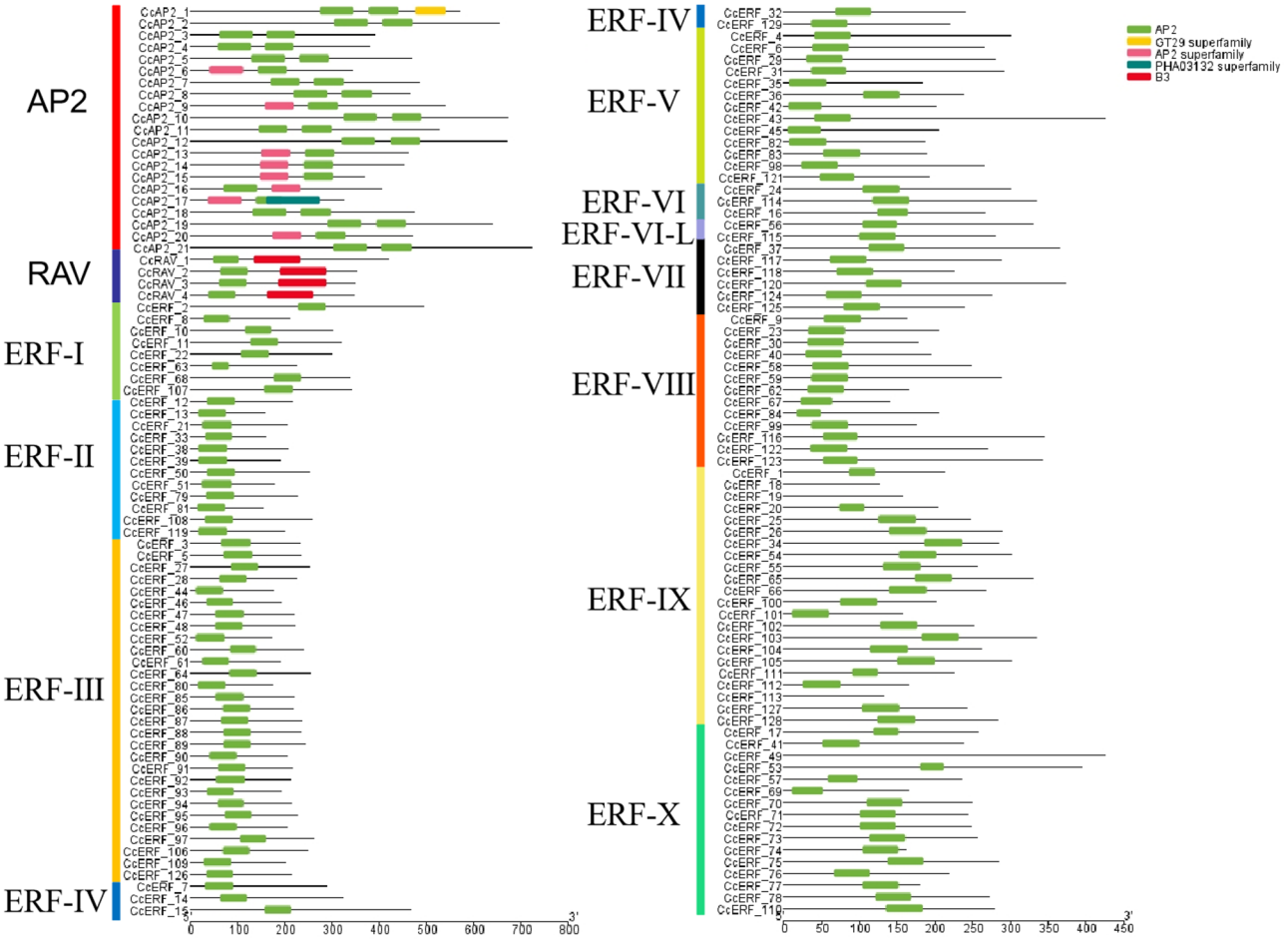

Fig. 4 Conserved domains of the AP2/ERF gene family in *C. camphora*

largely unexplored. Our present study provides a comprehensive analysis of the AP2/ERF family in *C. camphora* and explores their potential roles in terpene biosynthesis.

Extensive research on plant genomes has identified members of the AP2/ERF gene family in diverse species, including *Raphanus sativus* [37], *Zingiber officinale* [38], *Arachis hypogaea* [39], *Oryza sativa* [40], *Saccharum spontaneum* [41], *Zea mays* [42], and *Hordeum vulgare* [43]. In this study, we identified 154 unique *AP2/ERF* genes in the *C. camphora* genome. These genes were classified into four subfamilies (RAV, AP2, ERF, and DREB) based on their structural domain profiles and sequence homologies. The ERF subfamily was the largest, comprising 75 members, further subdivided into seven distinct subgroups. This classification aligns with findings in other plants, such as *Dendrobium catenatum* [44] and *Camellia sinensis* [45], where the ERF subfamily is also predominant. Also, the proportional distribution of subfamily genes within the AP2/ERF superfamily remains conserved across plant species, suggesting evolutionary stability in gene family composition. Furthermore, the gene structure, motif composition, conserved domains of CcAP2/ERFs within each subfamily is quite similar, suggesting that these proteins are likely to share similar functions and this observation is consistent with the outcomes of the phylogenetic analysis.

Cis-acting elements within gene promoters typically modulate gene expression levels, and their analysis can offer insights into gene regulation and stress response mechanisms [46]. To investigate the potential regulatory mechanisms of *CcAP2/ERF* genes, the cis-acting elements within their promoter regions were analyzed. It is noteworthy that the most prevalent cis-acting element among the promotor of 154 C*cAP2/ERF* genes is the MeJA-responsive elements (including the MYC element, CGTCA-motif and TGACG-motif), this finding aligns with our previous observations regarding trans-isoprenoid diphosphate synthases (CcTIDS) and terpene synthase in the *C. camphora*, which are involved in the biosynthesis of terpeneoid. We speculate that the AP2/ERF transcription factors in camphor trees can respond to MeJA signaling to regulate the biosynthesis of terpenoid secondary metabolites. Furthermore, the high number of stress responsiveness elements suggests that these

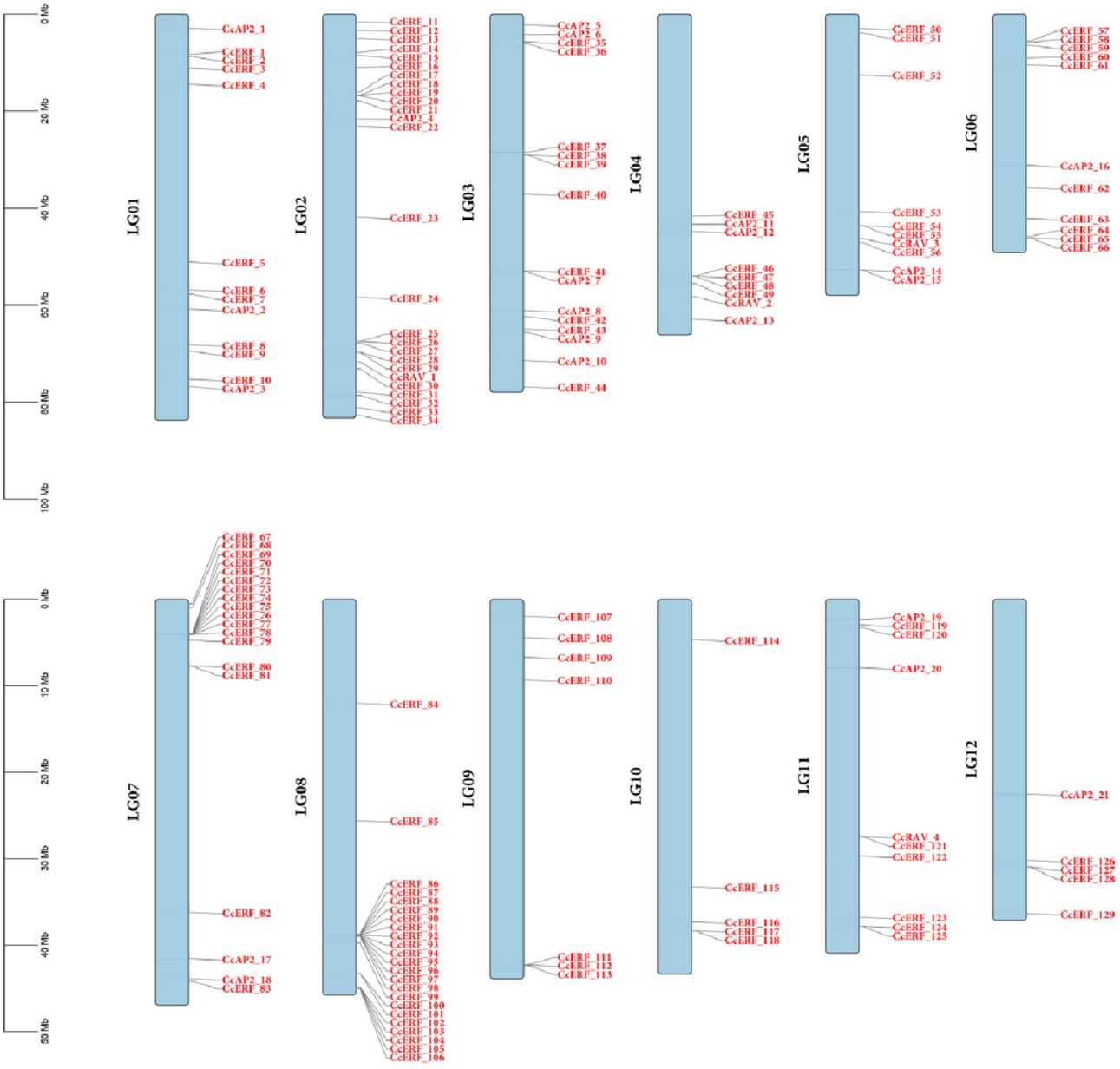

Fig. 5 Chromosomal localization of *AP2/ERF* genes from *C. camphora*

transcription factors may play significant roles in stress responses, consistent with their known functions in other plants [47].

To elucidate the potential regulatory functions of AP2/ERF genes in *C. camphora*, we examined their expression profiles in various tissues. The expression pattern analysis using transcriptomic data revealed that most *CcAP2/ERF* genes were highly expressed in roots, indicating tissue-specific functions. Given the established link between transcription factor expression and key metabolic enzyme genes [17, 48–51], we performed correlation analysis between *CcAP2/ERF* genes and *CcTPS* genes involved in monoterpene biosynthesis. This analysis identified 46 *CcAP2/ERFs* exhibited strong correlations with six TPS genes (*CcTPS23*, *CcTPS26*, *CcTPS44*, *CcTPS49*, *CcTPS72* and *CcTPS74*); among these, 32 showed positive correlations and 14 showed negative correlations. This suggests that these transcription factors predominantly have a positive impact on terpene biosynthesis. Similar findings have been reported in other plants, where AP2/ERF genes have been shown to regulate the expression of genes involved in terpene biosynthesis [52]. It is particularly noteworthy that the robust co-expression patterns observed between key regulators-such as *CcERF_104* and the borneol diphosphate synthase genes *CcTPS26/CcTPS49*, as well as between

Fig. 6 Collinearity analysis of the *AP2/ERF* TF family between *C. camphora* and *A. thaliana*, *C. camphora* and *S. lycopersicum*, *C. camphora* and *O. sativa*, *C. camphora* and *C. kanahirae*. Red lines link identified collinear genes

multiple transcription factors (e.g., *CcERF_27* and *CcAP2_42*) and the highly efficient BPPS gene *CcTPS72*-collectively provide compelling evidence for a sophisticated transcriptional regulatory network. This finding carries significant implications, given that borneol is one of the most valuable natural products in camphor trees, and its biosynthesis is directly governed by these BPPS enzymes. Future studies will focus on experimentally validating these regulatory hypotheses through the functional characterization of key *CcAP2/ERF* genes, using integrated genetic and biochemical approaches, to ultimately provide a genetic basis for enhancing terpenoid production in *C. camphora* via genetic engineering.

Conclusions

In summary, this study provides a comprehensive analysis of the AP2/ERF transcription factor family in *Cinnamomum camphora* and elucidates their potential roles in regulating terpene biosynthesis. We identified and classified 154 unique *AP2/ERF* genes into four subfamilies (RAV, AP2, ERF, and DREB) based on structural domain profiles and sequence homologies. The presence of MeJA-responsive elements and stress responsiveness elements in the promoters of these genes suggests their involvement in regulating terpene biosynthesis and stress responses. Expression profile analysis revealed tissue-specific expression patterns, with most *CcAP2/ERF* genes highly expressed in roots. Correlation analysis showed that 44 *CcAP2/ERF* genes were strongly correlated with six terpene synthase genes (*CcTPS23*, *CcTPS26*, *CcTPS44* *CcTPS49*, *CcTPS72*, and *CcTPS74*), indicating their potential regulatory roles in monoterpene biosynthesis. Crucially, our study pinpointed several high-potential transcriptional regulators, such as *CcERF_104*, *CcERF_27* and *CcAP2_42*, showing strong co-expression with key borneol diphosphate synthase (BPPS) genes (*CcTPS26*, *CcTPS49*, *CcTPS72*). This finding is of particular importance as it directly links these AP2/ERF members to the biosynthesis of borneol, the most economically valuable compound in *C. camphora*, offering precise targets for molecular breeding. Future research will focus on functional characterization of key *CcAP2/ERF* genes using genetic and biochemical approaches, aiming to enhance terpenoid content in *C. camphora* leaves through genetic improvement technologies.

Statements and Declarations.

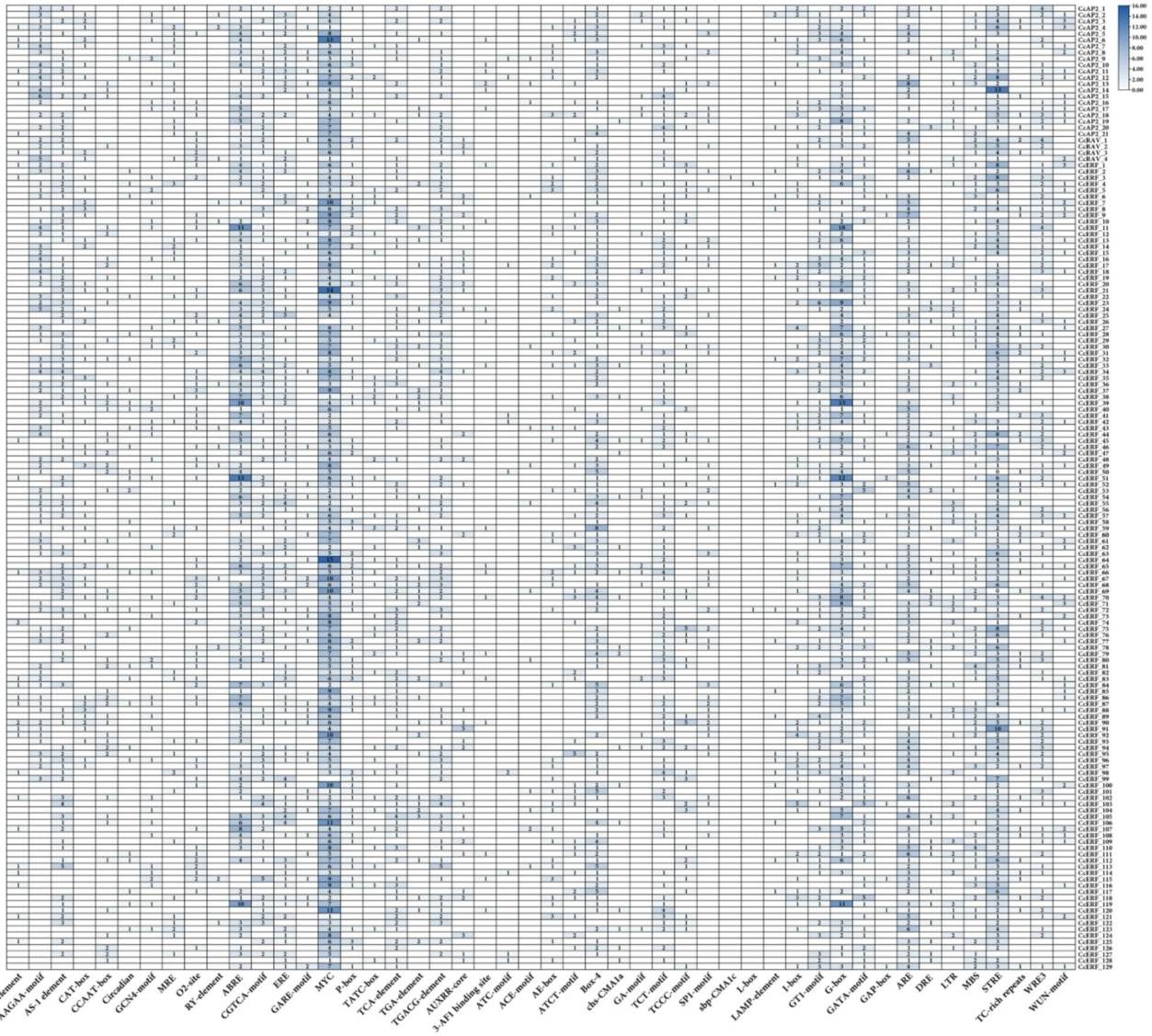

Fig. 7 Cis-acting elements of *CcAP2/ERF* promoters

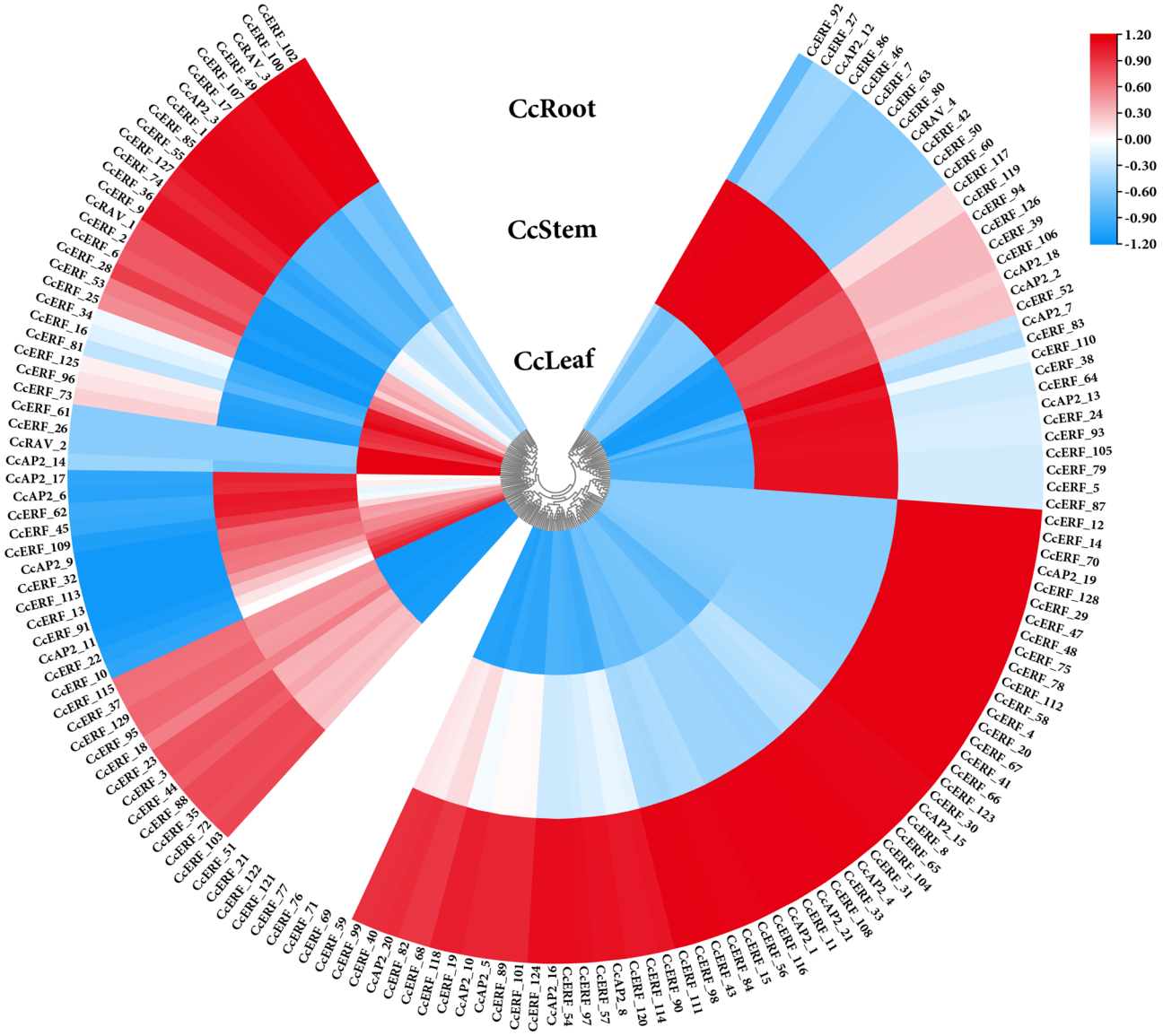

Fig. 8 Expression profiles of *CcAP2/ERE* genes in various tissues

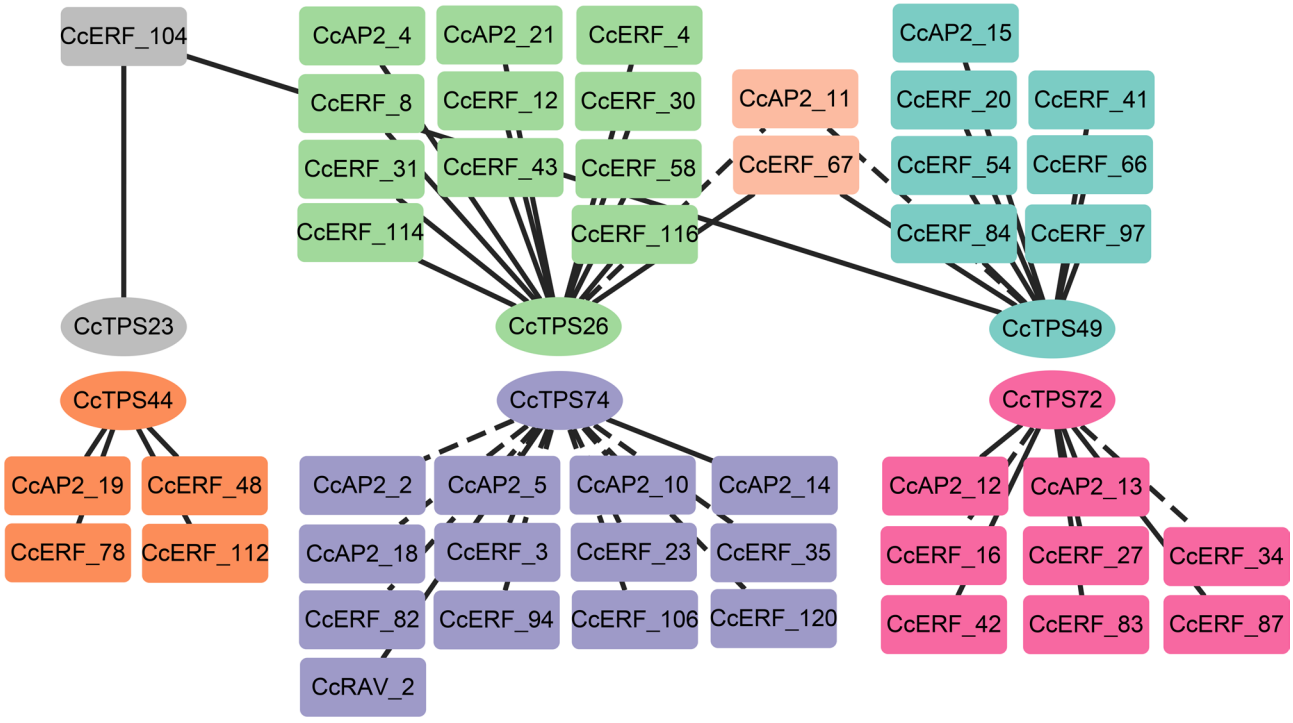

Fig. 9 Network diagram of the correlations between the expression patterns of *CcAP2/ERF* genes and nine *CcTPS* genes. The solid lines represent positive regulation, while the dashed lines represent negative regulation

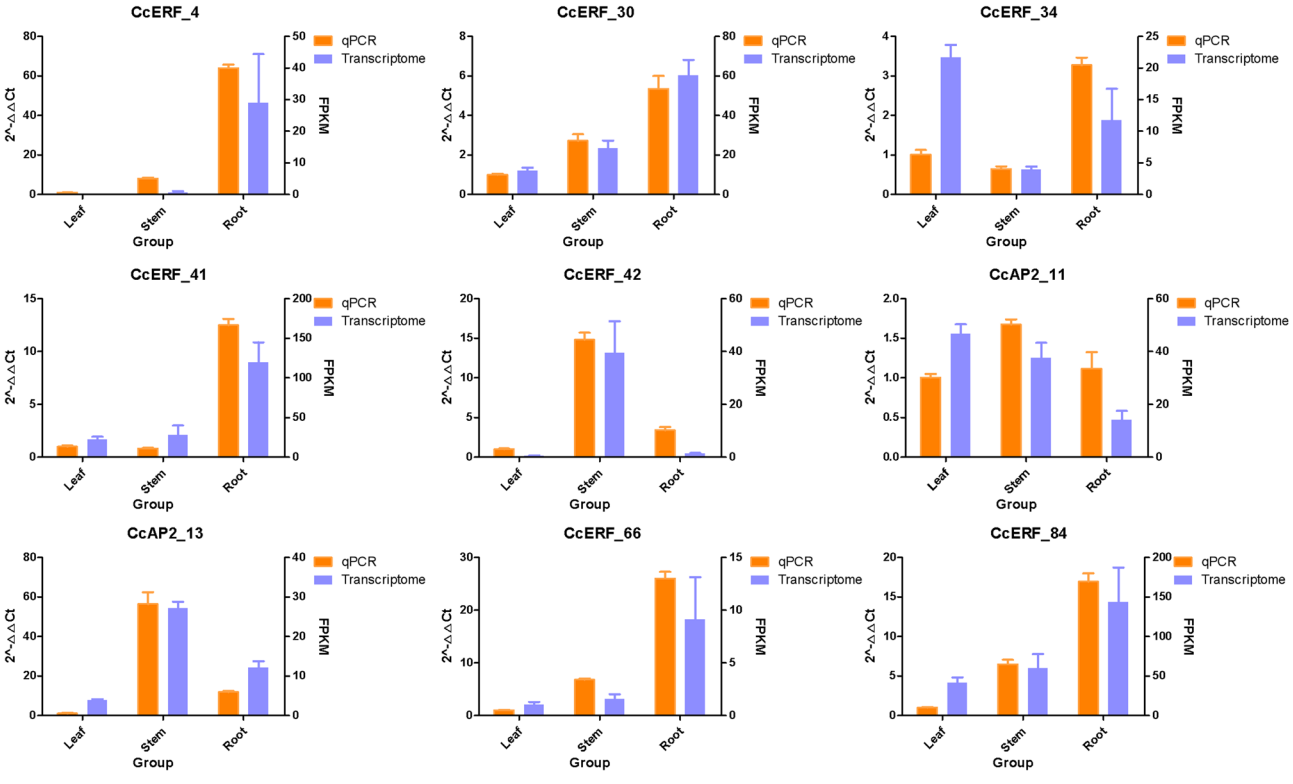

Fig. 10 Real-time PCR analysis was conducted on 9 *CcAP2/ERF* genes from transcriptomic data. The fold changes determined by qPCR are shown on the left vertical axis, whereas those calculated based on FPKM values are displayed on the right vertical axis. Error bars represent the standard error of the mean

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-025-12169-6.

Supplementary material 1.

Acknowledgements
Not applicable.

Authors' contributions
ZR Y performed most of the experiments, analyzed the data, prepared the figures and tables and wrote the manuscript. XN Y participated in analyzing the data, preparing the figures and performing the experiments. ZR Y and XS Z conceived and designed the experiments and revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding
This work was supported by grants from the National Natural Science Foundation of China (No.82304657), Guangzhou Science and Technology Plan Project (2024A04J4895), National Key Specialty Construction Project of Clinical Pharmacy, High Level Clinical Key Specialty of Clinical Pharmacy in Guangdong Province.

Data availability
The assembly of the borneol-type camphor tree genome was deposited on the Global Pharmacopoeia Genome Database (http://www.gpgenome.com/species/206). All RNA-Seq raw data has been submitted to CNGBdb: CNP0006412 (https://db.cngb.org/search/project/CNP0006412/) and in NCBI: PRJNA1189175 (https://dataview.ncbi.nlm.nih.gov/object/PRJNA1189175).

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 31 May 2025 / Accepted: 30 September 2025
Published online: 30 October 2025

References

1. Yang Z, Zhan T, Xie C, et al. Genome-wide analyzation and functional characterization on the TPS family provide insight into the biosynthesis of mono-terpenes in the Camphor tree. Plant Physiol Biochem. 2023;196:55–64.
2. Xiao S, Yu H, Xie Y, et al. The anti-inflammatory potential of *Cinnamomum Camphora* (L.) J.Presl essential oil in vitro and in vivo. J Ethnopharmacol. 2021;267:113516.
3. Chen C, Zheng Y, Zhong Y, et al. Transcriptome analysis and identification of genes related to terpenoid biosynthesis in *Cinnamomum camphora*. BMC Genomics. 2018;19(1):550.
4. Guo X, Cui M, Deng M, et al. Molecular differentiation of five *Cinnamomum camphora* chemotypes using desorption atmospheric pressure chemical ionization mass spectrometry of raw leaves. Sci Rep. 2017;7:46579.
5. Yang Z, An W, Liu S, et al. Mining of candidate genes involved in the biosynthesis of dextrorotatory Borneol in *Cinnamomum burmannii* by transcriptomic analysis on three chemotypes. PeerJ. 2020;8:e9311.
6. Huang J, Tang X, Ye F, et al. Clinical therapeutic effects of aspirin in combination with Fufang Danshen Diwan, a traditional Chinese medicine formula, on coronary heart disease: a systematic review and meta-analysis. Cell Physiol Biochem. 2016;39(5):1955–63.
7. Liang CJ, Li JH, Zhang Z, et al. Suppression of MIF-induced neuronal apoptosis may underlie the therapeutic effects of effective components of Fufang Danshen in the treatment of alzheimer's disease[J]. Acta Pharmacol Sin. 2018;39(9):1421–38.
8. Ren L, Wang J. Efficacy of suxiao jiuxin pill on coronary heart disease: a meta-analysis of randomized controlled trials. Evid Based Complement Alternat Med. 2018;2018:9745804.
9. Chai Y, Yin Z, Fan Q. Protective effects of Angong Niuhuang Pill on early atherosclerosis in ApoE(-/-) mice by reducing the inflammatory response. Evid Based Complement Alternat Med. 2019;2019:9747212.
10. Yang Z, Xie C, Zhan T, et al. Genome-wide identification and functional characterization of the trans-isopentenyl diphosphate synthases gene family in *Cinnamomum camphora*. Front Plant Sci. 2021;12:708697.
11. Tan ZW, Lu DD, Yu YL, et al. Genome-wide analysis of the APETALA2/Ethylene-Responsive factor gene family in *Carthamus tinctorius* L. Plant Direct. 2025;9(1):e70032.
12. Hu F, Zhang Y, Guo J. Identification and characterization of AP2/ERF transcription factors in yellow Horn[J]. Int J Mol Sci. 2022;23(23):14199.
13. Zhang D, Zeng B, He Y, et al. Genome-wide identification and comparative analysis of the AP2/ERF gene family in *Prunus dulcis* and *Prunus tenella*: expression of PdAP2/ERF genes under freezing stress during dormancy. BMC Genomics. 2025;26(1):95.
14. Zhang J, Liao J, Ling Q, et al. Genome-wide identification and expression profiling analysis of maize AP2/ERF superfamily genes reveal essential roles in abiotic stress tolerance. BMC Genomics. 2022;23(1):125.
15. Nakano T, Suzuki K, Fujimura T, et al. Genome-wide analysis of the ERF gene family in Arabidopsis and rice[J]. Plant Physiol. 2006;140(2):411–32.
16. Zhang M, Lu P, Zheng Y, et al. Genome-wide identification of AP2/ERF gene family in *Coptis chinensis* Franch reveals its role in tissue-specific accumulation of benzylisoquinoline alkaloids. BMC Genomics. 2024;25(1):972.
17. Liu X, Zhang W. Genomic-wide identification and expression analysis of AP2/ERF transcription factors in *Zanthoxylum armatum* reveals the candidate genes for the biosynthesis of terpenoids[J]. Plant Genome. 2024;17(1):e20422.
18. Ji AJ, Luo HM, Xu ZC et al. Genome-Wide identification of the AP2/ERF gene family involved in active constituent biosynthesis in *Salvia miltiorrhiza*[J]. Plant Genome. 2016;9(2):1–11.
19. Li X, Xu Y, Shen S, et al. Transcription factor CitERF71 activates the terpene synthase gene CitTPS16 involved in the synthesis of E-geraniol in sweet orange fruit. J Exp Bot. 2017;68(17):4929–38.
20. Wang M, Gao M, Zhao Y, et al. LcERF19, an AP2/ERF transcription factor from *Litsea cubeba*, positively regulates geranial and Neral biosynthesis[J]. Hortic Res. 2022;9:uhac093.
21. Li S, Wang H, Li F, et al. The maize transcription factor EREB58 mediates the jasmonate-induced production of sesquiterpene volatiles. Plant J. 2015;84(2):296–308.
22. Chen C, Wu Y, Li J, et al. TBtools-II: a one for all, all for one bioinformatics platform for biological big-data mining. Mol Plant. 2023;16(11):1733–42.
23. Duvaud S, Gabella C. Expasy, the Swiss bioinformatics resource Portal, as designed by its users[J]. Nucleic Acids Res. 2021;49(W1):W216–27.
24. Nielsen H. Practical applications of Language models in protein sorting prediction: signalp 6.0, deeploc 2.1, and deeplocPro 1.0[J]. Methods Mol Biol. 2025;2941:153–75.
25. Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. Mol Biol Evol. 2021;38(7):3022–7.
26. Subramanian B, Gao S, Lercher MJ, et al. Evolview v3: a webserver for visualization, annotation, and management of phylogenetic trees[J]. Nucleic Acids Res. 2019;47(W1):W270–5.
27. Bailey TL, Boden M, Buske FA et al. MEME SUITE: tools for motif discovery and searching[J]. Nucleic Acids Res, 2009, 37(Web Server issue): W202–208.
28. Wang Y, Tang H, Debarry JD, et al. MCscanx: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40(7):e49.
29. Lescot M, Déhais P, Thijs G, et al. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in Silico analysis of promoter sequences[J]. Nucleic Acids Res. 2002;30(1):325–7.
30. Zerui Yang XY, Xlasheng Z. Cloning and expression analysis of the phosphatidic acid phosphatase genesfrom *Cinnamomum camphora*[J]. Seed. 2025;44(6):6.
31. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498–504.

32. Liu J, Geng Z, Zhao G, et al. Genome-wide analysis of the FKBP gene family and the potential role of *GhFKBP* 13 in chloroplast biogenesis in upland cotton. BMC Genomics. 2025;26(1):125.
33. Parthasarathy R, Shanmuganathan R, Pugazhendhi A. Vinblastine production by the endophytic fungus *Curvularia verruculosa* from the leaves of catharanthus roseus and its in vitro cytotoxicity against HeLa cell line. Anal Biochem. 2020;593:113530.
34. Resetar M, Liu X. Polyacetylenes from *Oplopanax horridus* and *Panax ginseng*: relationship between structure and PPARγ activation. J Nat Prod. 2020;83(4):918–26.
35. Lee SH, Kim DS, Park SH, et al. Phytochemistry and applications of *Cinnamomum Camphora* essential Oils[J]. Molecules. 2022;27(9):2695.
36. Ma Z, Hu L, Jiang W. Understanding AP2/ERF transcription factor responses and tolerance to various abiotic stresses in plants: a comprehensive review. Int J Mol Sci. 2024. https://doi.org/10.3390/ijms25020893.
37. Karanja BK, Xu L, Wang Y, et al. Genome-wide characterization of the AP2/ERF gene family in radish (*Raphanus sativus* L.): unveiling evolution and patterns in response to abiotic stresses. Gene. 2019;718:144048.
38. Xing H, Jiang Y, Zou Y, et al. Genome-wide investigation of the AP2/ERF gene family in ginger: evolution and expression profiling during development and abiotic stresses. BMC Plant Biol. 2021;21(1):561.
39. Cui M, Haider MS, Chai P, et al. Genome-wide identification and expression analysis of AP2/ERF transcription factor related to drought stress in cultivated peanut (*Arachis hypogaea* L.). Front Genet. 2021;12:750761.
40. Xie W, Ding C, Hu H, et al. Molecular events of rice AP2/ERF transcription factors. Int J Mol Sci. 2022. https://doi.org/10.3390/ijms231912013.
41. Li P, Chai Z, Lin P, et al. Genome-wide identification and expression analysis of AP2/ERF transcription factors in sugarcane (*Saccharum spontaneum* L.). BMC Genomics. 2020;21(1):685.
42. Cheng C, An L, Li F, et al. Wide-range portrayal of AP2/ERF transcription factor family in maize (*Zea mays* L.) development and stress responses. Genes. 2023;14(1):194.
43. Guo B, Wei Y, Xu R, et al. Genome-Wide analysis of APETALA2/Ethylene-Responsive factor (AP2/ERF) gene family in barley (*Hordeum vulgare* L.)[J]. PLoS ONE. 2016;11(9):e0161322.
44. Han Y, Cai M, Zhang S, et al. Genome-wide identification of AP2/ERF transcription factor family and functional analysis of DcAP2/ERF#96 associated with abiotic stress in *Dendrobium catenatum*. Int J Mol Sci. 2022. https://doi.org/10.3390/ijms232113603.
45. Liu Y, Chen S, Chen J, et al. Comprehensive analysis and expression profiles of the AP2/ERF gene family during spring bud break in tea plant (*Camellia sinensis*). BMC Plant Biol. 2023;23(1):206.
46. Xie Z, Yang L, Fan M, et al. Genome-wide identification, characterization and expression analysis of the chalcone synthase gene family in Chinese cabbage. BMC Genomics. 2025;26(1):168.
47. Feng K, Hou XL, Xing GM, et al. Advances in AP2/ERF super-family transcription factors in plant[J]. Crit Rev Biotechnol. 2020;40(6):750–76.
48. Huang H, Qi Q, Yang Z, et al. Warm temperatures regulate Andrographolide biosynthesis through jasmonic acid signaling mediated by the ApMYC2-ApHSFB2b module in *Andrographis paniculata*. Plant J. 2025;123(3):e70419.
49. Li C, Jiao M, Zhao X, et al. bZIP transcription factor responds to changes in light quality and affects saponins synthesis in *Eleutherococcus senticosus*[J]. Int J Biol Macromol. 2024;279(Pt 2):135273.
50. Jiang T, Zhang Y. Transcription factor PgNAC72 activates DAMMARENEDIOL SYNTHASE expression to promote ginseng saponin biosynthesis. Plant Physiol. 2024;195(4):2952–69.
51. Chuang YC, Hung YC, Hsu CY, et al. A dual repeat cis-element determines expression of GERANYL DIPHOSPHATE SYNTHASE for monoterpene production in *Phalaenopsis Orchids*. Front Plant Sci. 2018;9:765.
52. Zhang Y, Ji A, Xu Z, et al. The AP2/ERF transcription factor SmERF128 positively regulates diterpenoid biosynthesis in *Salvia miltiorrhiza*[J]. Plant Mol Biol. 2019;100(1–2):83–93.

Publisher's Note