Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Genome-wide Identification and ABA-responsive Characterization of Calmodulin and Calmodulin-like Genes in Salvia miltiorrhiza

Yansong Zhang
Universities of Shaanxi Province
Yuanchu Liu
lyc09172022@163.com

Universities of Shaanxi Province
Yuan Yang
Xi'an International University

Article

Keywords: Calmodulin, calmodulin-like proteins, Salvia miltiorrhiza, phylogenetic analysis, promoter cis-elements, ABA response

Posted Date: September 11th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-7478111/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Calcium ions (Ca^{2+}) are universal secondary messengers that regulate plant growth, development, and responses to environmental stresses. Calmodulin (CaM) and calmodulin-like (CML) proteins are key calcium sensors, yet their roles remain poorly explored in Salvia miltiorrhiza, a traditional medicinal plant. Here, we performed a genome-wide analysis to identify and characterize the CaM and CML gene families in S. miltiorrhiza. We identified six SmCaM genes and twenty-six SmCML genes, revealing conserved EF-hand motifs in SmCaMs, whereas SmCMLs presented significant variability in protein length, domain composition, and gene structure. Phylogenetic analysis classified these proteins into eight subfamilies, suggesting functional divergence. Promoter analysis revealed abundant cis-elements related to light, hormone, and stress responses. Chromosomal mapping indicated nonrandom localization, with significant enrichment in gene-rich regions. Transcriptomic profiling under abscisic acid (ABA) treatment highlighted the stable expression of SmCaMs, whereas SmCMLs presented dynamic, tissue-specific responses. Notably, SmCML19 exhibited root-specific downregulation under ABA stress despite a lack of canonical ABA-responsive elements. These findings provide a comprehensive foundation for understanding the calcium signaling networks in S. miltiorrhiza and may facilitate future studies aiming to increase stress tolerance and secondary metabolite production.

1 Introduction

Calcium ions (Ca^{2+}) act as universal secondary messengers in plant signaling networks, modulating numerous physiological processes, including growth, development, and responses to environmental cues[1]. Upon stimulation, the transient increase in cytosolic Ca^{2+} levels is sensed by a variety of Ca^{2+} sensor proteins[2]. Among these proteins, calmodulin (CaM) and calmodulin-like proteins (CMLs) constitute a large family of EF-hand-containing Ca^{2+} sensors that transmit calcium signals by interacting with specific downstream targets[3].

CaMs are highly conserved across eukaryotes and typically contain four EF-hand domains, whereas CMLs are more variable in size and structure, often harboring different numbers and arrangements of EF-hands[4]. Functional studies have demonstrated the involvement of CaM/CMLs in diverse biological processes, such as plant immunity, ion transport, pollen development, and hormone signaling, especially abscisic acid (ABA) responses[5]. With the advent of plant genome sequencing, genome-wide identification and functional analysis of the CaM/CML gene family have been conducted in multiple species, including *Arabidopsis thaliana, Oryza sativa, Glycine max, and Triticum aestivum* L.[6, 7, 8, 9].

Salvia miltiorrhiza Bunge, commonly known as Danshen, is a model medicinal plant used in traditional Chinese medicine (TCM)[10]. Its bioactive compounds, including phenolic acids and diterpenoid tanshinones, are widely used in cardiovascular and anti-inflammatory treatments[11]. Emerging evidence suggests that calcium signaling plays a role in the biosynthesis and regulation of secondary metabolites in medicinal plants[12]. However, the CaM/CML gene families in *S. miltiorrhiza* have not yet been systematically characterized.

In this study, we performed a comprehensive genome-wide identification and characterization of the CaM and CML gene families in *S. miltiorrhiza*. We analyzed the gene structures, conserved motifs, phylogenetic relationships, chromosomal distributions, and promoter cis-elements of these genes. Furthermore, RNA-seq datasets were used to examine the expression profiles of these genes under ABA treatment. Our results provide fundamental insights into the evolutionary dynamics and potential functional roles of SmCaMs and SmCMLs, especially in hormone signaling and root-specific regulation, and lay the groundwork for future studies on the calcium-mediated regulation of medicinal compound biosynthesis.

2 Results

2.1 Identification of CaM and CML genes in *Salvia miltiorrhiza*

Through a genome-wide search via HMMER and BLAST approaches, we identified a total of 32 candidate calmodulin (CaM) and calmodulin-like (CML) genes in the *Salvia miltiorrhiza* genome. These genes included six SmCaM genes (SmCaM1 to SmCaM6) and twenty-six SmCML genes (SmCML1 to SmCML26). The identification process combined homology-based alignment with Arabidopsis CaM/CML sequences and the presence of conserved EF-hand motifs, which are critical for calcium-binding functionality.

The SmCaM proteins presented highly conserved features (Table 1). Their amino acid sequence lengths ranged narrowly approximately 149 amino acids, except for SmCaM4 (186 amino acids) and SmCaM5 (160 amino acids). Notably, SmCaM2 and SmCaM3 are identical,

whereas SmCaM1 differs in the presence of a single amino acid at position 8 (glutamic acid to aspartic acid). This observation mirrors the phenomenon in *Arabidopsis thaliana*, where AtCaM2, AtCaM3, and AtCaM5 encode identical proteins. All SmCaMs possessed four EF-hand motifs and conserved residues, including a cysteine at position 27 and a lysine at position 116, which are essential for calcium ion binding.

In contrast, SmCML proteins exhibit remarkable diversity in both sequence length and structural characteristics (Table 1). Their lengths spanned from 98 to 535 amino acids, and the number of EF-hand motifs varied between two and four. The sequence similarity of SmCMLs to the representative Arabidopsis CaM (AtCaM6) ranged from 40.146% to 79.054%, considerably lower than that of SmCaMs, reflecting the evolutionary divergence typical of CML family proteins. While certain SmCMLs retain key residues analogous to those of canonical CaMs, many lack either Cys27 or Lys116. Additionally, three SmCML proteins (SmCML10, SmCML18, and SmCML24) were predicted to contain potential myristoylation sites, suggesting possible roles in membrane associations or protein—protein interactions.

The identification of these genes established a foundational catalog of the CaM/CML gene family in S. miltiorrhiza, providing a valuable basis for subsequent analyses of their phylogenetic relationships, structural characteristics, chromosomal localization, and expression profiles.

2.2 Sequence alignment and structural analysis

To investigate the structural conservation and diversity of the CaM and CML proteins in S. miltiorrhiza, multiple sequence alignments were performed via ClustalX 2.1 and DNAMAN. Alignment of the SmCaM proteins with *A. thaliana* CaMs revealed high conservation across their entire amino acid sequence (Fig. 1A). In addition to minor amino acid substitutions, such as an isoleucine-to-valine change at position 118 in SmCaM1, the SmCaM sequences were nearly identical. SmCaM2 and SmCaM3 displayed complete sequence identity, which was consistent with their classification as redundant paralogs, mirroring the sequence redundancy observed in AtCaM2, AtCaM3, and AtCaM5.

Detailed analysis of EF-hand domains across SmCaM proteins revealed the presence of classical “Ehhhhh” α-helical structures, indicative of canonical calcium-binding motifs (Fig. 1B). Each EF-hand domain comprises highly conserved residues, including aspartic acid, asparagine, and glycine, which are known to participate directly in calcium ion coordination. This conservation suggests stable calcium signaling functions for SmCaMs in S. miltiorrhiza.

In contrast, the SmCML proteins exhibited considerable sequence variability, especially within their EF-hand regions. While some SmCMLs, such as SmCML1 and SmCML2, maintain four EF-hand motifs with conserved calcium-binding residues, other members display deletions, substitutions, or truncations in these motifs. For example, SmCML6 and SmCML15 were found to contain only three EF-hand motifs, and SmCML19 possessed only two EF-hand domains concentrated toward the C-terminal region of the protein. Despite this truncation, the EF-hand motifs present in SmCML19 retained characteristic residues, suggesting potential functionality in calcium binding, albeit possibly with altered affinity or specificity.

Further structural observations highlighted unique modifications in specific SmCML proteins (Fig. 1A). In SmCaM6, a glutamine substitution occurred at the fourth residue of the second EF-hand loop, replacing the more common glycine. This substitution could reduce loop flexibility and modify calcium-binding dynamics. These subtle changes might have functional implications, potentially affecting interactions with target proteins.

Moreover, the analysis revealed variations in the conserved regions flanking the EF-hand motifs among SmCMLs. While canonical CaMs typically exhibit glutamic acid-rich segments that facilitate helix—helix interactions, certain SmCMLs lack these regions, indicating possible divergence in structural conformation or partner-binding capabilities. The absence or alteration of these regions in several SmCMLs points toward potential functional specialization or subfunctionalization within the gene family.

Overall, the comparative sequence analysis demonstrated that while SmCaMs are highly conserved and structurally consistent with known CaMs from other plant species, SmCMLs display significant sequence diversity. This diversity, particularly in calcium-binding motifs and flanking regions, suggests potential variability in calcium-sensing capacities and downstream signaling roles among the SmCML family members in S. miltiorrhiza.

2.3 Phylogenetic analysis

To explore the evolutionary relationships of the CaM and CML proteins in S. miltiorrhiza, phylogenetic analyses were conducted via the neighbor-joining (NJ) method implemented in MEGA11. The analysis incorporated the full-length amino acid sequences of SmCaMs, SmCMLs, and their counterparts from *Arabidopsis thaliana* to establish a comprehensive evolutionary framework.

The constructed phylogenetic tree revealed clear separation of the CaM and CML proteins into eight distinct subfamilies, designated Groups I through VIII (Fig. 2). All six SmCaM proteins clustered tightly within subfamily I, forming a monophyletic group alongside AtCaMs. This strong clustering reflects the conserved nature of CaM proteins across plant species, supporting their fundamental roles in calcium signal sensing and transduction.

SmCML proteins, however, displayed a broader distribution across the remaining seven subfamilies, highlighting their greater evolutionary diversification. Subfamilies II and III each contained multiple SmCML members, suggesting possible gene expansion events specific to these groups in S. miltiorrhiza. Notably, subfamily II included SmCML6 and SmCML7, which are closely related to AtCML20, which was previously implicated in guard cell signaling and drought responses in Arabidopsis.

Certain subfamilies were represented by only one or a few SmCML members, such as subfamily IV, containing SmCML17, and subfamily V, consisting solely of SmCML14. These singleton or low-member clusters may represent specialized functions acquired through evolutionary divergence. For example, SmCML14 clustered closely with AtCML25, which is known to mediate pollen germination and pollen tube elongation in Arabidopsis, suggesting potential functional parallels.

The phylogenetic topology also revealed that some SmCMLs, including SmCML19, formed part of subfamily VIII along with larger proteins such as SmCML18 and SmCML24. Despite the significant differences in protein length and domain architecture among these members, their clustering indicates potential conservation of ancestral functional motifs, warranting further functional investigation.

Overall, the phylogenetic analysis underscores the evolutionary divergence of CML proteins relative to the conserved CaM family. While SmCaMs appear highly conserved and closely related to their Arabidopsis counterparts, the diversification observed among SmCMLs suggests adaptive evolution, potentially driven by the need for specialized roles in calcium-mediated signaling pathways in S. miltiorrhiza. These findings provide an essential context for interpreting subsequent structural and functional analyses of individual SmCaM/CML proteins.

2.4 Conserved Motifs, Domain Architecture, and Gene Structure

To gain insight into the structural organization of the SmCaM and SmCML genes, we conducted comprehensive analyses of conserved motifs, domain architecture, and gene structure. Conserved motif identification via the MEME suite revealed three highly conserved motifs consistently present across all SmCaM proteins (Fig. 3A). These motifs were arranged in a conserved sequential pattern corresponding to the EF-hand calcium-binding regions, underscoring the structural integrity and conserved functional roles of SmCaMs. The spacing and sequence composition of these motifs were nearly identical among SmCaMs, indicating minimal divergence within this subfamily.

In contrast, SmCML proteins displayed remarkable diversity in motif composition and organization. While many SmCMLs retained Motifs 1 and 2, which align with classical EF-hand regions, several members lacked Motif 3 entirely, as observed for SmCML8, SmCML11, SmCML12 (subfamily VI), and SmCML19 (subfamily VII), suggesting possible specialization or truncation during evolution. Despite these differences, the retained motifs in SmCMLs still mapped to known calcium-binding regions, indicating preserved core functionalities.

Domain prediction analyses confirmed the presence of various EF-hand architectures across SmCaMs and SmCMLs (Fig. 3B). All SmCaMs possessed four canonical EF-hand domains, classified within the PTZ00184 superfamily, characteristic of typical calmodulin proteins. Conversely, SmCMLs exhibited varied domain arrangements: some, such as SmCML1 and SmCML2, harbored four EF-hand domains similar to those of SmCaMs, whereas others contained only two or three EF-hand motifs or presented unique configurations. Notably, SmCML15 has features characteristic of the PEF (penta-EF-hand) family, which is often implicated in processes such as apoptosis and protein degradation, suggesting potential functional divergence beyond conventional calcium signaling. Similarly, SmCML7 possessed EF-hand arrangements reminiscent of AtCML19 and AtCML20, which are known for their roles in DNA repair and guard cell signaling in Arabidopsis, implying specialized functional pathways in S. miltiorrhiza. Subtle variations were observed within the EF-hand domains; for example, SmCaM6 presented a glutamine substitution for glycine at the fourth residue of the second EF-hand loop, potentially impacting flexibility and calcium-binding dynamics, thus highlighting the structural plasticity of the CaM/CML proteins.

Gene structure analysis based on S. miltiorrhiza genomic annotations provided further evidence of this dichotomy. All SmCaM genes contained multiple introns and presented conserved exon—intron structures, including a characteristic intron interrupting the coding sequence of the first EF—hand domain (Fig. 3C). This arrangement mirrors patterns observed in Arabidopsis and other angiosperms, suggesting evolutionary conservation within the CaM subfamily.

In contrast, SmCML genes presented significant variability in their gene structures. Among the 26 SmCML genes, 14 were intronless, consisting of single-exon coding sequences, exemplified by genes such as SmCML6, SmCML20, SmCML21, and SmCML19. This streamlined architecture may facilitate rapid transcriptional responses, a feature observed among certain stress-responsive plant genes. The remaining 12 SmCML genes contained varying numbers of introns, indicating structural diversity within the family. Some proteins, such as SmCML4 and SmCML17, exhibit complex exon—intron architectures, potentially enabling regulatory flexibility and alternative splicing. Notably, SmCML19, the shortest gene in the family, lacked introns entirely and encoded only two EF-hand motifs. Despite its truncated structure, SmCML19 retained conserved calcium-binding regions, suggesting potential functional significance despite its minimal architecture.

Collectively, these analyses underscore the dichotomy between the highly conserved SmCaM subfamily and the structurally diverse SmCML subfamily, which is evident in motif composition, domain architecture, and gene organization. This structural diversity among SmCMLs may underlie specialized regulatory and functional roles within calcium-mediated signaling pathways in S. miltiorrhiza, warranting further functional exploration.

2.5 Cis-Acting Element Analysis

To elucidate potential regulatory mechanisms governing the expression of the SmCaM and SmCML genes, we analyzed cis-acting regulatory elements within the 2 kb promoter regions upstream of each gene's transcription start site. Promoter sequences were extracted from the *S. miltiorrhiza* genome and subjected to motif scanning via the PlantCARE and PLACE databases, as described in the Methods section. Two datasets were generated: one quantifying the number of each cis-element type per gene and another detailing the positional distribution of individual motifs within the promoter regions.

The quantitative analysis revealed a diverse array of cis-elements, totaling 20 functional categories across all the promoters (Fig. 4A-B). Among these, light-responsive elements were the most prevalent and were detected in nearly all SmCaM and SmCML promoters with high frequency. Elements such as the GT1 motif and Box 4 were frequently observed, which is consistent with the potential roles of CaM/CML genes in light-mediated signaling pathways.

Hormone-responsive elements were abundantly represented. Notably, ABA-responsive elements, including ABRE motifs, were present in varying copy numbers across numerous SmCML promoters, particularly in SmCML6, SmCML7, SmCML17, and SmCML22, suggesting their involvement in ABA-mediated stress responses. Auxin-responsive elements, such as TGA elements and AuxRE cores, were also detected, suggesting the possible participation of these genes in auxin signaling and developmental processes. MeJA-related elements (CGTCA motif and TGACG motif) were frequently found in SmCML promoters, suggesting roles in defense and stress response pathways (Fig. 4A).

Elements associated with environmental stress responses, including anaerobic-responsive elements (AREs), low-temperature responsiveness motifs, and drought-inducible MYB-binding sites (MBSs), were identified in several SmCML promoters (Fig. 4B). For example, SmCML11 and SmCML13 contain multiple MBS motifs, suggesting potential responsiveness to drought stress.

Spatial distribution analysis of the cis-elements within the promoter regions revealed clustering of certain motifs near the transcription start sites, particularly within the first 500 bp upstream regions. This pattern was evident for many light- and hormone-responsive elements, suggesting a potential influence on transcription initiation and gene regulation.

Interestingly, SmCML19, despite its minimal gene architecture and shorter promoter length, contained several cis-elements related to meristem-specific expression and the stress response, including CTCC_motif and E2Fb binding sites. However, it lacked canonical ABA-responsive motifs such as ABREs, indicating possible regulation via alternative signaling pathways.

Overall, the updated cis-acting element analysis provides a more accurate and comprehensive view of the potential transcriptional regulation of the SmCaM and SmCML genes. These findings underscore the complexity of regulatory networks that may govern the diverse roles of CaM/CML family members in development and stress responses in *S. miltiorrhiza*.

2.6 Chromosomal distribution

To investigate the genomic organization of the SmCaM/CML gene family, chromosomal localization analysis was performed via TBtools, which mapped the physical positions of each gene onto the eight assembled chromosomes of S. miltiorrhiza. Among the 32 identified SmCaM/CML genes, 29 were successfully mapped to specific chromosomes, whereas three genes (SmCML13, SmCML15, and SmCML16) remained unanchored, likely residing on unassembled scaffolds (Fig. 5A).

The chromosomal distribution of the SmCaM and SmCML genes was uneven across the genome. Chromosome 1 harbored the greatest number of family members, with nine genes located on this chromosome. Chromosomes 3 and 6 also contained significant clusters, hosting five and four SmCaM/CML genes, respectively. In contrast, chromosomes 2, 5, 7, and 8 each contained fewer than four family members.

The SmCaM genes were predominantly located on chromosome 1, with SmCaM2 and SmCaM5 positioned in close physical proximity. This arrangement suggests a possible tandem duplication event, contributing to the expansion of the CaM subfamily within the S. miltiorrhiza genome. Other SmCaMs were distributed more distantly across the same chromosome or located on separate chromosomes, indicating the potential involvement of segmental duplications or chromosomal rearrangements in their evolutionary history.

The SmCML genes exhibited a more dispersed distribution pattern, with certain genes located in close proximity on individual chromosomes. For example, SmCML2 and SmCML4 were found adjacent to chromosome 3, suggesting tandem duplication. Interestingly, SmCML23 was the sole CML gene mapped to chromosome 7, which aligns with its unique phylogenetic placement and potentially distinct functional role.

To assess whether SmCaM/CML genes preferentially occupy regions of high gene density, genomic bins were analyzed for gene density, defining high-density regions as those above the third quartile (Q3) threshold of gene counts per bin. Among the 32 SmCaM/CML genes, 15 were located within high-density genomic regions. Permutation testing (10,000 simulations) revealed that this enrichment was statistically significant ($p = 0.0024$), indicating a nonrandom distribution favoring gene-rich areas (Fig. 5B-C).

These findings suggest that the SmCaM and SmCML genes are not randomly scattered across the genome but rather exhibit specific localization patterns. The clustering of family members, particularly within high-density genomic regions, may facilitate coordinated regulation or coexpression, potentially contributing to functional integration within calcium-mediated signaling pathways in S. miltiorrhiza.

2.7 Expression Profiles Under ABA Treatment

To elucidate the potential involvement of SmCaM and SmCML genes in hormonal signaling and stress responses, transcriptomic analyses were conducted via RNA-seq data derived from S. miltiorrhiza leaf and root tissues exposed to abscisic acid (ABA) treatment for 0 h, 12 h, and 24 h. Data processing included adapter trimming, quality filtering, and alignment to the S. miltiorrhiza reference genome, followed by the quantification of gene expression levels via DESeq2.

Principal component analysis (PCA) was performed to assess global expression patterns across the samples (Fig. 6A). The PCA plot revealed clear separation between the ABA-treated and control groups, indicating that significant transcriptomic changes were induced by ABA. Notably, samples from leaf and root tissues segregated distinctly along the primary principal component axis, reflecting tissue-specific transcriptional responses. Within tissue types, time-course analysis demonstrated further stratification, with the 12 h and 24 h treatment groups diverging from the 0 h control groups, particularly in the leaf samples.

Heatmap visualization of normalized expression values highlighted differential expression profiles among the SmCaM and SmCML genes across treatment conditions (Fig. 6B). Overall, SmCaM genes presented relatively stable expression levels in both leaf and root tissues, with minor fluctuations under ABA treatment. This consistent expression pattern suggests that SmCaMs may serve as constitutive calcium sensors involved in maintaining basal calcium signaling homeostasis, independent of acute stress stimuli.

Conversely, SmCML genes presented pronounced variability in expression in response to ABA treatment, indicating potential roles in stress-responsive signaling pathways. Several SmCML genes were significantly upregulated in leaf tissues following ABA exposure. For example, SmCML6, SmCML7, and SmCML17 presented substantial increases in transcript abundance at 24 h posttreatment. These

genes clustered together phylogenetically within the same subfamily, suggesting shared regulatory mechanisms and possibly redundant or cooperative functions in mediating ABA-induced signaling cascades.

Volcano plot analysis further confirmed the differential expression patterns of SmCML genes under ABA treatment (Fig. 6C). In the leaf tissues, eight SmCML genes were significantly upregulated after 24 h of ABA treatment (log2-fold change > 1, adjusted p value < 0.05). These genes included SmCML6, SmCML7, SmCML17, and several others, highlighting a coordinated transcriptional response potentially linked to adaptive stress signaling pathways. In contrast, only a few SmCML genes were significantly downregulated in leaves, indicating a predominantly activational transcriptional response in this tissue.

The expression landscape of root tissues differed. While several SmCML genes were differentially expressed following ABA treatment, the overall magnitude of expression changes was less pronounced than that in leaf tissues. Notably, SmCML19 and SmCML21 were consistently downregulated in roots at both the 12 h and 24 h time points relative to those in the controls. SmCML19, in particular, exhibited substantial downregulation. Updated promoter analysis revealed that while it contains several meristem- and stress-related cis-elements, it lacks canonical ABA-responsive motifs such as ABREs, suggesting that its regulation under ABA stress may involve alternative pathways or indirect regulatory networks.

Closer examination of SmCML19 revealed unique expression dynamics. Although it is characterized by a truncated protein structure and a relatively simple promoter architecture, SmCML19 was significantly downregulated exclusively in root tissues under ABA stress but maintained low expression levels in leaves across all conditions. This observation, coupled with its reduced EF-hand motif composition, raises the possibility that SmCML19 functions as an atypical calcium sensor or negative regulator specifically within root stress response pathways.

The expression patterns were also correlated with chromosomal localization. Several upregulated SmCML genes, such as SmCML6 and SmCML17, were mapped to gene-rich regions on chromosomes 1 and 4, suggesting that clustered gene organization may facilitate coordinated expression under stress conditions. This genomic arrangement could enable rapid and synergistic transcriptional responses during ABA-mediated signaling events.

Collectively, the expression profiling results indicate that while SmCaMs maintain steady expression levels, SmCML genes exhibit dynamic, tissue specific, and time-dependent responses to ABA treatment. These differential expression patterns imply that SmCMLs play diverse and potentially specialized roles in mediating ABA signaling and stress adaptation in S. miltiorrhiza. These data provide a valuable foundation for future functional studies aimed at unraveling the precise mechanisms through which individual SmCMLs contribute to calcium-dependent hormonal regulation and stress resistance.

3 Discussion

In this study, we performed a comprehensive genome-wide analysis of the calmodulin (CaM) and calmodulin-like (CML) gene families in Salvia miltiorrhiza, integrating sequence characterization, structural analysis, phylogenetic relationships, chromosomal localization, cis-regulatory element profiling, and transcriptomic expression patterns under ABA treatment. This work represents the first systematic exploration of this gene family in this important medicinal species, providing a valuable framework for future functional investigations.

Our findings demonstrate a clear dichotomy between the SmCaM and SmCML subfamilies in both structure and expression behavior. SmCaM proteins exhibit striking conservation across their amino acid sequences, domain architecture, and gene structures, which is consistent with previous observations in other plant species, such as *Arabidopsis thaliana* and wheat (*Triticum aestivum* L.) [7, 13]. This high degree of conservation underlines their fundamental roles as ubiquitous calcium sensors, which are likely involved in core cellular signaling processes that are maintained across diverse plant lineages. The identification of four EF-hand motifs in all SmCaM proteins, together with the preservation of critical residues such as Cys27 and Lys116, suggests functional stability in calcium binding and target protein interactions.

In contrast, SmCML proteins display marked differences in sequence length, EF-hand motif composition, exon—intron structure, and promoter cis-element content. These observations align with prior studies indicating that CML genes have evolved considerable functional diversification, allowing them to participate in specialized signaling pathways and environmental responses[5, 6]. The presence of SmCMLs lacking introns or exhibiting truncated structures, such as SmCML19, hints at possible adaptations for rapid transcriptional responses under stress conditions, as has been observed for other stress-inducible genes in plants.

A cis-acting element analysis based on promoter sequences revealed that both the SmCaM and SmCML genes harbor diverse regulatory motifs associated with developmental processes, hormone signaling, and environmental stress responses. Notably, light-responsive elements are highly prevalent across the promoters of both gene families, whereas ABA-responsive elements such as ABREs and MYB binding sites, including SmCML6, SmCML7, and SmCML17, have been identified as having varying copy numbers among SmCML promoters [14, 15]. These elements may contribute to the differential transcriptional regulation observed under ABA treatment.

Our expression data under ABA stress further highlight the functional divergence within the SmCaM/CML family. While SmCaM genes maintained relatively stable expression, suggesting their constitutive roles in maintaining calcium homeostasis, SmCML genes presented pronounced, tissue specific, and time-dependent expression changes in response to ABA[16]. This pattern implies that SmCMLs might act as fine-tuned modulators of calcium signaling, integrating environmental cues and hormonal signals to orchestrate stress adaptation processes. The significant upregulation of SmCML6, SmCML7, and SmCML17 in leaf tissues following ABA exposure is particularly intriguing, as these genes could be key players in ABA-mediated signaling networks, possibly contributing to stomatal regulation, secondary metabolism, or stress-induced gene expression pathways[17, 18].

SmCML19 has emerged as an especially compelling candidate for further study. Despite its short protein length and lack of classical ABA-responsive cis-elements, SmCML19 was consistently and significantly downregulated in root tissues under ABA treatment. Moreover, structural modeling revealed that SmCML19 retains two complete EF-hand domains, suggesting that it can maintain its calcium-binding ability despite its truncated sequence. These findings suggest that SmCML19 functions as an atypical calcium sensor or negative regulator of ABA signaling, perhaps through unique protein—protein interactions or the modulation of specific transcriptional networks. Similar cases have been reported in other systems, where truncated or divergent CMLs retain specialized regulatory functions distinct from those of canonical CaMs[19, 20].

Chromosomal localization analyses further revealed that SmCaM/CML genes are nonrandomly distributed within the S. miltiorrhiza genome and are significantly enriched in regions with high gene density. This clustering may facilitate coordinated gene regulation, enabling rapid, collective transcriptional responses during stress adaptation. This genomic organization is consistent with observations in other plant genomes, where gene clusters can contribute to functional synergy among related genes[21].

Taken together, our findings reveal the complexity and evolutionary plasticity of the CaM/CML gene family in S. miltiorrhiza[22, 23]. The diversity in gene structure, domain architecture, and expression dynamics points toward a finely tuned calcium signaling network capable of integrating multiple signals to modulate plant growth, development, and stress resilience[24]. These findings lay a robust foundation for future experimental validation, including functional analyses via CRISPR/Cas9-mediated gene editing, protein interaction studies, and calcium flux assays, to elucidate the precise biological roles of individual SmCaM and SmCML members[25].

In particular, further investigations into the unique features and regulatory mechanisms of SmCML19 could provide critical insights into noncanonical calcium signaling pathways and their roles in stress responses in medicinal plants[26, 27, 28]. This knowledge could ultimately inform efforts to increase the stress tolerance and secondary metabolite production of S. miltiorrhiza, contributing to the improvement of this valuable traditional medicinal resource.

4 Materials and methods

4.1 Identification and characterization of CaM and CML gene families in Salvia miltiorrhiza

The complete genome sequence of *S. miltiorrhiza* was downloaded from the China National Center for Bioinformation (https://www.cncb.ac.cn/)[29]. The sequences of seven CaM and 28 CML proteins from *Arabidopsis thaliana* were retrieved from TAIR (https://www.arabidopsis.org/)[30].

Candidate SmCaM and SmCML proteins were identified by searching the *S. miltiorrhiza* proteome against Arabidopsis sequences via BLAST implemented in TBtools-II and HMMER (for profile hidden Markov model analysis)[31,32]. The results from both methods were integrated for subsequent analyses.

Conserved domains and motifs of the candidate sequences were analyzed via InterPro (https://www.ebi.ac.uk/interpro/)[33], SMART (https://smart.embl.de/)[34], and Pfam (http://pfam-legacy.xfam.org/)[35]. Multiple sequence alignments were performed with

ClustalX 2.1[36]. Proteins were classified as SmCaMs if they contained exactly four EF-hand motifs, approximately 150 amino acids in length, with at least 90% sequence identity to AtCaM6, and included 27 cysteine and 116 lysine residues. Proteins containing one to six EF-hand motifs and at least 40% identity to AtCaM6 were defined as SmCMLs.

4.2 Multiple Sequence Alignment

Multiple sequence alignments of SmCaM and SmCML proteins with their Arabidopsis homologs were conducted via ClustalX 2.1, MEGA 11[37], and DNAMAN (Lynnon Corporation, Canada) under default parameters. EF-hand motifs were annotated within the sequences. Potential N-myristoylation sites were predicted with a myristoylator from ExPASy (http://www.expasy.org/myristoylator/)[38].

4.3 Phylogenetic analysis

Phylogenetic trees were constructed in MEGA 11 via the neighbor–joining method with 1,000 bootstrap replicates and the Poisson substitution model. Trees were visualized and annotated via iTOL (Interactive Tree Of Life, https://itol.embl.de/)[39]. Subfamily classification was based on references from Arabidopsis.

4.4 Gene Structure and Conserved Motif Analysis

Coding sequences (CDSs) of *S. miltiorrhiza* were retrieved from the China National Center for Bioinformation. Conserved domain information was integrated from analyses via InterPro, SMART, and Pfam, along with results from NCBI's CD-Search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) and MEME Suite (https://meme-suite.org/meme/tools /meme)[40,41].

Cis-acting regulatory elements were predicted with PlantCARE (https://bioinformatics.psb.ugent.be/webtools/ plantcare/html/) and PLACE (https://www.dna.affrc.go.jp/PLACE/?action=newplace)[42,43]. The data were visualized via TBtools-II and R to generate exon–intron structure diagrams, heatmaps of the cis-elements, and conserved domain maps. For the selected genes, three-dimensional protein structures were predicted online via AlphaFold2 via ColabFold v1.5. [44,45].

4.5 Chromosomal Localization

The chromosomal locations of the SmCaM and SmCML genes were determined with TBtools-II. Chromosome lengths were derived from genome data, and gene locations were visualized with the Gene Location Visualize (Advanced) tool. Gene density was analyzed from the GFF files. The enrichment patterns were examined in Python via permutation tests to assess nonrandom distributions, and the results are presented as frequency plots.

4.6 Expression analysis

Transcriptomic data of S. miltiorrhiza leaves and roots treated with abscisic acid (ABA) at different time points were downloaded from the NCBI Sequence Read Archive (SRA, https://www.ncbi.nlm.nih.gov/sra), comprising six groups with two biological replicates each, totaling twelve datasets[46].

The raw RNA-seq data were processed via the SeqKit under the Windows Subsystem for Linux (WSL)[47]. Adapter removal and quality filtering were conducted with Porechop and NanoFilt[48,49]. Read quality was assessed via NanoPlot[50,51]. Clean reads were aligned with Minimap2, and sorted BAM files were generated with SAMtools[52,53]. Gene expression levels were quantified via FeatureCounts[54]. Differential expression and normalization were performed in R via DESeq2[55,56]. Data visualizations, including volcano plots, principal component analysis (PCA) plots, and heatmaps, were produced via the R packages Pheatmap, ggplot2, EnhancedVolcano, and RColorBrewer[57,58,59].

4.7 Ethics approval

This article does not contain any studies with human participants or animals performed by any of the authors. Therefore, ethics approval and consent were not needed.

Abbreviations

ABA: abscisic acid; ABRE: ABA-responsive element; At: *Arabidopsis thaliana*; Ca^{2+}: calcium ion; CaM: calmodulin; CML: calmodulin-like; CDS: coding DNA sequence; CNCB: China National Center for Bioinformation; Cys27: cysteine at position 27; EF-hand: helix–loop–helix calcium-binding motif; FPKM: fragments per kilobase of transcript per million mapped reads; GF: general feature format; HMMER: hidden Markov model for protein domain detection; iTOL: Interactive Tree of Life; Lys116: lysine at position 116; MEME: Multiple EM for Motif Elicitation; NJ: neighbor–Joining; PCA: principal component analysis; qRT–PCR: quantitative real–time polymerase chain reaction; RNA–seq: RNA sequencing; SRA: Sequence Read Archive; Sm: Salvia miltiorrhiza; SMART: Simple Modular Architecture Research Tool; TBtools: Toolkit for Biologists; WSL: Windows Subsystem for Linux.

Declarations

Competing Interests

The authors have no relevant financial or nonfinancial interests to disclose.

Funding

The authors did not receive support from any organization for the submitted work.

Author Contribution

Yansong Zhang conceived the project, conducted the data analysis, and drafted the manuscript. Yuanchu Liu supervised the research and contributed to manuscript revision. Yuan Yang assisted with figure preparation, formatting, and proofreading. All the authors read and approved the final manuscript.

Acknowledgement

The authors would like to thank Dr. Yuanchu Liu for his guidance and constructive feedback throughout the development of this study. We are also grateful to the Applied Research Institute of Life Sciences at Xi'an International University for providing computational resources and technical support. Special thanks are extended to the team maintaining the Salvia miltiorrhiza genome database at the China National Center for Bioinformation (CNCB) for making the genomic resources publicly available.We also acknowledge the use of RNA-seq data retrieved from the NCBI Sequence Read Archive (SRA), which facilitated the transcriptomic analyses presented in this work.

Data Availability

The genomic and transcriptomic datasets analyzed during the current study are publicly available. The *Salvia miltiorrhiza* reference genome was obtained from the China National Center for Bioinformation (CNCB, https://ngdc.cncb.ac.cn/). The RNA-seq data used for expression analysis were retrieved from the NCBI Sequence Read Archive (SRA) under accession number PRJNA560452. All processed data supporting the findings of this study are available from the corresponding author upon reasonable request.

References

1. Rasmussen, H., Barrett, P., Smallwood, J., Bollag, W. & Isales, C. Calcium ion as intracellular messenger and cellular toxin. *Environ. Health Perspect.* **84**, 17–25 (1990).
2. Hoeflich, K. P. & Ikura, M. Calmodulin in action: diversity in target recognition and activation mechanisms. *Cell* **108**, 739–742 (2002).

3. Wang, L., Liu, Z., Han, S., Liu, P., Sadeghnezhad, E. & Liu, M. Growth or survival: what is the role of calmodulin-like proteins in plant? *Int. J. Biol. Macromol.* **242**, 124733 (2023).
4. Klee, C. B., Crouch, T. H. & Richman, P. G. Calmodulin. *Annu. Rev. Biochem.* **49**, 489–515 (1980).
5. Zeng, H., Zhu, Q., Yuan, P., Yan, Y., Yi, K. & Du, L. Calmodulin and calmodulin-like protein-mediated plant responses to biotic stresses. *Plant Cell Environ.* **46**, 3680–3703 (2023).
6. McCormack, E., Tsai, Y. C. & Braam, J. Handling calcium signaling: Arabidopsis CaMs and CMLs. *Trends Plant Sci.* **10**, 383–389 (2005).
7. Boonburapong, B. & Buaboocha, T. Genome-wide identification and analyses of the rice calmodulin and related potential calcium sensor proteins. *BMC Plant Biol.* **7**, 4 (2007).
8. Yadav, M., Pandey, J., Chakraborty, A., Hassan, M. I., Kundu, J. K., Roy, A. *et al.* A comprehensive analysis of calmodulin-like proteins of *Glycine max* indicates their role in calcium signaling and plant defense against insect attack. *Front. Plant Sci.* **13**, 817950 (2022).
9. Liu, Y., Chen, W., Liu, L., Su, Y., Li, Y., Jia, W. *et al.* Genome-wide identification and expression analysis of calmodulin and calmodulin-like genes in wheat (*Triticum aestivum* L.). *Plant Signal. Behav.* **17**, 2013646 (2022).
10. Su, C. Y., Ming, Q. L., Rahman, K., Han, T. & Qin, L. P. *Salvia miltiorrhiza*: traditional medicinal uses, chemistry, and pharmacology. *Chin. J. Nat. Med.* **13**, 163–182 (2015).
11. Yang, Y., Du, L., Zhao, H., Zhao, Y., Liao, S., Zhang, Z. *et al.* Danshensu derivatives: a series of promising drugs with protective effects against cardiocerebrovascular diseases and cancers. *Med. Res. Rev.* **45**, 1163–1183 (2025).
12. Ikram, M., Batool, M., Ullah, M., Khalid, B., El-Badri, A. M., Mohamed, I. A. A. *et al.* Molecular alchemy: converting stress into resilience via secondary metabolites and calcium signaling in rice. *Rice* **18**, 32 (2025).
13. Bouché, N., Yellin, A., Snedden, W. A. & Fromm, H. Plant-specific calmodulin-binding proteins. *Annu. Rev. Plant Biol.* **56**, 435–466 (2005).
14. Cai, K. ***et al.*** Calmodulin and calmodulin-like gene family in barley: identification, characterization and expression analyses. *Front. Plant Sci.* **13**, 964888 (2022).
15. Yoshida, T. ***et al.*** Four *Arabidopsis* AREB/ABF transcription factors function predominantly in gene expression downstream of SnRK2 kinases in abscisic acid signaling in response to osmotic stress. *Plant Cell Environ.* **38**, 35–49 (2015).
16. McCormack, E. & Braam, J. Calmodulins and related potential calcium sensors of *Arabidopsis*. *New Phytol.* **159**, 585–598 (2003).
17. Cutler, S. R., Rodriguez, P. L., Finkelstein, R. R. & Abrams, S. R. Abscisic Acid: Emergence of a Core Signaling Network. *Annu. Rev. Plant Biol.* **61**, 651–679 (2010).
18. Ma, Q. ***et al.*** Isolation and expression analysis of *CsCML* genes in response to abiotic stresses in the tea plant (*Camellia sinensis*). *Sci. Rep.* **9**, 8211 (2019).
19. La Verde, V., Trande, M., D'Onofrio, M., Dominici, P. & Astegno, A. Binding of calcium and target peptide to calmodulin-like protein CML19, the centrin 2 of *Arabidopsis thaliana*. *Int. J. Biol. Macromol.* **108**, 1289–1299 (2018).
20. Wu, X., Qiao, Z., Liu, H., Acharya, B. R., Li, C. & Zhang, W. CML20, an *Arabidopsis* calmodulin-like protein, negatively regulates guard cell ABA signaling and drought stress tolerance. *Front. Plant Sci.* **8**, 824 (2017).
21. Lu, Y., Truman, W., Liu, X., Bethke, G., Zhou, M., Myers, C. L., Katagiri, F. & Glazebrook, J. Different modes of negative regulation of plant immunity by calmodulin-related genes. *Plant Physiol.* **176**, 3046–3061 (2018).
22. Mohanta, T. K., Kumar, P. & Bae, H. Genomics and evolutionary aspect of calcium signaling event in calmodulin and calmodulin-like proteins in plants. *BMC Plant Biol.* **17**, 38 (2017)
23. Vandelle, E. *et al.* Identification, characterization, and expression analysis of calmodulin and calmodulin-like genes in grapevine (*Vitis vinifera*) reveal likely roles in stress responses. *Plant Physiol. Biochem.* **129**, 221–237 (2018).
24. Kudla, J. *et al.* Advances and current challenges in calcium signaling. *New Phytol.* **218**, 414–431 (2018).
25. Zhou, Z. *et al.* Multiplexed CRISPR/Cas9-mediated knockout of laccase genes in *Salvia miltiorrhiza* revealed their roles in growth, development, and metabolism. *Front. Plant Sci.* **12**, 647768 (2021).
26. Day, I. S., Reddy, V. S., Ali, G. S. & Reddy, A. S. N. Analysis of EF-hand-containing proteins in Arabidopsis. Genome Biol. **3**, research0056 (2002).
27. Zhou, L., Fu, Y. & Yang, Z. A genome-wide functional characterization of Arabidopsis regulatory calcium sensors in pollen tubes. J. Integr. Plant Biol. **51**, 751–761 (2009).

28. UniProt Consortium. Probable calcium-binding protein CML28 (*Arabidopsis thaliana*). UniProtKB/Swiss-Prot Q9SRP7. Available at: https://www.uniprot.org/uniprotkb/Q9SRP7 (Accessed 28 Aug 2025).
29. Song, Z., Lin, C., Xing, P., Fen, Y., Jin, H., Zhou, C. *et al.* A high-quality reference genome sequence of *Salvia miltiorrhiza* provides insights into tanshinone synthesis in its red rhizomes. *Plant Genome* **13**, e20041 (2020).
30. Berardini, T. Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E. & Huala, E. The Arabidopsis information resource: making and mining the 'gold standard' annotated reference plant genome. *Genesis* **53**, 474–485 (2015).
31. Chen, C., Wu, Y., Li, J., Wang, X., Zeng, Z., Xu, J. *et al.* TBtools-II: a 'one for all, all for one' bioinformatics platform for biological big-data mining. *Mol. Plant* **16**, 1733–1742 (2023).
32. Potter, S. C., Luciani, A., Eddy, S. R., Park, Y., Lopez, R. & Finn, R. D. HMMER web server: 2018 update. *Nucleic Acids Res.* **46**, W200–W204 (2018).
33. Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B. L., Salazar, G. A. *et al.* InterPro in 2022. *Nucleic Acids Res.* **51**, D418–D427 (2023).
34. Letunic, I., Khedkar, S. & Bork, P. SMART: recent updates, new developments and status in 2020. *Nucleic Acids Res.* **49**, D458–D460 (2021).
35. Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L. *et al.* Pfam: the protein families database in 2021. *Nucleic Acids Res.* **49**, D412–D419 (2021).
36. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
37. Tamura, K., Stecher, G. & Kumar, S. MEGA11: molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* **38**, 3022–3027 (2021).
38. Duvaud, S., Gabella, C., Lisacek, F., Stockinger, H., Ioannidis, V. & Durinx, C. ExPASy, the Swiss Bioinformatics Resource Portal, as designed by its users. *Nucleic Acids Res.* **49**, W216–W227 (2021).
39. Letunic, I. & Bork, P. Interactive Tree of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296 (2021).
40. Wang, J., Chitsaz, F., Derbyshire, M. K., Gonzales, N. R., Gwadz, M., Lu, S. *et al.* The conserved domain database in 2023. *Nucleic Acids Res.* **51**, D384–D388 (2023).
41. Bailey, T. L., Johnson, J., Grant, C. E. & Noble, W. S. The MEME suite. *Nucleic Acids Res.* **43**, W39–W49 (2015).
42. Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y. *et al.* PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* **30**, 325–327 (2002).
43. Higo, K., Ugawa, Y., Iwamoto, M. & Korenaga, T. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Res.* **27**, 297–300 (1999).
44. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
45. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S. & Steinegger, M. ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).
46. Leinonen, R., Sugawara, H., Shumway, M. & International Nucleotide Sequence Database Collaboration. The Sequence Read Archive. *Nucleic Acids Res.* **39**, D19–D21 (2011).
47. Shen, W., Sipos, B. & Zhao, L. SeqKit2: a swiss army knife for sequence and alignment processing. *iMeta* **3**, e191 (2024).
48. Wick, R. R., Judd, L. M., Gorrie, C. L. & Holt, K. E. Completing bacterial genome assemblies with multiplex MinION sequencing. *Microb. Genom.* **3**, e000132 (2017).
49. De Coster, W., D'Hert, S., Schultz, D. T., Cruts, M. & Van Broeckhoven, C. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* **34**, 2666–2669 (2018).
50. De Coster, W. & Rademakers, R. NanoPack2: population-scale evaluation of long-read sequencing data. *Bioinformatics* **39**, btad311 (2023).
51. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
52. Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O. *et al.* Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008 (2021).

53. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general-purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
54. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
55. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. https://www.r-project.org (2025).
56. Kolde, R. pheatmap: pretty heatmaps. GitHub, https://github.com/raivokolde/pheatmap (2025).
57. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis.* (Springer, New York, 2009).
58. Blighe, K., Rana, S. & Lewis, M. EnhancedVolcano: publication-ready volcano plots with enhanced coloring and labeling. Bioconductor, https://bioconductor.org/packages/EnhancedVolcano (2025).
59. Neuwirth, E. RColorBrewer: ColorBrewer palettes. CRAN, https://CRAN.R-project.org/package=RColorBrewer (2022).

Tables

Table 1: Sequence characteristics of Salvia miltiorrhiza CaM and CML proteins.

Name	Accession Number	Group number	Length of protein (bp)	Number of EF hands	Percentage of methionine	Presence of cysteine 27	Presence of lysine 116	Potential myristoylation site	Similarity to AtCaM6 protein
SmCaM1	EVM0001113.1	1	149	4	6.0%	+	+		99.329
SmCaM2	EVM0006388.1	1	149	4	6.0%	+	+		98.658
SmCaM3	EVM0002422.1	1	149	4	6.0%	+	+		98.658
SmCaM4	EVM0001121.1	1	186	4	4.8%	+	+		97.987
SmCaM5	EVM0000751.1	1	160	4	5.6%	+	+		91.875
SmCaM6	EVM0003400.1	1	149	4	6.0%	+	+		90.604
SmCML1	EVM0015279.1	2	150	4	6.7%	+	+		79.054
SmCML2	EVM0021986.1	2	150	4	6.0%	+	+		77.703
SmCML3	EVM0003732.1	2	151	4	7.3%	+			68.243
SmCML4	EVM0007053.2	2	180	4	5.0%		+		64.607
SmCML5	EVM0000268.1	2	145	3	3.4%	+	+		51.034
SmCML6	EVM0000662.1	2	145	3	3.4%		+		49.655
SmCML7	EVM0025465.1	2	169	4	7.1%				48.592
SmCML8	EVM0022857.1	6	220	2	2.7%				48.571
SmCML9	EVM0006715.1	3	147	4	4.8%		+		48.551
SmCML10	EVM0003909.1	2	147	3	3.4%		+	+	48.322
SmCML11	EVM0025629.1	6	188	2	2.1%		+		46.875
SmCML12	EVM0016198.1	6	185	2	5.4%		+		46.875
SmCML13	EVM0022270.1	3	152	4	3.3%		+		45.926
SmCML14	EVM0014301.1	5	190	4	4.7%				45.588
SmCML15	EVM0007945.1	3	216	3	2.8%		+		45.455
SmCML16	EVM0008161.1	7	164	4	2.4%				44.304
SmCML17	EVM0007809.1	4	154	4	6.5%		+		42.667
SmCML18	EVM0025367.1	8	521	4	4.0%			+	42.553
SmCML19	EVM0003329.1	8	98	2	4.1%				42.177
SmCML20	EVM0023528.1	3	140	3	2.9%				41.379
SmCML21	EVM0009975.1	3	202	3	2.5%				41.27
SmCML22	EVM0000136.1	3	205	3	2.4%				41.27
SmCML23	EVM0001245.1	3	185	4	5.4%				41.27
SmCML24	EVM0013168.1	8	535	4	3.9%			+	41.176
SmCML25	EVM0021133.1	7	166	4	5.4%				41.096
SmCML26	EVM0006109.1	3	222	4	3.2%				40.816

Figures

Fig. 1

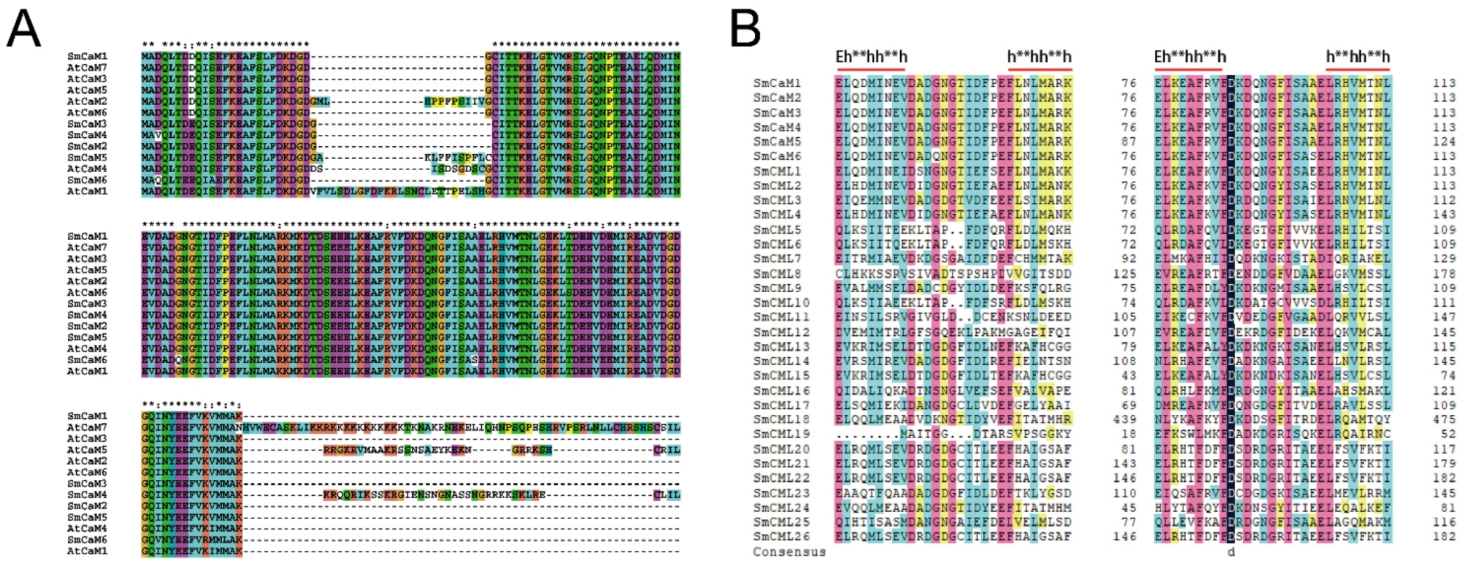


Figure 1

Comparative sequence alignment and EF-hand motif analysis of CaM and CML proteins in *Salvia miltiorrhiza* and *Arabidopsis thaliana*. (A) Multiple sequence alignment of the full-length amino acid sequences of seven AtCaMs and six SmCaMs performed via ClustalX 2.1. Conserved EF-hand structural motifs are indicated above the aligned sequences with the signature pattern “Eh**hh*h.” Residues in the alignment are colored according to the degree of conservation: black (highly conserved), pink, blue, yellow, and white (least conserved). (B) Multiple sequence alignment of the full-length amino acid sequences of six SmCaMs and twenty-six SmCMLs conducted with DNAMAN. EF-hand motifs are annotated as in (A), with color coding reflecting conservation levels. Sequence logos or conservation scoring methods from the alignment software were used to visualize the conservation pattern. At, *Arabidopsis thaliana*; Sm, Salvia miltiorrhiza

Fig. 2

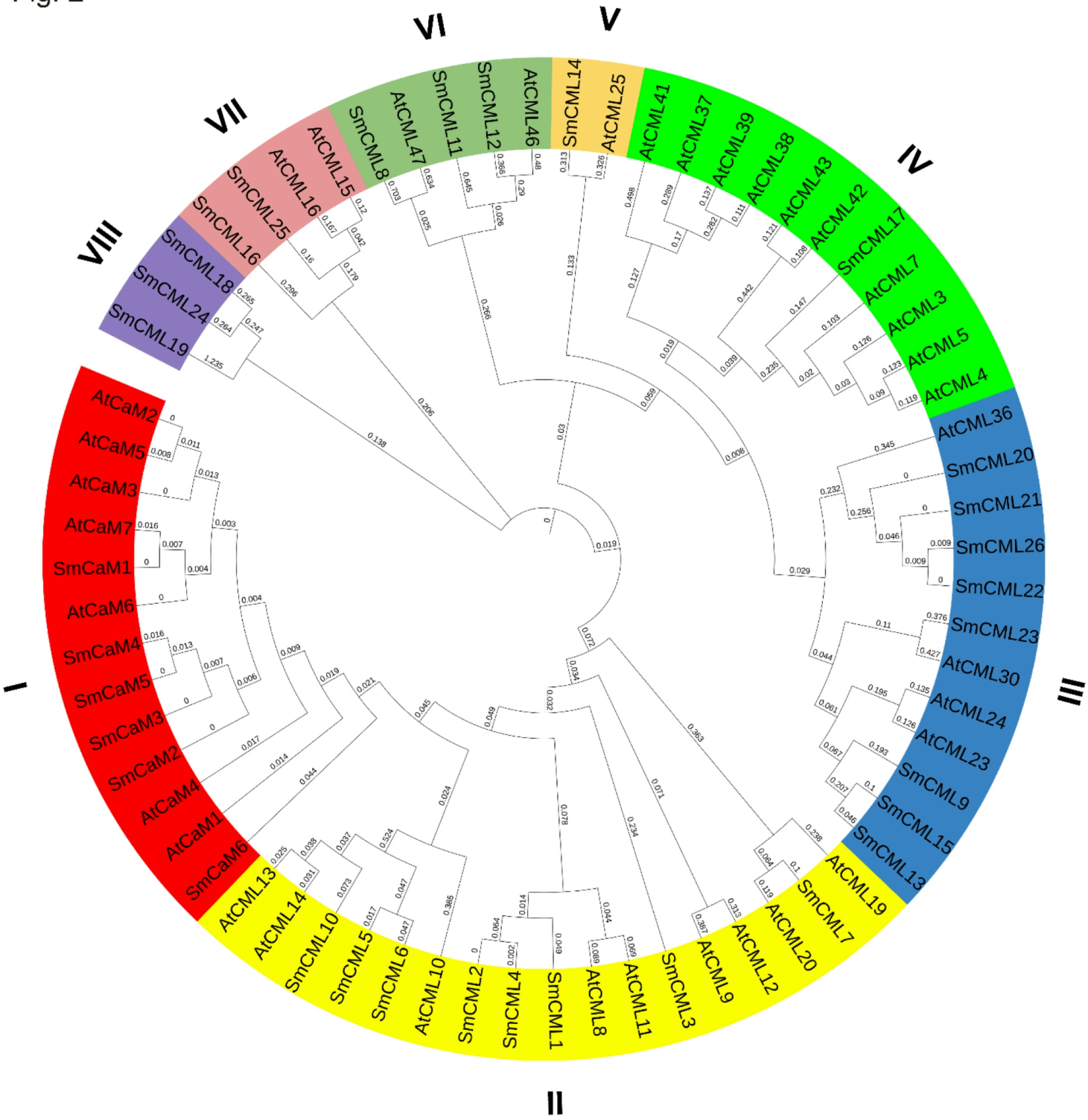


Figure 2

Phylogenetic relationships of CaMs and CMLs in *Salvia miltiorrhiza*and *Arabidopsis thaliana*. An unrooted phylogenetic tree was constructed via MEGA11 on the basis of the full-length amino acid sequences of seven *AtCaMs*and six *SmCaMs*, as well as their corresponding *CML* proteins. The tree was inferred via the neighbor—joining (NJ) method with 1,000 bootstrap replicates to assess node reliability. Pairwise deletion was applied to manage alignment gaps or missing data. The colored sectors distinguish different subfamilies identified in the tree, and the gene names are labeled accordingly. The numeric values shown on the branches represent the branch lengths, reflecting the evolutionary distances between sequences.

Fig. 3

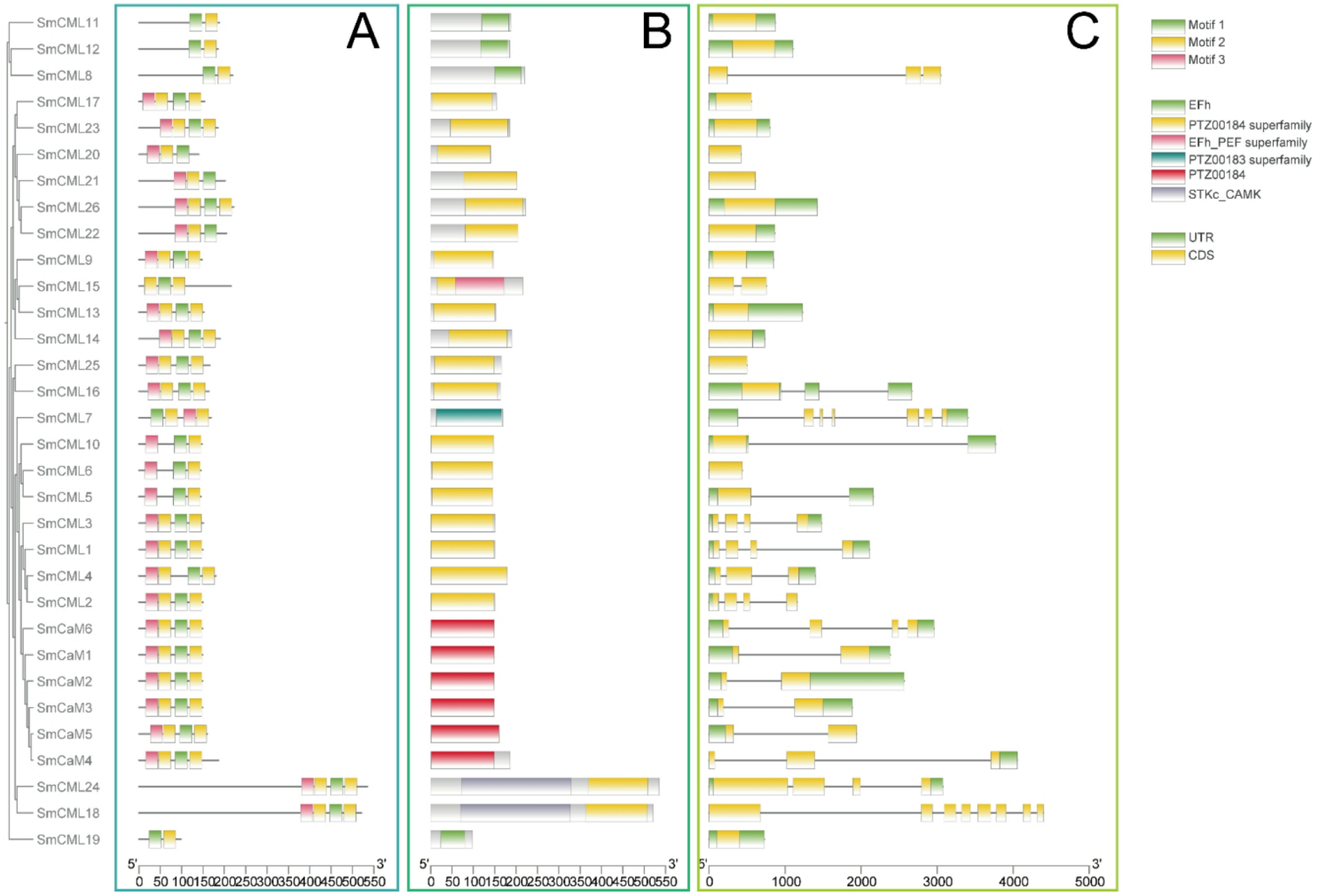


Figure 3

Structural analysis of the CaM and CML genes in *Salvia miltiorrhiza* and *Arabidopsis thaliana*. (A) Distribution of conserved motifs in the CaM and CML proteins from *S. miltiorrhiza* and *A. thaliana*, identified via the MEME Suite, with a maximum of 10 motifs and motif widths ranging from 6--50 amino acids. Both shared and unique motif patterns among family members are illustrated. (B) Predicted functional domain architectures of the CaM and CML proteins in both species, as determined via NCBI CD-Search and InterProScan. The figure highlights the presence, number, and arrangement of EF-hand motifs and other structural domains. (C) Exon—intron structures of the CaM and CML genes in *S. miltiorrhiza* and *A. thaliana*, visualized with TBtools, showing differences in gene organization, including exon count, intron presence, and distribution patterns. The subfigure labels are indicated in the top-left corner of each panel.

Fig. 4

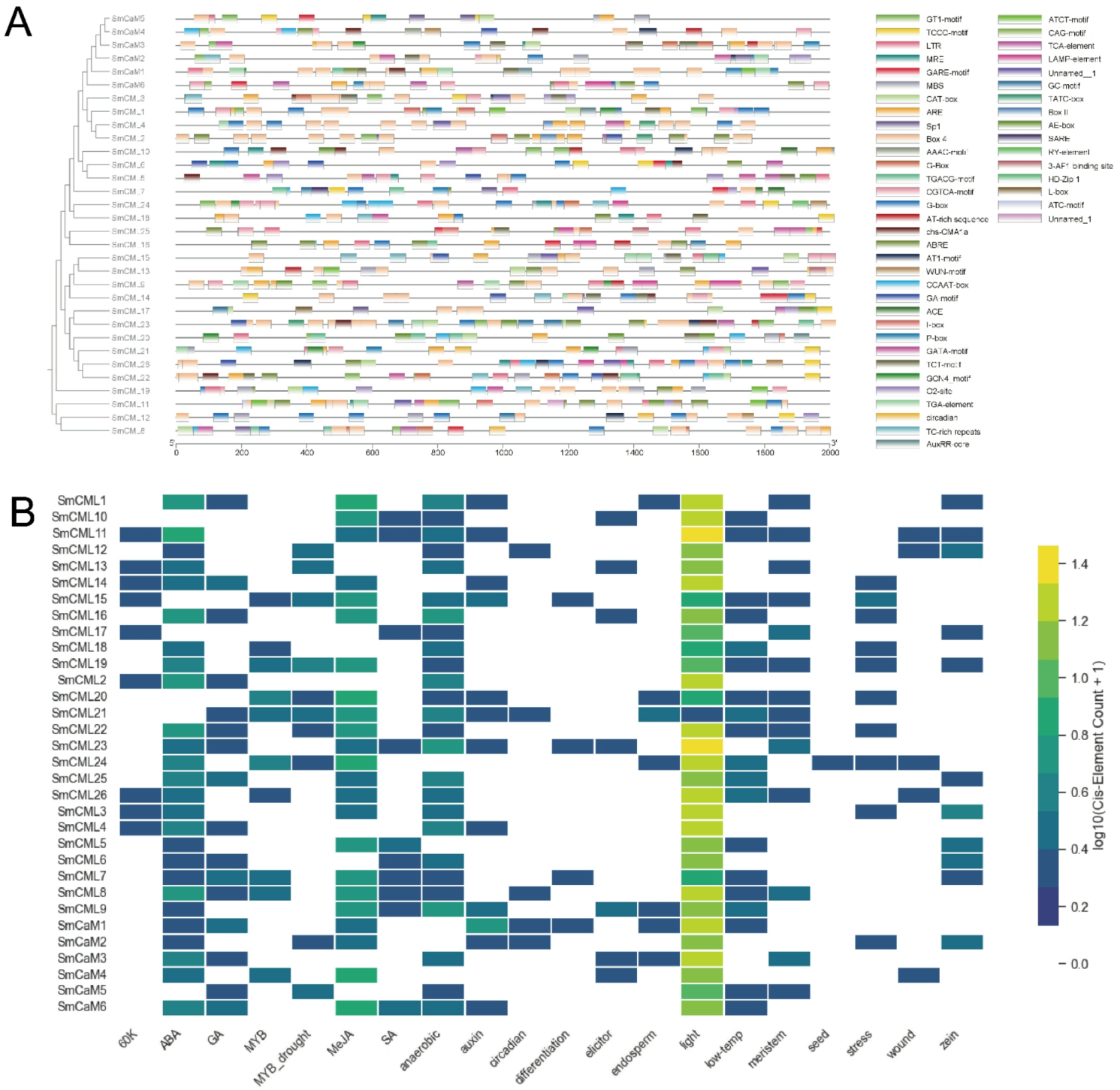


Figure 4

Predicted cis-acting elements in the promoter regions of the SmCaM and SmCML genes in *Salvia miltiorrhiza*. (A) Spatial distribution of cis-regulatory elements along the 2,000 bp upstream promoter regions of the SmCaM and SmCML genes. Each color corresponds to a specific type of cis-acting element, as indicated in the legend. (B) Heatmap depicting the abundance of various cis-regulatory element categories within the same promoter regions. Each row represents a gene, and each column represents a cis-element category, with color intensity indicating the number of elements detected per category. Cis-element prediction was performed via the PlantCARE database

Fig. 5

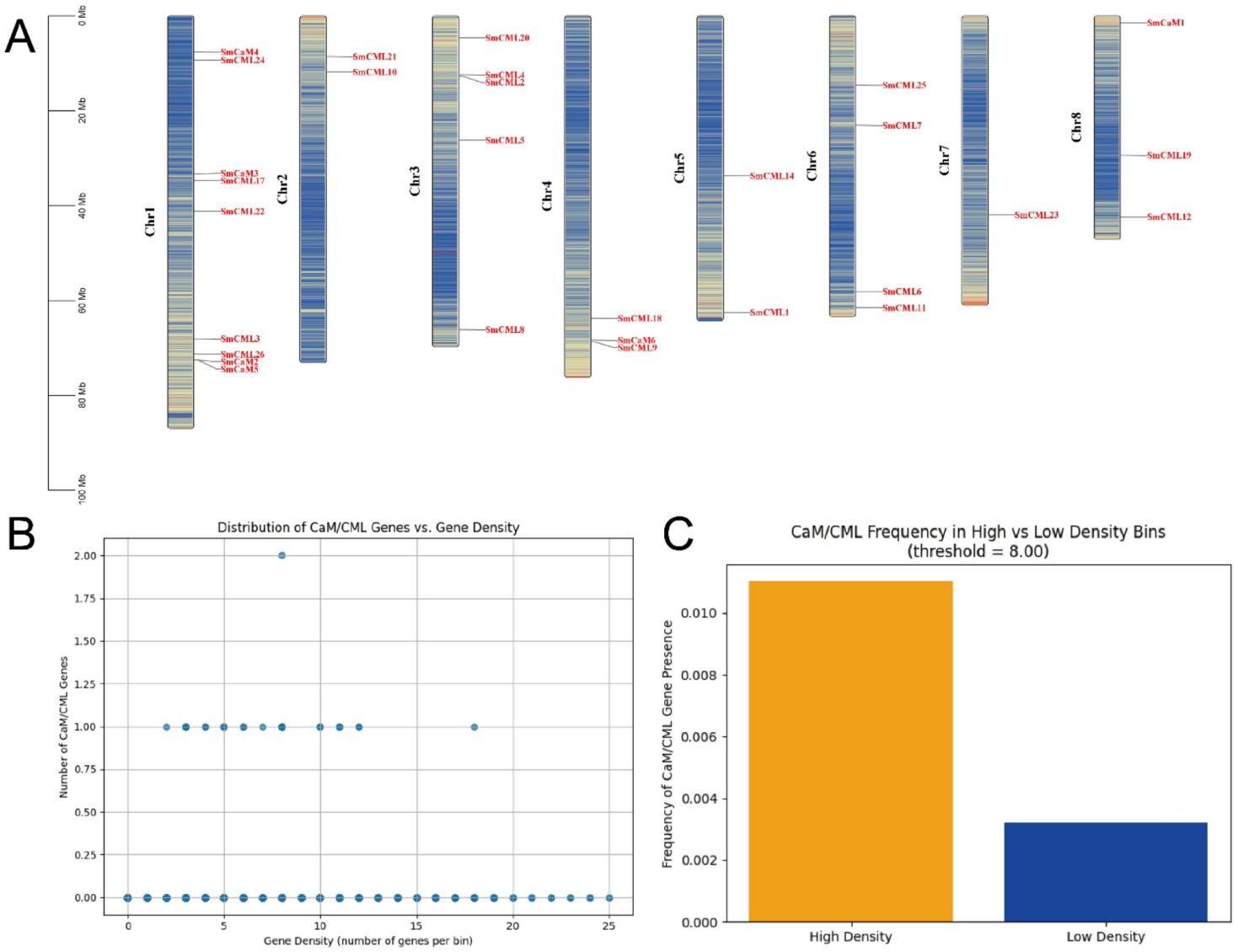


Figure 5

Chromosomal localization and distribution characteristics of the SmCaM and SmCML genes in *Salvia miltiorrhiza*. (A) Physical positions of the SmCaM and SmCML genes mapped onto the eight assembled chromosomes of *S. miltiorrhiza*, visualized via TBtools. Gene locations are shown alongside chromosome-wide gene density tracks to highlight regions of varying gene density. (B) Scatter plot illustrating the distribution of SmCaM and SmCML genes relative to genomic regions of different gene densities. Each point represents a gene, positioned according to its genomic bin's density classification. (C) Histogram depicting the frequency of SmCaM and SmCML genes located in high-density versus low-density genomic regions. High-density regions were defined as bins exceeding the third quartile (Q3) of gene counts per bin. A significant enrichment of SmCaM/CML genes in gene-rich regions was observed ($p = 0.0024$) on the basis of 10,000 permutation tests

Fig. 6

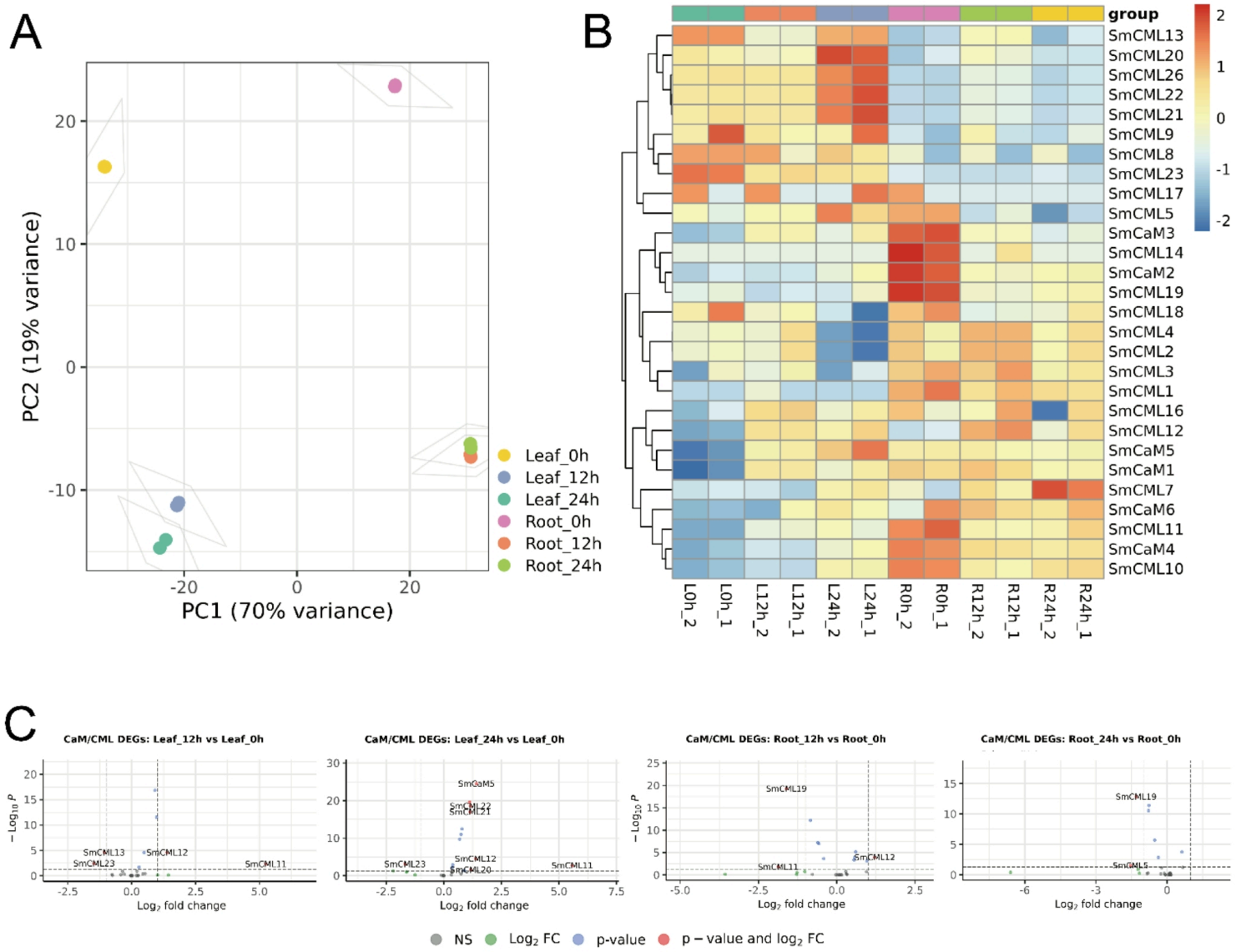


Figure 6

Expression profiling of the SmCaM and SmCML genes in *Salvia miltiorrhiza* under abscisic acid (ABA) treatment. (A) Principal component analysis (PCA) of transcriptomic profiles from *S. miltiorrhiza* leaf and root tissues subjected to ABA treatment for 0 h, 12 h, and 24 h. The first principal component (PC1) accounts for 70% of the total variance, whereas the second principal component (PC2) explains 19%. Six experimental groups, each with two biological replicates, are represented by distinct symbols and colors, demonstrating clear separation between treated and control samples as well as tissue-specific expression patterns. (B) Heatmap visualization of the differential expression of SmCaM and SmCML genes across different ABA treatment conditions in leaf and root tissues. The heatmap was generated via the R package *pheatmap*. Red denotes higher transcript abundance, whereas blue indicates lower expression levels. Genes are shown on the right Y-axis, while treatment groups are displayed along the X-axis, with biological replicates indicated by consistent color coding. (C) Volcano plots illustrating the differential expression of the SmCaM and SmCML genes under ABA treatment, comparing treated samples to untreated controls. The four panels correspond to (i) root tissue after 12 h of ABA treatment, (ii) root tissue after 24 h of ABA treatment, (iii) leaf tissue after 12 h of ABA treatment, and (iv) leaf tissue after 24 h of ABA treatment. Significantly differentially expressed genes were identified via DESeq2, with thresholds of $|\log_2\text{-fold change}| > 1$ and adjusted p value < 0.05. Notably, several SmCML genes, particularly in leaf tissues after 24 h of ABA exposure, exhibited substantial upregulation, suggesting potential roles in ABA-mediated stress response pathways.