

RESEARCH

Open Access



Genome-wide identification and expression analyses of *CYP450* genes in *Chrysanthemum indicum*

Yuyuan Chen^{1†}, Tongjun Zhou^{1†}, Jian Zhong², Yuxian Xu¹, Peng Zhang¹, Xiaoyu Yue¹, Hua Zhang¹, Ming Sun^{1*} and Xuehao Fu^{1*}

Abstract

Background The cytochrome P450 superfamily comprises a large group of enzymes crucial for the biosynthesis and metabolism of diverse endogenous and exogenous secondary metabolites in plants. *Chrysanthemum*, an ornamental genus with considerable medicinal value, is one of the most economically important floricultural crops in the world. The characteristics and functions of *CYP450* genes in *Chrysanthemum* species, however, remain largely unknown.

Results In this study, we identified 371 *CYP450* genes in the *Chrysanthemum indicum* genome, and categorized them into 8 clans and 44 families through phylogenetic analysis. Gene duplication analysis revealed 111 genes in 47 tandem duplicated clusters and 28 genes in 15 syntenic blocks, suggesting that extensive duplication events may account for the rapid expansion of *CYP450* superfamily. Additionally, extensive variations in gene structure, motif composition, and *cis*-regulatory element likely enhance the functional diversity of *CYP450* proteins. Volatile metabolomic analysis detected a total of 53 distinct volatile organic compounds across the leaves, stems, and roots of *C. indicum*, with 19 and 16 compounds being exclusive to leaves and stems, respectively. Transcriptomic analysis identified 248 expressed *CYP450* genes, with 31, 40, and 88 specifically or preferentially expressed in leaves, stems, and roots, respectively. Further correlation analyses between gene expression levels and compound contents highlighted 36 candidate *CYP450* genes potentially responsible for the biosynthesis of 47 volatile organic compounds.

Conclusions The genome-wide analyses of cytochrome P450 superfamily offers essential genomic resources for functional studies of *CYP450* genes, and is significant for the molecular breeding of *Chrysanthemum*.

Keywords *CYP450*, *Chrysanthemum indicum*, Phylogeny, Volatile metabolome, Gene expression

[†]Yuyuan Chen and Tongjun Zhou contributed equally to this work.

*Correspondence:

Ming Sun
sunming@bjfu.edu.cn
Xuehao Fu
xuehaofu@bjfu.edu.cn

¹State Key Laboratory of Efficient Production of Forest Resources, National Engineering Research Center for Floriculture, Beijing Key Laboratory

of Ornamental Plants Germplasm Innovation and Molecular Breeding, Beijing Laboratory of Urban and Rural Ecological Environment, Key Laboratory of Genetics and Breeding in Forest Trees and Ornamental Plants of Ministry of Education, School of Landscape Architecture, Beijing Forestry University, Beijing 100083, China

²Key Laboratory of Southwest China Wildlife Resources Conservation (Ministry of Education), College of Life Science, China West Normal University, Nanchong 637009, China



Background

Cytochromes P450 (CYP450s) are a superfamily of heme-thiolate enzymes found across all kingdoms of life [1]. These versatile enzymes can catalyze a wide range of biochemical reactions, including monooxygenation, oxidative rearrangement of carbon skeletons, and oxidative C-C bond cleavage, and thus transform endogenous or exogenous compounds into diverse secondary metabolites [2, 3, 4]. In plants, these metabolites include membrane sterols, phytohormones (e.g., gibberellins, strigolactones and abscisic acid), biopolymers (e.g., lignin and cutin), pigments (e.g., flavonoids and carotenoids), volatile organic compounds (e.g., terpenoids and fatty acid derivatives), and a wide range of lineage-specific secondary metabolites (e.g., taxol in *Taxus* and berberine in the Ranunculales) [3, 5, 6, 7]. These chemical components are crucial for plant development and adaptation to the environment. For example, lignin strengthens the cell wall, providing structural support, while cutin forms protective barriers to prevent water loss and protect against desiccation [8, 9]. Furthermore, compounds like cyanogenic glucosides, alkaloids, flavonoids, and terpenoids are essential for plant interactions with biotic and abiotic factors [10, 11, 12, 13].

The advent of sequencing technology has greatly accelerated the research progress of plant genomes, as well as the identification of CYP450 genes. To date, over 32,000 CYP450s, grouped into 670 families, have been named and annotated across the plant kingdom, with more than 800 of them being functionally characterized [3]. Notably, the CYPomes (i.e., complete set of CYP450s in a given species) vary widely between species due to frequent gene duplication and loss events [14]. For instance, only 39 CYP genes were detected in the green algae *Chlamydomonas reinhardtii*, whereas those in angiosperms account for up to 1% of protein-coding genes, with 236 in *Arabidopsis thaliana*, 309 in poplar (*Populus trichocarpa*), and 305 in rice (*Oryza sativa*) [15, 16]. The land plant CYP450s are divided into 11 clans based on homology and phylogeny: seven single-family clans (CYP51, CYP74, CYP97, CYP710, CYP711, CYP727, and CYP746) and four multi-family clans (CYP71, CYP72, CYP85, and CYP86) [3, 5]. Structurally, only a few protein sequences are highly conserved, including the heme-binding motif FxxGxRxCxG (CXG motif) in C-term, the ExxR motif in the K-helix, the PxRx motif and the oxygen binding and catalysis motif (A/G)Gx(E/D)T(T/S) in the I-helix [17]. The extensive diversity in the number and structure of plant CYP450s provide a molecular basis for the variability in metabolite profiles, which may lead to differences in morphological or developmental traits, and enable adaptation to specific ecological niches.

The genus *Chrysanthemum* (Asteraceae) consists of approximately 40 species and more than 30,000 cultivars,

most of which are cultivated as ornamental plants, with significant economic and cultural values [18]. Beyond their visual appeal, *Chrysanthemum* species are also renowned for their medicinal and culinary uses, with various plant parts being used in traditional medicine and tea beverage [19, 20, 21]. In *Chrysanthemum* species, the pharmacological activities are primarily linked to their component of flavonoids, alkaloids, and sesquiterpene lactones, while the distinctive floral scent is largely attributed to terpenoids, encompassing camphor, α -pinene, cineole, and caryophyllene [22, 23, 24, 25]. Although studies on closely related species highlighted the role of CYP450 genes in the biosynthesis of several important secondary metabolites, such as the artemisinin in *Artemisia annua* (CYP71AV1), the pyrethrin I in *Tanacetum cinerariifolium* (CYP82Q3), and the parthenolide in *T. parthenium* (CYP71CB1), researches on the characteristics and functions of CYP450s in *Chrysanthemum* remain limited [26, 27, 28]. Therefore, a comprehensive genome-wide identification and expression analysis of CYP450 gene family in *Chrysanthemum* is urgently needed for the effective gene utilization in breeding and metabolic engineering.

Chrysanthemum indicum, a diploid species ($2n = 2x = 18$) with the genome size of approximately 3.11 Gb, has been developed as a model of the genus *Chrysanthemum* [29, 30]. In this study, we identified and classified the CYP450 genes in the genome of *C. indicum* through phylogenetic analysis, and revealed their characteristics in physicochemical properties, gene structures, motif compositions, gene duplication events, and cis-regulatory elements. Furthermore, by integrating volatile metabolomic and transcriptomic data, we further explored the mechanisms underlying the volatile organic compounds of *C. indicum*. Our findings not only provide insights into the functions of CYP450 genes in *C. indicum*, but also offer valuable genetic resources for future molecular breeding efforts.

Methods

Plant materials

C. indicum was identified and kindly shared by Prof. Chao Ma in China Agricultural University [30], and conserved at the National Engineering Research Center for Floriculture, Beijing Forestry University, China. Voucher specimens (BJFC 00115566) were deposited at the Beijing Forestry University Herbarium (BJFC), Beijing. The rooted cuttings of *C. indicum* were grown in greenhouse under a 12-h-light /12-h-dark photoperiod at 25 °C and 60% relative humidity.

Identification of CiCYP450 genes

To identify the CYP450 genes in the *C. indicum* genome (available at *Chrysanthemum Genome Database*, <http://>

[//210.22.121.250:8880/asteraceae/download/downloadPage](https://210.22.121.250:8880/asteraceae/download/downloadPage)) (accessed on 27 April 2024), the Hidden Markov Model (HMM) file for the conserved domain of *CYP450* genes (PF00067) was downloaded from Pfam database (<http://pfam.xfam.org/>) (accessed on 28 April 2024), and the HMMER v3.3.2 (accessed on 28 April 2024) was used to search the protein sequences containing this domain [31]. At the same time, we obtained the *CYP450* protein sequences of *Arabidopsis thaliana*, poplar, and rice from TAIR (<https://www.arabidopsis.org/>) (accessed on 29 June 2024) and Cytochrome P450 Homepage (<https://drnelson.uthsc.edu/>) (accessed on 29 June 2024), and used them as queries for a local BLAST search against the *C. indicum* genome, with an E-value $\leq 10^{-5}$. Candidate protein sequences, ranging from 300 to 600 amino acids, were further filtered using the Conservative Domain Database tool (CCD, <https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>) (accessed on 29 June 2024) [32]. Finally, a total of 371 *CYP450* genes were identified in the *C. indicum* genome.

Analysis of physicochemical properties of *CiCYP450* genes

To characterize the identified *CiCYP450* protein sequences, we used the ExPASy ProtParam tool (<https://web.expasy.org/protparam/>) (accessed on 1 July 2024) to predict their molecular weights, isoelectric points, and grand average of hydrophilicity [33], and the WoLF PSORT tool (<https://wolfsort.hgc.jp/>) (accessed on 1 July 2024) to predict their subcellular localization [34].

Phylogenetic analysis of *CiCYP450* genes

To investigate the phylogenetic relationships of the *CiCYP450* genes, the protein sequences of *CYP450* genes from *C. indicum*, *A. thaliana*, poplar, and rice were aligned using MAFFT v7.490, with the default parameters [35]. A maximum likelihood tree was then constructed using IQ-TREE 2.1.4-beta, with the parameters of “-bb 1000 -nt 128” [36].

Chromosomal localization of *CiCYP450* genes

The locations of *CiCYP450* genes on chromosomes were obtained from the annotation file of the *C. indicum* genome, and visualized through TBtools (accessed on 6 July 2024) [37].

Genome collinearity analyses of *CiCYP450* genes

The *Chrysanthemum seticuspe* genome was available at PlantGarden (<https://plantgarden.jp/>) (accessed on 17 May 2024), and the *A. thaliana* genome was downloaded from NCBI (<https://www.ncbi.nlm.nih.gov/>) (accessed on 17 May 2024). Genome collinearity analysis was performed using MCScanX, with default parameters (accessed on 6 July 2024 and on 12 January 2025) [38], and visualized using TBtools.

Conserved motif of *CiCYP450* genes

To analyze the conserved motifs of the *CiCYP450* proteins, the MEME tool (<https://meme-suite.org/meme/>) (accessed on 8 July 2024) was employed, with the number of motifs being set to 15 [39].

Cis-regulatory elements analyses of *CiCYP450* genes

To identify potential *cis*-regulatory elements of *CiCYP450* genes, the 2 kb upstream of the translation start site of *CiCYP450* genes was retrieved from *C. indicum* genome, and submitted to the PlantCARE (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) (accessed on 8 July 2024) [40]. All results were visualized using TBtools.

GC-MS analysis of volatiles in *C. indicum*

The volatiles from roots, stems, and leaves were detected using headspace-solid phase micro extraction-gas chromatography-mass spectrometry (HS-SPME-GC-MS), with three biological replicates each. Fresh samples were separately collected and sealed in 20 ml glass bottles, with ethyl decanoate (1:1000) added as an internal standard. A pre-conditioned SPME fiber (DVB/CAR/PDMS, Sigma-Aldrich) was then inserted into the vial's headspace, and incubated at 40 °C for 40 min. Following extraction, the fiber was desorbed in the GC injection port at 250 °C for 5 min. The GC-MS analysis was performed using Agilent 7890B-5977 A GC/MSD (America) equipped with a DB-5MS capillary column (30 m × 0.25 mm × 0.25 μm, Agilent). The GC program started at 40 °C for 2 min, then increased to 180 °C at a rate of 5 °C/min, and finally rose to 270 °C at a rate of 20 °C/min. The mass spectra were acquired within the scanning range of 50–400 m/z.

Compounds were identified by comparing spectra with the National Institute of Standards and Technology database (NIST, 2017) using Qualitative Navigator B.08.00 (accessed on 13 July 2024). Volatile compounds were quantified relative to the internal standard, and peak areas were normalized to percentages to calculate the amounts of volatiles, following the method of Feng et al. [41].

RNA-seq and expression analysis of *CiCYP450* genes

The fresh leaves, stems, and roots of *C. indicum* were collected and stored at –80 °C for preservation, with three biological replicates each. Total RNA was separately extracted using the RNAprep Pure Plant Plus Kit (TIANGEN). RNA concentration was measured using Qubit RNA Assay Kit in Qubit 3.0 Fluorometer (Life Technologies, Carlsbad, CA, USA) and RNA integrity was assessed using the RNA Nano 6000 Assay Kit of the Bio-analyzer 2100 system (Agilent Technologies, Santa Clara, CA, USA). Libraries were generated using VAHTS Universal V6 RNA-seq Library Prep Kit for Illumina® (Vazyme

Biotech Co., Ltd), and were sequenced on Illumina Nova-seq X plus platform (Illumina, San Diego, CA) with 150 bp read length of paired-end (PE) reads by Annoroad Gene Technology (Beijing, China). Each sample obtained approximately 6.0 Gb raw data. To obtain the clean reads, all raw reads were filtered using fastp version 0.23.4, with default parameters. The clean data of RNA-seq were mapped to the *C. indicum* genome using HISAT2 v2.21, with default parameters [42]. The TPM values were calculated using StringTie v2.21 [43]. Correlation analyses between the gene expression levels and the contents of volatile compound was performed using the R v4.3.2 package *corrplot*.

KEGG enrichment analysis of *CiCYP450* genes

The annotation information of *CiCYP450s* for KEGG enrichment was downloaded from eggNOG-mapper (<http://eggno-mapper.embl.de/>) (accessed on 27 August 2024) [44]. Then we obtained the KEGG backend background files from TBtools. The KEGG path was visualized by TBtools.

Results

Identification and phylogenetic analysis of *CYP450* genes in *Chrysanthemum indicum*

A total of 371 *CYP450* genes were identified in the genome of *C. indicum*, with protein lengths ranging from 300 to 600 amino acids and molecular weights between 34.02 and 69.34 kDa (Table S1). The *CiCYP71-23* and *CiCYP71-24* protein have the lowest isoelectric points (PI = 4.96), while the *CiCYP71-44* protein has the highest (PI = 9.82) (Table S1). The proteins also have different levels of hydrophilicity, ranging from -0.42 of *CiCYP71-21* and *CiCYP76-4* to 0.307 of *CiCYP71-1* and *CiCYP75-1* (Table S1). Subcellular localization predictions showed that most of proteins were localized in chloroplast (223), followed by cytosol (65), plasma membrane (30), nucleus (13), endoplasmic reticulum (11), vacuolar membrane (9), extracellular (7), peroxisome (6), mitochondrion (6), and cytoskeleton (1) (Table S1).

To investigate the evolutionary relationships of *CiCYP450s*, the protein sequences of *CYP450* genes from *C. indicum*, *A. thaliana*, *O. sativa*, and *P. trichocarpa* were aligned to construct the phylogenetic tree. The analysis revealed that the *CiCYP450* genes could be divided into 8 clans and 44 families: clan710 (i.e., *CYP710*), clan85 (i.e., *CYP716*, *CYP90*, *CYP724*, *CYP720*, *CYP87*, *CYP85*, *CYP88*, *CYP707*, *CYP728*, *CYP729*, *CYP733*, and *CYP718*), clan74 (i.e., *CYP74*), clan711 (i.e., *CYP711*), clan97 (i.e., *CYP97*), clan72 (i.e., *CYP714*, *CYP715*, *CYP735*, *CYP749*, *CYP721*, *CYP734*, and *CYP72*), clan86 (i.e., *CYP94*, *CYP704*, *CYP96*, and *CYP86*), and clan71 (i.e., *CYP89*, *CYP77*, *CYP701*, *CYP79*, *CYP78*, *CYP73*, *CYP98*, *CYP76*, *CYP80*, *CYP706*, *CYP82*, *CYP81*, *CYP93*,

CYP75, *CYP84*, *CYP71*, and *CYP92*) (Fig. 1, Fig. S1). Among them, the *CYP71*, *CYP706*, and *CYP72* were the top three largest families with 58, 40, and 33 members, respectively (Fig. 1, Table S2). Conversely, the smallest families, each contained only one member, were *CYP85*, *CYP98*, *CYP711*, *CYP715*, *CYP721*, *CYP724*, *CYP733*, *CYP734* and *CYP735* (Fig. 1, Table S2). Notably, families involved in biosynthesis of flavonoids, alkaloids, and/or terpenoids, such as *CYP82*, *CYP706*, *CYP72*, and *CYP80*, were more abundant in *C. indicum* compared to the other three species (Table S2).

Chromosomal locations and gene duplication analysis of *CiCYP450* genes

To reveal the evolutionary mechanisms of *CiCYP450* genes, we first analyzed their chromosomal locations in *C. indicum* genome. Our analysis showed that 6 genes were located in unplaced contigs and 365 *CYP450* genes were unevenly distributed across nine chromosomes, with gene counts ranging from 16 on chromosome 8 to 56 on chromosome 4 (Fig. 2). Moreover, these genes were predominantly clustered near the chromosomal ends, resulting in 47 tandem duplicated clusters (Fig. 2, Table S3). Of these, 35 clusters had two genes and 12 clusters contained three or more members, with two largest cluster composed of five members (*CiCYP80-6* to *CiCYP80-8*, *CiCYP92-2* to *CiCYP92-8*) located on chromosome 4 (Fig. 2). The tandem duplicated genes spanned 17 families, including 18 in *CYP706*, 14 in *CYP71*, 12 in *CYP72*, 10 in *CYP80*, 7 in *CYP96*, *CYP92* and *CYP89*, 6 in *CYP82*, 5 in *CYP94* and *CYP704*, 4 in *CYP84*, *CYP729* and *CYP76*, and 2 each in *CYP93*, *CYP81*, *CYP707*, and *CYP718* (Fig. 2). These findings suggest that tandem duplication has played an important role in the expansion of the *CiCYP450* supergene family.

In addition to tandem duplication, segmental duplication has been proposed as another driver of gene family expansion. Through a collinearity analysis of the *C. indicum* genome, we further identified 16 pairs of segmentally duplicated *CiCYP450* genes within 15 syntenic blocks (Fig. 3, Table S4). These syntenic blocks were distributed across eight chromosomes, with the exception of chromosome 5, where no segmental duplicated genes were observed (Fig. 3). The 16 segmentally duplicated gene pairs were belonging to 11 families, including three pairs from the *CYP84* family, two pairs from the *CYP71*, *CYP82* and *CYP749* family, and one pair each from the *CYP77*, *CYP78*, *CYP90*, *CYP92*, *CYP94*, *CYP701*, and *CYP707* families (Fig. 3). These results together suggest that at least 35.58% (132/371) *CiCYP450* genes originated from duplication events, offering opportunities for the acquisition of new gene functions or the refinement of existing ones.

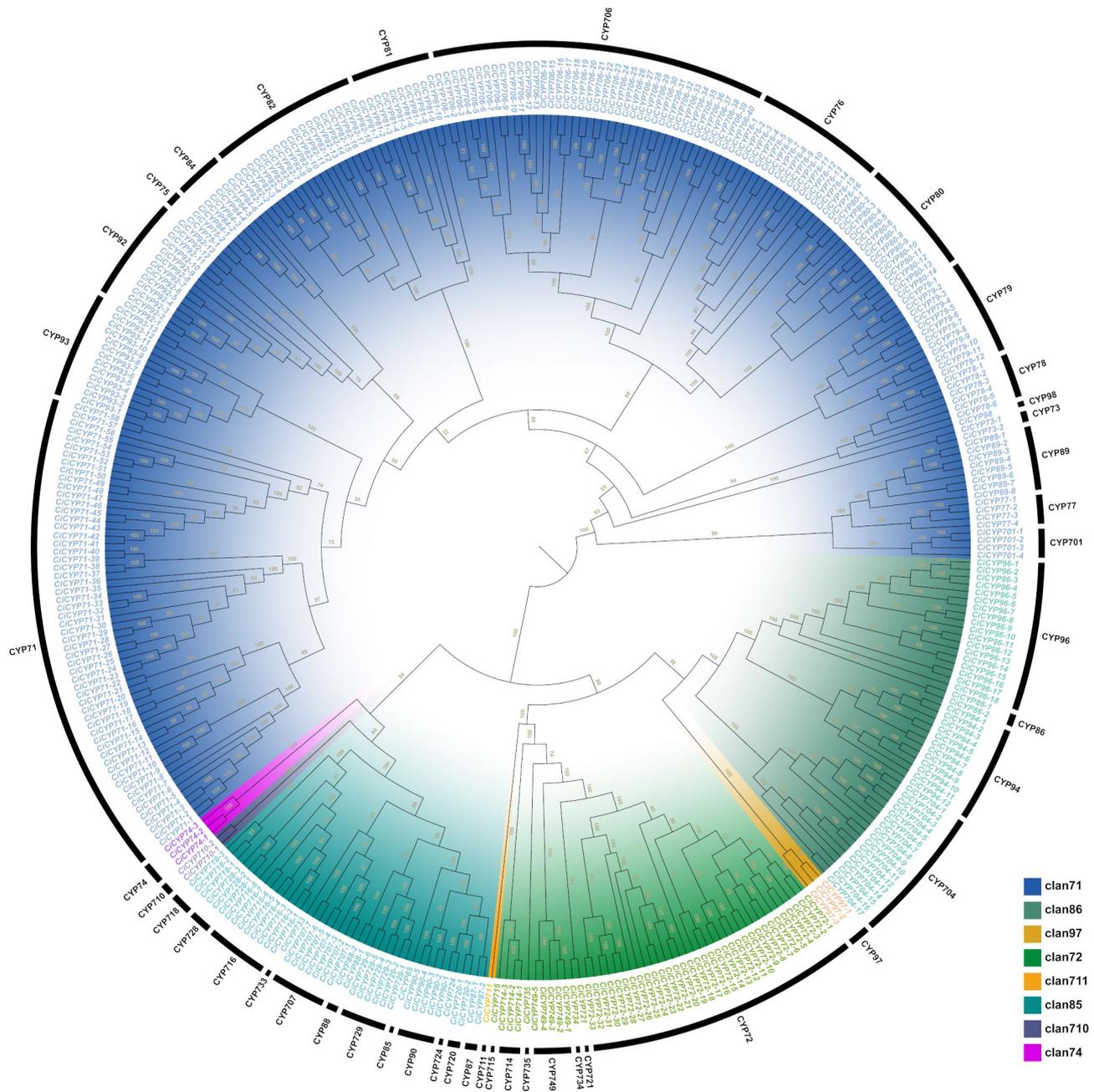


Fig. 1 Phylogenetic trees of *CiCYP450* genes. Phylogenetic tree is constructed by using ML algorithms, with 1000 bootstrap replicates. Different clans are represented by different colors

To further investigate the origin and evolutionary relationship of *CiCYP450s*, we performed collinearity analyses among *C. indicum*, *C. seticuspe* and *A. thaliana* (Fig. S3). The analyses showed that *C. indicum* shared 213 collinear genes with *C. seticuspe* and 47 with *A. thaliana* (Tables S5-S6), consistent with the closer evolutionary relationship between *C. indicum* and *C. seticuspe* compared to *C. indicum* and *A. thaliana*. In addition, we found that there were 22 and 7 *CYP450* genes with multiple copies in *C. indicum* were identified as homologous

to a single *CsCYP450* or *AtCYP450* gene, respectively, suggesting that these *CiCYP450* genes likely expanded after the divergence of *C. indicum* with *C. seticuspe* and *A. thaliana*.

Gene structure and conserved motif analyses of *CiCYP450* genes

To understand the structural characteristics and diversity of the *CiCYP450s*, we conducted a detailed analyses of the gene structure and motif composition of all

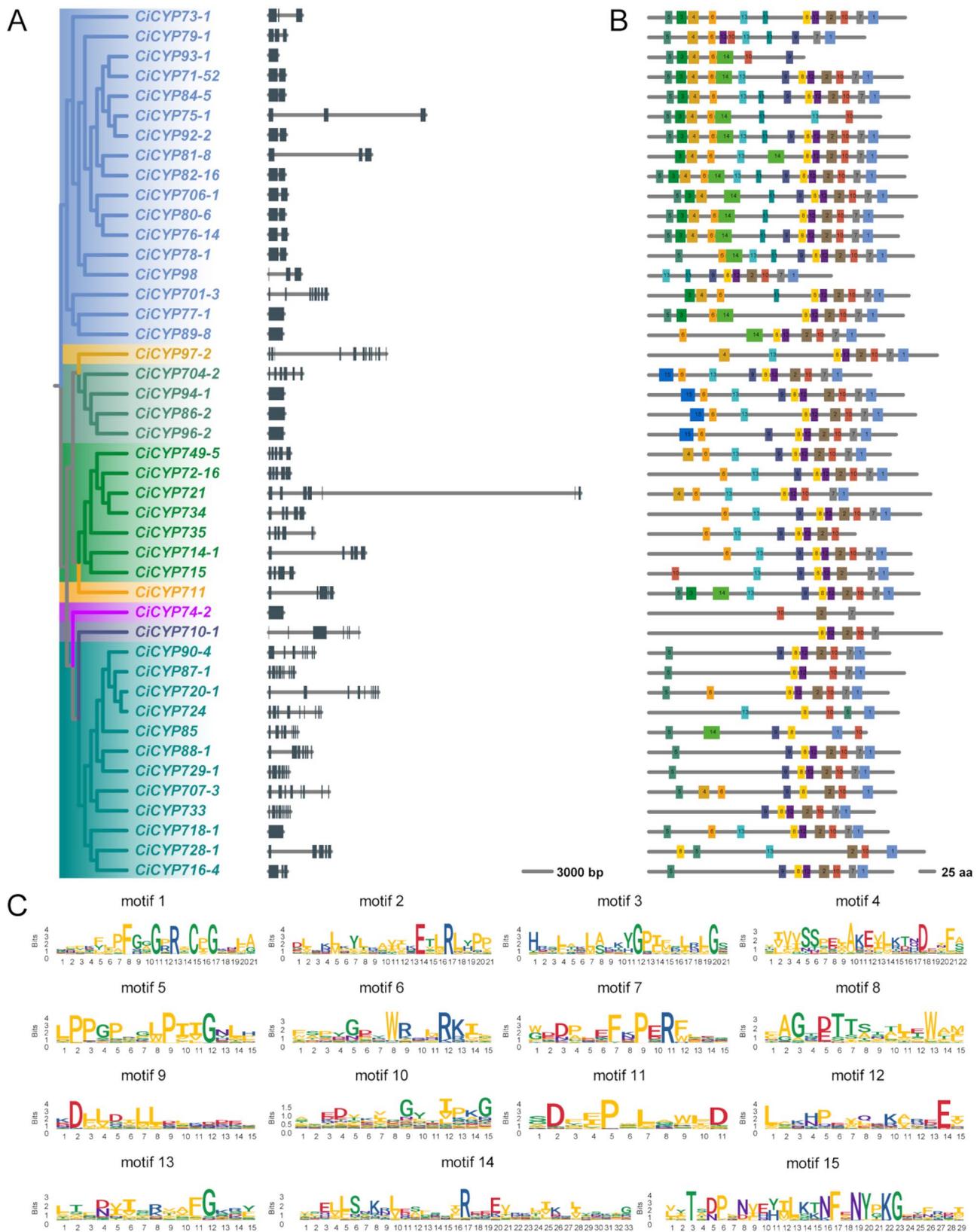


Fig. 4 Phylogenetic relationships, gene structures, and conserved motifs of representative *CiCYP450* genes in 44 families. The gray boxes represent CDS, and the colored and numbered boxes represent different motifs

genes, the promoter regions (2 kb upstream of the translation start site) were retrieved and analyzed using PlantCARE. A total of 25 putative *cis*-elements related to plant growth and development, stress response, and phytohormone response were identified across the 371 *CiCYP450* genes (Fig. S2). The *cis*-elements involved in plant growth and development included those for zein metabolism regulation (O2-site), meristem expression (CAT-box), endosperm expression (GCN4-motif and AACA-motif), circadian control (circadian), seed-specific regulation (RY-element), palisade mesophyll cells (HD-Zip 1), and cell cycle regulation (MSA-like) (Fig. S4). The stress response elements included those involved in anaerobic induction (ARE), low temperature responsiveness (LTR), drought inducibility (MBS), defense and stress responsiveness (TC-rich repeats), and anoxic specific inducibility (GC-motif) (Fig. S4). For phytohormone responses, *cis*-elements included MeJA responsive (CGTCA-motif and TGACG-motif), ABA responsive (ABRE), gibberellin responsive (TATC-box, GARE-motif, and P-box), auxin responsive (TGA-element, TGA-box, AuxRE, and AuxRR-core), and SA responsive (SARE and TCA-element) (Fig. S4). Our findings suggest that *CiCYP450* genes may play significant roles in regulating plant growth and development, as well as in responding to hormonal signals and both abiotic and biotic stresses.

Integrated volatile metabolomic and transcriptomic analyses of *CiCYP450* genes

The *Chrysanthemum* species are distinguished by their unique scent, setting them apart from other ornamental plants. To explore the physical and genetic basis of the chrysanthemum's unique aroma, we conducted integrated volatile metabolomic and transcriptomic analyses focusing on *CiCYP450* genes. Using HS-SPME coupled with GC-MS, we identified 53 distinct volatile organic compounds (VOCs) from *C. indicum* (Fig. 5A–B, Table S7). These VOCs could be categorized into three groups: 40 terpenoids (including 18 monoterpenoids and 22 sesquiterpenoids), 6 aromatic hydrocarbons, and 7 fatty acid derivatives (Fig. 5A, Table S7). Among the analyzed organs, 33 volatiles were detected in both leaves and stems, while only 10 were found in the roots, with most compounds being organ-specific. For instance, 19 compounds (e.g., 2-isopropyltoluene, 1,8-cineole, and calamenene) were unique to leaves, and 16 volatiles (e.g., β -sesquiphellandrene, α -zingiberene, and β -phellandrene) were exclusive to stems (Fig. 5A). Notably, even VOCs present in all three organs exhibited significant differences in concentration. For example, the content of caryophyllene in the stems was 67.01 $\mu\text{g/g}$ compared to only 2.78 $\mu\text{g/g}$ in leaves and 0.60 $\mu\text{g/g}$ in roots (Fig. 5A, Table S7).

We then analyzed the expression patterns of *CiCYP450* genes in leaves, stems and roots using RNA-seq data. Of 371 *CiCYP450* genes, 248 (66.85%) had a TPM value greater than 1 in at least one organ, indicating that they were expressed (Fig. 5C, Table S8). Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis showed that these genes mainly involved in the pathway of “metabolism of terpenoids and polyketides”, “isoflavonoid biosynthesis”, “cutin, suberine and wax biosynthesis”, and “diterpenoid biosynthesis” (Fig. 5D). Notably, we found 31, 40, and 88 genes were specifically (i.e., a gene whose TPM was ≥ 1.0 only in this organ but < 1.0 in any other organs) or preferentially (i.e., a gene whose TPM in this organ was at least two-fold higher than those in any other organs) expressed in leaves, stems, and roots, respectively (Fig. 5C, Table S8). Additionally, several genes were preferentially expressed in two of the three organs: 22 in leaves and stems, 5 in leaves and roots, and 25 in stems and roots (Fig. 5C, Table S8). The remaining 37 genes showed comparable expression levels across all three organs (Fig. 5C, Table S8). Interestingly, different members of the same family also displayed distinct expression patterns, suggestive of functional divergence. For example, *CiCYP706-13*, *CiCYP706-32*, *CiCYP706-35*, and *CiCYP706-36* were highly expressed in leaves, *CiCYP706-1*, *CiCYP706-2*, *CiCYP706-3*, and *CiCYP706-4* were preferentially expressed in stems, whereas *CiCYP706-12*, *CiCYP706-14*, *CiCYP706-15*, and *CiCYP706-20* were primarily expressed in roots (Fig. 5C).

The variations in VOC profiles and gene expression provide a valuable opportunity to identify candidate genes associated with the diverse volatile organic compounds in *C. indicum*. We therefore analyzed the correlations between gene expression levels and the contents of these VOCs. Among the 28 compounds that were either unique to or abundant in leaves, such as 2-isopropyltoluene, 1,8-cineole, germacrene D, and terpinen-4-ol, were positively correlated with 39 *CiCYP450* genes, including *CiCYP704-13*, *CiCYP706-35*, and *CiCYP71-23* (Fig. 6). Notably, the nine compounds (i.e., germacrene D, (-)- γ -cadinene, γ -terpinene, β -cadinene, copaene, pinocarvone, (+)-camphor, camphene, and terpinen-4-ol) detected in both leaves and stems, were also associated with additional 25 *CiCYP450* genes, including *CiCYP96-11*, *CiCYP96-13*, *CiCYP96-14*, *CiCYP72-5*, *CiCYP90-5*, and *CiCYP706-30* (Fig. 6). Specially, methyl isovalerate were associated with 9 *CiCYP450* genes, including *CiCYP82-5*, *CiCYP72-3*, *CiCYP74-3*, *CiCYP82-12*, and *CiCYP72-26* (Fig. 6). Similarly, the 16 volatiles exclusive to or enriched in stems, like β -sesquiphellandrene, α -zingiberene, β -phellandrene, and terpinolene, showed positive correlations with 43 *CiCYP450* genes, such as *CiCYP72-17*, *CiCYP706-4*, and *CiCYP77-4* (Fig. 6). The four compounds (i.e., myrcene,

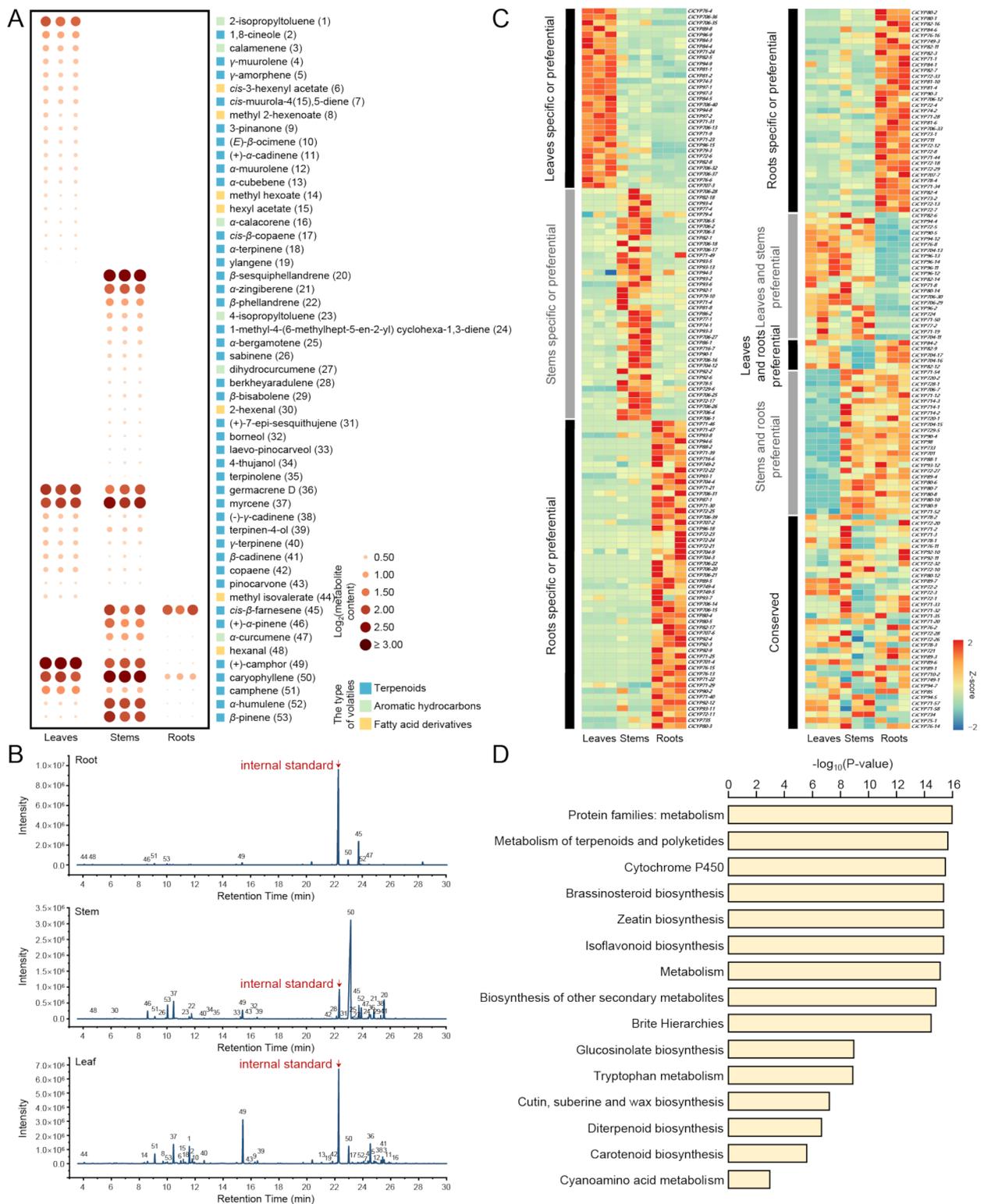


Fig. 5 Volatile organic compounds and expression patterns of *CiCYP450* genes. **(A)** Heatmap of 53 volatile organic compounds in roots, stems, and leaves of *C. indicum*. **(B)** Total ion chromatogram of different tissues in *C. indicum* volatile components. **(C)** Expression patterns of 248 *CiCYP450* genes in roots, stems, and leaves of *C. indicum*. **(D)** KEGG pathway enrichment analysis of 248 expressed *CiCYP450* genes

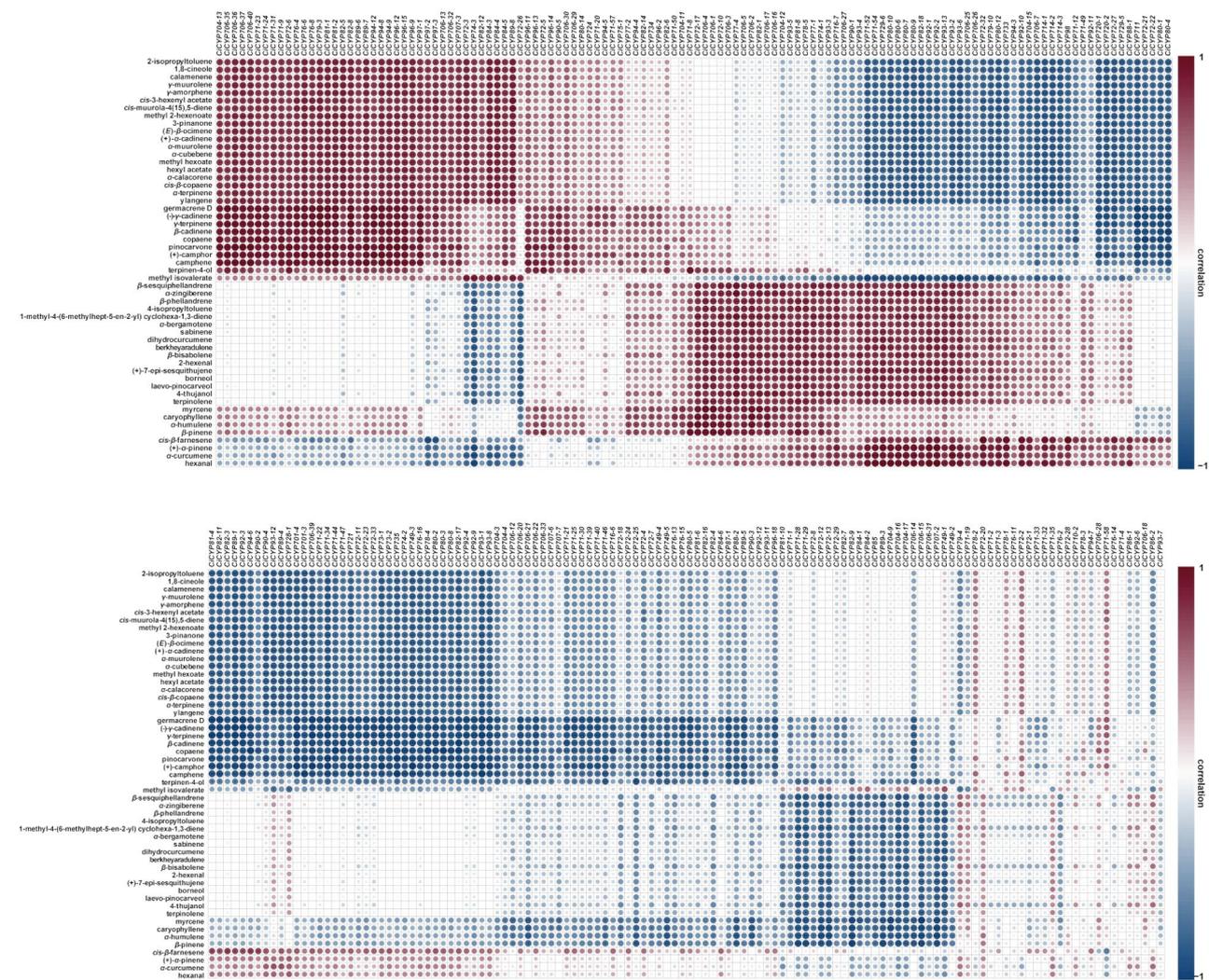


Fig. 6 Correlation analyses between the expression levels of *CYP450* genes and the contents of volatile compounds

caryophyllene, α -humulene, and β -pinene) were associated with 31 *CiCYP450* genes, including *CiCYP72-5*, *CiCYP71-8*, and *CiCYP72-17* (Fig. 6). And the four compounds (i.e., *cis*- β -farnesene, (+)- α -pinene, α -curcumene, and hexanal) detected in both stems and roots, were associated with 56 *CiCYP450* genes, including *CiCYP93-5*, *CiCYP81-8*, and *CiCYP78-5* (Fig. 6).

Discussion

Cytochrome P450 superfamily represents the largest enzyme family that plays crucial roles in metabolic diversification during plant evolution [5, 45]. Gene identification, classification, and expression analysis are essential for understanding the functions of this extensive gene family. With the increase of whole-genome sequences, *CYP450* genes have been identified in various plants, such as tea plant (*Camellia sinensis*, 273 genes), citrus (*Citrus clementina*, 301 genes), peanuts (*Arachis hypogaea*, 589

genes), sweet potatoes (*Ipomoea batatas*, 95 genes), pears (*Pyrus* spp., 419 genes), and *Brassica napus* (384 genes) [46, 47, 48, 49, 50, 51]. However, little is known about the *CYP450* gene family in *C. indicum*, a model species of the genus *Chrysanthemum*. In this study, we identified a total of 371 *CYP450* genes in the genome of *C. indicum*, and revealed their phylogenetic relationships, molecular characteristics, and expression patterns. Furthermore, we examined correlations between the expression levels of 248 *CiCYP450* genes and the contents of 53 VOCs to gain insight into their potential functions.

In this study, 371 *CYP450* genes were identified in *C. indicum*, which were grouped into 8 clans and 44 families. As in most angiosperms [52, 53], the CYP71 clan constituted the largest *CYP450* clan in *C. indicum*, comprising 17 families and 227 members. In contrast, although the CYP51 clan is generally considered evolutionarily conserved [3, 54], no members of this clan were

detected in *C. indicum*, which is likely due to incomplete annotation of protein-coding genes. In addition, *C. indicum* have experienced extensive gene expansions, resulting in multiple copies across several gene families, such as 58 members in CYP71, 40 members in CYP706, 33 members in CYP72, and 18 members in CYP82. The expansion appears to be primarily driven by gene duplication events, as 47 tandem duplicated clusters and 15 syntenic blocks of *CiCYP450* genes were identified within the *C. indicum* genome. Gene duplication, along with variations in gene structure, motif composition, and *cis*-regulatory elements, likely led to a wide range of substrate specificities of *CiCYP450* proteins, and thus provided *C. indicum* opportunities to meet diverse adaptive requirements [3, 55]. For instance, the most expanded gene families (e.g., CYP71, CYP82, CYP706, and CYP72) were known to participate in the biosynthesis of flavonoids, alkaloids, and terpenoids, which coincidentally coupled with the pharmacological activities and distinctive scent of *Chrysanthemum* [3, 5, 56, 57].

Previous studies have identified the major constituents of *Chrysanthemum* essential oils as camphor, borneol, camphene, α -pinene, β -cymene, 1,8-cineole, and caryophyllene [22, 25, 58, 59]. Our GC-MS analysis confirmed the presence of all these compounds, except for β -cymene, indicating the high feasibility and reproducibility of our results. In addition to the primary constituents, our analysis also detected 6 aromatic hydrocarbons and 7 fatty acid derivatives within the volatile metabolome, which was not reported in the aforementioned studies. These variations in chemical composition could stem from several factors, including species-specific differences in secondary metabolite production or disparities in experimental methodologies, such as the sampling of different tissues or variations in extraction and analytical techniques.

It has been suggested that volatile terpenoids, including monoterpenes, sesquiterpenes and their oxygen-containing derivatives, are mainly catalyzed by members of the CYP71, CYP76 and CYP706 families [3, 5, 60]. Among the 53 volatiles detected in *C. indicum*, 40 were terpenoids, suggesting that terpenoids were the dominant volatile components of this species. Furthermore, correlation analyses between gene expression levels and compound contents revealed that 21 terpenoids (e.g., 1,8-cineole, germacrene D, (+)-camphor, and camphene) were positively correlated with 17 genes, including *CiCYP71-23*, *CiCYP76-4* and *CiCYP706-35* (Table S9) [61, 62, 63, 64, 65], while the remaining 19 terpenoids (e.g., β -sesquiphellandrene, sabinene, borneol, and *cis*- β -farnesene) were associated with 17 genes, including *CiCYP706-5* and *CiCYP71-50* (Table S9) [61, 66], indicating that members of the same clan may be responsible for the biosynthesis of different terpenoids. In addition

to terpenoids, seven fatty acid derivatives were also identified in *C. indicum*, indicating their contribution to unique scent of *C. indicum*. Previous studies have found that volatile fatty acid derivatives are primarily catalyzed by members of the CYP74 family [67, 68]. Accordingly, we hypothesized that the production of five fatty acid derivatives (i.e., *cis*-3-hexenyl acetate, methyl 2-hexenoate, methyl hexoate, hexyl acetate, and methyl isovalerate) was likely regulated by *CiCYP74-3* [69, 70] while the biosynthesis of 2-hexenal and hexanal were possibly controlled by *CiCYP74-1* [71], due to the high correlation between gene expression and compound contents. Our predictions suggest that many genes within the same family are capable of catalyzing the production of different volatile compounds, indicating functional divergence among homologous genes [72, 73]. Nonetheless, further functional studies are necessary to confirm the roles of individual *CiCYP450* genes.

Conclusions

In summary, our work presents a comprehensive genome-wide analysis of the cytochrome P450 superfamily in *C. indicum* genome, identifying 371 *CYP450* genes across 44 families within 8 clans. Tandem duplication and segmental duplication appear to be the major drivers for the rapid expansion of *CiCYP450* gene family. Gene duplication, together with the variations in gene structure, motif composition, and *cis*-regulatory elements, likely contributed to the functional diversity of *CiCYP450* proteins. Integrated volatile metabolomic and transcriptomic analyses not only revealed several tissue-specific compounds and genes, but also identified 36 *CiCYP450* genes potentially responsible for the biosynthesis of 47 volatile organic compounds. Our findings provide a valuable foundation for functional study of *CiCYP450* genes and will accelerate genetic improvement of *Chrysanthemum*.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-025-11664-0>.

Supplementary Material 1

Supplementary Material 2

Acknowledgements

We thank members of the Sun Laboratory for helpful discussions and Prof. Chao Ma for sharing plant material.

Author contributions

YC, XF, and MS designed the research and wrote the manuscript. YC, TZ, JZ, XF, and HZ conducted experiments and analyzed the data. LL, YX, PZ, and XY revised the paper.

Funding

This work was supported by the Fundamental Research Funds for the Central Universities (QNTD202306 and BLX202306), the Science, Technology

& Innovation Project of Xiongan New Area (2022XAGG0100), and National Natural Science Foundation of China (32271947).

Data availability

The RNA-seq data have been deposited in NCBI Short Read Archive with accession number PRJNA1167174. All other data supporting the findings of this study are available within the paper and its Supplementary Information files.

Declarations

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 15 October 2024 / Accepted: 1 May 2025

Published online: 16 May 2025

References

- Nelson DR. Cytochrome P450 and the individuality of species. *Arch Biochem Biophys.* 1999;369:1–10.
- Mitchell AJ, Weng JK. Unleashing the synthetic power of plant oxygenases: from mechanism to application. *Plant Physiol.* 2019;179:813–29.
- Hansen CC, Nelson DR, Møller BL, Werck-Reichhart D. Plant cytochrome P450 plasticity and evolution. *Mol Plant.* 2021;14:1244–65.
- Yan Y, Wu J, Hu G, Gao C, Guo L, Chen X, et al. Current state and future perspectives of cytochrome P450 enzymes for C-H and C=C oxygenation. *Synth Syst Biotechnol.* 2022;7:887–99.
- Nelson D, Werck-Reichhart D. A P450-centric view of plant evolution. *Plant J.* 2011;66:194–211.
- Rasool S, Mohamed R. Plant cytochrome P450s: nomenclature and involvement in natural product biosynthesis. *Protoplasma.* 2016;253:1197–209.
- Singh A, Panwar R, Mittal P, Hassan MI, Singh IK. Plant cytochrome P450s: role in stress tolerance and potential applications for human welfare. *Int J Biol Macromol.* 2021;184:874–86.
- Alber AV, Renault H, Basilio-Lopes A, Bassard J-E, Liu Z, Ullmann P, et al. Evolution of coumaroyl conjugate 3-hydroxylases in land plants: lignin biosynthesis and defense. *Plant J.* 2019;99:24–36.
- Huang H, Wang Y, Yang P, Zhao H, Jenks MA, Lü S, et al. The *Arabidopsis* cytochrome P450 enzyme *CYP96A4* is involved in the wound-induced biosynthesis of cuticular wax and cutin monomers. *Plant J.* 2024;118:1619–34.
- Ayabe S, Akashi T. Cytochrome P450s in flavonoid metabolism. *Phytochem Rev.* 2006;5:271–82.
- Hansen CC, Sørensen M, Veiga TAM, Zibrandtsen JFS, Heskens AM, Olsen CE, et al. Reconfigured cyanogenic glucoside biosynthesis in *Eucalyptus cladocalyx* involves a cytochrome P450 *CYP706C55*. *Plant Physiol.* 2018;178:1081–95.
- Williams D, De Luca V. Plant cytochrome P450s directing monoterpene indole alkaloid (MIA) and benzylisoquinoline alkaloid (BIA) biosynthesis. *Phytochem Rev.* 2023;22:309–38.
- Chakraborty P, Biswas A, Dey S, Bhattacharjee T, Chakraborty S. Cytochrome P450 gene families: role in plant secondary metabolites production and plant defense. *J Xenobiot.* 2023;13:402–23.
- Nelson DR, Ming R, Alam M, Schuler MA. Comparison of cytochrome P450 genes from six plant genomes. *Trop Plant Biol.* 2008;1:216–35.
- Nelson DR. Plant cytochrome P450s from moss to poplar. *Phytochem Rev.* 2006;5:193–204.
- Nelson DR. The cytochrome P450 homepage. *Hum Genomics.* 2009;4:59.
- Sezutsu H, Le Goff G, Feyereisen R. Origins of P450 diversity. *Philos Trans R Soc B Biol Sci.* 2013;368:20120428.
- Hao DC, Song Y, Xiao P, Zhong Y, Wu P, Xu L. The genus *Chrysanthemum*: phylogeny, biodiversity, phytometabolites, and chemodiversity. *Front Plant Sci.* 2022;13:973197.
- Hadzadeh H, Samiei L, Shakeri A. *Chrysanthemum*, an ornamental genus with considerable medicinal value: a comprehensive review. *South Afr J Bot.* 2022;144:23–43.
- Li Y, Yang P, Luo Y, Gao B, Sun J, Lu W, et al. Chemical compositions of chrysanthemum teas and their anti-inflammatory and antioxidant properties. *Food Chem.* 2019;286:8–16.
- Kaneko S, Chen J, Wu J, Suzuki Y, Ma L, Kumazawa K. Potent odorants of characteristic floral/sweet odor in Chinese chrysanthemum flower tea infusion. *J Agric Food Chem.* 2017;65:10058–63.
- Youssef FS, Eid SY, Alshammari E, Ashour ML, Wink M, El-Readi MZ. *Chrysanthemum indicum* and *Chrysanthemum morifolium*: chemical composition of their essential oils and their potential use as natural preservatives with antimicrobial and antioxidant activities. *Foods.* 2020;9:1460.
- Mutiah R, Marsyah A, Saputra AF. Systematic review: phytochemical content and pharmacological effects of *Chrysanthemum* sp. *Proc Int Pharm Ulul Albab Conf Semin (PLANAR).* 2021;1:31–41.
- Chen L, Liu Y, Huang X, Zhu Y, Li J, Miao Y, et al. Comparison of chemical constituents and pharmacological effects of different varieties of *Chrysanthemum Flos* in China. *Chem Biodivers.* 2021;18:e2100206.
- Wang Z, Yuan Y, Hong B, Zhao X, Gu Z. Characteristic volatile fingerprints of four chrysanthemum teas determined by HS-GC-IMS. *Molecules.* 2021;26:7113.
- Teoh KH, Polichuk DR, Reed DW, Nowak G, Covello PS. *Artemisia annua* L. (Asteraceae) trichome-specific cDNAs reveal *CYP71AV1*, a cytochrome P450 with a key role in the biosynthesis of the antimalarial sesquiterpene lactone artemisinin. *FEBS Lett.* 2006;580:1411–6.
- Li W, Lybrand DB, Zhou F, Last RL, Pichersky E. Pyrethrin biosynthesis: the cytochrome P450 oxidoreductase *CYP82Q3* converts jasmolone to pyrethrolone. *Plant Physiol.* 2019;181:934–44.
- Liu Q, Manzano D, Tanić N, Pestic M, Banković J, Pateraki I, et al. Elucidation and in planta reconstitution of the parthenolide biosynthetic pathway. *Metab Eng.* 2014;23:145–53.
- Deng Y, Yang P, Zhang Q, Wu Q, Feng L, Shi W, et al. Genomic insights into the evolution of flavonoid biosynthesis and O-methyltransferase and glucosyltransferase in *Chrysanthemum indicum*. *Cell Rep.* 2024;43:113725.
- Liu L, Xue Y, Luo J, Han M, Liu X, Jiang T, et al. Developing a UV-visible reporter-assisted CRISPR/Cas9 gene editing system to alter flowering time in *Chrysanthemum indicum*. *Plant Biotechnol J.* 2023;21:1519–21.
- Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 2011;39 Suppl 2:W29–37.
- Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, et al. CDD: NCBI's conserved domain database. *Nucleic Acids Res.* 2015;43:D222–6.
- Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, et al. Protein identification and analysis tools on the expasy server. In: Walker JM, editor. *The proteomics protocols handbook*. Totowa, NJ: Humana; 2005. p. 571–607.
- Horton P, Park K-J, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, et al. WoLF PSORT: protein localization predictor. *Nucleic Acids Res.* 2007;35 Suppl 2:W585–7.
- Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 2013;30:772–80.
- Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol.* 2020;37:1530–4.
- Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol Plant.* 2020;13:1194–202.
- Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 2012;40:e49.
- Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME suite: tools for motif discovery and searching. *Nucleic Acids Res.* 2009;37 Suppl 2:W202–8.
- Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, et al. Plant-CARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* 2002;30:325–7.
- Feng L, Chen C, Li T, Wang M, Tao J, Zhao D, et al. Flowery odor formation revealed by differential expression of monoterpene biosynthetic genes and monoterpene accumulation in rose (*Rosa rugosa* Thunb). *Plant Physiol Biochem.* 2014;75:80–8.

42. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol*. 2019;37:907–15.
43. Shumate A, Wong B, Perteu G, Perteu M. Improved transcriptome assembly using a hybrid of long and short reads with StringTie. *PLOS Comput Biol*. 2022;18:e1009730.
44. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. EggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol Biol Evol*. 2021;38:5825–9.
45. Bak S, Beisson F, Bishop G, Hamberger B, Höfer R, Paquette S, et al. Cytochromes P450. *Arabidopsis Book*. 2011;9:e0144.
46. Shen C, Li X. Genome-wide analysis of the P450 gene family in tea plant (*Camellia sinensis*) reveals functional diversity in abiotic stress. *BMC Genomics*. 2023;24:535.
47. Liu X, Gong Q, Zhao C, Wang D, Ye X, Zheng G, et al. Genome-wide analysis of cytochrome P450 genes in *Citrus clementina* and characterization of a CYP gene encoding flavonoid 3'-hydroxylase. *Hortic Res*. 2023;10:uhac283.
48. Zhang K, Qin Y, Sun W, Shi H, Zhao S, He L, et al. Phylogenomic analysis of cytochrome P450 gene superfamily and their association with flavonoids biosynthesis in peanut (*Arachis hypogaea* L). *Genes*. 2023;14:1944.
49. Zhang W, Li H, Li Q, Wang Z, Zeng W, Yin H, et al. Genome-wide identification, comparative analysis and functional roles in flavonoid biosynthesis of cytochrome P450 superfamily in pear (*Pyrus* spp). *BMC Genom Data*. 2023;24:58.
50. Lin X, Tang B, Li Z, Shi L, Zhu H. Genome-wide identification and expression analyses of *CYP450* genes in sweet potato (*Ipomoea batatas* L). *BMC Genomics*. 2024;25:58.
51. Song H, Hua Y, Zhou T, Yue C, Huang JY, Feng Y. Genomic identification of *CYP450* enzymes and new insights into their response to diverse abiotic stresses in *Brassica napus*. *Plant Mol Biol Rep*. 2025;43:41–59.
52. Wei K, Chen H. Global identification, structural analysis and expression characterization of cytochrome P450 monooxygenase superfamily in rice. *BMC Genomics*. 2018;19:35.
53. Cheng Y, Liu H, Tong X, Liu Z, Zhang X, Li D, et al. Identification and analysis of *CYP450* and UGT supergene family members from the transcriptome of *Aralia elata* (Miq.) seem reveal candidate genes for triterpenoid saponin biosynthesis. *BMC Plant Biol*. 2020;20:214.
54. Vriet C, Russinova E, Reuzeau C. From squalene to brassinolide: the steroid metabolic and signaling pathways across the plant kingdom. *Mol Plant*. 2013;6:1738–57.
55. Darabi M, Seddigh S, Abarshahr M. Structural, functional, and phylogenetic studies of cytochrome P450 (CYP) enzyme in seed plants by bioinformatics tools. *Caryologia*. 2017;70:62–76.
56. Zheng X, Li P, Lu X. Research advances in cytochrome P450-catalysed pharmaceutical terpenoid biosynthesis in plants. *J Exp Bot*. 2019;70:4619–30.
57. Liu X, Ma Y, Bu J, Lian C, Ma R, Li Q, et al. Characterization of *CYP82* genes involved in the biosynthesis of structurally diverse benzylisoquinoline alkaloids in *Corydalis yanhusuo*. *Plant Mol Biol*. 2024;114:23.
58. Sun H, Zhang T, Fan Q, Qi X, Zhang F, Fang W, et al. Identification of floral scent in chrysanthemum cultivars and wild relatives by gas chromatography-mass spectrometry. *Molecules*. 2015;20:5346–59.
59. Guo Y, Zhang T, Zhong J, Ba T, Xu T, Zhang Q, et al. Identification of the volatile compounds and observation of the glandular trichomes in *Opisthopappus taihangensis* and four species of *Chrysanthemum*. *Plants*. 2020;9:855.
60. Zhang Z, Wu Q-Y, Ge Y, Huang Z-Y, Hong R, Li A, et al. Hydroxylases involved in terpenoid biosynthesis: a review. *Bioresour Bioprocess*. 2023;10:39.
61. Boachon B, Burdloff Y, Ruan J-X, Rojo R, Junker RR, Vincent B, et al. A promiscuous *CYP706A3* reduces terpene volatile emission from *Arabidopsis* flowers, affecting florivores and the floral microbiome. *Plant Cell*. 2019;31:2947–72.
62. Boachon B, Junker RR, Miesch L, Bassard J-E, Höfer R, Caillieudeaux R, et al. *CYP76C1* (cytochrome P450)-mediated linalool metabolism and the formation of volatile and soluble linalool oxides in *Arabidopsis* flowers: a strategy for defense against floral antagonists. *Plant Cell*. 2015;27:2972–90.
63. Höfer R, Boachon B, Renault H, Gavira C, Miesch L, Iglesias J, et al. Dual function of the cytochrome P450 CYP76 family from *Arabidopsis thaliana* in the metabolism of monoterpenols and phenylurea herbicides. *Plant Physiol*. 2014;166:1149–61.
64. Höfer R, Dong L, André F, Ginglinger J-F, Lukan R, Gavira C, et al. Geraniol hydroxylase and hydroxygeraniol oxidase activities of the CYP76 family of cytochrome P450 enzymes and potential for engineering the early steps of the (seco)iridoid pathway. *Metab Eng*. 2013;20:221–32.
65. Lückler J, Schwab W, Franssen MCR, Van Der Plas LHW, Bouwmeester HJ, Verhoeven HA. Metabolic engineering of monoterpene biosynthesis: two-step production of (+)-*trans*-isopiperitenol by tobacco. *Plant J*. 2004;39:135–45.
66. Deng X, Ye Z, Duan J, Chen F, Zhi Y, Huang M, et al. Complete pathway elucidation and heterologous reconstitution of (+)-nootkatone biosynthesis from *Alpinia oxyphylla*. *New Phytol*. 2024;241:779–92.
67. Vincenti S, Mariani M, Alberti J-C, Jacopini S, Brunini-Bronzini De Caraffa V, Berti L, et al. Biocatalytic synthesis of natural green leaf volatiles using the lipoxygenase metabolic pathway. *Catalysts*. 2019;9:873.
68. Hughes RK, De Domenico S, Santino A. Plant cytochrome CYP74 family: biochemical features, endocellular localisation, activation mechanism in plant defence and improvements for industrial applications. *ChemBiochem*. 2009;10:1122–33.
69. Yactayo-Chang JP, Hunter CT, Alborn HT, Christensen SA, Block AK. Production of the green leaf volatile (*Z*)-3-hexenal by a *Zea mays* hydroperoxide lyase. *Plants*. 2022;11:2201.
70. Ono E, Handa T, Koeduka T, Toyonaga H, Tawfik MM, Shiraiishi A, et al. *CYP74B24* is the 13-hydroperoxide lyase involved in biosynthesis of green leaf volatiles in tea (*Camellia sinensis*). *Plant Physiol Biochem*. 2016;98:112–8.
71. Duan H, Huang M-Y, Palacio K, Schuler MA. Variations in *CYP74B2* (hydroperoxide lyase) gene expression differentially affect hexenal signaling in the columbia and landsberg erecta ecotypes of *Arabidopsis*. *Plant Physiol*. 2005;139:1529–44.
72. Mau CJD, Croteau R. Cytochrome P450 oxygenases of monoterpene metabolism. *Phytochem Rev*. 2006;5:373–83.
73. Werck-Reichhart D. Promiscuity, a driver of plant cytochrome P450 evolution? *Biomolecules*. 2023;13:394.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.