Comprehensive genome-wide identification of *PgCYP450* genes and functional characterization of *PgCYP71BE217* in saponin biosynthesis of *Platycodon grandiflorus*

Wuhua Zhang[1,2#], Pinghui Li[1,2#], Jie Dong[1,2], Yao Song[1,2], DaiDi Che[1,2], Tao Yang[1,2*] and Jinzhu Zhang[1,2*]

[1] *College of Horticulture and Landscape Architecture, Northeast Agricultural University, Harbin 150030, China*
[2] *Key Laboratory of Cold Region Landscape Plants and Applications, Harbin 150030, China*
[#] Authors contributed equally: Wuhua Zhang, Pinghui Li
[*] Correspondence: 542131028@qq.com (Yang T); 13796622756@163.com (Zhang J)

Abstract

Platycodon grandiflorus (Jacq.) A. DC., a well-known traditional Chinese medicinal herb, is particularly abundant in saponins. The cytochrome P450 monooxygenase (CYP450) superfamily constitutes the key group of enzymes responsible for regulating saponin biosynthesis. To date, a comprehensive analysis of *CYP450* genes associated with saponin biosynthesis in *P. grandiflorus* has not been conducted. In the present study, 175 *PgCYP450* genes were systematically identified in *P. grandiflorus*. Phylogenetic analysis classified these genes into nine distinct clans. Transcriptome (RNA-seq) data further demonstrated that the majority of genes within the CYP71 family were upregulated in response to methyl jasmonate (MeJA) treatment. qRT-PCR revealed that *PgCYP71BE217* had the highest expression level after MeJA induction, which was over 60 times higher than the control, and displayed elevated expression primarily in the roots. Subcellular localization and yeast self-activation experiments indicated that PgCYP71BE217 was localized to both the plasma membrane and cytoplasm, and lacked self-activation activity. Overexpression of *PgCYP71BE217* using a hairy root transformation system enhanced the transcript levels of *PgHMGS* (3-hydroxy-3-methylglutaryl-CoA synthase), *PgbAS1* (β-amyrin synthase 1), and *PgSS* (squalene synthase), while simultaneously elevating the accumulation of saponins such as platycodin D, platycoside E, and platycodin D3. The identification and functional analysis of the CYP450 gene family in this study enhances understanding of saponin synthesis and provides valuable genetic resources for *P. grandiflorus* breeding.

Introduction

Platycodon grandiflorus (Jacq.) A. DC., commonly referred to as the balloon flower, is a perennial herb belonging to the family Campanulaceae and is indigenous to East Asia[1]. It is widely valued, not only as an ornamental plant, but also as a medicinal herb. Its ornamental appeal lies in its distinctive balloon-shaped buds that open into star-shaped flowers, available in various colours such as blue, purple, pink, and white. With a long blooming period from June through early autumn, it is popularly used in gardens, rockeries, and containers[2,3]. Meanwhile, the root of *P. grandiflorus* contains bioactive saponins with demonstrated efficacy against respiratory and inflammatory diseases[2]. In recent years, especially during the COVID-19 pandemic, its medicinal applications have attracted increasing attention, leading to growing demand[3].

Saponins, as the primary active components of *P. grandiflorus*, are important indicators for evaluating its medicinal quality[4]. The saponin skeleton of triterpenoid saponins is primarily synthesized through two pathways: the MVA and MEP pathways[5−7]. Several cytochrome P450 monooxygenases (CYP450) and glycosyltransferases (GT) catalyze the hydroxylation, oxidation, and glycosylation of the saponin skeleton to produce saponin monomers, including platycodin D, platycoside E, and platycodin D3[8,9]. Notably, saponin accumulation is subject to transcriptional regulation by phytohormonal elicitors. Among these, MeJA has emerged as a potent inducer of triterpenoid saponin biosynthesis in diverse medicinal species. Exogenous MeJA application can activate biosynthetic pathways by upregulating key genes involved in both early skeleton formation and late-stage modification steps, including

CYP450s[10]. This established role makes MeJA induction a valuable experimental approach for probing the regulatory mechanisms underlying saponin production, and for identifying candidate biosynthetic genes. Cytochrome P450 is one of the classes of heme-containing B-type cytochrome superfamily proteins[11]. Through evolutionary expansion, it is widely distributed in cells of animals, plants, bacteria, and fungi[12]. Currently, more than 300,000 *CYP450* genes, and over 50,000 CYP450 proteins have been discovered. They are divided into 11 clans, including four multi-family clans (CYP71, CYP72, CYP85, and CYP86), and seven single-family clans (CYP51, CYP74, CYP97, CYP710, CYP711, CYP727, and CYP746)[13].

CYP71 is a multi-family clan and contributes to approximately 50% of members in the plant CYP450 family[14]. It comprises several families that are heavily involved in the production of triterpenes, monoterpenes, and sesquiterpenes, including CYP71, CYP81-84, and CYP92[15,16]. The CYP71 gene family is the most abundant and functionally diverse family in the CYP71 clan[17]. Several members of the CYP71A, CYP71D, CYP71AR, and CYP76B subfamilies are involved in monoterpene metabolism[18]. Meanwhile, CYP71AV, CYP71BA, CYP71BL, CYP71D, and CYP71Z play a crucial role in sesquiterpene oxidation[19]. In addition to their involvement in the biosynthesis of monoterpenes and sesquiterpenes, genes from the CYP71 family participate in the biosynthesis of triterpene saponins. In *Sapindus mukorossi*, combined analysis of metabolites and transcriptome, and molecular biology experiments revealed that *SmbHLH2*, *SmTCP4*, and *SmWRKY27* directly regulate the transcription of *SmCYP71D-3* to affect the saponin content[20]. However, it is still unclear whether *CYP71* genes in *P. grandiflorus* is involved in saponin biosynthesis.

The secondary metabolites of traditional Chinese medicinal plants are difficult to obtain through artificial synthesis due to their complex structures[21]. Inducing hairy roots using *Agrobacterium rhizogenes*-mediated transformation is considered a sustainable and effective method to study the biosynthesis of secondary metabolites in plants[22]. In addition, this method is proven to be effective in validating gene function, particularly in the case of regulatory key genes that play a role in specific biosynthetic pathways of secondary metabolites[23]. To date, this method has been successfully employed in various plants, including *Panax ginseng*, *Salvia miltiorrhiza*, and *Astragalus membranaceus*[24,25]. However, it requires tissue culture, increasing the difficulty and cost of operation. To solve this problem, the cut-dip-budding (CBD) delivery system is established, which can generate transformed roots and transformed buds under non-sterile conditions, and without the need for tissue culture. This approach has been successfully applied to various plants, including *Aralia elata* and *Coronilla varia*[26]. However, there have been no studies using this method to obtain transgenic roots in *P. grandiflorus*.

Although *CYP450* genes have been identified and functionally verified in various plant materials, the identification of the CYP450 gene family in *P. grandiflorus* has not been conducted. Nine clans were formed from the identification of 175 *PgCYP450s* genes in *P. grandiflorus*. The expression patterns of 175 *PgCYP450s* genes under MeJA treatment were then investigated, and a hairy root transformation system was used to confirm *PgCYP71BE217* gene's function. This study offers valuable insights into the expression dynamics of the CYP450 gene family in response to MeJA induction, therefore establishing a foundation for future investigations into the roles of *CYP450* genes in saponin biosynthesis.

Materials and methods

Identification and phylogenetic analysis of the *PgCYP450* gene family in *P. grandiflorus*

The National Genomics Data Center (https://ngdc.cncb.ac.cn/gwh) provided the genome data for *P. grandiflorus*[27]. First, potential *PgCYP450s* in the genome of *P. grandiflorus* were found using the hidden Markov model profile of the p450 domain (PF00067) as a reference file. Second, the members of the CYP450 gene family in *P. grandiflorus* were identified using a BLASTP search with query sequences derived from *Arabidopsis thaliana* CYP450 genes. Ultimately, the P450 nomenclature committee (http://drnelson.uthsc.edu/CytochromeP450.html) received the acquired amino acid sequences of the *PgCYP450s* to assign nomenclature. For a thorough understanding, the ExPasy website (https://web.expasy.org/protparam/) was used to study the properties of the PgCYP450 sequences, including their amino acid length, molecular weight (MW), isoelectric point (pI), and other pertinent data. The phylogenetic tree with 1,000 bootstrap repetitions was built using the neighbor-joining (NJ) method in MEGA7.0 software. Based on the genomic annotation of the *P. grandiflorus* genome, the chromosomal position map was plotted using TBtools software[28].

Analysis of phylogenetic tree, *cis*-acting elements, and conserved domains of the *PgCYP71s*

The motif patterns of the PgCYP71s protein sequences were examined using the MEME website (https://meme-suite.org/tools/meme). The nmotif parameter was set to 10; the minw parameter

was set to 6, and the maxw parameter was set to 50. The ggtree platform was used to illustrate the conserved domains and phylogenetic tree of *PgCYP71s*. TBtools was employed for the extraction of a 2,000 bp sequence upstream of the start codon for every PgCYP450 member. The PlantCARE website (https://bioinformatics.psb.ugent.be/webtools/plantcare/html/) was used to identify the *cis*-acting elements in the promoter region.

Treatment of *P. grandiflorus* roots with MeJA

P. grandiflorus seedlings were cultured uniformly on 1/2 MS medium (pH 5.8) supplemented with 100 μM MeJA and 30 g/L sucrose, and 5 g/L agar[5]. Control groups received equivalent volumes of dimethyl sulfoxide. MeJA-treated samples were harvested at 0, 12, 24, and 48 h post-exposure.

qRT-PCR assay

Total RNA was isolated from *P. grandiflorus* tissues and MeJA-treated roots using TransZol UP (TransGen Biotech, Beijing, China). cDNA synthesis employed the SuperMix kit for qRT-PCR (TransGen Biotech). Gene expression quantification utilized the BIO-RAD CFX96 Touch system (Hercules, CA, USA), with 18S rRNA as the endogenous reference (Supplementary Table S1).

Establishment of a genetic transformation system for hairy roots

A hairy root system was developed for *P. grandiflorus* to speed up the evaluation of gene function. The TY solid medium (50 mg/L kanamycin) was used to cultivate *Agrobacterium rhizogenes* K599 transformed with the empty vector pBI121 for 36–48 h. A 5 mL of TY liquid medium was used to incubate a few selected single colonies for the entire night. A 100 μL culture was then added to 50 mL of TY liquid that had been treated with 200 μmol/L AS and 50 mg/L kanamycin. Centrifugation (for 10 min at 4,000 rpm) was used to pellet the cells once they reached an $OD_{600} \approx 0.8$. They were then resuspended in activation buffer (200 μmol/L AS, 10 mmol/L MES, and 10 mmol/L $MgCl_2$), and dark-incubated for two hours at room temperature. The activated solution was used to infect root-free sterile 30-day-old seedlings via vacuum infiltration[5]. After 60 d, transformed plants were moved to sterile soil, and positive transformants were found for target gene expression and saponin measurement.

Subcellular localization and self-activation of *PgCYP71BE217* in yeast

Cloning the *PgCYP71BE217* coding gene without a termination codon into the pCAMBIA1300-GFP *BamHI/KpnI* sites. In tobacco epidermal cells, the recombinant construct (pCAMBIA1300-PgCYP71BE217-GFP), and the empty vector were both simultaneously expressed. Subcellular localization was analyzed using confocal microscopy (Olympus FV3000; 488 nm excitation). For yeast two-hybrid assays, *PgCYP71BE217* CDS was fused to pGBKT7 and transformed into the Y2H Gold strain. Self-activation testing followed established protocols[29].

Analysis of the function of *PgCYP71BE217* in the hairy root of *P. grandiflorus*

The recombinant pCAMBIA1300-PgCYP71BE217-GFP construct was introduced into *Agrobacterium rhizogenes* K599 for *P. grandiflorus* transformation. Transgenic hairy roots were initially screened via GFP fluorescence detection using a chemiluminescence imaging system (Tanon 5200, Shanghai, China). Subsequently, genomic DNA

from positive lines was amplified with GFP-specific primers. Lastly, the transcript levels of PgCYP71BE217 and related pathway genes were measured using RT-qPCR assay (Supplementary Table S1), while HPLC assessed platycodin D, platycoside E, and platycodin D3 contents as described previously[30].
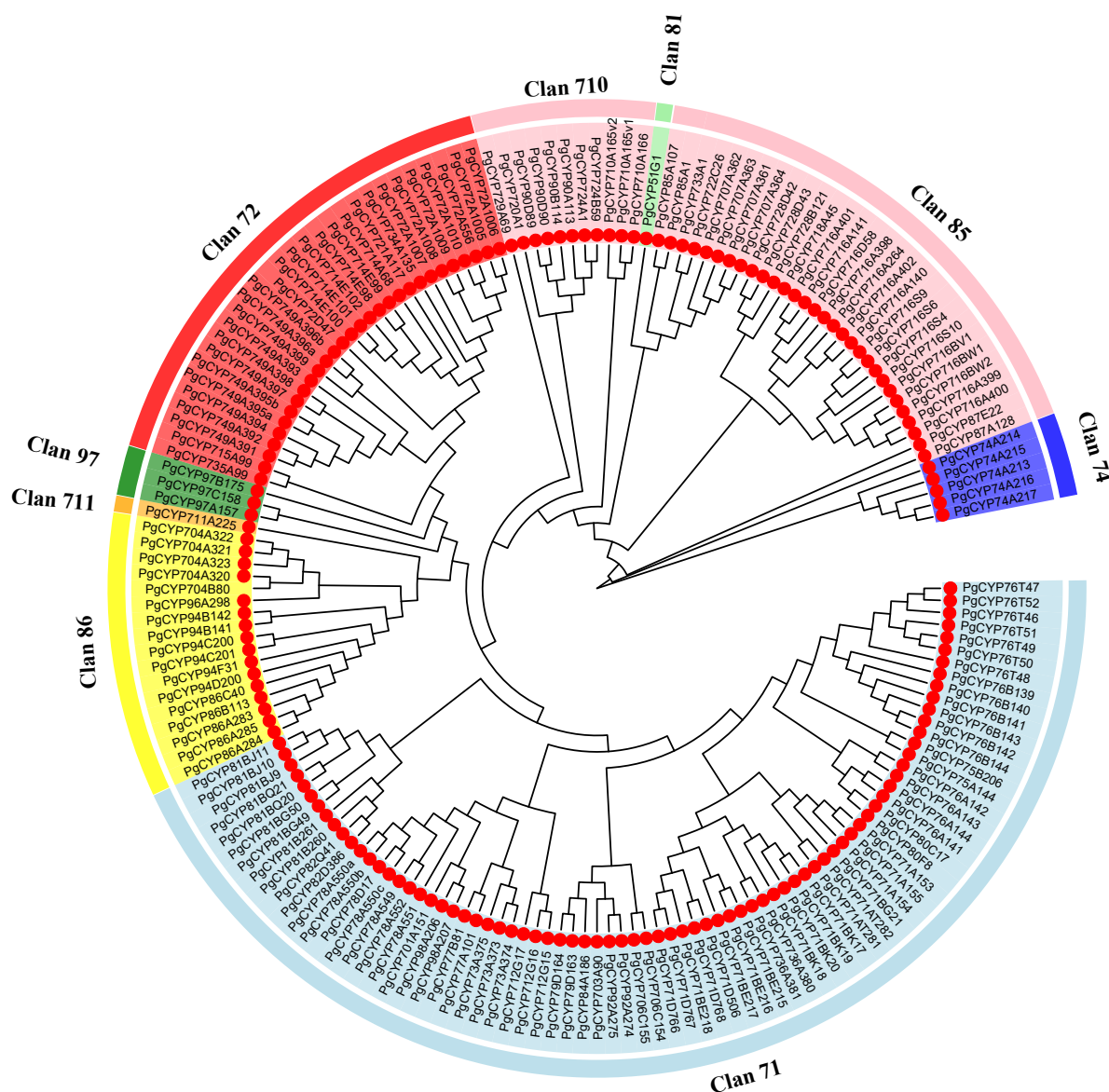
Results

PgCYP450 identification and phylogenetic analysis

In total, 175 PgCYP450s genes with conserved p450 domains were identified in *P. grandiflorus*. The P450 nomenclature committee received the protein sequences of 175 CYP450s to provide nomenclature. Their characteristics, including amino acid length, MW, and pI, are listed in Supplementary Table S2. The amino acid number varied from 219 (*PgCYP716S4*) to 940 (*PgCYP71AT281*); MW varied from 24.89 kDa (*PgCYP716S4*) to 107.32 kDa (*PgCYP71AT281*), and pI varied from 5.79 (*PgCYP71AT281*) to 9.41 (*PgCYP72A1005*).

Phylogenetic analysis divided *PgCYP450s* into nine clans (Fig. 1), including five single-family, and four multi-family clans. Among the nine clans, the clan with the largest number of members was the CYP71 clan, which consisted of 78 members, with the CYP71 family containing the highest members (18), followed by the CYP85 (30), CYP72 (29), and CYP86 (17) clans. The remaining clans exclusively consisted of single-family clans, with the CYP710 clan having the highest number of members.

Phylogenetic analysis and analyses of conserved domains and *cis*-acting elements of *PgCYP71* genes

MeJA has been demonstrated to enhance the accumulation of triterpenoid saponins[5]. MeJA significantly increased the expression level of the *PgCYP71s* gene in the PgCYP450 family, according to an analysis of transcriptome data from previously published studies (Supplementary Fig. S1)[5]. This finding suggested that *PgCYP71s* may play a role in the MeJA-mediated biosynthetic of triterpenoid

Fig. 1 Phylogenetic relationships of the *Platycodon grandiflorus* CYP450 proteins. A neighbor-joining phylogenetic tree was generated based on the alignment of full-length amino acid sequences of the 175 *PgCYP450* proteins in *P. grandiflorus*.
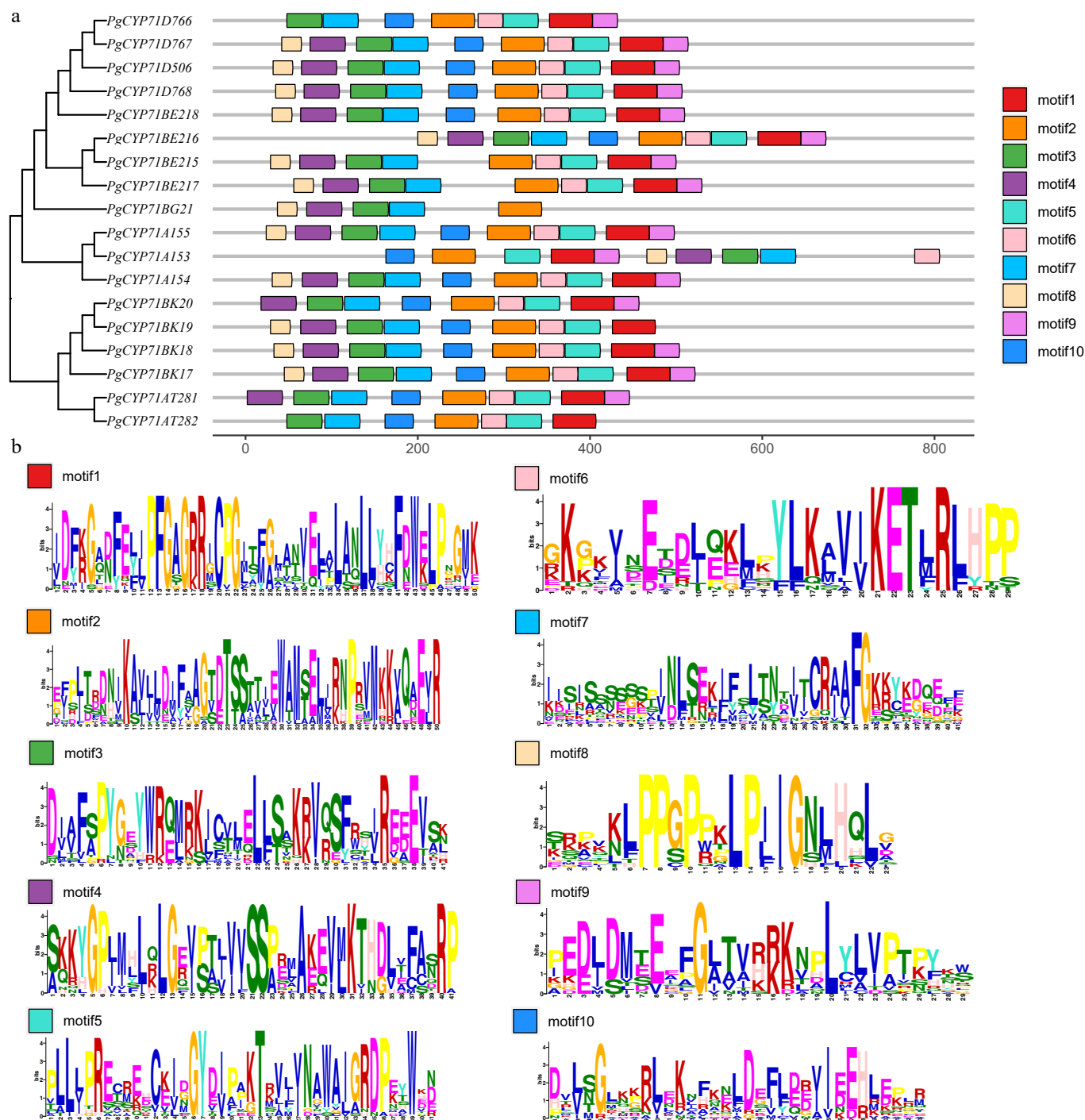
saponins. Moreover, 18 PgCYP71s were determined in *P. grandiflorus*, which were divided into four subfamilies, namely, *CYP71As*, *CYP71Ats*, *CYP71Ds*, and *CYP71BEs*. Analysis of the conserved domains of *PgCYP71* genes revealed that most genes had 9–10 motifs, except the *PgCYP71BG21* gene, exhibiting five motifs (Fig. 2). Promoter *cis*-acting element analysis revealed that *PgCYP71* genes are mainly involved in growth and development, light response, plant hormone signaling, and stress response. Among these genes, *PgCYP71BE217* contained the highest number of plant hormone
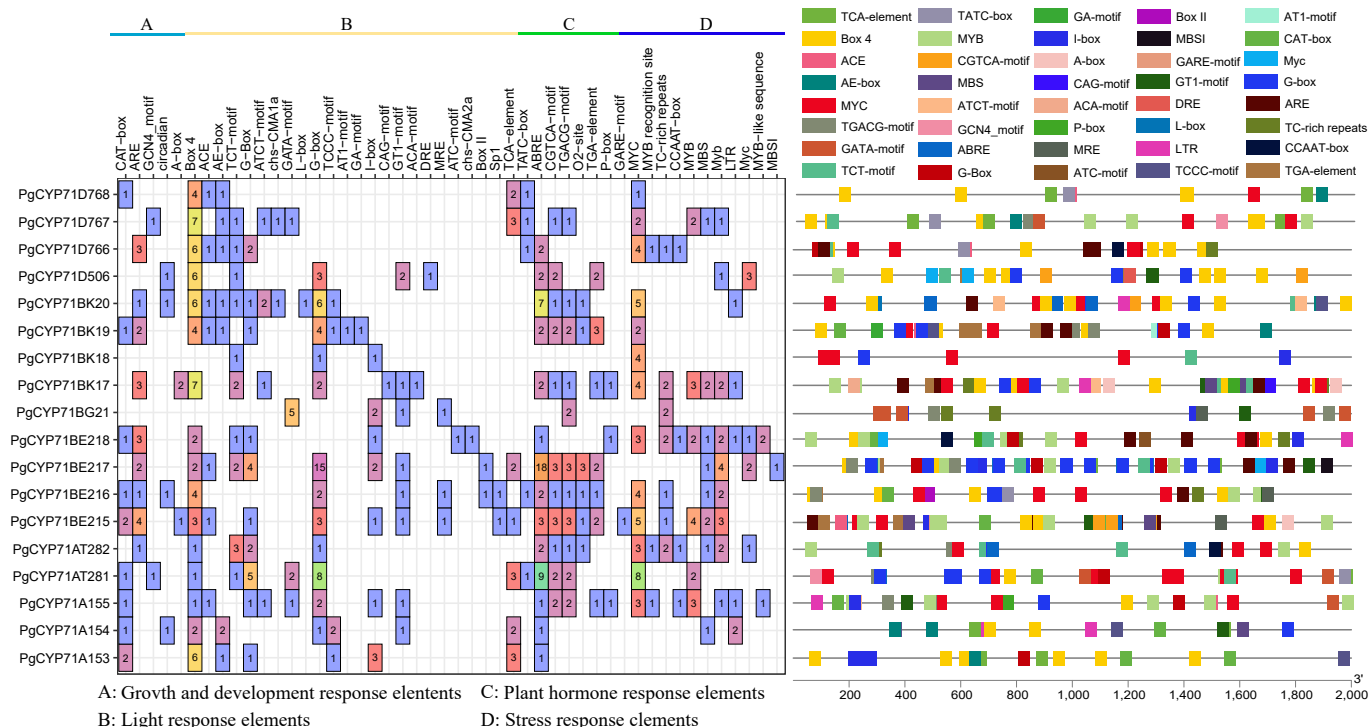
signaling elements (Fig. 3). These results revealed that MeJA appears to regulate saponin accumulation by inducing *PgCYP71s* expression.

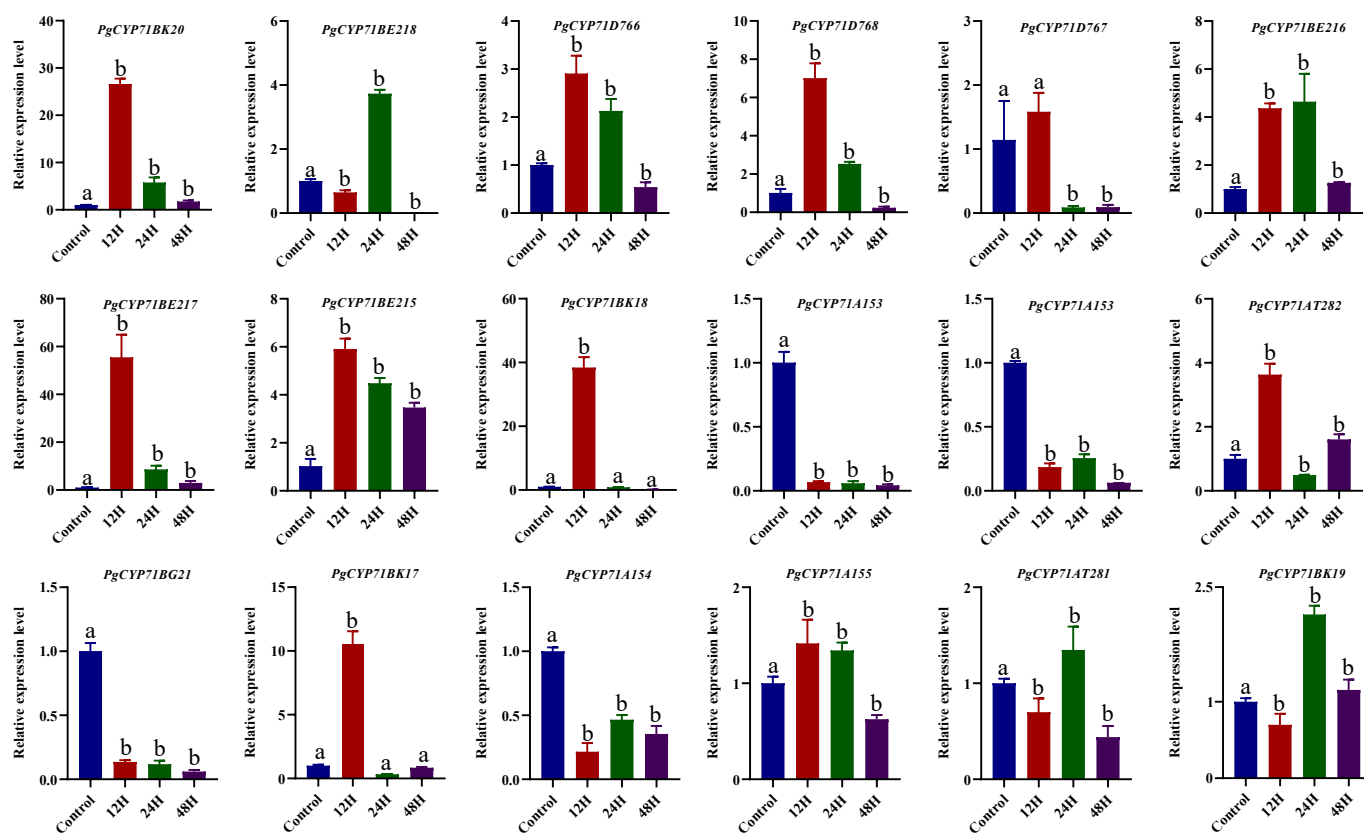## qRT-PCR analysis of the expression patterns of *PgCYP71s* under MeJA treatment

MeJA treatment could increase the total saponin and monomeric saponin contents in the root of *P. grandiflorus*[5,31]. The expression pattern of the *PgCYP71* gene family was analyzed using RT-qPCR. It can be loosely classified into three different classes on the basis

Fig. 2 Phylogenetic relationships and conserved motifs of *P. grandiflorus* PgCYP71 proteins. 18 *PgCYP71s* in P. *grandiflorus* have their full-length amino acid sequences aligned to generate a neighbor-joining phylogenetic tree. (a) Phylogenetic tree and onserved motifs distribution of PgCYP71 proteins. (b) Composition of the conserved motifs in PgCYP71 proteins.
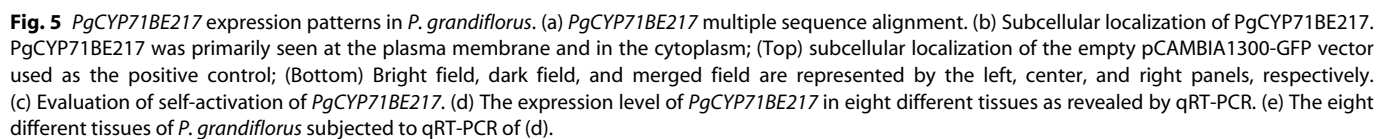
Fig. 3 *Cis*-acting elements in the promoter regions of *PgCYP71* genes. Left - Distribution and functional classification of *cis*-acting elements identified in each *PgCYP71* gene. Right - Visualization of major *cis*-acting elements, with each category represented by a distinct color.

A: Growth and development response elements   C: Plant hormone response elements
B: Light response elements   D: Stress response elements

Fig. 4 The qRT-PCR has shown the expression pattern of the *PgCYP71* gene family at 0 (control), 12, 24, and 48 h after 100 µmol MeJA treatment. Different letters indicate significant differences ($p < 0.05$).

of expression trend. MeJA upregulated *PgCYP71BK20* and *PgCYP71BE217*, which were included in the first category. The second category included *PgCYP71A153* and *PgCYP71D506*, which were inhibited by MeJA. *PgCYP71AT281* and *PgCYP71BE218* were included in the third category; their expression levels first decreased, then increased, and finally decreased (Fig. 4). Compared
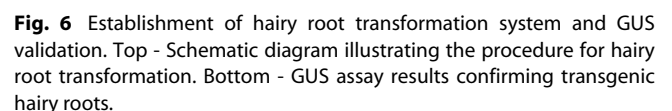
Fig. 5 *PgCYP71BE217* expression patterns in *P. grandiflorus*. (a) *PgCYP71BE217* multiple sequence alignment. (b) Subcellular localization of PgCYP71BE217. PgCYP71BE217 was primarily seen at the plasma membrane and in the cytoplasm; (Top) subcellular localization of the empty pCAMBIA1300-GFP vector used as the positive control; (Bottom) Bright field, dark field, and merged field are represented by the left, center, and right panels, respectively. (c) Evaluation of self-activation of *PgCYP71BE217*. (d) The expression level of *PgCYP71BE217* in eight different tissues as revealed by qRT-PCR. (e) The eight different tissues of *P. grandiflorus* subjected to qRT-PCR of (d).

with the control group, *PgCYP71BE217* exhibited upregulation by more than 60-fold under MeJA treatment at 12 h, implicating it as a pivotal MeJA-responsive regulator of saponin biosynthesis.

Expression analysis and self-activation of *PgCYP71BE217* in yeast
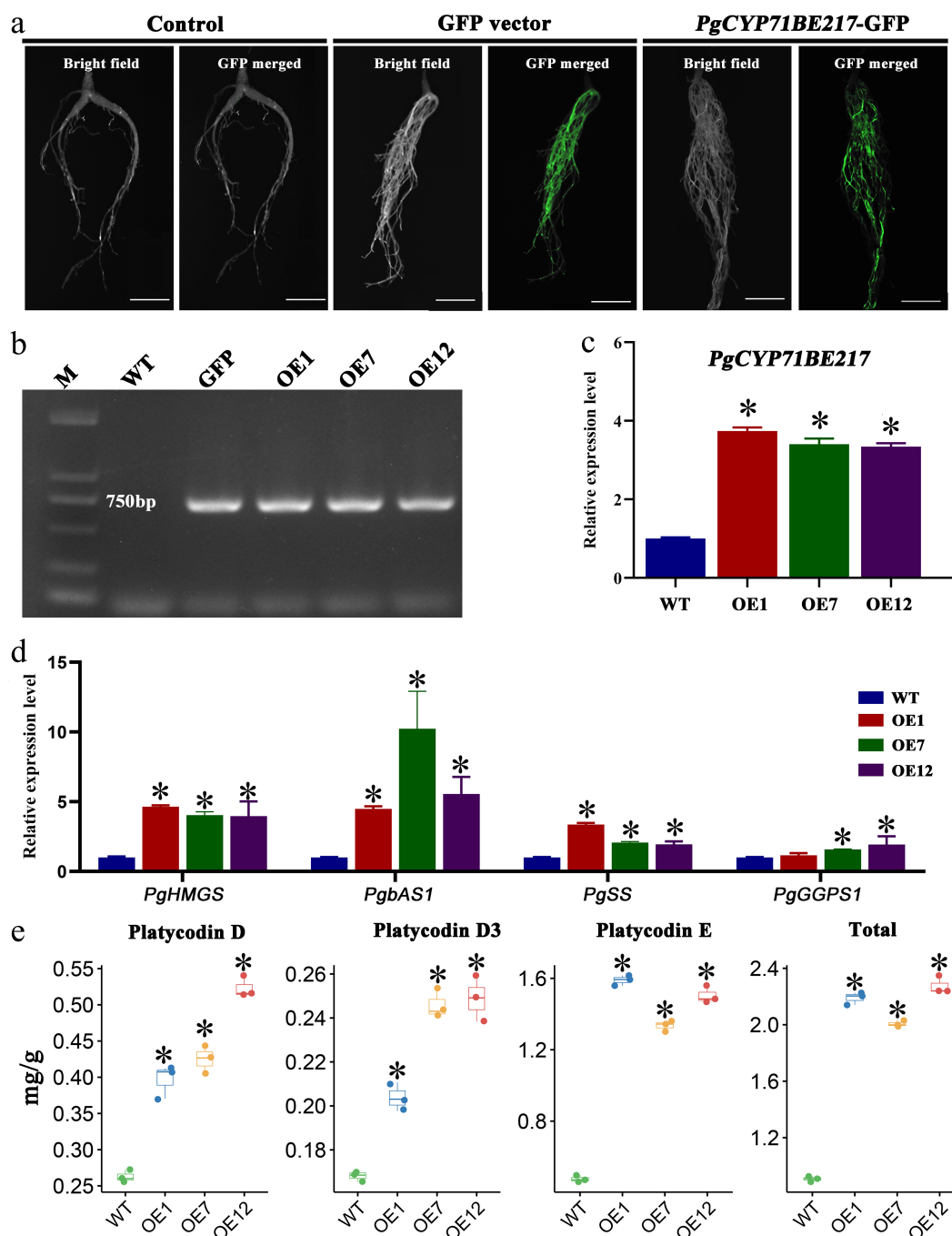
The 504 amino acid-encoding CDS section of the *PgCYP71BE217* gene was cloned. The presence of p450 domains in this gene was discovered through multiple sequence alignment (Fig. 5a). Subcellular localization analysis in tobacco leaf epidermal cells revealed that *PgCYP71BE217*-GFP fluorescence was predominantly observed at the plasma membrane and in the cytoplasm (Fig. 5b). PgCYP71BE217 did not exhibit self-activation activity (Fig. 5c). Moreover, tissue-specific analysis revealed that the *PgCYP71BE217* gene was highly expressed specifically in the roots (Fig. 5d, e).

PgCYP71BE217 overexpression enhanced *P. grandiflorus's* saponin content

P. grandiflorus was effectively used to develop a hairy root transformation system, and positive transgenic hairy roots with an 82%

transformation efficiency were produced (Fig. 6). *PgCYP71BE217* was successfully introduced into *P. grandiflorus* (Fig. 7a, b). *PgCYP71BE217* expression levels in OE1, OE7, and OE12 were

Fig. 6 Establishment of hairy root transformation system and GUS validation. Top - Schematic diagram illustrating the procedure for hairy root transformation. Bottom - GUS assay results confirming transgenic hairy roots.

Fig. 7 Functional verification of *PgCYP71BE217*. (a) Phenotypic characterization of *PgCYP71BE217*-OE transgenic hairy roots using GFP detection. (b) DNA verification of *PgCYP71BE217*-OE vector. (c) qRT-PCR analysis of *PgCYP71BE217* expression in transgenic and wild-type hairy roots. (d) The expression levels of *PgHMGS*, *PgbAS1*, and *PgSS* were significantly increased in the transgenic hairy roots of plants with *PgCYP71BE217* overexpression. (e) Changes in saponin content after *PgCYP71BE217* overexpression. The asterisk denotes a statistically significant difference compared to the WT ($p < 0.05$).

substantially higher than those in the wild type (Fig. 7c). Moreover, the expression levels of *PgHMGS*, *PgbAS1*, and *PgSS* (responsible for the biosynthesis of saponin) were significantly elevated compared with wild type (Fig. 7d). According to HPLC examination, compared to the roots of wild-type plants, the roots overexpressing *PgCYP71BE217* had significantly larger concentrations of platycoside E, platycodin D3, and platycodin D (Fig. 7e). These results clearly demonstrated that *PgCYP71BE217* participated in saponin biosynthesis.

Discussion

Identification of 175 *PgCYP450* genes in *P. grandiflorus*

In recent years, with the development of genome sequencing technology and bioinformatics, researchers could identify *CYP450* genes from 66 plant species, including *A. thaliana*, *Solanum lycopersicum*, and *Zea mays*[32]. The model plant *A. thaliana* has 286 *CYP450* genes, which can be categorized into nine clans consisting of

45 families, and 72 subfamilies[33]. A total of 175 *PgCYP450s* genes were found in this investigation. According to evolutionary analysis, the *PgCYP450* genes could be divided into nine clans: five single gene family clans (CYP51, CYP74, CYP97, CYP710, and CYP711), and four multiple gene family clans (CYP71, CYP72, CYP85, and CYP86). The CYP727 and CYP746 clans were not present. This is consistent with the majority of dicotyledonous plants[16].

PgCYP71BE217 controls *P. grandiflorus's* production of triterpenoid saponins

Among the key inducers that control plant growth and secondary metabolite synthesis is MeJA. Furthermore, it is crucial in controlling the production of saponins[34]. RNA-seq analysis revealed that *PgCYP71s* are induced by MeJA (Supplementary Fig. S1). Several plant hormone-responsive elements, including MeJA *cis*-acting elements, were present in the promoter of *PgCYP71* genes (Fig. 3). *PgCYP71BE217* was significantly induced by MeJA as revealed by qRT-PCR (Fig. 4). These data demonstrated that *PgCYP71BE217* is involved in regulating the synthesis of plant signaling molecules, which is consistent with the observations in *Ginkgo biloba* and *P. ginseng*[35,36]. In addition to the response elements for MeJA, the promoter of this gene contained response elements for other hormones such as abscisic acid and gibberellins, suggesting that *PgCYP71BE217* is regulated by other hormones (Fig. 3)[37].

The CYP71 clan, which was often considered the cradle for the synthesis of monoterpenes and sesquiterpenes, participates in the biosynthesis of diterpenes and triterpenes[38,39]. In *Lotus japonicus*, *CYP71D353* belonged to the CYP71 clan, which was significantly upregulated in the roots, and participated in triterpene biosynthesis[38]. In *P. grandiflorus*, *PgCYP71BE217* overexpression in transformed hairy roots improved the content of saponins, indicating that *PgCYP71BE217* participated in saponin biosynthesis (Fig. 7). These data demonstrated that the CYP71 clan in *P. grandiflorus* was involved in saponin biosynthesis. Our subsequent studies will be focused on functional validation of other potential genes in this clan.

Transcription factors specifically bind to DNA sequences to regulate transcription. Through attaching themselves to the promoters of structural genes, they might control the production of triterpenoid saponins. The promoter of *PgCYP71BE217* contained numerous transcription factor binding sites, such as AP2/ERF[40], bHLH[41], and WRKY transcription factor families[42]. This indicated that *PgCYP71BE217* may be regulated by these transcription factors to regulate saponin accumulation (Fig. 3; Supplementary Table S3).

Rapid analysis of gene function using the hairy root transformation system

Transgenic plants are typically obtained using two methods: *Agrobacterium*-mediated transformation or the gene gun. However, both methods have low transformation efficiency and require tissue culture[26]. Tissue culture is a significant obstacle for obtaining transgenic plants of species that are difficult or impossible to genetically transform. Hence, the development of an efficient method that does not rely on tissue culture is currently a prominent area of research[43].

For medicinal plants, the transformation of hairy roots is a fast and stable method, characterized by genetic stability, rapid growth, and easy cultivation. Related extensive research has been conducted on medicinal plants such as *Catharanthus roseus* and *P. ginseng*[44,45]. However, it requires tissue culture, increasing the difficulty and cost of operation. To reduce the complexity of the process and develop a transformation system that does not rely on tissue culture, a tissue-culture-independent hairy root transformation system was established (Fig. 6). The establishment of this method will provide technical support for the rapid analysis of gene function in *P. grandiflorus*.

Conclusions

In this study, 175 *PgCYP450* genes in *P. grandiflorus* were identified. They were divided into nine clans, excluding the CYP727 and CYP746 clans. Saponin accumulated after MeJA treatment, and *PgCYP71BE217* was clearly stimulated in roots, suggesting that *PgCYP71BE217* may be involved in the production of saponin in roots. The *PgCYP71BE217* overexpression may promote saponin accumulation in *P. grandiflorus*, according to the hairy root transformation. Furthermore, novel technical assistance for quick evaluation of gene function in *P. grandiflorus* was made possible by the development of a hairy root transformation system that does not require tissue culture.

Author contributions

The authors confirm their contributions to the paper as follows: investigation, data curation, formal analysis, writing – original draft, software: Zhang W, Li P; resources: Dong J; visualization: Che D; software: Song Y; visualization, conceptualization, methodology: Yang T, Zhang J. All authors reviewed the results and approved the final version of the manuscript.

Data availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Conflict of interest

The authors declare that they have no conflict of interest.

References

[1] Su X, Liu Y, Han L, Wang Z, Cao M, et al. 2021. A candidate gene identified in converting platycoside E to platycodin D from *Platycodon grandiflorus* by transcriptome and main metabolites analysis. *Scientific Reports* 11:9810

[2] Wu J, Yang G, Zhu W, Wen W, Zhang F, et al. 2012. Anti-atherosclerotic activity of platycodin D derived from roots of *Platycodon* grandiflorum

in human endothelial cells. *Biological and Pharmaceutical Bulletin* 35:1216−1221

[3] Kim TY, Jeon S, Jang Y, Gotina L, Won J, et al. 2021. Platycodin D, a natural component of *Platycodon grandiflorum*, prevents both lysosome- and TMPRSS2-driven SARS-CoV-2 infection by hindering membrane fusion. *Experimental & Molecular Medicine* 53:956−972

[4] Yao L, Lu J, Wang J, Gao WY. 2020. Advances in biosynthesis of triterpenoid saponins in medicinal plants. *Chinese Journal of Natural Medicines* 18:417−424

[5] Zhang W, Zhang J, Fan Y, Dong J, Gao P, et al. 2024. RNA sequencing analysis reveals *PgbHLH28* as the key regulator in response to methyl jasmonate − induced saponin accumulation in *Platycodon grandiflorus*. *Horticulture Research* 11:uhae058

[6] Zhao CL, Cui XM, Chen YP, Liang Q. 2010. Key enzymes of triterpenoid saponin biosynthesis and the induction of their activities and gene expressions in plants. *Natural Product Communications* 5:1147−1158

[7] Song Y, Zhang Y, Wang X, Yu X, Liao Y, et al. 2024. Telomere-to-telomere reference genome for *Panax ginseng* highlights the evolution of saponin biosynthesis. *Horticulture Research* 11:uhae107

[8] Mai Y, Hu H, Ji W, Xiao Y, Zhou H, et al. 2025. Evolution and functional characterization of a biosynthetic gene cluster for saponin biosynthesis in Sapindaceae. *Molecular Plant* 18:1089−1093

[9] Cheng Y, Liu H, Tong X, Liu Z, Zhang X, et al. 2020. 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 Biology* 20:214

[10] Lacchini E, Qu T, Moses T, Volkov AN, Goossens A. 2025. Engineering *Gypsophila elegans* hairy root cultures to produce endosomal escape-enhancing saponins. *Plant Biotechnology Journal* 23:3068−3082

[11] Finnigan JD, Young C, Cook DJ, Charnock SJ, Black GW. 2020. Cytochromes P450 (P450s): a review of the class system with a focus on prokaryotic P450s. *Chemistry and Structural Biology* 122:289−320

[12] Nelson DR. 1999. Cytochrome P450 and the individuality of species. *Archives of Biochemistry and Biophysics* 369:1−10

[13] Nelson DR, Goldstone JV, Stegeman JJ. 2013. The cytochrome P450 genesis locus: the origin and evolution of animal cytochrome P450s. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences* 368:20120474

[14] Bathe U, Tissier A. 2019. Cytochrome P450 enzymes: a driving force of plant diterpene diversity. *Phytochemistry* 161:149−162

[15] Zhang W, Li H, Li Q, Wang Z, Zeng W, et al. 2023. Genome-wide identification, comparative analysis and functional roles in flavonoid biosynthesis of cytochrome P450 superfamily in pear (*Pyrus* spp.). *BMC Genomic Data* 24:58

[16] Nelson D, Werck-Reichhart D. 2011. A P450-centric view of plant evolution. *The Plant Journal* 66:194−211

[17] Zhang C, Ni S, Zhang S, Nigarish M, Cheng C, et al. 2023. Genome-wide identification and expression of *CYP71* gene family in response to low-temperature stress in banana. *The Journal of Horticultural Science and Biotechnology* 98:159−177

[18] Ghosh S. 2017. Triterpene structural diversification by plant cytochrome P450 enzymes. *Frontiers in Plant Science* 8:1886

[19] Sui C, Zhang J, Wei J, Chen S, Li Y, et al. 2011. Transcriptome analysis of *Bupleurum chinense* focusing on genes involved in the biosynthesis of saikosaponins. *BMC Genomics* 12:539

[20] Xu Y, Zhao G, Ji X, Liu J, Zhao T, et al. 2022. Metabolome and transcriptome analysis reveals the transcriptional regulatory mechanism of triterpenoid saponin biosynthesis in soapberry (*Sapindus mukorossi* Gaertn.). *Journal of Agricultural and Food Chemistry* 70(23):7095−7109

[21] Wang T, Liu J, Luo X, Hu L, Lu H. 2021. Functional metabolomics innovates therapeutic discovery of traditional Chinese medicine derived functional compounds. *Pharmacology & Therapeutics* 224:107824

[22] Yang J, Yang X, Li B, Lu X, Kang J, et al. 2020. Establishment of *in vitro* culture system for *Codonopsis pilosula* transgenic hairy roots. *3 Biotech* 10:137

[23] Bosselut N, Van Ghelder C, Claverie M, Voisin R, Onesto JP, et al. 2011. *Agrobacterium rhizogenes* − mediated transformation of *Prunus* as an alternative for gene functional analysis in hairy-roots and composite plants. *Plant Cell Reports* 30:1313−1326

[24] Yu KW, Murthy HN, Hahn EJ, Paek KY. 2005. Ginsenoside production by hairy root cultures of *Panax ginseng*: influence of temperature and light quality. *Biochemical Engineering Journal* 23:53−56

[25] Zhang J, Zhou L, Zheng X, Zhang J, Yang L, et al. 2017. Overexpression of SmMYB9b enhances tanshinone concentration in *Salvia miltiorrhiza* hairy roots. *Plant Cell Reports* 36:1297−1309

[26] Cao X, Xie H, Song M, Lu J, Ma P, et al. 2023. Cut−dip−budding delivery system enables genetic modifications in plants without tissue culture. *The Innovation* 4:100345

[27] Jia Y, Chen S, Chen W, Zhang P, Su Z, et al. 2022. A chromosome-level reference genome of Chinese balloon flower (*Platycodon grandiflorus*). *Frontiers in Genetics* 13:869784

[28] Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, et al. 2020. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Molecular Plant* 13:1194−1202

[29] Sun HJ, Luo ML, Zhou X, Zhou Q, Sun YY, et al. 2020. PuMYB21/PuMYB54 coordinate to activate *PuPLDβ1* transcription during peel browning of cold-stored 'Nanguo' pears. *Horticulture Research* 7:136

[30] Yu H, Liu M, Yin M, Shan T, Peng H, et al. 2021. Transcriptome analysis identifies putative genes involved in triterpenoid biosynthesis in *Platycodon grandiflorus*. *Planta* 254:34

[31] Li W, Lin S, Wang R, Chen C, Ni L, et al. 2025. Regulation of plant hormones on the secondary metabolism of medicinal plants. *Medicinal Plant Biology* 4:e020

[32] Nelson DR. 2009. The cytochrome P450 homepage. *Human Genomics* 4:59

[33] Chapple C. 1998. Molecular-genetic analysis of plant cytochrome p450-dependent monooxygenases. *Annual Review of Plant Biology* 49:311−343

[34] Wang JW, Wu JY. 2013. Effective elicitors and process strategies for enhancement of secondary metabolite production in hairy root cultures. In *Biotechnology of Hairy Root Systems*, ed. Doran PM. Berlin, Heidelberg: Springer. pp. 55−89 doi: 10.1007/10_2013_183

[35] Kochan E, Balcerczak E, Lipert A, Szymańska G, Szymczyk P. 2018. Methyl jasmonate as a control factor of the *synthase squalene* gene promoter and ginsenoside production in American ginseng hairy root cultured in shake flasks and a nutrient sprinkle bioreactor. *Industrial Crops and Products* 115:182−193

[36] Vijendra PD, Jayanna SG, Kumar V, Sannabommaji T, Rajashekar J, et al. 2020. Product enhancement of triterpenoid saponins in cell suspension cultures of *Leucas aspera* Spreng. *Industrial Crops and Products* 156:112857

[37] Tian R, Zhang C, Gu W, Gu Y, Xu F, et al. 2021. Proteomic insights into protostane triterpene biosynthesis regulatory mechanism after MeJA treatment in *Alisma orientale* (Sam.) Juz. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 1869:140671

[38] Krokida A, Delis C, Geisler K, Garagounis C, Tsikou D, et al. 2013. A metabolic gene cluster in *Lotus japonicus* discloses novel enzyme functions and products in triterpene biosynthesis. *New Phytologist* 200:675−690

[39] Schröder G, Unterbusch E, Kaltenbach M, Schmidt J, Strack D, et al. 1999. Light-induced cytochrome P450-dependent enzyme in indole alkaloid biosynthesis: tabersonine 16-hydroxylase. *FEBS Letters* 458:97−102

[40] Deng B, Huang Z, Ge F, Liu D, Lu R, et al. 2017. An AP2/ERF family transcription factor PnERF1 raised the biosynthesis of saponins in *Panax notoginseng*. *Journal of Plant Growth Regulation* 36:691−701

[41] Song M, Wang H, Wang Z, Huang H, Chen S, et al. 2021. Genome-wide characterization and analysis of bHLH transcription factors related to anthocyanin biosynthesis in fig (*Ficus carica* L.). *Frontiers in Plant Science* 12:730692

[42] Singh AK, Kumar SR, Dwivedi V, Rai A, Pal S, et al. 2017. A WRKY transcription factor from *Withania somnifera* regulates triterpenoid withanolide accumulation and biotic stress tolerance through modulation of phytosterol and defense pathways. *New Phytologist* 215:1115−1131

[43] Altpeter F, Springer NM, Bartley LE, Blechl AE, Brutnell TP, et al. 2016. Advancing crop transformation in the era of genome editing. *The Plant Cell* 28:1510−1520

[44] Park NI, Tuan PA, Li X, Kim YK, Yang TJ, et al. 2011. An efficient protocol for genetic transformation of *Platycodon grandiflorum* with *Agrobacterium rhizogenes*. *Molecular Biology Reports* 38:2307−2313

[45] Yamazaki Y, Sudo H, Yamazaki M, Aimi N, Saito K. 2003. Camptothecin biosynthetic genes in hairy roots of *Ophiorrhiza pumila*: cloning, characterization and differential expression in tissues and by stress compounds. *Plant & Cell Physiology* 44:395−403