BMC Plant Biology

RESEARCH Open Access

Characterization of the *2ODD* genes of DOXC subfamily and its members involved in flavonoids biosynthesis in *Scutellaria baicalensis*

Sanming Zhu[1], Mengying Cui[1*] and Qing Zhao[1,2*]

Abstract

Background 2-oxoglutarate-dependent dioxygenase (2ODD) superfamily is the second largest enzyme family in the plant genome and plays diverse roles in secondary metabolic pathways. The medicinal plant *Scutellaria baicalensis* Georgi contains various flavonoids, which have the potential to treat coronavirus disease 2019 (COVID-19), such as baicalein and myricetin. Flavone synthase I (FNSI) and flavanone 3-hydroxylase (F3H) from the 2ODDs of DOXC subfamily have been reported to participate in flavonoids biosynthesis. It is certainly interesting to study the 2ODD members involved in the biosynthesis of flavonoids in *S. baicalensis*.

Results We provided a genome-wide analysis of the *2ODDs* of DOXC subfamily in *S. baicalensis*, a total of 88 *2ODD* genes were identified, 82 of which were grouped into 25 distinct clades based on phylogenetic analysis of At2ODDs. We then performed a functional analysis of Sb2ODDs involved in the biosynthesis of flavones and dihydroflavonols. Sb2ODD1 and Sb2ODD2 from DOXC38 clade exhibit the activity of FNSI (Flavone synthase I), which exclusively converts pinocembrin to chrysin. *Sb2ODD1* has significantly higher transcription levels in the root. While *Sb2ODD7* from DOXC28 clade exhibits high expression in flowers, it encodes a F3H (flavanone 3-hydroxylase). This enzyme is responsible for catalyzing the conversion of both naringenin and pinocembrin into dihydrokaempferol and pinobanksin, kinetic analysis showed that Sb2ODD7 exhibited high catalytic efficiency towards naringenin.

Conclusions Our experiment suggests that Sb2ODD1 may serve as a supplementary factor to SbFNSII-2 and play a role in flavone biosynthesis specifically in the roots of *S. baicalensis*. Sb2ODD7 is mainly responsible for dihydrokaempferol biosynthesis in flowers, which can be further directed into the metabolic pathways of flavonols and anthocyanins.

Keywords *Scutellaria baicalensis*, Flavonoid, Flavone synthase I, Flavanone 3-hydroxylase, Biosynthesis

*Correspondence:
Mengying Cui
cuimengying@outlook.com
Qing Zhao
zhaoqing@cemps.ac.cn
[1]Shanghai Key Laboratory of Plant Functional Genomics and Resources, CAS Center for Excellence in Molecular Plant Sciences Chenshan Science Research Center, Shanghai Chenshan Botanical Garden, Shanghai 201602, China
[2]State Key Laboratory of Plant Molecular Genetics, CAS Center for Excellence in Molecular Plant Sciences, Chinese Academy of Sciences, Shanghai 200032, China

Background

Scutellaria baicalensis Georgi is a perennial herb belonging to *Scutellaria* genus in the Labiaceae family. It is cultivated worldwide due to its remarkable therapeutic properties. *S. baicalensis* contains two types of flavones: the 4′-hydroxyflavones like scutellarein and scutellarin accumulate in aerial tissues, while 4′-deoxyflavones such as baicalein, baicalin, wogonin and wogonoside, are abundant in the roots [1]. In addition, *S. baicalensis* also contains various flavonoids like naringenin, pinocembrin, dihydrokaempferol, kaempferol, and so on [2, 3]. These flavonoids exhibit anti-bacterial, anti-inflammatory, anticancer and anti-viral effects [2]. The ongoing COVID-19 pandemic, caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) poses a serious threat to human health, and the 3 C-like protease ($3CL^{pro}$) of SARS-CoV-2 is a primary target for the development of broad-spectrum antiviral drugs [4]. Baicalein and myricetin, found in *S. baicalensis*, have been identified as inhibitors of the SARS-CoV-2 $3CL^{pro}$, demonstrating inhibition of viral replication in Vero cells [5]. It has been reported that a total of 100 flavonoids have been found in *S. baicalensis*, with flavones being the major compounds [3]. Recent studies have provided genomic and transcriptome data for *S. baicalensis*, which serve as a robust foundation for the analysis of flavonoid biosynthesis in the medicinal plant [6].

Flavonoids are a class of secondary metabolites widely distributed in plants. They possess a common C6-C3-C6 skeleton with two aromatic rings connected by a three-carbon chain, typically arranged in a phenylchromane configuration [7]. Based on the degree of unsaturation and the substitution pattern, the structures of flavonoids are diverse. The major types of flavonoids include flavones, flavonols, flavanones, flavanols, dihydroflavonols, anthocyanins, isoflavones and chalcones [8]. In addition to serving as the primary source of plant pigments, flavonoids play a crucial role in influencing the colors of flowers and fruits, thus contributing to the process of plant reproduction [9]. Furthermore, flavonoids offer protection against various external stresses, such as drought [10], low temperature [11], diseases [12] and UV-B radiation [13].

The classic flavonoid biosynthesis pathway starts with L-phenylalanine, which is catalyzed by phenylalanine ammonia lyase (PAL), cinnamate 4-hydroxylase (C4H), 4-coumarate CoA ligase (4CL), chalcone synthase (CHS) and chalcone isomerase (CHI), resulting in the formation of naringenin (4′-hydroxyflavanone), a crucial intermediate compound in the pathway (Fig. 1) [1]. Naringenin can subsequently serve as a substrate for various enzymes, leading to the production of different types of flavonoids, such as isoflavones, flavones, dihydroflavonols and so on [14]. However, *S. baicalensis* has evolved a specific 4′-deoxyflavones pathway for the biosynthesis of baicalein and wogonin (Fig. 1). In this pathway, L-phenylalanine is converted into 4′-deoxyflavanone pinocembrin by SbPAL, cinnamate CoA ligase (SbCLL-7), SbCHS-2 and SbCHI. Pinocembrin is also an important intermediate product, which is further converted into 4'-deoxyflavones through flavone synthase, flavone hydroxylases and methyl-transferases.

The 2ODD superfamily is the second largest enzyme family in the plant genome, after the cytochrome P450 superfamily (CYP450). Members of the 2ODD family catalyze various oxidative reactions in plants, such as hydroxylations, demethylations, desaturations, epimerization, rearrangement, halogenation ring closure and ring cleavage [15]. The plant 2ODD family can be categorized into three distinct evolutionary subfamilies based on amino acid sequence similarity: DOXA, DOXB and DOXC [16]. Plant homologs of *Escherichia coli* AlkB are DNA repair proteins [17], and they are classified into the DOXA subfamily. The DOXB subfamily includes Prolyl 4-hydroxylases (P4Hs), which participate in the synthesis of cell wall proteins in plants [18]. The DOXC subfamily is involved in secondary metabolisms, such as alkaloids and flavonoids [19]. The 2ODDs involved in flavonoid biosynthesis are further classified into two clades: DOXC28 and DOXC47 [16]. DOXC28 clade comprises the classic FNSI and F3H enzymes, while the DOXC47 clade consists of flavonol synthase (FLS) and anthocyanidin synthase (ANS), which function downstream of F3H in the biosynthesis of flavonoids. Early studies of classic FNSI enzymes primarily focused on Apiaceae and monocots [20, 21]. It has been reported that PcFNSI and PcF3H have a high level of sequence identity, and PcFNSI may have evolved from PcF3H by gene duplication [22]. However, there is another type of 2ODD subfamily that has been reported to have FNSI (AtDMR6) activity, which belongs to the DOXC38 clade in *Arabidopsis thaliana*. AtDMR6 is involved not only in salicylic acid catabolism but also in apigenin (4′-hydroxyflavone) biosynthesis [23, 24]. *S. baicalensis* is rich in various flavonoids. However, it is still unclear whether any member of the 2ODD family is involved in the biosynthesis of flavones in this plant. Furthermore, the role of F3H, an important enzyme involved in the flavonoid pathway, has not been studied in *S. baicalensis*.

Here, we carried out a genome-wide study of *2ODD* DOXC family members in *S. baicalensis* by phylogenetic analysis and expression profiles. Subsequently, we identified and characterized the Sb2ODD1-Sb2ODD7 of DOXC28 and DOXC38 clades, and performed enzyme assays. Our findings elucidated the relationship between the Sb2ODDs and flavonoids accumulation patterns in *S. baicalensis*.

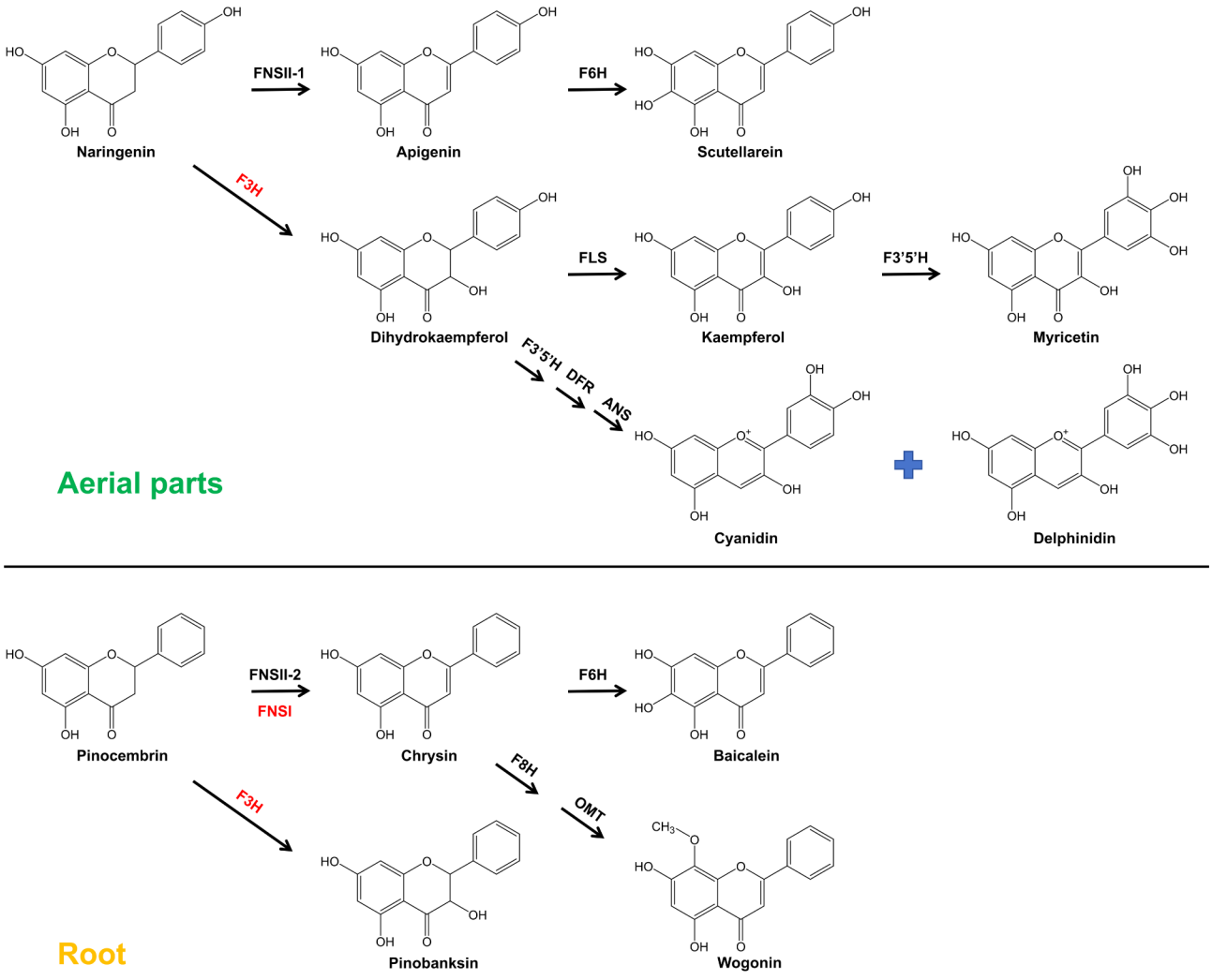


Fig. 1 Two Pathways for biosynthesis of Flavonoids in *S. baicalensis*
FNSII, flavone synthase II; FNSI, flavone synthase I; F3H, flavanone 3-hydroxylase; F6H, flavone 6-hydroxylase; F8H, flavone 8-hydroxylase; OMT, 8-*O*-methyl transferase; FLS, flavonol synthase; F3'5'H, flavonoid 3'5'-hydroxylase; DFR, dihydroflavonol 4-reductase; ANS, anthocyanidin synthase. The enzymes highlighted in red are still under investigation

Materials and methods

Genome-wide identification, sequence alignment and phylogenetic analysis of *Sb2ODD* genes.

The HMM profile of 2ODD domain (PF14226 and PF03171) from Pfam database (https://www.ebi.ac.uk/interpro) was used to extract full-length *2ODD* candidates from the *S. baicalensis* genome by the HMM algorithm (HMMER) [25]. Multiple sequence alignments and phylogenetic analysis were performed using MEGA 11 [26]. The neighbor-joining tree was constructed under the default parameters with Sb2ODD candidates and At2ODD sequences. The maximum-likelihood tree was constructed under the default parameters with Sb2ODD sequences and reported F3H and FNSI. The bootstrap statistics were calculated with 1,000 replications.

Gene location visualization

The location of the *Sb2ODD* genes on the chromosome was determined by TBtools [27].

Gene cloning and expression vector construction

Based on the transcriptome data of different organs and the genome of *S. baicalensis* [6], we amplified the full-length coding regions of *Sb2ODD1* using specific primers (Table S2). *Sb2ODD2*, *Sb2ODD3*, *Sb2ODD4*, *Sb2ODD5*, *Sb2ODD6* and *Sb2ODD7* were obtained by *de novo* synthesis (GenScript, Nanjing, China). According to the manufacturer's instructions, fragments were cloned into the entry vector pDONR207 using the Gateway BP Clonase II Enzyme Kit (Invitrogen, MA, USA). These fragments were then cloned into the yeast expression vector pYesdest52 and prokaryotic expression vector

pYesdest17 using the Gateway LR Clonase II Enzyme Kit, respectively.

In vivo yeast enzyme assays

S. cerevisiae WAT11 was used as the host strain for in vivo enzyme assays. The pYesdest52 empty vector or constructs were transformed into the yeast using the Yeast Transformation II Kit (ZYMO, CA, USA). Transformants were selected on synthetic drop-out medium without uracil (SD-Ura) containing 20 g/L glucose and grown at 28 °C for 48 h. The recombinant strains were initially grown in SD-Ura liquid medium with 20 g/L glucose at 28 °C for 24 h until the OD_{600} reached 2–3. Yeast cells were centrifuged at 4000 rpm for 10 min, then resuspended in the SD-Ura liquid medium with 20 g/L galactose to induce expression of the target proteins. Different substrates and α-ketoglutaric acid were supplemented in the medium. After fermentation for 48 h, yeast cells were harvested by centrifugation, extracted with 1 mL of 70% MeOH (pH 5.0) for metabolite analysis.

Enzyme assays and kinetic studies

The empty vector or constructed prokaryotic expression vector was transformed into *E. coli* Rosetta (DE3). Transformants were initially grown in 10 mL of LB liquid medium with 100 μg/mL ampicillin at 37 °C for 12 h and then transferred to 300 mL of LB liquid medium until the OD_{600} reached 0.6–0.8. Recombinant protein expression was then induced with 1 mM Isopropyl β-D-thiogalactopyranoside (IPTG) at 16 °C for 16 h. After harvesting the *E. coli* cells, high-pressure cell disruption equipment (Constant Systems, Northants, UK) was used to crush the cells. The crude protein lysate was centrifuged and purified by affinity chromatography with Ni-nitrilotriacetic acid (Ni-NTA) agarose (Qiagen, Germany). The protein concentration was determined using the Bradford method and analyzed by SDS-polyacrylamide gel electrophoresis. The target protein concentration was further determined by Image J software.

The enzyme assays in vitro was performed according to a previously described protocol with some modifications [23]. The reaction mixture contained 100 mM NaH_2PO_4 (pH 6.8), 2 mM DTT, 1 mM α-ketoglutaric acid, 2 mM ascorbic acid, 1 mM ATP, 0.25 mM ferrous sulfate, 50 μM substrate and 1 μg of recombinant purified protein in a final volume of 100 μL. Enzyme assay was performed at 37 °C for 1 h in open tubes with shaking. The reaction was initiated by the addition of the enzyme and terminated by adding methanol. After centrifugation at the top speed for 10 min, the supernatant was analyzed by HPLC.

For kinetics measurements, naringenin or pinocembrin were used at concentrations ranging from 1 to 250 μM. The reaction time was reduced to 10 min. *K*m and *V*max values were obtained by using GraphPad Prism version 8.0.2 for Windows (GraphPad Software, San Diego, California USA, www.graphpad.com).

Standard compounds

Naringenin and pinocembrin were purchased from Sigma-Aldrich (MO, USA), dihydrokaempferol, pinobanksin, salicylic acid and 2, 5-dihydroxybenzoic acid were purchased from Yuanye-Biotech (Shanghai, China). All of the above standard compounds were dissolved in dimethyl sulfoxide (DMSO).

Metabolite analysis

An Agilent 1260 Infinity II HPLC system was used for metabolite analysis. Flavones were detected at 280 nm. Separation was carried out on a 100×2 mm, 3 μ Luna C18 (2) column. The column was maintained at 35 °C. The flow rate of the mobile phase consisting of 0.1% (v/v) formic acid in water (A) and 1:1 MeOH/Acetonitrile + 0.1% formic acid (B) was set to 0.26 mL/min. The gradient program was as follows: 0–3 min, 20% B; 20 min, 50% B; 20–30 min, 50% B; 36 min, 70% B; 37 min, 20% B and 37–43 min, 20% B.

Salicylic acid and 2, 5-dihydroxybenzoic acid were detected at 300 nm. Separation was carried out on a 250×4.6 mm, 5 μm Eclipse XDB-C18 column. The column was maintained at 35 °C. The flow rate of the mobile phase consisting of 0.1% (v/v) formic acid in water (A) and 0.1% (v/v) formic acid in Acetonitrile (B) was set to 0.8 mL/min. The gradient program was as follows: 0–5 min, 10% B; 25 min, 40% B; 25–30 min, 40% B; 36 min, 60% B; 37 min, 10% B and 37–43 min, 10% B.

Based on the retention time of standard substances and standard curves, metabolites were confirmed and measured.

Results

S. baicalensis contains 88 2*ODD* genes belonging to the DOXC subfamily

It has been reported that *S. baicalensis* is rich in not only flavones, but also in flavonols and anthocyanins [3, 28, 29]. Considering the structures of these compounds (Fig. 1), it can be inferred that *S. baicalensis* harbors Sb2ODDs involved in their biosynthetic pathway. Based on a HMMER search of the *S. baicalensis* genome, a total of 88 protein sequences were identified as members of the 2ODD enzyme family. By conducting a phylogenetic analysis of Sb2ODDs and At2ODDs, 82 candidates were grouped into 25 clades while 6 candidates remained unclassified (Fig. 2). We annotated these clades based on those annotation of *A. thaliana* enzymes, including gibberellin biosynthesis (DOXC3, DOXC7 and DOXC22), gibberellin catabolism (DOXC12 and DOXC13), auxin metabolism (DOXC15), glucosinolate metabolism

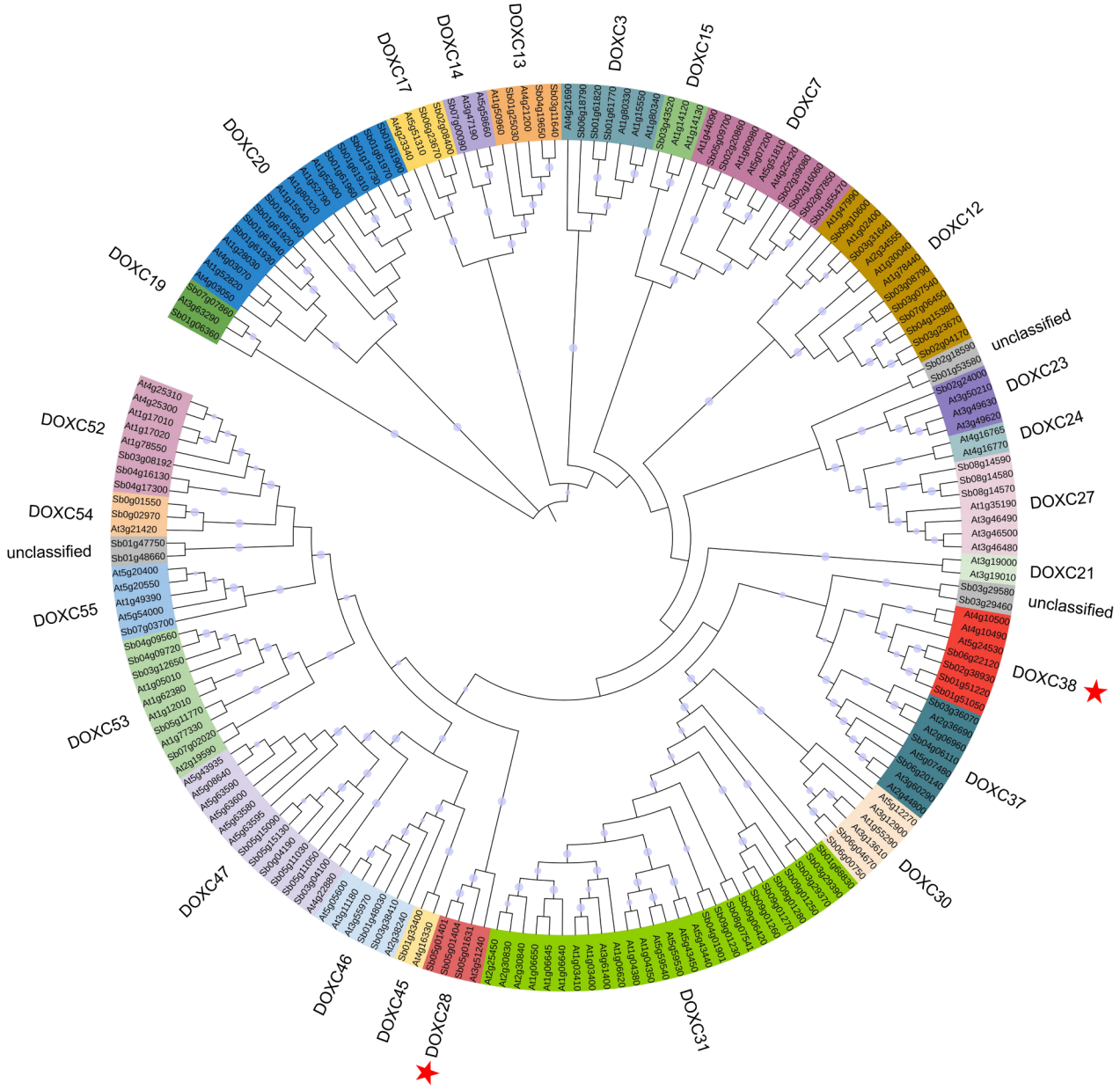


Fig. 2 Phylogenetic analysis of 2ODDs proteins
Clades were annotated with 2ODDs of AtDOXC. The neighbor joining method was used to construct the tree with bootstrap ($n = 1000$). A gray circle in the middle of each branch indicated bootstrap values greater than 0.5. Red asterisks indicated the DOXC28 and DOXC38 clades for further study

(DOXC20 and DOXC31), alkaloid metabolism (DOXC31, DOXC41 and DOXC52), salicylic acid catabolism (DOXC38), coumarin biosynthesis (DOXC30), flavonoid biosynthesis (DOXC28 and DOXC47), ethylene biosynthesis (DOXC53) [16]. However, further research is needed to determine the functional classification of other clades (DOXC14, DOXC17, DOXC19, DOXC21, DOXC23, DOXC24, DOXC27, DOXC37, DOXC45, DOXC46, DOXC54 and DOXC55).

Chromosomal location of *Sb2ODDs* in *S. baicalensis*

We conducted chromosomal localization mapping of *Sb2ODDs* using gene annotation files and found that they were distributed unevenly across the *S. baicalensis* genome (Fig. 3). Chromosomes 01–09 contained 22, 9, 15, 8, 9, 6, 5, 4 and 7 *Sb2ODDs*, respectively. Additionally, 3 *Sb2ODDs* were not anchored onto any specific chromosomes. Some *Sb2ODDs* were located in close proximity to specific regions on the chromosomes, indicating possible tandem gene duplication events. A gene cluster is defined as the existence of two neighboring *Sb2ODD*

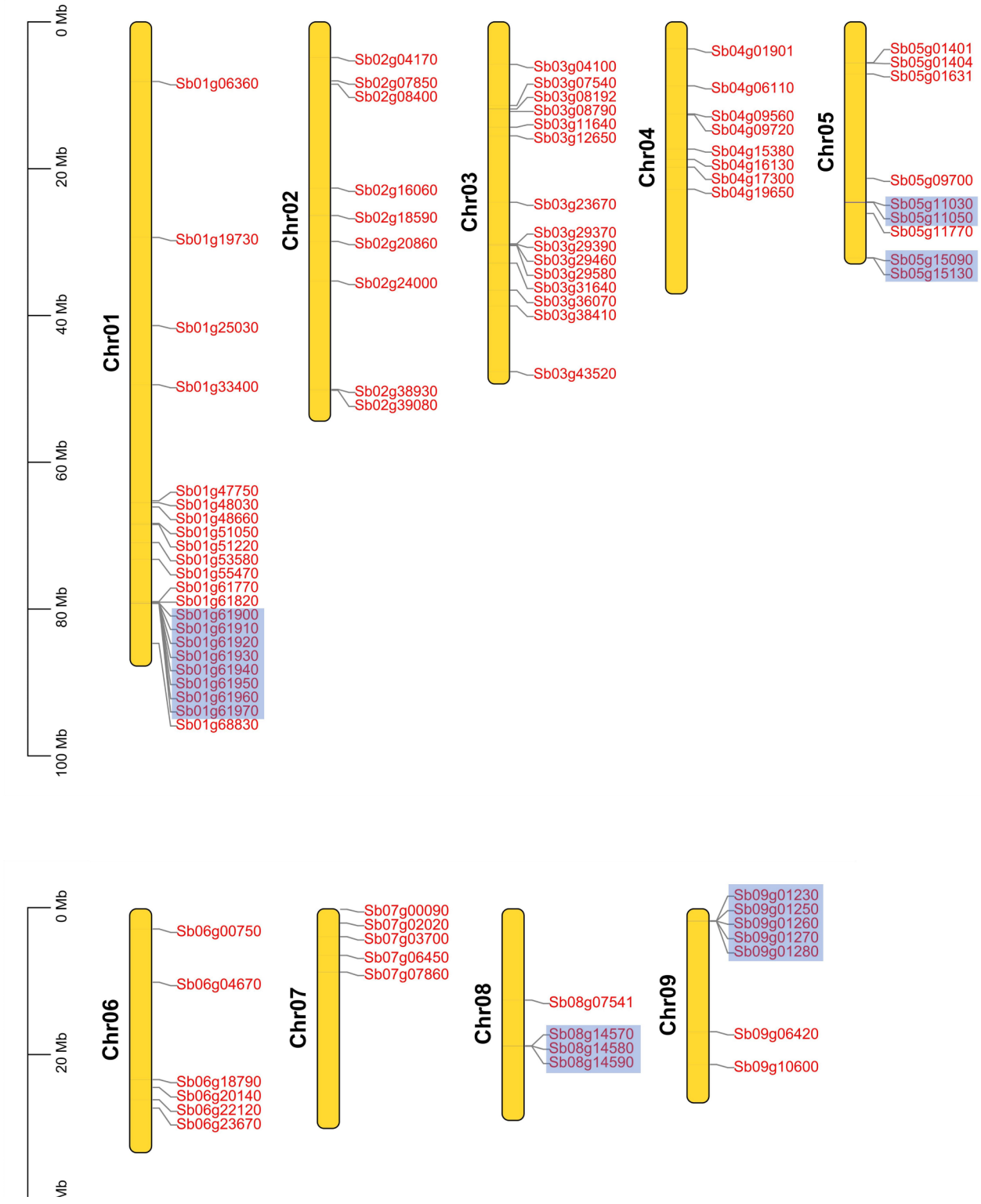


Fig. 3 Chromosomal distribution of *Sb2ODDs* in *S. baicalensis*
Yellow bars represent pseudochromosomes, blue rectangles represent that the distance between *Sb2ODD* genes in a gene cluster on the chromosome is less than 50 kb

genes on a chromosome with a distance of less than 50 kb [16]. According to the annotation of DOXC, these gene clusters located on chromosomes 01, 05, 08 and 09 were found to be involved in flavonoid biosynthesis, glucosinolate metabolism and alkaloid metabolism, respectively.

Tissue-specific expression patterns of *Sb2ODDs* in *S. baicalensis*

The expression patterns of the *Sb2ODDs* were determined based on FPKM values from the transcriptome of flowers, flower buds, leaves, stems, roots and MeJA-treated roots of *S. baicalensis* [6]. These *Sb2ODDs* were classified into five groups based on their tissue-specific expression patterns (Fig. 4 and Table S1). Group A consisted of only 1 *Sb2ODD*, which exhibited high expression levels across all the tissues. Group B showed relatively high expression levels specifically in flowers and flower buds. Notably, DOXC47, which includes anthocyanidin synthase (ANS), was clustered in Group B, suggesting a potential association between these genes and anthocyanidin biosynthesis in *S. baicalensis* flowers. Groups C and D displayed low transcript levels across all the tissues, with Group D showing relatively higher levels compared to Group C. Group E showed predominant expression in roots and their transcripts could be induced by MeJA treatment, indicating its potential involvement in the accumulation of specific flavones in the roots of *S. baicalensis*, as we previously showed that MeJA treatment leads to increased flavones in *S. baicalensis* roots [1].

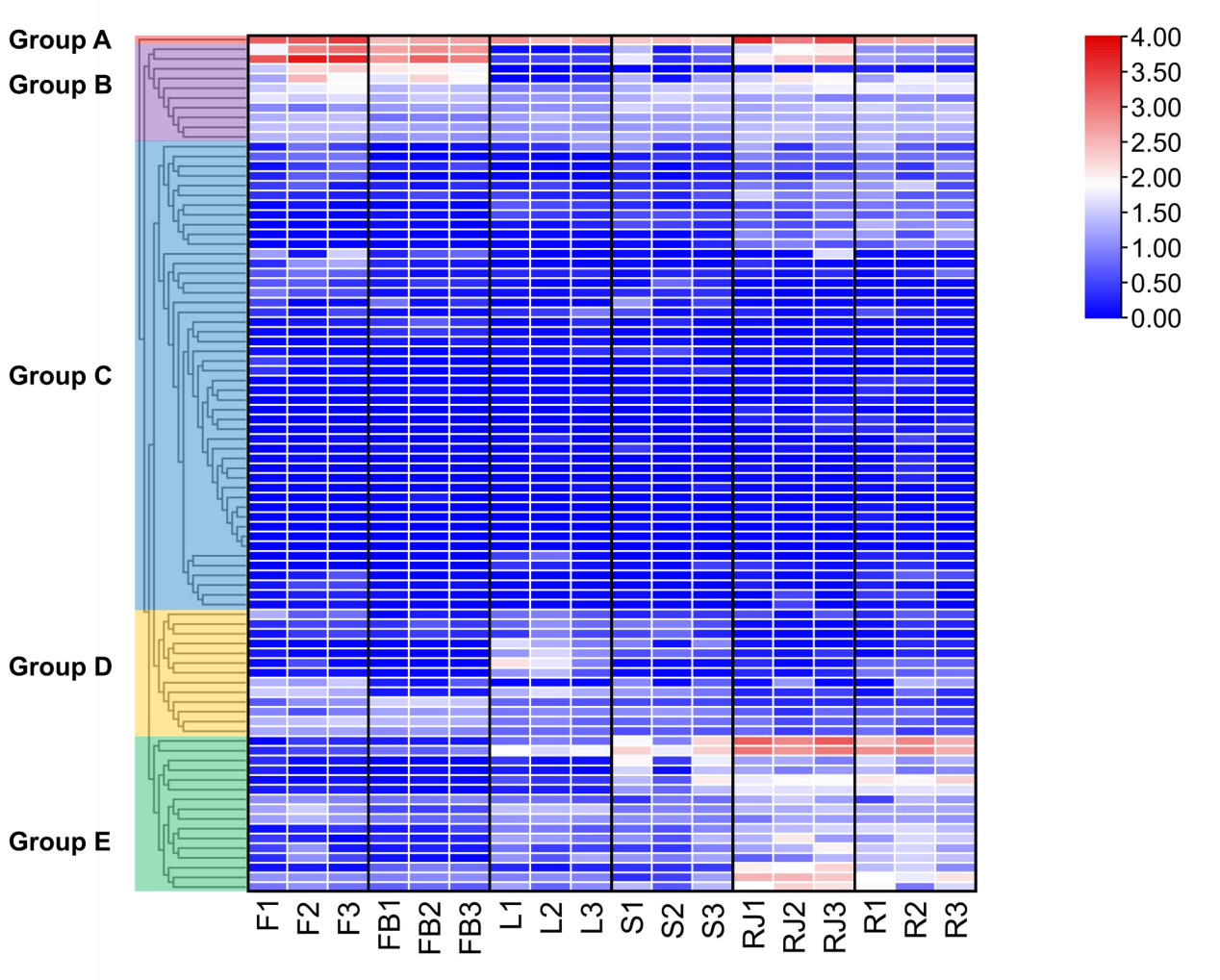


Fig. 4 Tissue-specific expression heatmap of *Sb2ODDs* in *S. baicalensis*
The color scale on the right represents the FPKM values normalized with log10. F, flower; FB, flower bud; L, leaf; S, stem; RJ, MeJA-treated root; R, root; the numbers behind indicated biological replicates. Colored rectangles represent genes clustered in different groups (group A-E) based on their expression patterns

Identification of the candidate genes encoding FNSI and F3H in *S. baicalensis*

The DOXC28 clade includes the classic FNSI and F3H, which are involved in the biosynthesis of flavones and dihydroflavonols, respectively [16]. It has been reported that AtDMR6 from the DOXC38 clade can also be involved in the biosynthesis of flavone and the catabolism of salicylic acid [23, 24]. To analyze the phylogenetic relationship of the Sb2ODDs in DOXC38 and DOXC28, a phylogenetic tree was constructed with previously studied enzymes from other species (Fig. 5A). The results showed that 4 putative Sb2ODDs (Sb2ODD1-Sb2ODD4) were clustered with AtDMR6 (AtFNSI/S5H) on the DOXC38 clade, while Sb2ODD5-Sb2ODD7 clustered with AtF3H and PcFNSI on the DOXC28 clade (Fig. 5A).

Sequence alignment of Sb2ODDs and other homologous proteins revealed that they all had ferrous iron binding domain (HxDxnH) and 2-oxoglutarate binding domain (RxS, Fig. S1). Expression heat-map based on FPKM values showed that *Sb2ODD1* was highly expressed in roots and JA-treated roots, while *Sb2ODD3* and *Sb2ODD4* had relatively higher transcripts in root tissues, *Sb2ODD2* had very low expression levels in all the tissues analyzed. The expression of *Sb2ODD5*, *Sb2ODD6* and *Sb2ODD7* were higher in flowers and flower buds than in other tissues. Therefore, Sb2ODD1, 3, 4 and Sb2ODD5-7 might be involved in the biosynthesis of flavones and anthocyanidins in roots and flowers, respectively. We isolated the ORFs of the 7 *Sb2ODDs* by RT-PCR (Table S2).

Enzyme assays of recombinant Sb2ODD proteins

To explore the activities of the Sb2ODDs, we expressed coding regions of the 7 enzymes in yeast. As AtDMR6 in DOXC38 can converse naringenin and salicylic acid to apigenin and 2, 5-dihydroxybenzoic acid (2, 5-DHBA) [23]. We assayed the enzymes of Sb2ODD1-4 by feeding naringenin, salicylic acid and pinocembrin as substrates. The yeast strains expressing Sb2ODD1 and Sb2ODD2 fermented with pinocembrin produced new compound (Peak I) with the same retention time as the chrysin standard (Fig. 6A). We then expressed the proteins in *E. coli* and purified them from the strains, the enzyme activities were also confirmed through in vitro enzyme assay. Liquid chromatography-mass spectrometry (LC-MS) analysis showed that peaks I had the same mass charge

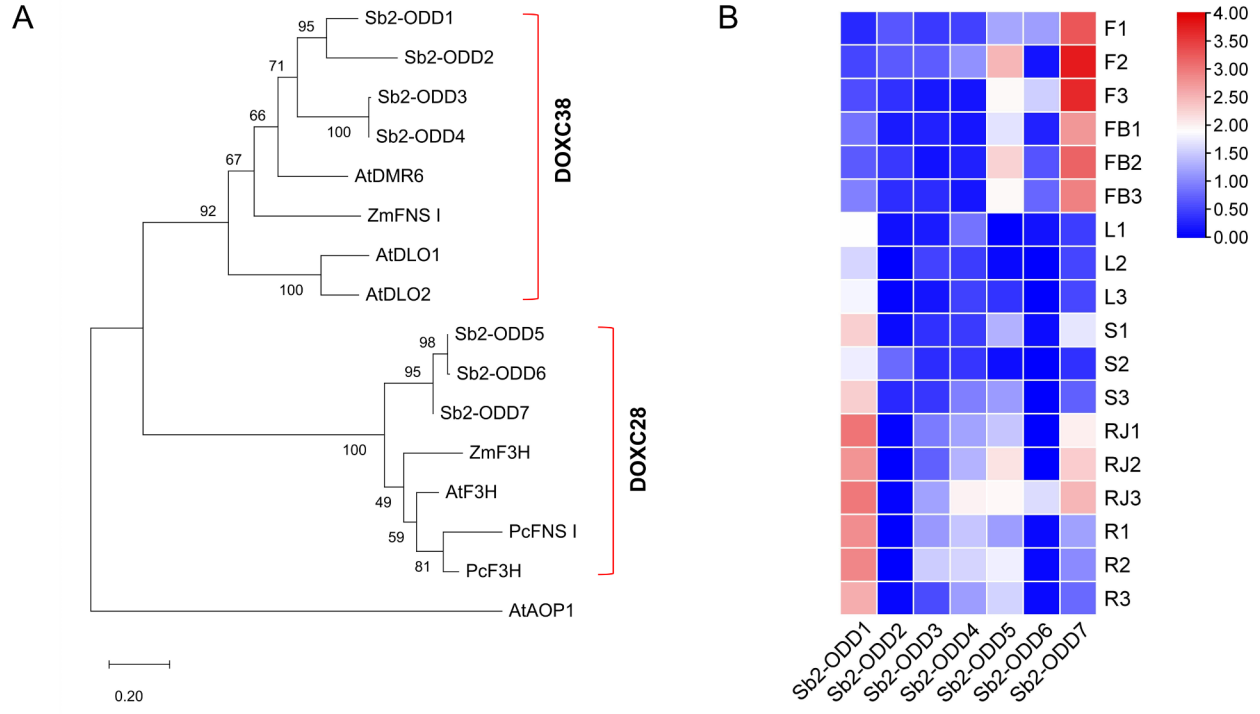
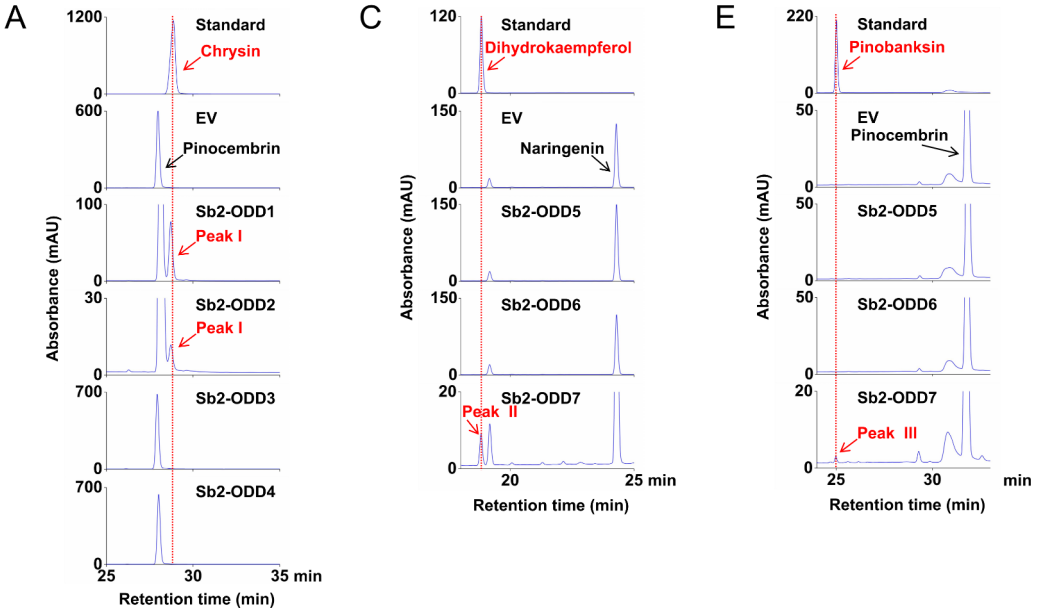
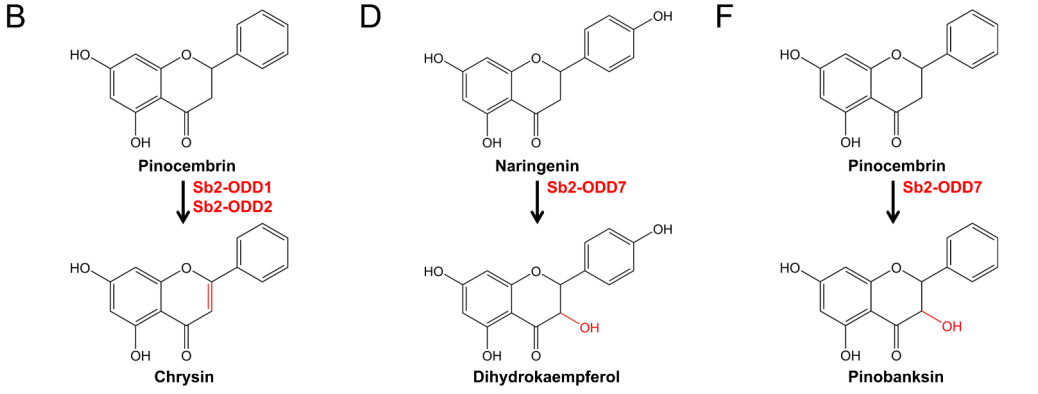


Fig. 5 Phylogenetic analysis and expression patterns of *SbDOXC28s* and *SbDOXC38s* in *S. baicalensis*
Phylogenetic tree of SbDOXC28 and SbDOXC38 proteins. The maximum likelihood method was used to construct the tree with bootstrap ($n = 1000$). AtAOP1 from At2ODD was used as an outgroup. At, *Arabidopsis thaliana*; Zm, *Zea mays*; Pc, *Petroselinum crispum*. Accession numbers are as follows: AtDMR6, NP_197841; AtDLO1, NP_192788; AtDLO2, NP_192787; ZmFNSI, NP_001151167; PcFNSI, AAX21541; AtF3H, NP_190692; ZmF3H, NP_001130275; PcF3H, AAP57394; AtAOP1 (NM_116541); Sb2ODD1, Sb06g22120; Sb2ODD2, Sb02g38930; Sb2ODD3, Sb01g51050; Sb2ODD4, Sb01g51220; Sb2ODD5, Sb05g01401; Sb2ODD6, Sb05g01404; Sb2ODD7, Sb05g01631. B. Tissue-specific expression heatmap of *SbDOXC28s* and *SbDOXC38s*. The color scale on the right represents the FPKM values normalized with log10. F, flower; FB, flower bud; L, leaf; S, stem; RJ, MeJA-treated root; R, root; the numbers behind indicated biological replicates





G

Enzyme	Substrate	*K*m (μM)	*V*max (pkat mg protein^{-1})	*V*max/*K*m
Sb2-ODD7	naringenin	57.48	2032	35.35
Sb2-ODD7	pinocembrin	36.55	114.1	3.12

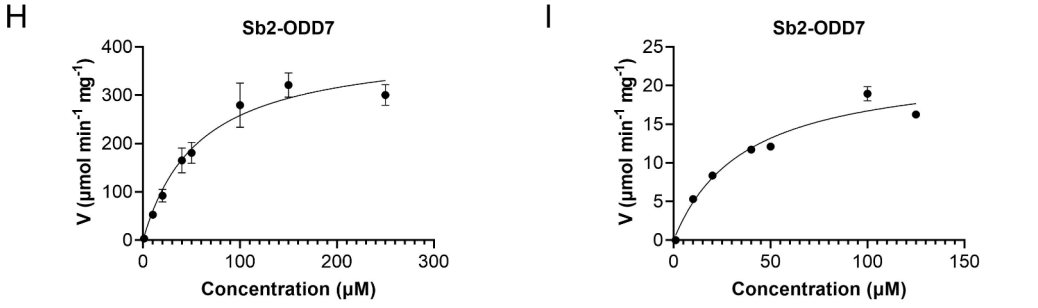


Fig. 6 (See legend on next page.)

(See figure on previous page.)

Fig. 6 Enzyme assays of Sb2ODDs

(**A**) HPLC analysis of Sb2ODD1-4 from DOXC38 using pinocembrin as a substrate in vivo yeast enzyme assays. Top, chrysin standard; EV, empty vector control; Sb2ODD1-4, assays with corresponding Sb2ODD proteins. (**B**) The reaction was catalyzed by Sb2ODD1-2 from DOXC38 using pinocembrin as a substrate. (**C**) HPLC analysis of Sb2ODD5-7 from DOXC28 using naringenin as a substrate in vivo yeast enzyme assays. Top, dihydrokaempferol standard; EV, empty vector control; Sb2ODD5-7, assays with corresponding Sb2ODD proteins. (**D**) The reaction was catalyzed by Sb2ODD7 from DOXC28 using naringenin as a substrate. (**E**) HPLC analysis of Sb2ODD5-7 from DOXC28 using pinocembrin as a substrate in vivo yeast enzyme assays. Top, pinobanksin standard; EV, empty vector control; Sb2ODD5-7, assays with corresponding Sb2ODD proteins. (**F**) The reaction was catalyzed by Sb2ODD7 from DOXC28 using pinocembrin as a substrate. (**G**) Kinetic analysis of Sb2ODD7 to naringenin and pinocembrin. (**H**) Enzymatic kinetic curve of Sb2ODD7, naringenin as substrate. (**I**) Enzymatic kinetic curve of Sb2ODD7, pinocembrin as substrate

ratio (m/z) and MS/MS patterns as the chrysin standard (Fig. S2A, B). Sb2ODD1 and Sb2ODD2 could specifically catalyze the conversion of pinocembrin to chrysin, but no new products were detected when naringenin was used as substrate (Fig. 6B and Fig. S3A, B). The FNSI's activity (catalyzed 4′-deoxyflavanone) of Sb2ODD1 and Sb2ODD2 in *S. baicalensis* were different from the classic FNSIs (catalyzed 4′-hydroxyflavanone) in other species, as the previously reported FNSIs use naringenin as natural substrate. When we fed yeasts with salicylic acid as a substrate, no new products were detected by high-performance liquid chromatography (HPLC), but a very tiny peak was detected in the yeast extracts of Sb2ODD1 and Sb2ODD2 in mass spectrometry (MS), which had the same MS patterns as 2, 5-DHBA standard (Fig. S3C, D, E). Compared with the empty vector (EV) control, the transformed yeast cells expressing Sb2ODD3 and Sb2ODD4 showed no activity to the 3 substrates (Fig. 6A and Fig. S3A, C).

We then supplemented flavanones naringenin and pinocembrin into the yeasts transformed with Sb2ODD5-7, respectively, and two new products (Peak II and Peak III) were detected in the strains expressing Sb2ODD7, which had the same retention time and MS/MS patterns as dihydrokaempferol and pinobanksin standards (Fig. 6C, E), showing it is a F3H. The results were also verified through in vitro enzyme activity experiments, (Fig. S2C, D, E, F). However, Sb2ODD5 and Sb2ODD6 had no activity to naringenin and pinocembrin substrates (Fig. 6C, E). Therefore, Sb2ODD7 was involved in the biosynthesis of dihydroflavonols, which not only catalyzed naringenin to dihydrokaempferol, but also catalyzed pinocembrin to pinobanksin (Fig. 6D, F).

Kinetic analysis of Sb2ODD7

We then conducted a study to explore the kinetic parameters of Sb2ODD7 to naringenin and pinocembrin (Fig. S4). The results demonstrated that the apparent Michaelis constant (*K*m) values of Sb2ODD7 were 57.48 μM and 36.55 μM for naringenin and pinocembrin, respectively. Additionally, the apparent maximal velocity (*V*max) values were 2032 pkat mg protein^{-1} and 114.1 pkat mg protein^{-1} for naringenin and pinocembrin, respectively (Fig. 6G, H, I). While Sb2ODD7 exhibited a lower *K*m for pinocembrin, it displayed significantly higher *V*max for naringenin, resulting in an 11.33-fold higher *V*max/*K*m ratio for naringenin compared to pinocembrin. So Sb2ODD7 had a higher catalytic efficiency to naringenin. Moreover, the expression level of Sb2ODD7 was found to be higher in flowers than in roots, so this enzyme is involved in the conversion of naringenin in flowers of *S. baicalensis*, subsequently entering the biosynthesis pathway of anthocyanidins and flavonols.

Discussion

S. baicalensis is known to contain a variety of flavonoids, including classic 4′-hydroxyflavonoids and root-specific 4′-deoxyflavonoids [1]. Flavonoids exhibit significant pharmacological activity and have been utilized to treat a variety of diseases. Recently, baicalein and myricetin have been reported to possess anti-COVID-19 properties [5]. The DOXC subfamily of 2ODD family is involved in the biosynthesis of the flavonoids, which are of particular interest to us. The number of *2ODD* genes varies in each plant genome, for example, 7, 49, 56, and 99 *2ODD* genes of the DOXC subfamily have been identified in *Chlamydomonas reinhardtii*, *Physcomitrella patens*, *Selaginella moellendorffii* and *Arabidopsis thaliana* respectively [16]. In our study, 88 *2ODDs* of the DOXC subfamily were identified by searching the genome of *S. baicalensis* [6, 30]. Compared to *A. thaliana*, *S. baicalensis* lacks the DOXC21 and DOXC24 clades (Fig. 2). Most members of the DOXC subfamily are involved in the biosynthesis of specialized metabolites, such as flavonoids and phytohormones [16]. This suggests that there has been a large-scale duplication of *2ODD* genes related to specific metabolism, occurring from green algae to higher plants. In *S. baicalensis*, we found *Sb2ODD* gene tandem duplications on chromosomes 01, 05, 08, and 09, which are annotated as being associated with flavonoid biosynthesis, glucosinolate metabolism, and alkaloid metabolism, respectively (Fig. 3). The *Sb2ODD* gene tandem duplication (Sb05g11050) on chromosome 05 shows relatively high expression levels in flowers and flower buds, it was annotated as FLS belonging to DOXC47 clade, which is most likely to be involved in flavonol biosynthesis (Fig. 4).

Pinocembrin and naringenin are central intermediates in the biosynthesis of root-specific flavonoids (4′-deoxyflavonoids) and classic flavonoids (4′-hydroxyflavonoids), respectively [1, 31]. There are two types of FNSs in

plants that convert flavanones to flavones, namely FNSI (2ODD) and FNSII (CYP450). In *S. baicalensis*, a FNSII member enzyme (SbFNSII-2) can specifically catalyze the conversion of pinocembrin into chrysin in the roots [1]. FNSI was firstly characterized in parsley (PcFNSI), this gene belongs to the DOXC28 clade, catalyzes the desaturation reaction of naringenin to form apigenin [21]. However, most members from the DOXC28 clade exhibit only F3H enzyme activity (Fig. 5A). On the other hand, Arabidopsis AtDMR6, belonging to the DOXC38 clade, has been discovered to hydroxylate salicylic acid and plays a role in the degradation of plant hormones [23, 24], while also exhibiting tiny FNSI activity. So the 2ODD enzymes from both DOXC28 and DOXC38 may be involved in flavonoids metabolism. We then isolated all the *Sb2ODDs* from the two clades. Phylogenetic analysis showed that four Sb2ODDs formed a cluster with AtDMR6 in the DOXC38 clade, while three Sb2ODDs were grouped with AtF3H and PcFNSI in the DOXC28 clade (Fig. 5A). Notably, *Sb2ODD1* and *Sb2ODD7* exhibited high expression levels in the roots and flowers, respectively, suggesting their different function in the biosynthesis of flavonoids in the different organs (Fig. 5B).

The *Sb2ODDs* that belong to the DOXC38 encode proteins homologous to the AtDMR6 protein (Table S2). Interestingly, while Sb2ODD1 and Sb2ODD2 were involved in the conversion of pinocembrin to chrysin, they did not exhibit any catalytic activity towards naringenin (Fig. 6A, B and Fig. S3A). Unlike AtDMR6, which mainly takes salicylic acid as the substrate, Sb2ODD1 and Sb2ODD2 can only produce a tiny amount of 2, 5-DHBA when salicylic acid is supplemented (Fig. S3C, D, E). Our previous studies have demonstrated that SbFNSII-2, belonging to the CYP450 family, is primarily responsible for the biosynthesis of chrysin in the *S. baicalensis* roots. Now, our results show that Sb2ODD1 and Sb2ODD2 also have flavone synthase activity (FNSI). As *Sb2ODD1* is highly expressed in the roots, while *Sb2ODD2* has very low transcripts in all the tissues studied, *Sb2ODD1* may function as a supplement to participate in the root-specific biosynthesis of 4′-deoxyflavones in *S. baicalensis*. The FNSI activity has been found in mosses, gymnosperms, monocotyledons, and dicotyledons, and the enzymes belong to different 2ODD subfamilies, indicating FNSI might have undergone multiple independent evolutionary events, to resist biotic or abiotic stress during land plant colonization and radiation [32].

There are 3 *Sb2ODDs* belonging to the DOXC28 clade (Table S2). Sb2ODD7 could catalyze the conversion of naringenin (4′-hydroxyflavanone) and pinocembrin (4′-deoxyflavanone) into dihydrokaempferol and pinobanksin, respectively, indicating that the enzyme is a typical F3H (Fig. 6C, E). Kinetic analysis revealed that Sb2ODD7 possessed high catalytic efficiency towards naringenin, producing dihydrokaempferol (Fig. 6G), the precursor of anthocyanidin and myricetin. *Sb2ODD7* was highly expressed in flowers and flower buds (where anthocyanins accumulate), suggesting this enzyme is involved in anthocyanins biosynthesis (Figs. 1 and 5B) in flowers of *S. baicalensis*.

Conclusions

Our work provides a genome-wide analysis of the *2ODDs* of DOXC subfamily in the *S. baicalensis* genome. We identified 88 *2ODD* genes of DOXC subfamily. We also performed functional analysis of Sb2ODDs involved in the biosynthesis of flavones and dihydroflanonols. These findings complement the flavonoid biosynthesis pathway in *S. baicalensis*, providing genetic resources for the modification of biological chassis and large-scale production of medicinal active ingredients in synthetic biology.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-024-05519-1.

Supplementary Material 1: **Table S1**. The list of gene locus and their groups from DOXC clade in *S. baicalensis*. **Table S2**. The list of enzyme names, gene locus and primers used for cloning of predicted Sb2ODD in *S. baicalensis*.

Supplementary Material 2: **Fig. S1.** The alignment of Sb2ODDs with FNSI and F3H sequences from other species. **Fig. S2.** *In vitro* enzyme assays of Sb2ODDs. **Fig. S3.** *In vivo* yeast enzyme assays of Sb2ODDs. **Fig. S4.** SDS PAGE analysis of purification of Sb2ODD1 and Sb2ODD7 proteins.

Acknowledgements

We greatly appreciate the experimental facilities and services provided by the office of Chenshan Plant Science Research Center. We also thank Xingguo Li from National Key Laboratory of Wheat Breeding, College of Life Sciences, Shandong Agricultural University for his suggestions on the manuscript.

Author contributions

Q.Z. conceptualized and designed the project. S.M.Z. and M.Y.C. performed the experiments. All the authors analyzed and interpreted the data. S.M.Z. wrote the manuscript. M.Y.C. and Q.Z. revised the manuscript. All authors read and approved the final manuscript.

Funding

This work is sponsored by Natural Science Foundation of Shanghai (22ZR1479500), Special Fund for Scientific Research of Shanghai Landscaping & City Appearance Administrative Bureau (G242403, G212401), Ministry of Science and Technology of China (YDZX20223100001003), and funding for Shanghai science and technology promoting agriculture from Shanghai Agriculture and Rural Affairs Commission (Hu Nong Ke Chan Zi(2023) No.8).

Data availability

The DNA and the protein sequences from *S. baicalensis* are provided in Table S2. Protein sequences from A. thaliana are available with the link of http://www.plants.ensembl.org/index.html. RNA sequencing data are available in the Sequence Read Archive (SRA) database with the link of www.ncbi.nlm.nih.gov/sra, under the accession number SRP156996. The genome of *S. baicalensis* is available in the National Genomics Data Center (https://bigd.big.ac.cn/gwh) with accession number GWHAOTC00000000.All the data supporting our findings are contained within the manuscript, in text, tables and figures and in the supplementary files.

Ethics approval and consent to participate
All methods were carried out in accordance with local and national guidelines and regulations.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 19 January 2024 / Accepted: 14 August 2024
Published online: 26 August 2024

References

1. Zhao Q, Zhang Y, Wang G, Hill L, Weng JK, Chen XY, et al. A specialized flavone biosynthetic pathway has evolved in the medicinal plant, *Scutellaria baicalensis*. Sci Adv. 2016;2(4):1501780.
2. Shang XF, He XR, He XY, Li MX, Zhang RX, Fan PC, et al. The genus *Scutellaria* an ethnopharmacological and phytochemical review. J Ethnopharmacol. 2010;128(2):279–313.
3. Wang ZL, Wang S, Kuang Y, Hu ZM, Qiao X, Ye M. A comprehensive review on phytochemistry, pharmacology, and flavonoid biosynthesis of *Scutellaria baicalensis*. Pharm Biol. 2018;56(1):465–84.
4. Liu HB, Ye F, Sun Q, Liang H, Li CM, Li SY, et al. *Scutellaria baicalensis* extract and baicalein inhibit replication of SARS-CoV-2 and its 3 C-like protease *in vitro*. J Enzyme Inhib Med Chem. 2021;36(1):497–503.
5. Su HX, Yao S, Zhao WF, Zhang YM, Liu J, Shao Q, et al. Identification of pyrogallol as a warhead in design of covalent inhibitors for the SARS-CoV-2 3CL protease. Nat Commun. 2021;12(1):3623.
6. Zhao Q, Yang J, Cui MY, Liu J, Fang YM, Yan MX, et al. The reference genome sequence of *Scutellaria baicalensis* provides insights into the evolution of wogonin biosynthesis. Mol Plant. 2019;12(7):935–50.
7. Wen W, Alseekh S, Fernie AR. Conservation and diversification of flavonoid metabolism in the plant kingdom. Curr Opin Plant Biol. 2020;55:100–8.
8. Santos-Buelga C, Feliciano AS. Flavonoids: from structure to health issues. Mol. 2017;22(3):477.
9. Sheehan H, Moser M, Klahre U, Esfeld K, Dell'Olivo A, Mandel T, et al. MYB-FL controls gain and loss of floral UV absorbance, a key trait affecting pollinator preference and reproductive isolation. Nat Genet. 2016;48(2):159–66.
10. Tian F, Jia TJ, Yu BJ. Physiological regulation of seed soaking with soybean isoflavones on drought tolerance of *Glycine max* and *Glycine soja*. Plant Growth Regul. 2014;74(3):229–37.
11. Bhatia C, Pandey A, Gaddam SR, Hoecker U, Trivedi PK. Low temperature-enhanced flavonol synthesis requires light-associated regulatory components in *Arabidopsis thaliana*. Plant Cell Physiol. 2018;59(10):2099–112.
12. Malhotra B, Onyilagha JC, Bohm BA, Towers GHN, James D, Harborne JB, et al. Inhibition of tomato ringspot virus by flavonoids. Phytochemistry. 1996;43(6):1271–6.
13. Bieza K, Lois R. An Arabidopsis mutant tolerant to lethal ultraviolet-B levels shows constitutively elevated accumulation of flavonoids and other phenolics. Plant Physiol. 2001;126(3):1105–15.
14. Liu WX, Feng Y, Yu SH, Fan ZQ, Li XL, Li JY, et al. The flavonoid biosynthesis network in plants. Int J Mol Sci. 2021;22(23):12824.
15. Farrow SC, Facchini PJ. Functional diversity of 2-oxoglutarate/Fe(II)-dependent dioxygenases in plant metabolism. Front Plant Sci. 2014;5:524.
16. Kawai Y, Ono E, Mizutani M. Evolution and diversity of the 2-oxoglutarate-dependent dioxygenase superfamily in plants. Plant J. 2014;78(2):328–43.
17. Kataoka H, Yamamoto Y, Sekiguchi M. A new gene (alkB) of *Escherichia coli* that controls sensitivity to methyl methane sulfonate. J Bacteriol. 1983;153(3):1301–7.
18. Keskiaho K, Hieta R, Sormunen R, Myllyharju J. Chlamydomonas reinhardtii has multiple prolyl 4-hydroxylases, one of which is essential for proper cell wall assembly. Plant Cell. 2007;19(1):256–69.
19. Araujo WL, Martins AO, Fernie AR, Tohge T. 2-oxoglutarate: linking TCA cycle function with amino acid, glucosinolate, flavonoid, alkaloid, and gibberellin biosynthesis. Front Plant Sci. 2014;5:552.
20. Lee YJ, Kim JH, Kim BG, Lim Y, Ahn JH. Characterization of flavone synthase I from rice. BMB Rep. 2008;41(1):68–71.
21. Martens S, Forkmann G, Matern U, Lukacin R. Cloning of parsley flavone synthase I. Phytochemistry. 2001;58(1):43–6.
22. Gebhardt YH, Witte S, Steuber H, Matern U, Martens S. Evolution of flavone synthase I from parsley flavanone 3beta-hydroxylase by site-directed mutagenesis. Plant Physiol. 2007;144(3):1442–54.
23. Ferreyra MLF, Emiliani J, Rodriguez EJ, Campos-Bermudez VA, Grotewold E, Casati P. The identification of maize and Arabidopsis type I flavone synthases links flavones with hormones and biotic interactions. Plant Physiol. 2015;169(2):1090–107.
24. Zhang YJ, Zhao L, Zhao JZ, Li YJ, Wang JB, Guo R, et al. S5H/DMR6 encodes a salicylic acid 5-hydroxylase that fine-tunes salicylic acid homeostasis. Plant Physiol. 2017;175(3):1082–93.
25. Eddy SR. Profile hidden Markov models. Bioinformatics. 1998;14(9):755–63.
26. Tamura K, Stecher G, Kumar S, Battistuzzi FU. MEGA11: molecular evolutionary genetics analysis version 11. Mol Biol Evol. 2021;38(7):3022–7.
27. 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(8):1194–202.
28. Cui MY, Lu AR, Li JX, Liu J, Fang YM, Pei TL, et al. Two types of *O*-methyltransferase are involved in biosynthesis of anticancer methoxylated 4'-deoxyflavones in *Scutellaria baicalensis* Georgi. Plant Biotechnol J. 2022;20(1):129–42.
29. Wang DF, Wang JR, Wang YF, Yao DZ, Niu YB. Metabolomic and transcriptomic profiling uncover the underlying mechanism of color differentiation in *Scutellaria baicalensis* Georgi. Flowers Front Plant Sci. 2022;13:884957.
30. Pei TL, Zhu SM, Liao WZ, Fang YM, Liu J, Kong Y, et al. Gap-free genome assembly and *CYP450* gene family analysis reveal the biosynthesis of anthocyanins in *Scutellaria baicalensis*. Hortic Res. 2023;10(12):uhad235.
31. Martens S, Mithofer A. Flavones and flavone synthases. Phytochemistry. 2005;66(20):2399–407.
32. Li DD, Ni R, Wang PP, Zhang XS, Wang PY, Zhu TT, et al. Molecular basis for chemical evolution of flavones to flavonols and anthocyanins in land plants. Plant Physiol. 2020;184(4):1731–43.

Publisher's note