Preprints.org

Article

Phylogeny Analysis and Functional Characterizatiion of CYP76AHs from *Salvia*

Zhenyu Zhao , Dongfeng Yang , Juan Guo , Xiuyu Liu , Qishuang Li , Ping Su , Jian Wang , Ying Ma [*] , Luqi Huang [*]

Article

Phylogeny Analysis and Functional Characterization of CYP76AHs from *Salvia*

Zhenyu Zhao [1], Dongfeng Yang [2], Juan Guo [1], Xiuyu Liu [3], Qishuang Li [4], Ping Su [1], Jian Wang [1], Ying Ma [1,*] and Luqi Huang [1,*]

[1] State Key Laboratory of Dao-di Herbs, National Resource Center for Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China; zhaozhenyu9607@163.com (Z.Z.); guojuanzy@163.com (J.G.); suping120@163.com (P.S.); 1157873501@qq.com (J.W.)

[2] College of Life Sciences and Medicine, Key Laboratory of Plant Secondary Metabolism and Regulation in Zhejiang Province, Zhejiang Sci-Tech University, Hangzhou, China; ydf807@sina.com (D.Y.)

[3] College of pharmacy, Henan university of chinese medicine, Zhengzhou 450008, China; liuxiuyuzy@163.com (X.L.)

[4] School of Traditional Chinese Pharmacy, China Pharmaceutical University, Nanjing 210009, China. liqs556@163.com (Q.L.)

* Correspondence: xiaoma1110@126.com (Y.M.); huangluqi01@126.com (L.H.)

Abstract: Salvia is a large genus in traditional Chinese medicine with hundreds of species. Tanshinones are a highly representative class of exclusive compounds found in the Salvia genus that exhibit significant biological activity. Tanshinone components have been identified in 16 species of the Salvia genus. The CYP76AH subfamily (P450) is crucial for the synthesis of tanshinone due to its catalytic generation of polyhydroxy structures. 420CYP76AH genes were obtained, clear clustering relationships. In this study 15 CYP76AH genes from 10 Salvia species were cloned and studied from the perspectives of evolution and catalytic efficiency. Structure-function relationship revealed several conserved residues which might be related to the function of CYP76AHs. Amino acids (97I, 116V, and 290T) are conserved in CYP76AH1 group, while amino acids (117G, 208S, 237Y, and 479F) are CYP76AH3 group protein sequence characteristics with unique enzyme function.This research holds significant implications for the classification of the Salvia genus and provide efficient catalytic elements for the synthetic biology production of tanshinones.

Keywords: Salvia; CYP76AH; P450; biosynthesis; functional gene
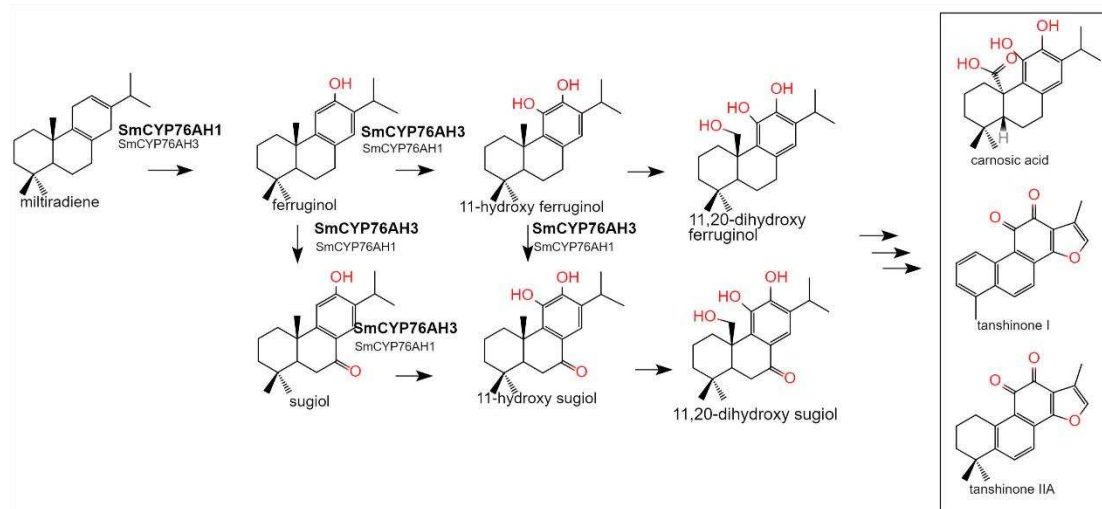
1. Introduction

Salvia is a vast genus of medicinal plants, comprising at least 1000 species and distributed worldwide, 78 of which are native to China. However, variations have been observed in the quantity and type of secondary metabolites found in different *Salvia* species [1]. The most characteristic compounds in *Salvia* is abietane diterpenoids, which comprises tanshinones and has a variety of biological activities, particularly antioxidant benefits and pharmacological actions for treating heart disease [2]. *Salvia* contains a variety of Chinese herbal medicines, including Danshen (*Salvia miltiorrhiza*), Diandanshen (*Salvia yunnanensis*) and so on, most of them have been reported the accumulation of tanshinones. Tanshinones are one of the main active ingredients of Danshen in the treatment of cardiovascular diseases, with various pharmacological activities [3]. Tanshinone IIA is a well-known monomer that exhibits substantial biological activity in the treatment of neurological illnesses and significant pharmacological effects in the treatment of cardiovascular diseases [4,5]. The differences in metabolic components are significantly correlated with the types and expression levels of biosynthetic pathway genes. Therefore, *Salvia* plants are very suitable for conducting correlation studies on chemical diversity and gene differences due to their large population and wide distribution.

CYP450 (Cytochrome P450) is essential in the biosynthesis of natural active ingredients in plants, participating in the biosynthesis of most terpenoid active ingredients [5]. Due to the high degree of

oxidation of tanshinones, multiple P450s are required to participate in their biosynthesis pathway, CYP76AH subfamily P450s plays a significant role in the synthesis of tanshinones [6]. The biological functions of CYP76AH subfamily genes in Labiatae medicinal plants have been continuously reported in recent years, SmCYP76AH1 in *S. miltiorrhiza* was identified as ferruginol synthase that catalyze the hydroxylation of C12 of miltiradiene. The biological functions of SmCYP76AH3 and SmCYP76AK1 were screened with engineering yeast construction and mature hairy root genetic transformation systems. SmCYP76AH3 is a hybrid catalytic P450[7], which can catalyze the hydroxylation of ferruginol at C11 and carbonylation of ferruginol at C7 to generate 11-hydroxy ferruginol, sugiol and 11-hydroxy sugiol. The carbonylation function of C7 leads to the possible extension of the biosynthetic pathway of tanshinone to another branch, thus forming a complex catalytic network (Figure 1). SmCYP76AK1 is one of the key enzymes in the biosynthesis pathway of tanshinone compounds, with specific C20 hydroxylation function [8]. CYP76AH4, CYP76AH22-24, and CYP76AH57 were discovered in *S. fruticosa* and *Rosmarinus officinalis*, the catalytic function are similar to that of SmCYP76AHs [9]. In addition, *CfCYP76AH15*, *CfCYP76AH11*, and *CfCYP76AH16*, which were cloned from the cortex of *Coleus forskohlii* and transiently expressed in tobacco [10], can catalyze the multi-stage reaction of the precursor 13R-mannoyl oxide to produce various labdane diterpenoids [11].

Miltiradiene is the biosynthetic precursor of many active diterpene [12,13], therefore, it is very important to study the catalytic function of the P450s that start the first modification steps of miltiradiene. Miltiradiene undergoes hydroxylation at the carbon atoms C7, C11, and C12 to form a number of different chemicals with the catalytic action of CYP76AHs. In this study, CYP76AHs homologous genes were isolated and subjected to tertiary structure comparison, bioinformatics study, and evolutionary analysis. 15 CYP76AH genes from 10 different species of *Salvia* were cloned, and miltiradiene was used as the substrate for the enzymatic reaction, and the catalytic efficiency of producing different C11, C7 and C12 oxidation products is analyzed. All of the enzyme genes had distinct effects and produced different products, key enzyme genes with higher catalytic efficiency than SmCYP76AH1 and SmCYP76AH3 were discovered [14].
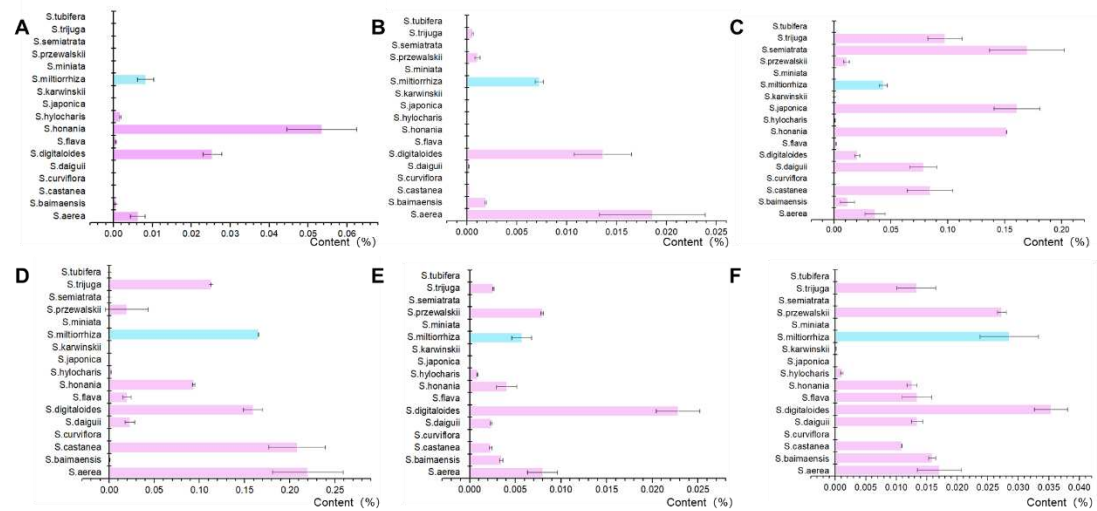

Figure 1. Biosynthetic pathway of tanshinones in S. miltiorrhiza mediated by CYP76AHs.

2. Results and Discussion

2.1. Chemical Constituents of Salvia

The biosynthetic pathway of tanshinones has been characterized in Danshen. Several components, including ferruginol, sugiol, miltirone, cryptotanshinone, and 2-isopropyl-8-methylphenanthrene-3,4-dione (Ro-09-0680, abbreviated as R09 in this article) and dihydrotanshinone I, were shown to be involved in the biosynthesis of tanshinone. According to previous research, the content of intermediates and end products in the tanshinone metabolism

pathway varies significantly between species (Figure 2). 16 distinct species from the *Salvia* genus to represent high, medium, and low levels of tanshinone based on the findings of previous studies were selected for further analysis. Among the *Salvia* species studied, *S. honania* has the greatest ferruginol concentration at 0.05352%, followed by *S. daiguii* at 0.02534%, both of which are much greater than the ferruginol content in Danshen. *S. aerea* has the highest content of sugiol at 0.0186%, followed by S. digitaloides at 0.01362%. The proportion of miltirone content in *Salvia* plants is significantly higher than that in ferruginol and sugiol and overall shows a high level. The tanshinone content of a total of 6 species was higher than that of *S. aerea*, *S. castanea*, *S. digitaloides*, *S. miltiorrhiza*. The content of cryptotanshinone in *S. milliorrhiza* is relatively high. The R09 content of various species in the *Salvia* genus is generally lower than that of miltirone and cryptotanshinone, and *S. digitaloides* has the highest R09 content. The content of dihydrotanshinone I in *S. digitaloides* is the highest. The results of this study suggest that there may be more alternative varieties of *S. miltiorrhiza* in *Salvia*. At the same time, it provides a material basis for the excavation of diterpenoid biosynthetic elements and the study of catalytic mechanisms in *Salvia*.
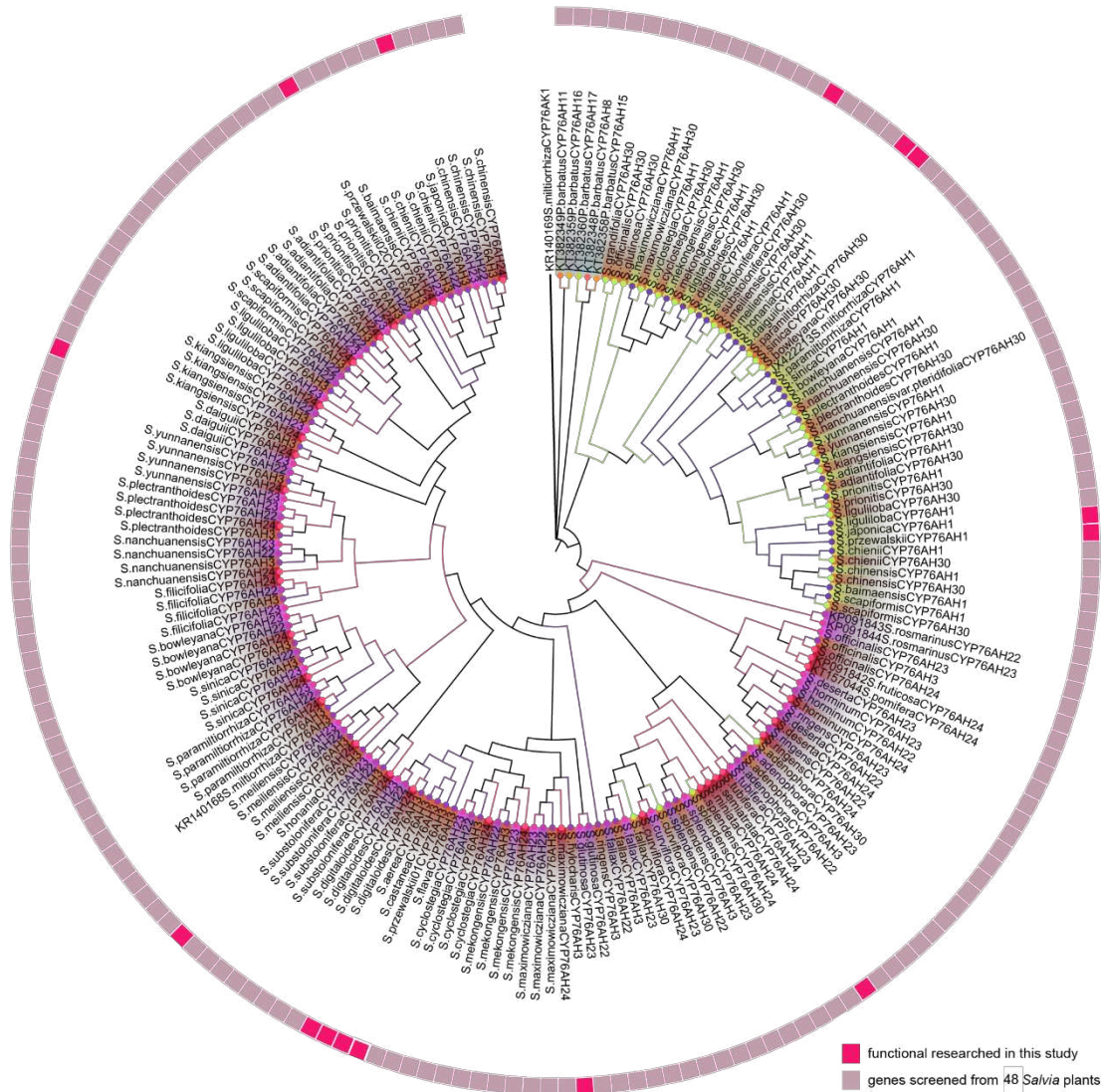

Figure 2. Content of key compounds in the biosynthesis pathway of tanshinone in Salvia plants (A: ferruginol; B: sugiol; C: miltirone; D: cryptotanshinone; E: R09; F: dihydrotanshinone I).

2.2. Phylogeny Analysis of Candidate CYP76AH Subfamily Genes

CYP76AHs were reported as the first CYP450 involved in biosynthesis of abidane-type diterpenoid, which catalyzed the carbon skeleton miltiradiene to produce ferruginol or further oxidation to sugiol [7]. Evolution of CYP76AH played essential role in diterpenoids biosynthesis in *Salvia*. In order to analyze the phylogeny relationship of CYP76AHs from different *Salvia* species. CYP76AH subfamily genes were screened from these 48 *Salvia* plants root transcriptome obtained previously and SmCYP76AH1 was used as a template for homology screening. 130 candidate CYP76AH genes with more than 55% homology and longer than 1400 bp expression frame were obtained. A phylogenetic tree was constructed to preliminarily investigate the potential catalytic function of candidate genes. Eleven functional screened CYP76AH subfamily genes were downloaded and CYP76AK1 was used as the root of the phylogenetic tree (Figure 3).

The discovered genes were named CYP76AH1, CYP76AH3, CYP76AH22-24, and CYP76AH30 with their species abbreviation prefixes respectively based on gene sequence similarity. According to the clustering results, CYP76AH1 and CYP76AH30 genes are clustered into one branch, genes of CYP76AH3 and CYP76AH22-24 clustered into a large branch, and the number of genes was significantly higher than other types of genes (Figure 3). The CYP76AH genes from *Salvia* plants is closely related to CYP76AH1, CYP76AH3, CYP76AH24, CYP76AH22, and CYP76AH23, but far from CYP76AH11, CYP76AH8, CYP76AH11, CYP76AH15, CYP76AH16, and CYP76AH17.

Figure 3. Figure 3. The phylogeny tree of *Salvia* CYP76AH subfamily genes . (GenBank accessions: SmCYP76H1, JX422213; SmCYP76AH3, KR140168; RoCYP76AH4; RoCYP76AH22, KP091843; RoCYP76AH23, KP091844; SfCYP76AH24, KP091842; SpCYP76AH24, KT157044S; PbCYP76AH8, KT382348; PbCYP76AH11, KT382349; PbCYP76AH15, KT382358; PbCYP76AH16, KT382359 and PbCYP76AH17, KT382360).
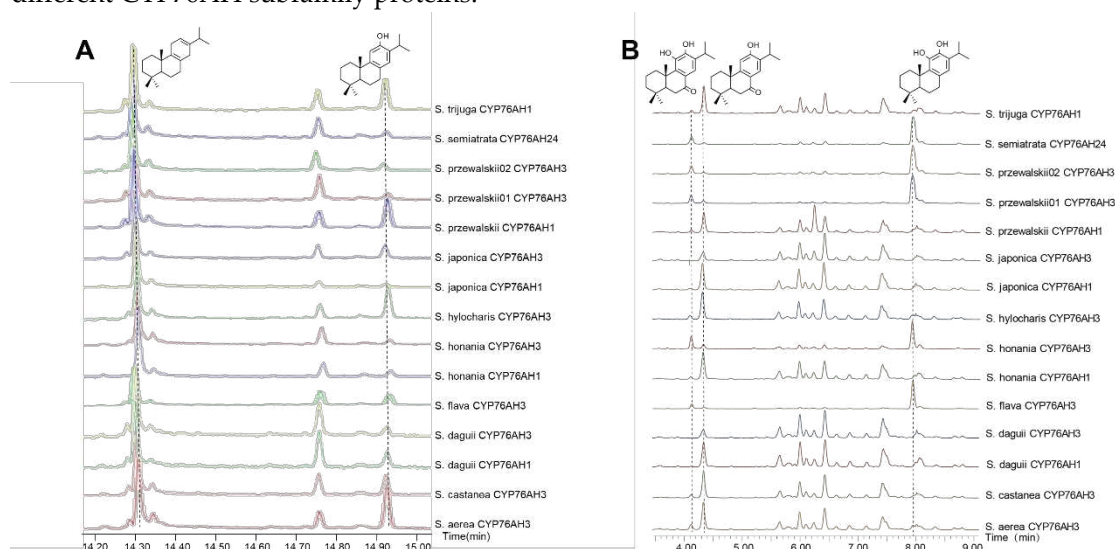
Next, the genetic distances between the species were calculated. The number of bootstrap replications was set to 50, while the gamma parameter was set to 1.00. The overall distance of *Salvia* is 0.20. The best maximum likelihood fits of 24 different nucleotide substitution models were selected. The BIC score was 35910.78 for TVM + F+I + G4. Next, the Xia test was performed using DAMBE 5.3.8, Iss< Iss.c with Prob (Two-tailed) being 0. CYP76AH were divided into 5 clades. Clade I includes CYP76AH1 from *Salvia*. Clade II includes CYP76AH22 and CYP76AH23 from *Salvia*. Clade III includes CYP76AH24 from *Salvia*. Clade IV includes CYP76AH24 from *Salvia*. Clade V was the root of the phylogenetic tree (Figure 3)

2.3. Biochemical Characterization of CYP76AHs

Among the CYP76AH homologous, ten *Salvia* plants with high abundance of tanshinones in the study of metabolic components were selected for CYP76AH gene cloning and functional characterization (Figure 2). Fifteen CYP76AH gene were cloned as candidate genes for functional research. To investigate the biochemical activity of these CYP76AHs, recombinant expression in yeast (*Saccharomyces cerevisiae*) was employed [7,8] . Full length cDNA of all CYP76Hs were cloned into the

yeast expression vector pESC-His, and the resulting constructs transformed into the WAT11 yeast strain in which the endogenous NADPH-CYP reductase has been replaced by one from *Arabidopsis thaliana* [15]. In vitro assays were then carried out with microsomal preparations from induced cultures of these recombinant yeast, using miltiradiene as substrates.
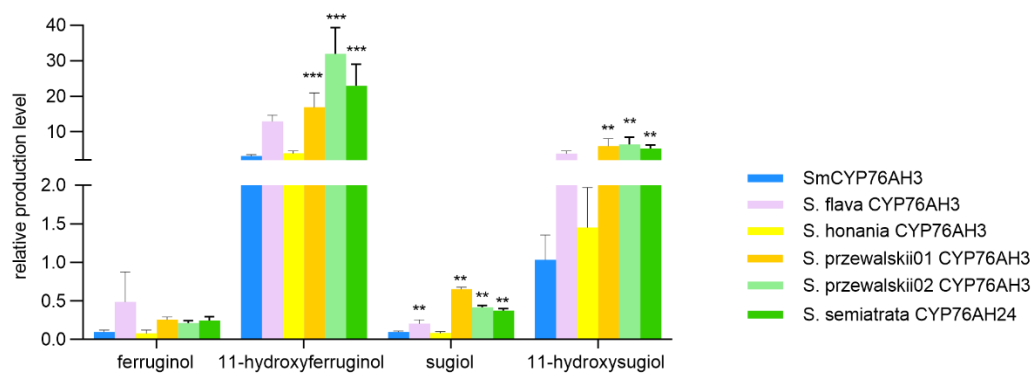
GC-MS and UPLC-Qtof-MS were employed for the detection of enzyme reaction products. All 15 CYP76AH proteins from different *Salvia* plants have the ability to catalyze the formation of ferruginol from miltiradiene, but the conversion efficiency of the products varies (Figure 4A). Moreover, due to the heterogeneric catalytic function of CYP76AH subfamily proteins [16,17], ferruginol was further oxidated at other carbon sites. Therefore, this study simultaneously detected other C7 and C11 oxidation products through UPLC-Qtof-MS (Figure 4B). Most CYP76AH proteins can catalyze the continuous oxidation of ferruginol to produce sugiol, 11-hydroxyferruginol, and 11-hydroxy sugiol. However, there are significant differences in the amount of each product catalyzed by different CYP76AH subfamily proteins.

Figure 4. Chromatograms of reaction products of CYP76AH protein with miltiradiene. A: GC-MS chromatogram of miltiradiene and ferruginol; B: UPLC-Qtof-MS chromatogram of 11-hydroxyferruginol, sugiol and 11-hydroxysugiol.

The main products generated by catalytic reaction of SaeCYP76AH3, ScaCYP76AH3, SdaCYP76AH1, SdaCYP76AH3, ShoCYP76AH1, ShyCYP76AH3, SjaCYP76AH1, SjaCYP76AH3, SprCYP76AH1 and StrCYP76AH1 with miltiradiene are ferruginol and sugiol, while the production of 11-hydroxyferruginol and 11- hydroxysugiol is relatively low or even undetectable. It indicates that the major catalytic sites of these CYP76AH subfamily proteins are C12 and C7, and the hydroxylation ability for C11 is weak. Among them, the amounts of each product of SdaCYP76AH3 and SjaCYP76AH3 are relatively low. In addition, the main products of SflCYP76AH3, ShoCYP76AH3, SprCYP76AH3-01, SprCYP76AH3-02, and SseCYP76AH24 are 11-hydroxyferruginol and 11-hydroxysugiol which can directly convert a large amount of catalyzed ferruginol and sugiol into their C11 hydroxylated products respectively, indicating that these CYP76AH subfamily proteins may have higher C11 hydroxylation activity than others.

For further study on the catalytic efficiency of each CYP76AHs, the same amount of each CYP76AH microsomes and miltiradiene were added in the enzyme reaction, and SmCYP76AH1 and SmCYP76AH3 proteins from *Salvia miltiorrhiza* were used as controls. Relative quantitative analysis was taken using the amount of each product generated by SaeCYP76AH3 as 1, and perform normalization analysis on the amount of each CYP76AH protein catalytic product (Figure 5).
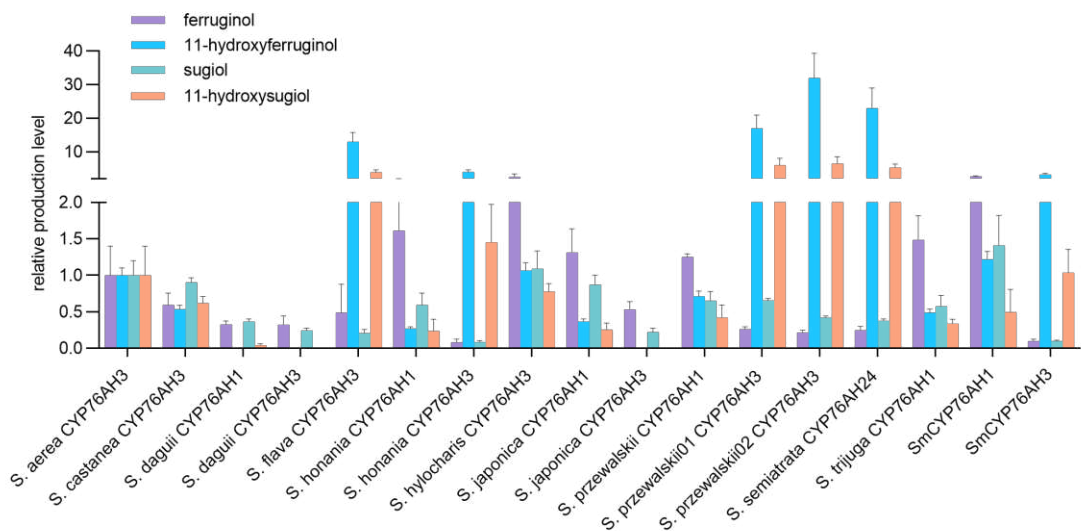
Figure 5. Relative quantitative analysis of the products from the reaction of miltiradiene catalyzed by CYP76AH microsomes.

The results of relative quantitative analysis showed that the main products of SflCYP76AH3, ShoCYP76AH3, SprCYP76AH3-01, SprCYP76AH3-02, and SseCYP76AH24 were C11 hydroxylated products, with a product ratio consistent with SmCYP76AH3. Further analysis of these protein and SmCYP76AH3 catalytic products revealed that the product yields of three proteins, SprCYP76AH3-01, SprCYP76AH3-02, and SseCYP76AH24, were significantly higher compared to SmCYP76AH3, especially the C11 hydroxylation products (Figure 6). The production of 11-hydroxy ferruginol and 11-hydroxysugiol in the catalytic reaction of SprCYP76AH3-02 are 10.67 and 6.31 times higher than that of SmCYP76AH3.

Table 1. The squence similarity of the Salvia CYP76AHs and identified CYP76AHs.

	Amino Acid Similarity%				Sequence Differences	
	CYP76 AH1	CYP76 AH3	SpCYP76 AH22	SpCYP76 AH24	CYP76AH1	CYP76AH3
SaeCYP76 AH3	99.60	79.60	78.26	77.73	S3Y, K329R	
ScaCYP76 AH3	95.91	77.00	75.76	75.24		S3Y, KS494NPRIRNTTHYRARASTWNRS
SdaCYP76 AH1	98.99	78.99	77.67	77.14		
SdaCYP76 AH3	80.00	99.39	86.93	86.85		S5P, Q161E, K320Q
SflCYP76 AH3	79.39	96.76	87.33	86.85		S5P, F12L, S14T, S21F, S22F, R157K, Q161K, A258T, D261N, N277Q, N279S, V315I, A322G, A394G, D466N, R488K
ShoCYP76 AH1	99.39	79.60	78.26	77.53	F4S, T18I, K384M	
ShoCYP76 AH3	79.80	99.39	86.73	86.65		S3N, S5P, K320E
ShyCYP76 AH3	99.80	79.80	78.46	77.93	S3I	
SjaCYP76 AH1	98.99	79.60	77.43	76.49	D2E, F4S, I13T, T18I, K156R	
SjaCYP76 AH3	80.00	98.58	86.93	86.65		S5P, H156D, R157K, D319H, K320Q, R345S, Q473G
SprCYP76 AH1	99.19	79.39	78.06	77.53	D2E, F4S, N279S	

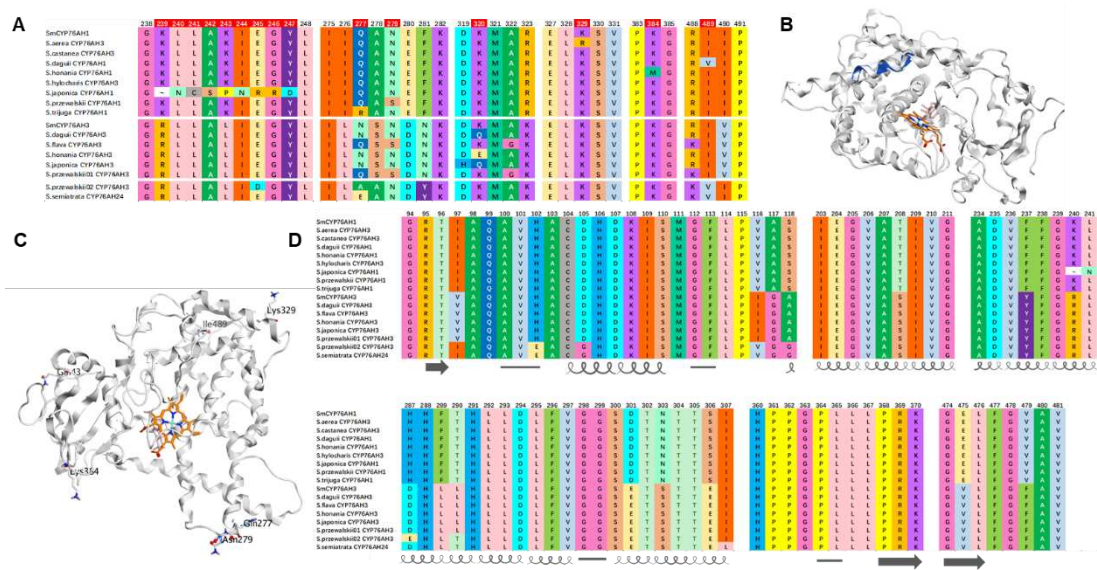|					
SprCYP76 AH3-02	80.30	96.76	87.52	87.65	S3Y, S5P, S14T, S21F, S22F, R157K, Q161K, D261N, N277Q, N279S, A322G, D406E, A407S, Q409E, D466N, R488K
SprCYP76 AH3-02	76.67	85.49	83.88	83.23	
SseCYP76 AH24	79.39	79.60	86.93	86.25	
StriCYP76 AH1	99.39	79.60	78.06	77.53	F3S, Q43, Q277R

Figure 6. Comparison of the product quantity of miltiradiene catalyzed by CYP76AH microsomes from six different sources.

2.4. Correlation Analysis of CYP76AH Protein Structure and Activity

In order to investigate the significant differences in the production of miltiradiene catalyzed by CYP76AH subfamily proteins, the amino acid sequences of each CYP76AH protein were compared. The sequence similarity of each protein with SmCYP76AH1, SmCYP76AH3, and SpCYP76AH22, SpCYP76AH24 were analyzed. Multiple CYP76AH sequences have a similarity of over 95% with the SmCYP76AH subfamily protein sequence (Table 1). Further analysis was conducted on the differential amino acid sites between these CYP76AH proteins and the SmCYP76AHs, these CYP76AH gene has extremely high similarity and can be called a natural multi mutant gene of SmCYP76AH1 and SmCYP76AH3.

There is only one amino acid difference between ShyCYP76AH3 and SmCYP76AH1 in the N-terminal transmembrane domain, which has little effect on catalytic function, so there is no significant difference in its catalytic products (Table 1, Figure 5). Amino acid sequence alignment analysis was conducted on each CYP76AH protein (Figure 7). Except for the N-terminal transmembrane domain, SjiaCYP76AH1 only had multiple consecutive amino acid sequences at 239–247 that differed from SmCYP76AH1 and were completely different from other CYP76AH proteins (Figure 7A). However, SjiaCYP76AH1 can still catalyze the production of four products of miltiradiene, but the catalytic efficiency is significantly reduced compared to SmCYP76AH1.

Figure 7. Amino acid and protein structure differences between the SmCYP76AHs and CYP76AHs in *Salvia* species.

Homologous modeling of CYP76AH protein was performed using the Swissmodel. Based on SmCYP76AH1 protein crystal structure, the feasibility of homologous modeling of CYP76AH subfamily proteins is high. From the 3D structure of the protein, SjiaCYP76AH1 has multiple consecutive amino acid sequences at 239–247 that differ from SmCYP76AH1, and do not cause any changes in the active pocket structure of the protein. However, at the outer helical structure of the protein, changes in this continuous amino acid sequence will cause twisting and folding in the middle of the helical structure (Figure 7B). From the analysis of amino acid properties, the change in the sequence at this location has little effect on the hydrophilicity of the protein, but it is possible that the protein substrate channel has changed, affecting the ability of the protein to grasp the substrate miltiradiene, thereby reducing its catalytic efficiency compared to SmCYP76AH1.

Compared with SmCYP76AH1, StrCYP76AH1 lacks 43Q and has a Q277R difference, SprCYP76AH1 has a N279S difference, while ShoCYP76AH1has a I489V difference (Figure 7C). It is possible that changes in amino acids in the external loose region affect the solubility of the protein, thereby affecting its catalytic efficiency and significantly reducing its catalytic product yield. Similarly, compared to all CYP76AH proteins in this study, SaeCYP76AH3 has a special the 329R differs from 329K and the ShoCYP76AH1 has a special 384M while others are 384K. These 'natural mutations' occurred in the loose structural area of the outer layer of the protein, which may have affected its solubility and weakened its catalytic efficiency.

The C11 hydroxylation catalytic efficiency of SprCYP76AH3-01, SprCYP76AH3-02, and SseCYP76AH24 was significantly higher compared to SmCYP76AH3. The amino acid sequences of the substrate binding sites in their protein activity pockets were compared, these three proteins with significantly higher catalytic efficiency have certain differences in amino acids in their active pockets compared to other CYP76AH subfamily proteins (Figure 7D), the amino acid sequences of the 102 E, 105 G, 118 G, and 307 L are different from those of other CYP76AH proteins. In addition, multiple amino acid sites have both CYP76AH1 and CYP76AH3 protein characteristics. For example, amino acids such as 97I, 116V, and 290T are same as that of CYP76AH1 group, while amino acids such as 117G, 208S, 237Y, and 479F have typical CYP76AH3 group protein sequence characteristics. The high catalytic efficiency of these CYP76AHs may be related to the advantageous amino acids that possess both CYP76AH1 and CYP76AH3 types of amino acids in the active pocket.

3. Discussion

Plant-derived diterpenoids are a class of compounds with diverse structures and functions. These compounds are widely used in pharmaceuticals, cosmetics and food additives industries because of their pharmacological properties such as anticancer, anti-inflammatory and antibacterial activities. In *Salvia* genus, diterpenoid components are widely distributed and diverse, therefore the proteins involved in the diterpenoid biosynthesis pathway are abundant and functionally diverse. In recent years, with the gradual discovery of functional genes in the biosynthetic pathway of plant-derived diterpenoids and the development of structural biology, more and more protein structures are being analyzed and designed rationally or semi rationally to obtain functional modified mutants. However, the natural environment, the natural driving force behind mutation and evolution, may be a natural treasure trove for us to explore more efficient catalytic components.

Salvia plants are mostly wild varieties, and for most species, obtaining materials sufficient for systematic chemical analysis and structural identification poses great difficulties. The complex wild environment provides mutation evolution pressure for plant gene evolution. Despite difficulties in obtaining materials, we still obtained a variety of wild *Salvia* plants and cloned a series of active P450s that can catalyze the production of various products from miltiradiene. Moreover, three high-efficiency catalytic CYP76AH proteins were obtained, with significantly higher catalytic efficiency than the same functional proteins in widely cultivated *Salvia miltiorrhiza*.

The modification of CYP76AHs in *Salvia miltiorrhiza* has been studied [18]. A series of mutant proteins with significantly higher catalytic efficiency designed and obtained than wild-type SmCYP76AH1 and SmCYP76AH3 through homologous modeling and molecular docking analysis. Reasonably designed CYP76AH mutant protein microsomes catalyze the production of multiple products of miltiradiene, resulting in a significant increase in yield in engineering strains. Compared with the designed CYP76AHs mutants which mutation sites are concentrated in the protein activity pocket, and the amino acid closest to the substrate and heme molecules in the substrate binding site, the amino acid differences between SmCYP76AHs and CYP76AHs in *Salvia* in this study mainly exist in areas with relatively loose structures such as protein outer walls or substrate channels. This may be due to the influence of differential amino acids on the solubility of proteins and the affinity of substrates entering the pocket. These differences can provide more reference for the rational design of such protein mutants.

4. Materials and Methods

4.1. Plant Material and Chemicals

Salvia species were collected in September 2021. Both leaves and roots of these species were obtained from Shanghai Chenshan Botanical Garden (Shanghai, China). The *Salvia* plant tissue was kept on dry ice. Then, the *Salvia* plants were snap frozen in liquid nitrogen and stored at −80 °C prior to total RNA extraction. Miltiradiene and ferruginol standards were purchased from Solarbio Bio-Technology (Beijing, China).

4.2. Chemical Constituents of Salvia

Standard amounts of 1.15 mg of ferruginol, 1.20 mg of sugiol, 1.13 mg of miltirone, 1.18 mg of cryptotanshinone, and 1.11 mg of dihydrotanshinone I were accurately weighed and placed in 2 mL centrifuge tubes. 200 μL of chromatographic grade dimethyl sulfoxide was added for dissolution and mixed well after 45 min of ultrasound. The concentration of the prepared solution was 20 mmol/L. The dried *Salvia* (4 replicates for each species) sample roots were crushed to 80 mesh. A total of 0.02 g of plant root powder was accurately weighed, extracted with 1 mL of 70% methanol ultrasound for 45 min, and centrifuged (12000 rpm min^{-1}) for 15 min. The supernatant was pipetted, passed through a 0.22 μm filter membrane, immediately transferred to a liquid phase bottle, and stored at 4 °C. The chemical constituents of *Salvia* were analyzed by HPLC.

4.3. Identification of CYP76AH Genes

Fifty-five percent were verified, and CYP76AH genes were collected from transcriptome data. CYP76AH sequences were aligned with CYP76AH1 of *S. miltiorrhiza* (BioEdit Sequence Alignment Editor 7.2.1), with 1000 bootstrap replicates. Due to sequencing problems, the gene for which full-length sequences could not be obtained was deleted.

4.4. Gene Diversity of the CYP76AH Genes

A phylogenetic tree was constructed using maximum likelihood (ML) method with 1000 bootstrap replications (Mega X). Base substitution saturation tests were performed (DAMBE5.3.8). The CYP76AH gene of *Salvia* has high variability. There are 609 mutation sites and 78 haplotypes. The haplotype diversity was approximately 0.991, and the gene diversity was 0.10771. (DNASP 5.10).

4.5. Total RNA Extraction

Ten *Salvia* species were chosen based on metabolomic data. Total RNA was extracted from the bulb of *Salvia* with a Quick RNA Isolation Kit (Huayueyang Biotechnology, Beijing, China). One microliter RNA integrity was detected immediately by 1% agarose gel electrophoresis with 1 μL DNA/RNA loading buffer. RNA purity and concentration were gauged by a NanoDrop 2000 (Thermo Fisher Scientific, USA). Each species of *Salvia* cDNA was reverse transcribed using Transcriptase (Jinsha Biotechnology, Beijing, China).

4.6. CYP76AH Gene Clone

The full-length cDNA of *Salvia* was amplified using Phusion® High-Fidelity DNA Polymerase (NEB, USA) with the primers listed in Table S1. The reaction conditions were 45 s at 94 °C for denaturation, then 15 s at 94 °C, 15 s at 62 °C, 2 min at 72 °C for 38 cycles, and finally, 10 min at 72 °C extension. The cloned CYP76AH DNA was purified and collected at −20 °C.

4.7. Heterologous Expression of CYP76AH in Yeast and In Vitro Activity Assays

The recombinant plasmid pESC-His-CYP76AH was transformed into yeast strain WAT11 using a Frozen EZ Yeast Transformation II Kit™ (ZYMO, USA) for heterologous expression, and the WAT11 strain was transformed with the pESC-His empty vector control. Both were cultured in SD-His medium at 30 °C for 48 h. A single colony was selected and cultured for 48 h, with an OD600 value greater than 2. Then, an equal volume of YPL (1% yeast extract, 2% peptone) was added for induction for 14 h, and the OD600 value reached 2–3. Combined with previous experimental experience, microsomes were chosen as the method for enzymatic reactions. Determination of in vitro activity was performed in a 500 μL reaction system, including 500 μM NADPH, 0.5 mg microsomal protein, 50 μM substrate, and regenerating system including 5 mM glucose-6-phosphate (with 1 unit glucose-6-phosphate dehydrogenase), 5 μM FAD, 5 μM FMN, and 1 μM DTT. The reaction was incubated at 30 °C and 200 rpm for 2.5 h. Then, 1000 μL ethyl acetate was used to extract the product (2 times). The first ethyl acetate extract (70 μL from 500 μL) was dispensed into gas phase vials. The product was resolubilized with 70 μL of methanol.

4.8. Ferruginol Analysis by GC–MS

Trace 1310 series GC & TSQ8000 MS were used for detection of substrates and products (Thermo Fisher Scientific Co. Ltd., Waltham, MA, USA). Chromatographic separation was performed on a TR-5ms capillary column (30 m9 0.25 mm ID DF = 0.25 lm (film thickness dimension); Thermo Fisher Scientific). The flow rate control was 1 mL min $^{-1}$.

4.9. 11-Hydroxyferruginol, Sugiol, and 11-Hydroxy Sugiol Analysis by UPLC-Qtof-MS

11-Hydroxyferruginol, sugiol, and 11-hydroxy sugiol was detected in vitro by UPLC-Qtof-MS (Waters Technologies, Milford, MA, USA). This BEH column (2.1×50 mm, 1.8 μm particle size; Waters

Technologies) was used for chromatographic separation. The mobile phase was acetonitrile and water (1% formic acid), with a UV absorption wavelength of 254 nm. Carnosic acid was used as an internal standard at 10 μM.

5. Conclusions

This study analyzed the functions of 11 CYP76AH genes derived from *Salvia* plants, all of which catalyze the formation of ferruginol from miltiradiene, and some can continue to catalyze the formation of three oxidation products at the C7 and C11 positions of ferruginol. By comparing catalytic efficiency, three active catalytic elements with significantly improved catalytic efficiency for C11 hydroxylation were identified, including SprCYP76AH3-01, SprCYP76AH3-02, and SseCYP76AH24. By amino acid sequence alignment and molecular docking analysis of the CYP76AH protein in these *Salvia* plants, potential key amino acid residues such as 117G, 208S, 237Y, and 479F were identified. However, their impact on activity still needs further experimental verification.

Salvia plants containing tanshinones are a natural gene element library for mining functional genes in the biosynthesis pathway of tanshinones. Comparing to the CYP76AH subfamily genes with verified functions in *Salvia miltiorrhiza*, mining proteins with similar catalytic functions, but higher expression level in yeast and higher catalytic efficiency can provide efficient and advantageous natural mutant catalytic elements for the synthetic biology production of tanshinones.

Author Contributions: Z.Z., X.L., J.W. and Q.L. conducted the experiments and completed the statistical and data analysis. D.Y. collected the plant materials. Z.Z., Y.M., D.Y. P.S. and wrote the manuscript. L.H. and J.G. were involved in planning the experiments. L.H., D.Y. and J.G. obtained the funding support. All authors have read and agreed to the published version of the manuscript.

References

1. Wu, Y.B.; Ni Z.Y. Shi Q.W.; Dong M.; Kiyota H.; Gu Y.C.. Constituents from Salvia species and their biological activities. Chem Rev. 2012, 11, 5967-6026.
2. Johnson, S.R.; Bhat, W.W.; Bibik, J.; Turmo, A.; Hamberger, B. Genomics Consortium EM. A database-driven approach identifies additional diterpene synthase activities in the mint family (Lamiaceae). Journal of Biological Chemistry. 2019, 294(4):1349-1362.
3. Subedi, L.; Gaire, B.; Tanshinone IIA: A phytochemical as a promising drug candidate for neurodegenerative diseases. Pharmacological research. 2021, 169, 105661.
4. Fang, Z.; Zhang, M.; Liu, J.N.; Zhao, X.; Fang, L.. Tanshinone IIA: A Review of its Anticancer Effects. Frontiers in Pharmacology. 2021, 11, 611087.
5. Meunier, B.; Visser, S.; Shaik, S.. Mechanism of Oxidation Reactions Catalyzed by Cytochrome P450 Enzyme. ChemInform. 2004, 9, 3947-3980.
6. Bathe, U.; Frolov, A.; Porzel, A.; Tissier, A. CYP76 Oxidation Network of Abietane Diterpenes in Lamiaceae Reconstituted in Yeast. J Agric Food Chem. 2019, 49, 13437-13450.
7. Guo, J.; Zhou, Y.J.; Hillwig, M.L.; Shen, Y.; Yang, L.; Wang, Y.. CYP76AH1 catalyzes turnover of miltiradiene in tanshinones biosynthesis and enables heterologous production of ferruginol in yeasts. Proc Natl Acad Sci U S A. 2013, 29, 12108-12113.
8. Guo, J.; Ma, X.; Cai, Y.; Ma, Y.; Zhan, Z.; Zhou, Y.J.. Cytochrome P450 promiscuity leads to a bifurcating biosynthetic pathway for tanshinones. New Phytol. 2016, 2, 525-34.
9. Scheler, U.; Brandt, W.; Porzel, A.; Rothe, K.; Manzano, D.; Bozic, D.. Elucidation of the biosynthesis of carnosic acid and its reconstitution in yeast. Nat Commun. 2016, 7, 12942.
10. Forman, V.; Bjerg-Jensen, N.; Dyekjr, JD.; Mller, BL.; Pateraki, I. Engineering of CYP76AH15 can improve activity and specificity toward forskolin biosynthesis in yeast. Microbial Cell Factories. 2018, 1, 181.

11. Janicsák, G.; Zupkó, I.; Nikolovac, MT.; Forgo, P.; Vasas, A.; Mathé, I.; Blunden, G.; Hohmann, J.. Bioactivity-guided study of antiproliferative activities of Salvia extracts. Nat Prod Commun. 2011, 5, 575-579.

12. Xiao, S.; Liu, C.; Chen, M.; Zou, J.; Zhang, Z.; Cui, X.; Jiang, S.; Shang, E.; Qian, D.; Duan, J.. Scutellariae radix and coptidis rhizoma ameliorate glycolipid metabolism of type 2 diabetic rats by modulating gut microbiota and its metabolites. Appl Microbiol Biotechnol. 2020, 1, 303-317.

13. Gohari, A.R.; Saeidnia, S.; Malmir, M.; Hadjiakhoondi, A.; Ajani, Y.. Flavones and rosmarinic acid from Salvia limbata. Nat Prod Res. 2010, 20, 1902-1906.

14. Lin, T.C.; Fang, J.M.; Cheng, Y.S.. Terpenes and lignans from leaves of Chamaecyparis formosensis. Phytochemistry 1999, 6, 793-801.

15. Urban, P.; Mignotte, C.; Kazmaier, M.; Delorme, F.; Pompon, D. Cloning, yeast expression, and characterization of the coupling of two distantly related Arabidopsis thaliana NADPH-cytochrome P450 reductases with P450 CYP73A5. Journal of Biological Chemistry 1997. 272, 19176-19186.

16. Chen, Yue.; Wang, Yanting.; Guo, Juan.; Yang, Jian.; Zhang, Xiaodan.; Wang, Zixuan.; Cheng, Ying.; Du, Zewei.; Qi, ZheChen.; Huang, YanBo.; Dennis, Man.; Wei, Yukun.; Yang, Dongfeng.; Huang, Luqi.; Liang, Zongsuo.. Integrated Transcriptomics and Proteomics to Reveal Regulation Mechanism and Evolution of SmWRKY61 on Tanshinone Biosynthesis in Salvia miltiorrhiza and Salvia castanea. Frontiers in plant science 2022, 12, 820582.

17. Fang, Y.M.; Hou, Z.N..; Zhang, X.D.; Yang, D.F.; Liang, Z.S.. Diverse specialized metabolism and their responses to lactalbumin hydrolysate in hairy root cultures of Salvia miltiorrhiza Bunge and Salvia castanea Diels f. tomentosa Stib. Biochemical Engineering Journal 2018, 131, 58-69

18. Mao, Y. P.; Ma, Y.; Chen, T.; Ma, X.H.; Xu Y.Q.; Bo, J.L.; Li, Q.S.; Jin, B.L.; Wang, Y.N.; Li, Y.; Cui, G.H.; Zhao, Y.G.; Tang, G.F.; Shen, Y.; Lai, C.G.S.; Zeng, W.; Chen, M.; Guo, J.; Huang L.Q.. Functional integration of two CYP450 genes involved in biosynthesis of tanshinones for improved diterpenoid production by synthetic biology. ACS Synth. Biol. 2020, 9, 1763-1770.