



# Potential role of two cytochrome P450s obtained from *Lithospermum erythrorhizon* in catalyzing the oxidation of geranylhydroquinone during Shikonin biosynthesis



Wan Song<sup>a,b,c</sup>, Yibin Zhuang<sup>a,b</sup>, Tao Liu<sup>a,b,\*</sup>

<sup>a</sup> Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, 300308, China

<sup>b</sup> Key Laboratory of Systems Microbial Biotechnology, Chinese Academy of Sciences, Tianjin, 300308, China

<sup>c</sup> University of Chinese Academy of Sciences, Beijing, China

## ARTICLE INFO

### Keywords:

*Lithospermum erythrorhizon*  
Boraginaceae  
Shikonin  
Cytochrome P450 monooxygenase  
Geranylhydroquinone

## ABSTRACT

Shikonin is a natural naphthoquinone derivative that specifically occurs in boraginaceous plants, and the major active ingredient of the medicinal plant *Lithospermum erythrorhizon*. Previously, a cytochrome P450 oxygenase (CYP) CYP76B74 catalyzing 3''-hydroxylation of geranylhydroquinone (GHQ) — a key intermediate of shikonin biosynthesis, was identified from cultured cells of *Arnebia euchroma*. However, the enzymes catalyzing oxidation of the geranyl side-chain of GHQ from *L. erythrorhizon* remain unknown. In this study, we performed transcriptome analysis of different tissues (red roots and green leaves/stems) from *L. erythrorhizon* using RNA sequencing technology. Highly expressed CYP genes found in the roots were then heterologously expressed in *Saccharomyces cerevisiae* and functionally screened with GHQ as the substrate. As the result, two CYPs of CYP76B subfamily catalyzing the oxidation of GHQ were characterized. CYP76B100 catalyzed the hydroxylation of the geranyl side-chain of GHQ at the C-3'' position to form 3''-hydroxyl geranylhydroquinone (GHQ-3''-OH). The enzyme CYP76B101 carried out oxidation reaction of GHQ at the C-3'' position to produce a 3''-carboxylic acid derivative of GHQ (GHQ-3''-COOH) as well as GHQ-3''-OH. This enzyme-catalyzed oxidation reaction with GHQ as the substrate is reported for the first time. This study implicates CYP76B100 and CYP76B101 as having a potential role in shikonin biosynthesis in *L. erythrorhizon*.

## 1. Introduction

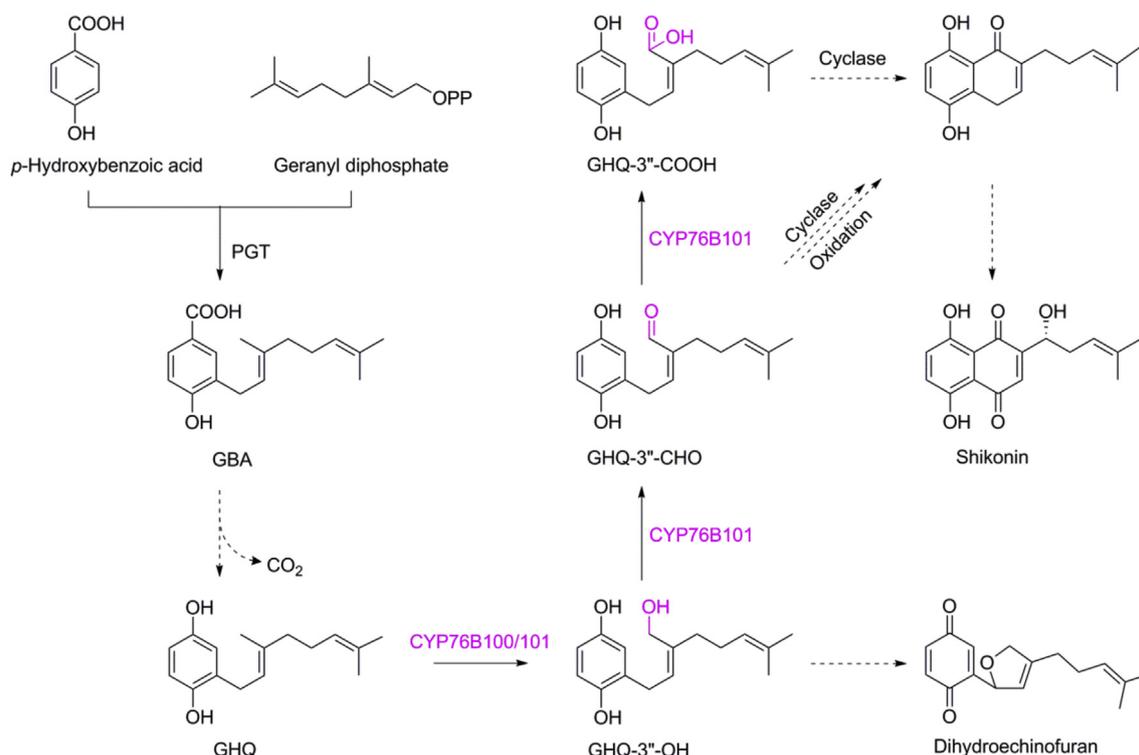
Shikonin and its derivatives are the natural naphthoquinone derivatives that have been considered as the primary active components of root tissues from *Lithospermum erythrorhizon*. Shikonins exhibit antimicrobial, wound-healing, anti-inflammatory, and anti-adenoviral properties (Andújar et al., 2013), and are currently used in food, cosmetics and pharmaceutical industry (Rai et al., 2018). Shikonins are only detected in boraginaceous plants belonging to the genera *Lithospermum*, *Arnebia*, *Anchusa*, *Echium*, *Onosma* and *Alkanna*, and have been used as natural dyes and herbal medicines in both Europe and the Orient for many centuries (Malik et al., 2016; Yazaki, 2017). The increasing demands for shikonins have made *L. erythrorhizon* at real risk of extinction in recent years (Yazaki, 2017). This heightened risk has inspired extensive efforts to focus on cell suspension cultures (Tabata and Fujita, 1985), biosynthetic pathways (Okamoto et al., 1995) and metabolic engineering production of 3-geranyl-4-hydroxybenzoic acid

(GBA) in yeast, an important intermediate in shikonin biosynthesis (Wang et al., 2017).

Shikonin is derived from *p*-hydroxybenzoic acid and geranyl diphosphate, originating from the shikimate pathway (Inouye et al., 1979) and mevalonate pathway (Li et al., 1998), respectively (Fig. 1). The coupling of two key precursors is catalyzed by geranyl diphosphate: 4-hydroxybenzoate 3-geranyltransferase (PGT) to yield the important intermediate GBA (Yazaki et al., 2002), which is subsequently converted to geranylhydroquinone (GHQ) by unknown enzymes. A cytochrome P450 monooxygenase (CYP) was identified from *L. erythrorhizon* suspension cultures catalyzing the C-3'' hydroxylation of GHQ to form a key intermediate 3''-hydroxy-geranylhydroquinone (GHQ-3''-OH) (Yamamoto et al., 2000). Recently, the first monooxygenase CYP76B74 of shikonin biosynthesis, which catalyzed the key 3''-hydroxylation step was cloned and identified from *Arnebia euchroma* (Wang et al., 2019). However, the enzymes involved in 3''-hydroxylation reaction of GHQ in *L. erythrorhizon* and the intriguing steps for

\* Corresponding author. Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, 300308, China.

E-mail address: [liu\\_t@tib.cas.cn](mailto:liu_t@tib.cas.cn) (T. Liu).



**Fig. 1.** The proposed biosynthetic pathway of shikonin in *L. erythrorhizon*. Single arrows represent one step reaction, while double arrows represent multiple step reactions. Dashed arrows signify undefined steps or the enzymes have not been verified yet. The steps with red-colored enzymes indicate the oxidation reactions of converting GHQ to GHQ-3'-OH or GHQ-3'-COOH by CYP76B100 and CYP76B101. PGT, geranyl diphosphate: 4-hydroxybenzoate 3-geranyltransferase; GBA, 3-geranyl-4-hydroxybenzoic acid; GHQ, geranylhydroquinone; GHQ-3'-OH, 3'-hydroxy-geranylhydroquinone; GHQ-3'-CHO, 3'-aldehyde-geranylhydroquinone; GHQ-3'-COOH, 3'-carboxyl-geranylhydroquinone.

naphthoquinone skeleton formation from GHQ-3'-OH remain unknown. A plausible mechanism for ring closure has been proposed via intermediate 3'-aldehyde-geranylhydroquinone (GHQ-3'-CHO), which could form the C-C bond with the aromatic nucleus by an electrophilic reaction (Yamamoto et al., 2000) (Fig. 1).

RNA sequencing technology (RNA-Seq) is a powerful tool for accelerating the discovery of novel genes through transcription pattern and functional analysis of genes thought to be involved in medicinal plant specialized metabolism. RNA-Seq thus plays important roles in elucidating pathways of natural products such as the seco-iridoid pathway in *Catharanthus roseus* (Krithika et al., 2015; Miettinen et al., 2014) and the tanshinones biosynthesis pathway in *Salvia miltiorrhiza* (Guo et al., 2012; Gao et al., 2014). The average lengths of the contigs generated by next-generation sequencing (NGS) based on RNA-Seq are < 500 bp, while long reads (average 4–8 kb) can be obtained by single-molecule real-time (SMRT) sequencing that is carried out in a third-generation sequencing platform (Xu et al., 2015). SMRT sequencing offers access to more complete transcriptome data, fully revealing the key advantage of obtaining full-length transcripts, and the associated problem of a higher error rate (up to 15%) can be solved by correction with NGS reads (Au et al., 2012). Thus, a hybrid sequencing approach combining NGS short reads and SMRT long reads will yield a more precise and extensive transcriptional database (Xu et al., 2015).

In plants, most of the reported terpene oxidations are carried out by CYPs to give rise to alcohols, aldehydes or carboxylic acids (Bathe and Tissier, 2019). In this work, we aimed to clone and functionally characterize CYP enzymes catalyzing the oxidation of GHQ at the C-3' position from *L. erythrorhizon*. The candidate genes were obtained by the transcriptome sequencing and differential gene expression analysis of green leaves/stems samples and red roots. Functional screening of the highly expressed CYP genes in yeasts combined with *in vitro* enzymatic activity assays resulted in identification of two CYP enzymes catalyzing

the oxidation of GHQ. CYP76B100 catalyzed the 3'-hydroxylation of GHQ to form GHQ-3'-OH. Interestingly, CYP76B101 catalyzed the oxidation of GHQ at the C-3' position to synthesize 3'-carboxyl-geranylhydroquinone (GHQ-3'-COOH) besides one-step oxidation product GHQ-3'-OH. The results might provide new insights into the biosynthesis of shikonin in *L. erythrorhizon* (Fig. 1).

## 2. Results

### 2.1. Discovery of CYP candidate genes

According to previous reports, shikonin is detected only in the red roots rather than the green leaves/stems of *L. erythrorhizon* and the genes related to shikonin biosynthesis are highly expressed in all red roots of three Lithospermeae plants, such as hydroxymethylglutaryl-CoA reductase gene (*HMGR*), phenylalanine ammonia lyase gene (*PAL*), *PGT* (Wu et al., 2017). Thus, the expression of the major genes that are related to shikonin downstream biosynthesis may be up-regulated in the roots. CYPs are frequently involved in the oxidation of terpenes, and GHQ hydroxylation at the C-3' position is catalyzed by CYPs (Yamamoto et al., 2000; Wang et al., 2019).

Combined the analysis of differentially expressed unigenes and the integrity of the open reading frames, a total of 40 CYP genes transcriptionally up-regulated in the red roots (Supplementary Table S1) were selected as the candidates for next step investigation. Cytochrome P450 reductase from *Arabidopsis thaliana* (*AtCPR1*), which is responsible for the electron transfer from NADPH to CYPs, was cloned into the yeast expression vector pCf302 under the control of the TEF1 promoter. Using the cDNA from red roots as the amplification template, the full-length sequences of 40 CYP candidates were obtained with the primers shown in Supplementary Table S2, and successfully cloned into pCf302-*AtCPR1* to yield CYP gene-carrying plasmids. These plasmids

were transformed into *Saccharomyces cerevisiae* BY4742 to yield engineered strains for further functional characterization using GHQ as the substrate.

## 2.2. Functional screening of the CYP candidates in recombinant yeast

To screen the encoded enzymes from the acquired candidate CYP genes to oxidize GHQ, we exploited the simplicity of heterologous expression of these cDNAs in engineered strains of *S. cerevisiae* harboring pCf302-AtCPR1-CYP. A control yeast strain expressed pCf302-AtCPR1 without carrying exogenous CYP. GHQ is not available commercially, and was synthesized according to the method reported in the literature (Baeza et al., 2012). The engineered strains were incubated with GHQ for 48 h, and then the ethyl acetate extracts of the yeast culture medium were analyzed by HPLC. As a result, two engineered strains harboring CYPs produced new products as compared with the strain harboring an empty plasmid. The two CYPs were respectively designated as CYP76B100 and CYP76B101 by the Cytochrome P450 Nomenclature Committee, and the gene sequence information of CYP76B100 and CYP76B101 from *L. erythrorhizon* were deposited into GenBank under accession numbers MN056183 and MN056184, respectively. A new peak (product 1) with a retention time (RT) of 21.1 min was observed on the chromatograph by analyzing metabolites of yeast harboring CYP76B100 (Fig. 2A). By co-expressing CYP76B101 and AtCPR1, two new products (2 and 3) with the corresponding retention time (RT) of 21.1 min and 21.6 min, were detected in the chromatogram (Fig. 3A). Meanwhile, the starting material GHQ completely disappeared, which meant total conversion into new products by CYP76B100 and CYP76B101 (Figs. 2A and 3A).

## 2.3. Products identification

To determine the molecular weight, the extracted metabolites were further subjected to LC-MS analyses. As shown in Fig. 2, product 1 exhibited the molecular ions of 263.1639 ( $[M + H]^+$ ) and 285.1460 ( $[M + Na]^+$ ), which corresponded to a hydroxylated GHQ derivative. The mass spectrum of product 2 that was formed by the strain carrying CYP76B101 was identical to that of product 1 (Fig. 3C). LC-MS analysis demonstrated that product 3 had the molecular ions of ( $[M +$

$H]^+ = 277.1453$ ,  $[M + Na]^+ = 299.1271$ ), indicating that product 2 was further oxidized by CYP76B101 into product 3 with a carboxyl group (Fig. 3D).

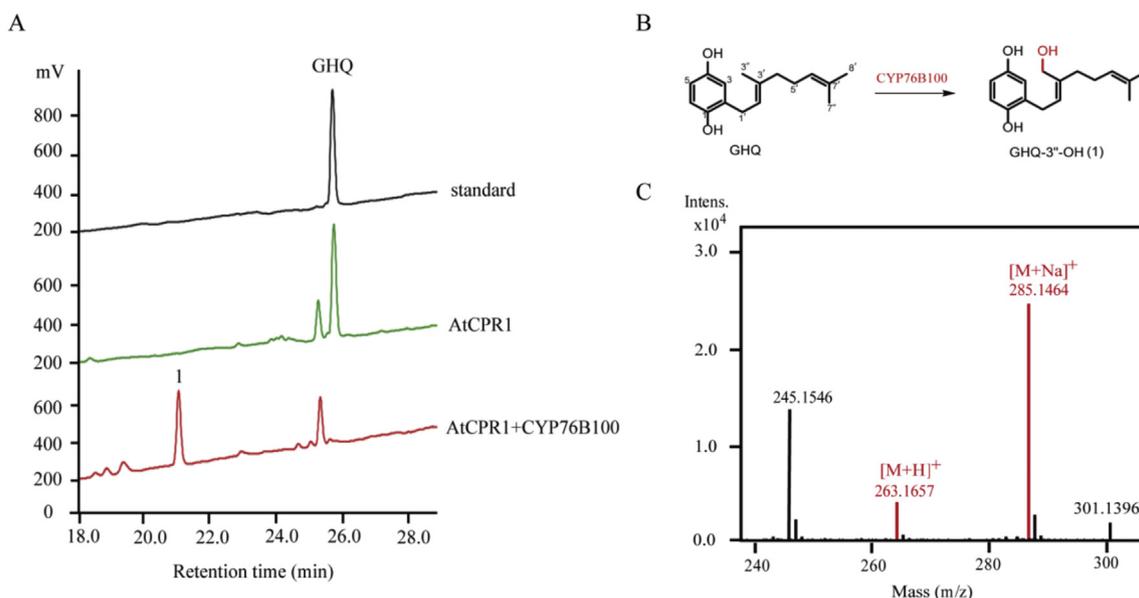
To isolate sufficient quantities of the oxidized products of CYP76B100 and CYP76B101 for chemical structure characterization, the *S. cerevisiae* BY4742 strains harboring pCf302-AtCPR1-CYP76B100 or pCf302-AtCPR1-CYP76B101 were cultured and incubated with GHQ for large-scale fermentation. The new products were purified and submitted for NMR analysis. Consequently,  $^1H$ -NMR and  $^{13}C$ -NMR spectra data of the purified products 1 and 2 (Table 1) were found to perfectly match with those of GHQ-3''-OH as previously reported (Wang et al., 2019).  $^1H$ -NMR and  $^{13}C$ -NMR spectra combined with the data of the two-dimensional experiment HMBC demonstrated that purified product 3 was GHQ-3''-COOH (Table 1 and Figs. S2, S3, S4). Full assignment of the  $^1H$  and  $^{13}C$  chemical shifts were achieved by evaluating the two-dimensional NMR spectra of product 3. Especially, the carbonyl carbon at  $\delta$  171.5 was assigned C-3'' on the basis of its  $^3J$ -HMBC correlation with the olefinic proton H-2' ( $\delta$  5.92) and the methylene protons H<sub>2</sub>-4' ( $\delta$  2.32). Thus, product 3 was identified as GHQ-3''-COOH, which was the three-step oxidation product of GHQ at the C-3'' position.

## 2.4. In vitro enzyme activity assays

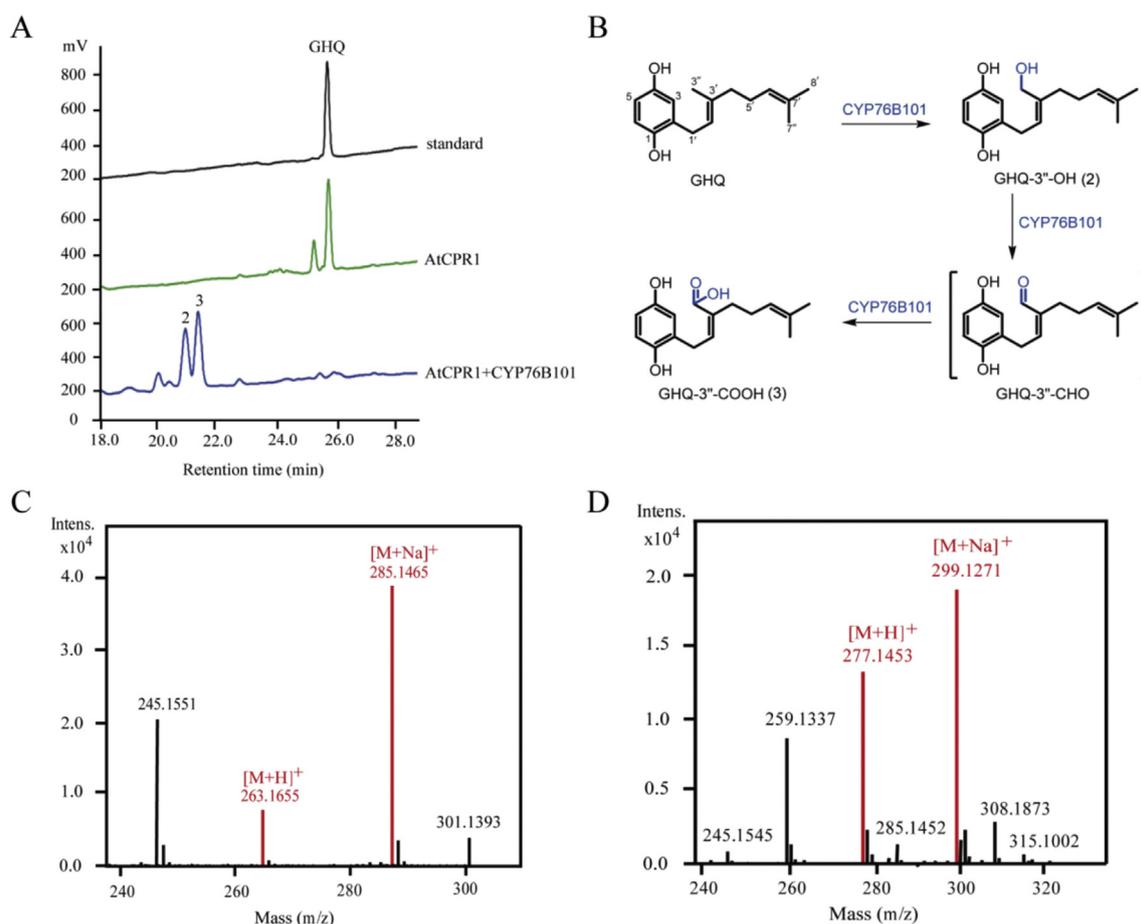
To verify the GHQ hydroxylation activity of CYP76B100 and CYP76B101, microsomal proteins were isolated from the engineered strains. With GHQ as the substrate, enzymatic assays of CYP76B100 or CYP76B101-containing microsomes were conducted. Compared with negative control assays, the enzymatically converted products showed an identical HPLC retention time to those of the GHQ-3''-OH and GHQ-3''-COOH produced in feeding experiments (Fig. 4). These results convincingly demonstrated that CYP76B100 catalyzed 3''-hydroxylation of GHQ to form GHQ-3''-OH. The other enzyme CYP76B101 catalyzed three-step oxidation of GHQ to form GHQ-3''-COOH.

## 2.5. Phylogenetic analysis of CYP76B100 and CYP76B101

To illustrate the evolutionary relationship among the CYP76B100, CYP76B101 and CYP76 family, phylogenetic analysis was constructed with the full-length amino acid sequences (Fig. S5 and Table S3). As



**Fig. 2.** LC-MS analysis of the biosynthetic products from GHQ catalyzed by CYP76B100 in feeding experiments. (A) The HPLC chromatogram of authentic standard GHQ (black line), the extracts from a control yeast strain expressing AtCPR1 alone (green line), and the extracts from the engineered yeast strain co-expressing CYP76B100 and AtCPR1 (red line). The labelled peaks correspond to GHQ-3''-OH (1), and authentic substrate GHQ. (B) Oxidation at the C-3'' position of substrate GHQ to GHQ-3''-OH (1) catalyzed by CYP76B100. (C) The mass spectrum for the product 1 eluting at 21.1 min.

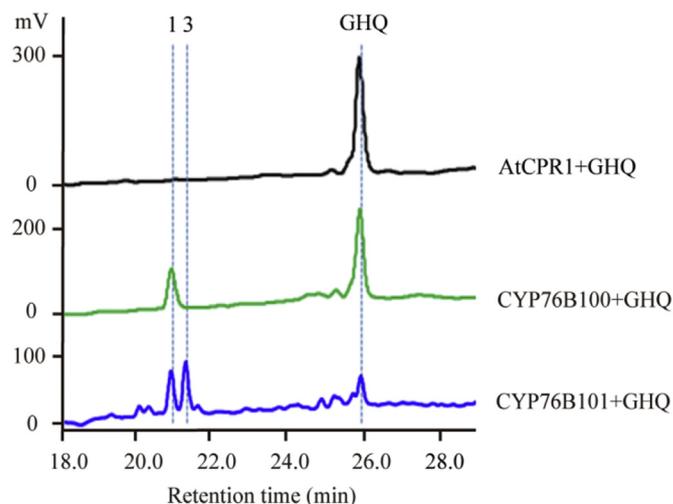


**Fig. 3.** LC-MS analysis of the biosynthetic products from GHQ catalyzed by CYP76B101 in feeding experiments. (A) The HPLC chromatogram of authentic standard GHQ (black line), the extracts from a control yeast strain expressing AtCPR1 alone (green line), and the extracts from the engineered yeast strain co-expressing CYP76B101 and AtCPR1 (blue line). The labelled peaks correspond to GHQ-3''-OH (2), GHQ-3''-COOH (3), and authentic substrate GHQ. (B) Oxidation at the C-3'' position of substrate GHQ to GHQ-3''-OH (2), to possible intermediate GHQ-3''-CHO (undetected) and finally to GHQ-3''-COOH (3), all catalyzed by CYP76B101. (C) The mass spectra for the product 2 eluting at 21.1 min. (D) The mass spectra for the product 3 eluting at 21.6 min.

**Table 1**  
<sup>1</sup>H and <sup>13</sup>C NMR spectroscopic data for products 2 and 3 (CD<sub>3</sub>OD, 400, 100 MHz, δ ppm).

position	products 2		Products 3	
	δ <sub>C</sub>	δ <sub>H</sub> (J in Hz)	δ <sub>C</sub>	δ <sub>H</sub> (J in Hz)
1	147.5, C		148.0, C	
2	128.2, C		126.6, C	
3	115.9, CH	6.54, d (3.0)	116.4, CH	6.61, d (3.0)
4	149.8, C		149.7, C	
5	112.8, CH	6.46, dd (8.5, 3.0)	113.4, CH	6.53, dd (8.5, 3.0)
6	115.2, CH	6.58, d (8.5)	115.5, CH	6.64, d (8.5)
1'	27.8, CH <sub>2</sub>	3.33, d (7.6)	30.6, CH <sub>2</sub>	3.67, d (8.0)
2'	126.1, CH	5.42, t (7.6)	137.8, CH	5.92, t (7.8)
3'	138.4, C		132.7, C	
4'	26.5, CH <sub>2</sub>	2.14, br s	34.6, CH <sub>2</sub>	2.32, t (7.5)
5'	34.9, CH <sub>2</sub>	2.15, br s	27.2, CH <sub>2</sub>	2.17, t (7.5)
6'	123.9, CH	5.11, br s	123.2, CH	5.12, t (7.2)
7'	130.9, C		131.6, C	
8'	24.4, CH <sub>3</sub>	1.64, s	24.4, CH <sub>3</sub>	1.66, s
9'	16.3, CH <sub>3</sub>	1.58, s	16.3, CH <sub>3</sub>	1.60, s
3''	58.8, CH <sub>2</sub> OH	4.20, s	171.5, COOH	

expected, the neighbor-joining tree confirmed that both CYP76B100 and CYP76B101 were members of the CYP76B subfamily, which includes CYP76B74 from *A. euchroma* and the enzyme geraniol 10-hydroxylase (G10H) hydroxylating geraniol to form 10-hydroxygeraniol. The three CYPs from Boraginaceae, CYP76B100, CYP76B101 and



**Fig. 4.** HPLC analysis of CYP76B100 and CYP76B101 microsomal enzyme assays *in vitro* incubated with GHQ. Microsomal preparation containing AtCPR1 alone was used as a negative control (black line). Microsomal proteins were respectively prepared from the engineered yeast strain co-expressing CYP76B100 and AtCPR1 (green line), and the engineered yeast strain co-expressing CYP76B101 and AtCPR1 (blue line).



64.7% identity with enzyme CYP76B74 from *A. euchroma* and 62.9% sequence identity with CYP76B101. In addition, CYP76B101 shares 93.6% identity with CYP76B74 and 60.5% sequence identity with G10H from *C. roseus*.

### 3. Discussion

Shikonins are the red naphthoquinone pigments produced in the root tissues of *L. erythrorhizon*, and display diverse medicinal properties. The long-standing and extensive uses, together with special formation of the naphthoquinone skeleton have aroused keen interest in shikonin biosynthesis. Despite the commercial importance of shikonin, the biosynthesis of the compound remains poorly understood. To illuminate shikonin biosynthesis, many whole-plants, suspension culture cell and hairy root culture-based transcriptome studies that rely on NGS have been previously reported (Takanashi et al., 2019; Wu et al., 2017; Rai et al., 2018). However, these studies were usually limited by incomplete transcriptomic analysis and inaccurate sequence information. To obtain more accurate and complete transcriptome data available for *L. erythrorhizon*, short-read NGS and long-read SMRT sequencing were combined. This strategy established the foundation for the discovery of genes involved in shikonin biosynthesis.

In this work, we focused on transcriptionally up-regulated candidate CYP genes in the red roots from differentially expressed unigenes, since shikonin accumulated in the red roots (Wu et al., 2017). Both heterologous expression of CYPs in yeasts and *in vitro* enzymatic assays revealed that CYP76B100 and CYP76B101 catalyzed the oxidation of GHQ to the hydroxyl or carboxyl group at the C-3" position. Milligram quantities of the new products were isolated and chemical structures of their oxidized products were identified by NMR spectroscopy.

It has been demonstrated that CYP76B74 catalyzes the C-3" hydroxylation of GHQ and is required in shikonin biosynthesis of *A. euchroma* (Wang et al., 2019). Phylogenetic analysis was carried out to illustrate the evolutionary relationship of the identified CYPs from CYP76 family. The neighbor-joining tree confirmed that CYP76B100 and CYP76B101 showed a proximal evolutionary relationship with CYP76B74, and G10H hydroxylating geraniol to form 10-hydroxygeraniol. The overall high sequence identity of CYP76B100, CYP76B101 and CYP76B74 indicates that these three enzymes may arise from a common ancestor. In the phylogenetic tree, the CYP76B subfamily is closest to the CYP76F subfamily, whose members involve in santalol and bergamotol biosynthesis in *Santalum album* (Diaz-Chavez et al., 2013). More distantly related CYPs of CYP76C subfamily from Arabidopsis are shown to be involved in the oxidation of monoterpenes (Boachon et al., 2015; Höfer et al., 2014; Wang et al., 2019). The members of CYP76 family play an important role in the synthesis of terpenoids (Wang et al., 2019), and especially in diterpene metabolism, more than half of all the identified angiosperm CYPs belong to the CYP76 family. It was found that the identified subfamilies of CYP76s underwent expansion in specific plant families that oxidized a group of diterpenes (Bathe and Tissier, 2019), such as tanshinone and forskolin (Fig. S5). Similarly, the CYP76Bs are expanded within Boraginaceae to form a Boraginaceae-specific clade, which are likely relevant to the biosynthesis of shikonin (Wang et al., 2019).

CYPs have been previously reported to catalyze multiple steps oxidation to form the carboxyl group, such as enzymes that are involved in the biosynthesis of artemisinic acid (CYP71AV1) (Teoh et al., 2006; Ro et al., 2006), abietic acid (CYP720B1) (Ro et al., 2005), and carnosic acid (C<sub>20</sub>Ox) (Scheler et al., 2016). In our current work, heterologous expression of CYP76B101 in yeast revealed the accumulation of GHQ-3"-OH and GHQ-3"-COOH, wherein the intermediate aldehyde GHQ-3"-CHO was not detected. The results suggest that CYP76B101 is ineffective at catalyzing the formation of the aldehyde GHQ-3"-CHO from GHQ-3"-OH. The product profile of CYP76B101 may be affected by heterologous expression systems, and it has been reported that CYP716A80 and CYP716A81 from *Barbarea vulgaris* transformed  $\beta$ -

amyrin into erythrodiol and oleanolic acid, and not oleanolic aldehyde in *S. cerevisiae* (Andre et al., 2016). In addition, a variant redox partner may modulate the catalytic activity of P450 enzymes by protein-protein interactions (Zhang et al., 2014).

It was previously speculated that the naphthalene ring could be formed by oxidation of GHQ-3"-OH to aldehyde, leading to C-C bond formation between the oxidized prenyl side-chain and the aromatic nucleus by an electrophilic reaction (Yamamoto et al., 2000). CYP76B101 catalyzed successive oxidations of GHQ at the C-3" position to synthesize GHQ-3"-COOH. Numerous studies have described the coupling of a carboxylic acid with an aromatic component by the Friedel-Crafts acylation reaction, which permits the preparation of aryl ketones using different chemical catalysts (Wilkinson, 2011; Babu et al., 2007; Kangani and Day, 2008; Li and Fuchs, 2003). Thus, naphthalene ring closure may be formed via GHQ-3"-COOH catalyzed by an unknown enzyme (Fig. 1). Recently, Friedel-Crafts alkylation and acylation have also been reported by employing enzymes as catalysts, such as prenyltransferase (Zhou et al., 2015) and acyltransferase (Schmidt et al., 2017; Pavkov-Keller et al., 2019). Even though it is still possible that the naphthoquinone skeleton can be formed directly from GHQ-3"-CHO, our work might provide another possible mechanism for the ring formation, which needs to be further confirmed by using multiple approaches including gene knockouts and characterization of downstream enzymes catalyzing the naphthoquinone formation in the future.

Plentiful quantities of dihydroechinofuran derived from GHQ were secreted and no naphthoquinone derivatives were detected in cell suspension cultures of *L. erythrorhizon* in a growth medium (Fukui et al., 1992). Since GHQ-3"-OH is the intermediate for the formation of both shikonin and dihydroechinofuran (Yamamoto et al., 2000), it is likely that the cyclization reaction that forms naphthoquinone skeleton or dihydroechinofuran may be regulated by fine-tuning the expression of CYP76B100 and CYP76B101 in cell suspension culture (Fig. 1). Further studies aimed at disrupting the function of CYP76B100 or CYP76B101 in plant tissues, for example, RNA interference (RNAi) may provide further insights into the physiological roles played by CYP76B100 and CYP76B101.

In summary, we report identification of two CYPs catalyzing oxidation at C-3" position of GHQ, which might provide more insights for shikonin biosynthesis in *L. erythrorhizon*.

### 4. Methods

#### 4.1. Chemicals and reagents

Chromatography-grade formic acid, isopropanol and methanol were obtained from Sigma-Aldrich, USA. Lithium acetate, PEG 3350, ssDNA and ampicillin were obtained from Solarbio, China. Geraniol, BF<sub>3</sub>·Et<sub>2</sub>O and 1, 4-hydroquinone were purchased from Sigma-Aldrich. Restriction enzymes, DNA polymerase and T4 ligase were purchased from Thermo Fisher Scientific, USA. The authentic standard of geranylhydroquinone (GHQ) was chemically synthesized in our lab as reported previously (Baeza et al., 2012). All other chemicals were of commercial reagent grade.

#### 4.2. Plants materials and growth conditions

The mature seeds of *Lithospermum erythrorhizon* Sieb. et Zucc. (Boraginaceae) were collected in the field of Chifeng city, Inner Mongolia Autonomous Region, China (located at 42°15'23.50" N; 118°52'58.14" E) in October 2016. The seeds were germinated on moist filter gauze as described in the reported methods (Fang et al., 2016). The seedlings with two cotyledons were subsequently transferred into 1.5 L circular plastic pots (15 cm in diameter) and filled with peat growth media. Plants were grown in a greenhouse at 23 ± 1 °C with a 16 h-light/8 h-dark photoperiod (Wu et al., 2017) and they were watered as needed. After 60 d, three plants at a height of 10 cm above

ground were harvested and respectively divided into three green leaves/stems samples and three red roots samples (each in triplicate). All tissues were flash frozen by liquid nitrogen and stored at  $-80\text{ }^{\circ}\text{C}$  till used.

#### 4.3. Transcriptome sequencing and bioinformatics analysis

Six total RNA samples were extracted respectively from green leaves/stems and red roots of *L. erythrorhizon* and used for subsequent cDNA library construction for transcriptome sequencing in Genewiz, China. Firstly, six mRNA samples from green leaves/stems and red roots tissues were subjected to paired-end sequencing on the Illumina HiSeq 2000 platform; Meanwhile, the six mixed poly(A) RNA samples were normalized and subjected to an SMRT sequencing using the Pacbio platform to generate the full-length transcriptome; Eventually, all the SMRT subreads were corrected using the NGS reads to resolve the high error rates of the subreads. The FPKM (fragments per kilobase of transcript per million mapped reads) value was used to quantify the expression level of each unigene and the DEGSeq program (Wang et al., 2009) was used to analyze the differentially expressed unigenes between green leaves/stems and red roots. With the set stringent threshold ( $|\text{Log}_2(\text{ratio})| \geq 1$  and  $\text{FDR} < 0.001$  and  $\text{Max}(\text{FPKM}) \geq 10$ ), there are 52 CYP genes transcriptionally up-regulated in red roots compared to green stem/leaf of *L. erythrorhizon*. Considering the integrity of open reading frame, a total of 40 up-regulated CYPs were selected as candidates for further functional screening (Supplementary Table S1).

#### 4.4. Strains, plasmid and growth conditions

*S. cerevisiae* BY4742 (*MAT $\alpha$  his3 $\Delta$ 1 leu2 $\Delta$ 0 lys2 $\Delta$ 0 ura3 $\Delta$ 0*), which is a derivative of S288C, was used as the parent strain for all engineered strains. Yeast strains were cultivated at  $30\text{ }^{\circ}\text{C}$  and 220 rpm in YPD medium containing  $10\text{ g l}^{-1}$  of yeast extract,  $20\text{ g l}^{-1}$  of beef peptone, and  $20\text{ g l}^{-1}$  of glucose or in SD-Ura medium (uracil-minus) containing  $6.7\text{ g l}^{-1}$  of yeast nitrogen base without amino acids,  $0.9\text{ g l}^{-1}$  of SD-Ura,  $20\text{ g l}^{-1}$  of glucose (Liu et al., 2018). *Escherichia coli* Trans-T1 (TransGen Biotech, China) was used for bacterial transformation and recombinant vectors construction. The *E. coli* strains with recombinant plasmids were grown at  $37\text{ }^{\circ}\text{C}$  and 200 rpm in Luria-Bertani medium with  $100\text{ mg l}^{-1}$  ampicillin.

#### 4.5. cDNA cloning and P450 gene mining

The total RNA from red roots was used as a template for first-strand cDNA synthesis using the PrimeScript RT reagent Kit with gDNA Eraser (TransGen Biotech) and the oligo(dT)15 primer following the manufacturer's recommended protocols. The deduced amino acid sequences of CYP genes were analyzed by the NCBI's BLAST tool and the full-length transcripts were cloned from the cDNA library of red roots. The amplified DNA fragments were cloned into pEASY-Blunt (TransGen Biotech) and the clones harboring the correct insert were confirmed by sequencing. The expression vectors harboring the candidate CYP genes were constructed for expression in yeasts.

#### 4.6. Construction of plasmids and strains

The yeast expression vector pCf302 with the constitutive promoters  $P_{\text{TEF1}}$ ,  $P_{\text{PGK1}}$ ,  $P_{\text{TDH3}}$  was constructed in our lab (Jiang et al., 2018). For the analysis of CYPs, the vector harboring the *A. thaliana* cytochrome P450 reductase gene (*AtCPR1*) was constructed. The *AtCPR1* gene with the restriction sites *Xho*I and *Bam*HI was synthesized by Generay Biotech Co., Ltd (Shanghai, China) with codon optimization for improving expression in *S. cerevisiae*. The *AtCPR1* gene was cloned into the plasmid pCf302 under the promoter  $P_{\text{TDH3}}$ , resulting in CPR expression vector pCf302-*AtCPR1*. The coding-region fragments for the candidate

CYPs were cloned into pCf302-*AtCPR1* to yield a series of CYPs carrying plasmids using primer pairs as described in Supplementary Table S2. All of the constructed CYPs expression plasmids were transformed into *S. cerevisiae* BY4742 and the transformations were carried out with the standard lithium acetate method (Gietz and Schiestl, 2007). The vector pCf302-*AtCPR1* was also transformed into *S. cerevisiae* BY4742 and served as a control. The recombinant yeast cells were selected on a uracil-minus plate (SD-Ura) at  $30\text{ }^{\circ}\text{C}$  for 3 d and verified by colony PCR.

#### 4.7. Feeding experiments

All the recombinant yeast strains carrying CYPs expression vectors and the empty vector were cultured in 2 ml SD-Ura medium at  $30\text{ }^{\circ}\text{C}$  and 220 rpm for 20 h, and three colonies were picked for each genotype to reduce errors. Subsequently, GHQ that was synthesized in our laboratory was added to the cultures to a final concentration of  $0.1\text{ g l}^{-1}$  and the cultures were further shaken at  $30\text{ }^{\circ}\text{C}$  and 220 rpm for 48 h. The yeast cells were harvested by centrifugation at 12,000 rpm for 10 min, and extracted twice with 1.5 ml methanol. Culture broth was extracted twice with an equivalent volume of ethyl acetate. The methanol and ethyl acetate extracts were mixed together, evaporated, and finally redissolved in methanol for HPLC analysis.

To isolate sufficient quantities of the oxidation products for chemical structure characterization, the *S. cerevisiae* BY4742 strains that respectively harbored pCf302-*AtCPR1*-CYP76B100 and pCf302-*AtCPR1*-CYP76B101 were cultivated in 1.5 L SD-Ura medium for purification of GHQ-3'-OH and GHQ-3'-COOH. The positive colonies were inoculated into a 250 ml flask containing 40 ml culture medium, and grown at  $30\text{ }^{\circ}\text{C}$  and 220 rpm for 24 h. The resulting seed cultures were transferred into 1.5 L fresh SD-Ura medium, and then cultivated them at  $30\text{ }^{\circ}\text{C}$  and 220 rpm for 24 h. Subsequently, the substrate GHQ (20 mg) was added to the cultures, which were further shaken at  $30\text{ }^{\circ}\text{C}$  and 220 rpm for 48 h. A total of 1.5 L of the culture broth was harvested and extracted twice with ethyl acetate as described above. The yeast cells were collected by centrifugation at 4000 rpm for 10 min and extracted twice with methanol. The methanol and ethyl acetate extracts were mixed together, evaporated, redissolved in methanol and further purified by semi-preparative HPLC.

#### 4.8. Synthesis of the substrate geranylhydroquinone (GHQ)

A solution of 1, 4-hydroquinone (0.15 g, 1.25 mmol) and geraniol (0.2 g, 1.25 mmol) was placed in a round bottom flask and dissolved in freshly distilled 1, 4-dioxane (5 ml). Under nitrogen gas and with vigorous stirring,  $\text{BF}_3 \cdot \text{Et}_2\text{O}$  (0.057 g, 0.4 mmol) was added dropwise at room temperature to the solution and the mixture was stirred at room temperature under a nitrogen atmosphere (Baeza et al., 2012). After 24 h, the mixture was poured onto crushed ice (about 10 g) and the organic layer extracted with ethyl acetate ( $3 \times 10\text{ ml}$ ). The organic solutions that were obtained after the extractions were mixed and dried over anhydrous sodium sulphate and filtered. The solvent was evaporated under reduced pressure and the crude residue was redissolved in methanol (2 ml). The resulting crude product was purified by semi-preparative HPLC separation to obtain the substrate GHQ.

Semi-preparative HPLC separation was performed using a Shimadzu LC-6 AD with a SPD-20A detector and a Shim-pack PREP-ODS (H) column ( $10 \times 250\text{ mm}$ ,  $5\text{ }\mu\text{m}$ ). The mobile phase contained 0.1% formic acid in water (A) and 100% HPLC grade methanol (B). The elution condition was adopted isocratically at 70% of B. Products were detected and quantified by UV absorption at 294 nm. The solvent flow rate was  $4.0\text{ ml min}^{-1}$ . The target fraction was collected manually, dried, and resuspended in  $\text{CD}_3\text{OD}$  for structural analysis.  $^1\text{H}$  NMR spectra was obtained on a 400 MHz Bruker Avance III spectrometer.

#### 4.9. HPLC, LC-MS and NMR analysis

The samples for HPLC analysis were cleaned off impurities by centrifugation (12,000 rpm) and by filtration using PTFE 0.2  $\mu\text{m}$  syringe filters (Axiva, Sigma Chemicals). In order to identify the function of the two new CYPs, 30  $\mu\text{l}$  methanol extracts from feeding experiments were analyzed by a Shimadzu HPLC system with a UV detector set at 294 nm. The analytical column used was an Agela Innoval C18 column (4.6  $\times$  250 mm, 5  $\mu\text{m}$ ). The mobile phase contained 0.1% formic acid in water (A) and 100% HPLC grade methanol (B). The gradient conditions were as follows: 0–3 min, 45% B; 3–22 min, a linear gradient of 45–100% B; 22–32 min, 100% B; 32–33 min, 100–45% B; 33–43 min, 45% B. The mobile phase flow was 1 ml  $\text{min}^{-1}$ .

To identify the oxidation products of CYP76B100 and CYP76B101, the LC-MS analysis was carried out on an Agilent 1200 HPLC system coupled with an Agilent Infinity UV detector and a Bruker-MicrOTOF-II mass spectrometer that was equipped with an electrospray ionization device. Data acquisition and processing were done with MicrOTOF control version 3.0/Data Analysis Version 4.0 software. For HPLC analysis, the Agela Innoval C18 column (4.6  $\times$  250 mm, 5  $\mu\text{m}$ ) was used. The mobile phase contained 0.1% formic acid in water (A) and 100% HPLC grade methanol (B). The gradient programs were as follows: 0–3 min, 45% B; 3–22 min, a linear gradient of 45–100% B; 22–32 min, 100% B; 32–33 min, 100–45% B; 33–43 min, 45% B. The mobile phase flow was 1 ml  $\text{min}^{-1}$ , and the column temperature was set at 40  $^{\circ}\text{C}$ . The products were detected at 294 nm. Optimized MS operating conditions were as follows: all spectra were obtained in positive mode over an m/z range of 50–1000 under a dry gas flow of 6.0 l  $\text{min}^{-1}$ , a dry temperature of 180  $^{\circ}\text{C}$ , a nebulizer pressure of 1 bar and a probe voltage of  $-4.5$  kV.

To purify sufficient oxidation products from CYP76B100 and CYP76B101, substances GHQ-3'-OH and GHQ-3'-COOH were separated using semi-preparative HPLC operated in a Shimadzu LC-6 AD with SPD-20A detector and a Shim-pack PREP-ODS (H) column (10  $\times$  250 mm, 5  $\mu\text{m}$ ). The mobile phase contained 0.1% formic acid in water (A) and 100% HPLC grade methanol (B). The column was equilibrated with 45% B, and then 200  $\mu\text{l}$  of the enriched sample were injected, reaching 100% B after 18 min and running it for 6 min. Subsequently, it was returned to 45% B to equilibrate the column for 6 min before the next injection. Products were detected and quantified by UV absorption at 294 nm. The solvent flow rate was 4.0 ml  $\text{min}^{-1}$ . The target fraction that was eluted between 15.6 and 16.3 min (GHQ-3'-OH) or between 16.4 and 17.0 min (GHQ-3'-COOH) was repeatedly collected, dried and resuspended in  $\text{CD}_3\text{OD}$  for structural analysis. The chemical structure of the newly acquired substance GHQ-3'-OH was verified by  $^1\text{H}$  NMR and  $^{13}\text{C}$  NMR spectra. The chemical structure of substance GHQ-3'-COOH was verified by  $^1\text{H}$  NMR,  $^{13}\text{C}$  NMR and heteronuclear multiple-bond correlation spectroscopy (HMBC) spectra. The NMR spectra were collected on a 400 MHz Bruker Avance III spectrometer.

#### 4.10. Microsome isolation and in vitro enzyme assays

The strain *S. cerevisiae* BY4742 harboring pCf302-AtCPR1-CYP76B100 or pCf302-AtCPR1-CYP76B101 was cultivated respectively for microsome isolation. Microsomes extracted from the yeast carrying the vector pCf302-AtCPR1 were assayed as a negative control. The strains were inoculated into culture tubes containing 5 ml of SD-Ura medium and grown at 30  $^{\circ}\text{C}$  and 220 rpm for 16 h. The seed cultures were inoculated into 200 ml SD-Ura medium and further cultivated at 30  $^{\circ}\text{C}$  and 220 rpm for 36 h. The microsomal isolation was carried out by differential centrifugation as described previously (Krithika et al., 2015; Scheler et al., 2016) with the following modifications. The cells were centrifuged (6000 g, 10 min) and washed twice with TEK buffer (50 mM Tris-HCl pH 7.5, 1 mM EDTA, 100 mM KCl) and centrifuged. Collected cells were resuspended in TES buffer (50 mM Tris-HCl pH 7.5,

1 mM EDTA, 600 mM Sorbitol, 1 mM PMSF) and disrupted on ice by using a high-pressure homogenizer (JNBIO). The lysed cells were centrifuged (11,000 g, 20 min) to remove the cell debris, mitochondria and nuclei. The supernatant was transferred to ultracentrifugation tubes and centrifuged at 150,000 g for 1.5 h. The supernatant was discarded and the harvested microsomes were resuspended in 1 ml TEG buffer (50 mM Tris-HCl pH 7.5, 1 mM EDTA, and 20% glycerol), aliquoted (200  $\mu\text{l}$ ) and stored at  $-80$   $^{\circ}\text{C}$ . All steps for microsomal preparation were carried out at 4  $^{\circ}\text{C}$ .

The enzymatic assays were conducted in 200  $\mu\text{l}$  reaction mixtures containing 50 mM sodium phosphate buffer (pH 7.5), 1 mM NADPH, 50  $\mu\text{l}$  microsomal proteins, and 100  $\mu\text{M}$  substrate GHQ. After incubation for 3 h at 30  $^{\circ}\text{C}$ , the reactions were terminated by adding 200  $\mu\text{l}$  methanol. The denatured proteins were removed by centrifugation at 12,000 rpm for 10 min and the supernatant was submitted to HPLC analysis.

#### 4.11. Phylogenetic analysis

The phylogenetic tree was constructed using Mega 6.0 software package (Tamura et al., 2013) and the neighbor-joining program based on the Poisson model and a bootstrap of 1000 replicates. All the amino acid sequences used for phylogenetic tree construction were listed in Supplementary Table S3 and aligned by ClustalX. The scale bar indicates 0.2 amino acid substitution per site. The numbers at the nodes of each branch indicate the percentage of bootstrap values.

#### Funding information

This work was supported by the National Key Research and Development Program (2019YFA0905700), the National Natural Science Foundation of China (31400026, 31970065, U1902214), the Sciences and Technology Planning Projects of Tianjin city (18YFZCSY01350), and the Yunnan Key Research and Development Program (2019ZF011-2).

#### Notes

The authors declare the following competing financial interest(s): Part of the work has been included in a patent application by the Tianjin Institute of Industrial Biotechnology.

#### Acknowledgements

We thank Dr. David R. Nelson (University of Tennessee) for naming of the new CYPs.

#### Appendix B. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.phytochem.2020.112375>.

#### Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### References

- Andre, C.M., Legay, S., Deleruelle, A., Nieuwenhuizen, N., Punter, M., Brendolise, C., Cooney, J.M., Lateur, M., Hausman, J.F., Larondelle, Y., Laing, W.A., 2016. Multifunctional oxidosqualene cyclases and cytochrome P450 involved in the biosynthesis of apple fruit triterpenic acids. *New Phytol.* 211, 1279–1294.
- Andújar, I., Ríos, J.L., Giner, R.M., Recio, M.C., 2013. Pharmacological properties of shikonin—a review of literature since 2002. *Planta Med.* 79, 1685–1697.
- Au, K.F., Underwood, J.G., Lee, L., Wong, W.H., 2012. Improving PacBio long read

- accuracy by short read alignment. *PLoS One* 7, e46679.
- Babu, S.A., Yasuda, M., Baba, A., 2007. In(III)-mediated chemoselective dehydrogenative interaction of  $\text{ClMe}_2\text{SiH}$  with carboxylic acids: direct chemo- and regioselective Friedel-Crafts acylation of aromatic ethers. *Org. Lett.* 9, 405–408.
- Baeza, E., Catalán, K., Peña-Cortés, H., Espinoza, L., Villena, J., Carrasco, H., 2012. Synthesis of geranylhydroquinone derivatives with potential cytotoxic activity. *Quim. Nova* 35, 523–526.
- Bathe, U., Tissier, A., 2019. Cytochrome P450 enzymes: a driving force of plant diterpene diversity. *Phytochemistry* 161, 149–162.
- Boachon, B., Junker, R.R., Miesch, L., Bassard, J.E., Höfer, R., Caillieudeaux, R., Seidel, D.E., Lesot, A., Heinrich, C., Ginglinger, J.F., Allouche, L., Vincent, B., Wahyuni, D.S., Paetz, C., Beran, F., Miesch, M., Schneider, B., Leiss, K., Werck-Reichhart, D., 2015. CYP76C1 (cytochrome P450)-mediated linalool metabolism and the formation of volatile and soluble linalool oxides in *Arabidopsis* flowers: a strategy for defense against floral antagonists. *Plant Cell* 27, 2972–2990.
- Diaz-Chavez, M.L., Moniodis, J., Madilao, L.L., Jancsik, S., Keeling, C.I., Barbour, E.L., Ghisalberti, E.L., Plummer, J.A., Jones, C.G., Bohlmann, J., 2013. Biosynthesis of sandalwood oil: *Santalum album* CYP76F cytochromes P450 produce santalols and bergamotol. *PLoS One* 8, e75053.
- Fang, R., Wu, F., Zou, A., Zhu, Y., Zhao, H., Zhao, H., Liao, Y., Tang, R.J., Yang, T., Pang, Y., Wang, X., Yang, R., Qi, J., Lu, G., Yang, Y., 2016. Transgenic analysis reveals LeACS-1 as a positive regulator of ethylene-induced shikonin biosynthesis in *Lithospermum erythrorhizon* hairy roots. *Plant Mol. Biol.* 90, 345–358.
- Fukui, H., Tani, M., Tabata, M., 1992. An unusual metabolite, dihydroechinofuran, released from cultured cells of *Lithospermum erythrorhizon*. *Phytochemistry* 31, 519–521.
- Gao, W., Sun, H.X., Xiao, H., Cui, G., Hillwig, M.L., Jackson, A., Wang, X., Shen, Y., Zhao, N., Zhang, L., 2014. Combining metabolomics and transcriptomics to characterize tanshinone biosynthesis in *Salvia miltiorrhiza*. *BMC Genom.* 15, 73.
- Gietz, R.D., Schiestl, R.H., 2007. Frozen competent yeast cells that can be transformed with high efficiency using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* 2, 1–4.
- Guo, J., Zhou, Y.J., Hillwig, M.L., Shen, Y., Yang, L., Wang, Y., Zhang, X., Liu, W., Peters, R.J., Chen, X., Zhao, Z.K., Huang, L., 2012. CYP76AH1 catalyzes turnover of miltiradiene in tanshinones biosynthesis and enables heterologous production of ferruginol in yeasts. *Proc. Natl. Acad. Sci. U.S.A.* 110, 12108–12113.
- Höfer, R., Boachon, B., Renault, H., Gavira, C., Miesch, L., Iglesias, J., Ginglinger, J.F., Allouche, L., Miesch, M., Grec, S., Larbat, R., Werck-Reichhart, D., 2014. Dual function of the cytochrome P450 CYP76 family from *Arabidopsis thaliana* in the metabolism of monoterpenols and phenylurea herbicides. *Plant Physiol.* 166, 1149–1161.
- Inouye, H., Ueda, S., Inoue, K., Matsumura, H., 1979. Biosynthesis of shikonin in callus cultures of *Lithospermum erythrorhizon*. *Phytochemistry* 18, 1301–1308.
- Jiang, J., Yin, H., Wang, S., Zhuang, Y., Liu, S., Liu, T., Ma, Y., 2018. Metabolic engineering of *Saccharomyces cerevisiae* for high-level production of salidroside from glucose. *J. Agric. Food Chem.* 66, 4431–4438.
- Kangani, C.O., Day, B.W., 2008. Mild, efficient Friedel-Crafts acylations from carboxylic acids using cyanuric chloride and  $\text{AlCl}_3$ . *Org. Lett.* 10, 2645–2648.
- Krithika, R., Srivastava, P.L., Rani, B., Kolet, S.P., Chopade, M., Soniya, M., Thulasiram, H.V., 2015. Characterization of 10-hydroxygeraniol dehydrogenase from *Catharanthus roseus* reveals cascaded enzymatic activity in iridoid biosynthesis. *Sci. Rep.* 5, 8258.
- Li, S.M., Hennig, S., Heide, L., 1998. Shikonin: a geranyl diphosphate-derived plant hemiterpenoid formed via the mevalonate pathway. *Tetrahedron Lett.* 39, 2721–2724.
- Li, W., Fuchs, P.L., 2003. Polyphosphoric acid trimethylsilyl ester promoted intramolecular acylation of an olefin by a carboxylic acid: convenient construction of C-18-functionalized delta14-hecogenin acetate. *Org. Lett.* 5, 4061–4064.
- Liu, X., Cheng, J., Zhang, G., Ding, W., Duan, L., Yang, J., Kui, L., Cheng, X., Ruan, J., Fan, W., Chen, J., Long, G., Zhao, Y., Cai, J., Wang, W., Ma, Y., Dong, Y., Yang, S., Jiang, H., 2018. Engineering yeast for the production of breviscapine by genomic analysis and synthetic biology approaches. *Nat. Commun.* 9, 448.
- Malik, S., Bhushan, S., Sharma, M., Ahuja, P.S., 2016. Biotechnological approaches to the production of shikonins: a critical review with recent updates. *Crit. Rev. Biotechnol.* 36, 327–340.
- Miettinen, K., Dong, L., Navrot, N., Schneider, T., Burlat, V., Pollier, J., Woittiez, L., Krol, S., Lugan, R., Ilc, T., Verpoorte, R., Oksman-Caldentey, K.M., Martinoia, E., Bouwmeester, H., Goossens, A., Memelink, J., Werck-Reichhart, D., 2014. The seco-iridoid pathway from *Catharanthus roseus*. *Nat. Commun.* 5, 3606.
- Okamoto, T., Yazaki, K., Tabata, M., 1995. Biosynthesis of shikonin derivatives from L-phenylalanine via deoxyshikonin in *Lithospermum* cell cultures and cellfree extracts. *Phytochemistry* 38, 83–88.
- Pavkov-Keller, T., Schmidt, N.G., Żądło-Dobrowolska, A., Kroutil, W., Gruber, K., 2019. Structure and catalytic mechanism of a bacterial Friedel-Crafts acylase. *ChemBiochem* 20, 88–95.
- Rai, A., Nakaya, T., Shimizu, Y., Rai, M., Nakamura, M., Suzuki, H., Saito, K., Yamazaki, M., 2018. *De novo* transcriptome assembly and characterization of *Lithospermum officinale* to discover putative genes involved in specialized metabolites biosynthesis. *Planta Med.* 84, 920–934.
- Ro, D.K., Arimura, G.I., Lau, S.Y., Piers, E., Bohlmann, J., 2005. Loblolly pine abietadienol/abietadienal oxidase PTAO (CYP720B1) is a multifunctional, multisubstrate cytochrome P450 monooxygenase. *Proc. Natl. Acad. Sci. U.S.A.* 102, 8060–8065.
- Ro, D.K., Paradise, E.M., Ouellet, M., Fisher, K.J., Newman, K.L., Ndungu, J.M., Ho, K.A., Eachus, R.A., Ham, T.S., Kirby, J., Chang, M.C., Withers, S.T., Shiba, Y., Sarpong, R., Keasling, J.D., 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440, 940–943.
- Scheler, U., Brandt, W., Porzel, A., Rothe, K., Manzano, D., Božič, D., Papaefthimiou, D., Balcke, G.U., Henning, A., Lohse, S., Marillonnet, S., Kanellis, A.K., Ferrer, A., Tissier, A., 2016. Elucidation of the biosynthesis of carnosic acid and its reconstitution in yeast. *Nat. Commun.* 7, 12942.
- Schmidt, N.G., Pavkov-Keller, T., Richter, N., Wiltschi, B., Gruber, K., Kroutil, W., 2017. Biocatalytic Friedel-Crafts acylation and fries reaction. *Angew. Chem. Int. Ed.* 56, 7615–7619.
- Tabata, M., Fujita, Y., 1985. Production of shikonin by plant cell cultures. In: Zaitlin, M., Day, P., Hollaender, A. (Eds.), *Biotechnology in Plant Science*, pp. 207–218 Orlando.
- Takanashi, K., Nakagawa, Y., Aburaya, S., Kaminade, K., Aoki, W., Saida-Munakata, Y., Sugiyama, A., Ueda, M., Yazaki, K., 2019. Comparative proteomic analysis of *Lithospermum erythrorhizon* reveals regulation of a variety of metabolic enzymes leading to comprehensive understanding of the shikonin biosynthetic pathway. *Plant Cell Physiol.* 60, 19–28.
- Tamura, K., Stecher, G., Peterson, D., Filipowski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729.
- Teoh, K.H., Polichuk, D.R., Reed, D.W., Nowak, G., Covello, P.S., 2006. *Artemisia annua* L. (Asteraceae) trichome-specific cDNAs reveal CYP71AV1, a cytochrome P450 with a key role in the biosynthesis of the antimalarial sesquiterpene lactone artemisinin. *FEBS Lett.* 580, 1411–1416.
- Wang, S., Wang, R., Liu, T., Lv, C., Liang, J., Kang, C., Zhou, L., Guo, J., Cui, G., Zhang, Y., Werck-Reichhart, D., Guo, L., Huang, L., 2019. CYP76B74 catalyzes the 3'-hydroxylation of geranylhydroquinone in shikonin biosynthesis. *Plant Physiol.* 179, 402–414.
- Wang, S., Wang, R., Liu, T., Zhan, Z., Kang, L., Wang, Y., Lv, C., Werck-Reichhart, D., Guo, L., Huang, L., 2017. Production of 3-geranyl-4-hydroxybenzoate acid in yeast, an important intermediate of shikonin biosynthesis pathway. *FEMS Yeast Res.* 17, fox065.
- Wang, L., Feng, Z., Wang, X., Wang, X., Zhang, X., 2009. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* 26, 136–138.
- Wilkinson, M.C., 2011. Greener" Friedel-Crafts acylations: a metal- and halogen-free methodology. *Org. Lett.* 13, 2232–2235.
- Wu, F.Y., Tang, C.Y., Guo, Y.M., Bian, Z.W., Fu, J.Y., Lu, G.H., Qi, J.L., Pang, Y.J., Yang, Y.H., 2017. Transcriptome analysis explores genes related to shikonin biosynthesis in *Lithospermeae* plants and provides insights into Boraginales' evolutionary history. *Sci. Rep.* 7, 4477.
- Xu, Z., Peters, R.J., Weirather, J., Luo, H., Liao, H., Zhang, X., Zhu, Y., Ji, A., Zhang, B., Hu, S., Au, K.F., Song, J., Chen, S., 2015. Full-length transcriptome sequences and splice variants obtained by a combination of sequencing platforms applied to different root tissues of *Salvia miltiorrhiza* and tanshinone biosynthesis. *Plant J.* 82, 951–961.
- Yamamoto, H., Inoue, K., Li, S.M., Heide, L., 2000. Geranylhydroquinone 3'-hydroxylase, a cytochrome P-450 monooxygenase from *Lithospermum erythrorhizon* cell suspension cultures. *Planta* 210, 312–317.
- Yazaki, K., 2017. *Lithospermum erythrorhizon* cell cultures: present and future aspects. *Plant Biotechnol.* 34, 131–142.
- Yazaki, K., Kunihisa, M., Fujisaki, T., Sato, F., 2002. Geranyl diphosphate: 4-hydroxybenzoate geranyltransferase from *Lithospermum erythrorhizon*. *J. Biol. Chem.* 277, 6240–6246.
- Zhang, W., Liu, Y., Yan, J., Cao, S., Bai, F., Yang, Y., Huang, S., Yao, L., Anzai, Y., Kato, F., Podust, L.M., Sherman, D.H., Li, S., 2014. New reactions and products resulting from alternative interactions between the P450 enzyme and redox partners. *J. Am. Chem. Soc.* 136, 3640–3646.
- Zhou, K., Ludwig, L., Li, S.M., 2015. Friedel-Crafts alkylation of acylphloroglucinols catalyzed by a fungal indole prenyltransferase. *J. Nat. Prod.* 78, 929–933.