Functional characterization of the cytochrome P450 monooxygenase CYP71AU87 indicates a role in marrubiin biosynthesis in the medicinal plant *Marrubium vulgare*

Prema S. Karunanithi¹, Puja Dhanota¹, J. Bennett Addison², Shen Tong³, Oliver Fiehn³,⁴ and Philipp Zerbe¹*

**Abstract**

**Background:** Horehound (*Marrubium vulgare*) is a medicinal plant whose signature bioactive compounds, marrubiin and related furanoid diterpenoid lactones, have potential applications for the treatment of cardiovascular diseases and type II diabetes. Lack of scalable plant cultivation and the complex metabolite profile of *M. vulgare* limit access to marrubiin via extraction from plant biomass. Knowledge of the marrubiin-biosynthetic enzymes can enable the development of metabolic engineering platforms for marrubiin production. We previously identified two diterpene synthases, *MvCPS1* and *MvELS*, that act sequentially to form 9,13-epoxy-labd-14-ene. Conversion of 9,13-epoxy-labd-14-ene by cytochrome P450 monooxygenase (P450) enzymes can be hypothesized to facilitate key functional modification reactions in the formation of marrubiin and related compounds.

**Results:** Mining a *M. vulgare* leaf transcriptome database identified 95 full-length P450 candidates. Cloning and functional analysis of select P450 candidates showing high transcript abundance revealed a member of the CYP71 family, CYP71AU87, that catalyzed the hydroxylation of 9,13-epoxy-labd-14-ene to yield two isomeric products, 9,13-epoxy labd-14-ene-18-ol and 9,13-epoxy labd-14-ene-19-ol, as verified by GC-MS and NMR analysis. Additional transient *Nicotiana benthamiana* co-expression assays of CYP71AU87 with different diterpene synthase pairs suggested that CYP71AU87 is specific to the sequential *MvCPS1* and *MvELS* product 9,13-epoxy-labd-14-ene. Although the P450 products were not detectable in planta, high levels of CYP71AU87 gene expression in marrubiin-accumulating tissues supported a role in the formation of marrubiin and related diterpenoids in *M. vulgare*.

**Conclusions:** In a sequential reaction with the diterpene synthase pair *MvCPS1* and *MvELS*, CYP71AU87 forms the isomeric products 9,13-epoxy labd-14-ene-18/19-ol as probable intermediates in marrubiin biosynthesis. Although its metabolic relevance in planta will necessitate further genetic studies, identification of the CYP71AU87 catalytic activity expands our knowledge of the functional landscape of plant P450 enzymes involved in specialized diterpenoid metabolism and can provide a resource for the formulation of marrubiin and related bioactive natural products.

**Keywords:** Diterpene synthase, Cytochrome P450 monooxygenase, Diterpenoid biosynthesis, Marrubiin, Plant natural products, *Marrubium vulgare*
Background

Medicinal plants and their natural products are a rich, yet underutilized, source for modern therapeutics that stems from a wealth of knowledge of their use as traditional medicines [1, 2]. Among the myriad natural products formed in plants, diterpenoids are a diverse group of more than 12,000 metabolites [3] that have proven valuable for drug discovery. Examples include the chemotherapeutic agent Taxol® naturally produced in yew trees (Taxus spp.) [4, 5], the psychoactive clerodane diterpenoid salvorin A [6, 7], ingenol mebutate from species of Euphorbia for the treatment of actinic keratosis [8, 9], and the cAMP-regulator forskolin from Coleus forskohlii [10–12].

This study focuses on the medicinal plant white horehound (Marrubium vulgare), a member of the mint family (Lamiaceae), that has been traditionally used to alleviate dermatological and respiratory ailments. M. vulgare extracts and metabolites have been shown to have potential for treating type II diabetes and cardiovascular diseases [13, 14]. The anti-diabetic efficacy of M. vulgare has been attributed to marrubiin, a furanoid diterpene lactone that represents the signature metabolite in M. vulgare [15–17]. In vitro and in vivo studies demonstrated that marrubiin stimulates insulin secretion under hyperglycemic conditions with a potentially higher efficacy as compared to established anti-diabetic drugs such as metformin [13, 14]. Marrubiin accumulates in M. vulgare leaf glandular trichomes and flowers at levels of up to 4 mg per gram fresh weight [17]. In addition, phytochemical analyses of M. vulgare and closely related Marrubium species have revealed a suite of structurally similar labdane diterpenoids that feature common furan ring and 19,6-lactone functional groups [15, 16]. These functional modifications have attracted increasing interest in recent years [7, 16, 18–20], since similar furan and lactone groups have been implicated with natural product bioactivity, such as the contribution of the furan ring in the κ-opioid receptor agonist activity of salvorin A or the efficacy of sesquiterpene lactones such as the anti-malaria drug artemisinin [21, 22].

Toward elucidating the marrubiin-biosynthetic pathway in M. vulgare, we previously reported on the identification of gene candidates for terpenoid biosynthesis in M. vulgare using a genomics-based gene discovery approach [16, 23]. A functional pair of diterpene synthases (diTPSs), designated PEREGRINOL DIPHOSPHATE SYNTHASE (MvCPS1) and 9,13-EPOXY-LABD-14-ENE SYNTHASE (MvELS), were identified that catalyze the first committed reactions in marrubiin biosynthesis [16]. MvCPS1 transforms the central diterpenoid precursor geranylgeranyl diphosphate (GGPP) into peregrinol diprophosphate, a bicyclic prenyl diphosphate intermediate that features a hydroxy group as C-9 characteristically present in marrubiin and related metabolites (Fig. 1). In a sequential reaction facilitated by MvELS, the diphosphate moiety is ionized and the resulting carbocation undergoes rearrangement to form 9,13-epoxy-14-labd-ene [16]. Downstream of this intermediate, several oxidative reactions catalyzed by members of the large family of cytochrome P450 monooxogenases (P450s) presumably facilitate the functional decoration of the diterpenoid scaffold to yield marrubiin and structurally similar bioactive diterpenoids (Fig. 1). Specifically, hydroxylation or carboxylation at C-6 and C-19 would have the potential to form the characteristic γ-lactone ring structure, and hydroxylation at C-16 and/or C-15 would facilitate formation of the furan ring.

In related Lamiaceae species, members of the CYP71 and CYP76 families of the P450 superfamily have been shown to catalyze position-specific oxygenation reactions in diterpenoid metabolism [24–28]. For example, CYP76AH1 and CYP76AH4 from Salvia miltiorrhiza and rosemary (Rosmarinus officinalis) catalyze oxygenation at C-11 and/or C-12 of abietatriene en route to feruginol and downstream tanshinone diterpenoids [29, 30]. Similarly, CYP76AH24, CYP71BE52 and CYP76AK6/8, were shown to facilitate hydroxylation and carboxylation at C-12, C-2 or C-20, respectively, of related labdane scaffolds [26, 31]. Although plant P450s directly involved in the formation of diterpenoid lactones have not yet been reported, in rice (Oryza sativa), CYP99A3 and CYP76M8 were demonstrated to catalyze C-19 carboxylation and C-6 hydroxylation respectively as a prerequisite for lactone ring formation in bioactive monomilacone diterpenoids [32, 33]. Additionally, P450s involved in furan ring formation have been reported, including CYP71A32 from peppermint (Mentha x piperita) involved in the biosynthesis of the monoterpene menthofuran [34], and CYP76BK1 from Vitex agnus-castus hydroxylates peregrol at C-16 potentially en route to furan ring closure in the biosynthesis of diterpenoids with structural similarity to marrubiin [18].

In this study, we combined the interrogation of an established leaf transcriptome inventory with phylogenetic and gene expression analyses to identify the CYP71 family member, Mv1270 (CYP71AU87), as a probable candidate for a function in marrubiin biosynthesis. Co-expression of CYP71AU87 with MvCPS1 and MvELS in Nicotiana benthamiana and yeast (Saccharomyces cerevisiae) resulted in the oxygenation of the MvELS product, 9,13-epoxy-labd-14-ene, at position C-18 or C-19 as verified by GC-MS and NMR analysis. Although these isomeric diterpene alcohols could not be identified in planta using GC-MS and LC-MS analysis, high expression levels of CYP71AU87 transcript in marrubiin-accumulating tissues of M. vulgare supports a possible role in marrubiin biosynthesis. Thus, the discovery of CYP71AU87
provides new resources toward elucidating and metabolic pathway engineering of the production of marrubiin and related natural products.

**Results**

Transcriptomics-enabled identification of P450 genes with possible roles in marrubiin biosynthesis

Previous studies demonstrated that marrubiin and related diterpenoid metabolites accumulate predominantly in leaves and leaf trichomes of *M. vulgare* [16, 17]. To identify P450 candidates with possible functions in marrubiin biosynthesis, we mined an established *M. vulgare* leaf transcriptome [23] against a P450-specific protein database (Additional file 1: Data file S1). Applying a minimal sequence length of 960 bp and an E-value cut-off of $\leq 1 \times 10^{-50}$, a total of 95 candidate genes with significant matches to known P450 enzymes were identified (Additional file 2: Table S1). Phylogenetic analysis placed these P450 candidates into 33 P450 families with members of the CYP71 family being the most represented (Fig. 2). To further triage high-probability P450 candidates, gene expression levels were assessed on the basis of FPKM (Fragments Per Kilobase of transcript per Million mapped reads) values obtained by mapping the raw reads against the assembled *M. vulgare* leaf transcriptome. Highest transcript abundance was observed for members of the CYP71 (*Mv*1270, *Mv*2069, *Mv*3392, *Mv*4213), CYP72 (*Mv*1123), CYP76 (*Mv*1960), CYP81 (*Mv*3239), CYP706 (*Mv*1545), and CYP97 (*Mv*2827) families (Fig. 2). Since members of the CYP71 and CYP76 families have been demonstrated to function in diterpenoid metabolism in Lamiaceae and other plant species [26, 27, 30, 31, 35], *Mv*1270, *Mv*2069, *Mv*3392 and *Mv*4213 were pursued for functional studies. Despite extensive efforts, *Mv*2069 and the CYP76 candidate *Mv*1960 could not be cloned successfully, preventing further analysis. Phylogenetic analysis placed *Mv*2069 and the CYP76 candidate *Mv*1960 could not be cloned successfully, preventing further analysis. Phylogenetic analysis placed *Mv*1270, albeit distantly, into a branch that also contained known terpenoid-metabolic enzymes of the CYP71 family, such as the menthofuran synthase CYP71A32 from *Mentha x piperita* [34] and *Tanacetum parthenium* CYP71CA1 and CYP71CB1 involved in costunolide biosynthesis [36] (Fig. 2). Another lower abundant P450 assigned to the CYP71 family, *Mv*6504, was also located on this branch and showed a close phylogenetic relationship to CYP71A32 and was selected for functional analysis, due to a possibly similar activity in furan ring formation. *Mv*3392 was placed into a different branch distantly related to members of the CYP71D subfamily with roles in monoterpene biosynthesis in species of mint.
addition, Mv1545 was chosen as a candidate enzyme, since it represented the most abundant transcript of all identified P450s and was assigned to the CYP706 family with known activities in sesquiterpenoid metabolism [37, 38]. Similarly, Mv4213 was selected for further study, due to its high transcript abundance and only distant relationship to other M. vulgare P450 candidates (Fig. 2). The remaining abundant P450 transcripts assigned to the CYP72, CYP97 and CYP81 families were not further investigated in this study, since members of these families more commonly function in brassinosteroid, carotenoid or phenylpropanoid metabolism as demonstrated in other plant species [27, 39].

Clarification of reaction products derived from the combined activity of MvCPS1 and MvELS

Our prior work on the in vitro pairwise reaction MvELS showed three reaction products (compounds 1–3) [16]. The major product (compound 1) was identified as 9,13-epoxy labd-14-ene, whereas low abundance of the remaining products prevented structural analysis [16]. Compound 3 likely represents an isomer of 9,13-epoxy labd-14-ene as based on a near identical retention time and fragmentation pattern. Conversely, the mass spectrum of compound 2 showed additional abundant ions of m/z 191 and 209, indicative of a more distinct structure (Fig. 3). To clarify the product profile of MvELS, we performed large-scale co-expression assays of MvCPS1 and MvELS with a GGPP synthase from Abies grandis in an E. coli platform engineered for diterpenoid production [40]. Purification of compound 3 remained unsuccessful, due to its co-elution with compound 1. In contrast, purification of compound 2 using an optimized protocol combining silica chromatography and semi-preparative HPLC enabled 1D 1H-NMR and 13C-NMR, as well as 2D NMR experiments: COSY, HSQC, HMBC and H2BC
This approach identified compound 2 as labda-13(16),14-dien-9-ol, which features a free hydroxyl group at C-9 (13C, 84.6 ppm) as compared to the 9,13-spiroether (13C, 93.2 ppm) observed in 9,13-epoxy labd-14-ene (Fig. 3 and Additional file 3: Figure S1) [16]. This was evidenced by the presence of two double bonds between C-13 (13C, 147.54 ppm) and C-16 (13C, 115.43 ppm) as well as between C-14 (13C, 138.88 ppm) and C-15 (13C, 113.35 ppm) in compound 2, whereas compound 1 showed one double bond between C-14 (13C, 147.2 ppm) and C-15 (13C, 110.7 ppm), but not between C-13 (13C, 84.6 ppm) and C-16 (13C, 23.5 ppm) (Additional file 3: Figure S1) [16].

Biochemical characterization of CYP71AU87 as a 9,13-epoxy labd-14-ene hydroxylase

Next, we probed the activity of the selected P450 candidates in converting the reaction products of MvELS. The full-length sequences of Mv1270, Mv3392, Mv4213, Mv1545, and Mv6504 were amplified from cDNA prepared from M. vulgaris leaf RNA, and cloned into the pLIFE expression vector for use in established Agrobacterium-mediated transient co-expression assays in Nicotiana benthamiana [16, 23]. P450 activity was then tested via co-expression of each P450 candidate with MvCPS1 and MvELS. As compared to compounds 1–3 produced by the sequential activity of MvCPS1 and MvELS alone, no additional diterpenoid products were detected when co-expressing Mv1545, Mv4213, Mv6504 or Mv3392 (Additional file 4: Figure S2). Co-expression of the CYP71 family member Mv1270 (hereafter designated CYP71AU87) with MvCPS1 and MvELS resulted in the formation of two products (compounds 4 and 5) with retention times of 14.12 min and 14.19 min, respectively, indicating two closely related structures of higher polarity than the diTPS products (Fig. 4). Both products showed near identical fragmentation patterns that featured the dominant m/z 151 mass ion characteristic for the premarrubiin scaffold, as well as an additional mass ion of m/z 306 indicative of the presence of an additional oxygen atom (Fig. 4). Although only low product quantities were observed, CYP71AU87-catalyzed formation of compounds 4 and 5 was accompanied by a near complete turnover of labda-13(16),14-dien-9-ol 2 and a 3-fold decrease in the abundance of 9,13-epoxy labd-14-ene 1 as compared to the activity of MvCPS1 and MvELS only (Fig. 4). This reduction in substrate abundance was largely consistent with the formation of compounds 4 and 5 up to 4 μg g⁻¹ DW in transfected N. benthamiana leaves (Fig. 4).

To determine the structure of the CYP71AU87 products, ~1 mg of compounds 4 and 5 were extracted from 150 g of agroinfiltrated N. benthamiana leaf tissue, followed by purification using silica chromatography and semi-preparative HPLC analysis (Fig. 4). Although a separation of both compounds could not be achieved, the mixture of products 4 and 5 was successfully isolated with >90% purity. Combining a suite of 1D and 2D NMR analyses identified the products as 9,13-epoxy labd-14-ene-18-ol and 9,13-epoxy labd-14-ene-19-ol (Fig. 4 and Additional file 5: Figure S3). Importantly, chemical shifts of C-9 (13C, 92.8 ppm) and C-13 (13C, 83.7) showed the presence of the 9,13-epoxy group rather than a free C-9 hydroxyl function in both CYP71AU87 products (Additional file 5: Figure S3). In addition, the NMR analysis demonstrated that
hydroxylation of the CYP71AU87 occurred at C-18 or C-19, whereas presence of the methyl groups at C-20, C-17 and C-16 could be verified by means of HSQC and HMBC correlations (Additional file 5: Figure S3). Additional HMBC correlation analysis between C-18 and C-19 supported the identity of compounds 4 and 5 as an isomeric pair featuring hydroxyl groups at C-18 or C-19 (Additional file 5: Figure S3). A definitive assignment of which structure represents compounds 4 and 5 was not possible on the basis of these analyses, 9,13-epoxy labd-14-ene-18-ol and 9,13-epoxy labd-14-ene-19-ol were tentatively assigned as compounds 4 and 5, respectively. Additional transient N. benthamiana co-expression assays of MvCPS1, MvELS and CYP71AU87 with Mv3392, Mv4213, Mv1545 or Mv6504 did not result in any new reaction products that would indicate the ability of these P450 candidates in utilizing compounds 4 and 5 rather than the MvELS products as substrates (Additional file 4: Figure S2).

To further validate the activity of CYP71AU87 in N. benthamiana co-expression assays, we next performed co-expression of MvCPS1, MvELS, a M. vulgaris cytochrome P450 reductase (MvCPR) and CYP71AU87 in yeast (Saccharomyces cerevisiae). For this purpose, the constructs pESC-HIS:MvCPS/ScBTS1, pESC-Trp:MvELS/ScCPR and pESC-Ura:CYP71AU87 were generated to enable the co-expression of MvCPS1, MvELS, MvCPR, CYP71AU87 and the endogenous yeast GGPP synthase BTS1 in the yeast strain AM94 that is engineered for enhanced diterpenoid production [41]. GC-MS analysis of the resulting enzyme products confirmed the production of compounds 4 and 5, although with relatively low product abundance (Additional file 6: Figure S4). Notably, yeast expression of MvCPS1 alone or in combination with MvELS resulted in the formation of two additional compounds, 6 and 7, which represent the MvCPS1 product peregrinol 6 (i.e. dephosphorylated peregrinol diphosphate) [16] and likely a degradation product thereof 7, indicating a lower catalytic activity of MvCPS1, MvELS and/or CYP71AU87 as compared to co-expression assays in N. benthamiana. This observation is consistent with the presence of unconverted GGPP 8 substrate in the metabolite extract.

CYP71AU87 exhibits substrate specificity for 9,13-epoxy labd-14-ene and labda-13(16),14-dien-9-ol Our earlier studies demonstrated that MvELS is a multi-functional dTTPS that not only produces 9,13-epoxy labd-14-ene and labda-13(16),14-dien-9-ol but also produces additional compounds that are not substrates of MvCPR. The results presented here suggest that CYP71AU87 is able to oxidize 9,13-epoxy labd-14-ene and labda-13(16),14-dien-9-ol at different positions, with substrate specificity for 9,13-epoxy labd-14-ene and labda-13(16),14-dien-9-ol.
labd-14-ene as a major product, but also accepts the class II diTPS products (+)-copalyl diphosphate (CPP) and labda-13-en-8-ol diphosphate (LPP) to form miltiradiene and manoyl oxide, respectively [16]. While no LPP synthase has thus far been identified in this species, M. vulgare does contain a functional (+)-CPP synthase (MvCPS3) [16]. Therefore, we probed the possible substrate promiscuity of CYP71AU87 for different labdane diterpenoid intermediates. This was achieved by taking advantage of the natural modular nature of labdane diterpenoid biosynthesis, where diTPSs and P450s from different species can be combined to form active pathways toward different products [10, 42, 43]. Transient N. benthamiana co-expression assays were performed by combining CYP71AU87 and MvELS with different class II diTPSs, including a LPP synthase from Grindelia robusta (GrLPPS) [23] to produce manoyl oxide (compound 10), and a (+)-CPP synthase from Isodon rubescens (IrTPS3) [44] to form miltiradiene (compound 11) (Fig. 5). In addition, transient N. benthamiana co-expression assays with maize (Zea mays) ent-CPP synthase ZmAN2 [45] and the G. robusta ent-kaurene synthase GrEKS [23] were used to test the ability of CYP71AU87 to convert the gibberellin precursor ent-kaurene (compound 9). Co-expression of CYP71AU87 with MvCPS1 and MvELS served as a positive control. Co-expression of CYP71AU87 and the diTPS combinations MvELS/GrLPPS, MvELS/IrTPS3 or ZmAN2/GrEKS resulted in the expected biosynthesis of manoyl oxide, miltiradiene and ent-kaurene, respectively (Fig. 5 and Additional file 7: Figure S5). However, no additional products indicative of a CYP71AU87-catalyzed conversion of these diterpenoids were detected.

**Analysis of CYP71AU87 products in planta**

To investigate if the CYP71AU87 products 9,13-epoxy labd-14-ene-18-ol 4 and 9,13-epoxy labd-14-ene-19-ol 5 are present in planta, metabolite extracts prepared from M. vulgare leaves and flowers were analyzed via GC-MS, since these tissues had been shown to be most abundant in premarrubiin, marrubiin and related diterpenoids [16]. This approach identified marrubiin, premarrubiin, 9,13-epoxy labd-14-ene 1 and labda-13(16),14-dien-9-ol 2 as well as the unidentified MvELS product 3 in both tissues, whereas neither of the P450 products were detectable (Fig. 6). To verify the results obtained by GC-MS analysis, metabolite extracts were further analyzed using UPLC-Q Exactive MS analysis (Additional file 8: Figure S6). While the purified CYP71AU87 products could be...
detected as ammonium adducts with a parental ion of $m/z$ 324.2893 (monoisotopic mass $[M] + m/z$ 306.2559) using this approach, neither 9,13-epoxy labd-14-ene-18-ol nor 9,13-epoxy labd-14-ene-19-ol were detectable, thus confirming the results derived from GC-MS analysis (Additional file 8: Figure S6).

**Discussion**

Marrubiin and related diterpenoids of *M. vulgare* are of potential medicinal value for their anti-diabetic and vasorelaxant properties [13, 15]. Further drug development relies on the availability of viable means for producing these metabolites. Isolation of bioactive compounds,
including marrubiin, from plant biomass is often insufficient for meeting this need, due to the typically complex chemical profile of the natural producers, limited plant cultivation, and necessary protection of endemic wild species. Although multi-step synthetic routes for marrubiin and select derivatives thereof have been reported, the inherent stereochemical complexity of terpenoid natural products also hampers chemical synthesis approaches [46, 47]. With rapid advances in gene and enzyme discovery, reconstruction of pathways utilizing genes from the same or even different species has become an alternative strategy for plant natural product manufacture as exemplified by the microbial manufacture of the anti-malaria drug artemisinin [48]. Toward realizing such metabolic engineering approaches for marrubiin and related furanoterpenoid lactones, knowledge of the biosynthesis of the characteristic furan and lactone groups is of particular interest, since these functional modifications likely contribute to the therapeutic activities, as has been shown for the psychotropic compound salvinorin A, the anti-inflammatory diterpenoid andrographolide, and the antimalarial agent artemisinin [49–52].

The presented study illustrates the utility of combining genomics-enabled gene discovery and multi-gene co-expression analyses to identify previously hidden functions among the diverse P450 superfamily. Characterization of CYP71AU87 adds an additional catalyst to the group of terpenoid-metabolic members of the CYP71 family. In a sequential reaction with the diTPS pair, MvCPS1 and MvELS, CYP71AU87 forms the isomeric products 9,13-epoxy labd-14-ene-18-ol and 9,13-epoxy labd-14-ene-19-ol, as verified by GC-MS and NMR analysis (Fig. 4). Oxidation at C-18 and C-19 distinguishes CYP71AU87 from other members of the CYP71 family that function in labdane diterpenoid metabolism, including CYP71BE52 from R. officinalis and Salvia pomifera that catalyze the hydroxylation of ferruginol at C-2 [26, 31], rice CYP71Z6 and CYP71Z7 that facilitate C-2 oxygenation of ent-isokaurene and ent-cassadiene in oryzalide and phytocassane biosynthesis [53], and CYP71Z16/18 that controls the sequential oxygenation of dolabradiene at C-16 and C-3 to form dolabralexins in maize [54]. However, rice CYP99A2/A3, members of the CYP99 sub-family that belongs to the larger CYP71 family, have been shown to catalyze the sequential oxygenation at C-19 of syn-pimaradiene [55, 56]. In addition, oxygenation at C-18 and C-19 of labdane diterpenes has been shown for other P450 families, such as members of the CYP701A family relevant for the oxidation of ent-kaurene in gibberellin biosynthesis [35] and the gymnosperm-specific CYP720 family of diterpene resin acids metabolism [57, 58]. However, unlike the CYP71AU87-catalyzed hydroxylation, these enzymes facilitate a step-wise oxidation of their respective substrates to form the corresponding acids. Thus, P450-catalyzed oxidation of the labdane scaffold at C-18 or C-19 likely evolved independently several times with further species-specific diversification of catalytic specificity.

Co-expression assays in N. benthamiana and yeast showed that, while CYP71AU87 is specific to converting the C-9-oxygenated MvELS products within the scope of substrates tested here (Fig. 5), the enzyme shows promiscuity toward hydroxylation the C-18 or C-19 position (Fig. 4). As these products were not detected in planta (Fig. 6), it is not discernable if formation of this isomeric product pair is the native function of CYP71AU87. Marrubiin and all described structurally related diterpenoids are functionalized at C-19, supporting a regio-specific biosynthetic reaction at this position [15]. Notably, a similar catalytic pattern has been described for CYP720B1 from Pinus taeda that forms C-18- and C-19-hydroxylated isomers of miltiradiene when co-expressed in S. cerevisiae [59], whereas the close homolog CYP720B4 from Sitka spruce (Picea sitchensis) yields the regio-selective product oxygenated at C-18 [57]. C-19-oxidized diterpene acids, such as communic acids and isocupressic acid in species of the genus Juniperus and other coniferous trees [60, 61], have also been observed, but the underlying metabolic enzymes are unknown. Against this background, it remains to be clarified whether CYP71AU87 forms both, the C-18- and C-19-hydroxylated products, in M. vulgare or if more efficient metabolite channeling in planta may result in the formation of marrubiin from only the C-19 isomer.

In a previous study, biochemical characterization of MvELS illustrated three products, two of which represented 9,13-epoxy labd-14-ene-18-ol and 9,13-epoxy labd-14-ene-19-ol, as verified by GC-MS and NMR analysis (Fig. 4). Oxidation at C-18 and C-19 distinguishes CYP71AU87 from other members of the CYP71 family that function in labdane diterpenoid metabolism, including CYP71BE52 from R. officinalis and Salvia pomifera that catalyze the hydroxylation of ferruginol at C-2 [26, 31], rice CYP71Z6 and CYP71Z7 that facilitate C-2 oxygenation of ent-isokaurene and ent-cassadiene in oryzalide and phytocassane biosynthesis [53], and CYP71Z16/18 that controls the sequential oxygenation of dolabradiene at C-16 and C-3 to form dolabralexins in maize [54]. However, rice CYP99A2/A3, members of the CYP99 sub-family that belongs to the larger CYP71 family, have been shown to catalyze the sequential oxygenation at C-19 of syn-pimaradiene [55, 56]. In addition, oxygenation at C-18 and C-19 of labdane diterpenes has been shown for other P450 families, such as members of the CYP701A family relevant for the oxidation of ent-kaurene in gibberellin biosynthesis [35] and the gymnosperm-specific CYP720 family of diterpene resin acids metabolism [57, 58]. However, unlike the CYP71AU87-catalyzed hydroxylation, these enzymes facilitate a step-wise oxidation of their respective substrates to form the corresponding acids. Thus, P450-catalyzed oxidation of the labdane scaffold at C-18 or C-19 likely evolved independently several times with further species-specific diversification of catalytic specificity.

Co-expression assays in N. benthamiana and yeast showed that, while CYP71AU87 is specific to converting the C-9-oxygenated MvELS products within the scope of substrates tested here (Fig. 5), the enzyme shows promiscuity toward hydroxylation the C-18 or C-19 position (Fig. 4). As these products were not detected in planta (Fig. 6), it is not discernable if formation of this isomeric product pair is the native function of CYP71AU87. Marrubiin and all described structurally related diterpenoids are functionalized at C-19, supporting a regio-specific biosynthetic reaction at this position [15]. Notably, a similar catalytic pattern has been described for CYP720B1 from Pinus taeda that forms C-18- and C-19-hydroxylated isomers of miltiradiene when co-expressed in S. cerevisiae [59], whereas the close homolog CYP720B4 from Sitka spruce (Picea sitchensis) yields the regio-selective product oxygenated at C-18 [57]. C-19-oxidized diterpene acids, such as communic acids and isocupressic acid in species of the genus Juniperus and other coniferous trees [60, 61], have also been observed, but the underlying metabolic enzymes are unknown. Against this background, it remains to be clarified whether CYP71AU87 forms both, the C-18- and C-19-hydroxylated products, in M. vulgare or if more efficient metabolite channeling in planta may result in the formation of marrubiin from only the C-19 isomer.

In a previous study, biochemical characterization of MvELS illustrated three products, two of which represented 9,13-epoxy labd-14-ene-18-ol and 9,13-epoxy labd-14-ene-19-ol, as verified by GC-MS and NMR analysis (Fig. 4). Oxidation at C-18 and C-19 distinguishes CYP71AU87 from other members of the CYP71 family that function in labdane diterpenoid metabolism, including CYP71BE52 from R. officinalis and Salvia pomifera that catalyze the hydroxylation of ferruginol at C-2 [26, 31], rice CYP71Z6 and CYP71Z7 that facilitate C-2 oxygenation of ent-isokaurene and ent-cassadiene in oryzalide and phytocassane biosynthesis [53], and CYP71Z16/18 that controls the sequential oxygenation of dolabradiene at C-16 and C-3 to form dolabralexins in maize [54]. However, rice CYP99A2/A3, members of the CYP99 sub-family that belongs to the larger CYP71 family, have been shown to catalyze the sequential oxygenation at C-19 of syn-pimaradiene [55, 56]. In addition, oxygenation at C-18 and C-19 of labdane diterpenes has been shown for other P450 families, such as members of the CYP701A family relevant for the oxidation of ent-kaurene in gibberellin biosynthesis [35] and the gymnosperm-specific CYP720 family of diterpene resin acids metabolism [57, 58]. However, unlike the CYP71AU87-catalyzed hydroxylation, these enzymes facilitate a step-wise oxidation of their respective substrates to form the corresponding acids. Thus, P450-catalyzed oxidation of the labdane scaffold at C-18 or C-19 likely evolved independently several times with further species-specific diversification of catalytic specificity.
in extracts from *M. vulgare* (Fig. 6) and similar metabolite profiles observed in related species [19].

Although direct evidence that C-18/19 hydroxylation of 9,13-epoxy labd-14-ene is a native function of CYP71AU87 in *M. vulgare* would require further genetic studies, co-occurrence of the *MvCPS1*, *MvELS* and CYP71AU87 transcripts in tissues where premarrubiin and the *MvELS* products are also present (Fig. 6) suggest that CYP71AU87 would readily encounter 9,13-epoxy labd-14-ene as a substrate. Substrate specificity for the MvELS products further support a role of CYP71AU87 in the biosynthesis of marrubiin or related diterpenoids (Fig. 5). Absence of the CYP71AU87 products in planta could be attributed to a rapid conversion of these functionalized diterpenoids in *M. vulgare*, as has similarly been proposed for other P450s for which enzyme products were not detectable in plant tissues, such as *Thapsia garganica* CYP76AE2 with a probable function in thapsigin biosynthesis [63] and tomato CYP71BN1 that catalyzes the formation of the diterpenoid lycosan-talonol [64].

CYP71AU87-catalyzed formation of especially 9,13-epoxy labd-14-ene-19-ol then represents a possible reaction step in the biosynthesis of the 19,6-lactone ring structure characteristic of premarrubiin and other diterpenoids present in the genus *Marrubium* (Fig. 1). Presence of marrubenol and premarrubenol, bioactive diterpenoids featuring hydroxyl groups at both C-19 and C-6b, in several species of *Marrubium* supports a sequential hydroxylation reaction toward lactone formation [15, 65, 66] (Fig. 1). Conceptually similar pathways have been described for the biosynthesis of costunolide sesquiterpenoids via sequential oxidation of C-12 and hydroxylation at C-6 of germacrene A facilitated by a pair of P450 enzymes [20, 67]. Alternatively, dehydrogenases or reductases could function in further oxidizing the C-19 hydroxy group in 9,13-epoxy labd-14-ene-19-ol and, combined with P450-catalyzed oxygenation at C-6b, result in lactone ring formation as similarly proposed for the rice diterpenoid mumilactone A and the sesquiterpenoid arte-misinin [33, 68, 69]. Formation of premarrubiin and related structures further requires formation of a furan ring via additional P450-enabled oxidation C-15 and/or C-16 and subsequent ring closure (Fig. 1). Although the sequence of lactone and furan ring formation in marrubiin biosynthesis is unresolved, a P450, CYP76BK1 from *V. agnus-castus*, that catalyzes the C-16-hydroxylation of peregrinol (a dephosphorylated derivative of peregrinol diphosphate) toward the predicted vitexilactal precursor labd-13Z-ene-9,15,16-triol has indeed recently been identified [18].

**Conclusions**

The functional characterization of CYP71AU87 exemplifies the utility of combining gene-specific query of transcriptome data with multi-enzyme co-expression assays to identify natural product pathway genes in non-model plant systems. CYP71AU87 adds an additional catalyst to the catalog of diterpenoid-biosynthetic P450s and provides enzyme resources for producing marrubiin and related bioactive diterpenoid lactones.

**Methods**

**Plant material**

Seeds of *Marrubium vulgare* were purchased from Baker Creek Heirloom Seeds. Seeds of *M. vulgare* as well as *Nicotiana benthamiana* were germinated in Conviron TCR120 growth chambers (www.conviron.com) under a photoperiod of 16 h, 60% relative humidity, 100 μmol m⁻² s⁻¹ light intensity, and a day/night temperature cycle of 21/18 °C. After two weeks, *M. vulgare* plants were maintained in university greenhouses for an additional 12 weeks unless otherwise stated.

**Phylogenetic and sequence analysis**

Protein sequence alignments were generated using clustalW2 and curated with Gblocks [70]. Maximum likelihood phylogenetic analyses were performed using PhyML-aBayes version 3.0.1 beta with four rate substitution categories, LG substitution model, BIONJ starting tree and 1000 bootstrap repetitions [71]. Phylogenies were visualized using Interactive Tree of Life (iTOL) v4.3 [72].

**Mapping of transcriptome data to candidate P450 genes**

Leaf-specific transcriptome resources for *M. vulgare* were described earlier [23]. P450 candidates were identified by querying this transcriptome inventory against a custom P450 protein database (Additional file 1: Data file S1). Relative transcript abundance of P450 candidates was assessed by mapping adapter-trimmed Illumina reads back to the assembled P450 transcripts using BWA version 0.5.9-r16. Reads were mapped as paired with maximum insert size of 350 bp with a threshold of one allowed mismatch at the alignment step.

**Gene cloning**

Total leaf RNA of 12-week old *M. vulgare* plants was converted to cDNA using the SuperScript III First-Strand Synthesis System (ThermoFisher). The *CYP71AU87* gene then was amplified using Phusion-HF polymerase (New England Biolabs) and gene-specific oligonucleotides (Additional file 9: Table S2) and ligated into the pJET vector (Clontech) for sequence verification. The resulting amplicon was inserted into the pLIFE and pESC:Trp vector for expression in *S. cerevisiae*. The *MvCPR* gene was cloned in the same manner and ultimately inserted into the pESC:Ura vectors for expression in *N. benthamiana* and *S. cerevisiae*, respectively. The *MvCPS1* gene was cloned in the same manner and ultimately inserted into the pESC:Trp vector for expression in *S. cerevisiae*. *MvCPS1*
was cloned into the first multiple cloning site of the pESC:His vector containing the yeast GGPP synthase BTS1. MvELS was cloned into the first multiple cloning site of a pESC:Trp vector already containing MvCPR in the second multiple cloning site.

**GC-MS analysis of terpenoid metabolites**

Extracts were prepared from ~150 mg fresh tissue from leaves, stems, flowers and roots of 12-week old *M. vulgar* plants. Tissues were ground to a fine powder in liquid N$_2$ and terpenoids extracted with 1.5 ml hexane for 16 h under vigorous shaking at room temperature. Extracts were freed from residual water by addition of anhydrous Na$_2$SO$_4$, dried under N$_2$ stream, and re-dissolved in hexane for further analysis. GC-MS analysis was performed on an Agilent 7890B GC interfaced with a 5977 Extractor XL MS Detector at 70 eV and 1.2 ml min$^{-1}$ He flow, using a HP5-ms column (30 m, 250 μm i.d., 0.25 μm film) and the following GC parameters: 50°C for 1 min, 20°C min$^{-1}$ to 300°C, hold 3 min with pulsed splitless injection at 250°C and 50°C oven temperature. MS data from 90 to 600 mass-to-charge ratio (m/z) were collected after a 4 min solvent delay. Identification of marrubiin was achieved by comparison to the authentic standard (www.chromadex.com).

**LC-MS analysis of terpenoid metabolites**

For LC-MS analysis, dried samples were re-dissolved in 300 μl acetonitrile:water (80:20; v/v). All measurements were carried out on a Q Exactive HF mass spectrometer coupled with a Vanquish LC system (Thermo Scientific). A sample volume of 5 μl were separated on a Waters Acquity UPLC CSH C$_{18}$ column (100 × 2.1 mm; 1.7 μm) coupled to an Acquity UPLC CSH C$_{18}$ VanGuard pre-column (5 × 2.1 mm; 1.7 μm). The column was maintained at 65°C at a flow rate of 0.6 ml min$^{-1}$. The mobile phases consisted of A: acetonitrile:water (60:40, v/v) with ammonium formate (10 mM) and formic acid (0.1%) and B: 2-propanol:acetonitrile (90:10, v/v) with ammonium formate (10 mM) and formic acid (0.1%). The 15 min separation was conducted under the following gradient: 0 min 15% B; 0–2 min 30% B; 2–2.5 min 48% B; 2.5–11 min 82% B; 11–11.5 min 99% B; 11.5–12 min 99% B; 12–12.1 min 15% B; 12.1–15 min 15% B. Orbitrap MS instrument was operated in electrospray ionization (ESI) in positive mode with the following parameters: mass range 60–900 m/z, spray voltage 3.6 kV, sheath gas (nitrogen) flow rate 60 units; auxiliary gas (nitrogen) flow rate 25 units, capillary temperature 320°C, full scan MS1 mass resolving power 120,000, data-dependent MSMS (dd-MSMS) 4 scans per cycle, dd-MSMS mass resolving power 30,000. Thermo Xcalibur 4.0.27.19 was used for data acquisition and analysis [73].

**Transient expression in Nicotiana benthamiana**

Full-length cDNA clones of *MvCPS1, MvELS and CYP71AU87* in the pLIFE expression vector were individually transformed into *A. tumefaciens* strain GV3101. Bacterial cultures were grown at 28°C in Luria-Bertani (LB) medium supplemented with 50 mg l$^{-1}$ kanamycin, pelleted and resuspended to a final OD$_{600}$ of 1.0 in 10 mM MES buffer with 10 mM MgCl$_2$. Following incubation for two hours at 22°C and gentle shaking, cell suspensions were mixed and syringe-infiltrated into the underside of the leaves of 5-week-old *N. benthamiana* plants. For all assays, target genes were further co-expressed with the plant viral protein p19 to suppress RNA silencing. [23] Transfected plants were maintained for four days before metabolites were extracted with 1.5 ml hexane from a single transformed leaf and analyzed via GC/LC-MS as described above. Quantification of metabolites via GC-MS was performed using an external standard curve using the structurally related diterpenoid sclareol as a standard. Statistically significant differences between metabolite levels across tobacco co-expression assays was calculated based on Welch Two Sample $t$-test ($P$-value < 0.05). Quantification of metabolites via LC-MS in tobacco co-expression assays was calculated using sclareol as internal standard at a concentration of 1 μg ml$^{-1}$ using the Thermo Xcalibur 4.0.27.19 software.

**Enzyme co-expression in engineered yeast**

The generated constructs pESC-HIS:MvCPS/ScBTS1, pESC-Trp:MvELS/MvCPR and pESC-Ura:CYP71AU87 were co-transformed *Saccharomyces cerevisiae* strain AM94 that was specifically engineered for diterpenoid production [59, 74]. Cells were grown in 50 ml of the corresponding selective dropout medium (~His, –Trp, –Leu, –Ura, 2% dextrose) at 30°C to an OD$_{600}$ of -0.6, followed by transfer of cells into 50 ml of corresponding selective dropout media with 2% galactose for induction. After 48 h, metabolites were extracted by vortexing the cell pellets with glass beads in 5 ml of diethyl ether, air-dried and re-suspended in 1 ml hexane for GC-MS analysis.

**Nuclear magnetic resonance (NMR) analysis of the CYP71AU87 product**

Approximately 1 mg of the CYP71AU87 product was extracted from ~50 N. benthamiana leaves transfected with MvCPS1, MvELS and CYP71AU87 using 400 ml hexane. Extracts were air-dried and purified through iterative chromatography on silica matrix (230–400 Mesh, Grade 60) using an ethyl acetate-hexane gradient as mobile phase (100% hexane, 10% ethyl acetate, and 20% ethyl acetate; v/v). Product-containing fractions were pooled, dried under N$_2$ stream, and resuspended in 100% acetonitrile. Further purification was achieved by
reverse-phase HPLC on an Agilent 1100 series HPLC equipped with an Agilent Zorbax Eclipse Plus C8 column, G1315B diode array detector (DAD) and G1364C-1260 FC-AS fraction collector. The mobile phases consisted of a mixture of A: water and B: acetonitrile. Product separation was conducted using the following gradient: 0 min 50% B; 0–7 min 50% B; 7–10 min 75% B; 10–20 min 90% B; 20–40 min 100% B. Purified products were pooled and resuspended in chloroform-D spiked with 0.03% Tetramethylsilane (TMS) as internal standard (Sigma). For structural identification of the CYP71AU87 products 1D (1H, 13C, NOE) and 2D (COSY, HSQC, HMBC, selective HSQC, selective HMBC) spectra were collected on Bruker Avance 800 MHz spectrometer equipped with a 5 mm CPTCI cryoprobe.

Quantitative PCR (qPCR)
Total RNA was isolated as previously described using approximately 100 mg tissue [75]. RNA integrity and concentration was measured using Bioanalyzer 2100 RNA Nano chip assays (Agilent) following the manufacturer’s protocol. Equal RNA amounts were used for cDNA synthesis with the Superscript III reverse transcriptase (ThermoFisher) and oligo(dT) primers. The subsequent qPCR reaction was performed on a Bio-Rad CFX96 Real-time system using the SsoFast kit (www.bio-rad.com) and target-specific oligonucleotides (Additional file 9: Table S2). Relative transcript abundance was calculated using efficiency corrected ACT using the amplification efficiency values (E-values) generated from primer efficiency calculations for each gene pair. Relative gene expression values were calculated based on M. vulgaris Elongation Factor 1α as reference gene and triplicate measurements with three biological replicates. Target specificity was confirmed by sequence verification of representative amplicons.

Additional files

---

### Additional file 1: Data file 1
P450 database used in this study. (DOCX 164 kb)

### Additional file 2: Table S1
List of transcript sequences of P450 candidate genes identified in the established Marrubium vulgaris leaf transcriptome. (XLSX 79 kb)

### Additional file 3: Figure S1
NMR analysis of labda-13(16),14-dien-9-ol (compound 2) formed by the coupled reaction of McVPS1 and McELS. (PDF 910 kb)

### Additional file 4: Figure S2
GC-MS analysis of reaction products derived from Nicotiana benthamiana co-expression assays of McVPS1 and McELS the P450 candidates McV170, McV1545, McV4213, McV6504, and McV3392, respectively. Shown are extracted ion chromatograms (m/z 151) and mass spectra of the reaction products resulting from co-expression of McVPS1, McELS, CYP71AU87, McCR1, and the endogenous yeast GGPP synthase B51 in the yeast strain AM94. (PDF 191 kb)

### Additional file 5: Figure S3
NMR analysis of 9,13-epoxy-labd-14-ene-19-ol (compound 4/5) formed by the coupled reaction of McVPS1, McELS and CYP71AU87. (PDF 111 kb)

### Additional file 6: Figure S4
Co-expression of McVPS1, McELS and CYP71AU87 in yeast (Saccharomyces cerevisiae). Shown are extracted ion chromatograms (m/z 151) and mass spectra of the reaction products resulting from co-expression of McVPS1, McELS, CYP71AU87, McCR1 and the endogenous yeast GGPP synthase B51 in the yeast strain AM94. (PDF 141 kb)

### Additional file 7: Figure S5
CG-MS analysis of reaction products derived from diterpenoids obtained in N. benthamiana co-expression assays to test CYP71AU87 substrate promiscuity (see Fig. 6). (PDF 205 kb)

### Additional file 8: Figure S6
LC-MS analysis of CYP71AU87 products in M. vulgaris leaves and flowers. (A) Leaf and flower tissue from 12-week-old M. vulgaris plants was extracted with 20% (v/v) ethyl acetate in hexane, and then dried and re-suspended in MeOH. Extracts were analyzed by LC-MS on a Q Exactive HF mass spectrometer coupled with a Vanquish LC system (Thermo Scientific). A volume of 5 µl was separated on a Waters Acquity UPLC CSH C18 column (100 × 2.1 mm; 1.7 µm) coupled to an Acquity UPLC CSH VanGuard precolumn (5 × 2.1 mm; 1.7 µm) at 65 °C and a flow rate of 0.6 ml min⁻¹. (PDF 322 kb)

### Additional file 9: Table S2
Oligonucleotides used in this study. (PDF 111 kb)

---

**Abbreviations**

CPP: Copalyl diphosphate; DTPS: Diterpene synthase; GC-MS: Gas chromatography mass spectrometry; GGPP: Geranylgeranyl diphosphate; LC-MS: Liquid chromatography mass spectrometry; NMR: Nuclear magnetic resonance; P450: Cytochrome P450 monooxygenase

**Acknowledgements**

We would like to thank We thank Dr. David Nelson (University of Tennessee) for assistance with the annotation of CYP71AU87.

**Funding**

Financial support for the development and completion of this project was provided by startup funds through the University of California Davis (to P.Z.) and a UC Davis Academic Senate Faculty Research Grant (to P.Z.) providing experimental support. Publication costs were funded by startup funds through
the University of California Davis (to P.Z.), in addition, graduate student salary and tuition support was provided by a NIGMS predoctoral fellowship (T32 GM007377, to P.S.K.).

Availability of data and materials
I can confirm I have included a statement regarding data and material availability in the declaration section of my manuscript. Nucleotide sequences reported in this study were submitted to the National Center for Biotechnology Information (NCBI) GenBank/EMBL Database with accession numbers: CYP71AU87 (MH593587), MvCPR (MH593588), MV1545 (MH824164), MV2413 (MH824162), MV6504 (MH824163), MV3392 (MH824164).

Authors’ contributions
P.Z. conceived the original research and oversaw data analysis; P.S.K. performed most experiments; P.D. assisted with sub-cloning and enzyme characterization experiments; T.S. and O.F. performed LC-MS analyses; J.B.A assisted with NMR most experiments; Hambreger B, Hamberger B, Motawia MS, Olsson CE, et al. Total biosynthesis of the cyclic AMP booster forskolin from Coleus forskohlii eLife. 2017;6:e23001.

Nielsen MT, Randberg J, Christensen U, Christensen HB, Harrison SJ, Olsson CE, Hambreger B, Maller BL, Naftolin MH. Microbial biosynthesis of the Forskolin precursor Manoyl oxide in an Enantioselectively pure form. Appl Environ Microbiol. 2014;80(23):7258–65.

Popoola OK, Elbagory AM, Ameer F, Hussein AA. Marrubium. Molecules. 2013; 18(8):9069–60.

Minopoli N, Levendal RA, Molikkar N, Frost CL. Marrubium, a constituent of Leonurus leonurus, alleviates diabetic symptoms. Phytochemistry. 2012;89:488–93.

Meve-Silva C, Cechinel-Filho V. A review of the chemical and pharmacological aspects of the genus marrubium. Pharm Cur Dis. 2010;16(3):5303–18.

Zerbe P, Chang A, Dullat H, O’Neill-Johnson M, Starks C, Hamberger B, Jitter syntheses of the biosynthetic system of medicinally active diterpenoids in Marrubium vulgare. Plant J. 2014;79(6):914–27.

Krois W, Reuter O, Zopp J. Biosynthesis of the labdane diterpene marrubin in Marrubium vulgare via a non-mevalonate pathway. Biochem J. 1997;326(Pt 2):2449–54.

Heskes AM, Sundram TCM, Boughton BA, Jensen NB, Hansen NL, Crocrol C, Cozzi F, Rasmussen S, Hamberger B, Hamberger B, et al. Biosynthesis of bioactive diterpenoids in the medicinal plant Vitex agnus-castus. Plant J. 2018;93(5):949–58.

Xie DF, Ma DM, Judd R, Jones AL. Artemisinin biosynthesis in Artemisia annua and metabolic engineering: questions, challenges, and perspectives. Phytochemistry. 2016;115(6):1093–114.

Ikezawa N, Gopfert JC, Nguyen DT, Kim SU, O’Maille P, Spring O, Ro DK. Lettuce costunolide synthase (CYP71BL2) and its homolog (CYP71BL1) from sunflower catalyze distinct regiono- and stereoselective hydroxylations in sesquiterpene lactone metabolism. J Biol Chem. 2011;286(24):21601–11.

Munro TA, Rizzacasa MA, Roth BL, Toth BA, Yan F. Studies toward the pharmacophore of salvinorin a, a potent kappa opioid receptor agonist. J Med Chem. 2005;48(2):345–52.

Eckstein-Ludwig U, Webb RJ, Van Goethem ID, East JM, Lee AG, Kimura M, O’Neill PM, Bray PG, Ward SA, Krishna S. Artemisins target the SERCA of plasmodium falciparum. Nature. 2003;424(6951):957–61.

Zerbe P, Hamberger B, Yuen MW, Chiang A, Sandhu HK, Madlao LL, Nguyen A, Hamberger B, Bach SS, Bohmann J. Gene discovery of modular diterpene metabolism in nonmodel systems. Plant Physiol. 2013;162(2):1073–91.

Toch KH, Polchuk DR, Reed DW, Nowak G, Covello PS, Artemisia annua L. (Asteraceae) trichome-specific cDNAs reveal CYP71AV1, a cytochrome P450 with a key role in the biosynthesis of the antimalarial sesquiterpene lactone artemisinin. FEBS Lett. 2006;580(1):1411–6.

Guo J, Ma K, Cai Y, Ma Y, Zhan Z, Zhou YJ, Liu W, Guan M, Yang J, Cui G, et al. Cytochrome P450promiscuity leads to a bifurcating biosynthetic pathway for tanhydroines. New Phytol. 2016;201(2):525–34.

Ignea C, Athanasakoglou A, Ioannou E, Georgantza P, Trikka FA, Lopuassaki S, Roussis V, Makris AM, Kampuran C. Carnosic acid biosynthesis elucidated by a synthetic biology platform. Proc Natl Acad Sci U S A. 2016;113(13):3681–6.

Zerbe P, Hamberger B, Sank P. Plant P450s as versatile drivers for evolution of species-specific chemical diversity. Philos Trans R Soc Lond Ser B. 2013;368(1621):20120426.

Hambreger B, Bak S. Plant P450s as versatile drivers for evolution of species-specific chemical diversity. Philos Trans R Soc Lond Ser B. 2013;368(1621):20120426.

Zhao YJ, Cheng QQ, Su P, Chen X, Wang XJ, Gao W, Huang LQ. Research progress relating to the role of cytochrome P450 in the biosynthesis of terpenoids in medicinal plants. Appl Microbiol Biotechnol. 2014;108(6):2371–83.

Guo J, Zhou YJ, Hillwig ML, Shen Y, Yang L, Wang X, Zhang X, Liu W, Peters RJ, Chen X, et al. CYP76AH1 catalyzes tunneling of mi trifidane in tanhydroines biosynthesis and enables heterologous production of fergunidin in yeasts. Proc Natl Acad Sci U S A. 2013;110(29):12089–13.

Zi J, Peters RJ. Characterization of CYP76AH4 clarifies phenolic diterpene biosynthesis in the Lamiaceae. Org Biomol Chem. 2013;11(44):7650–2.

Božič D, Papainthimou D, Brücker K, de Vos RC, Tsoleridis CA, Katsarou D, Papainthimou A, Patrakis I, Chatzopoulou FM, Dimidiou E, et al. Towards elucidating Camoscic acid biosynthesis in Lamiaceae: functional characterization of the three first steps of the pathway. Org Biomol Chem. 2015;13(10):e0124106.

Wang Q, Hillwig ML, Okada K, Yamazaki K, Wu Y, Swaminathan S, Yamane H, Peters RJ. Characterization of CYP76MS5-A indicates metabolic plasticity within a plant biosynthetic gene cluster. J Biol Chem. 2012;287(9):6159–68.
