Molecular cloning, bacterial expression and functional characterisation of cytochrome P450 monooxygenase, CYP97C27, and NADPH-cytochrome P450 reductase, CPR I, from *Croton stellatopilosus* Ohba

Siriluk Sintupachee [a,b], Nattaya Ngamrojanavanich [c], Worapan Sitthithaworn [d], Wanchai De-Eknamkul [b,e,*]

[a] Program in Biotechnology, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand
[b] Research Unit for Natural Product Biotechnology, Faculty of Pharmaceutical Sciences, Chulalongkorn University, Bangkok 10330, Thailand
[c] Department of Chemistry, Faculty of Science, Chulalongkorn University Bangkok 10330, Thailand
[d] Department of Pharmacognosy, Faculty of Pharmacy, Srinakharinwirot University, Nakonnayok 26120, Thailand
[e] Department of Phamacognosy and Pharmaceutical Botany, Faculty of Pharmaceutical Sciences, Chulalongkorn University, Bangkok 10330, Thailand

ARTICLE INFO

ABSTRACT

The cDNAs for cytochrome P450 monooxygenase (designated as CYP97C27 by D. Nelson's group) and NADPH-cytochrome P450 reductase (designated as CPR I based on its classification) were isolated from *Croton stellatopilosus* leaves, which actively biosynthesise plaunotol (18-OH geranylgeraniol). CYP97C27 and CPR I contain open reading frames encoding proteins of 471 and 711 amino acids with predicted molecular masses of 53 and 79 kDa, respectively. By aligning the deduced sequences of CYP97C27 and CPR I with other plant species, all functional domains of CYP97C27 (heme and oxygen binding) and CPR I (CYP- and FMN, FAD, and NADPH cofactor binding) were identified. Amino acid sequence comparison indicated that both CYP97C27 (85–93%) and CPR I (79–83%) share high sequence identities with homologous proteins in other plant species, suggesting that CYP97C27 belongs to the CYP97 C subfamily and that CPR I belongs to class I of the dicotyledonous CPR. Functional characterisation of both enzymes, produced in *Escherichia coli* (pET32a/BL21(DE3)) as recombinant proteins, showed that simultaneous incubation of CYP97C27 and CPR I with the substrate geranylgeraniol (GGOH) and coenzyme NADPH led to formation of the product plaunotol. In *C. stellatopilosus*, the levels of the CYP97C27 and CPR I transcripts were highly correlated with those of several mRNAs involved in the plaunotol biosynthetic pathway, suggesting that CYP97C27 and CPR I are the enzymes that catalyse the last hydroxylation step of the pathway.

1. Introduction

Croton stellatopilosus Ohba (Euphorbiaceae) is a Thai medicinal plant first reported in 1978 to contain the anti-peptic plaunotol as a major constituent [1]. This plant's leaves contain 0.3–0.5% dry weight of plaunotol [2] and have been used as raw materials for manufacturing the anti-peptic ulcer drug Kelnac®. Plaunotol is an acyclic diterpene alcohol derived from four isoprene units that are biosynthesised exclusively via the 2C-methyl-D-erythritol 4-phosphate (MEP) pathway [3] (Fig. 1). Of the isoprenes, one molecule of dimethylallyl diphosphate (DMAPP) is attached to the other three molecules of isopentenyl diphosphate (IPP) by a head-to-tail condensation, which is catalysed by the enzyme geranylgeranyl diphosphate synthase (GGPPS), yielding geranylgeranyl diphosphate (GGPP) [4]. GGPP is subsequently converted to geranylgeraniol (GGOH) by a two-step monodephosphorylation [5] catalysed by a membrane-bound GGPP phosphatase [6,7]. In the final step, GGOH is presumably hydroxylated specifically at the C-18 position to form 18-GGOH or plaunotol, the oil-like substance that accumulates in the chloroplast [8]. Of these steps, the enzyme that catalyses the last reaction of
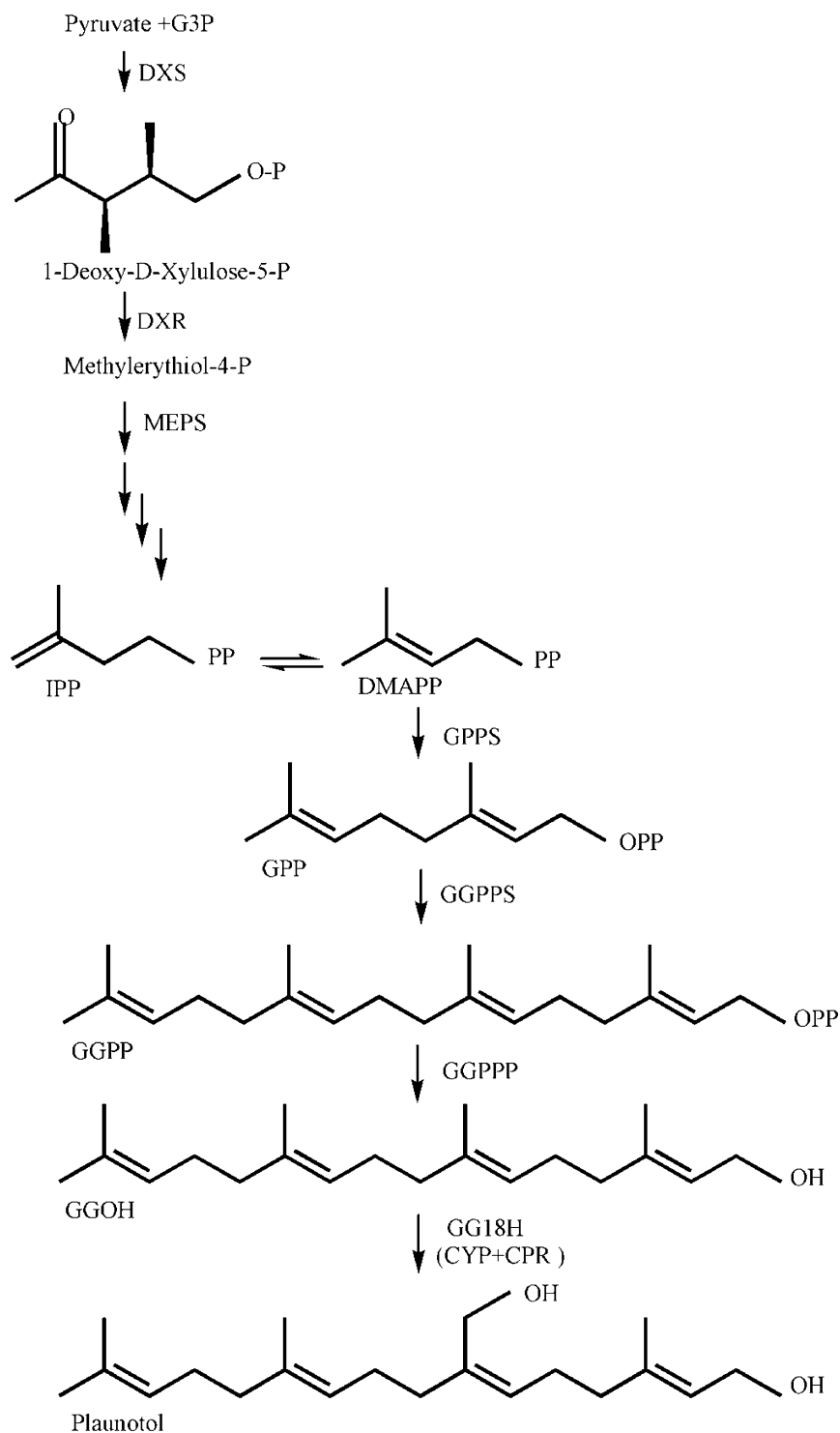
Fig. 1. Biosynthetic pathway of the acyclic diterpenoid plaunotol beginning with the 2C-methyl-D-erythritol 4-phosphate (MEP) pathway in *C. stellatopilosus*. DXS, 1-deoxy-D-xylulose 5-phosphate synthase; DXR, 1-deoxy-D-xylulose 5-phosphate reductase; MEPS, 2C-methyl-D-erythritol 4-phosphate synthase; GPPS, geranyl diphosphate synthase; GGPPS, geranylgeranyl diphosphate synthase; GGPPP, geranylgeranyl diphosphate phosphatase; GG18H, geranylgeraniol-18-hydroxylase.

GGOH-18-hydroxylation has not been identified in *C. stellatopi-losus*. Our previous studies indicated that although the enzyme activity of GGOH-18-hydroxylase was easily detected in the microsomal fraction in the presence of NADPH [9], purification of the enzyme to obtain either a membrane-bound or solu-bilised form for characterisation has not yet been achieved. However, partial characterisation of the microsomal fraction sug-gests that the hydroxylase enzyme is a plant cytochrome P450 [10].

Plant cytochrome P450s (CYPs) form a large superfamily of heme-containing monooxygenases that are metabolic intermedi-ate enzymes involved in both primary and secondary metabolism pathways. CYPs are heme-thiolate-proteins, which catalyse the majority of oxidative reactions, including hydroxylation, epoxida-tion, dealkylation, dehydration, and carbon–carbon bond cleavage [11]. CYPs are located in the endoplasmic reticulum (ER) and their catalytic activities rely strictly upon the supply of electrons from NADPH-cytochrome P450 reductases (CPRs) [12]. CPRs [EC 1.6.2.4]

are also membrane-bound proteins localised to the ER and contain an N-terminal positioned FMN binding domain linked to the NADPH binding domain via the FAD domain [13]. Therefore, during the reactions catalysed by the coordination of CYPs and CPRs, the CYPs catalyse a wide variety of different regio- and stereospecific reactions, and the CPRs catalyse the transfer of electrons from NADPH via FAD and FMN to the prosthetic heme group of the CYP proteins to complete the reaction. Through these CYP-catalysed reactions, the diverse oxygenation patterns of natural products generated by secondary metabolic pathways in plants and microorganisms are achieved [14].

Plant CYPs are grouped as either A-type or non-A-type based on both the available sequences in 11 phylogenetically distinct clans [15] and the phylogenetic trees [16]. The A-type CYPs constitute a monophyletic clade are currently the CYP71 clans whereas the more diverse non-A-type CYPs are the remaining 10 clans [15]. It was originally postulated that the A-type CYPs were involved in secondary metabolism, and the non-A-type CYPs were involved in primary metabolism [17,18]. However, this view is over simplified because there are several non-A-type CYPs that have been categorised in the secondary plant metabolism group [18]. Biochemically, plant CYPs are involved in a wide range of secondary metabolite biosynthetic reactions, which are grouped into three major classes, terpenoids, phenylpropanoids, and nitrogen-containing compounds, including alkaloids, cyanogenic glucosides, and glucosinolates [19].

For the genes encoding the plant CPRs, the number of CPR paralogues varies from one to three, depending on the species. A single CPR paralogue has been characterised from *Coleus blumei*, *Papaver somniferum*, *Taxus cuspidata*, and *Vigna radiata* [20–22], whereas two CPR paralogues have been characterised from *Helianthus tuberosus*, *Petroselinum crispum*, *Arabidopsis thaliana*, *Centaurium erythraea*, and *Gossipium hirsutum* [21,23–27], and three paralogues have been characterised from *Nothapodytes foetida*, *Populus trichocarpa*, and *Populus deltoides* [20–22,28]. Based on the N-terminal anchoring sequences, the CPRs are classified into two classes, class I (CPR I) and class II (CPR II) [29]. CPR I is expressed constitutively, whereas CPR II is transcribed under stress or wounding [30]. The presence of multiple CPRs in plants may reflect the diversity of CYPs [31].

To elucidate the entire biosynthetic pathway of plaunotol, we have identified the CPR and CYP genes potentially involved in the last GGOH-18-hydroxylation step. In this report, we describe the results of the cloning, expression, and characterisation of both genes from *C. stellatopilosus* Ohba. The obtained genes were submitted to the NCBI GenBank database with accession no. KF738254 for CYP97C27 and KF738256 for CPR I. In addition, the levels of CYP97C27 and CPR I transcripts in the shoots, stalks, and young leaves of *C. stellatopilosus* in relation to other mRNAs involved in the plaunotol biosynthetic pathway are reported.

2. Materials and methods

2.1. Plant collection

Fresh leaves from *C. stellatopilosus* Ohba were collected from plants growing in an open field at the Faculty of Pharmaceutical Sciences, Chulalongkorn University, Bangkok, Thailand, which is located at 13°44′33″ north latitude and 100°31′49″ east longitude. The voucher specimen of *C. stellatopilosus* (No. 184779) was deposited at the Office of the Forest Herbarium, Department of National Parks, Wildlife and Plant Conservation, Bangkok, Thailand. The leaves at the second position on the shoot were used for cDNA amplification and full-length gene determination. Portions of the shoot, leaves, and stalks from the same branch of the plant were collected to determine the transcript levels.

Table 1
Sequences of the primes used in this study.

Primer name	Sequence 5′→3′
Degenerate PCR	
CsCPR-de-F	TTTAR*A*RTTGTTGAT*W*TGGATGATTATGC
CsCPR-de-R	GCCAT*R*C*Y*TTGGCATCACCACA*S*ACATA
CsCYP-de-F	GT*R*TTCAACTA*Y*RA*Y*TTTGA
CsCYP-de-R	A*YS*AC*D*GG*D*GG*D*TG*W*GG*Y*TA
RACE-PCR	
CsCPR-3RACE	AACACCAGGAGGACGGCTACACAA
CsCPR-3RACEnested	GGTTGCAGGAACAGTAAGATGG
CsCPR-5RACE	CCTGCTCAGCAAGCAACTCATCCAACTC
CsCPR-5RACEnested	CTGTTGGCTCGCCATCTCCATATGT
CsCYP-3RACE	AAGTGATTGTGGAGGCCGAAGGTGAA
CsCYP-3RACEnested	AAGAGGCGGAGGCTCGTTCCACAGATCTT
CsCYP-5RACE	AACCAGTTGTCTCGTGGCCAGCAACTA
CsCYP-5RACEnested	CGAGCAAGAACCGAAGAATGCTTGGATCAG
GeneRacer™3′primer	GCTGTCAACGATACGCTACGTAACG
GeneRacer™3′nested primer	CGCTACGTAACGGCATGACAGGT
GeneRacer™5′primer	CGACTGGAGCACGAGGACACTGA
GeneRacer™5′nested primer	GGACACTGACATGGACTGAAGGAGTA
Full-length PCR	
CsCPR-ORF-F-*Bam*HI	GGATCCATGCAATCTTCGACTGGTGCAG
CsCPR-ORF-R-*Sal*I	GTCGACGTCGACTTAGAAATTGCCTGCTTC
CsCYP-ORF-F-*Not*I	GCGGCCGCATGGCAACCGCAAAACTGGATGAT
CsCYP-ORF-R-*Xho*I	CTCGAGCCTTAATGAAGATGAAGCAAATGCTTTA
Real-time PCR	
GGPPP-F	CCTCCATCATGCCATACTAGGTCTC
GGPPP-R	CAACATGGCCTCTGCGGTCAAATGC
GGPPS-F	GTCGTGGTAAGCCGACGAATCAT
GGPPS-R	TCCTTCAGCACCAATCGCCTTAG
DXS-F	TATCCCAGTAGCGCCACTCG
DXS-R	CCTCCTAACATGGTGGTGATGGCTCCTTC
MEPS-F	ATAATGCGCCGCAATGCACC
MEPS-R	ATAATGCGCCGCAATGCACC
CsCPR-RT-F	ATGCAACCATCATCATCGGCAGGTTGG
CsCPR-RT-R	GTGCCGTCCTCTTCATCAGTA
CsCYP-RT-F	CTTCGGTTCTTGCTCGCTAGTA
CsCYP-RT-R	GGCCTTCCTTGTAACACTTGGT
18s rRNA F	CAAAGCAAGCCTACGCTCTG
18s rRNA R	CGCTCCACCAACTAAGAACG

2.2. Total RNA extraction and cDNA amplification

The collected leaves were flash frozen with liquid nitrogen and ground with a pestle-mortar into powder. The total RNAs were extracted from the leaf powder (100 mg) using an RNeasy® Plant Mini kit (Qiagen, Valencia, CA) according to the manufacturer's instruction. The total RNA extracts were treated with a DNaseI reagent kit (Fermentus) and were amplified using oligo $(dT)_{18}$ with M-MLV reverse transcriptase. The reaction mixture consisted of 1 mM oligo$(dT)_{18}$ and 1 μg total RNAs. The reaction was initiated by heating at 72 °C to denature the double stranded RNAs. Next, 1× M-MLV buffer, 1 mM dNTPs mixture, and 200 units M-MLV reverse transcriptase were added. The reaction tube was incubated at 40 °C for 90 min before enzyme inactivation at 65 °C for 15 min. The primer sequences used for gene amplification are summarised in Table 1.

2.3. Molecular cloning of full-length cytochrome P450 and cytochrome P450 reductase

The A-type plant cytochrome P450 family clans were used to design primers for the core sequence of CYP97C27. Each forward primer contains the sequence of the oxygen-binding motif, and the reverse primer was designed to amplify the heme-binding motif. The primer set for the amplification of the core sequence of CPR I was designed based on the conserved sequence of plant CPRs. The primers used in this study are summarised in Table 1. The PCR reaction mixture consisted of 1× PCR buffer, 0.4 mM dNTP mixture, 2 mM MgCl$_2$, 0.4 mM each primer, 100 ng cDNA template, and 0.5

Unit Platinum® Taq DNA Polymerase High Fidelity (Invitrogen). PCR amplification (My Cycler™ Thermal Cycler, BioRad) was performed under the following conditions: 1 cycle at 94 °C for 2 min, followed by 35 cycles of 94 °C for 30 s, 55 °C for 30 s, and 68 °C for 30 s, with a final extension at 68 °C for 5 min. The PCR products were visualised on 1% agarose gels stained with 10 mg/mL ethidium bromide with UV-transluminescence (ChemiDoc XRS, BioRad). To obtain the 3′-end and 5′-end cDNA sequence, RACE-PCR was performed using the 3′-RACE and 5′-RACE primers listed in Table 1. The cDNA amplification reaction mixture consisted of 1× PCR buffer, 0.4 mM dNTP mixture, 2 mM MgCl₂, 0.4 mM each primer, 100 ng cDNA, and 0.5 Unit Platinum® Taq DNA Polymerase High Fidelity (Invitrogen). The PCR thermal profile was performed with two sequential steps. The first step of the PCR was 1 cycle at 94 °C for 2 min, followed by 5 cycles at 94 °C for 30 s and 72 °C for 1 min, followed by 5 cycles at 94 °C for 30 s and 70 °C 1 min, followed by 25 cycles at 94 °C for 30 s and 55 °C 1 min, and then 1 cycle at 72 °C for 1 min. The second nested step of the PCR was 1 cycle at 94 °C for 2 min, followed by 30 cycles at 94 °C for 30 s and 55 °C for 30 s, followed by 68 °C for 1 min, and then 68 °C for 5 min. The PCR products were confirmed using bidirectional sequencing of DNA freshly purified from the gel with the HiYield™ Gel/PCR DNA Fragment Extraction kit (RBC Bio-Science). The fragments were directly ligated into the pGEM-T easy vector (Promega).

2.4. Sequence analysis and phylogenetic relationships

The sequences and sequence chromatograms were obtained using the liquid polymer POP7 on an ABI730xl (96 capillary) and DNA analyser (Applied BioScience). Quality analysis was performed by trimming off the vector sequence using the Sequencer Scanner v1.1 program. The sequences were analysed, and the homology of the protein was identified using the blastx (translated nucleotide query blast) algorithm. The open reading frame sequence was predicted and translated in the CloneManager v9.2 program and then calculated for the monomeric molecular weights and theoretical pI using http://web.expasy.org/compute_pi/. The phylogenetic trees were constructed using homologue sequences with identity based on the blast algorithm. The sequences were collected and multiple alignments were performed in the clustalx program [32]. The phylogenetic tree was constructed with the MEGA6 software [33]. The neighbour-joining (NJ) tree was obtained using the algorithm to select the minimum evolution methods and evolution distance calculated for the substitution gene using the Jones-Taylors-Thomton (JTT) method with 1000 replication bootstrap value [34,35].

2.5. Expression of recombinant proteins in Escherichia coli

The genes encoding CYP97C27 and CPR I within pGEM-T Easy (Promega) vector constructs were subcloned with the appropriate restriction enzymes at the XhoI/NotI and BamHI/SalI (New England Biolabs) positions for the CYP97C27 and CPR I genes, respectively, into the pET32a expression plasmid (Novagen) and separately introduced into the BL21(DE3) bacterial stain (Novagen) using heat shock at 42 °C. The optimisation of the recombinant protein expression in the bacterial system was performed by monitoring the cell density. Overnight cultures of both pET32a-CYP97C27 and pET32a-CPR I were inoculated separately into 1 L LB broth containing 100 μg/mL ampicillin and grown at 37 °C with shaking at 250 rpm until the OD₆₀₀ reached 0.6. Expression of pET32a-CYP97C27 was induced in the recombinant culture with 1 mM IPTG and shaken continuously at 37 °C with 250 rpm, whereas expression of pET32a-CPR I was induced at 30 °C with 200 rpm shaking. A total of 1 mL of each culture was collected at various incubation times (1, 2, 6, 18, and 40 h) and proteins were separated by 12%

SDS-PAGE (Mini-Protein® TetraCell, BioRad) Western blot analysis was then used to confirm the target protein expression. The protein was incubated with a 1:2000 dilution of alkaline phosphatase (AP)-conjugated antibodies (Invitrogen) with a six-histidine residue tag at the N-terminus of CPR I and the C-terminus of CYP97C27 and then developed in substrate solution containing NBT/BICP. The target protein developed as a blue band.

2.6. Preparation of CYP97C27 and CPR I insoluble fractions

The insoluble fractions (20,000 × g pellet) of both CYP97C27 and CPR I were used for the detection of enzyme activities. The enzyme preparations were obtained as described previously [36]. For pET32a-CYP97C27, expression was induced in the bacterial cultures by adding 1 mM IPTG and 0.005 μM δ-ALA with shaking at 180 rpm at 37 °C until the OD₆₀₀ was 1.0. The cultures were then incubated at 4 °C for 18 h followed by incubation at 37 °C with shaking at 180 rpm for 3 h. For pET32a-CPR I, after incubation for 40 h, the insoluble fraction of CPR I was prepared by centrifugation at 20,000 × g at 4 °C for 20 min in 50 mM Tris-HCl pH 7.5 and 0.1 M NaCl. Protein quantitation was performed using Bradford's reagent solution (BioRad) with bovine serum albumin as the standard (Thermo Scientific).

2.7. Expression patterns of CYP97C27 and CPR I in C. stellatopilosus

The expression of the genes encoding CYP97C27 and CPR I in the shoots, leaves, and stalks of the plant was determined by semi-quantitative real-time PCR and was compared to the accumulation of plaunotol in C. stellatopilosus. Primers for real-time PCR were designed using the CloneManager v9.2 program and validated using the RTprimer database (http://medgen.ugent.be/rtprimerdb/) to ensure that the primers do not amplify other targets. The primers were used in a pre-amplification reaction consisting of normal PCR using LightCycler®480 SYBR Green I Master (Roche) to optimise the annealing temperature and the reaction as a whole. The load of detection (LOD) was validated using the 18S rRNA gene and determined as copies per nanogram, and Ct was optimised at 25 cycles. The RNA used in this study was extracted using the RNase™ Plant Mini Kit (QIAgen). After treatment with DNaseI (Fermentas), the RNA was reverse-transcribed to cDNA using M-MLV reverse transcriptase and an oligo(dT)₁₈ primer. Real-time PCR reactions were performed with a LightCycler®480 SYBR Green I Master (Roche) with gene-specific primers (Table 1). Within The reaction consisted of 1× LightCycler®480 SYBR Green I Master Mix, 20 ng cDNA and 1 mM primer pair (Table 1). The reaction progress was monitored in a Light Cycler 480 multiwell plate (Roche), performed in triplicate. The relative quantitative real-time PCR conditions were pre-incubation at 95 °C for 1 min, 44 cycles of amplification at 95 °C for 10 s, 55 °C for 30 s, 72 °C for 5 s, 1 cycle for the melting curve at 65 °C for 5 s, and 95 °C for 5 s. The relative expression fold was calculated with the CFX manager program (BioRad).

2.8. Plaunotol determination by thin layer chromatography (TLC)

The shoots, leaves, and stalks of C. stellatopilosus were ground into a powder using pestle-mortal after flash freezing with liquid nitrogen. Plaunotol extraction was performed as reported previously [2]. Briefly, each powdered sample was refluxed at 80 °C for 1 h in 10 mL methanol and evaporated to dryness. The dried extract was redissolved in 10 μL methanol and spotted onto a TLC silica gel 60 F254 aluminium plate (CAMAG). The plate was then developed in a saturated chamber with mobile phase containing toluene:acetonitrile:ethyl acetate:glacial acetic acid (35:5:15:0.15). The mobile phase reached the 90 mm developing
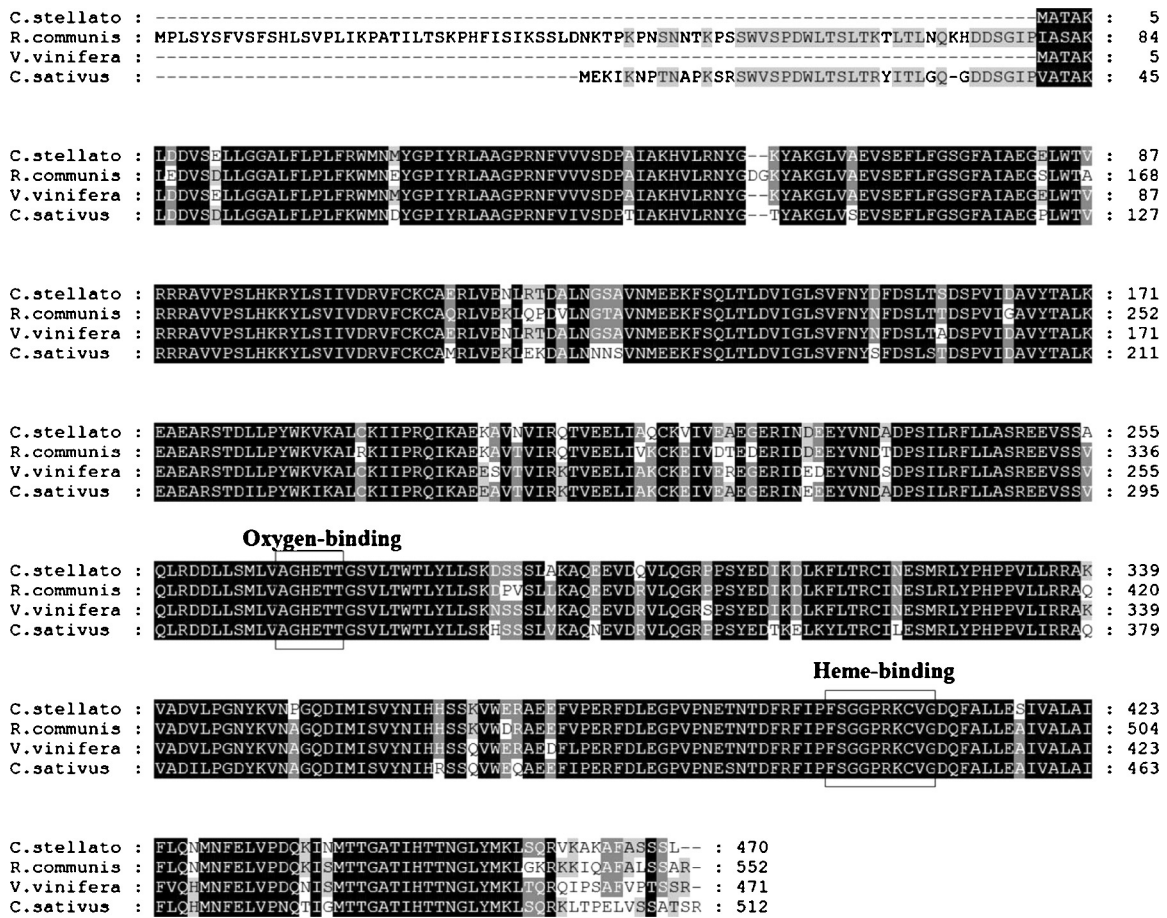
Fig. 2. Alignment of amino acid sequences of the plant CYPs. The alignment was performed based on the homologous amino acid sequences retrieved from other plant hosts using blastx (nucleotide translated blast) in NCBI database with the GeneDoc program [54]. The proteins shown in the alignment are from *Ricinus communis* (GenBank ID: XP002519427), *Vitis vinifera* (GenBank ID: CBI30186); CYP97C27, *Croton stellatopilosus* (GenBank ID: KF738256), and *Cucumis sativus* (GenBank ID: XP004156328). The boxes indicate a cluster of a heme-binding domain (FXXGXRXCXG) and an oxygen-binding domain ((A/G)GX(D/E)T(T/S)). Identical amino acid residues are boxed in black, and similar residues are shaded in grey. Gaps are inserted to maximise homology.

distance and was then visualised and documented at 254 nm. The plaunotol was identified by densitometry scanning at 210 nm.

2.9. Enzyme activity assay

A total of 100 mL (50–250 μg) of the crude CYP97C27 and CPR I enzymes was added to the plaunotol enzymatic reaction assay mixture containing 83 mM tricine pH 7.8, 0.8 mM NADPH, and 57 mM GGOH as the substrate. The reaction was incubated with vigorous shaking in a 30 °C incubator for 18 h. The product of the reaction was extracted twice with 300 μL ethyl acetate and dried under a vacuum for 2 h at 30 °C. The resulting product was resuspended in 5 μL ethyl acetate and spotted onto a TLC silica gel 60 F254 aluminium plate (CAMAG). The chromatogram was visualised after the TLC plate was developed in toluene:acetonitrile:ethyl acetate:glacial acetic acid (35:5:15:0.15) as the mobile phase.

2.10. Product identification by LC-MS

The band of the enzymatic product co-chromatographed with the standard plaunotol on TLC plates was scratched off as a powder and eluted with 5 mL methanol. The solution was evaporated to dryness, weighed, and redissolved with methanol to obtain a concentration of 1 mg/mL for further LC-MS analysis. The product solution (20 μL) was applied to a LC-APCIMS (Agilent 1200 series, Bruker MicroTOF) using a TSK gel Super-ODS column (4.6 × 200 mm, 2.3 μM particle size, Tosoh), and eluted for

20 min with methanol:water (7:3) (flow rate 0.8 mL/min, temperature 35 °C). The retention time and LC-APCIMS extracted ionised chromatogram at m/z = 200–350 were used to compare the product and authentic plaunotol.

3. Results

3.1. Gene cloning and phylogenetic tree analysis

In searching for a CYP candidate involved in the 18-hydroxylation of GGOH to form plaunotol, degenerate primers were designed from the conserved regions covering various CYP-family clans known to be involved in the hydroxylation of different substrates [36]. These clans include CYP51 (sterols), CYP71/76 (shikimate), CYP72 (isoprenoid hormones), CYP74 (allene), CYP85 (sterol and acyclic terpenes), CYP86 (fatty acid), and CYP97 (carotenoids). Based on these degenerate primer designs, it appeared that only the core fragments of CYP76, CYP86, and CYP97 genes could be amplified from the cDNAs isolated from *C. stellatopilosus* leaves. These three core fragments were then used as templates for primer designation of each 5′RACE and 3′RACE fragment, but only the core fragments of CYP76 and CYP97, not CYP86, could be amplified. The resulting consensus sequences of the two genes were then used as templates to obtain the full-length genes. The genes were designated as CYP76F45 and CYP97C27, respectively, by David Nelson (http://drnelson.utmem.edu/CytochromeP450.html), with
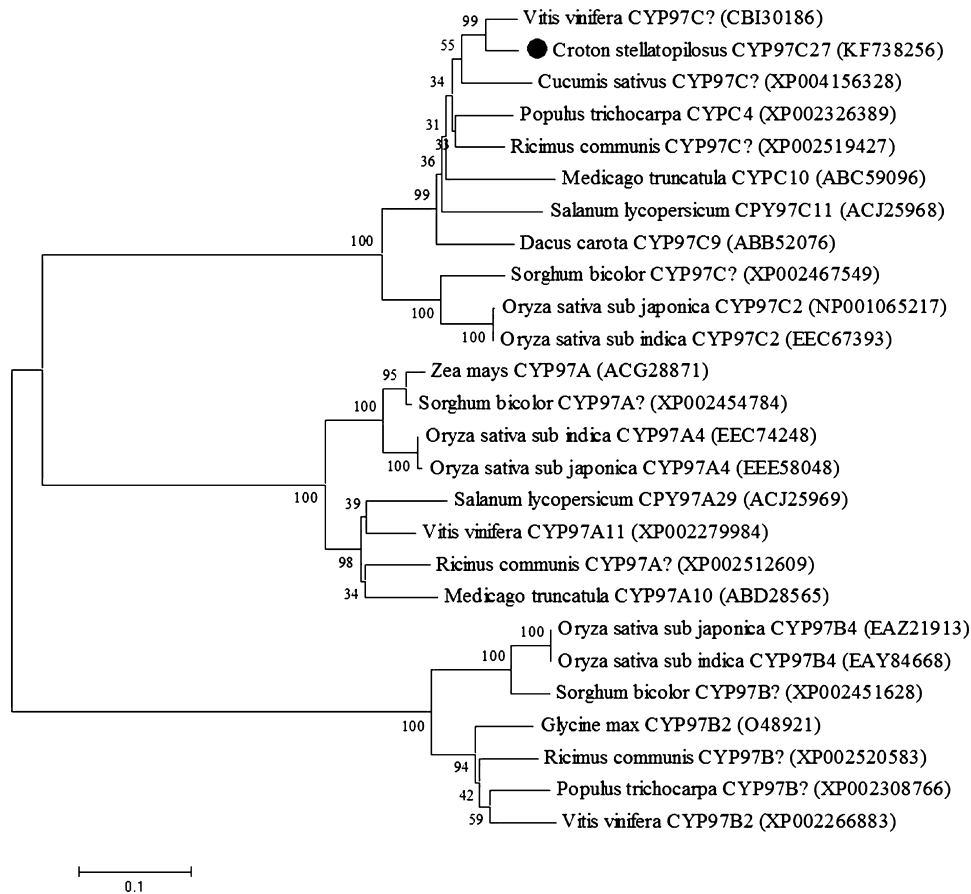
Fig. 3. Phylogenetic tree of the CYP family members. The mid-point root neighbour-joining (NJ) tree was generated using the MEGA6 program (http://www.megasoftware.net) [33]. The homologous sequences were obtained using the National Center for Biotechnology Information search engine (http://www.ncbi.nim.bih.gov/) [55]. The cytochrome P450s are listed by the species name followed by the GenBank accession numbers. CYP97C27 isolated from *C. stellatopilosus* is indicated by a dark spot. The number presented in the branch indicates the bootstrap values supporting the evolution [56,57].

the former showing high amino acid identity with the enzyme geraniol 8-hydroxylase (data not shown), and the latter was further characterised in this study. For CPR I, a similar homology approach was also used to identify homologous proteins. A degenerate primer pair was used to amplify the 1100 bp core fragment (the FMN-binding motif forward primer and NADPH-binding motif reverse primer). The obtained core fragment showing homology to plant cytochrome P450 reductases was subsequently used as a template for 5′-RACE and 3′-RACE fragments. The full-length cDNA was amplified using the primers from the consensus sequence of the three-fragment sequence assembly.

The open reading frames (ORFs) of CYP97C27 and CPR I contained 1413 and 2300 bp, which corresponded to 471 and 711 amino acid residues, respectively (ExPASy algorithmic tool at http://au.expasy.org/cgi-bin/pi_tool). The predicted monomeric molecular weights using the ExPASy-Compute pI/Mw tool (http://web.expasy.org/compute_pi) for CPR97C27 and CPR I were 52,834 and 78,744 Da with theoretical pIs of 5.27 and 5.94, respectively. Based on the PSI- and PHI-BLAST searches against the conserved domain database (CDD) of NCBI, it was confirmed that the obtained translated cDNAs were cytochrome P450 (CYP97C27) and cytochrome P450 reductase (CPR I).

The CYP97C27 gene showed high amino acid sequence similarity with the *CYPs* from several plants, particularly *Vitis vinifera* (93% identity), *Ricinus communis* (90% identity), and *Cucumis sativus* (85% identity) (Fig. 2). The cloned CYP7C27 contained characteristic

conserved motifs of plant cytochrome P450s that possess a heme-binding domain (FXXGXRXCXG) and an oxygen-binding domain ((A/G)GX(D/E)T(T/S)). As shown in Fig. 2, the heme-binding domain of FSGGPRKCVG is located at position 400–409, and the oxygen-binding domain AGHETT is at position 267–272. The results of the phylogenetic tree construction using 25 protein sequences and 12 different plant species from the CYP97 family (A, B, and C clans) suggest that CYP97C27 should be classified with the CYP97 C subfamily (Fig. 3) and is closely related to the CYP97 C isolated from *V. vinifera* (99% bootstrap supported). By personal communication, it was named CYP97C27 based on the nomenclature system established by David Nelson's committee group (http://drnelson.utmem.edu/CytochromeP450.html) [37].

The CPR I gene also showed high homology to the *CPR* genes from a number of plants, including *Ricinus communis* (83% identity), *Gossypium hirsutum* (81% identity), and *Populus trichocarpa* (79% identity) (Fig. 4). The alignment profile confirmed that CPR I, similar to other CPRs, had two binding sites each for FMN, FAD, and NADPH and one binding site for P450. The phylogenetic tree constructed with 24 protein sequences and 20 different plant species from the CPR family (class I and II) (Fig. 5) revealed that CPR I is classified in CPR class I and is closely related to the CPR isolated from *R. communis* with 99% bootstrap support. CPR class I consists of the CPR isoforms found in dicotyledon plants.

The nucleotide sequences of both genes were submitted to the NCBI database as GenBank accession no. KF738256 and KF738254 for CYP97C27 and CPR I, respectively.
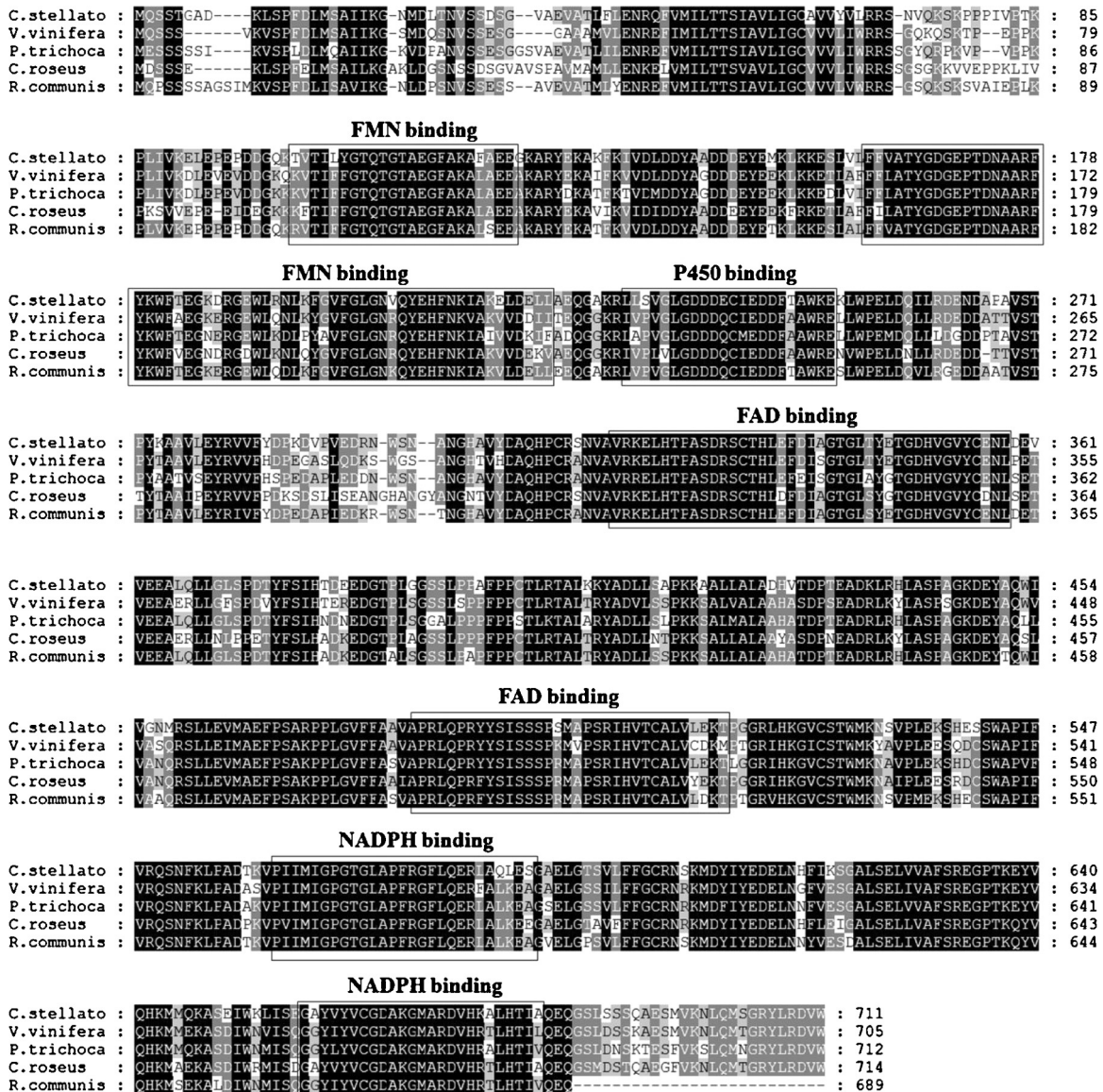
```
                                                   FMN binding
C.stellato : MQSSTGAD----KLSPFDIMSAIIKG-NMILTNVSSDSG--VAEVATIFLENRQFVMILTTSIAVLIGQAVVYVLRRS-NVQKSKPPPIVETK : 85
V.vinifera : MQSSS------VKVSPFDLMSAIIKG-SMDQSNVSSESG----GAAAMVLENREFIMILTTSIAVLIGCVVVLIWRRS-GQKQSKTP--EEPK : 79
P.trichoca : MESSSSSI----KVSPLDIMCAIIKG-KVDFANVSSESGGSVAEVATLILENREFVMILTTSIAVLIGCVVVLIWRRSSGYQFPKVP--VLPK : 86
C.roseus   : MDSSSE------KLSPFDLMSAILKG-AKLLDGSNSSDSGVAVFSAVMAMLLENKEIVMILTTSVAVLIGCVVVLIWRRSSGSGKVVEPPKLIV : 87
R.communis : MQFSSSSAGSIMKVSPFDLISAVIKG-NLDPSNVSSSESS--AVEVATMIYENREFVMILTTSIAVLIGCVVVLVWRRS-GSQKSKSVAIBELK : 89

                      FMN binding
C.stellato : PLIVKELEPEDDGQKTVTILYGTQTGTAEGFAKAFAEEGKARYEKAKFKIVDLDDYAADDDEYEMLKKESLVIFFVATYGDGEPTDNAARF : 178
V.vinifera : PLIVKDLEVEVDDGKQKVTIFFGTQTGTAEGFAKALAEEAKARYEKAIIKVVDLDDYAGDDDEYEKLKKETIAFFFIATYGDGEPTDNAARF : 172
P.trichoca : PLIVKDLEPEVDDGKKVTIFFGTQTGTAEGFAKALAEEAKARYDKATFKTVCMDDYAGDDDEYEKLKKDDLVIFFIATYGDGEPTDNAARF : 179
C.roseus   : PKSVVEPE-DIDEGKKKFTIFFGTQTGTAEGFAKAIAEEAKARYEKAVIKVIDIDDYAADDEYEKFRKETLAFFIATYGDGEPTDNAARF : 179
R.communis : PLVVKEPEPEPEDDGQKRVTIFFGTQTGTAEGFAKAISEEAKARYEKAIFKVVDLDDYAADDDEYETKLKKESIAFFVATYGDGEPTDNAARF : 182

                 FMN binding                                    P450 binding
C.stellato : YKWFTEGKDRGEWIRNLKFGVFGLGNVQYEHFNKIAKELDELIAEQGAKRLLSVGLGDDDECIEDDFTAWKEKLWPELDQILRDENDAFAVST : 271
V.vinifera : YKWFAEGKERGEWIQNLKYGVFGLGNRQYEHFNKVAKVVDELIIIEQGGKRIVPVGLGDDDQCIEDDFAAWREILWPELDQLLRDEDDATTVST : 265
P.trichoca : YKWFTEGNERGEWIKDLEYAVFGLGNRQYEHFNKIATVVDKIFALQGGKRIAFVGLGDDDQCMEDDFAAWRELLWPEMDQLILCGDPTAVST : 272
C.roseus   : YKWFVEGNDRGDWIRNLQYGVFGLGNRQYEHFNKIAKVVDEKVAEQGGKRIVPLVLGDDDQCIEDDFAAWRENVWPELDNLLRDEDD-TTVST : 271
R.communis : YKWFTEGKERGEWIQDLKFGVFGLGNKQYEHFNKIAKVLDELIAEQGAKRLVPFVGLGDDDQCIEDDFTAWKESLWPELDQVLFGEDDAATVST : 275

                                                        FAD binding
C.stellato : FYKAAVLEYRVVFYDPKLVPVEDRN-WSN--ANGHAVYLAQHPCRSNVAVRKELHTFASDRSCTHLEFDIAGTGLTYETGDHVGVYCENLDEV : 361
V.vinifera : FYTAAVLEYRVVFHDPEGASLQDKS-WGS--ANGHTVHLAQHPCRANVAVRKELHTFASDRSCTHLEFDISGTGLTYETGDHVGVYCENLPET : 355
P.trichoca : FYAATVSEYRVVFHSPEFAPLEDDN-WSN--ANGHAVYLAQHPCRANVAVRRELHTFASDRSCTHLEFDISGTGIAYETGDHVGVYCENLSET : 362
C.roseus   : TYMAAIEEYRVVFPLKSFSLISEANGHANGYANGNTVVLAQHPCRSNVAVRKELHTFASDRSCTHLEFDIAGTGLSYETGDHVGVYCENLSET : 364
R.communis : FYTAAVLERIVFYDPEDAPIEDKR-WSN--TNGHAVYLAQHPCRANVAVRKELHTFASDRSCTHLEFDIAGTGLSYETGDHVGVYCENLDET : 365

C.stellato : VEEALQLLGLSPDTYFSIHLREDGTFIGGSSLPAFPPCTLRTALKKYADLLSAPKKALLIALADFVTDPTEADKLRELASFAGKDEYAQWI : 454
V.vinifera : VEEAERLLGFSPDVYFSIHTEREDGTPLSGSSLSPPFPPCTLRTALTRYADVLSSPKKSALVALAAHASDPSEADRLKYLASFSGKDEYAQWV : 448
P.trichoca : VEEALQLLGLSPDTYFSIHLNENEDGTPLSGGALPPPFPPSTLKTALARYADLLSLPKKSALMALAAHATDPTEADRLRELASFAGKDEYAQLL : 455
C.roseus   : VEEAERLLNIFPETYFSLHADKEDGTFIAGSSLPPPFPPCTLRTALTRYADLLNTPKKSALLALAAYASDFNEADRLKYLASFAGKDEYAQSL : 457
R.communis : VEEALQLLGLSPDTYFSIHADKEDGTALSGSSLPAFPPCTLRTALTRYADLLSPKKSALLALAAHATDPTEADRLRELASFAGKDEYTQWI : 458

                                 FAD binding
C.stellato : VGNMRSLLEVMAEFPSARPPLGVFFAAVAPRLQPRYYSISSSPSMAPSRIHVTCALVLEKTEGGRLHKGVCSTWMKNSVPLEKSHESSWAPIF : 547
V.vinifera : VASQRSLLEIMAEFPSAKPPLGVFFAAVAPRLQPRYYSISSSPKVPSRIHVTCALVCDRMETGRIHKGICSTWMKYAVPLPESQDCSWAPIF : 541
P.trichoca : VANQRSLLEVMAEFPSAKPPLGVFFASVAPRLQPRYYSISSSPRMAPSRIHVTCALVEKTLGGRIHKGVCSTWMKNAVPLEKSHDCSWAPVF : 548
C.roseus   : VANQRSLLEVMAEFPSAKPPLGVFFAAIAPRLQPRFYSISSSPRMAPSRIHVTCALVYEKTEGGRIHKGVCSTWMKNAIPLEBSRDCSWAPIF : 550
R.communis : VAAQRSLLEVMAEFPSAKPPLGVFFASVAPRLQPRFYSISSSPRMAPSRIHVTCALVLDKTEGRVHKGVCSTWMKNSVPMEKSHECSWAPIF : 551

                                NADPH binding
C.stellato : VRQSNFKLFADTKVPIIMIGPGTGLAPFRGFLQERIAQLFSGAELGTSVLFFGCRNSKMDYIYEDELNHFIKSGALSELVVAFSREGPTKEYV : 640
V.vinifera : VRQSNFKLFALASVPIIMIGPGTGLAPFRGFLQERFALKEAGALGSSILFFGCRNRKMDYIYEDELNGFVESGALSELIVAFSREGPTKEYV : 634
P.trichoca : VRQSNFKLFALAKVPIIMIGPGTGLAPFRGFLQERIALKEAGSELGSSVLFFGCRNRKMDFIYEDELNNFVESGALSELVVAFSREGPTKEYV : 641
C.roseus   : VRQSNFKLFADPKVPVIMIGPGTGLAPFRGFLQERIALKEAGAELGTIAVFFGCRNRKMDYIYEDELNHFLEFIGALSELLVAFSREGPTKQYV : 643
R.communis : VRQSNFKLFADTKVPIIMIGPGTGLAPFRGFLQERIALKEAGVELGPSVLFFGCRNSKMDYIYEDELNNYVESDALSELIVAFSREGPTKQYV : 644

                             NADPH binding
C.stellato : QHKMMQKASEIWKLISEGAYYVVCGDAKGMARDVHKALHTIAQEQGSISSSCAESMVKNLQMSGRYLRDVW : 711
V.vinifera : QHKMMEKASDIWNLISQGGYIYVCGDAKGMARDVHRILHTILQEQGSIDSSKAESMVKNLQMTGRYLRDVW : 705
P.trichoca : QHKMMQKASDIWNMISQGGYLYVCGDAKGMAKDVHRALHTIVQEQGSIDNSKTESEVKSLQMNGRYLRDVW : 712
C.roseus   : QHKMAEKASDIWFMISDGAYYVVCGDAKGMARDVHRTLHTIAQEQGSMDSTCAEBGFVKNLQMTGRYLRDVW : 714
R.communis : QHKMSEKALDIWNMISGGGYIYVCGDAKGMARDVHRTLHTIVQEQG------------------------- : 689
```

Fig. 4. Alignment of the amino acid sequences of the plant CPRs. The alignment was performed based on the homologous amino acid sequences retrieved from other plant hosts using blastx (nucleotide translated blast) in the NCBI database with the GeneDoc program [54]. The proteins shown in the alignment are from *Gossypium hirsutum* (GenBank ID: ACN54324), *Populus trichocarpa* (GenBank ID: AF302497), *Croton stellatopilosus* (CPR I, GenBank ID: KF738254), and *Ricinus communis* (GenBank ID: XP002534464). The boxes indicate the typical motifs of the P450-, NADPH-, FMN-, and FAD-binding domains. Identical amino acid residues are boxed in black, and similar residues are shaded in grey. Gaps are inserted to maximise the homology.

3.2. Recombinant protein expression in the E. coli system

The putative genes encoding CYP97C27 and CPR I were constructed separately in the expression vector pET32a, and protein expression was placed under the T7-promoter in the *E. coli* BL21 (DE3) strain. Optimal expression conditions were first established by varying the induction time. The pET32a-CYP97C27 and pET32a-CPR I recombinants appeared to be expressed starting in the first hour, and expression levels increased with time. As shown in Fig. 6, protein expression was clearly observed for CPY97C27 (Fig. 6**A**) and CPR I (Fig. 6**B**). The expression of both proteins in *E. coli* was confirmed by Western blot analysis, which showed monomeric sizes of 56 kDa for CYP97C27 and 97 kDa for CPR I (Fig. 6**C**).

3.3. Functional identification of CYP97C27 and CPR I

To determine whether CYP97C27 and CPR I are the cytochrome-related enzymes involved in the last step of GGOH hydroxylation to form plaunotol, an *in vitro* enzyme assay was performed. The recombinant CYP97C27 and CPR I proteins prepared as insoluble fractions were incubated under various conditions to observe the potential coordination of CYP97C27 and CPR I in catalysing the C-18 hydroxylation of GGOH. The varied conditions included a set of incubation mixtures containing the CYP97C27 preparation alone and with or without NADPH, GGOH, and CPR I for 0 and 18 h incubations. The reaction mixtures were then each extracted with ethyl acetate and analysed for product formation by TLC. As shown in Fig. 7, no reaction mixture resulted in the enzymatic product except for the complete reaction mixture with CYP97C27, CPR I, GGOH, and NADPH, which showed time-dependent product formation (Fig. 7). Absolutely no product formation was observed in the boiled control of the same complete reaction mixture. The product positions on the TLC plate, lanes 12–14, appear to have the same R_f as the plaunotol standard (lane 15). The product band was then removed from the TLC plate, extracted with methanol and subjected to LC-MS analysis. Using the positive-ion mode, the fragment ions and
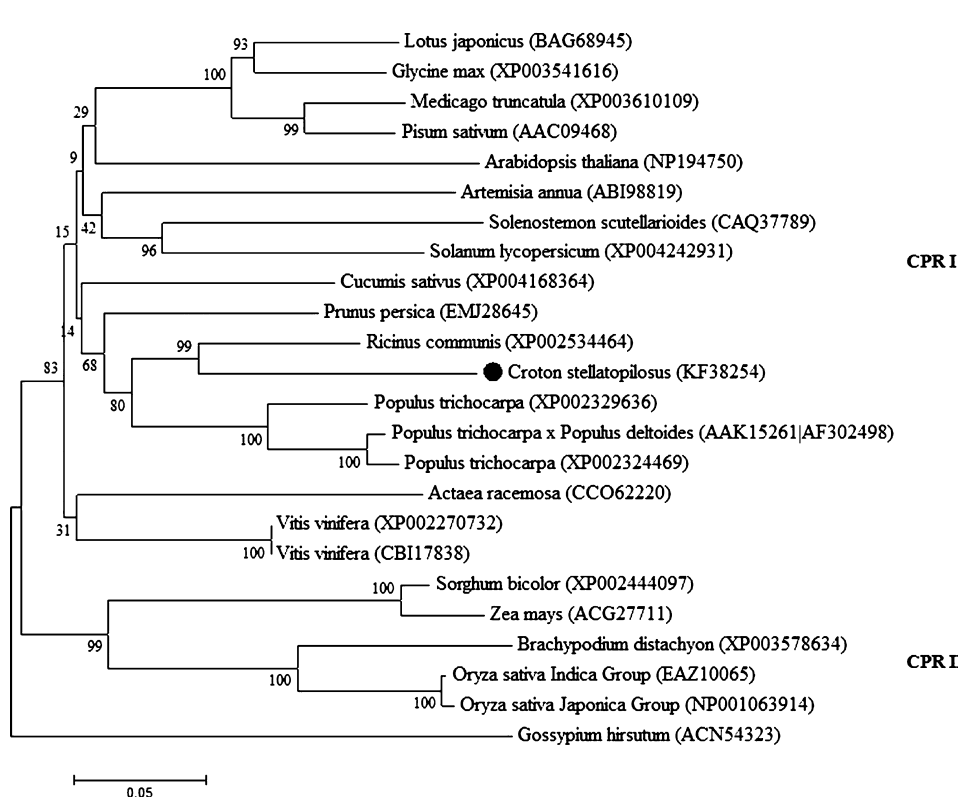
Fig. 5. Phylogenetic tree of the CPR family members. The mid-point root neighbour-joining (NJ) tree was generated using the MEGA6 program (http://www.megasoftware.net) [33,35]. The homologous sequences were obtained using the National Center for Biotechnology Information search engine (http://www.ncbi.nim.bih.gov/). The cytochrome P450 reductases are listed by the species name followed by the GenBank accession numbers. The CPR I isolated from *C. stellatopilosus* is indicated by a dark spot. The CPRs classified as CPR I and CPR II are labelled in the figure. The number presented at the branch indicates the bootstrap values supporting the evolution.

the retention time of the product (15 min) were consistent with the authentic plaunotol. The product was also analysed in ESI(+) mode with an observed mass of m/z 307 and the characteristic $[M + Na]^+$ (306 + 23) ion (Fig. 8B). These results confirm that the enzymatic product obtained from coordinated catalysis by CYP97C27 and CPR I was plaunotol.

3.4. Transcription profile analysis of CYP97C27 and CPR I

To confirm that CYP97C27 and CPR I are both involved in the biosynthetic pathway of plaunotol, the transcription profiles of both the CYP97C27 and CPR I mRNAs were examined in three different plant parts that may be the sites of plaunotol biosynthesis. The plant parts were the shoot, leaf and stalk, and the genes selected were DXS and MEPS, the first two genes of the MEP pathway, GGPPS and GGPPP, the first two genes of the plaunotol pathway, and CYP97C27 and CPR I, the last two genes of the pathway. All six genes were fully expressed in the leaf, whereas there was limited expression of GGPPS and relatively low levels of CYP97C27 in the shoot. In

the stalk, in contrast, there was extremely low expression of DXS, MEPS, and GGPPS, and limited expression of CYP97C27 (Fig. 8A). These transcription profiles were consistent with the plaunotol content profile in each of plant part (Fig. 8B).

4. Discussion

Plant cytochrome P450 monooxygenases (CYPs) and NADPH-cytochrome P450 reductases (CPRs) function coordinately to catalyse many oxidative reactions involved in both primary and secondary metabolism [38,39]. CYPs provide a considerable variety of different regio- and stereospecific reactions, and CPRs catalyse the transfer of electrons from NADPH via FAD and FMN to the prosthetic heme group of the CYP proteins to complete the reaction [40]. In this study, the cDNAs for a P450 and CPR from the same *C. stellatopilosus* plant were successfully isolated. The cDNAs, CYP97C27, and CPR I, were obtained from young leaves, which are known to be the site of active plaunotol biosynthesis and accumulation [41,42]. The amino acid sequences deduced for both genes confirmed that
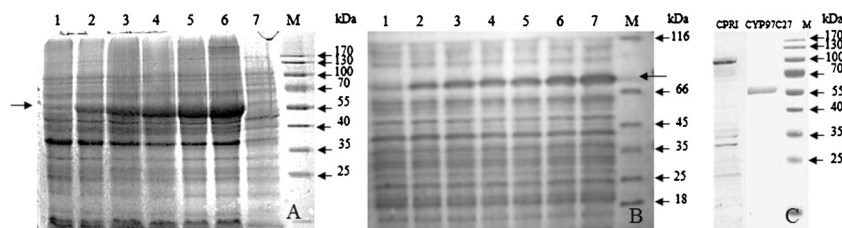

Fig. 6. Heterologous expression of the recombinant CYP97C27 (A) and CPR I (B) proteins in *E. coli* analysed by 12% SDS-PAGE. Lanes 1–7 show protein patterns of cell lysates obtained after IPTG induction for 0, 1, 2, 4, 12, 24, and 40 h, respectively. (C) Immunoblot with alkaline phosphatase (AP)-conjugated antibodies with a six-histidine residue tag at the N-terminus of CPR I and C-terminus of CYP97C27.
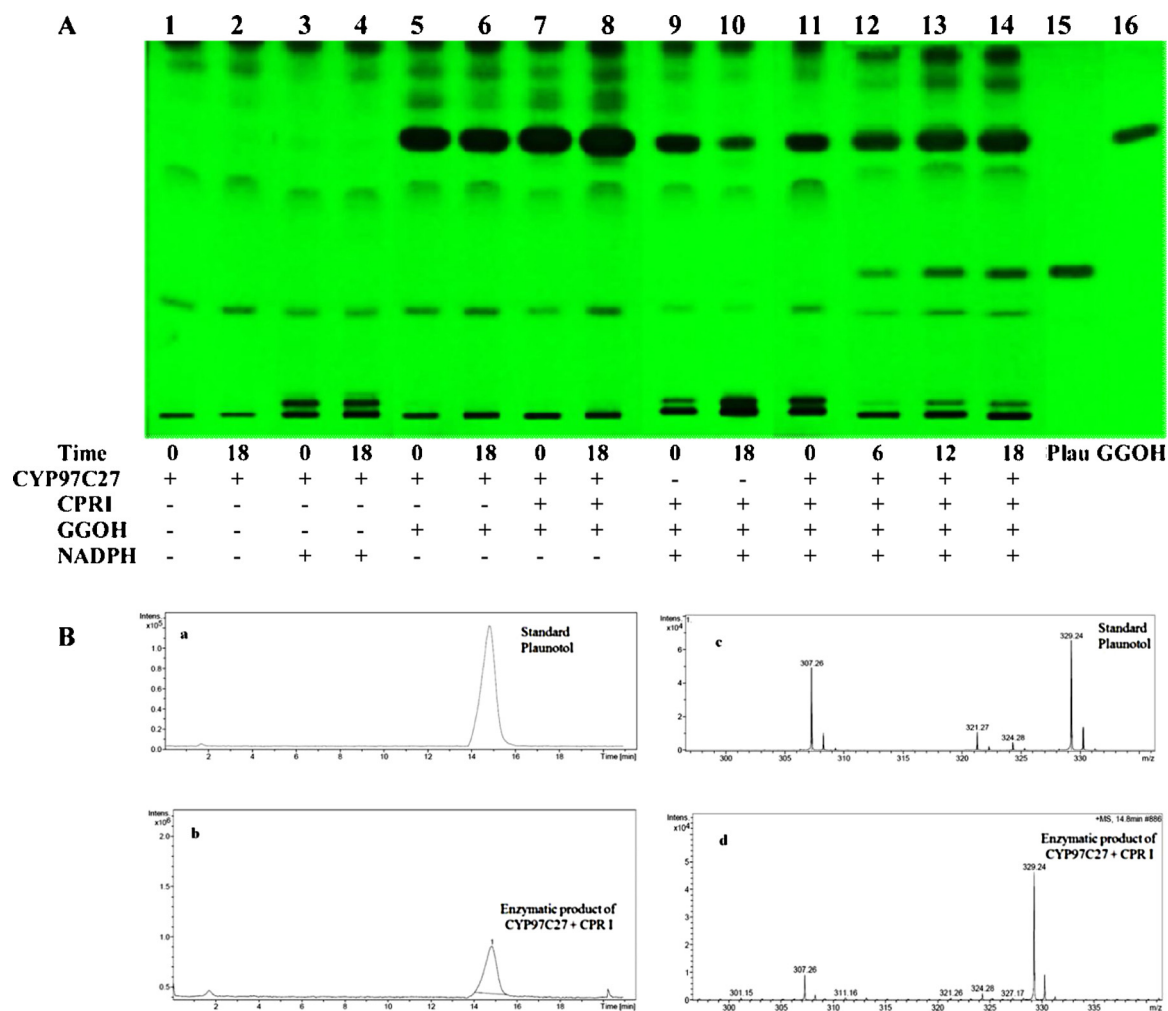
Fig. 7. (A) TLC detection of the enzymatic product of CYP97C27/CPR I using various reaction mixture ingredients. The insoluble fraction (250 μg) of CYP97C27, CPR I or both was used in each enzyme assay reaction. Each reaction mixture was incubated for 0 and up to 18 h at 30 °C before being analysed using TLC with plaunotol (plau, product) and geranylgeraniol (GGOH, substrate) as the standards. (B) LC-MS analysis of the enzymatic products (lanes 12–14) co-chromatographed with the plaunotol standard (lane 15). The retention times (a and b) and MS ions at m/z = 200–350 (c and d) were compared between the plaunotol standard (a and c) and the product (b and d).

the CYP97C27 had the characteristics of the plant cytochrome P450 monooxygenase superfamily (heme and oxygen binding) [18], and CPR I had characteristics of NADPH-cytochrome reductases (CYP and FMN, FAD and NADPH binding) [40].

The results of the phylogenetic analysis revealed that CYP97C27 was classified in the P450 family CYP97 under the subfamily CYP97 C. Its sequence was submitted to David Nelson's committee group (http://drnelson.utmem.edu/CytochromeP450.html) and was identified and named CYP97C27. The CYP97 C subfamily is grouped based on the function of the carotenoid epsilon-ring hydroxylases [43–45]. However, this reaction, which is proposed to be catalysed by CYP97 C, has been demonstrated in only *A. thaliana* [37] and *Solanum lycopericum* [46,47]. Additional studies are necessary to confirm the function of this enzyme.

Both CYP97C27 and CPR I were expressed as catalytically active enzymes using a simple prokaryotic *E. coli* expression system in the BL21(DE3) *E. coli* strain using the pET vector system. Both genes were presumably expressed as plasma membrane-bound enzymes because *E. coli* does not possess cellular compartmentation, and their activities were only detected in the insoluble 20,000 × *g* pellet fraction. Interestingly, the activity of both CYP97C27 and CPR I could only be detected in a reaction mixture containing both enzyme preparations based on conversion of GGOH to plaunotol. Absolutely no product formation was observed in any of the

incomplete reaction mixtures, including without NADPH and GGOH. This clearly suggested that both CYP97C27 and CPR I are directly involved in the catalysis of the C18-hydroxylation of GGOH.

Many plant CYPs and CPRs have been expressed in eukaryotic yeasts and insect cells, and a few have been expressed in prokaryotic *E. coli* [for review, [48]]. The two commonly used bacterial *E. coli* and yeast *Saccharomyces cerevisiae* systems differ from each other with respect to the presence of endogenous CYPs and CPRs. *E. coli* has no endogenous CYP or CPR genes, yeast contains three CYP and one associated CPR genes [49]. In addition, it is hypothesised that the expression of unmodified plant CYPs in *E. coli* may be more complicated than in yeast because of differences in both codon preference and the nature of the eukaryotic membrane-bound system, which is not present in bacteria [50,51]. Therefore, in the last few decades, P450 researchers have switched from an *E. coli* to yeast model system, and different yeast strains have been engineered to overcome the issue of expressing native CYPs and CPRs [48]. However, our successful functional expression of CYP97C27 and CPR I in *E. coli* suggest that the above-mentioned concerns are actually not applicable to these membrane-bound enzymes. The major advantages of our bacterial expression system are as follows: (1) lack of endogenous CYP orthologues capable of interfering with expression and activity assays compared to the yeast system, (2) large quantities of protein produced in *E. coli*, and importantly, (3) the pET
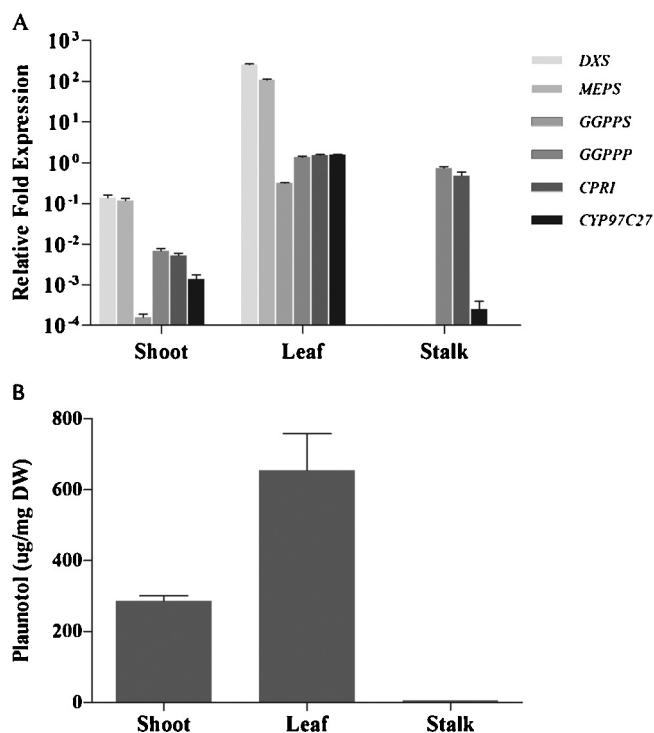
Fig. 8. mRNA-expression levels of several key enzymes involved in the biosynthesis of plaunotol in the shoots, leaves, and stalks of *C. stellatopilosus*. (A) Relative fold expression levels (relative to 18S rRNA) of the DXS, MEPS, GGPPS, GGPPP, CYP97C27, and CPR I genes in shoots, leaves, and stalks using real-time PCR. (B) The plaunotol content was also detected in the shoots, leaves, and stalks.

system developed to avoid toxic effects of overexpression of membrane proteins using the T7 bacteriophage promoter in the bacterial BL21 strain [50]. Previously, two *Arabidopsis* proteins, CYP79F1 and CYP79F2, both involved in the biosynthesis of aliphatic glucosinolates, have been characterised after recombinant expression in *E. coli* [51,52].

The observed enzymatic formation of plaunotol by the reconstitution of the two separate CYP97C27 and CPR I preparations also confirmed that both cytochrome P450 monooxygenase and NADPH-cytochrome P450 reductase work cooperatively as the GGOH 18-hydroxylase enzyme in *C. stellatopilosus*. Whether both are specific only for catalysis of the 18-hydroxylation or whether either can catalyse other hydroxylation reactions is still not known. Additionally, whether both are localised close to each other in situ to catalyse the specific 18-hydroxylation remains to be clarified. However, it is likely that fusion of CYP97C27 and CPR I would significantly enhance their catalytic activity, similar to the case of *Sorghum bicolor* CYP79A1 and CYP71E1 and *S. bicolor* CPR2. The modified pYeDP60 vector in all three cases produced high amounts of active protein [53].

The involvement of CYP97C27 and CPR I in the biosynthesis of plaunotol is also supported by the results of real-time expression analysis of the two genes in the shoots, leaves, and stalks. The expression of CYP97C27 clearly showed correlation with the accumulated content of plaunotol in these plant parts, although it appeared that GGPPS had limited expression in the shoots, and DXS and DXP had limited expression in the stalk. However, the observed parallel levels of CYP97C27 and CPR I expression in the shoot and stalk also suggested that these genes are not closely associated. It is likely that CYP97C27 specifically functions in plaunotol biosynthesis, whereas CPR I has a more general function with other CYPs in the same plant.

References

[1] A. Ogiso, E. Kitazawa, M. Kurabayashi, A. Sato, S. Takahashi, H. Noguchi, H. Kuwano, S. Kobayashi, H. Mishima, Isolation and structure of antipeptic ulcer diterpene from Thai medicinal plant, Chem. Pharm. Bull. (Tokyo) 26 (1978) 3117–3123.

[2] A. Vongchareonsathit, W. De-Eknamkul, Rapid TLC-densitometric analysis of plaunotol from *Croton sublyratus* leaves, Planta Med. 64 (1998) 279–280.

[3] J. Wungsintaweekul, W. De-Eknamkul, Biosynthesis of plaunotol in *Croton stellatopilosus* proceeds via the deoxyxylulose phosphate pathway, Tetrahedron Lett. 46 (2005) 2125–2128.

[4] W. Sitthithaworn, N. Kojima, E. Viroonchatapan, D. Suh, N. Iwanami, T. Hayashi, M. Noji, K. Saito, Y. Niwa, U. Sankawa, Geranylgeranyl diphosphate synthase from *Scoparia dulcis* and *Croton sublyratus* plastid localisation and conversion to a farnesyl diphosphate synthase by mutagenesis, Chem. Pharm. Bull. (Tokyo) 49 (2001) 197–202.

[5] N. Nualkaew, W. De-Eknamkul, T.M. Kutchan, M.H. Zenk, Geranylgeraniol formation in *Croton stellatopilosus* proceeds via successive monodephosphorylations of geranylgeranyl diphosphate, Tetrahedron Lett. 46 (2005) 8727–8731.

[6] N. Nualkaew, N. Guennewich, K. Springob, A. Klamrak, W. De-Eknamkul, T.M. Kutchan, Molecular cloning and catalytic activity of a membrane-bound prenyl diphosphate phosphatase from *Croton stellatopilosus* Ohba, Phytochemistry 91 (2013) 140–147.

[7] N. Nualkaew, W. De-Eknamkul, T.M. Kutchan, M.H. Zenk, Membrane-bound geranylgeranyl diphosphate phosphatases: purification and characterisation from *Croton stellatopilosus* leaves, Phytochemistry 67 (2006) 1613–1620.

[8] W.P. Sitthithaworn, B.W. De-Eknamkul, Localisation of Plaunotol in the leaf of *Croton stellatopilosus* Ohba, ScienceAsia 32 (2006) 17–20.

[9] P. Tansakul, W. De-Eknamkul, Geranylgeraniol-18-hydroxylase: the last enzyme on the plaunotol biosynthetic pathway in *Croton sublyratus*, Phytochemistry 47 (1998) 1241–1246.

[10] M. Chanama, T. Wunnakup, W. De-Eknamkul, S. Chanama, Improvement of thin-layer chromatography for enzyme assay of geranylgeraniol 18-hydroxylase from *Croton stellatopilosus* Ohba, J. Planar Chromatogr. Mod. TLC 22 (2009) 49–53.

[11] M.A. Schuler, D. Werck-Reichhart, Functional genomics of P450s, Annu. Rev. Plant Biol. 54 (2003) 629–667.

[12] S. Rana, S.K. Lattoo, N. Dhar, S. Razdan, W.W. Bhat, R.S. Dhar, R. Vishwakarma, NADPH-cytochrome P450 reductase: molecular cloning and functional characterisation of two paralogs from *Withania somnifera* (L.) dunal, PLoS One 8 (2013) e57068.

[13] T. Laursen, K. Jensen, B.L. Moller, Conformational changes of the NADPH-dependent cytochrome P450 reductase in the course of electron transfer to cytochromes P450, Biochim. Biophys. Acta 1814 (2011) 132–138.

[14] L.M. Podust, D.H. Sherman, Diversity of P450 enzymes in the biosynthesis of natural products, Nat. Prod. Rep. 29 (2012) 1251–1266.

[15] S. Bak, F. Beisson, G. Bishop, B. Hamberger, R. Hofer, S. Paquette, D. Werck-Reichhart, Cytochromes p450, Arabidopsis Book 9 (2011) e0144.

[16] F. Durst, D.R. Nelson, Diversity and evolution of plant P450 and P450-reductases, Drug Metabol. Drug Interact. 12 (1995) 189–206.

[17] C. Chapple, Molecular genetic analysis of plant cytochrome P450-dependent monooxygenases, Annu. Rev. Plant Physiol. Plant Mol. Biol. 49 (1998) 311–343.

[18] D. Werck-Reichhart, S. Bak, S. Paquette, Cytochromes p450, Arabidopsis Book 1 (2002) e0028.

[19] B. Hamberger, S. Bak, Plant P450s as versatile drivers for evolution of species-specific chemical diversity, Philos. Trans. R Soc. Lond. B Biol. Sci. 368 (2013) 20120426.

[20] M.S. Shet, K. Sathasivan, M.A. Arlotto, M.C. Mehdy, R.W. Estabrook, Purification, characterisation, and cDNA cloning of an NADPH-cytochrome P450 reductase from mung bean, Proc. Natl. Acad. Sci. U. S. A. 90 (1993) 2890–2894.

[21] D. Eberle, P. Ullmann, D. Werck-Reichhart, M. Petersen, cDNA cloning and functional characterisation of CYP98A14 and NADPH:cytochrome P450 reductase from *Coleus blumei* involved in rosmarinic acid biosynthesis, Plant. Mol. Biol. 69 (2009) 239–253.

[22] A. Rosco, H.H. Pauli, W. Priesner, T.M. Kutchan, Cloning and heterologous expression of NADPH-cytochrome P450 reductases from the Papaveraceae, Arch. Biochem. Biophys. 348 (1997) 369–377.

[23] A. Lesot, I. Benveniste, M.P. Hasenfratz, F. Durst, Production and characterisation of monoclonal antibodies against NADPH-cytochrome P-450 reductases from helianthus tuberosus, Plant Physiol. 100 (1992) 1406–1410.

[24] E. Koopmann, K. Hahlbrock, Differentially regulated NADPH:cytochrome P450 oxidoreductases in parsley, Proc. Natl. Acad. Sci. U. S. A. 94 (1997) 14954–14959.

[25] I. Benveniste, A. Lesot, M.P. Hasenfratz, G. Kochs, F. Durst, Multiple forms of NADPH-cytochrome P450 reductase in higher plants, Biochem. Biophys. Res. Commun. 177 (1991) 105–112.
[26] M. Mizutani, D. Ohta, Two isoforms of NADPH:cytochrome P450 reductase in *Arabidopsis thaliana*. Gene structure, heterologous expression in insect cells, and differential regulation, Plant Physiol. 116 (1998) 357–367.
[27] M. Mizutani, E. Ward, D. Ohta, Cytochrome P450 superfamily in *Arabidopsis thaliana*: isolation of cDNAs, differential expression, and RFLP mapping of multiple cytochromes P450, Plant Mol. Biol. 37 (1998) 39–52.
[28] S. Jennewein, H. Park, J.M. DeJong, R.M. Long, A.P. Bollon, R.B. Croteau, Coexpression in yeast of Taxus cytochrome P450 reductase with cytochrome P450 oxygenases involved in Taxol biosynthesis, Biotechnol. Bioeng. 89 (2005) 588–598.
[29] D.K. Ro, J. Ehlting, C.J. Douglas, Cloning, functional expression, and subcellular localisation of multiple NADPH-cytochrome P450 reductases from hybrid poplar, Plant Physiol. 130 (2002) 1837–1851.
[30] D. Werck-Reichhart, A. Hehn, L. Didierjean, Cytochromes P450 for engineering herbicide tolerance, Trends Plant Sci. 5 (2000) 116–123.
[31] K.A. Feldmann, Cytochrome P450s as genes for crop improvement, Curr. Opin. Plant Biol. 4 (2001) 162–167.
[32] M.A. Larkin, G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson, D.G. Higgins, Clustal W and clustal X version 2.0 bioinformatics, Bioinformatics 23 (2007) 2947–2948.
[33] K. Tamura, G. Stecher, D. Peterson, A. Filipski, S. Kumar, MEGA6. Molecular evolutionary genetics analysis version 6.0, Mol. Biol. Evol. 30 (2013) 2725–2729.
[34] J. Felsenstein, Estimating effective population size from samples of sequences: a bootstrap Monte Carlo integration method, Genet. Res. 60 (1992) 209–220.
[35] N. Saitou, M. Nei, The neighbour-joining method: a new method for reconstructing phylogenetic trees, Mol. Biol. Evol. 4 (1987) 406–425.
[36] D. Nelson, D. Werck-Reichhart, A P450-centric view of plant evolution, Plant J. 66 (2011) 194–211.
[37] D.R. Nelson, The cytochrome p450 homepage, Hum. Genomics 4 (2009) 59–65.
[38] M. Mizutani, Impacts of diversification of cytochrome P450 on plant metabolism, Biol. Pharm. Bull. 35 (2012) 824–832.
[39] M. Mizutani, F. Sato, Unusual P450 reactions in plant secondary metabolism, Arch. Biochem. Biophys. 507 (2011) 194–203.
[40] K. Jensen, B.L. Moller, Plant NADPH-cytochrome P450 oxidoreductases, Phytochemistry 71 (2010) 132–141.
[41] W. Sitthithaworn, N. Kojima, E. Viroonchatapan, D.Y. Suh, N. Iwanami, T. Hayashi, M. Noji, K. Saito, Y. Niwa, U. Sankawa, Geranylgeranyl diphosphate synthase from *Scoparia dulcis* and *Croton sublyratus*. Plastid localisation and conversion to a farnesyl diphosphate synthase by mutagenesis, Chem. Pharm. Bull. (Tokyo) 49 (2001) 197–202.
[42] W. Sitthithaworn, J. Wungsintaweekul, T. Sirisuntipong, T. Charoonratana, Y. Ebizuka, W. De-Eknamkul, Cloning and expression of 1-deoxy-d-xylulose 5-phosphate synthase cDNA from *Croton stellatopilosus* and expression of 2C-methyl-d-erythritol 4-phosphate synthase and geranylgeranyl diphosphate

[43] synthase, key enzymes of plaunotol biosynthesis, J. Plant Physiol. 167 (2010) 292–300.
[43] L. Tian, M. Magallanes-Lundback, V. Musetti, D. DellaPenna, Functional analysis of beta- and epsilon-ring carotenoid hydroxylases in *Arabidopsis*, Plant Cell 15 (2003) 1320–1332.
[44] L. Tian, D. DellaPenna, Progress in understanding the origin and functions of carotenoid hydroxylases in plants, Arch. Biochem. Biophys. 430 (2004) 22–29.
[45] J. Kim, J.J. Smith, L. Tian, D. Dellapenna, The evolution and function of carotenoid hydroxylases in *Arabidopsis*, Plant Cell Physiol. 50 (2009) 463–479.
[46] L. Tian, V. Musetti, J. Kim, M. Magallanes-Lundback, D. DellaPenna, The *Arabidopsis* LUT1 locus encodes a member of the cytochrome p450 family that is required for carotenoid epsilon-ring hydroxylation activity, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 402–407.
[47] A.L. Stigliani, G. Giorio, C. D'Ambrosio, Characterisation of P450 carotenoid beta- and epsilon-hydroxylases of tomato and transcriptional regulation of xanthophyll biosynthesis in root, leaf, petal and fruit, Plant Cell Physiol. 52 (2011) 851–865.
[48] H. Duan, M.A. Schuler, Hetrologous expression and strategies for encapulation of membrane-localised plant P450s, Phytochem. Rev. 5 (2006) 507–523.
[49] H.J. Barnes, M.P. Arlotto, M.R. Waterman, Expression and enzymatic activity of recombinant cytochrome P450 17 alpha-hydroxylase in *Escherichia coli*, Proc. Natl. Acad. Sci. U. S. A. 88 (1991) 5597–5601.
[50] S. Zelasko, A. Palaria, A. Das, Optimisations to achieve high-level expression of cytochrome P450 proteins using *Escherichia coli* expression systems, Protein Expr. Purif. 92 (2013) 77–87.
[51] S. Chen, E. Glawischnig, K. Jorgensen, P. Naur, B. Jorgensen, C.E. Olsen, C.H. Hansen, H. Rasmussen, J.A. Pickett, B.A. Halkier, CYP79F1 and CYP79F2 have distinct functions in the biosynthesis of aliphatic glucosinolates in *Arabidopsis*, Plant J. 33 (2003) 923–937.
[52] C.H. Hansen, U. Wittstock, C.E. Olsen, A.J. Hick, J.A. Pickett, B.A. Halkier, Cytochrome p450 CYP79F1 from arabidopsis catalyses the conversion of dihomomethionine and trihomomethionine to the corresponding aldoximes in the biosynthesis of aliphatic glucosinolates, J. Biol. Chem. 276 (2001) 11078–11085.
[53] T. Hamann, B.L. Moller, Improved cloning and expression of cytochrome P450s and cytochrome P450 reductase in yeast, Protein Expr. Purif. 56 (2007) 121–127.
[54] K.B. Nicholas, H.B. Nicholas, D.W. Deerfield, GeneDoc: Analysis and Visualisation of Genetic Variation, 1997 (EMBNEW).
[55] K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei, S. Kumar, MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods, Mol. Biol. Evol. 28 (2011) 2731–2739.
[56] J. Felsenstein, Confidence limits on phylogenies: an approach using the bootstrap, Evolution 39 (1985) 783–791.
[57] D.T. Jones, W.R. Taylor, J.M. Thornton, The rapid generation of mutation data matrices from protein sequences, Comput. Appl. Biosci. 8 (1992) 275.