Archives of Biochemistry and Biophysics 477 (2008) 121–130

Contents lists available at ScienceDirect

Archives of Biochemistry and Biophysics

journal homepage: www.elsevier.com/locate/yabbi

Isolation of cDNAs and functional characterisation of two multi-product terpene synthase enzymes from sandalwood, *Santalum album* L.

Christopher G. Jones [a,c,*], Christopher I. Keeling [b], Emilio L. Ghisalberti [c], Elizabeth L. Barbour [d], Julie A. Plummer [a], Jörg Bohlmann [b]

[a] School of Plant Biology (M084), Faculty of Natural and Agricultural Sciences, University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia
[b] Michael Smith Laboratories, University of British Columbia 301-2185 East Mall, Vancouver, BC, Canada V6T 1Z4
[c] School of Biomedical, Biomolecular and Chemical Sciences (M313), Faculty of Life and Physical Sciences, University of Western Australia, 35 Stirling Hwy, Crawley, WA 6009, Australia
[d] Forest Products Commission of Western Australia, 117 Great Eastern Highway, Rivervale, WA 6103, Australia

ARTICLE INFO

Article history:
Received 14 March 2008
and in revised form 14 May 2008
Available online 27 May 2008

Keywords:
Santalum album
Terpene synthase
Germacrene D-4-ol
α-Terpineol
Sesquiterpene

ABSTRACT

Sandalwood, *Santalum album* (Santalaceae) is a small hemi-parasitic tropical tree of great economic value. Sandalwood timber contains resins and essential oils, particularly the santalols, santalenes and dozens of other minor sesquiterpenoids. These sesquiterpenoids provide the unique sandalwood fragrance. The research described in this paper set out to identify genes involved in essential oil biosynthesis, particularly terpene synthases (TPS) in *S. album*, with the long-term aim of better understanding heartwood oil production. Degenerate TPS primers amplified two genomic TPS fragments from *S. album*, one of which enabled the isolation of two TPS cDNAs, *Sa*monoTPS1 (1731 bp) and *Sa*sesquiTPS1 (1680 bp). Both translated protein sequences shared highest similarity with known TPS from grapevine (*Vitis vinifera*). Heterologous expression in *Escherichia coli* produced catalytically active proteins. *Sa*monoTPS1 was identified as a monoterpene synthase which produced a mixture of (+)-α-terpineol and (−)-limonene, along with small quantities of linalool, myrcene, (−)-α-pinene, (+)-sabinene and geraniol when assayed with geranyl diphosphate. Sesquiterpene synthase *Sa*sesquiTPS1 produced the monocyclic sesquiterpene alcohol germacrene D-4-ol and helminthogermacrene, when incubated with farnesyl diphosphate. Also present were α-bulnesene, γ-muurolene, α- and β-selinenes, as well as several other minor bicyclic compounds. Although these sesquiterpenes are present in only minute quantities in the distilled sandalwood oil, the genes and their encoded enzymes described here represent the first TPS isolated and characterised from a member of the Santalaceae plant family and they may enable the future discovery of additional TPS genes in sandalwood.

Santalum album L. is a small hemi-parasitic tree found growing in southern India, Sri Lanka, eastern Indonesia and northern Australia. The timber is highly sought after for its fine grain, high density and excellent carving properties. The fragrant wood is usually ground and steam distilled, with the essential oil serving as a fixative for many high-end perfumes. Centuries of over-exploitation has led to the demise of sandalwood in natural stands. Large plantations are being established throughout northern Australia to satisfy demand and conserve remaining reserves. *Santalum album* heartwood contains up to 6% dry wt. sesquiterpene oils [1] predominantly α- and β-santalol, α-*trans*-bergamotol and *epi*-β-santalol, along with the sesquiterpene olefins α- and β-santalene, α-bergamotene and *epi*-β-santalene, β-bisabolene, α-, β- and γ-curcumene [2]. A series of aromatic and phenolic compounds have also been identified in the heartwood of *S. album* [3]. The amount of heartwood oil produced in a tree varies considerably, even under near-identical growing conditions [4]. Causes of this yield variation are not well understood, but it is likely to be the result of both genetic and environmental factors.

Almost nothing is known about the biosynthesis of sesquiterpenoids in *S. album* or how essential oil production is regulated. Biosynthetic schemes were proposed decades ago [5] but these schemes have not been studied biochemically. Recent investigations into co-occurrence patterns of several sesquiterpenes in wood extracts suggest multiple products are formed from common carbocation intermediates [1]. Multiple product formation by single terpene synthase (TPS)[1] enzymes is well documented in the literature [6–10]. Terpenoid biosynthesis is

* Corresponding author. Address: School of Plant Biology (M084), Faculty of Natural and Agricultural Sciences, University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia. Fax: +61 8 6488 1108.
E-mail address: jonescg@cyllene.uwa.edu.au (C.G. Jones).

[1] *Abbreviations used:* TPS, terpene synthases; GPP, geranyl diphosphate; FPP, farnesyl diphosphate; TEAS, tobacco *epi*-aristolochene synthase; NPP, nerolidyl diphosphate; PVPP, polyvinylpolypyrrolidone; IPTG, Isopropyl-β-D-thiogalactopyranoside.

0003-9861/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.abb.2008.05.008

widely understood to initiate through either the mevalonic acid pathway or the deoxy-D-xylulose pathway to isoprenoid diphosphate precursors like geranyl diphosphate (GPP) and farnesyl diphosphate (FPP), and these are in turn converted by TPS to the terpenoid products typical of many plant essential oils [11]. Plant terpenoids may be involved in defence, and the production of multiple compounds from a single TPS enzyme may provide an evolutionary advantage through increased resistance against potential herbivores or pathogens [12–14]. We proposed that TPS are active in the heartwood and leaves of mature sandalwood trees, and that the diverse mono- and sesquiterpenoid profile of *S. album* likely stems from several multi-product TPS enzymes collectively involved in total oil biosynthesis. Essential oil is located in the heartwood of stems and roots of sandalwood, with biosynthesis probably occurring in parenchyma. Microscopic sections of *S. album* wood show the accumulation of oil and other extractives in ray parenchyma (Fig. 1). The sapwood–heartwood transition zone is therefore a logical place for the isolation of TPS mRNAs and cDNA cloning. We describe here the functional characterization of the first two TPS cDNAs from *S. album*.

Results

Identification of TPS gene fragments amplified from genomic DNA of S. album

Degenerate primers (primers p5F and p8R) were based on conserved regions of several published angiosperm sesqui-TPS genes; valencene synthase from *Vitis vinifera* (grapevine; AAS66358), germacrene D synthase from *Populus trichocarpa x deltoides* (poplar; AAR99061), δ-cadinene synthase from *Gossypium hirsutum* (cotton; AAD51718) and β-caryophyllene synthase from *Artemisia annua* (AAL79181). These primers amplified an intron-free 180 bp fragment from *S. album* genomic DNA. Upon cloning and sequencing of the PCR product, two independent TPS gene sequences were apparent. Based on BLAST searches against the NCBI GenBank nr. database (www.ncbi.nih.gov) the deduced amino acid sequences from these fragments both showed strong similarity to known sesquiterpene synthases from grapevine and cotton. The predicted peptide sequences of the two gDNA fragments cloned from *S. album* differed by 13 amino acids (87% identity, 100% similarity).

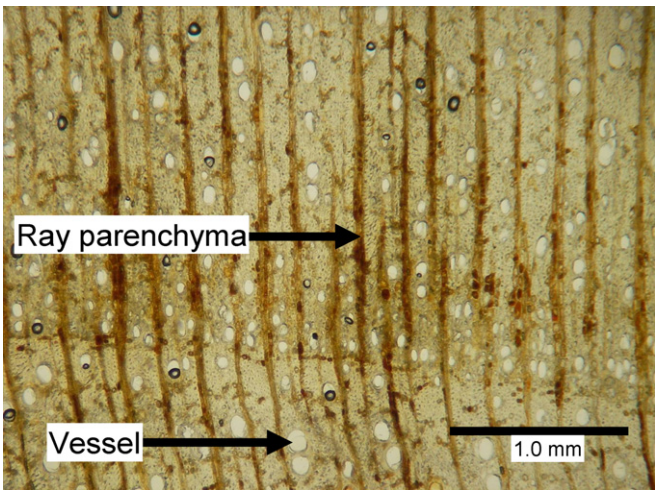

Fig. 1. Transverse section of *S. album* heartwood (lower)–sapwood (upper) transition zone (40× magnification). Essential oil and other extractives are deposited in the ray parenchyma (vertical features) during heartwood formation.

Full-length cDNA cloning of two TPS genes from S. album

The partial genomic TPS sequences enabled the design of PCR primers to be used for iterative rapid amplification of cDNA ends (RACE) which yielded the 3′- and 5′-cDNA ends and eventually the complete cDNA clones for two distinct TPS genes, *Sa*monoTPS1 (GenBank accession no. ACF24767) and *Sa*sesquiTPS1 (GenBank accession no. ACF24768) (Figs. 2 and 3). *Sa*sesquiTPS1 showed highest similarity to other angiosperm sesqui-TPS. *Sa*sesquiTPS1 was isolated from wood cDNA. Although all initial primer design was targeted at sesquiterpene synthases, the other cDNA *Sa*monoTPS1, isolated from leaf cDNA showed highest similarity to previously discovered monoterpene synthases, predominantly grapevine α-terpineol synthase. Alignments of the two *S. album* TPS deduced amino acid sequences with other angiosperm mono-TPS and sesqui-TPS are shown in Figs. 2 and 3, respectively.

*Sa*monoTPS1 features motifs typical of the monoterpene synthase gene TPS-b subfamily [15] including the aspartate rich DDXXD metal ion binding site (positions 323–327) and the RRX_8W motif (positions 33–44) which is implicated in diphosphate group migration [16]. *Sa*sesquiTPS1 also shares the $R(R/P)X_8W$ motif (positions 19–29) and the DDXXD motif. Y538 of *Sa*sesquiTPS1 resides in approximately the same position as Y520 of tobacco *epi*-aristolochene synthase (TEAS) [17,18] (Fig. 3). This residue has been shown to be responsible for protonation of germacrene A in TEAS [19].

Functional characterisation of SamonoTPS1 and SasesquiTPS1

The proteins encoded by cDNAs *Sa*monoTPS1 and *Sa*sesquiTPS1 were expressed as full-length His-tagged fusion proteins (N-terminal His_6) in *Escherichia coli*. Recombinant proteins were purified using Ni^{2+} affinity chromatography and confirmed to be of the expected size of ~66 kDa by SDS–PAGE and Western blot analysis using His_6-specific antibody.

Enzyme assays using *Sa*monoTPS1 recombinant enzyme with GPP as substrate in the presence of Mn^{2+} produced a mixture of mainly (+)-α-terpineol and (−)-limonene as well as a selection of minor monoterpene products as determined by GC–MS (Fig. 4; Table 1). Chemical ionisation assessment of α-terpineol in negative mode revealed an $[M-1]^-$ peak of $m/z = 153$, indicative of the alcohol. When (*E,E*)-FPP was used as substrate, only small amounts of one compound, β-bisabolene, was produced (data not shown). Under the chosen assay conditions, the catalytic activity and product profile of *Sa*monoTPS1 on either substrate was not substantially affected by exchanging the Mn^{2+} metal ion with Mg^{2+} (Table 1).

Enzyme assays of *Sa*sesquiTPS1 with FPP as substrate and Mg^{2+} ion containing buffer produced a complex mixture of mono- and bicyclic sesquiterpenes. The main products as analysed by GC–MS and GC-FID were the cyclic alcohol germacrene D-4-ol, and helminthogermacrene (Fig. 5 and Table 2). On-column injection with chemical ionization in negative mode using methane as the ionization gas confirmed the assignment of alcohol based upon the presence of $[M-1]^-$ of $m/z = 221$ for germacrene D-4-ol.

High GC injector temperatures resulted in a large proportion of β-elemene, and a suite of other bicyclic sesquiterpenes including δ- and γ-cadinene. Cold on-column GC–MS analysis revealed no β-elemene, and considerably less of the cadinenes. Several minor components with mass spectra and retention indices matching α- and β-selinene, α-humulene and nerolidol were detected in proportions below 5% (Fig. 5 and Table 2). In the presence of Mn^{2+}, a similar product profile was detected however the overall yield was almost 10-fold lower. When GPP was used as substrate, *Sa*sesquiTPS1 only produced very small amounts of linalool, regardless of which metal ion was used (data not shown).

Fig. 2. Alignment of *Sa*monoTPS1 amino acid sequence with other closely related monoterpene synthases; VVTPS, *Vitis vinifera* α-terpineol synthase [22]; CLLS, *Citrus limon* (−)-limonene synthase [43]; CTGS, *Cinnamomum tenuipilum* geraniol synthase [26]. Consensus indicated on bottom row. FEPQY motif, common to both *Sa*monoTPS1 and *Sa*sesquiTPS1 is indicated by an overhead line (−). Motifs common to all TPS indicated by a solid overhead line (■).

Discussion

With the plantation sandalwood industry growing rapidly in Asia and Australia, there is a substantial need for well-characterized germplasm. Knowledge and utilization of the genes involved in essential oil production may advance efforts towards sandalwood tree improvement, further development of sustainable of sandalwood plantations, and thereby conservation efforts to protect sandalwood trees in natural forests. In sandalwood trees, some environmental stimuli may induce terpenoid production, while specific genotypes of sandalwood may have a predisposition to high rates of heartwood oil deposition. Cloning of TPS involved in sandalwood oil production may enable a better molecular characterization of these genotypes and a better understanding of the molecular processes of phenotypic variability. In addition, the cloning of genes in sandalwood terpenoid biosynthesis may also enable the microbial production of sandalwood oil compounds.

Fig. 3. Alignment of *Sa*sesquiTPS1 with other closely related angiosperm sesquiterpene synthase sequences; VVGDS, *Vitis vinifera* (−)-germacrene D synthase and VVVS, *Vitis vinifera* (+)-valencene synthase [23]; PBGDS, *Populus trichocarpa x deltoides* (−)-germacrene D synthase [44]. Consensus indicated on bottom row. FEPQY motif highlighted with an overhead line (−). Double dagger indicates Y538 of *Sa*sesquiTPS1. Motifs common to all TPS indicated by a solid overhead line (■).

We describe the cDNA cloning and functional characterization of two TPS genes, *Sa*monoTPS1 and *Sa*sesquiTPS1, isolated from leaf and wood tissues of the essential oil-producing sandalwood tree *S. album*. Functional characterisation of *Sa*monoTPS1 and *Sa*sesquiTPS1 identified both enzymes as multi-product mono- and sesquiterpene synthases, respectively. The terpenoids produced by *Sa*monoTPS1 and *Sa*sesquiTPS1 are not found in substantial quantities in the distilled oil of sandalwood, suggesting that additional TPS genes are expressed in *S. album* which may contribute to the majority of heartwood essential oil. Since TPS make up large gene families [8,15] and since cloned TPS can be employed for the characterization of terpenoid modifying cytochrome P450 enzymes [20], cloning of the first TPS genes from *S. album* will enable the discovery of additional TPS as well as other metabolic pathway enzymes for sandalwood essential oil biosynthesis. With the long-term goal of sandalwood tree improvement and enhanced production of sandalwood oil, our ongoing cloning efforts are focusing on the TPS genes expressed in the heartwood–sapwood transition zone.

Based on sequence inference, *Sa*monoTPS1 is a member of the TPS-b subfamily and *Sa*sesquiTPS1 is a member of the TPS-a subfamily of the plant TPS gene family [8,15,21,22]. Both *Sa*monoTPS1 and *Sa*sesquiTPS1 share highest similarity with known mono-TPS and sesqui-TPS, respectively, from grapevine. *Sa*monoTPS1 shares 45% identity and 66% similarity with grape (−)-α-terpineol synthase [22], while sesquiterpene synthase *Sa*sesquiTPS1 shares 59% identity and 76% similarity with grapevine (+)-valencene synthase [23]. The next most closely related TPS gene sequences to

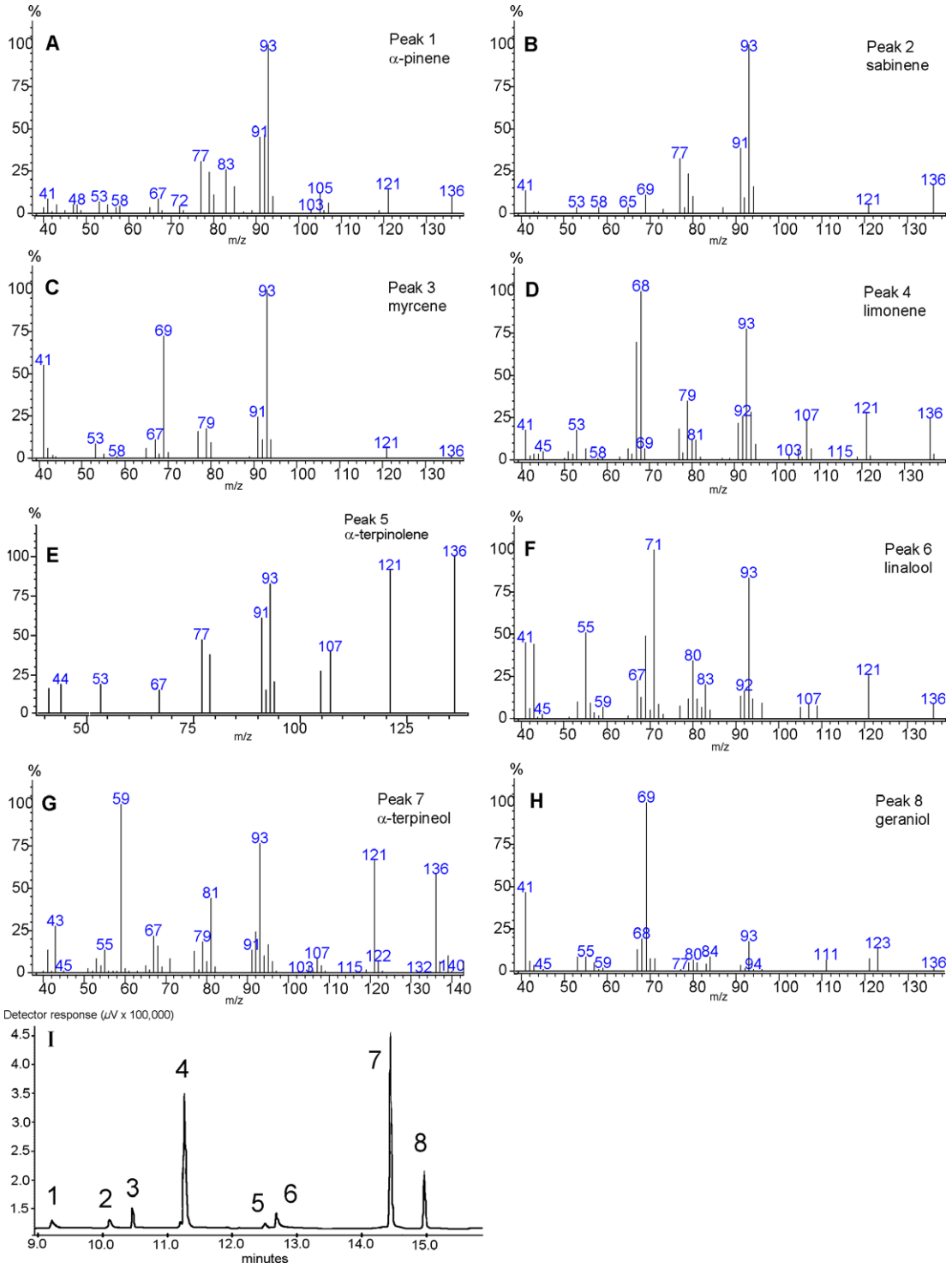

Fig. 4. Mass spectra of monoterpenes identified in the product mixture of *Sa*monoTPS1, incubated with GPP in the presence of Mn^{2+}. A, α-pinene; B, sabinene; C, myrcene; D, limonene; E, α-terpinolene; F, linalool; G, α-terpineol; H, geraniol; I, GC-FID chromatogram (DB-WAX) with peak numbering for reference.

*Sa*sesquiTPS1 are from poplar and cotton. These similarities are reasonable given the relatively close phylogenetic relationship between the Santalaceae family and the Rosid subclass of angiosperms, which includes the Vitales, Malpighiales and Malvales orders [24].

Fig. 6 outlines the likely mechanisms for transformation of GPP into monoterpenes found in the product profile of *Sa*monoTPS1. These transformations follow similar mechanisms to those described for previously characterised plant monoterpene synthases [25]. Geraniol may be produced from mild acid hydrolysis of GPP or through water quenching of the initial carbocation intermediate [26]. Intermediate (**9**) is responsible for the formation of a myriad of monoterpene compounds [27,28]. α-Terpineol may result from nucleophilic attack of (**9**) by water. The circumstances in which water is available to the active site are not known, and future efforts of enzyme structure analysis are needed to resolve this question. Some clues may come from a comparison with 1,8-cineole synthase from *Salvia fruticosa*, where water is presumed to access

Table 1
Product profile of *Sa*monoTPS1 when incubated with GPP in the presence of either Mn^{2+} or Mg^{2+} ions

	Compound	RI (HP5-MS)	Lit. RI[a] (DB-5)	RI (DB-WAX)	Composition (%)	
					Mn^{2+}	Mg^{2+}
1	(−)-α-Pinene	935	936	1013	3.3	3.3
2	(+)-Sabinene	974	973	1112	3.0	2.6
3	Myrcene	992	987	1158	3.4	5.5
4	(−)-Limonene	1030	1025	1188	33.6	35.3
5	α-Terpinolene	1098	1084	1236	1.6	1.0
6	Linalool	1100	1086	1551	5.0	3.1
7	(+)-α-Terpineol	1193	1176	1694	44.0	38.7
8	(*E*)-Geraniol	1224	1235	1849	5.9	10.5

Composition data based on GC-FID analysis using HP5-MS column.
[a] Taken from MassFinder terpenoid RI library and Adams [41]. RIs were also compared to an in-house monoterpene standard collection.

the active site through deformation of the α14 helix. Water may be H-bonded to a neighbouring Asn residue [29]. Although an Asn residue is not present in the corresponding region in *Sa*monoTPS1, a Gln residue is located slightly downstream at position 346 which may play a similar role.

Biosynthesis of the sesquiterpenes catalysed by *Sa*sesquiTPS1 can be explained by several rearrangements of a germacrenyl cation (Fig. 7). Cyclisation of FPP between C1 and C10 affords the (*E,E*)-germacradienyl cation intermediate (**1**). However, FPP may be converted to (6*E*)-nerolidyl diphosphate (NPP) prior to cyclisation between C1 and C10, leading to the (*Z,E*)-germacradienyl cation (**2**) [30]. As sesquiterpenes derived from both (*E,E*)-FPP and (6*E*)-NPP are found in the product mixture of *Sa*sesquiTPS1, the active site must accommodate both configurations. Germacrene D-4-ol may be formed through several possible mechanisms, requiring further investigation, all involving water attacking the carbocation. To our knowledge, germacrene D-4-ol as the product of a TPS enzyme has previously only been found with an engineered TPS enzyme [31].

Intermediate (**1**) may lose a proton from C12, resulting in the formation of germacrene A (**6**). No free germacrene A was found in the product mixture, even under cold on-column conditions. Under hot injector conditions, β-elemene was a major product, indicating operation of a thermocyclic Cope rearrangement of (**6**) to β-elemene. β-Elemene is rarely found in nature [32] and is generally accepted to be an artefact of high injector temperatures during GC-analysis [33]. Furthermore, under hot injector conditions there was a distinct increase in both δ and γ-cadinene, and a decrease in γ-muurolene compared to cold on-column injection. There is the possibility of a dehydration route from germacrene D-4-ol to δ- and γ-cadinene via intermediates (**3**) and (**4**) under hot GC conditions, given their absence under milder GC conditions (broken arrows in Fig. 7).

Formation of another component, α-bulnesene, can only occur through the re-ionisation of (**6**), which requires the addition of a proton to C3. The majority of germacrene A may not leave the active site, as its hydrophobic nature would prevent re-entry. Nonetheless, the role of germacrene A as an intermediate in sesquiterpene biosynthesis is well established [34]. Concerted ring closure between C2 and C6 generates a tertiary carbocation at C7 (**7**), a transformation which has been studied in tobacco *epi*-aristolochene A synthase (TEAS). Addition of a proton from Y520 of TEAS to germacrene A whilst located at the active site enables further transformations to occur via the eudesmane carbocation [19]. This may be occurring in *Sa*sesquiTPS1, however the exact location of the proton donor remains unknown. Based on the protein alignment in Fig. 3, Y538 is likely the same residue in *Sa*sesquiTPS1 and would be a potential candidate for site-directed mutational analysis to test this concept. Also present in the product mixture of *Sa*sesquiTPS1 are α- and β-selinene. Both of these compounds may be formed from free germacrene A under mildly acidic conditions [33] however enzyme catalysed conversion of (**6**) to the selinenes is possible.

In conclusion, we have described the first cloning and functional characterization of two multi-product TPS genes from sandalwood. This work is relevant for research in an important tropical tree species, as it enables the future cloning and characterization of other members of the TPS gene family and potentially members of the cytochrome P450 family involved in essential oil biosynthesis in *S. album* and related species. The results of this work may have important long-term implications for sandalwood plantations, silviculture, and tree breeding as well as for the fragrance industry.

Experimental

Chemicals and reagents

All reagents, solvents, antibiotics and precursor chemicals were purchased from commercial sources. Farnesyl diphosphate and geranyl diphosphate were from Sigma (St. Louis, MO, USA). Restriction endonucleases and T4 DNA ligase were from New England Biolabs (Ipswich, MA). Chiral standards (+) and (−)-pinenes, (+) and (−)-limonenes were purchased from Aldrich, (−)-linalool from Fluka (−)-α-terpineol from SAFC (via Sigma/Aldrich) and (−)-sabinene was purchased from Chroma-Dex (Irvine, CA).

Plant material collection, DNA and RNA extraction

Five increment cores were taken from three mature *S. album* trees growing in Perth, Western Australia. Cores went through the diameter of the trunk at approximately 30 cm above ground level. Once taken from the tree, the core was snap frozen in liquid nitrogen and stored in a cryogenic shipper. Mature and emerging leaves were also collected, frozen and stored at −150 °C. The samples were transported to UBC Vancouver, Canada for nucleic acid extraction.

Genomic DNA was extracted from leaves using the technique outlined by Byrne et al. [35]. Leaves (1 g) were ground in a chilled mortar and pestle with liquid nitrogen before being added to DNA extraction buffer (0.18 M sorbitol, 5% polyethylene glycol, 25 mM Tris pH 7.5, 2.5 mM EDTA pH 8, 0.05% BSA, 0.025% spermidine, 0.1% $NaSO_3$ and 0.5 mg l^{-1} β-mercaptoethanol). Total RNA was extracted from leaves and wood using a modified version of Kolosova et al. [36]. Leaves (3 g) were ground in liquid nitrogen and added to RNA extraction buffer (200 mM Tris–HCl, pH 8.5, 1.5% lithium dodecyl sulphate, 300 mM LiCl, 10 mM EDTA, 1% w/v sodium deoxycholate, 1% w/v Tergitol Nonidet® P-40). 5 mM Thiourea, 1 mM aurintricarboxylic acid, 10 mM dithiothreitol, and 2% (w/v) polyvinylpolypyrrolidone (PVPP) were added just prior to use. Wood cores (15 g) were ground into a fine powder in a coffee grinder with liquid nitrogen and solid CO_2 and added to extraction buffer. Several tubes were combined at the TE/NaCl resuspension step to concentrate the RNA sample. All solutions were prepared from DEPC treated and/or autoclaved water.

Primer design and genomic DNA amplification

Alignment of several previously characterised angiosperm sesquiterpene synthases enabled degenerate primer design based on several highly conserved regions, as per Chang et al. [37]. Forward primer p5F (5′-TGG TGG AAR GAY TTI GAY TTY-3′) and reverse primer p8R (5′-DGT WCC RTA CGC RTC RTA CGT RTC RTC-3′) were based on the WWK(E/D/S)LDF and DD(T/I)YDAYGT motifs, respec-

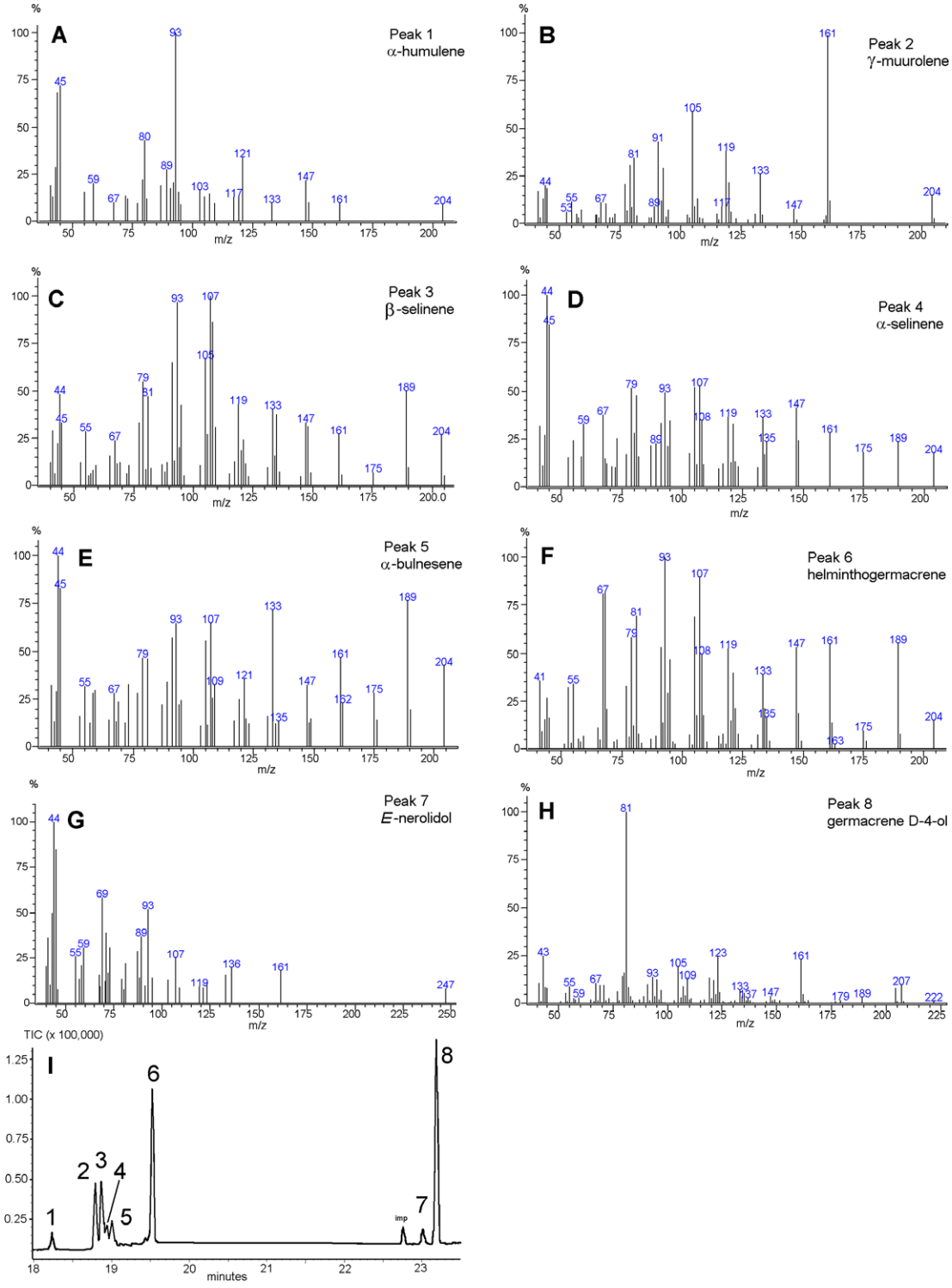

Fig. 5. Mass spectra of sesquiterpenoids identified in the product mixture of *Sa*sesquiTPS1 when incubated with FPP in the presence of Mg^{2+}. A, α-humulene; B, γ -muurolene; C, β-selinene; D, α-selinene; E, α-bulnesene; F, helminthogermacrene; G, *trans*-nerolidol; H, germacrene D-4-ol; I, GC–MS chromatogram (cold on-column injection, DB-WAX) with peak numbering for reference.

tively. All primers were purchased from IDT Pty. Ltd. (Coralville, IA). Annealing temperatures for all primers were calculated using OligoAnalyser (IDT Pty. Ltd.) and Primer3 [38]. Degenerate primers p5F and p8R (2 μM), along with a 2× mixture of thermostable *Taq* Jumpstart™ DNA polymerase (Invitrogen, Carlsbad CA) were added to 20 ng of genomic DNA in a final volume of 20 μl. The mixture was cycled using the following conditions; initial denaturing step of 4 min at 94 °C, then 35 cycles of 2 min at 94 °C, 30 s at 50 °C, and 1.5 min at 72 °C. A final extension time of 7 min completed the cycle. The PCR product was run on a 1% agarose gel in TAE buffer and visualised with SybrSafe® stain (Invitrogen). A band of 180 bp was gel purified, cloned into Topo T/A pCR 2.1 vector (Invitrogen) and sequenced using the BigDye™ terminator method. BLAST searches of the 180 bp fragment revealed 2 unique clones (from 4 colonies) with high similarity to δ-cadinene synthase from cotton.

RACE strategies

Single strand cDNA for 3′ RACE was generated from *S. album* leaf and wood total RNA using the First Choice® RLM-RACE kit (Ambion, Austin TX) as per the manufacturers instructions. 3′ RACE nested PCR was performed using universal

Table 2
GC–MS profile of *Sa*sesquiTPS1 when incubated with FPP

	Compound	RI (HP-5)	Lit. RI[a] (DB-5)	RI (DB-WAX)	Composition (%)	
					GC-FID (HP5MS)	On-column (DB-WAX)
1	β-Elemene	1400	1389	1594	37.6	–
2	α-Guaiene	1447	1440	–	1.6	–
3	α-Humulene	1465	1455	1680	2.2	2.2
4	Patchoulene?	1485	1473	1695	0.9	–
5	γ-Muurolene	1492	1474	1718	5.9	11.4
6	β-Selinene	1497	1486	1729	2.7	1.8
7	α-Selinene	1506	1494	1733	4.1	3.4
8	α-Bulnesene	1515	1503	1723	4.1[b]	10.0
9	Helminthogermacrene	1516	1503	1772	4.1[b]	29.4
10	γ-Cadinene	1527	1507	1764	1.4	–
11	δ-Cadinene	1532	1520	1768	3.8	–
12	*E*-Nerolidol	1570	1550	2037	0.7	2.5
13	Germacrene D-4-ol	1590	1574	2057	19.1	38.9

Composition determined by cold on-column GC–MS analysis as well as GC-FID (splitless injector).
[a] MassFinder terpenoid RI library and Adams [41].
[b] Overlapping peaks.

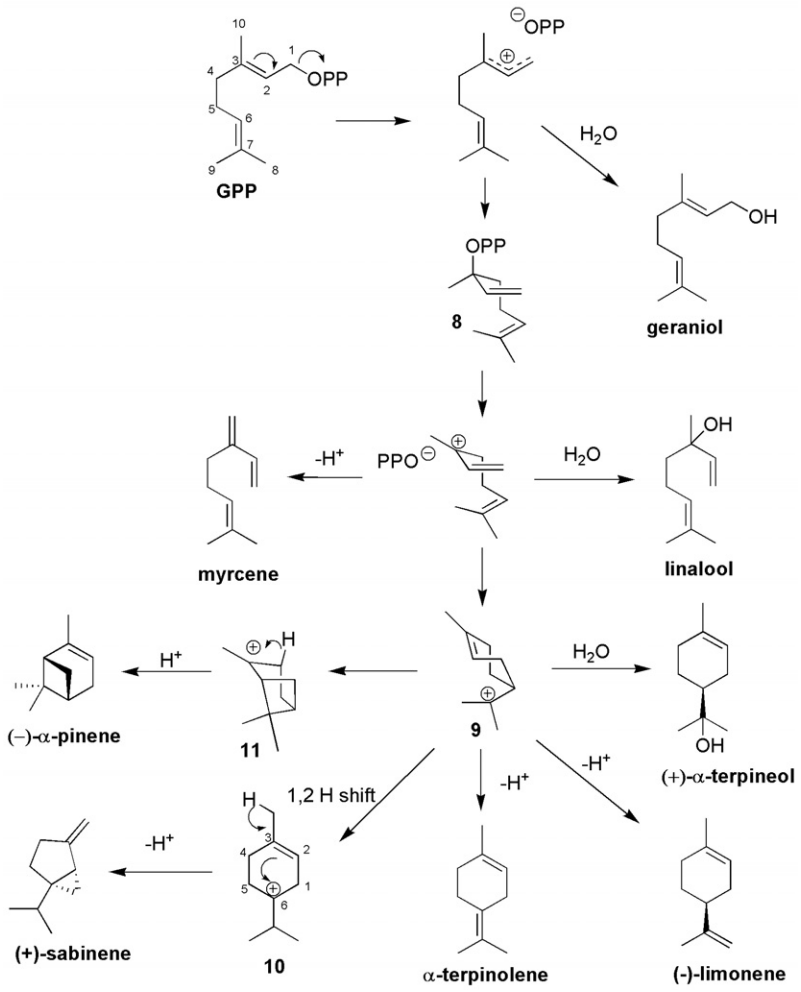

Fig. 6. Scheme showing hypothesised transformation of GPP by *Sa*monoTPS1.

3′ end anchor primers and gene-specific primers (outer primer 5′-AGA CTA GTT GAG GGT TAT TTT TGG GTA C-3′, inner primer 5′-TTG AGC CAC AGT ACT CCC TTG CAC G-3′). The 1.1 kb PCR product was gel purified, cloned into pCR 2.1 vectors and sequenced. The deduced amino acid sequence from this gene fragment closely resembled other previously published sesquiterpene synthases and was designated SesquiTPS-3. PCR program was as follows: initial denature step of 2 min at 94 °C, then 35 cycles at 94 °C for 30 s, annealing temperature of 57 °C for 30 s (increased to 60 °C in round 2) followed by 2 min extension at 72 °C.

5′ RACE was performed using SMART™ RACE (Clontech, Mountain View CA). Leaf and wood total RNA was reverse transcribed with PowerScript™ M-MLV reverse transcriptase and SMART II™ oligonucleotide as per the manufacturers instructions.

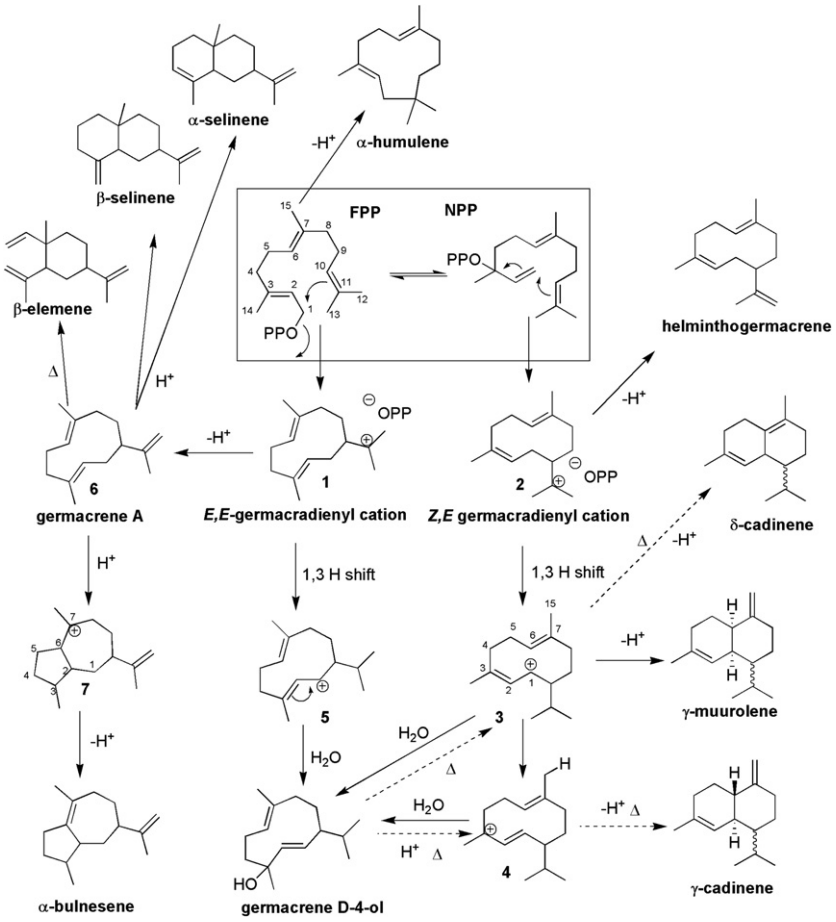

Fig. 7. Scheme showing hypothesised biosynthetic transformations of farnesyl and nerolidyl diphosphate (boxed) into compounds found in the product mixture of *Sa*sesquiTPS1. Broken arrows indicate possible heat induced dehydration and rearrangement reactions.

The resulting single strand cDNA was used as a template for nested PCR with universal 5′ end anchor primers and gene specific primers (outer primer 5′-AGC TCA TCT ACT GTA CCG TAT GCA TC-3′, inner primer 5′-CGT GCA AGG GAG TAC TGT GGC TCA AAG-3′). A BLAST search of the cloned and sequenced 1.1 kb 5′ gene fragment closely resembled published monoterpene synthases and was designated MonoTPS-5′. This fragment was probably from another gene which had serendipitously been amplified instead of the original putative sesquiterpene synthase gene. The nucleotide sequence translates to the FEPQY motif present in both *S. album* TPS genes (Figs. 2 and 3). A second pair of nested gene specific primers (outer primer 5′-CCA AGC TCC CAA CAT CCC ACC-3′, inner primer 5′-CTT TAG CGT AGT GCA TCC GAA AC-3′) were used to amplify a 1.2 kb product. The cloned and sequenced gene fragment was homologous to previously published sesquiterpene synthases, and designated SesquiTPS-5′. Concurrently, 3′ RACE using monoterpene synthase gene specific primer 5′-CAA TCA ATT CAT CAG CAG GAG CTT GGC AAC CTA GC-3′ and 3′ RACE universal anchor primers revealed a 1.2 kb product which was homologous to the 3′ end of other published angiosperm monoterpene synthase genes. This 3′ end fragment was designated MonoTPS-3′.

Full-length clones of both the putative monoterpene and sesquiterpene synthases were prepared from SMART™ RACE cDNA template. For the putative sesquiterpene synthase gene, a pair of nested start primers (outer forward primer 5′-GCT CCC GTC TAT GCT TTG AAG CAT TCC-3′, inner forward primer 5′-C ATC ATA CTG AAC TTC TCA TTT GCC CTC ATG-3′) and nested stop primers (outer reverse primer 5′-CTG AGG AAA AGC TAT AAA TCC AAC GGA TGC-3′, inner reverse primer 5′-TCAAAC AAT CAT GGG ATC TAT AAG TAG GGA TGC-3′) were used to amplify an open reading frame of 1680 bp incorporating start and stop codons (underlined). Nested PCR conditions were as follows: 94 °C for 2 min, then 35 rounds of 94 °C for 30 s, annealing for 30 s at 58 °C and extension at 68 °C. A final extension period of 7 min at 68 °C completed the cycle. Accuprime™ proofreading DNA polymerase was used in all reactions. Likewise, forward (5′-GCT TGA TCC GCC ATGGAT GCC TTT GCC ACT-3′) and reverse (5′-C TCA ATC CTC CTC GTT CAG TGG AAT AGG GTG GAT G-3′) primers were used to amplify the putative monoterpene synthase gene from leaf and wood cDNA. One round of 40 cycles of PCR (using identical conditions as the sesquiterpene synthase gene, except annealing temperature was set at 62 °C) produced the full-length gene with an open reading frame of 1731 bp. Both genes were cloned into Topo pCR II Zero Blunt™ vectors (Invitrogen).

Bacterial expression and protein isolation

Both putative terpene synthase genes were cloned into a pET28b (+) expression vector (Novagen, San Diego, CA) using sticky end PCR [39]. For the putative sesquiterpene synthase, sesquiTPS1 primers F1 (5′-TATG GAA AAT CAA AAA GTG CCT ATT TCT TCT GTC C-3′) and R1 (5′-T TCAAAC AAT CAT GGG ATC TAT AAG TAG GGA TGC AAT G-3′) were cycled 35 times with Accuprime™ proofreading DNA polymerase using the cloned full-length gene as template. PCR conditions were as follows, denature for 2 min at 94 °C, then 35 cycles of 30 s 94 °C, annealing for 30 s at 58 °C and extension for 2 min at 68 °C. For the second round, identical conditions were used with primers F2 (5′-TGG AAA ATC AAA AAG TGC CTA TTT CTT CTG TCC) and R2 (5′-AG CTT TCA AAC AAT CAT GGG ATC TAT AAG TAG GGA TGC-3′). Two micrograms of DNA from each PCR product were combined and denatured at 95 °C for 5 min, then re-annealed at 20 °C for 20 min. This mixture now contained hybridised PCR product bearing NdeI and HindIII overhangs suitable for ligation into similarly digested pET28b (+) vector. Ligation of the sticky end PCR product was done at 16 °C overnight with T4 DNA ligase. The circular plasmid, which contained the sesquiterpene synthase gene in frame with a Ni-affinity tag (His_6) was first transformed into Top 10 chemically competent cells and grown on LB plates with kanamycin (50 μg ml^{-1}).

Directional cloning of the putative monoterpene synthase gene, monoTPS1 was carried out using the same method, except primers F1 (5′-T ATG GAT GCC TTT GCC ACT TCT CCG ACC TC-3′), R1 (5′-TTC AAT CCT CCT CGT TCA GTG GAA TAG GGT GGA TG-3′) and F2 (5′-TGG ATG CCT TTG CCA CTT CTC CGA CCT C-3′), R2 (5′-AGC TTT CAA TCC TCC TCG TTC AGT GGA ATA GGG TG-3′) were used to create sticky ends suitable for ligation into NdeI and HindIII digested pET28b (+) vector. PCR conditions were the same as for the sesquiterpene synthase gene, except the annealing temperature was raised to 64 °C.

For both genes, transformants were grown on selective LB plates (kanamycin 50 μg ml^{-1}), and the DNA extracted via plasmid preparation (Invitrogen). pET28b (+) vectors containing the insert were sequence verified and transformed into chemically competent C41 *E. coli* cells (Avidis, Saint-Beauzire, France) containing the pRARE 2 plasmid isolated from Rosetta 2 competent cells (Novagen). Colonies were grown on LB plates containing kanamycin and chloramphenicol (50 μg ml^{-1}). Three independent colonies were picked and grown in a shaker overnight at 37 °C in 5 ml of LB with the same antibiotics. Two millilitres of this culture was added to

400 ml of TB containing kanamycin and chloramphenicol. The cell suspension was grown for 4 h with shaking at 37 °C until the OD_{600} = 0.8. Isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.2 mM and the mixture was shaken overnight at 16 °C. Cell suspension was harvested by centrifugation at 4 °C and the cell pellet frozen at −80 °C for future use.

Cell pellets (~7 g) were resuspended in 15 ml of cold binding buffer (20 mM HEPES pH 7.5, 150 mM NaCl and 20 mM imidazole), 1 ml of 10 mg ml^{-1} lysozyme in binding buffer and 500 μl of 200 mM PMSF. While on ice, the mixture was stirred thoroughly with a glass rod for 30 min. The lysate was then sonicated on ice using a Bronson ultrasonic probe until translucent. The lysate was centrifuged at 20,000*g* at 2 °C for 30 min. Supernatant was collected and filtered through a 0.22 μm syringe filter on ice.

Protein purification and enzyme assays

The cleared lysate (~20 ml) was loaded onto an ÄKTApurifier 10 FPLC with a 1 ml HisTrap HP column (GE Healthcare, Piscataway, NJ, USA). After washing with 60 ml binding buffer, the proton was eluted with elution buffer (20 mM HEPES pH 7.5, 150 mM NaCl and 350 mM imidazole) and then desalted on a BioRad Econo-Pac 10 DG desalting column (BioRad, Hercules, CA, USA) into elution buffer which lacked imidazole. The remaining eluent (4 ml) was spun in a 30-kDa Amicon Ultra-4 Centrifugal Filter Unit (Millipore, Billerica, MA, USA) to concentrate the protein. Aliquots of the purified proteins were frozen in liquid N_2 and stored at −80 °C until use. Proteins were analysed using SDS–PAGE, Western blotting and immunolabelling with His_6 antibody (Sigma). Antibody was visualised by the NBT/BCIP assay (Roche, USA).

Enzyme assays for both recombinant proteins were done in triplicate using GC vials as per O'Maille et al. [40]. For each assay, 10 μl of protein (25 mg ml^{-1}) was added to a reaction buffer containing 25 mM HEPES, 10% glycerol, 5 mM DTT and 10 mM of either Mg^{2+} or Mn^{2+}. Twenty-five microlitres of substrates FPP and GPP (~1 mg ml^{-1}) were added and the mixture gently shaken. Final enzyme concentration was 0.5 mg ml^{-1}. Vials were overlaid with 500 μl of pentane to trap volatile products and incubated at 30 °C for 1 h.

GC–MS analysis and product identification

Product mixtures of both enzymes were analysed on an Agilent 6890 GC with a 5973 MSD using helium as carrier gas. Two columns were used for peak identification; DB-WAX capillary column (30 m, 0.25 mm ID, 0.25 μm film thickness) and an HP5-MS capillary column (30 m, 0.25 mm ID, 0.25 mm film thickness). For the DB-WAX column, injector temperature was 240 °C, detector temperature was set at 250 °C, injection volume 1 μl and column flow rate was 1 ml min^{-1}. Oven temperature program was as follows; held for 3 min at 40 °C, then 3 °C min^{-1} to 110 °C, 10 °C min^{-1} to 180 °C, then a final ramp of 15 °C min^{-1} to 240 °C where this temperature was held for 5 min. For the HP5-MS column, conditions were the same, except the oven program was held for 1 min at 40 °C, then ramped at 7.5 °C min^{-1} to 250 °C and held. All mass spectra were detected at 70 eV in multiple ion scan mode.

Cold on-column GC–MS analysis of *Sa*sesquiTPS1 product mixture was performed on a DB-WAX column using the same instrument with the following conditions: Injector initial temperature was 20 °C (cooled below ambient with liquid CO_2) and tracked with the oven temperature program; initial 20 °C for 3 min, then ramped at 8 °C min^{-1} to 240 °C and held for 10 min. Injection volume 1 μl and column flow rate was 1 ml min^{-1}.

Percent composition of the enzyme product mixtures was confirmed by GD-FID using splitless injection of 1 μl onto a 30 m HP5-MS capillary column. Conditions were as follows: Injector 250 °C, detector 250 °C, oven initial temperature was 40 °C for 3 min, then ramped at 8 °C min^{-1} to 250 °C and held for 10 min. Column flow was 1 ml min^{-1}. Chiral GC analysis of *Sa*MonoTPS1 was by GC–MS using a 30 m Cyclodex B capillary column and splitless injection (1 μl). Injector: 230 °C, detector 230 °C, oven 55 °C for 1 min, then ramped at 10 °C min^{-1} to 230 °C, and held for 10 min. Column flow was 1 ml min^{-1}.

The identities of products from *Sa*monoTPS1 incubations were confirmed by comparison of retention times to an in-house collection of authentic chiral standards using HP5, DB-WAX and Cyclodex B stationary phases. Mass spectra were compared to the NIST 2005 library and [41]. Sesquiterpene product identification was more difficult as several compounds share identical fragmentation patterns. Since authentic standards were unavailable the product mixtures were run on both stationary phases and the retention indices recorded. These were compared to published retention indices, e.g. [41,42] and MassFinder (www.massfinder.com). All spectra were compared to the NIST 2005 spectral library. Likely fragmentation patterns of germacrene D-4-ol were interpreted, and were in agreement with the major peaks found in the spectra (Fig. 5 H) as well as those found by Yoshikuni et al. [31].

Acknowledgments

This research was funded by the Australian Research Council and the Forest Products Commission of Western Australia as part of linkage project LP0454919. Work in the laboratory of J.B. was supported in part with funds from the Natural Sciences and Engineering Research Council of Canada (NSERC). C.G.J. is grateful for scholarships received for research in the UBC Michael Smith Laboratories in Vancouver Canada, and to complete this work, particularly the DAFF Science and Innovation Award for 2005, the Mike Carroll travelling fellowship and the UWA Post Graduate Students Association travel award. We thank Britta Hamberger, Harpreet K. Sandhu and Lina Madilao, all are in the UBC Michael Smith Laboratories, for excellent technical assistance.

References

[1] C.G. Jones, E.L. Ghisalberti, J.A. Plummer, E.L. Barbour, Phytochemistry 67 (2006) 2463–2468.
[2] M.-J.R. Howes, M.S.J. Simmonds, G.C. Kite, J. Chromatogr. A 1028 (2004) 307–312.
[3] T.H. Kim, H. Ito, K. Hayashi, T. Hasegawa, T. Machiguchi, T. Yoshida, Chem. Pharm. Bull. 53 (2005) 641–644.
[4] C.G. Jones, J.A. Plummer, E.L. Barbour, J. Essent. Oil Res. 19 (2007) 157–164.
[5] W. Parker, J.S. Roberts, R. Ramage, Q. Rev. 21 (1967) 331–363.
[6] F. Deguerry, L. Pastore, S. Wu, A. Clark, J. Chappell, M. Schalk, Arch. Biochem. Biophys. 454 (2006) 123–136.
[7] C.L. Steele, J. Crock, J. Bohlmann, R. Croteau, J. Biol. Chem. 273 (1998) 2078–2089.
[8] D. Martin, J. Faldt, J. Bohlmann, Plant Physiol. 135 (2004) 1908–1927.
[9] D. Tholl, F. Chen, J. Petri, J. Gershenzon, E. Pichersky, Plant J. 42 (2005) 757–771.
[10] T.G. Kollner, P.E. O'Maille, N. Gatto, W. Boland, J. Gershenzon, J. Degenhardt, Arch. Biochem. Biophys. 448 (2006) 83–92.
[11] E.M. Davis, R. Croteau, Top. Curr. Chem. 209 (2000) 53–95.
[12] J. Gershenzon, N. Dudareva, Nat. Chem. Biol. 3 (2007) 408–414.
[13] D.P.W. Huber, S. Ralph, J. Bohlmann, J. Chem. Ecol. 30 (2004) 2399–2418.
[14] C.I. Keeling, J. Bohlmann, New Phytol. 170 (2006) 657–675.
[15] J. Bohlmann, G. Meyer-Gauen, R. Croteau, Proc. Natl. Acad. Sci. USA 95 (1998) 4126–4133.
[16] D.C. Williams, D.J. McGarvey, E.J. Katahira, R. Croteau, Biochemistry 37 (1998) 12213–12220.
[17] P. Facchini, J. Chappell, Proc. Natl. Acad. Sci. USA 89 (1992) 11088–11092.
[18] C.M. Starks, K. Back, J. Chappell, J.P. Noel, Science 277 (1997) 1815–1820.
[19] K.A. Rising, C.M. Starks, J. Noel, J. Chappell,J. Am. Chem. Soc. 122 (2000) 1861–1866.
[20] D.-K. Ro, G.-I. Arimura, S.Y.W. Lau, E. Piers, J. Bohlmann, Proc. Natl. Acad. Sci. USA 102 (2005) 8060–8065.
[21] N. Dudareva, D. Martin, C.M. Kish, N. Kolosova, N. Gorenstein, J. Faldt, B. Miller, J. Bohlmann, Plant Cell 15 (2003) 1227–1241.
[22] D. Martin, J. Bohlmann, Phytochemistry 65 (2004) 1223–1229.
[23] J. Lucker, P. Bowen, J. Bohlmann, Phytochemistry 65 (2004) 2649–2659.
[24] B. Bremer, K. Bremer, M.W. Chase, J.L. Reveal, D.E. Soltis, P.S. Soltis, P.F. Stevens, Bot. J. Linn. Soc. 141 (2003) 399–436.
[25] M.L. Wise, R. Croteau, in: D.E. Cane (Ed.), Comprehensive Natural Products Chemistry, Elseiver, Oxford, 1999, pp. 97–153.
[26] T. Yang, J. Li, H.-X. Wang, Y. Zeng, Phytochemistry 66 (2005) 285–293.
[27] J. Bohlmann, C.L. Steele, R. Croteau, J. Biol. Chem. 272 (1997) 21784–21792.
[28] D.C. Hyatt, R. Croteau, Proc. Natl. Acad. Sci. USA 104 (2007) 5360–5365.
[29] S.C. Kampranis, D. Ioannidis, A. Purvis, W. Mahrez, E. Ninga, N.A. Katerelos, S. Anssour, J.M. Dunwell, J. Degenhardt, A.M. Makris, P.W. Goodenough, C.B. Johnson, Plant Cell 19 (2007) 1994–2005.
[30] D.E. Cane, Chem. Rev. 90 (1990) 1089–1103.
[31] Y. Yoshikuni, V.J.J. Martin, T.E. Ferrin, J.D. Keasling,Chem. Biol. 13 (2006) 91–98.
[32] V. Herout, in: T.W. Goodwin (Ed.), Aspects of Terpenoid Chemistry and Biochemistry, Academic Press, London, 1971, pp. 53–94.
[33] J.-W. de Kraker, M.C.R. Franssen, A. de Groot, W.A. Konig, H.J. Bouwmeester, Plant Physiol. 117 (1998) 1381–1392.
[34] D.E. Cane, Y.S. Tsantrizos, J. Am. Chem. Soc. 118 (1996) 10037–10040.
[35] M. Byrne, B. MacDonald, J.E. Brand, Theor. Appl. Genet. 107 (2003) 1208–1214.
[36] N. Kolosova, B. Miller, S. Ralph, B. Ellis, C. Douglas, K. Ritland, J. Bohlmann, BioTechniques 36 (2004) 821–824.
[37] Y.-J. Chang, S.-H. Song, S.-H. Park, S.-U. Kim, Arch. Biochem. Biophys. 383 (2000) 178–184.
[38] S. Rozen, H.J. Skaletsky, in: S. Krawetz, S. Misener (Eds.), Bioinformatics Methods and Protocols: Methods in Molecular Biology, Humana Press, Totowa, NJ, 2000, pp. 365–386.
[39] G. Zeng, BioTechniques 25 (1998) 206–208.
[40] P.E. O'Maille, J. Chappell, J.P. Noela, Anal. Biochem. 335 (2004) 210–217.
[41] R.P. Adams, Identification of Essential Oil Components by Gas Chromatography/Mass Spectrometry, seconnd ed., Allured Publishing Corporation, Carol Stream, IL, 1995.
[42] J. Pala-Paul, M.J. Perez-Alonso, A. Velasco-Negueruela, J. Vadare, A.M. Villa, J. Sanz, J.J. Brophy, J. Chromatogr. A 1074 (2005) 235–239.
[43] J. Lucker, M.K. El Tamer, W. Schwab, F.W.A. Verstappen, L.H.W. van der Plas, H.J. Bouwmeester, H.A. Verhoeven, Eur. J. Biochem. 269 (2002) 3160–3171.
[44] G.-I. Arimura, D.P.W. Huber, J. Bohlmann, Plant J. 37 (2004) 603–616.