Phytochemistry 101 (2014) 52–64

Contents lists available at ScienceDirect

Phytochemistry

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

Characterization of two genes for the biosynthesis of abietane-type diterpenes in rosemary (*Rosmarinus officinalis*) glandular trichomes

Kathleen Brückner [a], Dragana Božić [b,1], David Manzano [c,d], Dimitra Papaefthimiou [b], Irini Pateraki [c,d,2], Ulschan Scheler [a], Albert Ferrer [c,d], Ric C.H. de Vos [e,f], Angelos K. Kanellis [b], Alain Tissier [a,*]

[a] Leibniz Institute of Plant Biochemistry, Department of Cell and Metabolic Biology, Weinberg 3, 06120 Halle-Saale, Germany
[b] Group of Biotechnology of Pharmaceutical Plants, Laboratory of Pharmacognosy, Department of Pharmaceutical Sciences, Aristotle University of Thessaloniki, 541 24 Thessaloniki, Greece
[c] Department of Molecular Genetics, Centre for Research in Agricultural Genomics (CSIC-IRTA-UAB-UB), Bellaterra-Cerdanyola del Vallés, 08193 Barcelona, Spain
[d] Department of Biochemistry and Molecular Biology, Faculty of Pharmacy, University of Barcelona, 08028 Barcelona, Spain
[e] Plant Research International, Wageningen University and Research Centre, The Netherlands
[f] Netherlands Metabolomics Centre, The Netherlands

ARTICLE INFO

Article history:
Received 12 November 2013
Received in revised form 15 January 2014
Available online 22 February 2014

Keywords:
Rosmarinus officinalis
Lamiaceae
Phenolic diterpenes
Carnosic acid
Miltiradiene
Diterpene synthases

ABSTRACT

Rosemary (*Rosmarinus officinalis*) produces the phenolic diterpenes carnosic acid and carnosol, which, in addition to their general antioxidant activities, have recently been suggested as potential ingredients for the prevention and treatment of neurodegenerative diseases. Little is known about the biosynthesis of these diterpenes. Here we show that the biosynthesis of phenolic diterpenes in rosemary predominantly takes place in the glandular trichomes of young leaves, and used this feature to identify the first committed steps. Thus, a copalyl diphosphate synthase (*RoCPS1*) and two kaurene synthase-like (*RoKSL1* and *RoKSL2*) encoding genes were identified and characterized. Expression in yeast (*Saccharomyces cerevisiae*) and *Nicotiana benthamiana* demonstrate that RoCPS1 converts geranylgeranyl diphosphate (GGDP) to copalyl diphosphate (CDP) of normal stereochemistry and that both RoKSL1 and RoKSL2 use normal CDP to produce an abietane diterpene. Comparison to the already characterized diterpene synthase from *Salvia miltiorrhiza* (SmKSL) demonstrates that the product of RoKSL1 and RoKSL2 is miltiradiene. Expression analysis supports a major contributing role for RoKSL2. Like SmKSL and the sclareol synthase from *Salvia sclarea*, RoKSL1/2 are diterpene synthases of the TPS-e group which have lost the internal gamma-domain. Furthermore, phylogenetic analysis indicates that RoKSL1 and RoKSL2 belong to a distinct group of KSL enzymes involved in specialized metabolism which most likely emerged before the dicot-monocot split.

Introduction

Rosemary (*Rosmarinus officinalis*) is a Mediterranean shrub of the Lamiaceae family, which has been used for a long time as food spice and folk medicine. More recently, thanks to the anti-oxidative properties of some of its constituents, rosemary extracts have been commercialized as food and cosmetics preservatives (Yanishlieva et al., 2006). Furthermore, some of these compounds are now investigated for their anti-inflammatory and anti-cancer activities (Ngo et al., 2011; Yanishlieva et al., 2006). The principal components which are responsible for the antioxidant activity of rosemary, as well as of several *Salvia* species, belong to the class of phenolic diterpenes. Carnosic acid (CA) and carnosol were shown to be the major phenolic diterpene constituents in leaves of both *R. officinalis* and *Salvia officinalis* (Schwarz and Ternes, 1992), but have not been found in their essential oils. Recent investigations have demonstrated that these compounds exhibit biological activity that may be useful for the treatment of neurodegenerative diseases (Fischedick et al., 2013; Satoh et al., 2008). CA and carnosol were able to activate the Keap1/Nrf2/ARE pathway *in vitro* thereby protecting cortical neurons against oxidative cell death. The potential of phenolic diterpenes from *R. officinalis* as ingredients for the treatment or prevention of neurodegenerative diseases thus makes them an attractive target

* Corresponding author. Tel.: +49 (0) 345 55821500; fax: +49 34555821009.
E-mail address: Alain.Tissier@ipb-halle.de (A. Tissier).
[1] Present address: Institute for Biological Research "Siniša Stanković", University of Belgrade, 142 Despot Stefan Boulevard, 11000 Belgrade, Serbia.
[2] Present address: Department of Plant Biology and Biotechnology, Faculty of Life Sciences, University of Copenhagen, Thorvaldsensvej 40, 1871 Frederiksberg C, Copenhagen, Denmark.

http://dx.doi.org/10.1016/j.phytochem.2014.01.021
0031-9422/

for developing strategies aimed at producing more efficiently these high-value plant compounds. This could be done by breeding (Tounekti and Munne-Bosch, 2012) or metabolic engineering, which both would require the elucidation of the biosynthesis pathway of these compounds.

CA and carnosol belong to the large family of labdane-related diterpenoids which include among many others the precursor of gibberellins, *ent*-kaurene, the rice diterpenoid phytoalexins and the diterpene resin acids in conifers (Hedden and Thomas, 2012; Keeling and Bohlmann, 2006; Toyomasu, 2008). Plant diterpenoids are synthesized from two C_5 building blocks, isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), which, because of the plastidial localization of diterpene synthases, normally originate from the plastidic 1-deoxyxylulose-5-phosphate (DXP) pathway. However, a contribution of the cytosolic mevalonate (MVA) pathway is not to be excluded, given the increasing number of reports showing the existence of cross-talk between the two pathways (Bick and Lange, 2003; Yang et al., 2012). The 20-carbon atom backbone of diterpenoids is provided by a prenyltransferase, geranylgeranyl diphosphate synthase (GGDPS), which sequentially combines one DMAPP and three IPP units in a head-to-tail fashion. Although no biosynthesis pathway for the phenolic diterpenes of rosemary has yet been proposed, one reasonable intermediate is abietatriene, a tricyclic abietane diterpene with an aromatized C-ring. Double hydroxylation of abietatriene on the C-ring would produce the typical catechol group, and multiple oxidations on C-20 would provide the carboxyl group present in CA. Subsequent hydroxylation at C-7 followed by internal lactonization would then result in the formation of carnosol. Since the cyclization of GGDP by olefin-producing diterpene synthases typically results in products with a parent mass of 272 rather than 270, it seems unlikely that abietatriene (molecular mass 270) is a direct product of such diterpene synthases. Thus, the precursor of abietatriene remains to be determined. The cyclization of GGDP into labdane-type diterpenes is a two-step process. The first is a protonation initiated cyclization, catalyzed by class II terpene synthases resulting in the production of copalyl diphosphate (CDP). Several stereoisomers of CDP are found, among which the most common is *ent*-CDP, the precursor of gibberellins. Other isomers are *syn*-CDP and normal-CDP, the latter being utilized for the biosynthesis of diterpene resin acids in Gymnosperms and abietane-type tanshinones in *Salvia miltiorrhiza* (Gao et al., 2009; Peters, 2010). The second step of the cyclization is initiated through the formation of a carbocation upon ionization of the diphosphate linkage of CDP. This is catalyzed by class I terpene synthases which can transform CDP into a variety of polycyclic diterpenoids. Several class I diterpene synthases, whose substrate is one of the stereoisomers of CDP, have been identified in higher plants, including the *ent*-kaurene synthase of the gibberellin pathway and enzymes of specialized metabolism, such as the diTPS involved in the biosynthesis of momilactones or phytocassanes in rice (Chen et al., 2011; Peters, 2006; Toyomasu, 2008). In Gymnosperms and lower plants, bifunctional TPS with both class I and class II activities are thought to represent an ancestral form of the plant terpene synthases (Hall et al., 2013; Hayashi et al., 2006; Peters et al., 2000). Dissociation of the class I and class II activities probably occurred concurrently with the appearance of gibberellins, which are present only in vascular plants (Chen et al., 2011). All Gymnosperm and Angiosperm terpene synthases share a common structure derived from a three-domain (alpha three-domain (alpha 'α' beta 'β' and gamma 'γ') ancestral TPS such as the bifunctional CPS/KS enzyme of the moss *Physcomitrella patens* (Cao et al., 2010; Hillwig et al., 2011; Zerbe et al., 2013). In class I TPS of the Angiosperms, selective loss of the 'γ'-domain required for class II TPS activity together with the domain, occurred several times during evolution. The most recent event appears to be the emergence of the Angiosperm-specific TPS-a clade and contains mostly sesqui- and diterpene synthases (Chen et al., 2011). Class I TPS that can take CDPs as substrates belong to the TPS-e class of enzymes, which include the *ent*-kaurene synthase of the gibberellin pathway and largely retained the three-domain structure. Some enzymes of the TPS-e class, however, were found to have lost their 'γ'-domain independently in Monocots and Dicots. In Monocots, these are the wheat and maize sesquiterpene synthases ZmTPS1 and TaKSL5, respectively (Hillwig et al., 2011; Schnee et al., 2002). In Dicots, "truncated" TPS-e enzymes have been found so far exclusively in the Lamiaceae family with two reported examples, a miltiradiene synthase from *S. miltiorrhiza* and a sclareol synthase from *Salvia sclarea* (Caniard et al., 2012; Gao et al., 2009).

By analogy to what has been found for other labdane-type diterpenoids, we hypothesized that the biosynthesis of the diterpene precursor of the CA pathway in rosemary should proceed in two sequential steps catalyzed by two distinct enzymes. The first, a class II TPS enzyme should yield normal CDP, which would then be taken by a KSL enzyme to yield an abietane diterpene (Fig. 1). Here we describe the cloning and characterization of three genes from *R. officinalis*, *RoCPS1*, *RoKSL1* and *RoKSL2*, and show that RoCPS1 in combination with either RoKSL1 or RoKSL2 produce miltiradiene from GGDP.

Results

Identification of CA producing plant organs

To localize the production site of CA in rosemary, young and old *R. officinalis* leaves as well as isolated trichomes were analyzed for their phenolic diterpene (PD) content. LC–MS analysis showed CA and carnosol to be the major PDs produced in rosemary leaves (Fig. 2A). CA and carnosol were the most abundant in young leaves (CA, 10.5 mg/g fr. wt; carnosol, 6.7 mg/g fr. wt) and decreased in older leaves. On a cell-specific level, trichomes appeared to accumulate CA more than other leaf tissues. While both CA and carnosol accumulated inside leaves, isolated rosemary trichomes primarily contained CA (Fig. 2B). In addition to the PDs indicated, 12-*O*-methylcarnosic acid could be detected in trace amounts in both leaves and trichomes.

Identification and cloning of candidate genes encoding terpene synthases

Following the hypothesis that the first steps of the CA biosynthesis should occur via two steps involving CPS and KSL enzymes, a search for rosemary genes encoding enzymes of these families was undertaken. BLAST searches on available EST databases generated from *R. officinalis* trichomes (http://www.terpmed.eu/) and leaves (http://medicinalplantgenomics.msu.edu/) revealed the existence of partial cDNA sequences coding for a CPS and a KSL enzymes. A full-length sequence coding for a closely related KSL enzyme was also found. Nucleotide sequences were aligned to the related Lamiaceae *SmCPS* [GenBank ABV57835] and *SmKSL* [GenBank ABV08817] from *S. miltiorrhiza* (Gao et al., 2009) in order to design degenerate primers on consensus sequences for the partial cDNA sequences. All possible combinations of degenerated primers were used to amplify cDNA fragments by RT-PCR with total RNA extracted from rosemary leaves. Single sequences were found for the CPS and the KSL products, and RACE PCR was used to amplify the full length coding sequences. The genes, from which these products were amplified, were named *RoCPS1* and *RoKSL1*. In addition, specific primers were designed to amplify full-length cDNA of the second KSL by RT-PCR from total RNA extracted from leaves. This gene was named *RoKSL2*.

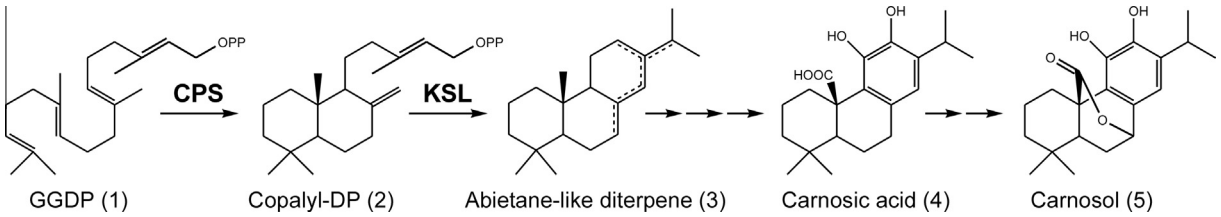

Fig. 1. Hypothetical biosynthetic pathway of carnosic acid (4) and carnosol (5) produced in *Rosmarinus officinalis*. Based on the universal diterpenoid precursor geranylgeranyl diphosphate (GGDP) (1), the first steps of carnosic acid biosynthesis are proposed to proceed via a copalyl diphosphate intermediate (copalyl-DP) (2) provided by copalyl diphosphate synthase (CPS), and followed by the production of an abietane-like diterpene (3) catalyzed by a kaurene synthase-like enzyme (KSL).

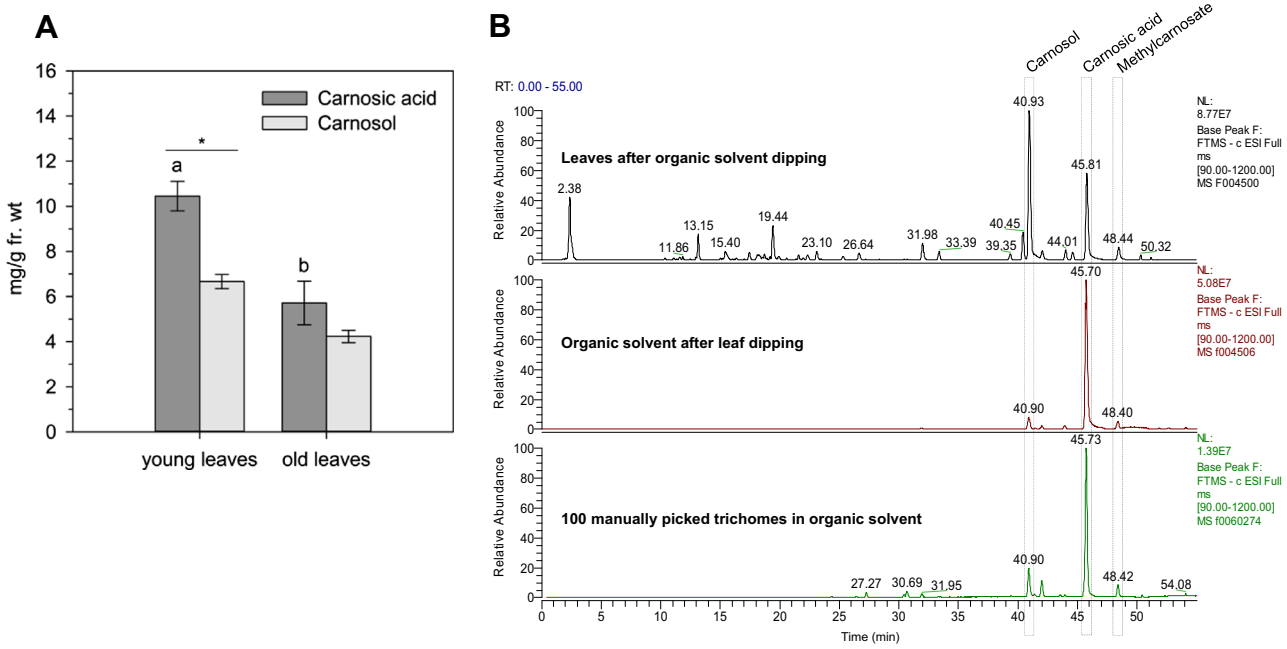

Fig. 2. Major phenolic diterpenes present in leaves and trichomes of *R. officinalis* analyzed by LC-MS. (A) PD content in young and old leaves of *R. officinalis* var. *Alderney*. Columns represent mean PD content from three independent biological samples ± s.e. A two-way ANOVA statistical analysis was applied to the data and resulted in significant differences between the mean CA and carnosol values in young leaves ($P < 0.05$, indicated by an asterisk) as well as in significant differences in the mean CA values regarding the leaf age ($P < 0.05$, as indicated by different letters a and b). (B) Comparative analysis of extracts from leaves after organic solvent dipping, the organic solvent after leaf dipping and extracts from 100 manually picked trichomes.

RoCPS1 and RoKSL2 expression is elevated in trichomes from young leaves

To further validate the isolated rosemary *CPS* and *KSL* genes as candidates, their specific expression in trichomes and leaves of different developmental ages was evaluated. First, quantitative RT-PCR on whole leaves, showed that *RoCPS1*, *RoKSL2* and *RoKSL1* are preferentially expressed in young leaves, with *RoKSL1* being much less expressed than either *RoCPS1* or *RoKSL2* (Fig. S1, Method S1). Next, quantitative RT-PCR assays were used to compare the expression level of *RoCPS1* and of both *RoKSL* genes in the trichomes to that in leaves from which trichomes were completely removed (Fig. 3). This revealed that both *RoCPS1* and *RoKSL2* exhibit preferential expression in trichomes from young leaves, with trichome/leaf expression ratios of 93 and 3972, respectively. The findings are in agreement with our results showing that CA accumulates to trichomes (Fig. 2B). Furthermore, a direct comparison of the expression levels between *RoKSL1* and *RoKSL2* in isolated trichomes and in the remaining leaves indicates that *RoKSL1* is not specific to the trichomes and significantly less expressed than *RoKSL2* (*RoKSL1* transcript level $\leqslant$1% of expression of *RoKSL2*).

Sequence and phylogenetic analysis of RoCPS1, RoKSL1 and RoKSL2

Based on the isolated full length amino acid sequences of RoCPS1, RoKSL1 and RoKSL2, a phylogenetic tree was generated (Fig. 4). RoCPS1 appears to be closely related to the labda-13-en-8-ol diphosphate (or 8-hydroxy-copalyl diphosphate) synthase from *S. sclarea* (Caniard et al., 2012; Schalk et al., 2012) and clusters further with SmCPS which is known to produce CDP of normal stereochemistry in *S. miltiorrhiza* (Gao et al., 2009). Based on the stereochemistry of the PDs from rosemary, RoCPS1 and SmCPS are expected to have the same enzymatic activity. Noticeably, the Lamiaceae CPS enzymes constitute together with NtCPS2 from tobacco and CcCLS from *Cistus creticus*, which both catalyze the formation of labda-13-en-8-ol diphosphate a clade distinct from the other Angiosperm CPS enzymes (Falara et al., 2010; Sallaud et al., 2012). In a multiple alignment, protein sequences corresponding to the enzymes in that dicot CPS cluster display conserved regions over the entire length of the proteins (see Supplemental Fig. S2).

Likewise, RoKSL1 and RoKSL2 cluster into a discrete clade, together with other KSL proteins from the Lamiaceae and Solanaceae (NtABS, ShSBS and SlPHS) (Sallaud et al., 2009, 2012; Schilmiller

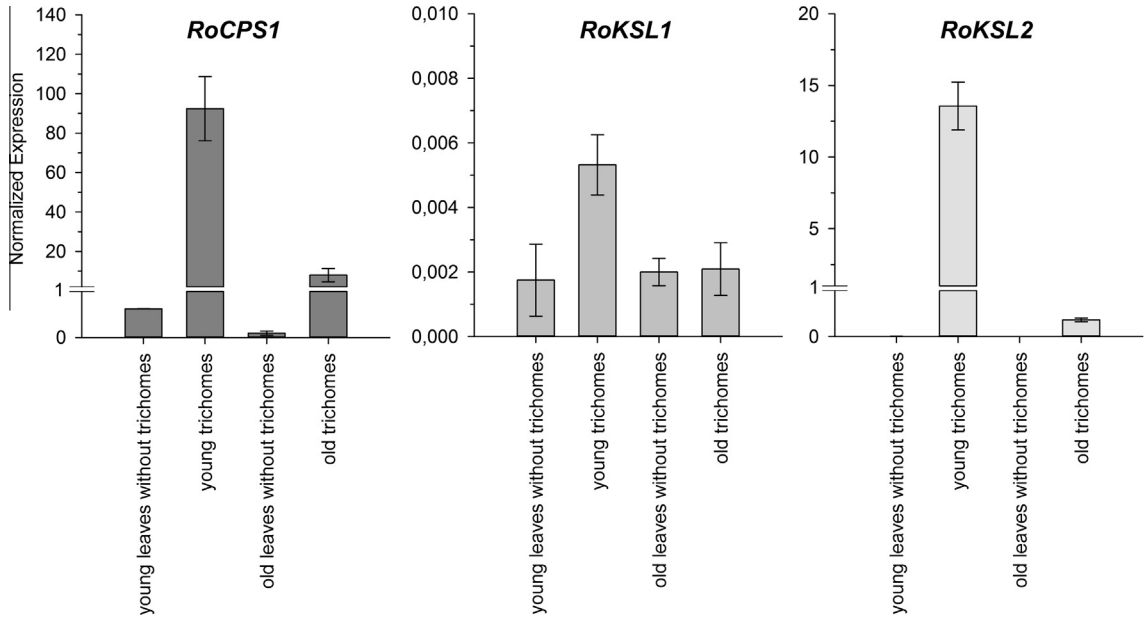

Fig. 3. Expression of *RoCPS1*, *RoKSL1* and *RoKSL2* genes in isolated rosemary trichomes and leaves from which trichomes were completely removed. The trichomes were abraded either from young or from aged rosemary leaves. Columns represent mean expression values from three technical replicates ± s.d. Transcript levels were normalized to eukaryotic elongation factor 4a (*elf*4a).

et al., 2009). In this branch, the Lamiaceae sequences form a subgroup, with RoKSL1 being more related to SmKSL (miltiradiene synthase from *S. miltiorrhiza*) than RoKSL2. The sclareol synthase from *S. sclarea* (SsSS) also belongs to this same branch, although it is more distantly related to SmKSL than RoKSL1/2. Interestingly, like SmKSL and SsSS, both RoKSLs are shorter than the other KS and KSL proteins by approximately 210 amino acids (see Fig. S3), corresponding to the loss of the 'γ'-domain.

Functional characterization of RoCPS1, RoKSL1 and RoKSL2 in planta

To test the activity of the RoCPS1 enzyme, its coding sequence was cloned into a T-DNA expression vector under control of the CaMV-35S promoter. The gene was transiently expressed in leaves of *Nicotiana benthamiana* via agro-infiltration, which after 5 days were extracted by dipping in hexane and extracts were analyzed by GC–MS using a method we recently developed (Brückner and Tissier, 2013). A novel peak was detected in tissue expressing RoCPS1 but not in extracts from leaves agro-infiltrated with the empty vector as a control (Fig. 5A). A direct comparison of its mass spectrum with previously published mass spectra of a dephosphorylated *ent*-CPP standard and the SmCPS-produced CDP (Gao et al., 2009) indicated that the compound produced by RoCPS1 is a copalyl diphosphate (Fig. 5B). The RoCPS1 produced CDP seemed to undergo endogenous dephosphorylation reactions by enzymes in *N. benthamiana* which resulted in the formation of a dephosphorylated derivative similar to copalol (290 Da) (Li et al., 2012).

To probe the function of RoKSL1 and RoKSL2, their complete cDNAs were cloned into plant T-DNA expression vectors and transformed into *Agrobacterium tumefaciens*. *N. benthamiana* transient expression of RoCPS1 with RoKSL1 and of RoCPS1 with RoKSL2 resulted in the production of two novel compounds (Fig. 6A). While mass signals (Fig. 6B) extracted for peak 1 displayed high similarity to the published mass spectra for miltiradiene (Gao et al., 2009; Guo et al., 2013), peak 2 could be identified as abietatriene based on similarity searches with the NIST mass spectral library (match 83.7%), and to a published mass spectrum (Kitadani et al., 1970). To further support the hypothesis that RoKSL1 and RoKSL2 function as miltiradiene synthases, the RoCPS1 was co-expressed in *N. benthamiana* with the previously characterized miltiradiene synthase (SmKSL) from *S. miltiorrhiza*. A direct comparison of retention time and mass spectrum recorded for miltiradiene in assays with RoCPS1 and SmKSL to those with RoCPS1 and RoKSL1 or RoCPS1 and RoKSL2 confirmed that miltiradiene is produced by the enzymatic activity of RoKSL1 or RoKSL2 using a copalyl diphosphate substrate synthesized by RoCPS1. Since the SmKSL was shown to use only CDP of normal stereochemistry as substrate (neither *ent*- nor *syn*-CDP) (Gao et al., 2009), the CDP substrate formed by RoCPS1 is also most likely of normal stereochemistry.

Reaction runs with RoKSL1, RoKSL2 or SmKSL alone did not result in any detectable products by GC–MS (see Supplemental Fig. S4). The transient assay was further applied to the quantification of miltiradiene (Fig. 6D). Whatever the miltiradiene synthase that was co-expressed with RoCPS1 in *N. benthamiana*, yields of ~100 ng/cm^2 miltiradiene and of ~30 ng/cm^2 abietatriene could be reached. We also investigated whether miltiradiene and abietatriene are present in leaf exudates of *R. officinalis*. Rosemary leaves of var. *Alderney* were extracted by dipping in hexane and analyzed by GC–MS (Fig. 6C). Both miltiradiene and abietatriene could be detected, strongly supporting that they are the first committed intermediates in CA biosynthesis.

Further examination of the GC–MS chromatograms revealed that additional minor peaks could be detected in samples where RoCPS1 and KSL were co-expressed (see Supplemental Fig. S5). Although the identity of these products was not yet defined, the corresponding mass spectra indicated that they were monohydroxylated derivatives of miltiradiene and of abietatriene and were likely the products of endogenous *N. benthamiana* enzymes.

Heterologous expression of RoCPS1 and RoKSL1 in S. cerevisiae

Because heterologous expression of the RoCPS1, RoKSL1 and RoKSL2 proteins in *Escherichia coli* failed (data not shown), we expressed them in an engineered yeast strain (AM113) that provides an increased supply of GGDP, a necessary substrate for RoCPS1, by overexpressing heterologous GGDPS1 from *C. creticus* (Pateraki and

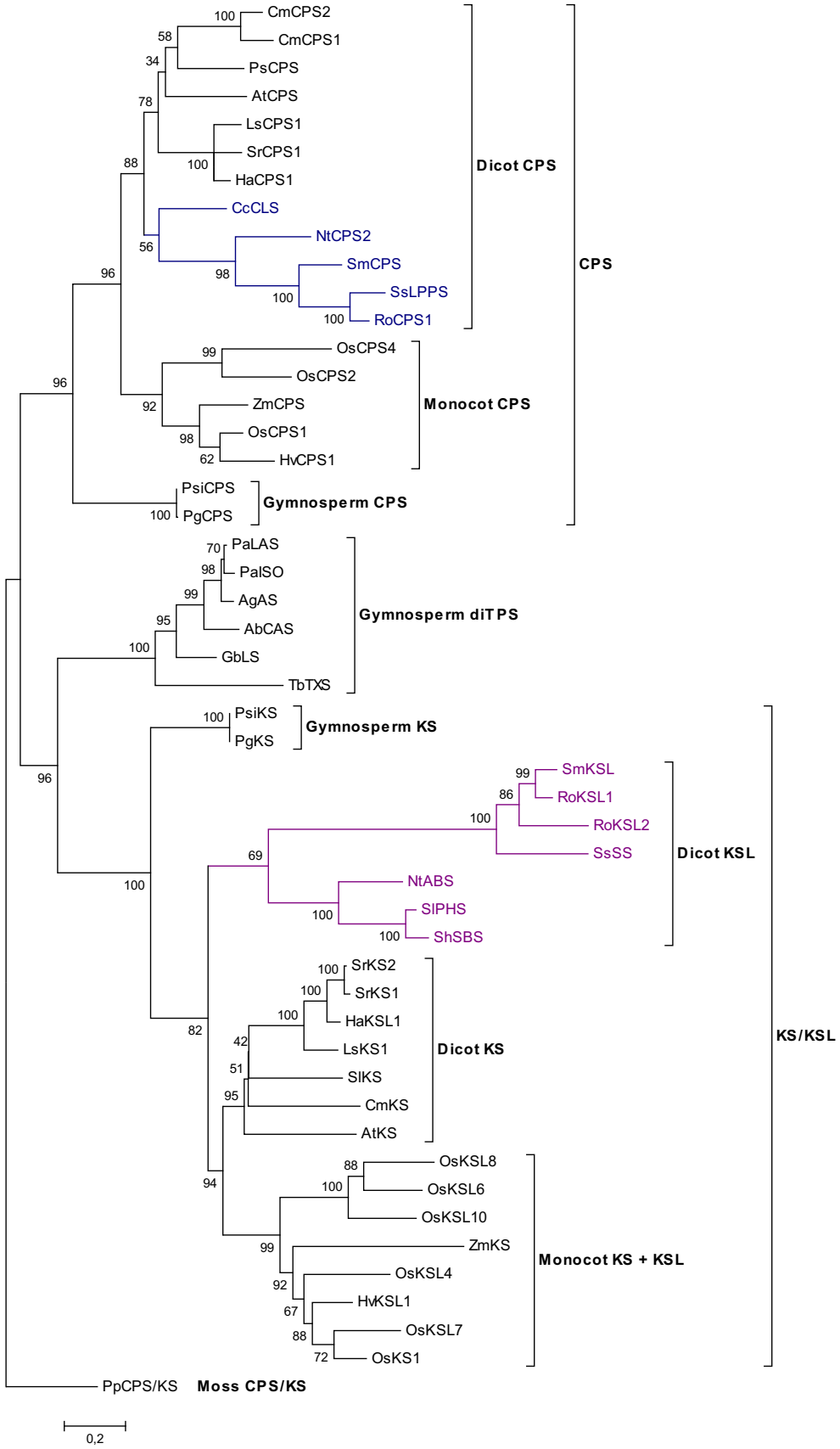

Fig. 4. Phylogenetic analysis of CPS and KSL enzymes from *Rosmarinus officinalis*. The Maximum likelihood tree illustrates the phylogenetic relationship of rosemary CPS and KSL candidates relatively to diterpene synthases from various plant species. The *Physcomitrella patens* bifunctional *ent*-copalyl-diphosphate/*ent*-kaurene synthase (PpCPS/KS) has been selected as tree root.

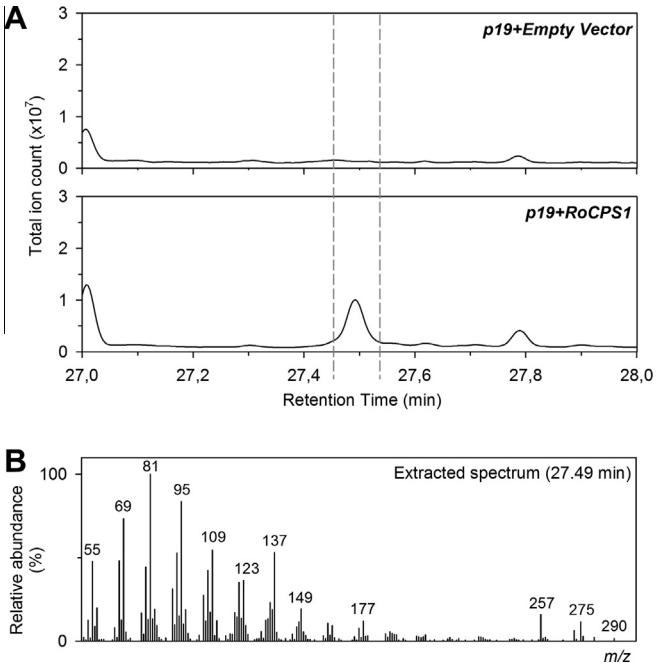

Fig. 5. RoCPS1 enzymatic activity. (A) GC–MS analysis of extracts derived from leaves co-infiltrated with *p19* and *RoCPS1*. Leaves transformed with *p19* and the empty T-DNA vector were used as control. Samples were taken five days post-infiltration. (B) Extracted mass spectrum for the novel peak at 27.49 min, which is indicated by dashed lines in the total ion chromatogram.

Kanellis, 2008). GC–MS analysis of the hexane extracts from recombinant yeast cultures revealed that miltiradiene could be detected as the product and only in strains that contained the full-length versions of RoKSL1 (Fig. 7). On the other hand, both full length and mature versions of the RoCPS1 gene seem to be functional, providing copalyl diphosphate as a substrate for RoKSL1. Besides miltiradiene, another compound was detected in hexane extracts of recombinant yeast cultures whose mass spectrum matched that of abietatriene.

Substrate specificity of RoKSL1 and RoKSL2

Because 8-hydroxy-copalyl diphosphate (8-OH-CDP) has the same stereochemistry as normal CDP, we wanted to see if RoKSL1/2 could accept it as substrate. Thus, the previously isolated tobacco gene *NtCPS2* which synthesizes 8-OH-CDP (Sallaud et al., 2012) was transiently co-expressed with either *RoKSL1*, *RoKSL2* or *SmKSL* in *N. benthamiana* leaves. Hexane extracts were prepared from treated leaves and analyzed by GC–MS. As expected, neither miltiradiene nor abietatriene could be detected. The analysis revealed a novel product that appears upon co-expression of NtCPS2 and RoKSL1 and of NtCPS2 and RoKSL2 (Fig. 8). The compound was identified as manoyl oxide based on mass spectra comparison to the NIST library. Expression of NtCPS2 alone resulted in the production of 13(E)-labdene-8,15-diol (Fig. 8), as expected from the *in planta* dephosphorylation of 8-OH-CDP (Sallaud et al., 2012). Also, as expected, co-expression of NtCPS2 and the *cis*-abienol synthase from tobacco (NtABS) resulted in the production of Z-abienol (Fig. 8), but no Z-abienol could be found after co-expression of NtCPS2 with RoKSL1/2 or SmKSL. Thus, the results indicate that rosemary RoKSL1/2 are functional enzymes able to convert either 8-OH-CDP or CDP into manoyl oxide or miltiradiene respectively, and that they likely function *in planta* as miltiradiene synthases.

Discussion

Phenolic diterpene constituents from *R. officinalis* have attracted interest because of their use as preservatives in the food and cosmetic industries, as well as potential drugs in the pharmaceutical industry. CA and carnosol are the most abundant abietane diterpenes found in rosemary leaves (Munne-Bosch and Alegre, 2001; Schwarz and Ternes, 1992) and have been shown to have anti-bacterial and neuroprotective properties (Fischedick et al., 2013; Jordan et al., 2012). There is therefore an increasing demand and interest in this class of compounds. Elucidation of their biosynthesis pathway is a prerequisite to implement breeding strategies to improve product yield *in planta* or for metabolic engineering in other hosts, plants or microorganisms, which would allow higher yield and productivity. To this day, however, and to the best of our knowledge, there is no report on the biosynthesis of PDs in rosemary. We here report the functional characterization of two sequentially acting diterpene synthases, RoCPS1 and RoKSL1/2, which catalyze the conversion of GGDP into miltiradiene, likely the first committed intermediate in the pathway of CA and related compounds.

To elucidate the site of CA accumulation, we measured the relative amounts of phenolic diterpenes in leaves of different developmental stages and in trichomes, isolated either manually or by gentle washing in an organic solvent. LC–MS analysis showed that CA is preferentially stored in the glandular trichomes of young leaves, while carnosol is stored in other leaf cells as well (Fig. 2). CA may give rise to carnosol after enzymatic dehydrogenation, which suggests that carnosol may be synthesized in the trichomes and then transported into the leaf, or alternatively that CA is transported to the leaf and then converted into carnosol. Nonetheless, glandular trichomes are likely the main organs that contribute to the biosynthesis of CA. Therefore, to identify genes involved in the CA biosynthesis, next generation sequencing was performed with RNA from trichomes. From our EST database and another publicly available NGS resource for rosemary, one CPS and two KSL candidates were identified and further characterized. From the candidates identified through these criteria, we were able to demonstrate that RoCPS1 readily converts GGDP into CDP of, in all likelihood, normal stereochemistry, as attested by the production of miltiradiene when *RoCPS1* is co-expressed with *SmKSL*, which is specific for normal CDP as substrate (Gao et al., 2009). RoCPS1 is similar (63.7% identity of amino acid sequence) to the previously characterized SmCPS, which also makes normal CDP (Gao et al., 2009; Peters and Croteau, 2002). Interestingly, phylogenetic analysis indicates that RoCPS1 belongs to a distinct clade of dicotyledon Angiosperm CPS sequences which includes other CPS enzymes involved in specialized diterpene metabolism with all enzymes producing CDP, or derivatives thereof, of the normal stereochemistry (see Fig. 4). Some of these enzymes, such as the tobacco NtCPS2, the Cistus CcCLS or the clary sage SsLPPS, produce 8-OH-CDP, also of normal stereochemistry. This clustering, although no definitive proof, reinforces the notion that RoCPS1 produces normal CDP and suggests that RoCPS1 and the related CPS enzymes arose via duplication from an ancestral CPS gene likely involved in general gibberellin metabolism and acquired a novel function. The topology of the phylogenetic tree suggests that this clade of CPS enzymes dedicated to specialized metabolism emerged after the dicot-monocot split. This indicates that specialized labdane diterpene metabolism has evolved independently in dicots and monocots. Furthermore, the fact that CPS and KSL enzymes of specialized metabolism in dicots show a different evolutionary history is intriguing. According to our phylogenetic tree, KSL enzymes of the dicots would have first emerged (even before the monocot/dicot split), subsequently followed by the appearance of

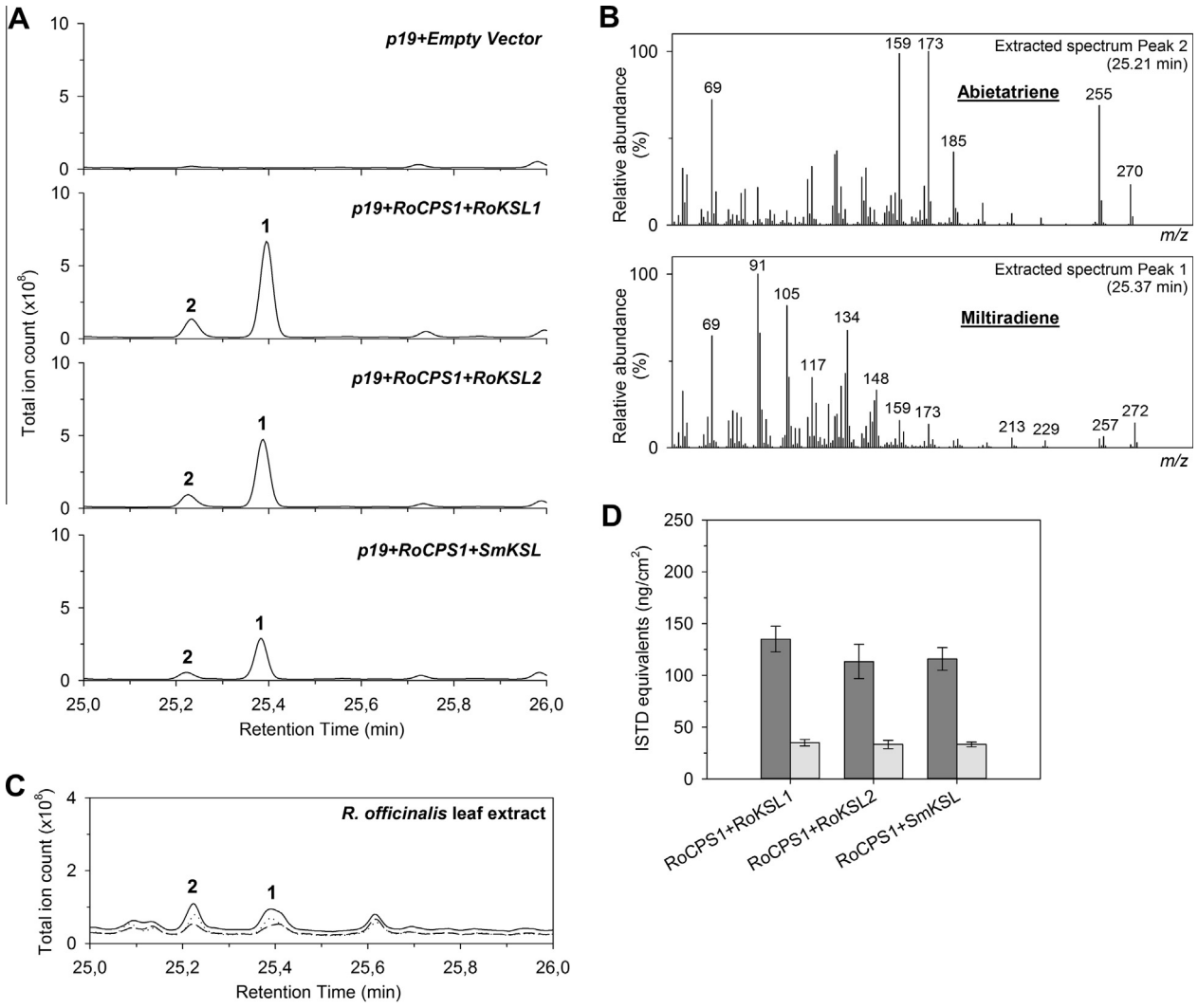

Fig. 6. Production of miltiradiene. (A) Total ion chromatograms of surface extracts of *N. benthamiana* leaves agro-infiltrated with *RoCPS1* and either *RoKSL1*, *RoKSL2* or *SmKSL*. Leaves transformed with *p19* and the empty T-DNA vector were used as control and did not contain any miltiradiene or abietatriene. Peaks (1) and (2) correspond to miltiradiene and abietatriene, respectively. (B) Extracted mass spectrum for miltiradiene (peak 1) and for abietatriene (peak 2). (C) Detection of miltiradiene and abietatriene in extracts prepared from leaves of three independent *R. officinalis* var. *Alderney* plants using hexane dipping procedure. (D) Amount of miltiradiene (dark grey bars) and abietatriene (light grey bars) in agro-infiltrated *N. benthamiana* leaves. Samples were taken five days post-infiltration from two individual *N. benthamiana* plants, each of them having two completely infiltrated leaves. Mean values ± s.e. are presented (n = 16, two experiments). Sclareol was used as internal standard (ISTD) for quantification.

specialized CPS enzymes in dicots. This would suggest that the ancestral specialized KSL enzymes of the dicots initially had access to *ent*-CDP, the precursor of the gibberellins, and that in the course of evolution novel CDP substrates (based on normal CDP) appeared through the emergence of a diversified CPS class. It is also noteworthy that the labdane-type diterpenes that have been identified in monocots (essentially in rice) come from either *ent*- or *syn*-CDP, and not from normal CDP (Morrone et al., 2011). However, in wheat, normal CDP synthase (TaCPS2) as well as normal CDP metabolizing KSL enzymes (TaKSL1/2/4) were identified, although no natural metabolites derived from normal CDP have been isolated from that monocot species (Wu et al., 2012; Zhou et al., 2012).

The subsequent RoKSL catalyzed class I reaction generates the diterpenoid product miltiradiene (Fig. 9). Interestingly, we have found two functional miltiradiene synthases encoded by differentially regulated genes. While *RoKSL2* transcript levels were high in young leaves and their corresponding trichomes, *RoKSL1* was poorly expressed regardless of the plant tissue and developmental stages tested (Fig. 3). Thus, *RoKSL2* is likely to be the main contributor to PD biosynthesis in glandular trichomes. *RoKSL1* expression could be regulated differentially, for example in response to abiotic or biotic stress conditions that are known to influence CA concentration in rosemary plants (Tounekti and Munne-Bosch, 2012). The combined expression of RoCPS1 and RoKSL1/RoKSL2 in *N. benthamiana* or in yeast gave the same product profile with miltiradiene being the major compound. Expression of RoKSL1/RoKSL2 in yeast was successful only with the full length version. The reason for this may be improper folding of the protein in the absence of the plastid targeting transit peptide, or that the transit peptide was not correctly predicted resulting in a non-functional protein. However, when the transit peptide of RoKSL1 was fused to the truncated version of SmKSL which is functional in *E. coli* (Gao et al., 2009) the resulting chimeric protein was also functional in *N. benthamiana*, indicating that the transit peptide was correctly predicted. Thus, the hypothesis of improper folding in the absence of transit peptide is more likely to be correct. As for the CPS, the analysis presented in Fig. 4 shows that RoKSL enzymes form a distinct clade of the KSL family that also includes KSL proteins that are involved in synthesis of specialized terpenoids in

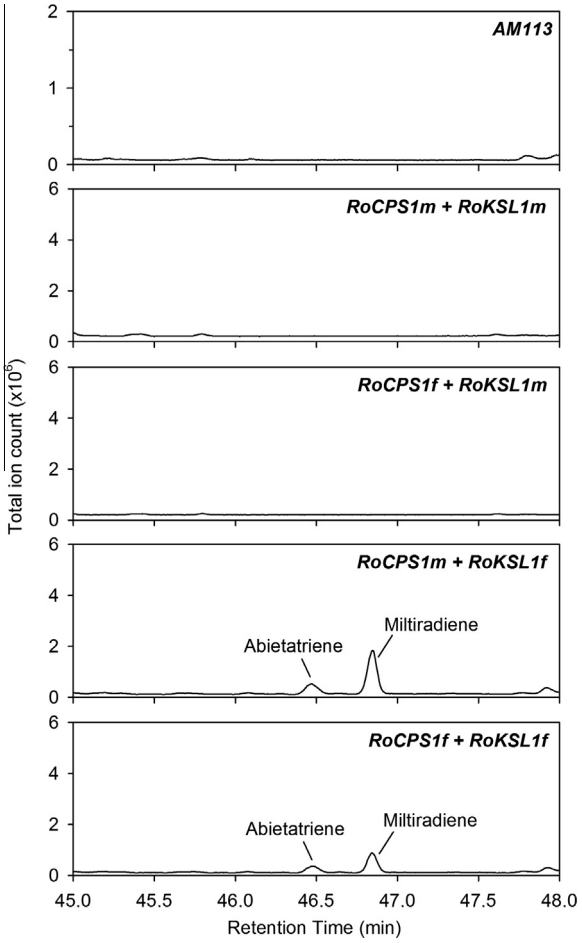

Fig. 7. Heterologous expression of *RoCPS1* and *RoKSL1* in yeast strain AM113. Total ion chromatograms showing miltiradiene and abietatriene secreted into the culture media of yeast co-expressing RoCPS1m + RoKSL1m, RoCPS1f + RoKSL1m, RoCPS1m + RoKSL1f and RoCPS1f + RoKSL1f. The AM113 yeast strain alone was used as non-transformed control. *f* – full length protein, *m* – mature protein.

Solanaceae. As observed earlier (Sallaud et al., 2012), and in contrast to the CPS case, this KSL clade apparently evolved early in Angiosperm evolution, prior to the dicot–monocot split. The fact that KSL enzymes of specialized metabolism in monocots co-cluster with the *ent*-KS of gibberellin metabolism (see Fig. 4), indicates that diversification of KS function occurred several times during land plant evolution. Both RoKSL enzymes appear closely related to the recently characterized miltiradiene synthase (SmKSL) from *S. miltiorrhiza* (Gao et al., 2009), as it could be expected from our results confirming the same function for RoKSL. This high sequence similarity (see Supplemental Fig. S3) between SmKSL and RoKSL1/2 combined with the same product profile suggests that RoKSL1/2 also share the same substrate specificity, namely for normal CDP, with SmKSL. However, this will need to be tested experimentally in further experiments. These enzymes together with the sclareol synthase from *S. sclarea* (SsSS) form a sub-branch of this clade of KSL enzymes of specialized metabolism, suggesting a Lamiaceae specific evolution. Remarkably, all the Lamiaceae KSL have lost the internal 'γ'-domain. Because Lamiaceae are known to produce a wide range of labdanoid diterpenes (Alvarenga et al., 2001), it will be interesting to see whether or not this loss is a conserved feature of such KSL proteins involved in specialized metabolism in the Lamiaceae family.

In addition to miltiradiene, abietatriene could also be detected in *N. benthamiana* leaves co-expressing *RoCPS1* and *RoKSL*. Both compounds were recovered in leaf exudate of *R. officinalis* plants as well (Fig. 6). In *in vitro* assays with SmKSL only miltiradiene was recovered (Gao et al., 2009 and Reuben Peters, personal communication). Since expression of SmKSL in *N. benthamiana* also results in co-production of abietatriene and miltiradiene, this implies that abietatriene is formed *in planta* by a spontaneous oxidation rather than by the enzymatic activity of KSL. Recently, it was shown that miltiradiene can spontaneously oxidize to abietatriene (Zi and Peters, 2013). Furthermore, the recently identified CYP76AH oxygenases from *S. miltiorrhiza* and rosemary, which produce ferruginol (Guo et al., 2013; Zi and Peters, 2013), were shown to act only on abietatriene and not on miltiradiene (Zi and Peters, 2013).

Terpene synthases carry out complex reactions thereby creating diverse structures from simple isoprenoid precursors. Most mechanistic studies of terpene synthases so far have focused on how these enzymes specify their product outcome. Recent reports have demonstrated that specificity can be switched by changes in a minimum number of amino acid residues (Kampranis et al., 2007; Xu et al., 2007). In the present paper, we report that RoKSL and SmKSL exhibit dual reactivity with normal CDP and 8-hydroxy-CDP, to yield miltiradiene and manoyl oxide, respectively. Dual substrate specificity has been reported for members of the *Oryza sativa* KSL family, in particular of OsKSL4, OsKSL11 and OsKSL10, as well as for wheat KSL enzymes which are able to utilize CDPs of different stereochemistry (Morrone et al., 2011; Zhou et al., 2012). In contrast, the class I activity of the bifunctional miltiradiene synthase from *Selaginella moellendorffi* is specific for normal CDP (Sugai et al., 2011). Further phylogenetic and sequence analysis revealed that RoKSLs and SmKSL proteins have ~60% identity of amino acid sequence with sclareol synthase from *S. sclarea* (SsSS), which catalyzes the production of sclareol but also trace amounts of manoyl oxide (Caniard et al., 2012). Conversely, the activity of SsSS with normal CDP was not tested (Caniard et al., 2012), and it is not possible to state if this substrate promiscuity is a general feature of Lamiaceae KSL enzymes. It is nonetheless tempting to postulate that such promiscuity would allow altering the profile of diterpenes produced by simply turning on or off the expression of specific CPS enzymes.

In conclusion, the functional characterization of RoCPS1 and RoKSLs reported here has revealed a novel CPS producing normal CDP and novel KSL enzymes subsequently acting as miltiradiene synthases. The dual substrate reactivity of the miltiradiene synthases provides scope for further research on structure–function mechanisms of terpene synthases. Furthermore, based on the current prevailing view on the biosynthesis of tanshinones in *S. miltiorrhiza* (Guo et al., 2013), miltiradiene in rosemary may undergo similar transformations en route to CA. Finally, our results provide the basis for investigation of the critical downstream CA biosynthetic enzymes and the first marker genes for the selection of cultivars of rosemary with improved phenolic diterpenes productivity.

Experimental

Plant material

Adult plants of different *R. officinalis* ecotypes were obtained from the Conservatoire National des Plantes à Parfum, Médicinales, Aromatiques et Industrielles (http://www.cnpmai.net/). Plants were grown on soil in greenhouse with a minimum of 16 h light. Air conditioning in the greenhouse allowed adjusting the temperature to 25 °C and the humidity to 65%. The *N. benthamiana* plants used for *in planta* transient expression in this study were grown from seeds on soil in climate-controlled chambers with a

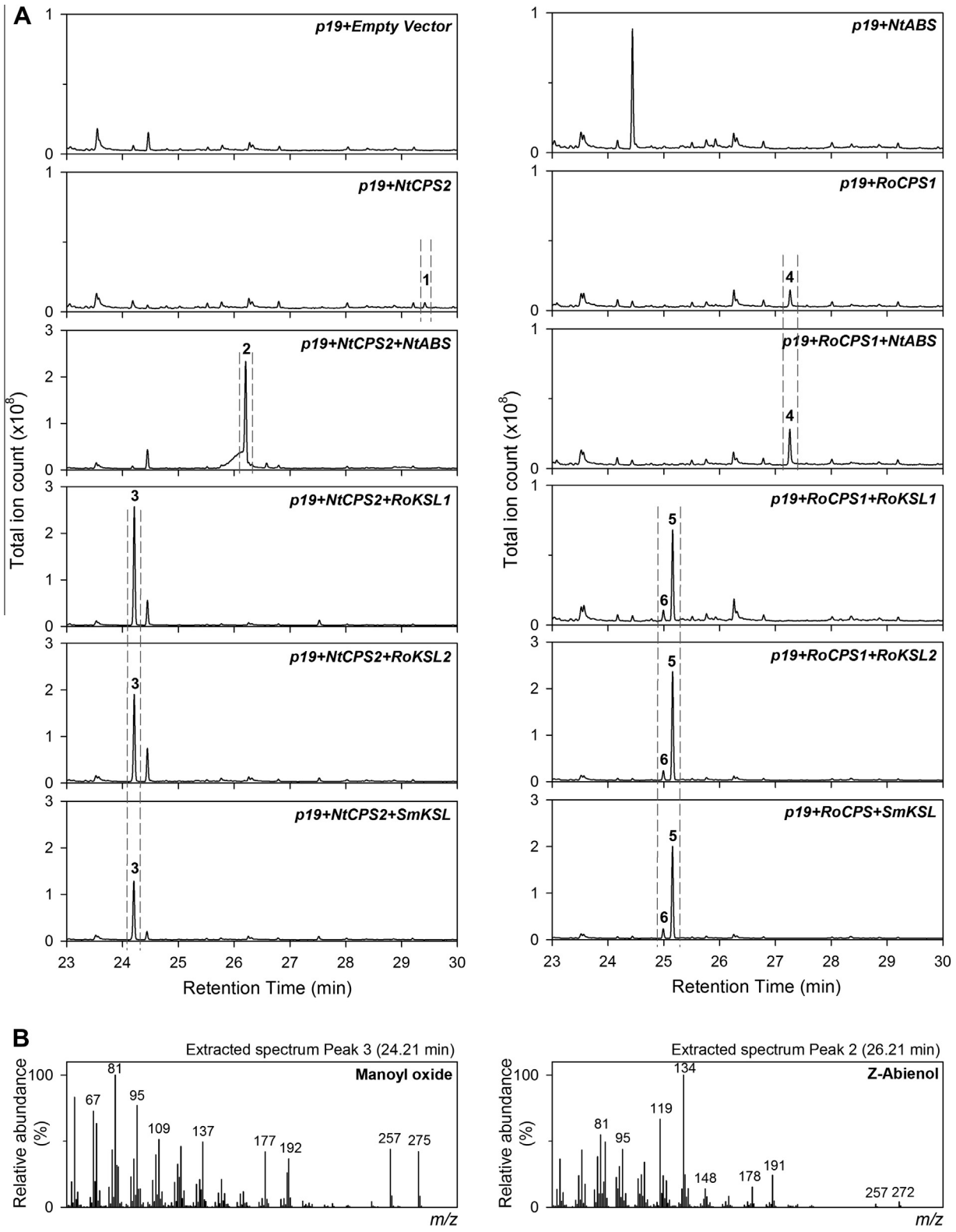

Fig. 8. Product formation of RoKSL enzymes with 8-hydroxy-copalyl diphosphate as substrate. (A) Total ion chromatograms of hexane extracts from *N. benthamiana* leaves transiently co-infiltrated with *p19* and *NtCPS2*, *NtABS*, *RoCPS1*, *RoKSL1*, *RoKSL2* and *SmKSL* in different combinations. Samples were taken from three fully infiltrated leaves per *N. benthamiana* plant. The *in planta* transient expression experiment was performed twice. 1 – 13(E)-labdene-8,15-diol, 2 – Z-Abienol, 3 – Manoyl oxide, 4 – *in planta* dephosphorylated product of the RoCPS1-produced CDP (copalol), 5 – Miltiradiene, 6 – Abietatriene; (B) Experimental mass spectrum recorded for manoyl oxide and for Z-abienol.

16-h day/8-h night photoperiod under an illumination of 90 μmol/s/m^2 (MT250L metal halide lamps). Growth temperatures were adjusted to 20 °C during the day and to 18 °C in the night. Humidity was set to 60%.

Oligonucleotide primers

Sequences of oligonucleotide primers that were used for isolation of cDNA, cloning and RACE-PCR reactions are listed in

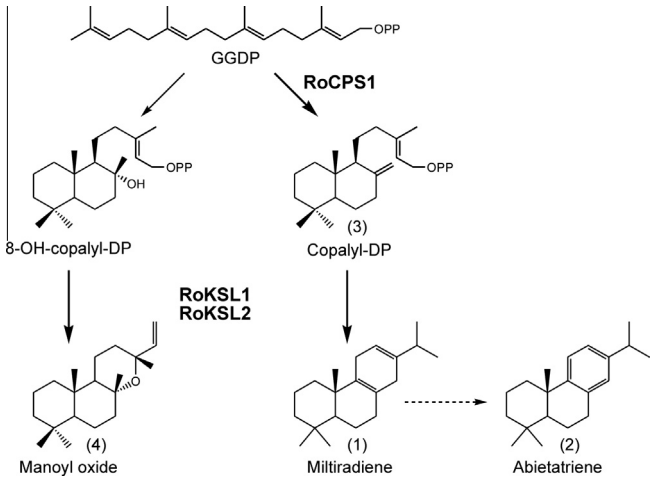

Fig. 9. Enzymatic reactions catalyzed by CPS1 and KSLs from *Rosmarinus officinalis*. RoKSL1 and RoKSL2 are functional miltiradiene synthases which catalyze the production of miltiradiene (1) from a copalyl diphosphate substrate (3). The substrate is provided by RoCPS1 enzyme that consumes GGDP. Abietatriene is proposed to arise non-enzymatically from miltiradiene.

Supplemental Table S1. The cloning primers were designed by hand with artificial *BpiI* restriction sites at the 5′ and 3′ end.

Trichome isolation

R. officinalis was grown in soil at the Aristotle University farm, where they were kept inside a greenhouse at a 16 h day/8 h photoperiod and temperatures of 24 °C day and 18 °C night. About 5–10 g of fresh leaves from two developmental stages were collected, young stage 2 (leaf length 1–2 cm) and old stage 4 (leaf length 3–4 cm) from several independently growing plants. Leaves were placed in liquid nitrogen immediately after harvest and were stored at −80 °C until further analysis.

For trichome isolation, the previously developed dry-ice abrasion method (Yerger et al., 1992) was used with slight modifications. Liquid nitrogen was applied to collect all trichomes abraded from the leaf surface and remaining within the sides of a 50 ml plastic tube. Complete trichome removal of the abraded rosemary leaves was achieved under a light microscope by using a synthetic brush to sweep off trichomes from leaves while constantly immersed in liquid nitrogen. Finally, these leaves were instantly immersed in cold pure ethanol to make sure that all trichomes were removed from their both abaxial and adaxial sides. After several consequent repeats of brushing off trichomes, completely trichome-free leaves were obtained. All tissues were immediately kept at minus 80 °C until further processing. For subsequent RNA extraction trichomes and leaves without trichomes were separately homogenized with liquid nitrogen using mortar and pestle.

Trichome specific gene expression analysis

High quality total RNA was isolated from trichomes using the commercial Spectrum™ Plant Total RNA Kit (Sigma–Aldrich), according to the manufacturers' specifications with additional ethanol washes. The cDNA was constructed from 1 μg of total RNA using the Superscript® III Reverse Transcriptase (Invitrogen) according to manufacturers' instructions with random primers.

For gene expression analysis, primer pairs were designed by hand to provide the best possible distinction of candidate gene EST contigs. Oligonucleotides for real-time PCR were as follows: RoCPS1 (Roff.N01.C023303, http://www.terpmed.eu/), 166 bp: *forward* 5′-TGCTGGTTCTGGCGACGTATGG-3′, *reverse* 5′-ACTAAC GCTGCTTCTGTGATACC-3′; RoKSL1 (Roff.N01.C024308), 333 bp: *forward* 5′-AGAGAATAGATGAAGAAGAGAGCAG-3′, *reverse* 5′-TTTTACCTACTTTGTCATACCTCTG-3′; RoKSL2 (Roff.N01.C014863/65/67/69), 183 bp: *forward* 5′-AGCGATGTCCCTTTGTCGAACC-3′, *reverse* 5′-ACGTCCTTTGCCTTGAACTTCTC-3′. The eukaryotic elongation factor 4a (*elf4a*) was selected as endogenous reference for normalization of transcript levels, and PCR amplified using *forward* 5′-AAGCCTTCCGCCATCCAACAGC-3′ and reverse 5′-ATCCTCACCCA-CACTTTTGCCTC-3′ primers. Samples were prepared using the KAPA SYBR® FAST qPCR Kit (Kapa Biosystems), and quantitative real-time PCR analysis was performed in technical triplicates using the Applied Biosystems® 7500 Real-Time PCR System (Life technologies). Expression of transgenes in isolated trichomes was calculated relative to leaves without trichomes based on an efficiency-corrected delta–delta Ct method (Livak and Schmittgen, 2001).

Isolation and cloning of candidate genes from R. officinalis

RNA was extracted from rosemary leaves using TRIzol® reagent (Ambion) according to the manufacturer's specifications, and cDNA was synthesized from 1 μg of total RNA using the RevertAid Premium First Strand cDNA Synthesis Kit (Fermentas, http://www.thermoscientificbio.com/fermentas/). Partial coding sequences for the candidates were PCR amplified with Phusion™ High Fidelity DNA Polymerase (Thermo Scientific, http://www.thermoscientificbio.com/) using degenerated primers, and PCR products were purified using QIAquick PCR Purification Kit (Qiagen, http://www.qiagen.com/) before sequencing. RACE-PCR was used for amplification of cDNA ends following a self-made protocol similar to commercial Roche 5′/3′ Race Kit. Prior to 5′-RACE, Terminal Transferase enzyme (Roche, http://www.roche-applied-science.com/) was used to add a homopolymeric A-tail to the 5′-end of the purified cDNA. Tailed and non-tailed cDNA were then amplified by PCR using a gene-specific SP1/SP5 primer and an Oligo-(dT) anchor primer. In addition, a second PCR run was performed using nested SP2/SP6 and anchor primer, and resulting products were verified by sequencing.

Full length coding sequences were cloned into Golden Gate entry vector pL0-SC (Engler et al., 2008). To remove any gene specific type II enzyme restriction sites, a single point mutation was introduced into the enzyme restriction sites and appropriate gene fragments were PCR amplified, subcloned into pUC18 vector (Engler and Marillonnet, 2011), sequenced and the full-length gene assembled into pL0-SC using a one-step restriction-ligation reaction with *BpiI* (Fermentas) and *T4*-HC-Ligase (Promega, http://www.promega.com/). Full-length coding sequence for KSL2 was PCR amplified from the same cDNA using specific oligonucleotide primers and AccuPrime High Fidelity Taq Polymerase (Invitrogen). The amplified fragment was gel-purified and subcloned into pENTR-D-TOPO vector (Invitrogen). After sequencing to check for mutations the KSL2 cDNA was transferred into pEarleyGate100 binary vector (Earley et al., 2006) using Gateway LR Clonase II Enzyme mix (Invitrogen). Final construct was confirmed both by restriction digestion and sequencing.

Cloning of SmKSL for in planta expression

Recombinant SmKSL was kindly provided by the lab of Prof. Reuben J. Peters, as a clone in pDest17 vector and optimized for expression in *E. coli* (Gao et al., 2009). To use it for transient expression in *N. benthamiana*, the transit peptide had to be restored. Thus, the translated nucleotide sequences of RoKSL1 and SmKSL were aligned to estimate homologous regions flanking the putative

transit peptide of RoKSL1. As a result, a fragment of 162 bp (=54 amino acids) downstream the ATG start codon was PCR amplified from the RoKSL1 entry clone using Phusion™ High Fidelity DNA Polymerase (Thermo Scientific). A truncated nucleotide sequence of SmKSL (1626 bp) was PCR amplified from recombinant pDest17 plasmid and ligated to the putative RoKSL1 transit peptide into Golden Gate entry vector pL0-SC.

Cloning and expression of recombinant RoCPS1 and RoKSL1 in yeast

Full length (*f*) as well as truncated sequences lacking putative transit peptide (*m*) of RoCPS1 and RoKSL1 were amplified by high fidelity PCR using Phusion Hot Start Flex DNA polymerase (New England Biolabs, https://www.neb.com/), cloned into pCRII vector (Invitrogen) and transferred into yeast expression vector using *BamHI* and *NotI* cohesive ends. While for RoCPS1 the transit peptide was predicted to be 67 amino acids (TargetP 1.1 server, http://www.cbs.dtu.dk/services/TargetP/), it was estimated to be 46 amino acids for RoKSL1. The following primers with added *BamHI*/*NotI* restriction sites were used for the amplification: *BamHI*: RoCPS1(*f*) forward 5′-GCATGAGGATCCATGACCTCTATGTC CTCTCTAAATTTG-3′, RoCPS1 (*m*) forward 5′-GCATGAGGATC CATGGCGTCACAAGTGAGTGAAAAAGG-3′, RoKSL1 (*f*) forward 5′-GCATGAGGATCCATGTCTCTCGCCTTCAACC-3′, RoKSL1 (*m*) forward 5′-GCATGAGGATCCATGTGCAACCTTACTACAACAGATTTG-3′; *NotI*: RoCPS1 reverse 5′-GCATGAGCGGCCGCTTCATACGACCGGTCCAAA CAG-3′ and RoKSL1 reverse 5′-GCATGAGCGGCCGCTTCATTTGC CACTCACTTTTTTG-3′. Plasmids used for expression in yeast cells were pUTDH3 (PTDH3 2 μ URA3) for the RoCPS1 and pHTDH3myc (PTDH3 2 μ HIS3) for RoKSL1 (both provided by Antonios Makris), yielding the final plasmids pHTDH3-*RoCPS1f*, pHTDH3-*RoCPS1m*, pUTDH3-*RoKSLf* and pUTDH3-*RoKSLm*. The plasmids were co-transformed into AM113 yeast strain (provided by Antonios Makris) in the following combinations: RoCPS1*f* + RoKSL1*f*, RoCPS1*m* + RoKSL1*f*, RoCPS1*f* + RoKSL1*m* and RoCPS1*m* + RoKSL1*m*. The AM113 strain harbors heterozygous deletions for *ERG9*, *UBC7*, *SSM4*, *MCT1*, *WHI2*, and expresses chromosomal inserted *GGDPS1* from *C. creticus*, the *FDPS1* from *Salvia fruticosa* and further contains two copies of a stabilized *HMG2* variant. Positive transformed clones were selected on glucose CM media lacking uracil and histidine. The final yeast strains were cultivated in 25 ml liquid media at 30 °C and 250 rpm. After 48 h of fermentation, the liquid cultures were extracted three times with equal volumes of hexane. The extraction was performed by sonication in an ultrasonic water bath at 40 °C. The hexane extract was dried using anhydrous MgSO4, evaporated under a stream of air to a final volume of 100 μl and analyzed by GC–MS. The GC–MS analyses were carried out using a Shimadzu GC/MS-QP2010 system (Shimadzu) consisting of GC-2010 gas chromatograph and QP-2010 electron impact (EI) mass spectrometer. Chromatographic separation was achieved on a 30-m × 0.25-mm × 0.25-μm HP-5-MS capillary column (Agilent Technologies). Ultra-high pure helium was used as carrier gas at a flow rate of 1 ml/min. Splitless mode was used for the injector with the inlet temperature set to 230 °C. The column temperature was initially hold at 60 °C for 1 min, then rose to 240 °C with a rate of 3 °C/min and hold for 15 min. The injection volume was 1 μL.

Transient expression of recombinant proteins in N. benthamiana

For expression in *N. benthamiana*, the full coding sequences of RoCPS1 and RoKSL1 were cloned into binary vector pL1F-1 (Weber et al., 2011) and fused to CaMV35S promoter by a single restriction-ligation reaction using 10 units *BsaI* (New England Biolabs) and 10 units *T4*-HC-Ligase (Promega). *A. tumefaciens* strain GV3101::pMP90 was transformed with the recombinant plasmids by using electroporation, and cells were grown on LB agar plates supplemented with rifampicin (25 μg/ml) and gentamycin (50 μg/ml) for 48 h at 28 °C. *Agrobacterium* cells harboring *RoCPS1* or *RoKSL1* constructs were selected in the presence of carbenicillin (50 μg/ml), and those with *RoKSL2* in the presence of kanamycin (50 μg/ml). Single transformed colonies were used to inoculate a 5 ml batch of LB broth containing the appropriate antibiotics, and cells were grown for 24 h at 28 °C and 180 rpm. The pre-culture was used to inoculate 50 ml LB broth with antibiotics in a ratio of 1:50 and growth was continued for 16–20 h at 28 °C and 180 rpm. Cells were harvested by centrifugation for 90 min at 1800*g* and 4 °C, and suspended in 0.25 volumes LB broth, 0.5 volumes infiltration buffer (20 mM glucose, 10% (w/v) sucrose, 8.6 g/L Murashige and Skoog basal salt mixture) and filled up with sterile water to reach a final OD_{600} of 0.5 in a total volume of 20 ml. When different constructs were co-infiltrated, the corresponding recombinant agrobacteria were mixed before centrifugation and so that OD_{600} = 0.5. *Agrobacterium* harboring a T-DNA with viral-encoded p19 protein was co-expressed in order to suppress gene silencing in the host (Voinnet et al., 2003). 20 μM acetosyringone were added to the agrobacteria suspension and mixtures were immediately infiltrated into leaves of 4–5 week old *N. benthamiana* plants by pressing a 1 mL syringe against the abaxial side of the leaf, and slowly injecting the agrobacteria suspension into the mesophyll. Treated plants were returned to growth chambers for additional 5 days until further analysis.

GC–MS analysis of N. benthamiana extracts

To analyze the exudate of agro-infiltrated *N. benthamiana* leaves, six leaf discs (9 mm in diameter) were punched out from infiltrated regions and extracted by dipping in hexane for 2 min. The extracts were evaporated completely under a stream of nitrogen, and resolved in 200 μl of hexane supplemented with internal standard sclareol (20 μM). After centrifugation for 5 min at 16,000*g*, the hexane extracts were transferred to GC vials, and 1 μl portions injected directly into GC–MS for analysis on a Trace GC Ultra gas chromatograph coupled to ISQ mass spectrometer (Thermo Scientific). Separation was performed in a 30-m × 0.32-mm diameter capillary, with a 0.25-μm film of ZB-5 ms (phenomenex). Splitless mode was used for injection with the inlet temperature set to 250 °C. The oven was programmed to start at 50 °C held for 1-min after which temperature was increased to 300 °C at a rate of 7 °C min^{-1} and further to 330 °C with 20 °C min^{-1} and a 5-min hold. Helium was used as carrier gas at flow rate of 1 ml/min. Electron impact was recorded at 70 eV and MS data were collected from 50 to 450 *m*/*z* during the temperature ramp. The concentration of target compounds was calculated from their total ion count (TIC) peak areas relative to the peak area of internal standard sclareol, and quantities were given as ISTD equivalents.

LC–MS analysis of phenolic diterpenes in rosemary trichomes and leaves

Manual isolation of trichomes was performed under a binocular. Per plant, a total of 100 trichomes were picked from the abaxial side of 10 leaves using the tip of a small spatula and collected in 1 mL ethanol in an Eppendorf tube. Solvent extraction of trichome content was performed by gently washing 10 leaves in 1 ml of chloroform in an Eppendorf tube for 30 s. The chloroform was then poured over into a clean tube and dried using a nitrogen air flow. Extracted compounds were redissolved in 75% methanol containing 0.1% formic acid. The washed leaves were frozen in liquid nitrogen, ground into a fine powder and extracted with 75% methanol containing 0.1% formic acid by sonication for 15 min. After extraction, the conversion of CA into carnosol was less than 20% over

a period of 24 h, as was determined by repeated analysis of the same extract (results not shown).

Phenolic diterpenes were separated, identified and quantified by using a LC-PDA-LTQ-Orbitrap FTMS system (Thermo Scientific) consisting of an Accela HPLC and an Accela photodiode array detector (240–600 nm) connected to an LTQ/Orbitrap hybrid mass spectrometer equipped with an ESI source. Injection volume was 5 µl. Chromatographic separation was on a reversed phase column (Luna C18/2, 3 µm, 2.0 × 150 mm; Phenomenex, USA) at 40 °C, using a linear gradient from 5% to 75% acetonitrile (acidified with 0.1% formic acid) in 45 min and a flow rate of 0.19 ml/min. FTMS full scans (*m*/*z* 90–1200) were recorded in negative ionization mode with a resolution of 60,000. Phenolic diterpenes were unambiguously identified based on their corresponding absorbance spectra, elemental formula within 3 ppm mass accuracy and co-elution with authentic standards. Both CA and carnosol were quantified based on the area of their specific accurate mass, while other phenolic diterpenes were quantified based on their absorbance at 280 nm using carnosic acid as a reference.

Phylogenetic analysis

A set of 50 protein sequences, as indicated in Supplemental Table S2, were aligned using ClustalW (Larkin et al., 2007) and the alignment was trimmed using GBlocks, with the least stringent criteria (Castresana, 2000). There were a total of 447 positions in the final data set. To find the most appropriate model of protein evolution that fits the data, an evolutionary model prediction was carried out using ProtTest 2.4 (http://darwin.uvigo.es/software/prottest2_server.html), (Abascal et al., 2005). The phylogenetic tree was implied by using the maximum-likelihood method, based on the JTT/*freq.* (*+F*) model (Jones et al., 1992). A discrete gamma distribution with invariant sites was used to model evolutionary rate differences among sites (4 categories; alpha (+G+I) 2.17). The bootstrap consensus tree was inferred from 1000 replicates and the percentage of replicate trees in which the associated taxa clustered together in the bootstrap test is shown next to the branches. Phylogenetic tree was drawn to scale, with branch lengths measured in the number of substitutions per site. Analyses were conducted in Mega 5 (Tamura et al., 2011).

Nucleotide sequence accession numbers

EST sequences for rosemary *CPS* and *KSL* were obtained from http://www.terpmed.eu/: CPS1, Roff.N01.C023303-05; KSL1, Roff.N01.C024308-09; KSL2, Roff.N01.C014863-70 and from http://medicinalplantgenomics.msu.edu/ (new sequence IDs since Oct 7, 2011): CPS1, roa_locus_11113_iso_1_len_2465_ver_1 (old ID) or roa_locus_14797_iso_1_len_2577_ver_2 (current sequence ID); KSL, roa_locus_5134_iso_4_len_1944_ver_1 / roa_locus_5134_iso_8_len_1935_ver_1 (old sequence IDs) or roa_locus_92_iso_30_len_1221_ver_2 (current ID).

The final isolated nucleotide sequences of *RoCPS1*, *RoKSL1* and *RoKSL2* genes have been deposited in GenBank under the accession numbers KF805857 through KF805859.

Acknowledgments

This work was supported by E.U. Grant agreement KBBE-227448 for the TERPMED project (http://www.terpmed.eu). Ric C. H. de Vos acknowledges the Netherlands Metabolomics Centre, which is part of the Netherlands Genomics Initiative/Netherlands Organization for Scientific Research, for additional funding. We further thank Prof. Reuben Peters for providing the *SmKSL* plasmid.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.phytochem.2014.01.021.

References

Abascal, F., Zardoya, R., Posada, D., 2005. ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21, 2104–2105.
Alvarenga, S.A.V., Gastmans, J.P., Rodrigues, G.D., Moreno, P.R.H., Emerenciano, V.D., 2001. A computer-assisted approach for chemotaxonomic studies – diterpenes in Lamiaceae. Phytochemistry 56, 583–595.
Bick, J.A., Lange, M.B., 2003. Metabolic crosstalk between cytosolic and plastidial pathways of isoprenoid biosynthesis: unidirectional transport of intermediates across the chloroplast envelope membrane. Arch. Biochem. Biophys. 415, 146–154.
Brückner, K., Tissier, A., 2013. High-level diterpene production by transient expression in *Nicotiana benthamiana*. Plant Methods 9, 46.
Caniard, A., Zerbe, P., Legrand, S., Cohade, A., Valot, N., Magnard, J.L., Bohlmann, J., Legendre, L., 2012. Discovery and functional characterization of two diterpene synthases for sclareol biosynthesis in *Salvia sclarea* (L.) and their relevance for perfume manufacture. BMC Plant Biol. 12, 119.
Cao, R., Zhang, Y., Mann, F.M., Huang, C., Mukkamala, D., Hudock, M.P., Mead, M.E., Prisic, S., Wang, K., Lin, F.Y., Chang, T.K., Peters, R.J., Oldfield, E., 2010. Diterpene cyclases and the nature of the isoprene fold. Proteins 78, 2417–2432.
Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17, 540–552.
Chen, F., Tholl, D., Bohlmann, J., Pichersky, E., 2011. The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. Plant J. 66, 212–229.
Earley, K.W., Haag, J.R., Pontes, O., Opper, K., Juehne, T., Song, K., Pikaard, C.S., 2006. Gateway-compatible vectors for plant functional genomics and proteomics. Plant J. 45, 616–629.
Engler, C., Kandzia, R., Marillonnet, S., 2008. A one pot, one step, precision cloning method with high throughput capability. PLoS One 3, e3647.
Engler, C., Marillonnet, S., 2011. Generation of families of construct variants using golden gate shuffling. Methods Mol. Biol. 729, 167–181.
Falara, V., Pichersky, E., Kanellis, A.K., 2010. A copal-8-ol diphosphate synthase from the angiosperm *Cistus creticus* subsp. creticus is a putative key enzyme for the formation of pharmacologically active, oxygen-containing labdane-type diterpenes. Plant Physiol. 154, 301–310.
Fischedick, J.T., Standiford, M., Johnson, D.A., Johnson, J.A., 2013. Structure activity relationship of phenolic diterpenes from *Salvia officinalis* as activators of the nuclear factor E2-related factor 2 pathway. Bioorg. Med. Chem. 21, 2618–2622.
Gao, W., Hillwig, M.L., Huang, L., Cui, G., Wang, X., Kong, J., Yang, B., Peters, R.J., 2009. A functional genomics approach to tanshinone biosynthesis provides stereochemical insights. Org. Lett. 11, 5170–5173.
Guo, J., Zhou, Y.J., Hillwig, M.L., Shen, Y., Yang, L., Wang, Y., Zhang, X., Liu, W., Peters, R.J., Chen, X., Zhao, Z.K., Huang, L., 2013. CYP76AH1 catalyzes turnover of miltiradiene in tanshinones biosynthesis and enables heterologous production of ferruginol in yeasts. Proc. Natl. Acad. Sci. U.S.A. 110, 12108–12113.
Hall, D.E., Zerbe, P., Jancsik, S., Quesada, A.L., Dullat, H., Madilao, L.L., Yuen, M., Bohlmann, J., 2013. Evolution of conifer diterpene synthases: diterpene resin acid biosynthesis in lodgepole pine and jack pine involves monofunctional and bifunctional diterpene synthases. Plant Physiol. 161, 600–616.
Hayashi, K., Kawaide, H., Notomi, M., Sakigi, Y., Matsuo, A., Nozaki, H., 2006. Identification and functional analysis of bifunctional ent-kaurene synthase from the moss *Physcomitrella patens*. FEBS Lett. 580, 6175–6181.
Hedden, P., Thomas, S.G., 2012. Gibberellin biosynthesis and its regulation. Biochem. J. 444, 11–25.
Hillwig, M.L., Xu, M., Toyomasu, T., Tiernan, M.S., Wei, G., Cui, G., Huang, L., Peters, R.J., 2011. Domain loss has independently occurred multiple times in plant terpene synthase evolution. Plant J. 68, 1051–1060.
Jones, D.T., Taylor, W.R., Thornton, J.M., 1992. The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. 8, 275–282.
Jordan, M.J., Lax, V., Rota, M.C., Loran, S., Sotomayor, J.A., 2012. Relevance of carnosic acid, carnosol, and rosmarinic acid concentrations in the in vitro antioxidant and antimicrobial activities of *Rosmarinus officinalis* (L.) methanolic extracts. J. Agric. Food Chem. 60, 9603–9608.
Kampranis, S.C., Ioannidis, D., Purvis, A., Mahrez, W., Ninga, E., Katerelos, N.A., Anssour, S., Dunwell, J.M., Degenhardt, J., Makris, A.M., Goodenough, P.W., Johnson, C.B., 2007. Rational conversion of substrate and product specificity in a *Salvia monoterpene* synthase: structural insights into the evolution of terpene synthase function. Plant Cell 19, 1994–2005.
Keeling, C.I., Bohlmann, J., 2006. Diterpene resin acids in conifers. Phytochemistry 67, 2415–2423.
Kitadani, M., Yoshikos, A., Kitahara, Y., Campello, J.D., Mcchesne, Jd, Watts, D.J., Wenkert, E., 1970. Natural Ar-Abietatriene. Chem. Pharm. Bull. 18, 402–405.
Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947–2948.

Li, G., Köllner, T.G., Yin, Y., Jiang, Y., Chen, H., Xu, Y., Gershenzon, J., Pichersky, E., Chen, F., 2012. Nonseed plant *Selaginella moellendorffii* has both seed plant and microbial types of terpene synthases. Proc. Natl. Acad. Sci. U.S.A. 109, 14711–14715.

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25, 402–408.

Morrone, D., Hillwig, M.L., Mead, M.E., Lowry, L., Fulton, D.B., Peters, R.J., 2011. Evident and latent plasticity across the rice diterpene synthase family with potential implications for the evolution of diterpenoid metabolism in the cereals. Biochem. J. 435, 589–595.

Munne-Bosch, S., Alegre, L., 2001. Subcellular compartmentation of the diterpene carnosic acid and its derivatives in the leaves of rosemary. Plant Physiol. 125, 1094–1102.

Ngo, S.N., Williams, D.B., Head, R.J., 2011. Rosemary and cancer prevention: preclinical perspectives. Crit. Rev. Food. Sci. Nutr. 51, 946–954.

Pateraki, I., Kanellis, A.K., 2008. Isolation and functional analysis of two *Cistus creticus* cDNAs encoding geranylgeranyl diphosphate synthase. Phytochemistry 69, 1641–1652.

Peters, R.J., 2006. Uncovering the complex metabolic network underlying diterpenoid phytoalexin biosynthesis in rice and other cereal crop plants. Phytochemistry 67, 2307–2317.

Peters, R.J., 2010. Two rings in them all: the labdane-related diterpenoids. Nat. Prod. Rep. 27, 1521–1530.

Peters, R.J., Croteau, R.B., 2002. Abietadiene synthase catalysis: conserved residues involved in protonation-initiated cyclization of geranylgeranyl diphosphate to (+)-copalyl diphosphate. Biochemistry 41, 1836–1842.

Peters, R.J., Flory, J.E., Jetter, R., Ravn, M.M., Lee, H.J., Coates, R.M., Croteau, R.B., 2000. Abietadiene synthase from grand fir (*Abies grandis*): characterization and mechanism of action of the "pseudomature" recombinant enzyme. Biochemistry 39, 15592–15602.

Sallaud, C., Rontein, D., Onillon, S., Jabès, F., Duffé, P., Giacalone, C., Thoraval, S., Escoffier, C., Herbette, G., Leonhardt, N., Causse, M., Tissier, A., 2009. A novel pathway for sesquiterpene biosynthesis from Z, Z-farnesyl pyrophosphate in the wild tomato *Solanum habrochaites*. Plant Cell 21, 301–317.

Sallaud, C., Giacalone, C., Topfer, R., Goepfert, S., Bakaher, N., Rosti, S., Tissier, A., 2012. Characterization of two genes for the biosynthesis of the labdane diterpene Z-abienol in tobacco (*Nicotiana tabacum*) glandular trichomes. Plant J. 72, 1–17.

Schilmiller, A.L., Schauvinhold, I., Larson, M., Xu, R., Charbonneau, A.L., Schmidt, A., Wilkerson, C., Last, R.L., Pichersky, E., 2009. Monoterpenes in the glandular trichomes of tomato are synthesized from a neryl diphosphate precursor rather than geranyl diphosphate. Proc. Natl. Acad. Sci. U.S.A. 106, 10865–10870.

Satoh, T., Kosaka, K., Itoh, K., Kobayashi, A., Yamamoto, M., Shimojo, Y., Kitajima, C., Cui, J., Kamins, J., Okamoto, S., Izumi, M., Shirasawa, T., Lipton, S.A., 2008. Carnosic acid, a catechol-type electrophilic compound, protects neurons both in vitro and in vivo through activation of the Keap1/Nrf2 pathway via S-alkylation of targeted cysteines on Keap1. J. Neurochem. 104, 1116–1131.

Schalk, M., Pastore, L., Mirata, M.A., Khim, S., Schouwey, M., Deguerry, F., Pineda, V., Rocci, L., Daviet, L., 2012. Toward a biosynthetic route to sclareol and amber odorants. J. Am. Chem. Soc. 134, 18900–18903.

Schnee, C., Kollner, T.G., Gershenzon, J., Degenhardt, J., 2002. The maize gene terpene synthase 1 encodes a sesquiterpene synthase catalyzing the formation of (E)-beta-farnesene, (E)-nerolidol, and (E, E)-farnesol after herbivore damage. Plant Physiol. 130, 2049–2060.

Schwarz, K., Ternes, W., 1992. Antioxidative constituents of *Rosmarinus officinalis* and *Salvia officinalis*. II. Isolation of carnosic acid and formation of other phenolic diterpenes. Z. Lebensm. Unters. Forsch. 195, 99–103.

Sugai, Y., Ueno, Y., Hayashi, K., Oogami, S., Toyomasu, T., Matsumoto, S., Natsume, M., Nozaki, H., Kawaide, H., 2011. Enzymatic (13)C labeling and multidimensional NMR analysis of miltiradiene synthesized by bifunctional diterpene cyclase in *Selaginella moellendorffii*. J. Biol. Chem. 286, 42840–42847.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.

Tounekti, T., Munne-Bosch, S., 2012. Enhanced phenolic diterpenes antioxidant levels through non-transgenic approaches. Crit. Rev. Plant Sci. 31, 505–519.

Toyomasu, T., 2008. Recent advances regarding diterpene cyclase genes in higher plants and fungi. Biosci. Biotechnol. Biochem. 72, 1168–1175.

Voinnet, O., Rivas, S., Mestre, P., Baulcombe, D., 2003. An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. Plant J. 33, 949–956.

Weber, E., Engler, C., Gruetzner, R., Werner, S., Marillonnet, S., 2011. A modular cloning system for standardized assembly of multigene constructs. PLoS One 6, e16765.

Wu, Y., Zhou, K., Toyomasu, T., Sugawara, C., Oku, M., Abe, S., Usui, M., Mitsuhashi, W., Chono, M., Chandler, P.M., Peters, R.J., 2012. Functional characterization of wheat copalyl diphosphate synthases sheds light on the early evolution of labdane-related diterpenoid metabolism in the cereals. Phytochemistry 84, 40–46.

Xu, M., Wilderman, P.R., Peters, R.J., 2007. Following evolution's lead to a single residue switch for diterpene synthase product outcome. Proc. Natl. Acad. Sci. U.S.A. 104, 7397–7401.

Yang, D., Du, X., Liang, X., Han, R., Liang, Z., Liu, Y., Liu, F., Zhou, J., 2012. Different roles of the mevalonate and methylerythritol phosphate pathways in cell growth and tanshinone production of *Salvia miltiorrhiza* hairy roots. PLos One 7, e46797.

Yanishlieva, N.V., Marinova, E., Pokorny, J., 2006. Natural antioxidants from herbs and *spices*. Eur. J. Lipid Sci. Technol. 108, 776–793.

Yerger, E.H., Grazzini, R.A., Hesk, D., Coxfoster, D.L., Craig, R., Mumma, R.O., 1992. A rapid method for isolating glandular trichomes. Plant Physiol. 99, 1–7.

Zerbe, P., Hamberger, B., Yuen, M.M., Chiang, A., Sandhu, H.K., Madilao, L.L., Nguyen, A., Bach, S.S., Bohlmann, J., 2013. Gene discovery of modular diterpene metabolism in nonmodel systems. Plant Physiol. 162, 1073–1091.

Zhou, K., Xu, M., Tiernan, M., Xie, Q., Toyomasu, T., Sugawara, C., Oku, M., Usui, M., Mitsuhashi, W., Chono, M., Chandler, P.M., Peters, R.J., 2012. Functional characterization of wheat ent-kaurene(-like) synthases indicates continuing evolution of labdane-related diterpenoid metabolism in the cereals. Phytochemistry 84, 47–55.

Zi, J., Peters, R.J., 2013. Characterization of CYP76AH4 clarifies phenolic diterpenoid biosynthesis in the Lamiaceae. Org. Biomol. Chem. 11, 7650–7652.