Received Date : 25-Mar-2014
Revised Date : 27-May-2014
Accepted Date : 10-Jun-2014
Article type : Original Article

Diterpene synthases of the biosynthetic system of medicinally active diterpenoids in *Marrubium vulgare*

Philipp Zerbe[1], Angela Chiang[1], Harpreet Dullat[1], Mark O'Neil-Johnson[2], Courtney Starks[2], Björn Hamberger[3] and Jörg Bohlmann[1,*]

[1]Michael Smith Laboratories, University of British Columbia, 301-2185 East Mall, Vancouver, BC, Canada, V6T 1Z4

[2]Sequoia Sciences, 1912 Innerbelt Business Center Drive, Saint Louis, MO 63114, USA

[3]Department of Plant and Environmental Sciences, Faculty of Science, University of Copenhagen, Thorvaldsensvej 40, 1871 Copenhagen, Denmark

[*]To whom correspondence should be addressed: Jörg Bohlmann, Michael Smith Laboratories, University of British Columbia, 301-2185 East Mall, Vancouver BC Canada V6T 1Z4, Phone: 604-822-0282, Fax: 604-822-2114, e-mail: bohlmann@mail.ubc.ca

Running head: *Marrubium* diterpene synthases

Key words: Diterpenoid biosynthesis, diterpene synthase, *Marrubium vulgare*, medicinal plants, plant specialized metabolism, marrubiin, natural product biosynthesis

Accession numbers: Nucleotide sequences reported in this study were submitted to the NCBI GenBankTM/EBI Data Bank with accession numbers: *Mv*CPS1 [KJ584450], *Mv*CPS2 [KJ584451], *Mv*CPS3 [KJ584452], *Mv*EKS [KJ584453], *Mv*ELS [KJ584454].

SUMMARY

Marrubium vulgare (Lamiaceae) is a medicinal plant whose major bioactive compounds, marrubiin and other labdane-related furanoid diterpenoids, have potential applications as anti-diabetics, analgesics or vasorelaxants. Metabolite and transcriptome profiling of *M. vulgare* leaves identified five different candidate diterpene synthases (diTPSs) of the TPS-c and TPS-e/f clades. We describe the *in vitro* and *in vivo* functional characterization of the *M. vulgare* diTPS family. In addition to *Mv*EKS *ent*-kaurene synthase of general metabolism, we identified three diTPSs of specialized metabolism: *Mv*CPS3 (+)-copalyl diphosphate synthase, and the functional diTPS pair *Mv*CPS1 and *Mv*ELS. In a sequential reaction, *Mv*CPS1 and *Mv*ELS produce a unique oxygenated diterpene scaffold 9,13-epoxy-labd-14-ene en route to marrubiin and an array of related compounds. In contrast to previously known diTPSs that introduce a hydroxyl group at carbon C-8 of the labdane backbone, the *Mv*CPS1-catalyzed reaction proceeds via oxygenation of an intermediate carbocation at C-9, yielding the bicyclic peregrinol diphosphate. *Mv*ELS belongs to a subgroup of the diTPS TPS-e/f clade with unusual βα-domain architecture. *Mv*ELS is active *in vitro* and *in vivo* with three different prenyl diphosphate substrates forming the marrubiin precursor 9,13-epoxy-labd-14-ene, as identified by NMR, manoyl oxide and miltiradiene. *Mv*ELS fills a central position in the biosynthetic system that forms the foundation for the diverse repertoire of *Marrubium* diterpenoids. Co-expression of *Mv*CPS1 and *Mv*ELS in

engineered *E. coli* and *Nicotiana benthamiana* offers opportunities for producing precursors for an array of biologically active diterpenoids.

INTRODUCTION

Medicinal plants and their metabolites have been used since ancient times to treat illness and disease and continue to serve as a major source for pharmaceuticals in modern times (De Luca *et al*., 2012). These include many plant species that produce terpenoids. Diterpenoids in particular have a wide range of pharmaceutical uses in addition to their manifold other industrial applications (Bohlmann and Keeling, 2008; Hillwig *et al*., 2011a; Zerbe and Bohlmann, 2014). *Marrubium vulgare* L. (Lamiaceae), commonly known as white horehound, is distributed across Mediterranean Europe, Southern America and Northern Africa, and used for medicinal purposes and as ingredient in skin cosmetics and cough lozenges (*e.g*. Ricola®). Biological activities of *M. vulgare* are attributed to an array of diterpenoids, sterols, flavonoids and phenylpropanoids (Meyre-Silva and Cechinel-Filho, 2010).

Marrubiin and a large group of related furanoid labdane diterpene lactones with characteristic oxygenation at carbon C-9 (Fig. 1) are the principal constituents of *M. vulgare* and appear to be unique to the genus *Marrubium* and a few closely related species of the Lamiaceae, including *Leonorus cardiaca*, *Leonotis leonorus* and *Phlomis bracteosa* (Knöss and Zapp, 1997; Mnonopi *et al*., 2011; Hussain *et al*., 2011). Biological activities of marrubiin and chemically modified derivatives such as marrubiinic acid include analgesic and anti-diabetic properties (Meyre-Silva *et al*., 2005; Boudejal *et al*., 2012; Mnonopi *et al*., 2012; Popoola *et al*., 2013). In addition, vasorelaxant and anti-hypertensive activity has been reported for the related diterpenoid marrubenol (El Bardai *et al*., 2004). Knowledge of the *Marrubium* diterpene biosynthetic system

can pave the way to better harness the medicinal potential of its metabolites. To this end, we recently established a functional genomics strategy, combining metabolite profiling and transcriptome sequencing of non-model plants including *M. vulgare*, to build a comprehensive sequence inventory for discovery of diterpenoid pathways (Zerbe *et al*., 2013).

Plant diterpenoids are produced from geranylgeranyl diphosphate (GGPP) through modular pathways with different combinations of diterpene synthases (diTPSs), cytochrome P450-dependent monooxygenases (P450s), and various other enzymes that increase structural diversity. DiTPSs share a conserved α-helical fold with variations of three common domains γ, β and α, and characteristic functional motifs (DxDD, DDxxD, NSE/DTE) that define the enzymes' catalytic activity (Gao *et al*., 2012). DiTPSs of angiosperms predominantly belong to the TPS-c and TPS-e/f clades of monofunctional diTPSs (Peters, 2010; Chen *et al*., 2011; Zerbe *et al*., 2013). The large group of labdane-related diterpenoids in angiosperms is formed by pairs of class I and class II diTPSs (Harris *et al*., 2005; Xu *et al*., 2007; Falara *et al*., 2010; Caniard *et al*., 2012, Sallaud *et al*., 2012; Zerbe *et al*., 2013; Brückner *et al*., 2014; Pateraki *et al*., 2014). Class II diTPSs catalyze the protonation-dependent cyclization of GGPP into bicyclic prenyl diphosphates assisted by a DxDD functional motif in the N-terminal γβ-domain. Known bicyclic prenyl diphosphate products of class II diTPSs include copalyl diphosphate (CPP) of *ent* (9*R*,10*R*), (+) (*i.e*. normal; 9*S*,10*S*) or *syn* (9*S*,10*R*) stereochemistry and the oxygenated labda-13-en-8-ol diphosphate (LPP). These class II diTPS products are substrates for class I diTPSs, which have characteristic DDxxD and NSE/DTE motifs in the α-domain. Class I diTPSs cleave the diphosphate moiety, followed by various possible rearrangements to afford different diterpene scaffolds. *Ent*-copalyl diphosphate synthases (*ent*-CPSs) and *ent*-kaurene synthases (*ent*-KSs) involved in GA biosynthesis of general metabolism are archetypical examples of such a pair of

class II and class I diTPSs (Sun and Kamiya, 1994; Yamaguchi *et al*., 1998; Xu *et al*., 2007; Keeling *et al*., 2010). Lineage-specific expansion and diversification of diTPS gene families appears to be the result of gene and genome duplications and subsequent neo-functionalization (Peters, 2010; Chen *et al*., 2011).

Based on the common architecture of labdane diterpenoid biosynthesis, we hypothesize a pair of specialized monofunctional diTPSs and several P450s are required for marrubiin biosynthesis. Using known structures of diterpenoids identified in *Marrubium* (Henderson and McCrindle, 1969; Meyre-Silva and Cechinel-Filho, 2010; Dendougui *et al*., 2011), we propose a pathway that proceeds *via* diTPS-catalyzed regio-specific cyclohydration of GGPP into a C-9 oxygenated intermediate, followed by addition of a furan ring and γ-lactone to yield marrubiin (Fig. 1). Additionally, various modifications of marrubiin or pathway intermediates may give rise to the array of related *Marrubium* diterpenoids. Here, we describe the functional characterization of four different diTPSs in *M. vulgare*, which comprise *Mv*EKS *ent*-KS, *Mv*CPS3 as a specialized (+)-CPS, and the pair of *Mv*CPS1 and *Mv*ELS that produce specialized C-9 oxygenated diterpene intermediates of marrubiin biosynthesis. Substrate promiscuity of the class I *Mv*ELS enzyme also places this enzyme into three different routes of a diterpene biosynthetic system that explains different components of the *M. vulgare* diterpenoid metabolite profile. Combinatorial expression of *M. vulgare* diTPSs illustrates opportunities for metabolic engineering and synthetic biology platforms for marrubiin and other diterpene phytotherapeutics.

RESULTS

Tissue Profiles of Diterpenoids in M. vulgare

To evaluate the spatial distribution of diterpenoids in *M. vulgare*, we established terpenoid

profiles for extracts of roots, stems, leaves, stem trichomes and flower calyces and corolla (Fig. 2 and Table S1). Consistent with earlier studies (Knöss and Zapp, 1997; Belhattab and Larous, 2006), we found terpenoids only in the above ground tissues with marrubiin as the most abundant terpenoid in calyces, leaves and trichomes. We also identified premarrubiin, a metabolite with similar pattern of spatial distribution but containing a 9,13-epoxy bridge instead of the C-9 hydroxy group present in marrubiin (Fig. 1). Premarrubiin has a nearly identical mass fragmentation pattern as marrubiin and converts into marrubiin by heating (Henderson and McCrindle, 1969) (Fig. S1). Other detected diterpenoids comprised manoyl oxide, most abundant in trichomes, and four diterpenes (A-D) exclusively detected in flowers and with high abundance in corolla tissue. Mass spectra of diterpenes A-D showed significant similarity to peregrinol (labda-13-en-9-ol; Salei *et al*., 1967), suggesting a labdane-type diterpene structure bearing an oxygen functionality at carbon C-9 characteristic for marrubiin and many of the specialized diterpenoids in the genus *Marrubium* (Fig. S1).

The Diterpene Synthase Family of M. vulgare

We previously developed a 454 and Illumina transcriptome resource from leaves of *M. vulgare* (Zerbe *et al*., 2013). Querying the transcriptome assemblies identified 16 TPS candidate genes (Zerbe *et al*., 2013) including five putative diTPSs (Table 1). Three sequences (*Mv*CPS1, *Mv*CPS2, and *Mv*CPS3) resembled class II diTPSs featuring a DxDD motif, and two transcripts (*Mv*EKS and *Mv*ELS) were annotated as class I diTPSs based on DDx_2D and $(N/D)x_2(S/T/Q)x_3E$ signature motifs. Phylogenetic analysis placed the five diTPSs with TPS-c and TPS-e/f clades (Fig. 3). *Mv*CPS2 and *Mv*EKS belong to subclades that primarily comprise known *ent*-CPSs and *ent*-KSs, suggesting a role in GA biosynthesis. *Mv*CPS1, *Mv*CPS3 and

*Mv*ELS were closely related to known diTPSs with functions in specialized metabolism (Fig. 3). Specifically, within the TPS-c clade, *Mv*CPS3 was most closely related to (+)-CPSs from *Salvia miltiorhizza* and *Coleus forskohlii* involved in specialized diterpenoid biosynthesis (Gao *et al*., 2009; Pateraki *et al*., 2014). *Mv*CPS1 clustered with diTPSs from other species catalyzing the cyclo-oxygenation of GGPP to LPP (Falara *et al*., 2010; Caniard *et al*., 2012; Sallaud *et al*., 2012; Zerbe *et al*., 2013; Pateraki *et al*., 2014). However, branch length and a low protein sequence identity of 37-50% between *Mv*CPS1 and LPP synthases (LPSs) from *Salvia sclarea*, *Cistus creticus* and tobacco (*Nicotiana tabacum*) suggested a related but distinct function of *Mv*CPS1. *Mv*ELS represents a member of a recently discovered diTPS group within the TPS-e/f clade that adopt an atypical βα-domain architecture through loss of the N-terminal γ-domain (Hillwig *et al*., 2011b; Caniard *et al*., 2012; Zerbe *et al.,* 2013). Known diTPSs of this group are involved in specialized metabolism, including *S. sclarea* sclareol synthase (Caniard *et al*., 2012), *S. miltiorhizza* miltiradiene synthase (Gao *et al*., 2009), and *C. forskohlii* manoyl oxide / miltiradiene synthases (Pateraki *et al*., 2014).

DiTPS Transcript Levels in Leaves and Flowers

DiTPS transcript levels were measured by quantitative reverse transcription PCR (qRT-PCR) in fully developed leaves and flowers (corolla). Transcripts of all five diTPSs were present in both leaves and flowers (Fig. 4). While transcript abundance of *MvCPS2* and *MvEKS* was similar in leaves and flowers, transcript levels of *MvCPS1*, *MvCPS3* and *MvELS* were, respectively, 49-fold, 28-fold, and 4-fold higher in leaves compared to flowers. Highest relative transcript abundance was detected with *MvCPS3* and *MvELS* in leaves, and lowest relative transcript levels were detected with *MvCPS1* and *MvCPS3* in flowers. Low RNA yields and

integrity prevented cell type specific transcript measurements in isolated trichomes.

Functions of Class II Diterpene Synthases

Full length (FL) cDNAs of the five different diTPSs were PCR amplified using gene-specific oligonucleotides based on assembled transcript sequences and completed, where required, by 5'-RACE. For *in vitro* functional characterization, recombinant diTPS proteins were expressed in *Escherichia coli* C41, affinity-purified and assayed with GGPP as substrate. FL or N-terminally truncated proteins (lacking the predicted plastidial transit peptide) were tested for enzymatic activity.

Products of the three class II diTPS (*Mv*CPS1, *Mv*CPS2, *Mv*CPS3) activities were dephosphorylated by alkaline phosphatase treatment prior to gas chromatography-mass spectrometry (GC-MS) analysis and were identified by comparison to appropriate reference assay products. DiTPS activities were found for *Mv*CPS1 and *Mv*CPS3, while no activity was detected for *Mv*CPS2. Chiral GC-MS analysis identified the *Mv*CPS3 product as copalol (*i.e.* dephosphorylated CPP) when compared to the products of *Zea mayz ent*-CPS (*Zm*AN2; Harris *et al*., 2005) and a protein variant of Norway spruce (*Picea abies*) levopimaradiene/abietadiene synthase (*Pa*LAS:D611A) that produces (+)-CPP (Zerbe *et al*., 2012a) (Fig. 5A). To affirm stereoselectivity of *Mv*CPS3, we conducted combinatorial assays with either white spruce (*P. glauca*) *ent*-KS (*Pg*EKS) specific to *ent*-CPP (Keeling *et al*., 2010) (Fig. 5B), or a monofunctional pimaradiene synthase from lodgepole pine (*Pinus contorta*; *Pc*PIM1) accepting (+)-CPP as substrate (Hall *et al*., 2013) (Fig. 5C). While pairwise activity of *Zm*AN2 and *Pg*EKS yielded *ent*-kaurene (positive control), combination of *Mv*CPS3 with *Pg*EKS did not result in any detectable product formation (Fig. 5B). In contrast, coupled assays of *Mv*CPS3 and *Pc*PIM1

afforded pimaradiene, whereas no activity was observed when combining *Zm*AN2 and *Pc*PIM1 (Fig. 5C). From these experiments, and considering that *syn*-CPP is to current knowledge restricted to members of the Poaceae (Peters, 2010), it can be concluded that *Mv*CPS3 produces (+)-CPP. The close phylogenetic relationship between *Mv*CPS3 and known monofunctional (+)-CPSs supports this result (Fig. 3).

The dephosphorylated product profile of *Mv*CPS1 consisted of four compounds showing different GC retention times but almost identical mass spectra (Fig. 6A-C; peaks *f*, *g*, *h* and *i*). Cool-on-column GC-MS analysis of the dephosphorylated product identified peak *i* as the major *Mv*CPS1 product (Fig. 6D), accounting for >90% of the product profile, whereas abundance of compounds corresponding to peaks *f*, *g* and *h* was largely diminished under these milder analysis conditions. Since peregrinol is not available as an authentic standard, compound *i* was tentatively identified as peregrinol based on a characteristic mass fragmentation pattern of *m/z* 308(0.2), 290(25), 275(1), 257(1), 209(3), 191(13), 151(100), 123(32), 109(27), 95(25), 81(30), 69(39), 55(23), 41(22) (*cf.* Salei *et al*., 1967), and a consistent parent mass of *m/z* 308 as verified by chemical ionization GC-MS analysis (Fig. S2). *Mv*CPS1 adds a new function to the portfolio of known class II diTPSs, transforming (*E*,*E*,*E*)-GGPP into a bicyclic prenyl diphosphate with regio-specific oxygenation at C-9, peregrinol diphosphate (PPP) (Fig. 6). Notably, peaks *f-i* detected in the *Mv*CPS1 product profile showed identical retention times and signature mass fragments as diterpenoids A-D uniquely detected in flower tissue, suggesting that the most polar metabolite (diterpenoid D) represents peregrinol (Fig. S2).

MvELS is a multi-functional βα-Domain Class I Diterpene Synthase

For *in vitro* functional characterization of the two class I diTPSs (*Mv*EKS and *Mv*ELS) we

performed combinatorial assays with class II diTPSs, since the relevant prenyl diphosphate substrates are not commercially available. Truncated proteins of *Mv*EKS (Δ32) and *Mv*ELS (Δ38) lacking N-terminal plastid transit peptides were assayed in coupled reactions with *Mv*CPS1, *Mv*CPS3, *Zm*AN2, or coupled with LPS from *C. forskohlii* (*Cf*TPS2; Pateraki *et al*., 2014). In these assays *Mv*CPS1 was used to produce PPP, *Mv*CPS3 to produce (+)-CPP, *Zm*AN2 to produce *ent*-CPP, and *Cf*TPS2 to produce LPP as candidate substrates for *Mv*EKS and *Mv*ELS.

*Mv*EKS showed detectable activity only in combination *Zm*AN2, yielding *ent*-kaurene as single product (Fig. 7A). In contrast, *Mv*ELS showed no detectable activity when combined with *Zm*AN2, but converted (+)-CPP, LPP and PPP as observed with the following assays: In combination with the (+)-CPP-producing *Mv*CPS3, *Mv*ELS formed miltiradiene as identified by comparison to the product of the combined activity of (+)-CPP synthase (*Cf*TPS1) and miltiradiene/manoyl oxide synthase (*Cf*TPS4) from *C. forskohlii* as recently published (Fig. 7C and E, peak *k*; Pateraki *et al*., 2014). When coupled with *Cf*TPS2, *Mv*ELS converted LPP into manoyl oxide as compared to the product of the coupled reaction of *Cf*TPS2 with *Cf*TPS4 and characteristic reference mass fragmentation patterns (Fig. 7D and E, peak *l*; Pateraki *et al*., 2014; Demetzos *et al*., 2002). The coupled reaction of *Mv*CPS1 and *Mv*ELS afforded three products (Fig. 7B and E, peaks *f*, *g* and *h*), which were also present in the dephosphorylated product profile of *Mv*CPS1. In contrast to the *Mv*CPS1 reaction alone, the additional activity of *Mv*ELS yielded a product profile that lacked the compound corresponding to peak *i*, and the compound corresponding to peak *g* became the major reaction product accounting for 80% of the total product profile (Fig. 7F).

To afford suitable amounts of the product corresponding to peak *g* for NMR analysis, we generated a fusion protein of *Mv*ELS and *Mv*CPS1, in this order from the N- to the C-terminus, using a previously described linker sequence (GGGS; Zhou *et al*., 2012a). The *Mv*ELS-GGGS-*Mv*CPS1 fusion protein was expressed in an *E. coli* system engineered for diterpene production (Morrone *et al*., 2010). Following extraction and chromatography of the reaction product on silica gel, 1D and 2D NMR spectra were acquired, identifying compound *g* as 9,13-epoxy-labd-14-ene (Fig. 7G). While overlapping signals prevented an unambiguous assignment of stereochemistry at positions C-8 and C-13, the obtained structure clearly revealed the characteristic 9,13-epoxy bridge also found in premarrubiin and many of the related diterpenoids in *Marrubium* (Knöss and Zapp, 1997). This result is consistent with a class I diTPS ionization of the diphosphate group of the *Mv*CPS1 product PPP and subsequent deprotonation to afford an exocyclic double bond as illustrated in Figure 1. Based on their mass spectral similarity, the *Mv*ELS by-products peaks *f* and *h* are likely related structures with distinct double bond arrangements; however, amounts of these minor products were insufficient for structural analysis.

In Vivo Co-expression in Nicotiana benthamiana Confirmed DiTPS Product Formation

To substantiate the results of diterpene product formation of *M. vulgare* diTPSs determined in *in vitro* assays with recombinant proteins, we used *Agrobacterium*-mediated transient (co-)expression in tobacco (*Nicotiana benthamiana*) for *in planta* analyses. FL class II diTPSs (*Mv*CPS1, *Mv*CPS2, *Mv*CPS3) were expressed alone or in combination with class I enzymes (*Mv*EKS, *Mv*ELS) along with the RNA silencing suppressor protein p19 (Voinnet *et al*., 2003). The *in planta* product profiles generally confirmed products found with *in vitro* assays (Fig. 8).

No product formation was detected with *Mv*CPS2 alone or in combination with *Mv*EKS or *Mv*ELS (Fig. 8E-G). Expression of *Mv*CPS3 alone (Fig. 8H) afforded copalol (peak *b*; *i.e*. dephosphorylated CPP) and an unidentified diterpene (peak *m*; Fig. S3). Since *in vivo* expression assays in *N. benthamiana* did not include dephosphorylation prior to GC-MS analysis, it is plausible that endogenous phosphatases convert the class II diTPS *Mv*CPS3 product into a corresponding diterpenol, as was previously suggested for labdane-diol formation in tobacco (Sallaud *et al*., 2012). Both *Mv*CPS3 products (peak *b* and peak *m*) were absent when *Mv*CPS3 was co-expressed with *Mv*ELS, while miltiradiene (peak *k*) was detected as a single product (Fig. 8J). Combination of *Mv*CPS3 and *Mv*EKS afforded the *Mv*CPS3 products (peak *b* and peak *m*) and traces of an additional unknown diterpene (peak *n*) not detected in *in vitro* assays (Fig. 8I and S3). Given the very low abundance of peak *n*, this product may not represent a major product of *Mv*EKS converting (+)-CPP. Expression of *Mv*CPS1 alone yielded a product profile consistent with *in vitro* assays, with peregrinol (*i.e*. dephosphorylated PPP, peak *i*) as the major product (Fig. 8B). Correspondingly, co-expression of *Mv*CPS1 with *Mv*ELS resulted in 9,13-epoxy-labd-14-ene (peak *g*) as the major product (Fig. 8D). In contrast, co-expression of *Mv*CPS1 with *Mv*EKS yielded no additional product (Fig. 8C), consistent with the substrate specificity of *Mv*EKS for *ent*-CPP to form *ent*-kaurene.

DISCUSSION

Furanic labdane diterpenoids of *Marrubium vulgare* are of potential industrial value for their anti-diabetic, analgesic and vasorelaxant properties (Meyre-Silva and Cechinel-Filho, 2010; Popoola *et al*., 2013). Marrubiin and related C-9 oxygenated diterpenoids form the plant's major diterpene constituents and are unique to the genus *Marrubium* and a few closely related species

(Meyre-Silva and Cechinel-Filho, 2010). Knowledge of the biosynthetic genes and enzymes offers opportunity to implement metabolic engineering and synthetic biology strategies for producing marrubiin and other bioactive diterpenoids. Here we report the functional characterization of four diTPSs of *M. vulgare*, *Mv*CPS1 peregrinol diphosphate synthase, *Mv*CPS3 (+)-CPP synthase, *Mv*EKS *ent*-kaurene synthase, and the multi-functional class I diTPS *Mv*ELS, which form the foundation of the biosynthetic system leading to the diverse array of oxygenated diterpenoids in the genus *Marrubium* (Fig. 1). This work was building on a recently described genomics approach of gene and enzyme discovery of diterpenoid specialized metabolism in non-model plant species (Zerbe *et al*., 2013)

Phylogenetic relatedness of *Mv*EKS and *Mv*CPS2 with known *ent*-KSs and *ent*-CPSs of general metabolism indicated a role in GA biosynthesis (Fig. 3). While *Mv*EKS was confirmed as an *ent*-KS (Fig. 7), *Mv*CPS2 was inactive in both *in vitro* and *in planta* assays. Sequence comparison of *Mv*CPS2 with known *ent*-CPSs revealed that a conserved aspartate residue, which was recently shown to be involved in the reprotonation of the DxDD motif during catalysis in *Arabidopsis ent*-CPS (Asp503; Köksal *et al*., 2014), is substituted for a valine (Val506) in *Mv*CPS2, which may explain the lack of activity.

In contrast to *Mv*CPS2 and *Mv*EKS, *Mv*CPS1, *Mv*CPS3 and *Mv*ELS showed high transcript levels in leaf tissue (Fig. 4), suggesting a role in specialized diterpene metabolism consistent with the predominant abundance of marrubiin and related diterpenoids in leaves and trichomes (Fig. 2). Combinatorial assays with class I diTPSs specific to *ent*-CPP and (+)-CPP, respectively, identified *Mv*CPS3 as a (+)-CPS functionally orthologous to diTPSs from other Lamiaceae lineages (Gao *et al*., 2009; Pateraki *et al*., 2014; Brückner *et al*., 2014). Biosynthesis of (+)-CPP has also been demonstrated for a monofunctional wheat class II diTPS (Wu *et al*., 2012), and as

intermediate of bifunctional class II/I diTPSs in *Selaginella moellendorfii* (Sugai *et al*., 2011) and various gymnosperms (Peters *et al*., 2000; Martin *et al*., 2004; Keeling *et al*., 2011; Zerbe *et al*., 2012a; Hall *et al*., 2013). Formation of (+)-CPP may represent a phylogenetically older functionality that evolved repeatedly and independently across the plant kingdom. In the angiosperms *S. miltiorhizza*, *Rosmarinus officinalis* and *C. forskohlii* (+)-CPP is converted into miltiradiene en route to tanshinones and possibly other phenolic diterpenoids (Guo *et al*., 2013; Brückner *et al*., 2014, Pateraki *et al*., 2014). In *M. vulgare*, *Mv*ELS has the capacity to convert (+)-CPP into miltiradiene (Fig. 7). However, miltiradiene was not detected *in planta* and its metabolic fate in *Marrubium* is unclear.

With *Mv*CPS1, we discovered in *M. vulgare* a previously unknown class II diTPS, which produces a unique C-9 oxygenated prenyl diphosphate intermediate identified as peregrinol diphosphate (PPP) based on mass spectral comparison to peregrinol (Salei *et al*., 1967) (Fig. 6). PPP is converted by *Mv*ELS into 9,13-epoxy-labd-14-ene as verified by NMR analysis (Fig. 7G). The 9,13-epoxy bridge in the product of the coupled reaction of *Mv*CPS1 and *Mv*ELS, together with the detection of this product in *M. vulgare* flower extracts (Fig. S2), supports a role of these two enzymes in marrubiin biosynthesis. Based on previous phytochemical studies (Henderson and McCrindle, 1969; Knöss and Zapp, 1997), it is thought that marrubiin and a set of other specialized diterpenoids found in the genus *Marrubium* contain the epoxy bridge in the native form *in planta*, with break down of the epoxy bridge to a C-9 hydroxy group upon extraction.

The *Mv*CPS1-catalyzed biosynthesis of PPP adds a new function to the TPS-c clade of diTPSs. Formation of PPP from GGPP presumably proceeds via formation of a labda-13-en-8-yl diphosphate cation, which undergoes additional 1,2 hydride shift between C-8 and the neighboring methine group, and termination of the reaction by water quenching of the

intermediate carbocation (Fig. 6). This functionality sets *Mv*CPS1 apart from previously known class II diTPSs which neutralize the labda-13-en-8-yl diphosphate cation either through hydroxylation at carbon C-8 or direct deprotonation at C-8 or C-7 (Mafu *et al*., 2011; Peters, 2010). The narrow taxonomic distribution of marrubiin and related diterpenoids in *Marrubium* and a few other Lamiaceae (Knöss and Zapp, 1997; Mnonopi *et al*., 2011; Hussain *et al*., 2011), suggests that the *Mv*CPS1 function evolved more recently, possibly early in the speciation of the genus. Its phylogenetic placement together with several dicot LPSs further suggests the emergence of *Mv*CPS1 through duplication and neo-functionalization of an ancestral LPS (Fig. 3).

The activity of the class I diTPS *Mv*ELS with multiple substrates, PPP, (+)-CPP, and LPP produced by different class II diTPSs (Fig. 1), supports a model according to which diterpenoid biosynthesis is a modular system in which different combinations of class I and class II enzymes increase the diversity of the chemical space of diterpenoid intermediates and products. Similar biosynthetic systems have been emerging with recently described diTPS families in other Lamiaceae (Caniard *et al*., 2012; Pateraki *et al*., 2014), members of the Poaceae (Xu *et al*., 2007; Morrone *et al*., 2011; Zhou *et al*., 2012b) and species of pine (*Pinus*; Hall *et al*., 2013). Our findings place *Mv*ELS in a critical position of the diterpene biosynthetic system in *M. vulgare*, due to its capacity to co-function with different class II diTPSs to form the large array of C-8 and C-9 oxygenated diterpene scaffolds (Fig. 1). The ability to form functional pairs with different class II diTPSs appears to be prevalent among the group of atypical βα-domain class I diTPSs within the TPS-e/f clade, with similar functions shown for other Lamiaceae diTPSs, including *Salvia sclarea* sclareol synthase (Caniard *et al*., 2012), and manoyl oxide/miltiradiene synthases from *C. forskohlii* and *R. officinalis* (Pateraki *et al*., 2014; Brückner *et al*., 2014). Presence of

*Mv*ELS reaction products (except for miltiradiene formed through pairwise reaction with *Mv*CPS3) in tissue extracts of *M. vulgare* supports the presence of a modular diterpene biosynthetic system (Fig. 2). Given that activities for formation of *ent*-CPS and LPS (for manoyl oxide biosynthesis; Zerbe *et al*., 2013; Pateraki *et al*., 2014) were not among the diTPSs described here, it appears that additional *M. vulgare* diTPSs remain to be discovered.

The naturally occurring pairwise activity of functionally distinct diTPSs in modular diterpene biosynthetic systems, inspires opportunities for metabolic engineering and synthetic biology platforms through combinatorial expression of diTPSs and possibly P450 enzymes selected from the large suite of cloned enzymes for the production of known natural, novel and non-natural diterpene bioproducts (Ajikumar *et al*., 2010; Schalk *et al*., 2012; Guo *et al*., 2013, Brückner and Tissier, 2013). Although not at production scale, the formation of 9,13-epoxy-labd-14-ene by co-expression of *Mv*CPS1 and *Mv*ELS in microbial and plant systems shown here, illustrates the possibility to produce marrubiin and other pharmaceutically active diterpenes in engineered biological systems as an alternative to multi-step total syntheses (Mangoni *et al*., 1972; Deschamp *et al*., 2012).

EXPERIMENTAL PROCEDURES

Plant Material

Seeds of *M. vulgare* were purchased from B&T World Seeds (www.b-and-t-world-seeds.com). *Nicotiana benthamiana* seeds were provided by Dr. Dominique Michaud (Université Laval, Canada). Plants were grown in a Cenviron ATC10 phytochamber under long-day conditions (16 h light at 80 µmol $m^{-2}s^{-1}$ irradiance) and a day/night temperature of 22/25ºC.

Metabolite Analysis

Extracts were prepared from 25-50 mg fresh tissue of flowering plants. Glandular and non-glandular trichomes were carefully isolated from stem surfaces using a razor blade without damaging the stem tissue. Tissues were pulverized under liquid N_2 and terpenoids extracted with 1.5 ml pentane for 16 h under vigorous shaking at 4 °C. Extracts were freed from residual water using anhydrous Na_2SO_4 and further passed through 0.22 µm GHP membrane filters (www.pall.com). After extraction, plant materials were dried for tissue quantification. GC-MS analysis was performed on an Agilent (www.chem.agilent.com) 7890A-GC with a 7683B series autosampler and a 7000A TripleQuad MS Detector at 70 eV and 1.2 ml min^{-1} He flow, using a HP-5MS column (30 m, 250 µm i.d., 0.25 µm film) and the following GC program: 50 °C for 1 min, 15 °C min^{-1} to 320 °C, hold 5 min, pulsed splitless injector at 250 °C. Identification of marrubiin was achieved by comparison to the authentic standard (www.chromadex.com). Other terpenoid metabolites were identified based on comparison to reference mass spectra from databases of the National Institute of Standards and Technology (NIST) MS library searches (Wiley W9N08L) and relevant literature. Measurements were performed in duplicate with three biological replicates and 1-eicosene as internal standard (www.sigma.com).

Gene cloning

Full length (FL) cDNAs were amplified from total leaf RNA of 10-week-old *M. vulgare* plants using gene-specific oligonucleotides (Table S2). Rapid amplification of cDNA ends (RACE) using the SMARTer cDNA RACE kit (www.clontech.com) was used to obtain 5'-sequences as needed. Putative plastidial transit peptides were predicted using the ChloroP and TargetP servers (Emanuelsson *et al*., 1999; Emanuelsson *et al*., 2007). FL and N-terminally truncated cDNAs

were then retrieved by touchdown PCR and ligated into pJET (Clontech) followed by sequence verification and subcloning into expression vectors (see below). A fusion cDNA of *MvELS* and *MvCPS1* was generated by overlap extension PCR with partially complementary oligonucleotides as described by Wurch *et al*. (1998).

Quantitative real-time PCR (qRT-PCR)

Total RNA was isolated according to Kolosova *et al*. (2004) using ~100 mg tissue. RNA integrity and concentration was measured using Bioanalyzer 2100 RNA Nano chip assays (Agilent) following the manufacturer's protocol. Equal RNA amounts were used for cDNA synthesis with the Superscript III reverse transcriptase (www.invitrogen.com) and random hexamer oligonucleotides. The subsequent qRT-PCR reaction was performed on a Bio-Rad CFX96 Real-time system using the SsoFast kit (www.bio-rad.com) and target-specific oligonucleotides (Table S2). Relative transcript abundance was calculated using efficiency corrected ΔC_T and $\Delta\Delta C_T$ values based on actin as reference gene and triplicate measurements with three biological replicates. Statistically significant differences between transcript levels in leaves and flowers were calculated for each diTPS based on a Student's t-test (p-value < 0.05). Target specificity was confirmed by sequence verification of representative amplicons.

***In vitro* enzyme assays**

N-terminally truncated (lacking sequences coding for transit peptides) or FLcDNAs were subcloned into the pET28b(+) vector (www.emdmillipore.ca) and transformed into *E. coli* BL21DE3-C41 for heterologous expression and Ni^{2+}-affinity purification as described in Zerbe *et al*. (2012b). Single or coupled diTPS enzyme assays were conducted as previously reported,

using 50-100 µg of recombinant protein and 15 µM of (*E*,*E*,*E*)-GGPP (Sigma) as substrate with gentle shaking for 1 h at 30 ºC (Zerbe *et al*., 2012b). Class II diTPS products were dephosphorylated by treatment with 10 U of calf intestinal phosphatase (Invitrogen) over night at 37 °C. Reaction products were extracted with 500 µl pentane and analyzed via GC-MS on a HP-5MS column as described above or using a DB-wax column (30 m, 250 µm i.d., 0.25 µm film) on an Agilent 6890N-GC, 7683B series autosampler, and a 5975 Inert XL MS Detector with electron ionization (EI) at 70 eV and appropriate temperature programs. Chemical ionization (CI) analysis was conducted with methane on a HP-5MS column at 40 °C for 1 min, 40°C min^{-1} to 300 °C, hold 6 min, pulsed splitless injection at 275 °C. Chiral analysis was performed on a Cyclodex-B column (30 m, 250 µm i.d., 0.25 µm film) at 100 °C for 2 min, 2 °C min^{-1} to 230 °C, hold 8 min, and pulsed splitless injection at 240 °C; or under cool-on-column conditions at 30 °C for 1 min, 40 °C min^{-1} to 250 °C, hold 30 min, split injection at 230 °C. Product quantification was based on three independent assays, each with two technical replicates, and 1-eicosene as internal standard.

Diterpene synthase co-expression in *Nicotiana benthamiana*

FLcDNA constructs were ligated into the pLIFE33 vector and transformed into *A. tumefaciens* strain AGL-1. Cells were grown at 28 °C in Luria-Bertani media supplemented with 50 mg L^{-1} kanamycin, pelleted and resuspended to a final OD_{600} of 0.5 in 10 mM MES buffer with 10 mM $MgCl_2$. Following 60 min incubation at 22°C and gentle shaking, cell suspensions were mixed with one volume of the plant viral protein p19 to suppress RNA silencing (Voinnet *et al*., 2003) for single construct transformations, or using equal volumes for dual expression with one volume of p19 before infiltration into the leaf underside of 5-week-old *N. benthamiana* plants.

Transfected plants were maintained for four days before metabolites were extracted with 1.5 ml hexane from a single transformed leaf and analyzed via GC-MS on a DB-wax column as described for *in vitro* assays.

Nuclear magnetic resonance (NMR) analysis

For NMR analysis diterpenes were produced by co-expression of the plasmid pET28-Mv-fusion producing the *M*vELS-GGGS-*Mv*CPS1 fusion protein with the plasmid pGG in *E. coli* BL21-C41 as described in Morrone *et al*., (2010). Cultures were grown at 37 °C and 150 rpm in 500 ml Terrific Broth medium to an OD_{600} of ~0.6. The temperature was then reduced to 16 °C and the pH adjusted to 7.0 before induction with 1 mM isopropyl-β-D-1-thiogalactopyranoside for 48-72 h. Diterpenes were extracted with 500 ml hexane after sonication of cultures and purified through iterative separation on silica matrix with hexane as mobile phase. Proton, gCOSY, ROESY, gHSQC, and gHMBC spectra were acquired in MeOD at 600 MHz on a Bruker Avance 600 MHz spectrometer equipped with a Bruker BioSpin TCI 1.7 mm MicroCryoProbe.

Phylogenetic analysis

Protein sequence alignments were generated using the Dialign 2.2.1 server (Morgenstern, 2004) and further manual inspection. Maximum likelihood phylogenetic analyses were performed using PhyML-aBayes version 3.0.1 beta (Anisimova *et al*., 2011) with four rate substitution categories, LG substitution model, BIONJ starting tree and 100 bootstrap repetitions.

CONFLICT OF INTEREST STATEMENT

The authors confirm that they have no conflict of interest to be declared in accordance with the

journal policy.

ACKNOWLEDGEMENTS

We thank Ms. Karen Reid for excellent laboratory and project management support, and Ms. Lina Madilao for support with metabolite analysis. We gratefully acknowledge Dr. Reuben Peters (Iowa State University, USA) for providing the *Zm*AN2 and pGG plasmids, and Dr. Dominique Michaud (Université Laval, Canada) for providing the *A. tumefaciens* strain AGL-1. We also thank Dr. Barbara Ann Halkier (University of Copenhagen, Denmark) for providing the pLIFE33 vector and Dr. David Baulcombe (University of Cambridge, UK) for sharing the p19 plasmid. Financial support was provided by the Natural Sciences and Engineering Research Council of Canada (NSERC; to JB) and funds in support of the PhytoMetaSyn Project (to JB) from Genome Canada, Genome Alberta, Government of Alberta, Genome British Columbia, Genome Prairie, Genome Quebec, National Research Council of Canada, Canada Foundation for Innovation, Ontario Ministry of Research and Innovation, McGill Innovation Centre, MDEIE (Ministere du developement economique, innovation et exportation) and other government and private sector partners. JB is a UBC Distinguished University Scholar.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. GC-MS analysis of diterpene metabolites detected in *Marrubium vulgare* tissue extracts.

Figure S2. GC-MS analysis of *Mv*CPS1 products compared with diterpenes detected in *M. vulgare* flower tissue extracts.

Figure S3. Mass spectra of products obtained by combinatorial transient expression of *M. vulgare* diterpene synthases in *N. benthamiana*.

Table S1. Terpenoid profile of *Marrubium vulgare.*

Table S2. Oligonucleotides used in this study.

Table S3. Abbreviations and accession numbers of proteins used for phylogenetic analysis.

REFERENCES

Ajikumar, P.K., Xiao, W.H., Tyo, K.E., Wang, Y., Simeon, F., Leonard, E., Mucha, O., Phon, T.H., Pfeifer, B., Stephanopoulos, G. (2010) Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. *Science* **330**, 70-74.

Anisimova, M., Gil, M., Dufayard, J.F., Dessimoz, C. and Gascuel, O. (2011) Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. *Syst. Biol*. **60**, 685-699.

Belhattab, R. and Larous, L. (2006) Essential oil composition and glandular trichomes of *Marrubium vulgare* L. growing in wild Algeria. J. Essent. Oil Res. **18**, 369-373.

Bohlmann, J. and Keeling, C.I. (2008) Terpenoid biomaterials. *Plant J*. **54**, 656-669.

Boudejal, A., Henchiri, C., Siracusa, L., Sari, M. and Ruberto, G. (2012) Compositional analysis and *in vivo* anti-diabetic activityof wild Algerian *Marrubium vulgare* L. infusion. *Filoterapia* **83**, 286-292.

Brückner, K. and Tissier, A. (2013) High-level diterpene production by transient expression in *Nicotiana benthamiana*. *Plant Methods* **9**,46

Brückner, K., Božić, D., Manzano, D., Papaefthimiou, D., Pateraki, I., Scheler, U., Ferrer, A., de Vos, R.C., Kanellis, A.K., Tissier, A. (2014) Characterization of two genes for the biosynthesis of abietane-type diterpenes in rosemary (*Rosmarinus officinalis*) glandular trichomes. *Phytochemistry* doi: 10.1016/j.phytochem.2014.01.021.

Caniard, A., Zerbe, P., Legrand, S., Cohade, A., Valot, N., Magnard, J.L., Bohlmann, J. and 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.

Chen, F., Tholl, D., Bohlmann, J. and 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.

Dai, Z., Liu, Y., Huang, L. and Zhang, X. (2012) Production of miltiradiene by metabolically engineered *Saccharomyces cerevisiae*. *Biotechnol. Bioeng*. **109**, 2845-2853.

De Luca, V., Salim, V., Masada, Atsumi, S. and Yu, F. (2012) Mining the biodiversity of plants, a revolution in the making. *Science* **336**, 1658-1661.

Demetzos, C., Kolocouris, A. and Anastasakia, T. (2002) A simple and rapid method for the

differentiation of C-13 manoyl oxide epimers in biologically important samples using GC-MS analysis supported with NMR spectroscopy and computational chemistry results. *Bioorg. Med. Chem. Lett*. **12**, 3605-3609.

Dendougui, H., Seghir, S., Belloum, Z., Benayache, F., Leon, F., Brouard, I., Bermejo, J. and Benayache, S. (2011) A new labdane diterpene and other constituents from *Marrubium deserti* Noe ex coss. *Rec. Nat. Prod*. **5**, 300-304.

Deschamp, J., Hermant, T. and Riant, O. (2012) An easy route toward enantio-enriched polycyclic derivatives via an asymmetric domino conjugate reduction-aldol cyclization catalyzed by a chiral Cu(I) complex. *Tetrahedron* **68**, 3457-3467.

El Bardai, S., Hamaide, M.-C., Lyoussi, B., Quetin-Leclercq, J., Morel, N. and Wibo, M. (2004) Marrubenol interactis with the phenylalkylamine binding site of the L-type calcium channel. *Eur. J. Pharmacol*. **492**, 269-272.

Emanuelsson O, Nielsen H, von Heijne G (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.* **8**, 978-984.

Emanuelsson O, Brunak S, von Heijne G, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP, and related tools. *Nature Protoc*. **2**, 953-971.

Falara, V., Pichersky, E. and 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.

Gao, W., Hillwig, M.L., Huang, L., Cui, G., Wang, X., Kong, J., Yang, B. and Peters, R.J. (2009) A functional genomics approach to tanshinone biosynthesis provides stereochemical insights. Org. Lett. **11**, 5170-5173.

Gao, Y., Honzatko, R.B. and Peters, R.J. (2012) Terpenoid synthase structures, a so far incomplete view of complex catalysis. *Nat. Prod. Rep*. **29**, 1153-1175.

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. and Huang, L. (2013) CYP76AH1 catalyzes turnover of miltiradiene in tanshinones biosynthesis and enables heterologous production of ferruginol in yeasts. *Proc. Natl. Acad. Sci*. USA **110**, 12108-12113.

Hall, D.E., Zerbe, P., Jancsik, S., Quesada, A.L., Dullat, H., Madilao, L.L., Yuen, M. and 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.

Harris, L.J., Saparno, A., Johnston, A., Prisic, S., Xu, M., Allard, S., Kathiresan, A., Ouellet, T. and Peters, R.J. (2005) The maize An2 gene is induced by *Fusarium* attack and encodes an *ent*-copalyl diphosphate synthase. *Plant Mol. Biol.* **59**, 881-894.

Henderson, M.S. and McCrindle, R. (1969) Premarrubiin. A diterpenoid from *Marrubium vulgare* L. *J. Chem. Soc*. **15**, 2014-2015.

Hillwig, M.L., Mann, F.M. and Peters, R.J. (2011a) Diterpenoid biopolymers, new directions for renewable materials engineering. *Biopolymers* **95**, 71-76.

Hillwig, M.L., Xu, M., Toyomasu, T., Tiernan, M.S., Wei, G., Cui, G., Huang, L. and Peters, R.J. (2011b) Domain loss has independently occurred multiple times in plant terpene synthase evolution. Plant J **68**, 1051-1060.

Hussain, J., Ullah, R., Khan, A., Mabood, F., Shah, M.R., Al-Harrasi, A. and Gilani, A.H. (2011) Antispasmodic and Ca^{++} antagonist potential of marrubiin, a labdane type diterpene from *Phlomis bracteosa*. *J. Pharmacy Res*. 4, 178-180.

Keeling, C.I., Dullat, H.K., Yuen, M., Ralph, S.G., Jancsik, S. and Bohlmann, J. (2010) Identification and functional characterization of monofunctional *ent*-copalyl diphosphate and *ent*-kaurene synthases in white spruce reveal different patterns for diterpene synthase evolution for primary and secondary metabolism in gymnosperms. *Plant Physiol*. **152**, 1197-1208.

Keeling, C.I., Weisshaar, S., Ralph, S.G., Jancsik, S., Hamberger, B., Dullat, H.K. and Bohlmann, J. (2011) Transcriptome mining, functional characterization, and phylogeny of a large terpene

synthase gene family in spruce (*Picea spp*.). *BMC Plant Biol*. **11**, 43.

Knöss, W. and Zapp, J. (1997) Accumulation of furanic diterpenes in *Marrubium vulgare* and *Leonorus cardiaca*. *Planta Med*. **64**, 357-361.

Köksal, M., Potter, K., Peters, R.J. and Christianson, D.W. (2014) 1.55Å-resolution structure of *ent*-copalyl diphosphate synthase and exploration of general acid function by site-directed mutagenesis. *Biochim. Biophys. Acta*. 1840, 184-190.

Kolosova, N., Miller, B., Ralph, S., Ellis, B.E., Douglas, C., Ritland, K. and Bohlmann, J. (2004) Isolation of high-quality RNA from gymnosperm and angiosperm trees. *BioTechniques* **36**, 821-824.

Mafu, S., Hillwig, M.L. and Peters, R.J. (2011) A novel labda-7,13*E*-dien-15-ol-producing bifunctional diterpene synthase from *Selaginella moellendorffii. Chembiochem* **12**, 1984-1987.

Mangoni, L., Adinolfi, M., Laonigro, G. and Caputo, R. (1972) Synthesis of marrubiin. *Tetrahedron* **28**, 611-621.

Martin, D.M., Fäldt, J. and Bohlmann, J. (2004) Functional characterization of nine Norway Spruce TPS genes and evolution of gymnosperm terpene synthases of the TPS-d subfamily. *Plant Physiol*. **135**, 1908-1927.

Meyre-Silva, C., Yunes, R.A., Schlemper, V., Campos-Buzzi, F. and Cechinel-Filho, V. (2005) Analgesic potential of marrubiin derivatives, a bioactive diterpene present in *Marrubium vulgare* (Lamiaceae). *Il Farmaco* **60**, 321-326.

Meyre-Silva, C. and Cechinel-Filho, V. (2010) A review of the chemical and pharmacological aspects of the genus *Marrubium. Curr. Pharm. Des*. **16**, 3503-3518.

Mnonopi, N., Levendal, R.A., Davies-Coleman, M.T. and Frost, C.L. (2011) The cardioprotective effects of marrubiin, a diterpenoid found in *Leonotis Leonurus* extracts. *J. Ethnopharmacol*. **138**, 67-75.

Mnonopi, N., Levendal, R.A., Mzilikazi, N. and Frost, C.L. (2012) Marrubiin, a constituent of *Leonotis leonurus*, alleviates diabetic symptoms. *Phytomedicine* **19**, 488-493.

Morgenstern, B. (2004) DIALIGN, multiple DNA and protein sequence alignment at BiBiServ. *Nucl. Acids Res.* **32**, W33-36.

Morrone, D., Lowry, L., Determan, M.K., Hershey, D.M., Xu, M. and Peters, R.J. (2010) Increasing diterpene yield with a modular metabolic engineering system in *E. coli*, comparison of MEV and MEP isoprenoid precursor pathway engineering. *Appl. Microbiol. Biotechnol*. **85**, 1893-1906.

Morrone, D., Hillwig, M.L., Mead, M.E., Lowry, L., Fulton, D.B. and 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.

Pateraki, I., Andersen-Ranberg, J., Hamberger, B., Heskes, A.M., Martens, H.J., Zerbe, P., Bach, S.S., Møller, B.L., Bohlmann, J., Hamberger, B. (2014) Manoyl oxide (13*R*), the biosynthetic precursor of forskolin, is synthesized in specialized root cork cells in *Coleus forskohlii*. *Plant Physiol*. **164**, 1222-1236.

Peters, R.J., Flory, J.E., Jetter, R., Ravn, M.M., Lee, H.J., Coates, R.M. and 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.

Peters, R.J. (2010) Two rings in them all, The labdane-related diterpenoids. *Nat. Prod. Rep*. **27**, 1521.

Popoola, O.K., Elbagory, A.M., Ameer, F. and Hussein, A.A. (2013) Marrubiin. *Molecules* **18**, 9049-9060.

Salei, L.A., Popa, D.P., Doleish, L. and Lazur'evskii, G.W. (1967) The structure of peregrinol, a diterpenoid from *Marrubium peregrinum. Khimiya Prirodnykh Soedinenii* **8**, 90-95.

Sallaud, C., Giacalone, C., Töpfer, R., Goepfert, S., Bakaher, N., Rösti, S. and 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.

Schalk M, Pastore L, Mirata MA, 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.

Sugai, Y., Ueno, Y., Hayashi, K., Oogami, S., Toyomasu, T., Matsumoto, S., Natsume, M., Nozaki, H. and 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.

Sun, T.P. and Kamiya, Y. (1994) The *Arabidopsis* GA1 locus encodes the cyclase *ent*-kaurene synthetase A of gibberellin biosynthesis. *Plant Cell* **6**, 1509-1518.

Voinnet, O., Rivas, S., Mestre, P. and 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.

Wu, Y., Zhou, K., Toyomasu, T., Sugawara, C., Oku, M., Abe, S., Usui, M., Mitsuhashi, W., Chono, M., Chandler, P.M. and 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.

Wurch, T., Lestienne, F., and Pauwels, P.J. (1998) A modified overlap extension PCR method to create chimeric genes in the absence of restriction enzymes. *Biotechnol. Tech.* **12**, 653-657.

Xu, M., Wilderman, P.R., Morrone, D., Xu, J., Roy, A., Margis-Pinheiro, M., Upadhyaya, N.M., Coates, R.M. and Peters, R.J. (2007) Functional characterization of the rice kaurene synthase-like gene family. *Phytochemistry* **68**, 312-326.

Yamaguchi, S., Sun, T.P., Kawaide, H. and Kamiya, Y. (1998) The GA2 locus of *Arabidopsis thaliana* encodes *ent*-kaurene synthase of gibberellin biosynthesis. *Plant Physiol*. **116**, 1271-1278.

Zerbe, P., Chiang, A., Yuen, M., Hamberger, B., Hamberger, B., Draper, J.A., Britton, R. and Bohlmann, J. (2012a) Bifunctional *cis*-abienol synthase from *Abies balsamea* discovered by transcriptome sequencing and its implications for diterpenoid fragrance production. *J. Biol. Chem.* **287**, 12121-12131.

Zerbe, P., Chiang, A. and Bohlmann, J. (2012b) Mutational analysis of white spruce (*Picea glauca*) *ent*-kaurene synthase (*Pg*KS) reveals common and distinct mechanisms of conifer diterpene synthases of general and specialized metabolism. *Phytochemistry* **74**, 30-39.

Zerbe, P. and Bohlmann, J. (2014) Metabolic engineering of terpene synthases of conifer defense, bioproducts, biofuels and perfumes. *Recent Adv. Phytochem.* (in press).

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

Zhou, Y.J., Gao, W., Rong, Q., Jin, G., Chu, H., Liu, W., Yang, W., Zhu, Z., Li, G., Zhu, G., Huang, L. and Zhao, Z.K. (2012a) Modular pathway engineering of diterpenoid synthases and the mevalonic acid pathway for miltiradiene production. *J. Am. Chem. Soc*. **134**, 3234-3241.

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

TABLES

Table 1: Diterpene synthase candidate sequences identified in the leaf transcriptome of *Marrubium vulgare*

Gene ID	FL	NCBI Annotation	*E*-value (closest match)	Functional Motifs
***Mv*CPS1**	+	Copalyl diphosphate synthase, *Salvia miltiorrhiza* [ABV57835.1]	0	DIDD
***Mv*CPS2**	-	Copalyl diphosphate synthase, *Scoparia dulcis* [BAD91286.1]	0	DIDD
***Mv*CPS3**	+	Copalyl diphosphate synthase, *Salvia miltiorrhiza* [ABV57835.1]	0	DIDD
***Mv*EKS**	-	Kaurene synthase 2, *Salvia miltiorrhiza* [AEZ55689.1]	$6*10^{-155}$	DDFFD, NDVQTFKKE
***Mv*ELS**	+	Kaurene synthase, *Salvia miltiorrhiza* [ABV08817.1]	0	DDFFD, NDIQTCERE

FIGURE LEGENDS

Figure 1: Diterpene synthases in the diterpenoid biosynthetic system of the medicinal plant *Marrubium vulgare*

Illustrated is a general proposed biosynthesis system of marrubiin and related diterpenoids based on known compounds from different *Marrubium* species. Diterpene synthases form the foundation of the diterpene chemical diversity in *Marrubium*, producing distinct diterpene scaffolds. DiTPS products and other downstream pathway intermediates are further modified by cytochromes P450 and other enzyme classes to afford the array of specialized diterpenoids. Structures are illustrated as published in the literature. Marrubiin and related compounds are thought to occur with a 9,13-epoxide bridge *in planta*, and produce a free hydroxyl group upon metabolite extraction (Henderson and McCrindle, 1969; Knöss and Zapp, 1997).

Figure 2: Terpenoid profiles of different organs of *Marrubium vulgare*

Heat map showing abundance of terpenoids detected in above-ground organs and stem trichomes of *M. vulgare*. Metabolites were isolated from ground tissues by pentane extraction and identified by GC-MS analysis and comparison to authentic marrubiin and reference mass spectra databases (Wiley W9N08L). Quantitation was based on 1-eicosene as internal standard (n=3). Detailed results of the quantitative analysis are given in Table S1.

Figure 3: Phylogeny of *Marrubium vulgare* diterpene synthases

The maximum likelihood tree illustrates the phylogenetic relatedness of *M. vulgare* diterpene synthases and known diterpene synthases of other species of the TPS-c and TPS-e/f clades. The ancestral *Physcomitrella patens ent*-kaurene/kaurenol synthase was used to root the tree. Branches with bootstrap support of >80% (100 repetitions) are highlighted with grey dots. Descriptions of diterpene synthases used in the phylogeny are listed in Table S3.

Figure 4: Relative transcript abundance of *Marrubium vulgare* diterpene synthases in flowers (corolla) and leaves

Transcript abundance was measured by qRT-PCR and normalized to the internal reference actin. Error bars represent standard errors based on triplicate measurements of three biological replicates. Statistically significant differences between transcript abundance in leaves and flowers were calculated using a Student's t-test ($p<0.05$), and two-tailed p-values are highlighted. Reaction specificity was verified by dissociation curves and sequence verification of representative products.

Figure 5: Diterpene product formation by *Mv*CPS3

[**A**] Chiral GC-MS analysis of the dephosphorylated reaction product of recombinant *Mv*CPS3 with GGPP as substrate, compared to dephosphorylated (+)-CPP and *ent*-CPP produced by *Picea abies* PaLAS:D611A (+)-CPP synthase (Zerbe *et al*., 2012a) and *Zea mayz Zm*AN2 *ent*-CPP synthase (Harris *et al*., 2005), respectively. [**B**] GC-MS analysis of *Picea glauca ent*-KS (*Pg*EKS; Keeling *et al*., 2010) in coupled *in vitro* assays with *Mv*CPS3 or *Zm*AN2. Comparison with authentic *ent*-kaurene showed *Pg*EKS activity when combined with *Zm*AN2, but not in coupled assays with *Mv*CPS3, suggesting *Mv*CPS3 does not produce *ent*-CPP. [**C**] GC-MS analysis of the reaction product of the (+)-CPP-specific *Pinus contorta* pimaradiene synthase (*Pc*PIM1; Hall *et al*., 2013) in coupled *in vitro* assays with *Mv*CPS3 or *Zm*AN2. Comparison with authentic pimaradiene showed *Pc*PIM1 activity when coupled with *Mv*CPS3, but not in assays with *Zm*AN2, suggesting *Mv*CPS3 produces (+)-CPP. Peak *a*, geranylgeraniol (dephosphorylated GGPP), peak *b*, (+)-copalol; peak *c*, *ent*-copalol (dephosphorylated *ent*-CPP); peak *IS*, internal standard; peak *d*, *ent*-kaurene; peak *e*, pimaradiene.

Figure 6: Diterpene product formation by *Mv*CPS1

[**A**] Extracted ion chromatograms (EIC) of products from *in vitro* assays of *Mv*CPS1 with GGPP as substrate. Product formation was observed only after dephosphorylation and consisted of four compounds (peaks *f-i*) with near identical mass spectra matching the published fragmentation pattern of peregrinol (Salei *et al*., 1967) [**B** and **C**]. [**D**] Cool-on-column GC-MS analysis verified compound *i* as major *Mv*CPS1 product. [**E**] Proposed *Mv*CPS1-catalyzed reaction mechanism via cyclo-isomerization of GGPP to a labda-13-en-8-yl^{+} intermediate, followed by a

hydride transfer between carbons C-8 and C-9 and carbocation neutralization by water capture to afford peregrinol diphosphate.

Figure 7: Diterpene product formation by *Mv*EKS and *Mv*ELS

[**A**] Total ion chromatograms (TIC) of *ent*-kaurene produced in coupled *in vitro* assays of *Zm*AN2 and *MvE*KS, compared to the authentic standard. [**B-D**] Extracted ion chromatograms (EIC) of products from coupled *in vitro* assays of *Mv*ELS with *Mv*CPS1, *Mv*CPS3 or *C. forskohlii Cf*TPS2 (Pateraki *et al*., 2014). [**B**] Coupled assays with *Mv*CPS1 (producing PPP) resulted in the formation of 9,13-epoxy-labd-14-ene. [**C**] Combined assays of the (+)-CPS *Mv*CPS3 with *Mv*ELS yielded miltiradiene verified by comparison to the pairwise reaction of the (+)-CPP synthase *Cf*TPS1 and the miltiradiene/manoyl oxide synthase *Cf*TPS4 from *C. forskohlii* (Pateraki *et al*., 2014). [**D**] Combination of the LPS *Cf*TPS2 with *Mv*ELS afforded manoyl oxide as identified by comparison to the known coupled activity of *Cf*TPS2 and *Cf*TPS4 (Pateraki *et al*., 2014). [**E**] Mass spectra of 9,13-epoxy-labd-14-ene (peak *g*), miltiradiene (peak *k*) and manoyl oxide (peak *l*) produced by *Mv*ELS. [**F**] Relative abundance of products formed by *Mv*CPS1 alone (after dephosphorylation) and in coupled assays with *Mv*ELS as based on three independent duplicate measurements. [**G**] NMR analysis of 9,13-epoxy-labd-14-ene produced in the pairwise reaction of *Mv*CPS1 and *Mv*ELS. ^{1}H-NMR (600 MHz, MeOD): 6.07 (1H, dd, J=17.4, 10.9 Hz, H-14), 5.16 (1H, dd, J=17.4, 1.6 Hz, H-15), 4.94 (1H, dd, J=10.9, 1.6 Hz, H-15), 1.30 (3H, s, H-16), 0.92 (3H, s, H‑20), 0.87 (3H, s, H-18), 0.83 (3H, s, H-19), 0.82 (3H, d, J=6.8 Hz, H-17). ^{13}C-NMR (MeOD) ^{13}C obtained from HSQC and HMBC spectra: 147.2 (1C, C-14), 110.7 (1C, C15), 93.8 (1C, C-9), 84.6 (1C, C-13), 48.4 (1C, C-5), 43.0 (1C, C-10) 42.7 (1C, C‑4), 38.0 (2C, C-8, C12), 34.0 (2C, C-1, C-3), 33.7 (1C, C-18), 32.5 (1C, C-7), 30.5

(1C, C-11), 23.5 (1C, C-16), 22.9 (1C, C-2), 22.2 (1C, C-19), 19.9 (1C, C-6), 18.4 (1C, C-17), 17.8 (1C, C-20). Arrows illustrate the HMBC correlations of the extracyclic C-14-C-15 double bond of the 9,13-epoxy bridge. Peak *IS*, internal standard; peak *j*, *ent*-kaurene; peak *f* and *h*, unidentified diterpenes produced by *Mv*CPS1 and *Mv*ELS; peak *g*, 9,13-epoxy-labd-14-ene; peak *i*, dephosphorylated PPP; peak *k*, miltiradiene; peak *l*, manoyl oxide.

Fig. 8: Diterpene synthase co-expression in *Nicotiana benthamiana*

Extracted ion chromatograms (EIC) of extracts from *N. benthamiana* leaves after *Agrobacterium*-mediated transient (co-)expression of *Mv*CPS1, *Mv*CPS2 and *Mv*CPS3 with *Mv*EKS or *Mv*ELS. Expression was performed together with the plant viral protein p19 to suppress RNA silencing and compared to p19 expression alone [**A**]. Product formation was evaluated four days post infiltration with identification based on reference mass spectra or authentic standards. [**B**] Formation of dephosphorylated PPP and the three additional products *f-h* after expression of *Mv*CPS1. [**C**] Co-expression of *Mv*CPS1 and *Mv*EKS showing no additional class I diTPS products. [**D**] Co-expression of *Mv*CPS1 and *Mv*ELS yielded compounds *f-h* with compound *g* as major product. [**E-G**] No product formation was detected after expression of *Mv*CPS2 alone or in combination with *Mv*EKS or *Mv*ELS. [**H**] Expression of *Mv*CPS3 resulted in (+)-copalol and an unidentified diterpene not observed in *in vitro* assays. [**I**] Co-expression of *Mv*CPS3 and *Mv*EKS showed an additional minor diterpene product (peak *n*). [**J**] When co-expressing *Mv*CPS3 and *Mv*ELS no (+)-copalol was detectable, while miltiradiene was produced. Peak *f* and *h*, unidentified diterpenes produced by *Mv*CPS1 and *Mv*ELS; peak *g*, 9,13-epoxy-labd-14-ene; peak *i*, dephosphorylated PPP; peak *b*, (+)-copalol (dephosphorylated (+)-CPP); peak *m* and *n*, unidentified diterpene product; peak *k*, miltiradiene; peak *o*, ethanone.

Figure 1

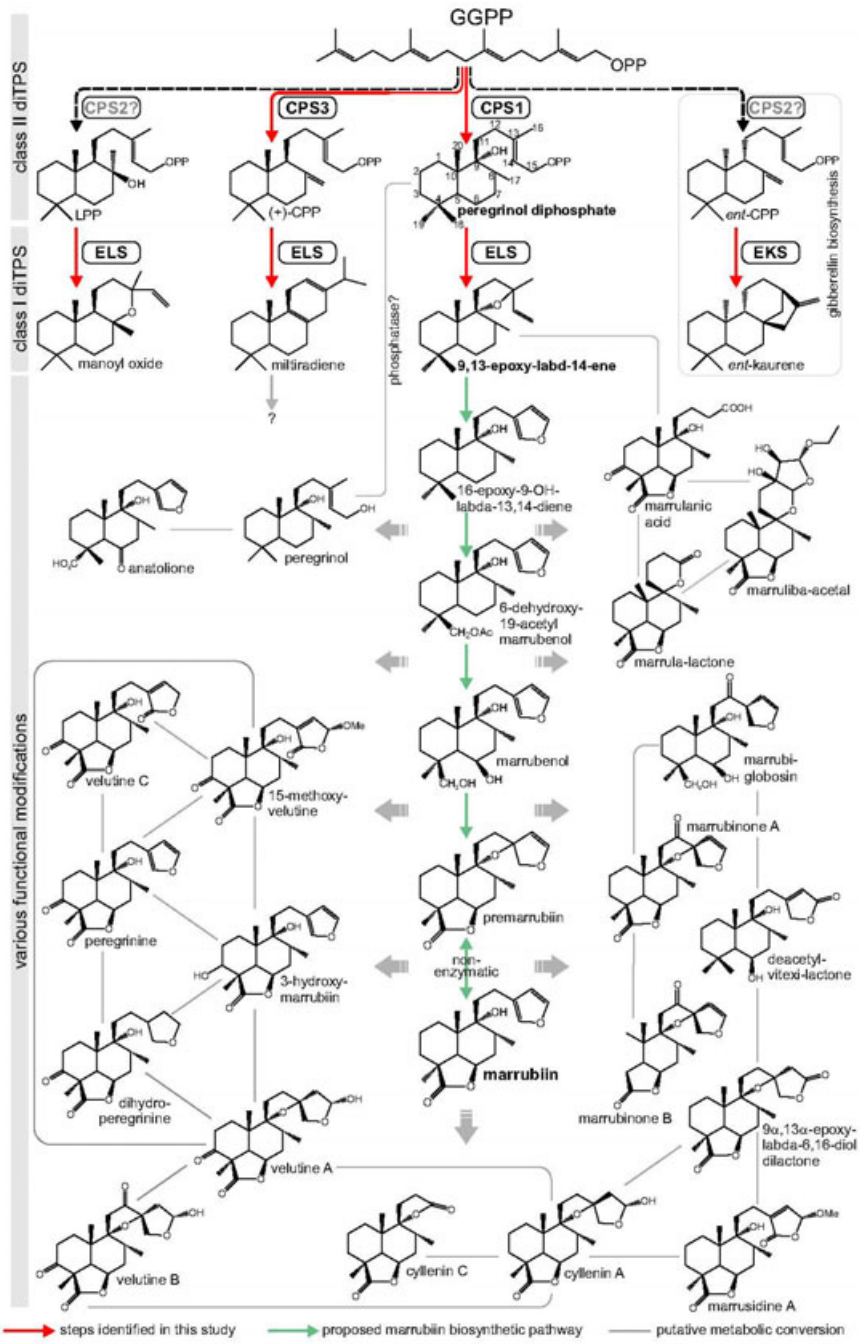

Fig. 1: Diterpene synthases in the diterpenoid biosynthetic system of the medicinal plant *Marrubium vulgare*

Illustrated is a general proposed biosynthesis system of marrubiin and related diterpenoids based on known compounds from different *Marrubium* species. Diterpene synthases form the foundation of the diterpene chemical diversity in *Marrubium*, producing distinct diterpene scaffolds. DiTPS products and other downstream pathway intermediates are further modified by cytochromes P450 and other enzyme classes to afford the array of specialized diterpenoids. Structures are illustrated as published in the literature. Marrubiin and related compounds are thought to occur with a 9,13-epoxide bridge *in planta*, and produce a free hydroxyl group upon metabolite extraction (Henderson and McCrindle, 1969; Knöss and Zapp, 1997).

Figure 2

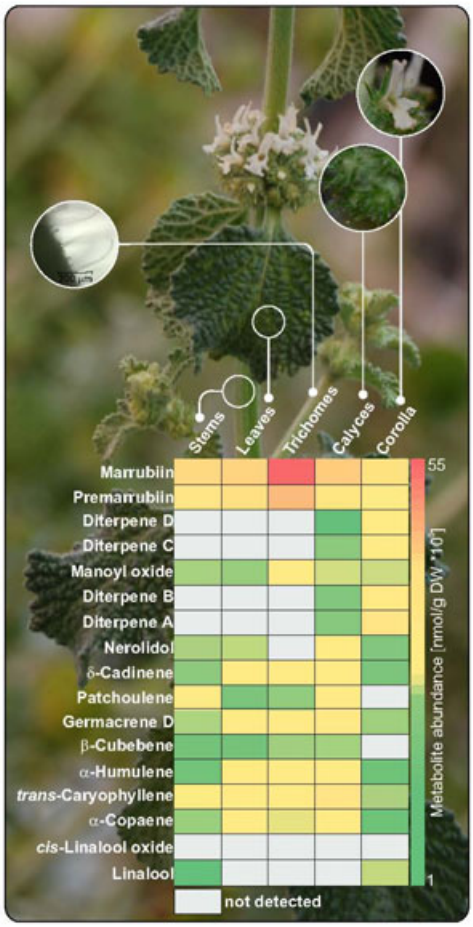

Fig. 2: Terpenoid profiles of different organs of *Marrubium vulgare*
Heat map showing abundance of terpenoids detected in above ground organs and stem trichomes of *M. vulgare*. Metabolites were isolated from ground tissues by pentane extraction and identified by GC-MS analysis as compared to authentic marrubiin and reference mass spectra databases (Wiley W9N08L). Quantitation was based on 1-eicosene as internal standard (n=3). Detailed results of the quantitative analysis are given in Table S1.

Figure 3

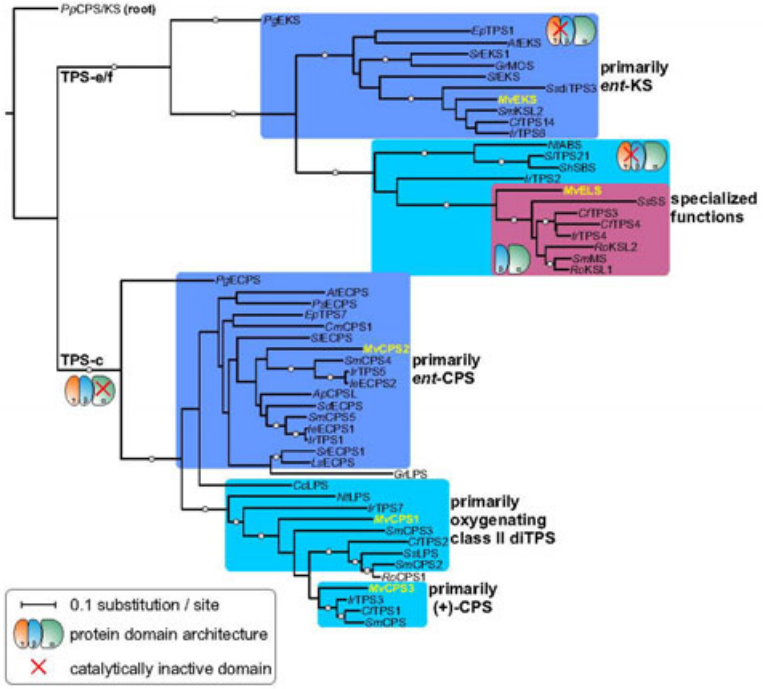

Fig. 3: Phylogeny of *Marrubium vulgare* diterpene synthases
The maximum likelihood tree illustrates the phylogenetic relatedness of *M. vulgare* diterpene synthases and known diterpene synthases of other species of the TPS-c and TPS-e/f clades. The ancestral *Physcomitrella patens ent*-kaurene/kaurenol synthase was used to root the tree. Branches with bootstrap support of >80% (100 repetitions) are highlighted with grey dots. Descriptions of diterpene synthases used in the phylogeny are listed in Table S3.

Figure 4

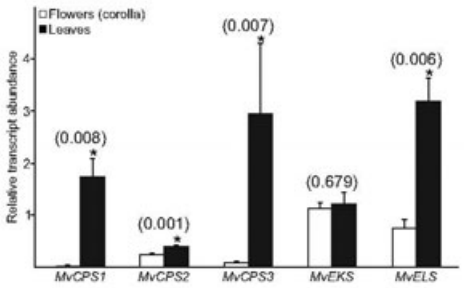

Fig. 4: Relative transcript abundance of *Marrubium vulgare* diterpene synthases in flowers (corolla) and leaves

Transcript abundance was measured by qRT-PCR and normalized to the internal reference actin. Error bars represent standard errors based on triplicate measurements of three biological replicates. Statistically significant differences between transcript abundance in leaves and flowers were calculated using a student's t-test (p <0.05) and two-tailed p-values are highlighted. Reaction specificity was verified by dissociation curves and sequence verification of representative products.

Figure 5

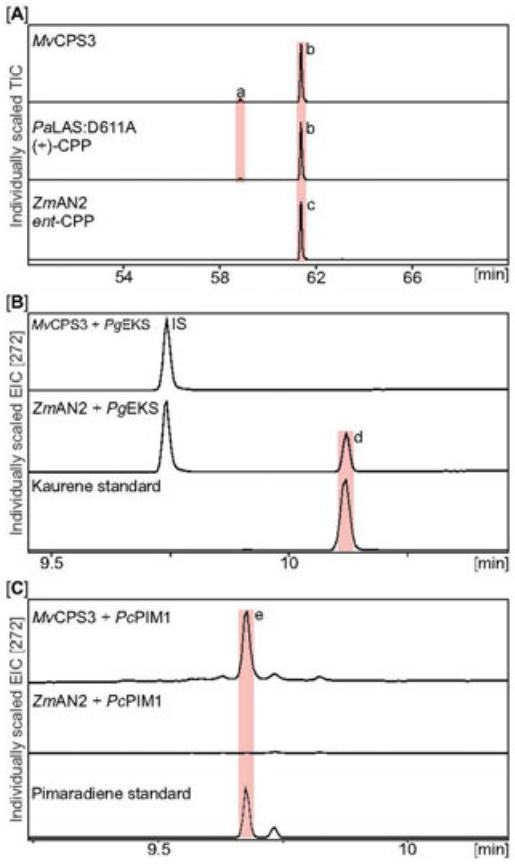

Fig. 5: Diterpene product formation by *Mv*CPS3
[**A**] Chiral GC-MS analysis of the dephosphorylated reaction product of recombinant *Mv*CPS3 with GGPP as substrate, compared to dephosphorylated (+)-CPP and *ent*-CPP produced by *Picea abies Pa*LAS:D611A (+)-CPP synthase (Zerbe *et al*., 2012a) and *Zea mayz Zm*AN2 *ent*-CPP synthase (Harris *et al*., 2005), respectively. [**B**] GC-MS analysis of *Picea glauca ent*-KS (*Pg*EKS; Keeling *et al*., 2010) in coupled *in vitro* assays with *Mv*CPS3 or *Zm*AN2. Comparison with authentic *ent*-kaurene showed *Pg*EKS activity when combined with *Zm*AN2, but not in coupled assays with *Mv*CPS3, suggesting *Mv*CPS3 does not produce *ent*-CPP. [**C**] GC-MS analysis of the reaction product of the (+)-CPP-specific *Pinus contorta* pimaradiene synthase (*Pc*PIM1; Hall *et al*., 2013) in coupled *in vitro* assays with *Mv*CPS3 or *Zm*AN2. Comparison with authentic pimaradiene showed *Pc*PIM1 activity when coupled with *Mv*CPS3, but not in assays with *Zm*AN2, suggesting *Mv*CPS3 produces (+)-CPP. Peak *a*, geranylgeraniol (dephosphorylated GGPP), peak *b*, (+)-copalol; peak *c*, *ent*-copalol (dephosphorylated *ent*-CPP); peak *IS*, internal standard; peak *d*, *ent*-kaurene; peak *e*, pimaradiene.

Figure 6

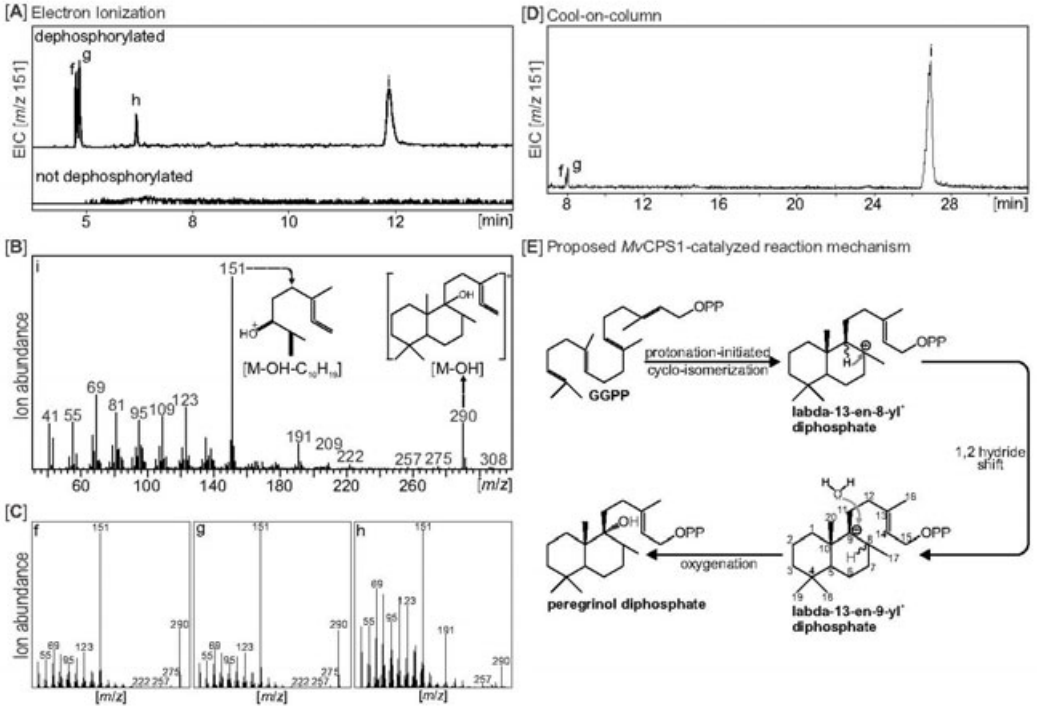

Fig. 6: Diterpene product formation by *Mv*CPS1

[**A**] Extracted ion chromatograms (EIC) of products from *in vitro* assays of *Mv*CPS1 with GGPP as substrate. Product formation was observed only after dephosphorylation and consisted of four compounds (peaks *f-i*) with near identical mass spectra matching the published fragmentation pattern of peregrinol (Salei *et al.*, 1967) [**B** and **C**]. [**D**] Cool-on-column GC-MS analysis verified compound *i* as major *Mv*CPS1 product. [**E**] Proposed *Mv*CPS1-catalyzed reaction mechanism via cyclo-isomerization of GGPP to a labda-13-en-8-yl+ intermediate, followed by C-9,8 hydride transfer between carbons C-8 and C-9 and carbocation neutralization by water capture to afford peregrinol diphosphate.

Figure 7

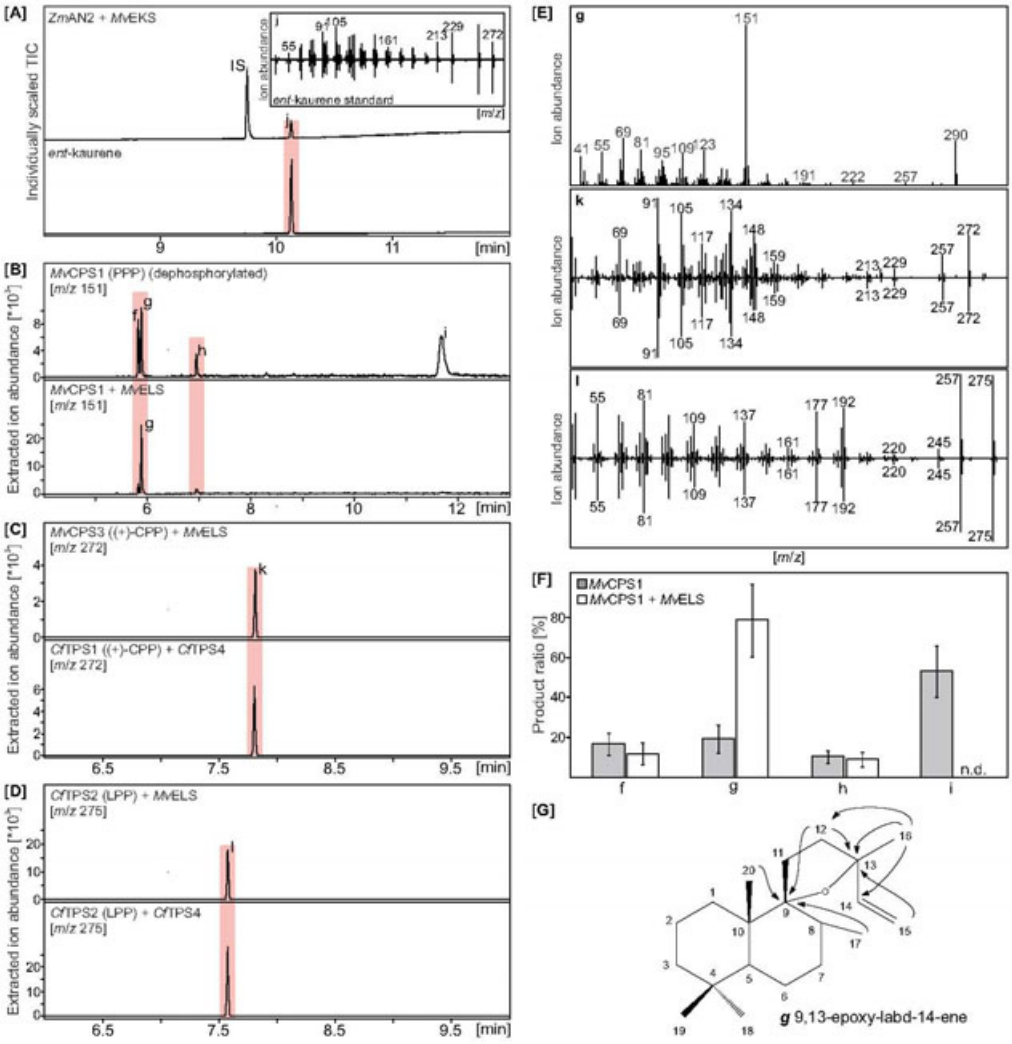

Fig. 7: Diterpene product formation by *Mv*EKS and *Mv*ELS.
[**A**] Total ion chromatograms (TIC) of *ent*-kaurene produced in coupled *in vitro* assays of *Zm*AN2 and *Mv*EKS, compared to the authentic standard. [**B-D**] Extracted ion chromatograms (EIC) of products from coupled *in vitro* assays of *Mv*ELS with *Mv*CPS1, *Mv*CPS3 or *C. forskohlii Cf*TPS2 (Pateraki *et al*., 2014). Coupled assays with *Mv*CPS1 (producing PPP) resulted in the formation of 9,13-epoxy-labd-14-ene [**B**]. Combined assays of the (+)-CPS *Mv*CPS3 with *Mv*ELS yielded miltiradiene verified by comparison to the pairwise reaction of the (+)-CPP synthase *Cf*TPS1 and the miltiradiene/manoyl oxide synthase *Cf*TPS4 from *C. forskohlii* (Pateraki *et al*., 2014) [**C**]. Combination of the LPS *Cf*TPS2 with *Mv*ELS afforded manoyl oxide as compared to the known coupled activity of *Cf*TPS2 and *Cf*TPS4 (Pateraki *et al*., 2014) [**D**]. [**E**] Mass spectra of 9,13-epoxy-labd-14-ene (peak *g*), miltiradiene (peak *k*) and manoyl oxide (peak *l*) produced by *Mv*ELS. [**F**] Relative abundance of products formed by *Mv*CPS1 alone (after dephosphorylation) and in coupled assays with *Mv*ELS as based on three independent duplicate measurements. [**G**] NMR analysis of 9,13-epoxy-labd-14-ene produced in the pairwise reaction of *Mv*CPS1 and *Mv*ELS. ^{1}H-NMR (600 MHz, MeOD): 6.07 (1H, dd, J=17.4, 10.9 Hz, H-14), 5.16 (1H, dd, J=17.4, 1.6 Hz, H-15), 4.94 (1H, dd, J=10.9, 1.6 Hz, H-15), 1.30 (3H, s, H-16), 0.92 (3H, s, H-20), 0.87 (3H, s, H-18), 0.83 (3H, s, H-19), 0.82 (3H, d, J=6.8 Hz, H-17). ^{13}C-NMR (MeOD) ^{13}C obtained from HSQC and HMBC spectra: 147.2 (1C, C-14), 110.7 (1C, C15), 93.8 (1C, C-9), 84.6 (1C, C-13), 48.4 (1C, C-5), 43.0 (1C, C-10) 42.7 (1C, C-4), 38.0 (2C, C-8, C12), 34.0 (2C, C-1, C-3), 33.7 (1C, C-18), 32.5 (1C, C-7), 30.5 (1C, C-11), 23.5 (1C, C-16), 22.9 (1C, C-2), 22.2 (1C, C-19), 19.9 (1C, C-6), 18.4 (1C, C-17), 17.8 (1C, C-20). Arrows illustrate the HMBC correlations of the extracyclic C-14-C-15 double bond of the 9,13-epoxy bridge.
Peak **IS**, internal standard; peak *j*, *ent*-kaurene; peak *f* and *h*, unidentified diterpenes produced by *Mv*CPS1 and *Mv*ELS; peak *g*, 9,13-epoxy-labd-14-ene; peak *i*, dephosphorylated PPP; peak *k*, miltiradiene; peak *l*, manoyl oxide.

Figure 8

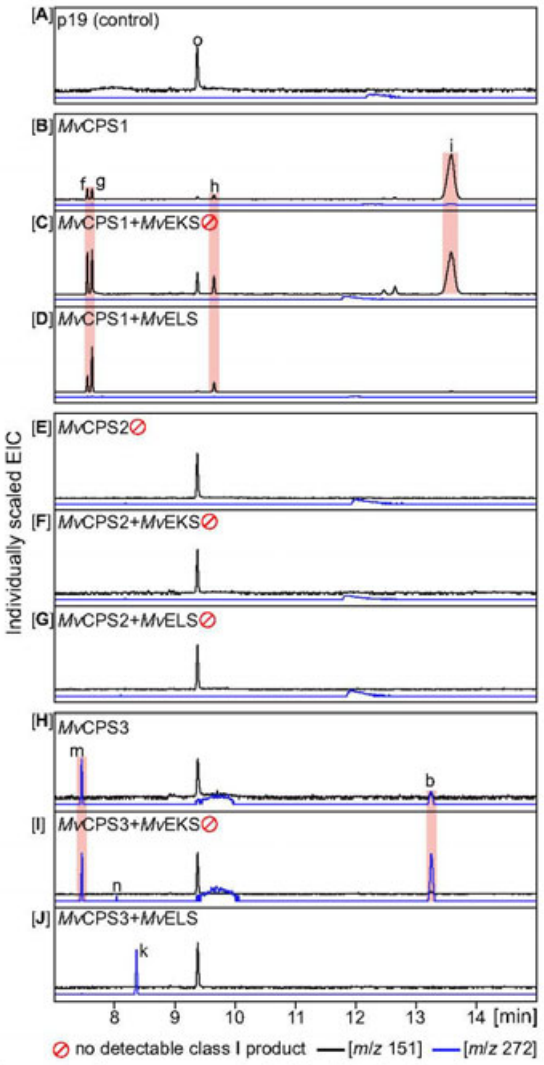

Fig. 8: Diterpene synthase co-expression in *Nicotiana benthamiana*
Extracted ion chromatograms (EIC) of extracts from *N. benthamiana* leaves after *Agrobacterium*-mediated transient (co-)expression of *Mv*CPS1, *Mv*CPS2 and *Mv*CPS3 with *Mv*EKS or *Mv*ELS. Expression was performed together with the plant viral protein p19 to suppress RNA silencing and compared to p19 expression alone [**A**]. Product formation was evaluated four days post infiltration with identification based on reference mass spectra or authentic standards. [**B**] Formation of dephosphorylated PPP and the three additional products *f-h* after expression of *Mv*CPS1. [**C**] Co-expression of *Mv*CPS1 and *Mv*EKS showing no additional class I diTPS products. [**D**] Co-expression of *Mv*CPS1 and *Mv*ELS yielded compounds *f-h* with compound *g* as major product. [**E-G**] No product formation was detected after expression of *Mv*CPS2 alone or in combination with *Mv*EKS or *Mv*ELS. [**H**] Expression of *Mv*CPS3 resulted in (+)-copalol and an unidentified diterpene not observed in *in vitro* assays. [**I**] Co-expression of *Mv*CPS3 and *Mv*EKS showed an additional minor diterpene product (peak *n*). [**J**] When co-expressing *Mv*CPS3 and *Mv*ELS no (+)-copalol was detectable, while miltiradiene was produced. Peak *f* and *h*, unidentified diterpenes produced by *Mv*CPS1 and *Mv*ELS; peak *g*, 9,13-epoxy-labd-14-ene; peak *i*, dephosphorylated PPP; peak *b*, (+)-copalol (dephosphorylated (+)-CPP); peak *m* and *n*, unidentified diterpene product; peak *k*, miltiradiene; peak *o*, ethanone.