Biosynthesis of bioactive diterpenoids in the medicinal plant Vitex agnus-castus

Heskes, Allison Maree; Sundram, Tamil C M; Boughton, Berin Alain; Jensen, Niels Bjerg; Hansen, Nikolaj Lervad; Crocoll, Christoph; Cozzi, Federico; Rasmussen, Simon; Hamberger, Britta; Hamberger, Björn; Stærk, Dan ; Møller, Birger Lindberg; Pateraki, Irini

Published in:
Plant Journal

Link to article, DOI:
10.1111/tpj.13822

Publication date:
2018

Document Version
Peer reviewed version

Link back to DTU Orbit

Citation (APA):
This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/tpj.13822
This article is protected by copyright. All rights reserved.
SUMMARY

_Vitex agnus-castus_ L. (Lamiaceae) is a medicinal plant historically used throughout the Mediterranean region to treat menstrual cycle disorders and is still used today as a clinically effective treatment for premenstrual syndrome. The pharmaceutical activity of the plant extract is linked to its ability to lower prolactin levels. This feature has been attributed to the presence of dopaminergic diterpenoids that can bind to dopamine receptors in the pituitary gland. Phytochemical analyses of _V. agnus-castus_ show that it contains an enormous array of structurally related diterpenoids and, as such holds potential as a rich source of new dopaminergic drugs. The present work investigated the localisation and biosynthesis of diterpenoids in _V. agnus-castus_. With
the assistance of matrix assisted laser desorption ionization-mass spectrometry imaging (MALDI-MSI), diterpenoids were localised to trichomes on the surface of fruit and leaves. Analysis of a trichome-specific transcriptome database, coupled with expression studies, identified seven candidate genes involved in diterpenoid biosynthesis: three class II diterpene synthases (diTPSs), three class I diTPSs, and a cytochrome P450 (CYP). Combinatorial assays of the diTPSs resulted in the formation of a range of different diterpenes that can account for several of the backbones of bioactive diterpenoids observed in *V. agnus-castus*. The identified CYP, *VacCYP76BK1*, was found to catalyse 16-hydroxylation of the diol-diterpene, peregrinol, to labd-13Z-ene-9,15,16-triol when expressed in *Saccharomyces cerevisiae*. Notably, this product is a potential intermediate in the biosynthetic pathway towards bioactive furan and lactone containing diterpenoids that are present in this species.

**INTRODUCTION**

Plants are a rich source of bioactive compounds, with numerous species used for millennia in traditional medicines. Advances in bioassay technology coupled with modern analytical techniques have now made it relatively straightforward to identify specific bioactivities and to pinpoint the responsible active components in medicinal plants (Wubshet *et al.*, 2016). The subsequent drug development process and (potential) wide-scale use of any identified compounds requires a reliable and scalable source of material that is both ethically and commercially viable. This requirement is often not met by the natural plant source due to difficulties with cultivation, endangered wild populations, and the inherent chemical complexity found in most medicinal plant species. While chemical synthesis can potentially provide a solution to this problem, the large number of stereocenters present in natural products makes this route challenging technically and economically. Another approach to solving this problem is via the engineering of microbial organisms to produce specific natural products using biosynthetic pathways reconstructed from the original or even a
combination of plant species. Examples using *Saccharomyces cerevisiae* as a host to make important plant derived pharmaceuticals and nutraceuticals are already available and include: artemisinic acid, a precursor of the anti-malarial drug artimisinin naturally produced by the plant *Artemisia annua* (Paddon *et al.*, 2013), the antioxidant compound carnosic acid from *Rosmarinus officinalis* (Ignea *et al.*, 2016; Scheler *et al.*, 2016), and the cAMP booster, forskolin, from *Coleus forskohlii* (Pateraki *et al.*, 2017). Through metabolic engineering of host organisms, routes to specific compounds can be optimised to increase yields and the generation of further chemical diversity explored through the application of combinatorial techniques (Ignea *et al.*, 2015; Andersen-Ranberg *et al.*, 2016; Jia *et al.*, 2016). For such an approach to work, the enzymes involved in the biosynthesis of a given compound are required and this has spurred interest in biosynthetic pathway discovery in medicinal plants.

The present work focuses on diterpenoid pathway discovery in *Vitex agnus-castus* L. (Lamiaceae), a shrub that grows throughout the Mediterranean region and parts of Asia. It has a long history of use in the treatment of female reproductive conditions and is still used today to treat premenstrual syndrome (PMS). The efficacy of crude fruit extracts for relieving symptoms of PMS and related syndromes is strongly supported by evidence from multiple clinical trials (Schellenberg, 2001; Atmaca *et al.*, 2003; He *et al.*, 2009) and include relief from latent hyperprolactinemia and cyclic mastalgia (Kilicdag *et al.*, 2004; Carmichael, 2008). In addition, extracts are considered to be safe and better tolerated than conventional hormone or selective serotonin reuptake inhibitor based treatments (Daniele *et al.*, 2005). The pharmacological properties of *V. agnus-castus* are attributed to the presence of dopamine, opioid, and oestrogen receptor ligands that are involved in modulating hormone levels that impact PMS (Wuttke *et al.*, 2003; Webster *et al.*, 2011). Specifically, those compounds thought to be responsible for the observed prolactin lowering effect are diterpenoids that interact with dopamine receptors in the pituitary gland (Hoberg *et al.*, 1999; Jarry *et al.*, 2006; Brattström, 2014). The ability of *V. agnus-castus* diterpenoids to interact with both D2 and D3
dopamine receptors in the brain has also flagged it as a potential source of new drugs that could treat other illnesses in which dopaminergic pathways play key roles (Brattström, 2014).

A large variety of labdane-related diterpenoids has been isolated from the Vitex genus, with a recent review listing 114 unique structures covering a range of classes including labdanes, abietanes and clerodanes (Yao et al., 2016). The majority of diterpenoids isolated from V. agnus-castus incorporate hydroxyl groups at C-9 and C-13 (Figure 1a; Hoberg et al., 1999; Ono et al., 2008; Ono et al., 2009; Ono et al., 2011). Common structural features of V. agnus-castus bioactive diterpenoids are the presence of furan, lactone or 3-hydroxy-3-methyl-pent-4-enyl groups (König, 2014). Furan or lactone groups, in particular, are frequent features of bioactive diterpenoids and appear, in some cases, to be important for their activity (Lim et al., 2012, Shul’ts et al. 2014). For example, work on elucidating the pharmacophore of salvinorin A, indicates that the furan ring influences its kappa opioid receptor agonist activity (Munro et al., 2005; Riley et al., 2014). Because of their potential pharmaceutical applications, the elucidation of biosynthetic pathways towards furan and lactone diterpenoids is of particular interest (Zerbe et al., 2014; Pelot et al., 2017).

Terpene biosynthesis is carried out in plants by a family of related enzymes known as terpene synthases (TPSs). These divide into subfamilies based on phylogenetic relationships, with some degree of functional conservation (Bohlmann et al., 1998; Chen et al., 2011). In angiosperms, the core structure of labdane-related diterpenoids is typically formed from geranylgeranyl diphosphate (GGPP) via the sequential action of two diterpene synthases (diTPS; Peters, 2010). The first enzyme, a class II diTPS (TPS-c subfamily), carries out a protonation-initiated cyclization of GGPP to generate the characteristic decalin core. The conformation of GGPP upon cyclization—as determined by the active site of the specific class II diTPS—determines the stereochemistry of the decalin core and this generally remains unchanged in later biosynthetic steps (Andersen-Ranberg et al., 2016). It is possible for water to capture the cyclized diphosphate carbocation resulting in the introduction of a
hydroxyl group (Caniard et al., 2012; Zerbe et al., 2014). The second enzyme, a class I diTPS (usually from the TPS-e/f subfamily), then ionizes the cyclic intermediate through removal of the diphosphate group and often guides further carbocation driven rearrangements. Again, water capture of the carbocation intermediate can occur resulting in the introduction of a (second) hydroxyl group (Caniard et al., 2012). From this scheme, the initial steps on the biosynthetic route to dopaminergic diterpenoids in V. agnus-castus can be predicted to proceed via an initial class II diTPS catalysed cyclization of GGPP to generate bicyclic intermediates, potentially oxygenated at C-9 followed by rearrangement and dephosphorylation by a class I diTPS to yield either bi- or tricyclic labdane-related diterpenoids. Subsequent oxidative steps in the pathway towards the highly functionalised diterpenoids of V. agnus-castus are then likely to be carried out by members of the large cytochrome P450 (CYP) family (Ignea et al., 2016; Scheler et al., 2016; Pateraki et al., 2017).

To identify enzymes involved in the biosynthesis of V. agnus-castus diterpenoids we first localised diterpenoids to glandular trichomes on the surface of fruit and leaves with the aid of matrix assisted laser desorption ionization-mass spectrometry imaging (MALDI-MSI). A leaf-trichome transcriptome database was then generated and mined for candidate diTPSs and CYPs. Functional characterisation of the identified diTPSs resulted in the discovery of two diTPS combinations capable of producing a series of C-9 hydroxylated diterpenes. A CYP with high relative expression levels in trichomes was also characterised and found to hydroxylate the diol-diterpenoid, peregrinol, to yield labd-13Z-ene-9,15,16-triol, a potential intermediate in the biosynthetic pathway towards bioactive diterpenoids with lactone and furan groups.

RESULTS AND DISCUSSION

Tissue specific localization of diterpenoids in V. agnus-castus

Terpenoids often accumulate in plant tissues at or near their site of biosynthesis. This tendency has been exploited in biosynthetic pathway discovery efforts as a way to indirectly find tissues enriched

This article is protected by copyright. All rights reserved.
in transcripts of relevant genes that can then be targeted for transcriptomics (Trikka et al., 2015; Luo et al., 2016; Pateraki et al., 2017). Consequently, as a first step towards diterpenoid pathway discovery in V. agnus-castus, we focused our investigation on the metabolite profile and localisation of diterpenoids in fruit—the part of the plant traditionally used for treating PMS (Figure 1b). To determine if diterpenoids are localised to specific tissues within fruit, MALDI-MSI was used to map the spatial distribution of metabolites across longitudinal sections (Figure 1c). This technique has been used to localise diverse specialised metabolites in plant tissues (for a comprehensive review see: Boughton et al., 2015) including terpenoids in the medicinal plants Tripterygium wilfordii (Lange et al., 2017) and Salvia divinorum (Chen et al., 2017).

Analysis of fruit sections revealed the presence of many compounds at the outer edge of the fruit with predicted formulae consistent with diterpenoids previously reported from this species (Figure 1c, Table 1). Their observed distribution was found to overlap with that of glandular trichomes (Figure 1d), which were subsequently isolated, extracted and analysed by LC-HRMS to test for the presence of diterpenoids. Co-elution with a standard confirmed one of the major peaks to be vitexilactone (1), a labdane type diterpenoid with a five-membered lactone ring, and of interest for its potential bioactivity (Figure 1a; Meier et al., 2000). A lack of other available reference compounds and databases precluded the assignment of other peaks to specific diterpenoids. However, in the LC-HRMS analysis of trichome extracts, Na$^+$ adducts were observed that corresponded to the [M + K]$^+$ ions of the putative diterpenoids found to co-localise in the MALDI-MSI analysis (Table 1). In many cases, ions with the same mass were detected at multiple retention times. This can be explained by the presence of more than one compound with the same formula, which, for Vitex spp., could be expected given the multiple structural and stereoisomers reported in the literature for diterpenoids (Yao et al., 2016).
LC-HRMS analysis of leaf extracts demonstrated that leaves have a very similar diterpenoid profile to fruit and, additionally, contain a comparable level of vitexilactone on a dry weight (DW) basis (with an average value (± SD) of 4.4 ± 1.0 mg g⁻¹ DW in leaves compared to 2.9 ± 1.5 mg g⁻¹ DW in fruit; Figure 1e, Table 1). Isolation and analysis of glandular trichomes from leaves showed that, as for fruit, diterpenoids are also present in these structures (Table 1). The predicted formulae of ions detected in the trichome extracts suggest diterpenoids ranging in structural complexity from simple to more complex oxidised and acetylated compounds. Some of these simple diterpenoids may represent pathway intermediates and their presence in trichomes along with more functionalised forms supports the idea that trichomes are a site of both biosynthesis and storage. The localisation of diterpenoids to glandular trichomes is consistent with findings from other Lamiaceae species, including Rosmarinus officinalis, Marrubium vulgare and S. divinorum. In trichomes of these species, both diterpenoid accumulation and high expression levels of relevant pathway enzymes have been reported (Brückner et al., 2014; Zerbe et al., 2014; Scheler et al., 2016; Chen et al., 2017; Pelot et al., 2017).

**Generation of a V. agnus-castus leaf-trichome transcriptome and identification of candidate diterpene synthases**

Taking into consideration metabolite accumulation patterns, as well as fresh tissue availability, RNA samples from isolated leaf trichomes and whole leaves were sequenced using Illumina HiSeq technology. Reads from the two samples were de novo assembled using the Trinity pipeline (Haas et al., 2013). Homology searches of the transcriptome databases using characterised plant diTPSs as queries resulted in the identification of six full-length transcripts. Phylogenetic analysis of the translated coding sequences revealed that VacTPS1, VacTPS3 and VacTPS5 are members of the angiosperm class II diTPS clade assigned as TPS-c (Figure 2a; Chen et al., 2011). VacTPS1 and VacTPS3 clustered with other Lamiaceae diTPSs from specialised metabolism: VacTPS1 was most closely related to a peregrinol diphosphate synthase from Marrubium vulgare (MvCPS1; Zerbe et al.,
2014) to which it shares 61% sequence identity at the amino acid level, while VacTPS3 was placed sister to this clade. In contrast, VacTPSS clustered in a clade of diTPSs with mixed activities including ent-copalyl diphosphate synthases (ent-CPS) involved in the biosynthesis of gibberellins but also several diTPSs involved in specialised metabolism. The most closely related diTPSs were an ent-CPS speculated to be involved in specialised metabolism from *Isodon eriocalyx* (IeCPS2; Li et al. 2012) and a (-)-kolavenyl diphosphate synthase from *S. divinorum* (SdKPS; Chen et al. 2017; Pelot et al. 2017), to both of which VacTPSS shares 73% sequence identity at the amino acid level. The remaining three diTPSs identified from the leaf/trichome transcriptome database clustered with other class I diTPSs from the TPS-e/f sub-family (Figure 2b). VacTPS2 and VacTPS6 clustered in a Lamiaceae specific group involved in specialised metabolism. Like other members of this group, both enzymes are missing the N-terminal γ-domain, with only the βα bi-domain present (Hillwig et al., 2011; Pateraki et al., 2014; Zerbe et al., 2014). VacTPS4 was found to be most closely related to ent-kaurene synthases involved in gibberellin biosynthesis, sharing 78% sequence identity at the amino acid level with the closest included relative, MvEKS, from *M. vulgare* (Zerbe et al., 2014).

To identify terpene synthase candidates likely to be involved in trichome specific diterpenoid metabolism, transcript levels in leaves, fruit and trichomes were quantified using qRT-PCR. All candidates, with the exception of VacTPS4, had higher transcript levels in trichomes isolated from leaves and fruit relative to whole leaves and fruit, while VacTPS4 expression was similar across all tissues tested (Figure 2c).

**Functional characterisation VacTPSs**

Characterisation of the enzymatic activity of VacTPSs was carried out using the *Nicotiana benthamiana*-Agrobacterium-mediated transient expression system (Bach et al., 2014). Hexane extracts of *N. benthamiana* leaves expressing VacTPSs were analysed by GC-MS. Product identity
was determined by NMR and/or by comparison of product profiles with those of parallel assays of previously characterised enzymes (details are given in Table S2 in Supporting Information).

**Class II diterpene synthases: VacTPS1, VacTPS3 and VacTPS5**

Expression of VacTPS1 alone in *N. benthamiana* resulted in a single major peak (2; Figure 3) with a mass spectrum and retention time matching that of peregrinol (also confirmed by NMR analysis; Table S3). When coupled with MvELS the product profile was dominated by 9,13-epoxy-labd-14-ene (3b) as per the native *M. vulgare* combination of MvCPS1 and MvELS (Zerbe et al., 2014). Combined, these data indicate that VacTPS1 is a peregrinol diphosphate synthase, and is the second example of a class II diTPS that gives rise to a C-9 hydroxylated product. Expression of VacTPS3 in *N. benthamiana* resulted in a product profile matching that of a syn-CPS from rice, OsCPSSyn (Figure 4; Xu et al., 2004). Coupling of VacTPS3 with a sclareol synthase from *S. sclarea* (SsSS; Caniard et al. 2012) gave the identical product profile of OsCPSSyn and SsSS—a single major peak corresponding to syn-manool/vitexifolin A (5; Andersen-Ranberg et al., 2016; Jia et al., 2016) based on retention time and mass spectrum (Figure 4). Expression of VacTPS5 in *N. benthamiana* gave a product profile matching that of TwTPS14, a kolavenyl diphosphate synthase characterised from *Tripterygium wilfordii* (Figure 5; Andersen-Ranberg et al., 2016). Correspondingly, the coupling of VacTPS5 with SsSS resulted in the same product profile as the characterised coupling of TwTPS14 and SsSS, namely the production of the clerodane type diterpene, kolavelool (7; Andersen-Ranberg et al., 2016). To provide further evidence for the identity of the VacTPS3 and VacTPS5 products, the alcohol derivatives of the respective diphosphate products were isolated. Their structures were determined to be syn-copalol (4; Table S4) and kolavenol (6; Table S5) based on comparison of data from \(^1\)H NMR and HSQC experiments with data from literature (Yee and Coates, 1992; Nakano et al., 2015), confirming that VacTPS3 is a syn-CPP synthase and VacTPS5 is a kolavenyl diphosphate synthase. To address the absolute configuration of the class II diTPS products, optical rotation measurements were attempted. Low quantities of isolated material did not allow for accurate calculation of the
specific optical rotation, however, it should be noted that both 4 and 6 showed weak dextrorotatory (+) properties. Because the absolute configurations of the class II diphosphate derivatives, 2, 4 and 6, were not established, the shown structures could be either of the two possible enantiomers for each compound, with the downstream class I diTPS products expected to share the same configuration.

Class I diterpene synthases: VacTPS2, VacTPS6 and VacTPS4

To characterise the activity of the three identified class I enzymes, all *V. agnus-castus* class II/class I enzyme combinations were tested. Additionally, to examine the range of accepted substrates two characterised class II diTPSs were included in the combinatorial assays: EpTPS7, an *ent*-CPP synthase from *Euphorbia peplus* (Zerbe et al., 2013), and CfTPS2, a (+)-8-LPP synthase from *C. forskholii* (Pateraki et al., 2014). Co-expression of VacTPS2 with VacTPS1 in *N. benthamiana* resulted in a shift in the VacTPS1 product profile from 2 to two main products, 3a and 8 (Figure 3). Structural analysis by NMR identified 3a as 9,13(\(\text{R}\))-epoxy-labd-14-ene (Table S6; presumably the C-13 epimer of 3b) and 8 as viteagnusin D (Table S7). Derivatives of 3 are known from the Vitex genus (e.g. Ono et al., 2008; Ono et al., 2009) and 8 has been isolated from fruits of *V. agnus-castus* with a series of other structurally related diterpenoids (Ono et al., 2008). VacTPS2 was also found to couple with VacTPS3 to produce 5—a diterpenoid previously isolated from *V. rotundifolia* (Figure 4; Ono et al., 2002)—as per the characterised enzyme combination of OsCPPsyn and SsSS (Andersen-Ranberg et al., 2016). A weak coupling between VacTSP2 and VacTPS5 to generate 7 was suggested by the presence of the characteristic m/z 257 and 189 ions at the correct retention time for this compound (when compared to the product profile of TwTPS14 and SsSS) but their intensity levels were too close to the background noise level for a definitive result. VacTPS2 was able to accept *ent*-CPP to produce *ent*-manool (10) and accept (+)-8-LPP to produce 13R-(+)-manoyl oxide (12a; Figure S1 and S2). A new product was observed when VacTPS6 was coupled with VacTPS1 which, after isolation and structural elucidation by NMR, was identified as labd-13(16),14-dien-9-ol (13; Figure 3 and Table S8), a recently reported diterpenoid previously only generated by an artificial combination of a bacterial
enzyme, KgTS and MvCPS1 (Jia et al., 2016). Co-expression of VacTPS6 with VacTPS3 resulted in two main components identified as dehydroabietadiene (15) and syn-isopimara-7,15-diene (14) by comparison to the products of the characterised enzyme combination, OsCPPsyn and CfTPS3 (Andersen-Ranberg et al., 2016). VacTPS6 did not couple with VacTPS5 (Figure 5) but could accept (+)-8-LPP to produce 12a (Figure S2). VacTPS4 was not found to couple with any of the V. agnus-castus class II enzymes or CfTPS2 but when co-expressed with EpTPS7 a peak corresponding to ent-kaurene (16) was observed (Figure S3). This activity would fit with a role of VacTPS4 in gibberellin biosynthesis.

A complex array of diterpenoids, with different labdane-related backbones is present in V. agnus-castus (Yao et al., 2016). The molecular basis of much of this diversity can be explained by the combined activities of the class II and class I diTPSs reported here (summarised in Figure 6). The diTPSs, VacTPS1, VacTPS2, and VacTPS6, are of particular interest as together they provide enzymatic access to C-9 hydroxylated backbones that are a feature of the bioactive preparations of this species (Jarry et al., 2006). Furthermore, VacTPS2 and VacTPS6 display a high level of substrate promiscuity and, as such, are valuable additions to the combinatorial toolbox of class I diTPS capable of catalysing the formation of valuable diterpenoids (Andersen-Ranberg et al., 2016; Jia et al., 2016). Clerodane type diterpenoids are also found in several Vitex species (Ono et al., 2002; Lee et al., 2013) and there is evidence that these may also be responsible for the dopaminergic activity observed in V. agnus-castus extracts (Wuttke et al., 2003). Hence, the lack of a clear coupling between VacTPS5 and any of the tested class I diTPS suggests there are possibly other diTPSs not characterised in this study. A question mark also remains over the direct products of VacTPS6 when supplied with syn-copalyl diphosphate. Dehydroabietadiene (15) was one of two products detected but it is unlikely to be a direct product of VacTPS6 due to a higher than usual level of oxidation. Instead, 15 is likely to be the dehydrogenated derivative of an abietane-type diterpene that has an already existing degree of unsaturation in the C-ring. Zi and Peters (2013) have demonstrated that
miltiradiene easily undergoes spontaneous oxidation to 15 and, as suggested by the presence of 15 in *N. benthamiana* leaves and *S. cerevisiae* cultures expressing a levopimaradiene synthase from *Ginkgo biloba* (Matsuda and Schepmann, 2008; Brückner and Tissier, 2013), levopimaradiene can apparently also undergo a similarly facile aromatisation. In the present work no such abietane type product was detected, and so at least one of the direct products of *VacTPS6* (when coupled with *VacTPS3*) remains to be determined.

**Identification of a cytochrome P450 capable of peregrinol oxidation**

Formation of the furan and lactone rings characteristic of bioactive Vitex diterpenoids is likely to be, at least in part, catalysed by the activity of cytochrome P450s (CYPs). To identify possible candidate enzymes for these reactions, the transcriptome databases were searched for sequences with similarity to members of the CYP76 family—a family within the large CYP71 clan involved in the biosynthesis of specialised diterpenoids in other Lamiaceae species (Guo *et al.*, 2013; Zi and Peters, 2013; Pateraki *et al.*, 2017). A candidate gene (*VacCYP76BK1*) was identified with high relative expression levels in leaf and fruit trichomes (Figure 7a). Phylogenetic analysis classified the candidate as the first characterised member of a new sub-family, CYP76BK, sister to the CYP76AK clade, which comprises CYPs involved in the biosynthesis of tanshinones and carnosic acid (Figure 7b; Guo *et al.*, 2016; Ignea *et al.*, 2016; Scheler *et al.*, 2016).

The ability of the identified CYP to oxygenate the C-9 hydroxylated backbones produced by the characterised *VacTPSs* was tested using *in vitro* microsomal assays. The microsomal fraction of a *S. cerevisiae* strain expressing *VacCYP76BK1* (along with a cytochrome P450 reductase (CPR) from poplar; Ro *et al.*, 2002) was isolated and incubated with 2, 3a, 8 or 13. LC-HRMS analysis of assays incubated with 2 revealed the presence of a new component (17) with an [M + Na]$^+$ ion of *m/z* 347.2558, which suggested the molecular formula C$_{20}$H$_{36}$O$_3$ (calculated [M + Na]$^+$ *m/z* 347.2557, −0.04 ppm difference; Figure 7c). No new components were observed in assays incubated with 3a, 8
or 13 compared to control assays (Figure S4a). To confirm the activity of VacCYP76BK1 in vivo and to obtain sufficient amounts of 17 for structural elucidation, VacTPS1, SsSS (previously shown in Jia et al. 2016 to catalyse the formation of 8 from peregrinol diphosphate), VacCYP76BK1 and a CPR from C. forskohlii (Pateraki et al., 2017) were integrated into the genome of S. cerevisiae (Jensen et al., 2013). Analysis of culture extracts by LC-HRMS showed production of 2, 8 and 17 (along with additional minor components only observed when VacCYP76BK1 was present; Figure 7c). 17 was isolated from an ethyl acetate extract of an engineered S. cerevisiae culture using a combination of silica chromatography and RP-TLC. The structure of 17 was determined by NMR analysis to be labd-13Z-ene-9,15,16-triol (Table S10), a known diterpenoid isolated from another species belonging to the Lamiaceae family, Leucas stelligera (Kulkarni et al., 2013). 17 was also found in the present study in fruit and leaf trichome extracts along with 2, 8 and an unidentified diterpenoid produced by S. cerevisiae strains expressing the diTPSs (Figures S4b and c).

That 2 appears to be a substrate for VacCYP76BK1 is interesting considering it is the simple dephosphorylated form of peregrinol diphosphate, the product of the class II enzyme, VacTPS1. From our experimental data it could not be determined if either of the characterised class I diTPSs are involved in its formation. 2 was observed in extracts of N. benthamiana leaves transiently expressing VacTPS1 both with and without co-expression of class I diTPSs. The non-quantitative nature of the expression system made it difficult to determine whether either VacTPS2 or VacTPS6 possess simple phosphatase activity (alongside their other characterised activities) or whether the presence of 2 in these assays was entirely due to the activity of endogenous phosphatases in N. benthamiana, an activity previously observed in this expression system (Andersen-Ranberg et al., 2016).

Without direct evidence for the function of VacCYP76BK1 in V. agnus-castus we can only speculate that C-16 hydroxylation of 2 is a native function of the enzyme in planta. Nevertheless,
VacCYP76BK1 was found to be expressed in the same tissue where 2 was detected, indicating that this diterpenoid would likely be available as a substrate. The addition of a hydroxyl group at this specific position of 2 to yield 17 could be considered a step towards the formation of both the lactone and furan rings characteristic of diterpenoids present in V. agnus-castus (Figure 1a). Further oxidation at either C-15 or C-16 of 17 to form a carboxylic acid has the potential to lead to formation of a lactone by spontaneous cyclization (Figure 8). A similar pathway has been demonstrated for the biosynthesis of costunolide—a sesquiterpene lactone found in several Asteraceae species. Here, the C-13 methyl of the sesquiterpene backbone, germacrene A, is successively oxidised to a carboxylic acid via an aldehyde by a single CYP (de Kraker et al., 2001). Hydroxylation of the nearby C-6 by a second CYP then leads to spontaneous cyclization, resulting in lactone formation (de Kraker et al., 2002; Liu et al., 2011). In the case of artemisinic acid biosynthesis, an alcohol dehydrogenase and an aldehyde dehydrogenase are responsible for catalysing these steps from the sesquiterpene alcohol to the acid (Teoh et al., 2009; Paddon et al., 2013). A route to furan ring formation from 17 is also feasible via the aldehyde intermediate involved in lactone ring formation. In this scheme tautomerization of the aldehyde in a keto-enol equilibrium reaction followed by simple dehydration would afford the furan ring (Figure 8). Alternative routes to furan formation have been proposed and involve the formation and subsequent loss of a 9,13-epoxy bridge in conjunction with an additional oxidation of C-15 or C-16 (Henderson and McCrindle, 1969; Zerbe et al., 2014). It seems, however, the presence of the C-9 hydroxyl (and by extension the 9,13-epoxy bridge) is not essential for furan or lactone ring formation as multiple furan and lactone bearing diterpenoids without C-9 hydroxylation exist (Smith et al., 1982; Omosa et al., 2014; Anwar et al., 2017). Nevertheless, a potential intermediate of this alternative route was identified in the present study as a major product of VacTPS2 when coupled with VacTPS1 (3a), and similar compounds are often encountered in other furan/lactone diterpenoid producing species (Ono et al., 1999; Tesso and König, 2004; Kulkarni et al., 2013). The sheer number of structurally related diterpenoids found in V. agnus-castus, points to the existence of a complex metabolic grid comprised of (multifunctional) enzymes,
which combine to generate the observed chemical diversity. In such a system, multiple routes to furan and lactone diterpenoids may exist.

In the present work, we have described the functional characterisation of a family of diTPSs in *V. agnus-castus*. Combined, the diTPSs could generate a range of diterpene backbones that form the biosynthetic basis for the multitude of highly functionalised labdane and clerodane-type diterpenoids observed in this species. Several of these functionalised diterpenoids possess unique therapeutic properties and the described CYP offers a potential enzymatic step towards their biosynthesis. Our findings accordingly show how *V. agnus-castus* offers a valuable source of enzymes which can be implemented in the engineering of biosynthetic pathways for the production of high-value diterpenoids.

**EXPERIMENTAL PROCEDURES**

**Plant material**

*Vitex agnus-castus* seeds were obtained from a plant growing in a private garden in Chania (Crete, Greece) and were germinated and grown in the greenhouse at the University of Copenhagen, Denmark (minimum day length/temperature: 11 h/12 °C). Samples of leaves and semi-mature fruit where harvested from three individual plants. Tissues were ground in liquid nitrogen and stored at −80 °C until use or prepared for MALDI-MS Imaging (see below).

**MALDI-MS Imaging and sample preparation**

*Vitex agnus-castus* fruit were embedded in denatured albumin (boiled egg-white) and frozen over a bath of liquid nitrogen. Albumin embedded fruit were cryosectioned (Leica CM3050 S cryostat, Leica Microsystems, Wetzler, Germany) and sections gently pressed onto pre-chilled double-sided carbon tape attached to glass slides. Slides were freeze-dried overnight and stored in a vacuum desiccator. 2,5-Dihydroxybenzoic acid (DHB) was used as matrix and sublimed onto tissue sections using a
custom-built sublimation apparatus at temperatures of 130-140 °C at vacuum pressures of less than 0.1 mbar for a period of 4-6 min. A Bruker Solarix 7 XR Tesla Hybrid ESI/MALDI-FT-ICR-MS (Bruker Daltonik, Bremen, Germany) was used to acquire MS images. For data acquisition and analysis details, see Methods S1.

**Metabolite analysis**

For metabolite analyses, ground *V. agnus-castus* tissues were extracted with n-hexane or acetonitrile with shaking incubation at 25 °C for 1 hour. Hexane samples were analysed by gas chromatography-mass spectrometry (GC-MS) using a Shimadzu GCMS-QP2010 Ultra (Shimadzu, Kyoto, Japan) with an HP-5MS UI column (20 m x 0.18 mm i.d., 0.25 μm film thickness; Agilent, Santa Clara, CA) with H₂ as the carrier gas. The injection port was operated in splitless mode at 250 °C and the GC program was as follows: 60 °C for 1 min, ramp to 150 °C at 30 °C min⁻¹, ramp to 250 °C at 15 °C min⁻¹, ramp to 290 °C at 30 °C min⁻¹, hold for 3 min. The MS used electron impact (EI) ionization with the ion source voltage and temperature set to 70 eV and 300 °C, respectively. Programmed Temperature Vaporization method (PTV) was used to analyse VacTPS1 products due to known degradation issues with peregrinol-type diterpenoids (Zerbe et al., 2014). Here initial injection port temperature was held at 40 °C for 1 min then ramped to 250 °C over 4 min. LC-HRMS analyses of acetonitrile extracts were performed using the coupled Ultimate 3000 UHPLC+ Focused system (Dionex Corporation, Sunnyvale, CA) and Bruker Compact ESI-QTOF-MS (Bruker) system described in Pateraki et al. (2017) with the following gradient method: 0-1 min, 20% B; 1-23 min, 20-100% B; 23-25 min, 100%; 25-25.5 min, 100-20%; 25.5-30.5 min, 20% B. Mass spectra were acquired in positive ion mode with the following ESI settings: capillary voltage, 4000 V; end plate offset, –500 V; dry gas temperature, 220 °C; dry gas flow of 8 L min⁻¹; nebulizer pressure, 2 bar. Data were analysed using DataAnalysis 4.1 (Bruker).

**Quantification of vitexilactone**

This article is protected by copyright. All rights reserved.
Quantification of vitexilactone was carried out on diluted acetonitrile extracts by LC-MS/MS using an external standard curve of vitexilactone. For details see, Methods S1.

**Transcriptome sequencing and assembly**

*Vitex agnus-castus* leaf total RNA was extracted from expanding leaves using the Spectrum Plant Total RNA Kit (Sigma-Aldrich, St Louis, MO). RNA was also extracted from a trichome enriched leaf tissue sample achieved by shaking leaves in a dry ice filled 50 mL tube. Trichomes adhered to the tube walls along with some broken leaf material were washed off with lysis buffer and extracted as above. Transcriptome sequencing was carried out by Macrogen (Seoul, South Korea) on an Illumina HiSeq2000 sequencer (Illumina, San Diego, CA) to generate paired end libraries (2 × 100 bp). The reads were *de novo* assembled and relative transcript abundance estimated using the Trinity pipeline (Haas *et al.*, 2013).

**Phylogenetic analysis**

Amino acid sequences from *V. agnus-castus* were aligned in MEGA 6.06 (Hall, 2013) with functionally characterised diterpene synthase or cytochrome P450 amino acid sequences from other plants species retrieved from GenBank (http://www.ncbi.nlm.nih.gov/; accession numbers are given in Tables S1 and S9) using MUSCLE with default settings (Edgar, 2004). The alignments were manually trimmed to remove terminal gaps and in the case of diterpene synthase alignments, N-terminal plastid targeting peptides were also removed. Phylogenetic trees were constructed using the Maximum Likelihood method based on the JTT matrix-based model with a site coverage cut-off of 90%. To model evolutionary rate differences among sites a discrete Gamma distribution was used with 5 categories. Trees are drawn to scale, with branch lengths representing the number of substitutions per site. Trees were tested by bootstrapping with 1000 repetitions. The class I and class II dITPS trees were rooted with PpCPS/EKS from *Physcomitrella patens* and the CYP tree was rooted with CYP71D18 from *Mentha spicata*.
RNA extraction and quantitative Real-Time PCR

RNA was extracted from whole leaves, fruit and trichomes (isolated by brushing the surface of frozen tissue samples) using the Ambion RNAqueous-Micro Total RNA Isolation Kit with Plant RNA isolation Aid (Thermo Fisher Scientific, Waltham, MA) and DNase treated using the Turbo DNA-free Kit (Thermo Fisher Scientific). Quality control of RNA samples was performed with a 2100 Bioanalyzer Instrument (Agilent). cDNA libraries were generated using an iScript cDNA Synthesis Kit (Bio-Rad Laboratories, Hercules, CA). qRT-PCR reactions were performed with gene specific primers (Table S11) and the SensiFAST SYBR® No-ROX Kit master mix (Bioline Reagents, UK) on a CFX384 Touch Real-Time PCR Detection System (Bio-Rad) using the following two step cycling parameters: 95 °C for 2 min, followed by 45 cycles of 95 °C for 5 s and then 60 °C for 30 s. Melt curve analysis was performed by increasing the temperature from 65 °C to 95 °C in 0.5 °C increments. RNA polymerase II subunit 3 was used as the reference gene. Gene expression levels were calculated using the efficiency corrected delta Ct (EΔCt) method with the specific efficiency (E) of each primer pair calculated using a standard curve based on four log dilutions of cDNA template. All reactions were carried out in three technical replicates and values reported for leaves and leaf trichomes are the mean of three biological replicates. For fruit trichomes tissue was pooled from three different plants to obtain sufficient trichome RNA for expression studies. A Student’s t-test was used to test for statistically significant differences in gene expression levels between leaves and leaf trichomes using SigmaPlot 13.0 (Systat Software Inc., San Jose, CA).

Cloning and transient expression of candidate genes from *Vitex agnus-castus* in *Nicotiana benthamiana*

Candidates were amplified with gene specific primers (Table S11) from cDNA generated from leaf trichome total RNA using SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen, Carlsbad, CA). The cloning procedure and transient expression of TPSs in *N. benthamiana* was performed as described in Andersen-Ranberg *et al.* (2016). Five to seven days after infiltration, leaf
discs were extracted and analysed by GC-MS as per “Metabolite analysis”. Characterised diTPSs used in the reference assays were provided by Johan Andersen-Ranberg (University of Copenhagen) in pCAMBIA130035Su.

Microsome preparation and in vitro assays

The S. cerevisiae strain AM94 (Ignea et al., 2012) was transformed with the expression vectors pYX143 carrying the poplar CYP reductase, CPR2 (Ro et al., 2002), under the control of the ADH1 promoter and either PYeDP60 carrying VacCYP76BK1 under the control of the GAL10-CYC1 promoter or without the insert (empty vector) as control. Transformants were grown for 24 h at 30 °C and 110 rpm in 50 mL of synthetic complete (SC) selection media (minus leucine and uracil) with 2% glucose. The cells were pelleted and resuspended in 200 mL SC selection media with 2% galactose and grown for a further 24 h. Microsomes were prepared as described in Pompon et al. (1996) and in vitro assays performed as described in Hamann and Møller (2007).

Saccharomyces cerevisiae strain construction and testing

The S. cerevisiae strains were constructed and grown for metabolite analysis as described in Andersen-Ranberg et al. (2016). Codon optimised VacTPS1 and SsSS (GeneArt, Thermo Fisher Scientific) were used to generate 2 and also viteagnusin D (8) (Jia et al., 2016) and were cloned into S. cerevisiae genome integration plasmids for integration into the XI-5 locus along with codon optimised VacCYP76BK1 and CfCPR (Pateraki et al., 2017). VacTPS1 and SsSS (Caniard et al., 2012) were under the control of the TEF1 and PGK1 promoters, respectively, VacCYP76BK1, was under the control of the TDH3 promoter, and CfCPR was under the control of the TEF2 promoter. For metabolite analysis of transformants, culture aliquots of 200 μL were extracted with an equal volume of ethyl acetate (EtOAc), vortexed and centrifuged. The EtOAc layer was dried under N₂, re-suspended in 150 μL acetonitrile and analysed by LC-HRMS as per “Metabolite analysis”.
Compound isolation and NMR experiments

For details, see Methods S1.

Accession numbers: Nucleotide sequences reported in this study were submitted to the National Center for Biotechnology Information (NCBI) GenBankTM/EBI Data Bank with accession numbers: VacTSP1, XXXXX; VacTSP2, XXXXX; VacTSP3, XXXXX; VacTSP4, XXXXX; VacTSP5, XXXXX; VacTSP6, XXXXX; VacCYP76BK1, XXXXX.

ACKNOWLEDGEMENTS

The authors thank Dan Luo and Qing Liu for assistance with microsomal assays, Codruta Ignea for AM94 S. cerevisiae strain, and the greenhouse personnel, specifically Theodor Emil Bolsterli, at the University of Copenhagen for growing and caring for our plants. We gratefully acknowledge Dr. David Nelson (University of Tennessee) for CYP naming. This work was supported by the Center for Synthetic Biology (University of Copenhagen Excellence Program for Interdisciplinary Research), by a European Research Council Advanced Grant to BLM (ERC-2012-ADG_20120314) and by the Danish Innovation Foundation funded project “Plant Power: Light-driven synthesis of complex terpenoids using cytochrome P450s” (12-131834; project lead, Dr. Poul Erik Jensen, University of Copenhagen). AMH was support by a Marie Skłodowska Curie Individual Fellowship. TCMS was supported by a SLAI grant from the Ministry of Higher Education Malaysia. MALDI-MSI was conducted at Metabolomics Australia (School of BioSciences, The University of Melbourne, Australia), a NCRIS initiative under Bioplatforms Australia Pty Ltd. The authors declare no competing financial interests.
SHORT SUPPORTING INFORMATION LEGENDS

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

**Figure S1.** *In planta* assays of class I enzymes VacTPS2 and VacTPS6 with a class II ent-copalyl diphosphate synthase from *Euphorbia peplus* (EpTPS7).

**Figure S2.** *In planta* assays of class I enzymes VacTPS2 and VacTPS6 with a class II labd-13-en-8-ol diphosphate synthase from *Coleus forskohlii* (CfTPS2).

**Figure S3.** *In planta* assays of class I enzyme VacTPS4.

**Figure S4.** (a) LC-HRMS total ion chromatograms of microsomal assay extracts. (b) and (c) LC-HRMS extracted ion chromatograms (EIC) comparing the product profile of the VacCYP76K1 expressing *S. cerevisiae* strain with *V. agnus-castus* leaf and fruit trichome extracts.

**Table S1.** Genes used in the diterpene synthase phylogenetic analysis.

**Table S2.** Compounds reported in this paper and the method of identification.

**Table S3.** Peregrinol (2) NMR data.

**Table S4.** syn-Copalol (4) NMR data.

**Table S5.** Kolavenol NMR data.

**Table S6.** 9,13-Epoxylabd-14-ene (3a) NMR data.

**Table S7.** Viteagnusin D (8) NMR data.

**Table S8.** Labda-13(16),14-dien-9-ol (13) NMR data.

**Table S9.** Genes used in the cytochrome P450 phylogenetic analysis.

**Table S10.** Labd-13Z-ene-9,15,16-triol (17) NMR data.

**Table S11.** Primers used for isolation and cloning of *V. agnus-castus* genes, CfCPR and SsSS, and qRT-PCR experiments.

**Methods S1.**
REFERENCES


dehydrogenase(s) involved in sesquiterpene lactone biosynthesis. *Plant Physiology*, **125**, 1930-1940.


This article is protected by copyright. All rights reserved.


This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/tpj.13822

This article is protected by copyright. All rights reserved.
TABLES

Table 1. Ions detected by MALDI-MSI found to colocalise with vitexilactone ([M + K]\(^+\) 417.2044, calculated 417.2038, 1.0 ppm error) in a fruit section of *V. agnus-castus* with associated calculated molecular formulae, and corresponding detection in LC-HRMS analyses of fruit and leaf trichome extracts as either [M + Na]\(^+\) or [M + H]\(^+\) ions. Colocalisation analysis was carried out using SCiLS lab software (p ≤ 0.05, threshold value of 0.5). Mass errors between observed LC-HRMS m/z values and calculated molecular formulae are < 3 ppm. Specific ions detected in only fruit or only leaf trichomes have retention times denoted by F or L, respectively, all other ions were detected in both tissues. Annotations are based on comprehensive searches of the SciFinder Scholar online chemical database for compounds with matching formulae reported from *Vitex* spp., with the exception of the component (m/z 401.2301) eluting at 12.6 min, which was identified as vitexilactone (1) by comparison to an authentic standard. Ions with the same mass were detected at different retention times, and may represent structural and stereoisomers and/or are the result of in-source fragmentation of different but related parent compounds.

<table>
<thead>
<tr>
<th>Observed mass MALDI-MSI (m/z)</th>
<th>Calculated molecular formula</th>
<th>Error (ppm)</th>
<th>Adduct</th>
<th>Calculated mass for [M + Na](^+) or [M + H](^+)</th>
<th>Observed mass LC-HRMS (m/z)</th>
<th>Retention time (min)</th>
<th>Compounds reported from Vitex spp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>345.2204</td>
<td>C(<em>{20})H(</em>{34})O(_{2})</td>
<td>3.8</td>
<td>K(^+)</td>
<td>329.2451</td>
<td>329.2455</td>
<td>16, 16.4</td>
<td>sclareol (Eryigit et al., 2015), viteagnusin C (Ono et al., 2008), viteagnusin D (Ono et al., 2008), peregrinol (this report)</td>
</tr>
<tr>
<td>347.2357</td>
<td>C(<em>{20})H(</em>{36})O(_{2})</td>
<td>2.9</td>
<td>K(^+)</td>
<td>331.2607</td>
<td>331.2612</td>
<td>16.7, 17.4, 17.8</td>
<td>viteagnusin C (Ono et al., 2008), viteagnusin D (Ono et al., 2008), peregrinol (this report)</td>
</tr>
<tr>
<td>359.1993</td>
<td>C(<em>{20})H(</em>{32})O(_{3})</td>
<td>2.7</td>
<td>K(^+)</td>
<td>343.2245</td>
<td>343.2249</td>
<td>16.15, 25.2</td>
<td>vitexolin A (Corlay et al., 2015)</td>
</tr>
<tr>
<td>373.1789</td>
<td>C(<em>{20})H(</em>{34})O(_{4})</td>
<td>-3.7</td>
<td>K(^+)</td>
<td>335.2216 [M + H](^+)</td>
<td>335.2216</td>
<td>common fragment ion</td>
<td>vitexolide A (Corlay et al., 2015), vitexolide B (Corlay et al., 2015), 12-epivitexolide A (Corlay et al. 2015), 12S,16S/R-dihydroxy-ent-labda-7,13-dien-15,16-oxide (Nyiligira et al., 2008)</td>
</tr>
<tr>
<td>M / z</td>
<td>Formula</td>
<td>Charge</td>
<td>Retention Time (min)</td>
<td>Molecular Mass</td>
<td>Mass Accuracy</td>
<td>Assignment</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>------------------</td>
<td>--------</td>
<td>----------------------</td>
<td>----------------</td>
<td>---------------</td>
<td>---------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>375.1943</td>
<td>C_{20}H_{32}O_{4}</td>
<td>2.9</td>
<td>359.2193</td>
<td>359.2189</td>
<td>20.4</td>
<td>vitexifolin F (Zheng et al., 2013), vitexifolin E (Ono et al., 2002)</td>
<td></td>
</tr>
<tr>
<td>383.0543</td>
<td>C_{23}H_{16}O_{6}</td>
<td>1.81</td>
<td>405.0370</td>
<td>not observed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>387.2308</td>
<td>C_{22}H_{38}O_{3}</td>
<td>3.0</td>
<td>371.2557</td>
<td>371.2570</td>
<td>17.2</td>
<td>viteagnusin A (Ono et al., 2008), viteagnusin B (Ono et al., 2008)</td>
<td></td>
</tr>
<tr>
<td>399.1946</td>
<td>C_{22}H_{32}O_{4}</td>
<td>3.5</td>
<td>383.2193</td>
<td>383.2192</td>
<td>16.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>401.2101</td>
<td>C_{22}H_{34}O_{4}</td>
<td>3.0</td>
<td>385.2349</td>
<td>385.2352</td>
<td>15.8 (F), 16.4, 17.1, 17.9, 18.5</td>
<td>vitetrifolin B (Ono et al., 2000), rotundifuran (Ono et al. 2000)</td>
<td></td>
</tr>
<tr>
<td>401.2306</td>
<td>C_{22}H_{34}O_{5}</td>
<td>1.8</td>
<td>Na^{+}</td>
<td>401.2298</td>
<td>see ion 417</td>
<td></td>
<td></td>
</tr>
<tr>
<td>403.2257</td>
<td>C_{22}H_{36}O_{2}</td>
<td>2.9</td>
<td>387.2506</td>
<td>387.2514</td>
<td>17.5, 18.2</td>
<td>vitetrifolin E (Ono et al., 2001), vitetrifolin F (Ono et al. 2001)</td>
<td></td>
</tr>
<tr>
<td>403.0658</td>
<td>C_{19}H_{14}O_{10}</td>
<td>0.3</td>
<td>H^{+}</td>
<td>403.0659 [M + H]^{+}</td>
<td>not observed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>417.2044</td>
<td>C_{22}H_{34}O_{5}</td>
<td>1.5</td>
<td>K^{+}</td>
<td>401.2298</td>
<td>401.2301</td>
<td>12.6</td>
<td>vitexilactone (confirmed by coelution with a standard; Hoberg et al., 1999)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>vitesin A (Alam et al., 2002), previtexilactone (Kiuchi et al., 2004),</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>vitexilactone B (Zheng et al., 2010), negundoin A (Zheng et al. 2010),</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>viterotulin B (Lee et al., 2013)</td>
</tr>
<tr>
<td>419.2053</td>
<td>C_{18}H_{36}O_{4}</td>
<td>2.7</td>
<td>K^{+}</td>
<td>403.2302</td>
<td>not observed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>419.2201</td>
<td>C_{22}H_{36}O_{5}</td>
<td>1.56</td>
<td>K^{+}</td>
<td>403.2455</td>
<td>403.2466</td>
<td>13.2, 13.8, 14, 15.1 (L), 16 (L), 16.8 (L)</td>
<td>negundoin E (Zheng et al. 2010)</td>
</tr>
<tr>
<td>451.2116</td>
<td>C_{27}H_{36}O_{6}</td>
<td>0.3</td>
<td>H^{+}</td>
<td>473.1935</td>
<td>not observed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>477.4104</td>
<td>C_{34}H_{52}O</td>
<td>2.8</td>
<td>H^{+}</td>
<td>499.3910</td>
<td>not observed</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This article is protected by copyright. All rights reserved.
**FIGURE LEGENDS**

**Figure 1.** Metabolite profiling of *Vitex agnus-castus* tissues. (a) Diterpenoids reported from *V. agnus-castus* fruit (1Hoberg et al. 1999, 2Ono et al. 2011, 3Ono et al. 2009). (b) Maturing fruit. (c) Longitudinal section through a semi-mature fruit prepared for MALDI-MSI (1000 μm scale bar) and corresponding ion maps of colocalising m/z values of putative diterpenoids (see Table 1, all images TIC normalised). (d) Bright field image of fruit epidermis with glandular trichomes (indicated with an arrowhead, 100 μm scale bar). (e) Quantification of vitexilactone (1) in leaf and fruit tissue (n = 3, error bars represent ± 1 SD).

**Figure 2.** Phylogenetic analysis and relative transcript abundance of *Vitex agnus-castus* diterpene synthases (diTPSs). Maximum likelihood trees of class II (a) and class I (b) diTPSs based on aligned protein sequences. Trees are drawn to scale, with branch lengths representing the number of substitutions per site. Branches with > 80% bootstrap support are indicated with a filled circle. Trees were rooted with PpCPS/EKS. Genbank accession numbers are listed in Table S1. (c) Relative transcript abundance of *V. agnus-castus* diTPSs in leaves, fruit and trichomes as determined by qRT-PCR. For leaf data, the mean of three biological replicates (± 1 SD) is shown; for fruit, tissue was combined from three different plants to obtain sufficient trichome RNA for expression studies. Two-tailed P-values < 0.05 are shown with an asterisk.

**Figure 3.** In planta functional characterization of class II enzyme, *VacTPS1*, and class I enzymes, *VacTPS2* and *VacTPS6*. (a) GC-MS extracted ion chromatograms (EIC) of hexane extracts of *N. benthamiana* leaves transiently expressing combinations of diTPSs. (b) Diterpene synthase products with representative mass spectra. Characterised diTPS combinations used as references and compound identification details are given in Table S2.
Figure 4. *In planta* functional characterization of class II enzyme, *VacTPS3*, and class I enzymes, *VacTPS2* and *VacTPS6*. (a) GC-MS extracted ion chromatograms (EIC) of hexane extracts of *N. benthamiana* leaves transiently expressing combinations of diTPSs. (b) Diterpene synthase products with representative mass spectra. Characterised diTPS combinations used as references and compound identification details are given in Table S2.

Figure 5. *In planta* functional characterization of class II enzyme, *VacTPS5*, and class I enzymes, *VacTPS2* and *VacTPS6*. (a) GC-MS extracted ion chromatograms (EIC) of hexane extracts of *N. benthamiana* leaves transiently expressing combinations of diTPSs. (b) Diterpene synthase products with representative mass spectra. Characterised diTPS combinations used as references and compound identification details are given in Table S2.

Figure 6. Summary of identified *Vitex agnus-castus* diterpene synthase (diTPS) activities. (a) Observed enzymatic activities of *V. agnus-castus* diTPSs. Dashed arrow indicates 15 may not be the direct product of *VacTPS6*. (b) Examples of functionalised diterpenoids from *Vitex* spp. representing the different diterpene backbones generated by the class II and class I diTPS combinations.

Figure 7. Phylogenetic analysis, relative transcript abundance in *Vitex agnus-castus* tissues and activity of *VacCYP76BK1*. (a) Relative transcript abundance of *VacCYP76BK1* in leaves, fruit and trichomes as determined by qRT-PCR. For leaf data, the mean of three biological replicates (± 1 SD) is shown; for fruit, tissue was combined from three different plants to obtain sufficient trichome RNA for expression studies. Two-tailed *P*-values < 0.05 are shown with an asterisk. (b) Maximum Likelihood tree of cytochrome P450s (CYPs) based on aligned protein sequences. The tree is drawn to scale, with branch lengths representing substitutions per site. Branches with > 80% bootstrap support are indicated with a filled circle. The tree was rooted with *MsCYP71D18*. Genbank accession numbers are listed in Table S9. (c) LC-HRMS total ion chromatograms of microsomal assay extracts.
prepared from a *S. cerevisiae* strain expressing VacCYP76BK1 supplied with peregrinol (2), and culture extracts of *S. cerevisiae* strains expressing either diTPSs alone or in combination with VacCYP76BK1.

**Figure 8.** Proposed pathway for the formation of lactone and furan derivatives of peregrinol (2) via labd-13Z-ene-9,15,16-triol (17). Lactone formation could proceed from 17 via a dehydrogenase catalysed oxidation giving rise to an aldehyde (a) and to a carboxylic acid (b), followed by an enzyme catalysed or spontaneous dehydration reaction to form the lactone ring. Alternatively, dehydration of the tautomeric enol form (c) of the aldehyde intermediate (a) would afford furan ring formation. Dashed arrows indicate putative steps.
Figure 1. Metabolite profiling of *Vitex agnus-castus* tissues. (a) Diterpenoids reported from *V. agnus-castus* fruit (Hobberg et al. 1996, Ono et al. 2011, Ono et al. 2009). (b) Maturing fruit. (c) Longitudinal section through a semi-mature fruit prepared for MALDI-MSI (1000 μm scale bar) and corresponding ion maps of colocalising m/z values of putative diterpenoids (see Table 1, all images TIC normalised). (d) Bright field image of fruit epidermis with glandular trichomes (indicated with an arrowhead, 100 μm scale bar). (e) Quantification of vitexilactone (1) in leaf and fruit tissue (n = 3, error bars represent ± 1 SD).
Figure 2. Phylogenetic analysis and relative transcript abundance of *Vitex agnus-castus* diterpene synthases (dTTPSs). Maximum likelihood trees of class II (a) and class I (b) dTTPSs based on aligned protein sequences. Trees are drawn to scale, with branch lengths representing the number of substitutions per site. Branches with > 80% bootstrap support are indicated with a filled circle. Trees were rooted with PpCPS/EKS. GenBank accession numbers are listed in Table S1. (c) Relative transcript abundance of *V. agnus-castus* dTTPSs in leaves, fruit and trichomes as determined by qRT-PCR. For leaf data, the mean of three biological replicates (± 1 SD) is shown; for fruit, tissue was combined from three different plants to obtain sufficient trichome RNA for expression studies. Two-tailed P-values < 0.05 are shown with an asterisk.
Figure 3. *In planta* functional characterization of class II enzyme, VacTPS1, and class I enzymes, VacTPS2 and VacTPS6. (a) GC-MS extracted ion chromatograms (EIC) of hexane extracts of *N. benthamiana* leaves transiently expressing combinations of dITPSs. (b) Diterpene synthase products with representative mass spectra. Characterised dITPS combinations used as references and compound identification details are given in Table S2.
Figure 4. In planta functional characterization of class II enzyme, VacTPS3, and class I enzymes, VacTPS2 and VacTPS6. (a) GC-MS extracted ion chromatograms (EIC) of hexane extracts of *N. benthamiana* leaves transiently expressing combinations of dTPSs. (b) Diterpene synthase products with representative mass spectra. Characterised dTPS combinations used as references and compound identification details are given in Table S2.
Figure 5. In planta functional characterization of class II enzyme, VacTPS5, and class I enzymes, VacTPS2 and VacTPS6. (a) GC-MS extracted ion chromatograms (EIC) of hexane extracts of N. benthamiana leaves transiently expressing combinations of diTPSs. (b) Diterpene synthase products with representative mass spectra. Characterised diTPS combinations used as references and compound identification details are given in Table S2.
Figure 6. Summary of identified *Vitex* agnus-castus diterpene synthase (dTPS) activities. (a) Observed enzymatic activities of *V. agnus-castus* dTPSs. Dashed arrow indicates 15 may not be the direct product of VacTPS6. (b) Examples of functionalised diterpenoids from *Vitex* spp. representing the different diterpene backbones generated by the class II and class I dTPS combinations.
Figure 7. Phylogenetic analysis, relative transcript abundance in Vitis agnus-castus tissues and activity of VacCYP76BK1. (a) Relative transcript abundance of VacCYP76BK1 in leaves, fruit and trichomes as determined by qRT-PCR. For leaf data, the mean of three biological replicates (± 1 SD) is shown; for fruit, tissue was combined from three different plants to obtain sufficient trichome RNA for expression studies. Two-tailed P-values < 0.05 are shown with an asterisk. (b) Maximum Likelihood tree of cytochrome P450s (CYPs) based on aligned protein sequences. The tree is drawn to scale, with branch lengths representing substitutions per site. Branches with > 80% bootstrap support are indicated with a filled circle. The tree was rooted with MsCYP71D18. Genbank accession numbers are listed in Table S9. (c) LC-HRMS total ion chromatograms of microsomal assay extracts prepared from a S. cerevisiae strain expressing VacCYP76BK1 supplied with peregrol (2), and culture extracts of S. cerevisiae strains expressing either diTPSs alone or in combination with VacCYP76BK1.
Figure 8. Proposed pathway for the formation of lactone and furan derivatives of peregrinol (2) via labd-13Z-ene-9,15,16-triol (17). Lactone formation could proceed from 17 via a dehydrogenase catalysed oxidation giving rise to an aldehyde (a) and to a carboxylic acid (b), followed by an enzyme catalysed or spontaneous dehydration reaction to form the lactone ring. Alternatively, dehydration of the tautomeric enol form (c) of the aldehyde intermediate (a) would afford furan ring formation. Dashed arrows indicate putative steps.