The biosynthesis of thymol, carvacrol, and thymohydroquinone in Lamiaceae proceeds via cytochrome P450s and a short-chain dehydrogenase

Sandra T. Krause[a,1], Pan Liao[b,1], Christoph Crocoll[c,2], Benoît Boachon[b,d], Christiane Förster[a,3], Franziska Leidecker[a], Natalie Wiese[a], Dongyan Zhao[e,4], Joshua C. Wood[e], C. Robin Buell[e,f,g], Jonathan Gershenzon[c], Natalia Dudareva[b,h,i], and Jörg Degenhardt[a,5]

[a]Institute of Pharmacy, Department of Pharmaceutical Biotechnology, Martin Luther University Halle-Wittenberg, 06120 Halle, Germany; [b]Department of Biochemistry, Purdue University, West Lafayette, IN 47907-2063; [c]Department of Biochemistry, Max Planck Institute for Chemical Ecology, 07745 Jena, Germany; [d]University of Lyon, UJM-Saint-Etienne, CNRS, BVpam Formation de Recherche en Evolution 3727, 42000 Saint-Etienne, France; [e]Department of Plant Biology, Michigan State University, East Lansing, MI 48824; [f]Plant Resilience Institute, Michigan State University, East Lansing, MI 48824; [g]Michigan State University AgBioResearch, Michigan State University, East Lansing, MI 48824; [h]Purdue Center for Plant Biology, Purdue University, West Lafayette, IN 47907; and [i]Department of Horticulture and Landscape Architecture, Purdue University, West Lafayette, IN 47907-2010

Edited by Rodney Croteau, Washington State University, Pullman, WA; received June 2, 2021; accepted October 31, 2021

Thymol and carvacrol are phenolic monoterpenes found in thyme, oregano, and several other species of the Lamiaceae. Long valued for their smell and taste, these substances also have antibacterial and anti-spasmolytic properties. They are also suggested to be precursors of thymohydroquinone and thymoquinone, monoterpenes with anti-inflammatory, antioxidant, and antitumor activities. Thymol and carvacrol biosynthesis has been proposed to proceed by the cyclization of geranyl diphosphate to γ-terpinene, followed by a series of oxidations via *p*-cymene. Here, we show that γ-terpinene is oxidized by cytochrome P450 monooxygenases (P450s) of the CYP71D subfamily to produce unstable cyclohexadienol intermediates, which are then dehydrogenated by a short-chain dehydrogenase/reductase (SDR) to the corresponding ketones. The subsequent formation of the aromatic compounds occurs via keto–enol tautomerisms. Combining these enzymes with γ-terpinene in in vitro assays or in vivo in *Nicotiana benthamiana* yielded thymol and carvacrol as products. In the absence of the SDRs, only *p*-cymene was formed by rearrangement of the cyclohexadienol intermediates. The nature of these unstable intermediates was inferred from reactions with the γ-terpinene isomer limonene and by analogy to reactions catalyzed by related enzymes. We also identified and characterized two P450s of the CYP76S and CYP736A subfamilies that catalyze the hydroxylation of thymol and carvacrol to thymohydroquinone when heterologously expressed in yeast and *N. benthamiana*. Our findings alter previous views of thymol and carvacrol formation, identify the enzymes involved in the biosynthesis of these phenolic monoterpenes and thymohydroquinone in the Lamiaceae, and provide targets for metabolic engineering of high-value terpenes in plants.

aromatic monoterpenes | Lamiaceae | carvacrol | thymol | thymohydroquinone

The phenolic monoterpenes of the Lamiaceae are widely used constituents of pharmaceuticals, cosmetics, and food products (1). Extracts of plants containing thymol or carvacrol are employed in medicine for their antibacterial, anti-spasmodic, antioxidant, and anti-cancer properties. Because of their pungent, warm, and aromatic odors, they also serve as additives to cosmetics and are used in aromatherapy. Thymol and carvacrol are best known as the aroma compounds of oregano and thyme, in which they provide the herbal, pizza-like tastes that are traditionally used in Mediterranean cuisine and food preservation (2). The occurrence of phenolic monoterpenes is restricted to a few genera in the Lamiaceae (*Thymus*, *Origanum*, *Satureja*, and *Thymbra*), Apiaceae (*Trachyspermum*), and Verbenaceae (*Lippia*). Of these, the essential oils of *Thymus* are the most important commercial source of phenolic monoterpenes (3). *Thymus vulgaris* L. and *Origanum* species also produce the structurally related monoterpenes thymohydroquinone and thymoquinone, which were first described in the essential oil of *Nigella sativa* L. black seed (4). Thymohydroquinone was shown to exhibit anticancer activity (5), and thymoquinone displays anti-inflammatory, hepatoprotective, antioxidant, cytotoxic, and anti-cancer activities (6).

To date, only a few biosynthetic pathways to pharmaceutically valuable, oxidized terpenes have been completely elucidated, such as those leading to artemisinin, paclitaxel, and the phenolic, labdane-type diterpenes of sage and rosemary (7–9). For monoterpenes, a complex biosynthetic pathway has only been described for menthol and its derivatives in *Mentha* (10).

Significance

The monoterpene alcohols thymol, carvacrol, and thymohydroquinone are characteristic flavor compounds of thyme, oregano, and other Lamiaceae. These specialized metabolites are also valuable for their antibacterial, anti-spasmolytic, and antitumor activities. We elucidated the complete biosynthetic pathway of these compounds, which starts with the formation of γ-terpinene from geranyl diphosphate. The aromatic backbone of thymol and carvacrol is formed by P450 monooxygenases in combination with a dehydrogenase via an unstable intermediate. Additional P450s hydroxylate thymol and carvacrol to form thymohydroquinone. Our findings demonstrate a mechanism for the formation of phenolic monoterpenes that differs from previous predictions and provides targets for metabolic engineering of high-value terpenes in plants.

Author contributions: S.T.K., C.C., J.G., N.D., and J.D. designed research; S.T.K., P.L., C.C., B.B., C.F., F.L., N.W., D.Z., J.C.W., and C.R.B. performed research; S.T.K., P.L., C.C., B.B., and C.F. analyzed data; and S.T.K., J.G., N.D., and J.D. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]S.T.K. and P.L. contributed equally to this work.

[2]Present address: DynaMo Center, Department of Plant and Environmental Science, University of Copenhagen, 1871 Frederiksberg C, Denmark.

[3]Present address: Department of Biochemistry, Max Planck Institute for Chemical Ecology, 07745 Jena, Germany.

[4]Present address: Biotechnology Center, Cornell University, Ithaca, NY 14853.

[5]To whom correspondence may be addressed. Email: joerg.degenhardt@pharmazie.uni-halle.de.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2110092118/-/DCSupplemental.

Published December 20, 2021.

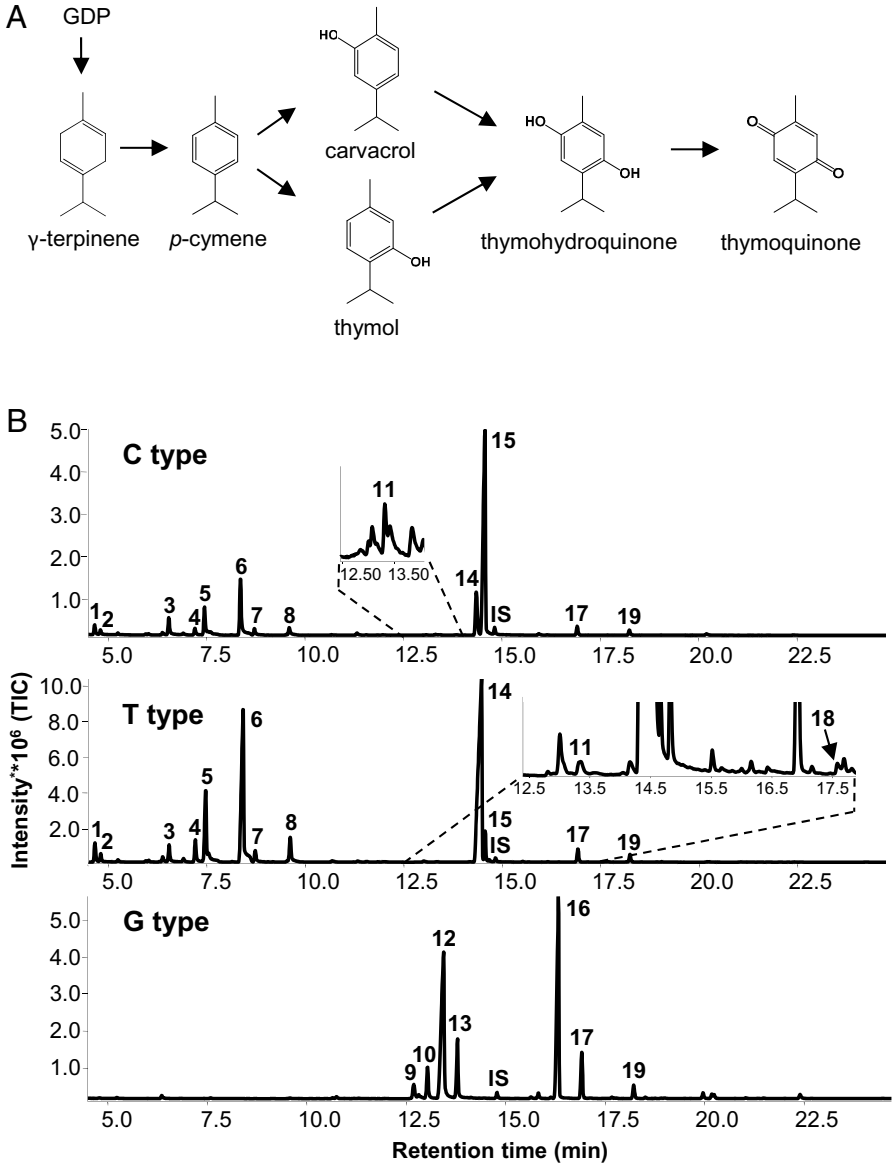

Fig. 1. *T. vulgaris* is composed of different monoterpene chemotypes. (*A*) Proposed biosynthetic pathway to the phenolic monoterpenes thymol and carvacrol as well as thymohydroquinone and thymoquinone. γ-terpinene and *p*-cymene were suggested as intermediates in the formation of thymol and carvacrol (17). (*B*) Essential oil composition of the *T. vulgaris* chemotypes dominated by carvacrol (C type), thymol (T type), and geraniol (G type). Terpenes were extracted with hexane and analyzed by GC-MS. The following terpenes were identified: 1, α-thujene; 2, α-pinene; 3, myrcene; 4, α-terpinene; 5, *p*-cymene; 6, γ-terpinene; 7, *cis*-sabinene hydrate; 8, linalool; 9, nerol; 10, neral; 11, thymoquinone; 12, geraniol; 13, geranial; 14, thymol; 15, carvacrol; 16, geranyl acetate; 17, (E)-β-caryophyllene; 18, thymohydroquinone; and 19, germacrene D. Nonyl acetate (10 μl/mL) was added as internal standard (IS) for quantification.

However, the biosynthetic pathways for phenolic monoterpenes like thymol or carvacrol remain uncharacterized. Most monoterpenes are biosynthesized by fusion of the ubiquitous C_5 intermediates, isopentenyl diphosphate, and its isomer dimethylallyl diphosphate, resulting in the formation of a C_{10} compound, geranyl diphosphate (GDP). This acyclic intermediate is the substrate for the large enzyme family of monoterpene synthases that form cyclic or acyclic products with an enormous variety of carbon skeletons (11, 12). In previous studies in thyme and oregano, the cyclic monoterpene olefin γ-terpinene was proposed as a precursor of thymol and carvacrol (Fig. 1*A*). Studies with ^{3}H-labeled γ-terpinene showed that this compound was converted into thymol and carvacrol after incubation with young thyme leaves (13). Furthermore, the essential oils of plant species rich in either thymol or carvacrol have always been reported to contain substantial amounts of γ-terpinene (3). Terpene synthases forming γ-terpinene have been identified and characterized from various Lamiaceae species (14–19). In oregano, the expression of the γ-terpinene synthase *OvTPS2* was found to correlate with thymol and carvacrol content in leaves (14). Beyond γ-terpinene, however, no further precursors of phenolic monoterpenes have been identified. The aromatic hydrocarbon *p*-cymene was suggested to be an intermediate in thymol and carvacrol formation from γ-terpinene (20), but its participation in the phenolic monoterpene pathway and the nature of the enzymes involved in formation of the aromatic ring still remain unknown. Moreover, there is little information about the conversion of the thymol and carvacrol to thymohydroquinone.

In this study, we investigated the biosynthetic pathway leading to the formation of thymol, carvacrol, and thymohydroquinone in thyme and oregano. We isolated and characterized six cytochrome P450 monooxygenases (P450s) of the CYP71D

subfamily from thyme and oregano accessions producing high levels of thymol and carvacrol. When these *CYP* genes were heterologously expressed and combined with a short-chain dehydrogenase from thyme in vitro or coexpressed in vivo in *Nicotiana benthamiana*, thymol or carvacrol were formed. Based on the characteristics of the expressed enzymes and their reaction with other substrates, we constructed the biosynthetic pathway leading to thymol and inferred the nature of unstable intermediates. Furthermore, we identified and characterized two P450s of the CYP76S and CYP736A subfamilies that hydroxylate thymol and carvacrol to thymohydroquinone when expressed in vivo in yeast and in *N. benthamiana*.

Results

Isolation of a γ-Terpinene Synthase and P450 Monooxygenases of the CYP71D Subfamily from Thyme and Oregano Varieties with High Levels of Thymol and Carvacrol. To investigate the biosynthetic route leading to thymol and carvacrol, we first identified plants rich in these phenolic monoterpenes. Three *T. vulgaris* accessions were chosen from the originally described monoterpene chemotypes (3), two of which contained high levels of either thymol (T type) or carvacrol (C type) in their leaves, while the third contained high levels of geraniol and geranyl acetate (G type) and only low amounts of phenolic monoterpenes (Fig. 1*B*). These chemotypes were used to identify the first step of the pathway, the formation of the monocyclic diene γ-terpinene from GDP (Fig. 1*A*). In analogy to previously identified γ-terpinene synthases from oregano (14, 17) and related thyme species (15, 16, 18, 19), we isolated a gene with >90% amino acid similarity. *TvTPS2* transcript levels correlated positively with the high-carvacrol and -thymol levels in the C and T chemotypes but were very low in plants of the G chemotype (Fig. 2*A*). Expression of *TvTPS2* in a bacterial system confirmed that this gene indeed encodes a functional γ-terpinene synthase (Fig. 2*B*).

The next step of the predicted pathway for phenolic monoterpenes requires the oxidation of γ-terpinene. Thus, we searched Expressed Sequence Tag (EST) databases (14) and available RNA sequencing (RNAseq) data of oregano and thyme for cytochrome P450s (CYPs) of the CYP71D subfamily, which were previously implicated in monoterpene oxidation. The limonene-6-hydroxylase from *Mentha spicata* (CYP71D18) (21) was used as a query in a blastp search of an EST database generated from peltate trichomes of *Origanum vulgare* (14). Completion of matching EST sequences by rapid amplification of complementary DNA (cDNA) ends–PCR in two oregano accessions led to the identification of a gene designated as *CYP71D178* (22). A search for transcripts with similarities to *CYP71D178* in the C and T chemotypes of *T. vulgaris* as well as other accessions of thyme and oregano (*SI Appendix*, Table S1) resulted in 14 different sequences belonging to the *CYP71D* subfamily. Assuming that *CYP* sequences with amino acid identities of 97% or higher represent alleles of the same gene (23), these sequences were assigned to five additional gene names and designated as *CYP71D179*, *CYP71D180*, *CYP71D181*, *CYP71D182* (22), and *CYP71D507*, according to the nomenclature of D. R. Nelson (24) (*SI Appendix*, Table S1).

All of the isolated *CYP71D* sequences encoded proteins with the typical amino acid motifs responsible for CYP catalysis, including a proline-rich hinge, a lipophilic membrane anchor at the N-terminal end, and a heme-binding motif in the C-terminal region, as evident from an amino acid alignment of the CYP71D enzymes, including CYP71D18 from *M. spicata* shown in *SI Appendix*, Fig. S1. A dendrogram analysis including members of the CYP71D subfamily from other plants showed that all identified thyme and oregano enzymes grouped with monoterpene-metabolizing enzymes from other Lamiaceae species (*SI Appendix*, Fig. S2). One cluster contained CYP71D178, CYP71D179, and CYP71D182, while a second was formed by CYP71D180 and CYP71D181. Both clusters showed a close relationship to the limonene-hydroxylating enzymes CYP71D18 and CYP71D13 from *Mentha*. CYP71D507 did not cluster with the other monoterpene-metabolizing enzymes and displayed only about 58% amino acid similarity to them.

The Activity of the CYP71D Enzymes Suggests the Oxidation of γ-Terpinene to a Cyclohexadienol Intermediate. To examine the potential function(s) of these CYPs, we chose the CYP71D enzymes present in the T, C, and G chemotypes of *T. vulgaris*. Expression of *TvCYP71D179-T* was highest in the T type and lowest in the C type, with intermediate expression levels in the G type despite the lack of phenolic monoterpene alcohols in this chemotype (Fig. 2*C*). Conversely, *TvCYP71D180-C* was almost exclusively expressed in the C type, while only trace transcript amounts were detected in the G and T types. The transcript levels of *TvCYP71D507-T* were similar in all three chemotypes. Moreover, a positive correlation between the expression of these CYPs and the levels of thymol and carvacrol was also found to be present in oregano (Fig. 3*C*). Metabolic profiling and transcriptomic analysis of different tissues of oregano (accession "USA") lead to the isolation of *OvCYP71D507*, which exhibits 90.1% amino acid identity with *TvCYP71D507-T* and was well correlated with thymol levels. In addition, an oregano *OvCYP71D180* gene encoding a protein with 91% amino acid identity to *TvCYP71D180-C* appeared to be highly expressed in tissues producing high amounts of carvacrol (Fig. 3 *A* and *B*).

After expression of *TvCYP71D179-T*, *TvCYP71D180-C*, and *TvCYP71D507-T* in yeast and incubation of the microsomal fractions with γ-terpinene, the only monoterpene product detected was *p*-cymene (Fig. 2*D*), an aromatic hydrocarbon previously proposed as an intermediate in phenolic monoterpene biosynthesis (20). However, this product was not accepted as a substrate for further oxidations by these CYPs, as would have been expected for the predicted pathway (Fig. 1*A*).

Because of their similarity with limonene-6-hydroxylase from *M. spicata*, we tested the thyme enzymes with (+)-limonene as a substrate. Limonene is a cyclohexanoid diene monoterpene similar to γ-terpinene but with one of its double bonds in the isopropyl side chain instead of the ring. Both TvCYP71D179 and TvCYP71D507 hydroxylated (+)-limonene at an allylic position, C-3, to form (+)-trans-isopiperitenol (*SI Appendix*, Fig. S3 *A* and *B*). In addition, TvCYP71D179 produced minor amounts of the allylic C-6 hydroxylation products, (+)-trans-carveol and (+)-cis-carveol. In contrast, the major product of TvCYP71D180 was the 6-hydroxylated (+)-cis-carveol, in addition to minor amounts of (+)-trans-carveol and (+)-trans-isopiperitenol (*SI Appendix*, Fig. S3 *A* and *B*). Based on the conversion of limonene to either a C-3 or a C-6 hydroxylated product, and the fact that the CYP71D enzymes characterized previously carry out hydroxylations at either an allylic position or on an aromatic ring, we hypothesized that TvCYP71D179, TvCYP71D180, and TvCYP71D507 produce allylic alcohols from γ-terpinene. If these cyclohexadienol intermediates were unstable under the in vitro assay conditions employed, they could have been converted to *p*-cymene. In this interpretation, *p*-cymene is an artifact rather than a true intermediate of thymol and carvacrol biosynthesis. Thus, we further hypothesized that the cyclohexadienol intermediate might be converted directly to thymol or carvacrol via oxidation to the respective ketone, followed by aromatization via keto–enol tautomerism (Fig. 4*A*). This proposed pathway has some similarities to that of menthol biosynthesis in *Mentha* × *piperita*, in which the CYP71D enzyme catalyzes the allylic hydroxylation of limonene at C-3 to form (−)-trans-isopiperitenol (Fig. 4*B*), which is

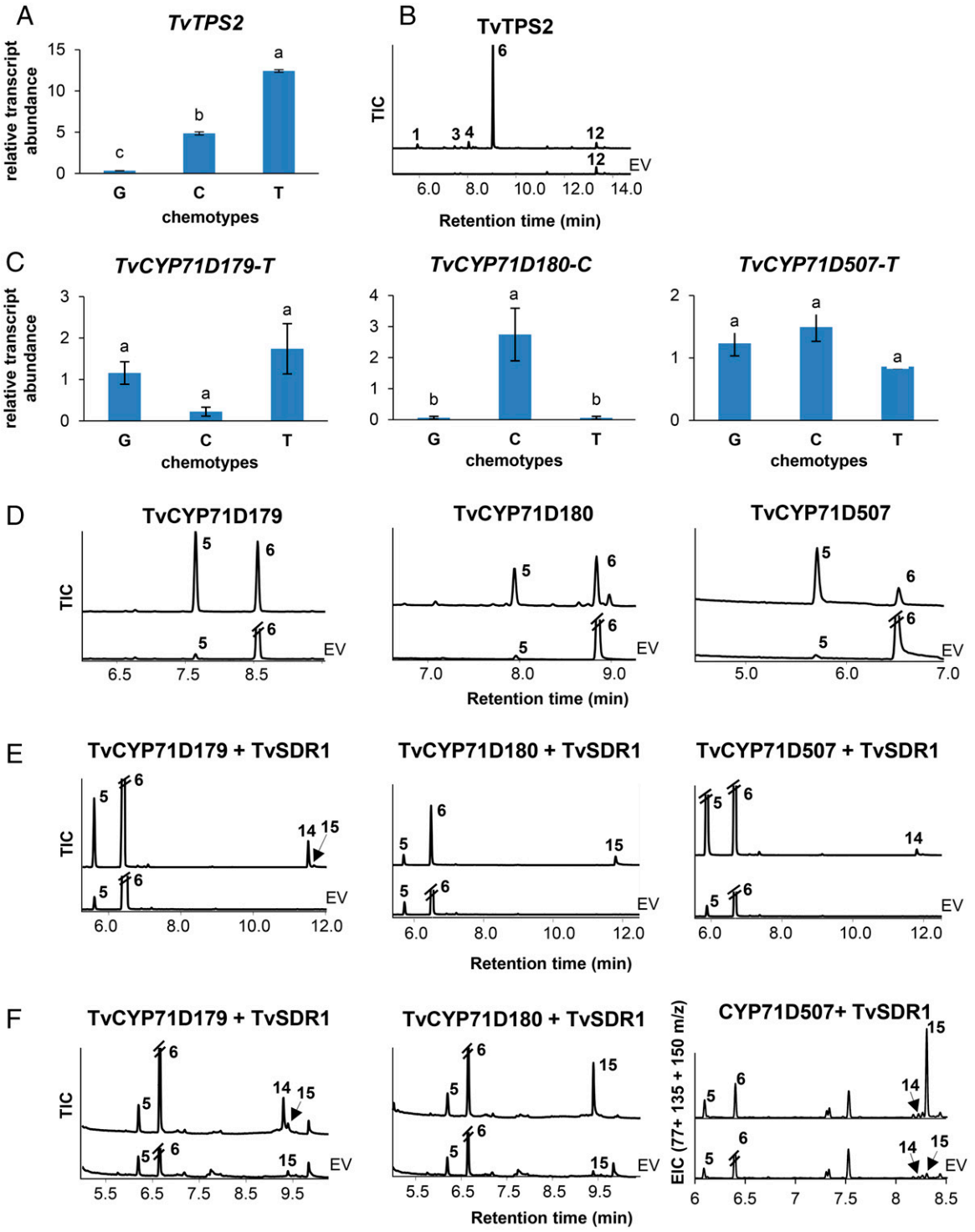

Fig. 2. Expression of four genes involved in the production of phenolic monoterpenes. (*A*) Relative transcript levels of the γ-terpene synthase gene *TvTPS2* in leaves of the *T. vulgaris* chemotypes geraniol (G), carvacrol (C), and thymol (T). (*B*) GC-MS chromatogram of products of TvTPS2 after heterologous expression and incubation with the substrate GDP. (*C*) Relative transcript levels of the CYPs involved in phenolic monoterpene biosynthesis in leaves of the *T. vulgaris* chemotypes C, G, and T. (*D*) GC-MS chromatograms of CYP products after heterologous expression in yeast in the presence of the substrate γ-terpinene. (*E*) GC-MS chromatograms of products after coincubation of CYP and SDR enzymes in the presence of the substrate γ-terpinene. (*F*) GC-MS chromatograms of products after transient transformation of CYP71Ds and *Tv*SDR1 in *N. benthamiana* in the presence of the substrate γ-terpinene. Enzyme products were identified by comparison of their retention times and mass spectra with those of authentic standards. EIC, extracted ion chromatogram; EV, empty vector control; TIC, total ion chromatogram. The following terpenes were identified: 1, α-thujene; 2, α-pinene; 3, myrcene; 5, *p*-cymene; 6, γ-terpinene; 12, geraniol; 14, thymol; and 15, carvacrol. In the quantitative real-time PCR experiments, each bar represents the mean value of three biological replicates ± SE ($n = 3$). Mean values were tested by one-way ANOVA for significant differences ($P < 0.05$). Tukey's post hoc test was used for pairwise comparisons; different letters indicate significant differences.

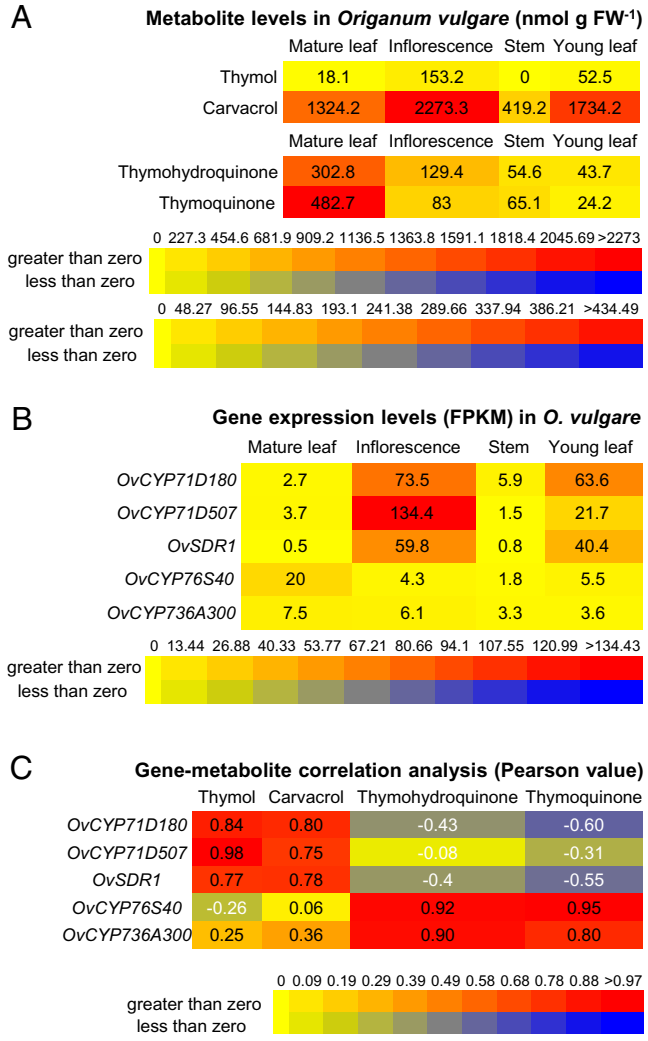

Fig. 3. Representative Pearson's correlation coefficient (PCC) analysis to identify candidate genes for biosynthesis of phenolic monoterpenes in tissues of *O. vulgare* (USA). (*A*) Heatmap showing levels of thymol, carvacrol, thymohydroquinone, and thymoquinone in mature leaf, inflorescence, stem, and young leaf of *O. vulgare*. (*B*) Expression of five candidate genes (four P450s and an SDR, SDR1). Expression levels on heatmap are presented as fragments per kilobase of transcript per million mapped reads (FPKM) from RNAseq datasets. RNAseq data from *O. vulgare* (USA) was validated by qRT-PCR in *SI Appendix*, Fig. S14. (*C*) Representative PCC analysis of gene expression and metabolite levels. Values are means from three biological replicates.

subsequently oxidized by a short-chain dehydrogenase/reductase (SDR), leading to formation of the corresponding ketone (25). To test our hypotheses, we searched for an SDR in thyme that would accept the cyclohexadienol intermediate.

The Short-Chain Dehydrogenase TvSDR1 Catalyzes the Oxidation of Allylic Monoterpene Alcohols to their Corresponding Ketones. To identify genes with similarity to the *M.* × *piperita* trans-isopiperitenol dehydrogenase (ISPD), we searched RNAseq data of the C, T, and G chemotypes from *T. vulgaris*. The sequences *TvSDR1-C, TvSDR1-T*, and *TvSDR1-G* were identified, which encoded the same deduced amino acid sequence with 79% amino acid identity to the trans-ISPD from mint. The deduced amino acid sequence of TvSDR1 contained the motifs typical for this enzyme class, including TGxxxGXG, (C)NAG, YXXXK, and the catalytic tetrad Asn-Ser-Tyr-Lys (*SI Appendix*, Fig. S4). While it would have once been considered a "classical" SDR, this enzyme can now be assigned to the family SDR110C, according to the latest hidden Markov model–based nomenclature system (26, 27). The Asp residue at position 52 in the protein indicated a preference for NAD(H) over NADP(H) as a cofactor. No signal peptides were predicted by SignalP 4.1 (28) or iPSORT (29). Analysis of *TvSDR1* expression revealed no differences in transcript levels in leaves of the C, T and G chemotypes of *T. vulgaris* (*SI Appendix*, Fig. S5). A homologous gene was found in transcriptomic data of oregano (accession "USA"). *OvSDR1*, which exhibits 93.6% amino acid identity with thyme *TvSDR1*, was well correlated with thymol levels (Fig. 3).

The recombinant protein TvSDR1 accepted the alcohols trans- and cis-(−)-isopiperitenol and converted them to the corresponding ketone isopiperitenone (*SI Appendix*, Fig. S6*A*). It also converted trans- and cis-(−)-carveol to the ketone carvone. *p*-Cymene was not accepted as a substrate (*SI Appendix*, Fig. S6*B*).

γ-Terpinene Is Converted to Thymol or Carvacrol by the Combined Activity of a CYP71D Monooxygenase and the Short-Chain Dehydrogenase TvSDR1. To test whether the combination of the CYPs with TvSDR1 produces thymol and carvacrol, microsomes containing TvCYP71D179, TvCYP71D180, or TvCYP71D507 were coincubated with recombinant TvSDR1 in vitro. In assays with TvCYP71D179 or TvCYP71D507, the fed γ-terpinene substrate was converted to thymol, along with minor amounts of carvacrol in case of TvCYP71D179 (Fig. 2*E*). Both minor amounts of thymol and carvacrol are produced by TvCYP71D507. Conversely, carvacrol was produced as the major product when TvCYP71D180 was used instead of TvCYP71D179 in assays. All other CYPs identified in *T. vulgaris* and *O. vulgare* also formed either thymol (OvCYP71D178) or carvacrol (OvCYP71D181 and TvCYP71D182) after coincubation in the presence of the substrate γ-terpinene (*SI Appendix*, Fig. S7). All assays including the empty vector control contained some *p*-cymene, which is formed spontaneously from the γ-terpinene substrate. However, the concentration of *p*-cymene is increased approximately fivefold in the presence of the CYP71D enzymes, suggesting that *p*-cymene might also be formed via the degradation of the unstable dienol intermediate. To test this hypothesis, the γ-terpinene substrate was preincubated with the CYP71D enzymes for 15 min before the addition of TvSDR1. Preincubation increased formation of *p*-cymene (*SI Appendix*, Fig. S8), indicating the instability of the CYP71D products under these assay conditions.

To determine whether the CYP71D enzymes and the identified SDR cooperate to produce phenolic monoterpenes in planta, as they do in an in vitro assay, the corresponding genes were coexpressed in *N. benthamiana*, a species that does not naturally produce any phenolic monoterpenes. After infiltration with the γ-terpinene substrate, leaves transiently expressing *TvCYP71D179-T* and *TvSDR1* produced thymol with small amounts of carvacrol (Fig. 2*F*). Likewise, leaves expressing both *TvCYP71D180-C* and *TvSDR1* formed carvacrol (Fig. 2*F*). The combination of *TvCYP71D507* and *TvSDR1* produced both carvacrol and low levels of thymol in the tobacco expression system and thereby differed from the production in yeast. To reconstitute the complete thymol biosynthetic pathway in *N. benthamiana*, we transiently expressed the γ-terpinene synthase *OvTPS2* from *O. vulgare* (14) together with *TvCYP71D179-T* and *TvSDR1*. OvTPS2 converts the endogenous *N. benthamiana* GDP to γ-terpinene, a substrate for the CYP enzymes. Five d after infiltration, thymol was detected in the *N. benthamiana* leaves (*SI Appendix*, Fig. S9), providing strong evidence that these three enzymes are sufficient for the biosynthesis of aromatic monoterpenes in plants. The relatively low levels of product could be due to the low availability of GDP in *N. benthamiana* leaves or the lack of glandular trichomes in this species. The biosynthesis and accumulation of

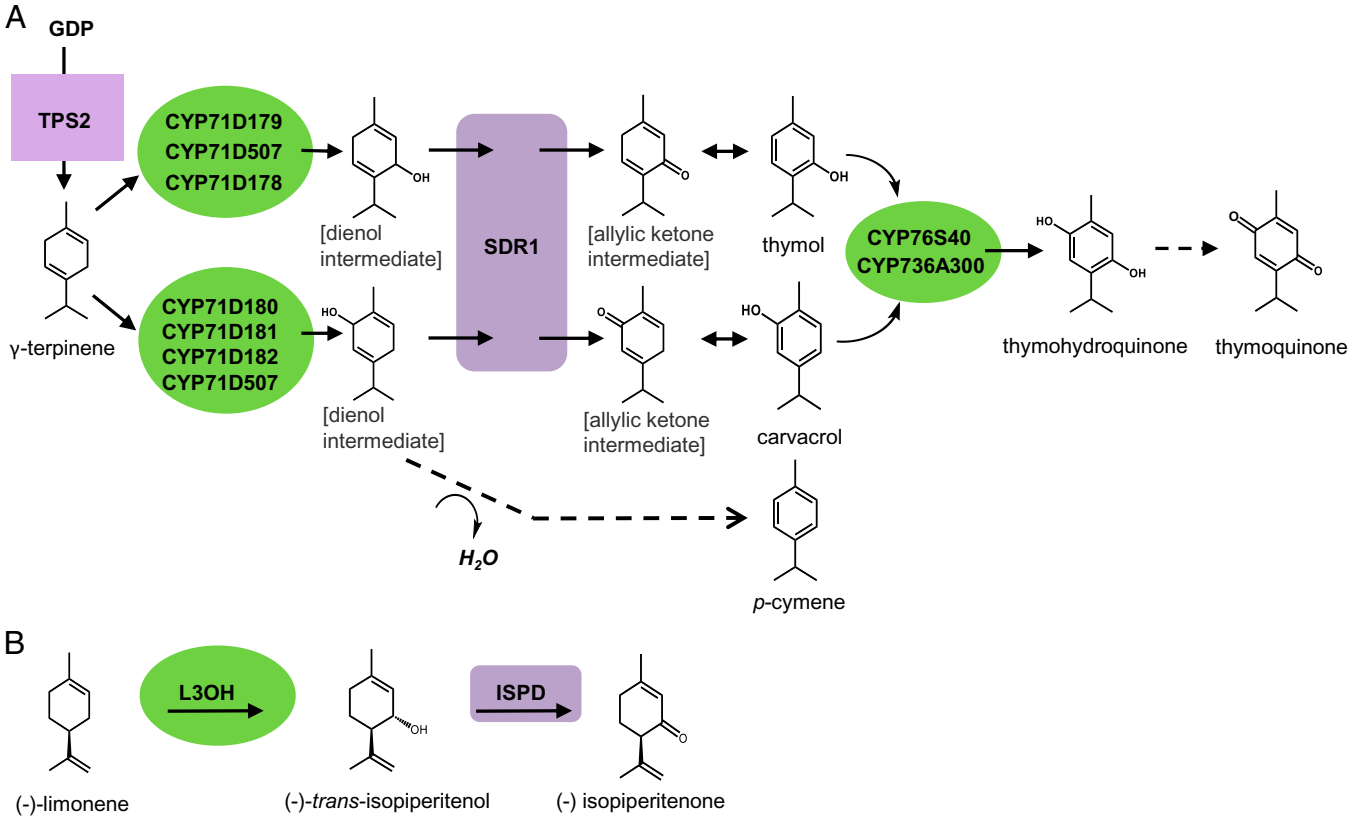

Fig. 4. Scheme of thymol, carvacrol, *p*-cymene, and thymohydroquinone biosynthesis in thyme and oregano in comparison to the first steps of menthol biosynthesis in *Mentha* × *piperita*. (*A*) In thymol, carvacrol, *p*-cymene and thymohydroquinone biosynthesis, geranyl diphosphate (GDP) is first cyclized to γ-terpinene by the terpene synthase (TvTPS2). P450s hydroxylate γ-terpinene either at C-3 or C-6, and the dienol intermediates are converted by a short-chain dehydrogenase (SDR) to the corresponding ketones. These allylic ketones undergo keto–enol tautomerisms to form thymol and carvacrol. The formation of *p*-cymene results from the spontaneous rearrangement of the dienol intermediates due to their instability in aqueous conditions. Thymol and carvacrol are converted to thymohydroquinone by CYP76S40 or CYP736A300 and, subsequently, to thymoquinone by spontaneous conversion or enzymatic action. (*B*) First steps of the pathway from (−)-limonene to menthol in *Mentha* × *piperita*, which also starts with an allylic P450 oxidation. The resulting alcohol is converted to the corresponding ketone by a short-chain dehydrogenase (10). Enzyme abbreviations are the following: L3OH: (−)-limonene-3-hydroxylase (CYP71D13) and ISPD: (−)-(trans)-isopiperitenol dehydrogenase.

monoterpenes in the Lamiaceae and other plant families commonly occurs in glandular trichomes, structures that may promote biosynthesis while minimizing autotoxicity and further metabolism (30). Only traces of monoterpenes were detected in empty vector control transformed *N. benthamiana* leaves.

Gene Expression–Metabolite Correlation Analysis Identified P450 Candidates for Thymohydroquinone Biosynthesis. For the further conversion of thymol and carvacrol to thymohydroquinone, an additional hydroxylation reaction was proposed to be catalyzed by a cytochrome CYP450 hydroxylase. Next, the conversion of thymohydroquinone to thymoquinone involves an oxidation reaction that could be catalyzed by an alcohol dehydrogenase (31). As *O. vulgare* produces both thymohydroquinone and thymoquinone (32), transcriptomics analysis (Dataset S1, Tabs 1 through 3) and metabolic profiling (Dataset S1, Tab 4) of different tissues, including young leaves, mature leaves, stems, and inflorescences, were performed on oregano accession "USA" to discover the biosynthetic genes involved. Gas chromatography–mass spectrometry (GC-MS) analysis revealed that thymohydroquinone and thymoquinone levels were highest in mature leaves, followed by inflorescences, stems, and young leaves, while thymol and carvacrol amounts were highest in inflorescences, followed by young leaves, mature leaves, and stems (Fig. 3*A* and Dataset S1, Tab 4). Several criteria were applied to search and narrow down the candidate genes. 1) As the biosynthesis of thymohydroquinone from thymol and carvacrol requires a cytochrome CYP450 hydroxylase, we searched for "cytochrome P450" and found 450 candidates out of 70,725 transcripts (Datasets S1, Tabs 5 through 7 and S2, Tab 1). 2) Pearson's correlation coefficient (PCC) analysis has been performed between expression of all genes and levels of all metabolites in four different tissues of oregano using the Psych 2.1.3 package available in R (version R 4.0.3). When Pearson's r value was set > 0.7 and $P < 0.01$, the number of candidates was reduced to 42 out of 450 (Dataset S1, Tab 8). 3) Given the average length of plant CYPs, 24 transcripts with an open reading frame less than 1,300 bp were excluded, reducing the number of candidates to 18 (Dataset S1, Tab 9). 4) Two candidates that expressed at very low levels (average fragments per kilobase of transcript per million mapped reads < 5) were excluded, resulting in total of 16 candidates (Dataset S1, Tab 10). 5) Seven candidates grouped with a gene involved in triterpene biosynthesis and one candidate distant from all three groups CYP71, CYP76s, and CYP706s (*SI Appendix*, Fig. S10) were also excluded, resulting in eight candidates left (Dataset S1, Tab 11). 6) The candidate (Vulgare_DN23873_c0_g1_i2_len_1937) was also excluded, since its full-length gene was not found in the assembled oregano genome (32), and its predicted coding sequence based on transcriptomics data could not be amplified by PCR (Dataset S1, Tab 12). These criteria left us with a total of seven candidates, the correlation analysis for which is shown on *SI Appendix*, Fig. S11 and Dataset S1, Tab 12.

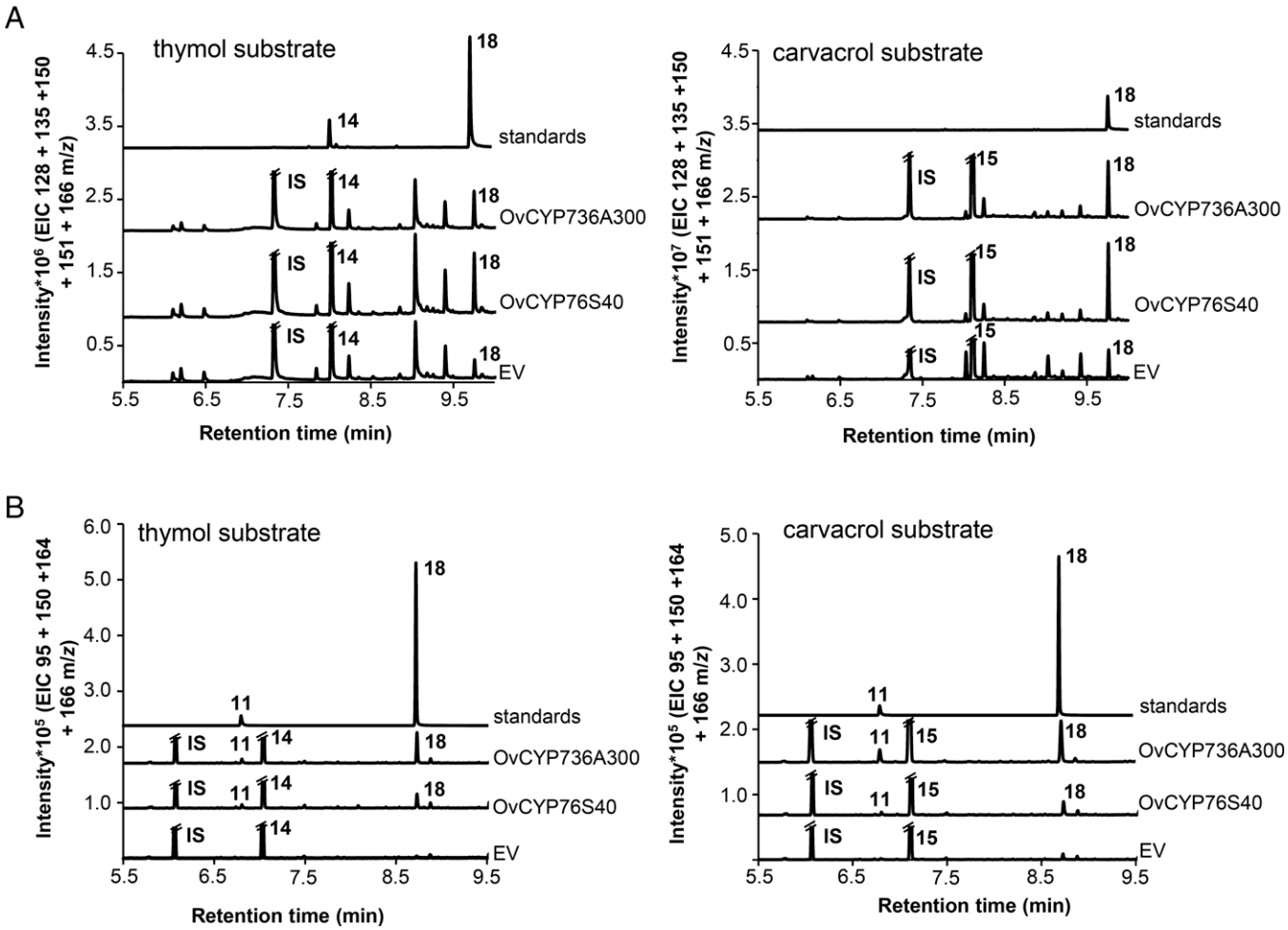

Fig. 5. Thymol and carvacrol are converted to thymohydroquinone by OvCYP76S40 and OvCYP736A300. (*A*) Leaf disks of *N. benthamiana* transiently transformed with *OvCYP76S40* or *OvCYP736A300* were incubated with the substrates thymol or carvacrol for 24 h. Products were extracted and analyzed by GC-MS. Chromatograms show the sum of extracted ion chromatogram (EIC) (*m/z* 128 + 135 + 150 + 151 + 166). IS, internal standard naphthalene. (*B*) Yeast transformed with *OvCYP76S40* or *OvCYP736A300* was incubated with the substrates thymol or carvacrol for 3 d. Products were extracted and analyzed by GC-MS. Chromatograms show the sum of EIC (*m/z* 95 + 150 + 164 + 166). EV, empty vector control. IS, internal standard camphor. The following terpenes were identified: 11, thymoquinone; 14, thymol; 15, carvacrol; and 18, thymohydroquinone.

CYP76S40 and CYP736A300 Catalyze the Formation of Thymohydroquinone from Thymol and Carvacrol in Planta. To test whether the seven selected *OvCYP* candidates convert thymol and carvacrol to thymohydroquinone, the genes were transiently expressed in *N. benthamiana*. Five d after infiltration, leaf discs were fed with thymol or carvacrol for 24 h, and terpenoids were extracted for GC-MS analysis. Only leaves transformed with *OvCYP1* (designated as *OvCYP76S40*) and *OvCYP2* (designated as *OvCYP736A300*) produced significantly higher levels of thymohydroquinone after feeding with thymol and carvacrol than the empty vector–transformed control plants (*SI Appendix*, Fig. S12 and Fig. 5*A*). The small amounts of thymohydroquinone detected in empty vector control leaves are likely the result of unspecific native *N. benthamiana* enzymes that hydroxylate carvacrol and thymol. Analysis of *OvCYP76S40* and *OvCYP736A300* messenger RNA transcripts by RT-qPCR in different oregano tissues further confirmed their expression in thymohydroquinone-producing tissues (*SI Appendix*, Fig. S13), which was consistent with the transcriptomic data. To validate the functions of OvCYP76S40 and OvCYP736A300 independently of a plant background, the proteins were expressed in yeast, followed by feeding with thymol or carvacrol. In yeast, both enzymes were also able to convert thymol and carvacrol into thymohydroquinone. A small amount of thymoquinone was also detected (Fig. 5*B*). Similar alleles of these genes (*TvCYP1* and *TvCYP2*) were also identified in the RNAseq datasets from the *T. vulgaris* "English" accession (*SI Appendix*, Fig. S14). When *TvCYP1* and *TvCYP2* genes were transiently expressed in *N. benthamiana*, leaf discs produced significantly higher levels of thymohydroquinone after feeding with thymol and carvacrol, relative to the empty vector–transformed control plants (*SI Appendix*, Fig. S12 *A* and *B*).

Discussion

A CYP71D Monooxygenase and a Short-Chain Dehydrogenase Are Required for Thymol and Carvacrol Formation from γ-Terpinene. Thymol, carvacrol, thymohydroquinone, and thymoquinone are valuable pharmaceuticals and flavor ingredients that are almost exclusively produced by species of the Lamiaceae. Previous studies suggested that the biosynthetic pathway to these compounds begins with the cyclization of GDP to γ-terpinene by a terpene synthase, and γ-terpinene synthases have been identified in various thyme (15, 16, 18, 19) and oregano (14, 17) species. However, little was known about the later steps, except from the pioneering work of Poulose and Croteau (20), who proposed *p*-cymene as an intermediate from γ-terpinene en route to aromatic monoterpenes. Here, we show that thymol

and carvacrol formation proceeds via hydroxylation of γ-terpinene at either the 3- (to thymol) or 6- (to carvacrol) position catalyzed by CYPs of the 71D subfamily to give unstable products that are likely to be cyclohexadienols. These intermediates are next converted by an SDR, TvSDR1, to their corresponding ketones, which can then aromatize via keto–enol tautomerism, forming thymol and carvacrol (Fig. 4).

Although none of the apparent TvCYP71D179, TvCYP71D180, TvCYP71D181, TvCYP71D182, and TvCYP71D507 products were stable enough for chemical characterization, the formation of cyclohexadienols after allylic oxidation is supported by the fact that these enzymes convert the monoterpene limonene to allylic hydroxylated products (*SI Appendix*, Fig. S3). Enzymes of the CYP71D subfamily from other plants also produce exclusively hydroxylated products (33–37). In the in vitro assays, we detected only the aromatic hydrocarbon, *p*-cymene, a dehydration product of the unstable α,β,α',β'-unsaturated dienol enzyme product (Fig. 2*D*). The other unstable intermediates of the phenolic monoterpene pathway are the TvSDR1 products, which we propose as α,β,α',β'–cross-conjugated cyclohexadienones, since TvSDR1 was found to convert other allylically hydroxylated terpenes to their corresponding ketones (*SI Appendix*, Fig. S6). Following their formation, these ketones readily aromatize to phenols.

CYP enzymes play a major role in the formation of aromatic rings in other terpenes as well. For example, the formation of gossypol, a well-studied sesquiterpene dimer from cotton, requires CYP enzymes of the 71BE, 82, and 706 subfamilies (38). These enzymes all catalyze allylic hydroxylations that facilitate the eventual formation of a naphthalene ring system with three phenolic hydroxyl groups. In the biosynthesis of the abietane diterpene ferruginol, a member of the CYP76 family carries out both a hydroxylation and a dehydrogenation leading to aromatization (39).

The roles of the enzymes involved in this pathway to phenolic monoterpenes were also determined by coexpression of the γ-terpinene synthase *OvTPS2* with *TvCYP71D*179-*T* and *TvSDR1* in *N. benthamiana* in vivo (*SI Appendix*, Fig. S9), as well as the combination of the CYP71Ds with TvSDR1 in vitro (Fig. 2*E*). These experiments resulted in the production of thymol and carvacrol, respectively. The proposed pathway is analogous to that converting limonene to isopiperitenone in *M.* × *piperita* en route to menthol (Fig. 4*B*). Thyme and oregano CYP71D178-182 share 73 to 78% amino acid identity to *M.* × *piperita* limonene-6-hydroxylase (CYP71D18) (*SI Appendix*, Figs. S1 and S2). Similarly, TvSDR1 shares 79% amino acid identity to the (−)-trans-ISPD from *M.* × *piperita*.

Formation of either Thymol or Carvacrol Depends on the Regiospecificity of the CYP71D-Catalyzed Reaction. The pathways to thymol and carvacrol diverge at the CYP71D-catalyzed step, in which some of the characterized enzymes produce a C-3–oxygenated product as an intermediate to thymol, while others produce a C-6–oxygenated product as an intermediate to carvacrol. The regiospecificity of the enzymes was supported by gene expression–metabolite correlations, sequence alignment, and the biochemical characterization of the heterologously expressed proteins. The genes encoding CYP71D178-182 clustered in two groups. The group with CYP71D179 had higher-transcript levels in thyme of the T chemotype and the "English" accession and was associated with thymol production (Fig. 2*C* and *SI Appendix*, Fig. S14). On the other hand, the second group, which contained CYP71D180 and CYP71D181, was correlated with both high-transcript levels in thyme of the C chemotype and high-carvacrol production in *O. vulgare* (Figs. 2*C* and 3). Moreover, TvCYP71D179 produced the C-3–oxygenated thymol when assayed in combination with TvSDR1, while TvCYP71D180 produced the C-6–oxygenated carvacrol under the same conditions (Fig. 2 *E* and *F*). TvCYP71D507 catalyzed a C-3 oxygenation in the yeast expression system (Fig. 2*E* and *SI Appendix*, Fig. S3) and both C-3 and C-6 oxygenations in after infiltration in tobacco leaves (Fig. 2*F*).

The position specificity of the CYP71D enzymes was also evident when the monoterpene (+)-limonene was used as a substrate in the reaction. The thymol-associated TvCYP71D179 formed chiefly the C-3–oxygenated product (+)-trans-isopiperitenol, while the carvacrol-associated TvCYP71D180 produced mainly the C-6–oxygenated product (+)-cis-carveol (*SI Appendix*, Fig. S3). Because of the high-sequence similarity among these enzymes, only a few amino acid residues may be crucial for the product outcome. Among the *Mentha* CYP71D enzymes, one amino acid position was shown to determine the regiospecificity of CYP71D13 (C-3) and CYP71D18 (C-6). In CYP71D18, limonene hydroxylation was altered from C-6 to C-3 when phenylalanine 363 was substituted by isoleucine in the substrate recognition site 5 (21). However, in the thyme and oregano CYP71D enzymes (CYP71D178-182), a phenylalanine was always present at this position with an isoleucine substitution in TvCYP71D507, which may be responsible for variability of hydroxylation among expression systems (Fig. 2 *E* and *F* and *SI Appendix*, Figs. S1 and S3*A*). Thus, phenylalanine 363 is not crucial for regiospecificity in this group. Further sequence analysis and site-directed mutagenesis are required to determine which amino acid(s) is responsible for the position-specific hydroxylation of γ-terpinene en route to thymol and carvacrol.

Interaction of a Short-Chain Dehydrogenase with the CYP71D Enzymes Drives Phenolic Monoterpene Formation. The short-chain dehydrogenase TvSDR1 shares an amino acid sequence identity of 79% to trans-ISPD from mint, which catalyzes the oxidation of trans-isopiperitenol to the respective ketone (25). SDRs have diverse sequences often sharing amino acid sequence identities as low as 15 to 30% (40, 41). Thus, with a sequence identity of 79% between TvSDR1 and the *M.* × *piperita* dehydrogenase, it is not surprising that both enzymes are associated with monoterpene metabolism and catalyze very similar reactions (Fig. 4). Short-chain dehydrogenases are also known to participate in the biosynthesis of other terpenoids, including iridoids and cardenolides (42). Our pathway reconstruction demonstrated the ability of CYP71D and SDR enzymes to work together in planta (Fig. 2*F* and *SI Appendix*, Fig. S9). Similar interactions of a CYP enzyme and a dehydrogenase in terpene metabolism were also reported from other biosynthetic pathways. For example, the formation of perillaldehyde in different *Perilla* species was shown to involve a P450-mediated C-7 hydroxylation of limonene followed by the oxidation of perilla alcohol to perillaldehyde by alcohol dehydrogenases (43). In caraway, the biosynthesis of the oxygenated monoterpene carvone is realized by the hydroxylation of limonene to trans-carveol by a limonene-6-hydroxylase and the subsequent oxidation by a SDR (44). In *Salvia officinalis*, the pathway to sabinol is conducted via a P450-mediated hydroxylation of sabinene to sabinol, which is subsequently dehydrogenated to sabinone (45, 46). A similar pathway for sabinone formation en route to thujone was described for western redcedar (47). The biosynthetic pathway to camphor is realized by hydroxylation of bornyl diphosphate to borneol and subsequent oxidation to camphor in sage, tansy, and ginger (*Zingiber zerumbet*) (46, 48–50). These examples illustrate that P450-mediated hydroxylations are commonly followed by ketone formation through dehydrogenation in biosynthetic pathways leading to derivatives of oxygenated monoterpenes among various plant species.

Similar reaction patterns have been found for the formation of sesquiterpene derivatives. Artemisinin, a pharmaceutically very important sesquiterpenoid lactone, is formed via hydroxylation of amorpha-4,11-diene by CYP71AV1 to artemisinic

alcohol, which is then oxidized to artemisinic aldehyde by an alcohol dehydrogenase (7). The first part of the biosynthesis of costunolide in chicory was found to comprise the hydroxylation of sesquiterpene germacrene A to germacrene alcohol and subsequent oxidation of the hydroxyl group to the respective ketone (germacrene aldehyde). In this case, both oxidation steps were shown to be mediated by the P450 enzyme germacrene A oxidase from the CYP71 family (51).

In the present study, the CYP71D enzymes all produced *p*-cymene in the in vitro assay in absence of the dehydrogenase, likely through dehydration of the initially formed enzyme product, an unstable cyclohexadienol (Fig. 4). Yet *p*-cymene is a frequent constituent of plant essential oils, including those of phenolic monoterpene-accumulating plants, such as thyme and oregano. Therefore, the ratio of *p*-cymene versus thymol or carvacrol in the plant may depend on the relative activities of the CYP71D and SDR enzymes and the physical distance between the two proteins in the cell. If the dehydrogenase is closely associated with the cytochrome P450s, there would be less opportunity for the dienol to rearrange to *p*-cymene. To facilitate the efficient transport of the unstable dienol intermediate and prevent its release into the cytosol, spatial proximity of the CYP71D enzymes to TvSDR1 might be essential. In this context, a localization of TvSDR1 in the mitochondria, as it was determined for ISPD in mint (52), seems not reasonable. The CYP71D enzymes are most likely located at the cytosolic surface of the endoplasmic reticulum (53). Therefore, we assume a cytosolic subcellular location of TvSDR1. According to the prediction server SignalP, TvSDR1 lacks a transit signal peptide, which supports this assumption. A cytosolic location was also reported for other SDRs involved in monoterpene metabolism, for example, isopiperitenone reductase and menthone reductase in peppermint (54) or (+)-trans-carveol dehydrogenase from *Carum carvi* (44).

Thus, both enzymes could be associated via noncovalent interactions, such as hydrogen bonds, hydrophobic bonds, or van der Waals forces (55, 56). The formation of an association with SDR could also occur dynamically by change of conformation of the cytochrome P450 enzyme upon substrate binding (57). Localization studies or yeast two-hybrid experiments could be employed to study the interaction between these enzymes. However, *p*-cymene may not only be formed by spontaneous rearrangement of the CYP71D cyclohexadienol products but also directly by the CYP71D enzymes characterized here, because TvCYP71D507 and TvCYP1D179 produced higher concentrations of *p*-cymene in the presence of TvSDR1 than the empty vector (Fig. 2*E*).

Thymohydroquinone Is Formed by the Subsequent Oxidation of Thymol and Carvacrol. While thymol and carvacrol are formed by the selective oxygenation at C-3 or C-6, we did not observe such a specificity for the subsequent hydroxylation of thymol or carvacrol to thymohydroquinone by OvCYP76S40 or OvCYP736A300. The two enzymes are structurally diverse and fall into different groups of CYPs involved in secondary metabolism (*SI Appendix*, Fig. S2). Nevertheless, they both hydroxylate at C-3 or C-6, the position opposite to the existing hydroxyl group on the aromatic ring, to form the hydroquinone. Although these enzymes accept both thymol and carvacrol substrates in vitro and in vivo after heterologous expression in *N. benthamiana*, we cannot exclude minor differences in substrate preference. When these enzymes were expressed in yeast, thymoquinone was detected as a minor product (Fig. 5*B*). Thymoquinone could be a genuine side product of the enzyme itself but could also arise from spontaneous conversion of thymohydroquinone, as we observed with a thymohydroquinone standard at different temperatures over times ranging from 1 to 7 d (*SI Appendix*, Fig. S15). Therefore, we assume that at least some of the thymoquinone observed was generated nonenzymatically from the relatively large quantities of thymohydroquinone produced by the yeast extracts and likely in planta. Alternatively, the conversion to thymoquinone could be catalyzed by alcohol dehydrogenase activities or other CYPs in the plant. The almost 10-fold difference in thymoquinone to thymohydroquinone ratio, observed among the *O. vulgare* and *T. vulgaris* varieties (Fig. 3*A* and *SI Appendix*, Fig. S14), suggests that the conversion of thymohydroquinone to thymoquinone in the glandular trichomes of these plants depends on more than one mechanism.

Materials and Methods

Plant Material and Cultivation. The origins of the lines and chemotypes of *O. vulgare* and *T. vulgaris* as well as their respective growth conditions are described in detail in *SI Appendix, SI Materials and Methods*. The lines are listed in *SI Appendix*, Table S1.

Molecular Methods and Gene Identification. The identification of *TPS2*, *CYP71D178-182*, *CYP71D507*, and *SDR1* from both *T. vulgaris* and *O. vulgare* lines utilized cDNA libraries, transcriptome sequencing, and gene–metabolite PCC analysis, which are described in detail in *SI Appendix, Materials and Methods*. The genes are listed in *SI Appendix*, Table S1, and primer sequences for the generation of constructs are listed in *SI Appendix*, Table S2. The data of PCC analysis are provided in Datasets S1–S4.

Expression of CYP and SDR Enzymes. Details of the expression of the CYP enzymes in yeast, expression of SDR1 in *Escherichia coli*, enzyme activity assays, and *Agrobacterium tumefaciens*–mediated expression of CYP enzymes in *N. benthamiana* are described in *SI Appendix, Materials and Methods*.

Analysis of Plant Volatile Terpenes. Leaves of *O. vulgare*, *T. vulgaris*, and *N. benthamiana* were frozen in liquid nitrogen. Volatile terpenes were extracted and analyzed by gas chromatography coupled to mass spectrometry, according to the description in *SI Appendix, Materials and Methods*.

Data Availability. All study data are included in the article and/or supporting information. Sequence data from this article can be found under the following GenBank accession numbers: *OvCYP71D178-d* (MK209612); *OvCYP71D178-f* (MK209613), *OvCYP71D179-f* (MK209614), *TvCYP71D179v1-T* (MK209616), *OvCYP71D180-Ct* (MK209618), *TvCYP71D180v1-T* (MK209619), *TvCYP71D180v1-C* (MK209621), *OvCYP71D181-Ct* (MK209622), *TvCYP71D507-T* (MK209625), *TvSDR1-C* (MK174258), *TvSDR1-G* (MK174259), *TvSDR1-T* (MK174260), *OvCYP76S40* (MW890013), and *OvCYP736A300* (MW890014). Raw RNAseq reads have been deposited in the National Center for Biotechnology Information Sequence Read Archive under BioProject PRJNA623602 and PRJNA592145. All other study data are included in the article and/or supporting information.

ACKNOWLEDGMENTS. We thank John Thompson (CNRS, Montpellier, France) for kindly providing the thyme chemotypes, David Nelson (University of Tennessee) for naming of P450 genes, Markus Lange (Washington State University) for providing (−)-trans-isopiperitenol as GC standard and enzyme substrate, and Fabian Bull (Martin Luther University Halle-Wittenberg) for helping with the analysis of thyme RNAseq data. The WAT11 expression strain was kindly provided by P. Urban. We thank Julia Asbach for her help with extracting the first P450 from thyme. This work was supported by a grant from the NSF Plant Genome Research Program (IOS-1444499) to C.R.B. and N.D. and by United States Department of Agriculture National Institute of Food and Agriculture Hatch Project (Grant No. 177845) to N.D.

1. G. Nieto, A review on applications and uses of thymus in the food industry. *Plants (Basel)* **9**, 961 (2020).
2. A. Zarzuelo, E. Crespo, "The medicinal and non-medicinal uses of thyme" in *Thyme: The Genus Thymus*, E. Stahl-Biskup, F. Sáez, Eds. (Taylor & Francis, 2002), pp. 263–292.
3. E. Stahl-Biskup, "Essential oil chemistry of the genus Thymus—A global view" in *Thyme: The Genus Thymus*, E. Stahl-Biskup, F. Saez, Eds. (Taylor & Francis, 2002), pp. 75–124.
4. R. Schneider-Stock, I. H. Fakhoury, A. M. Zaki, C. O. El-Baba, H. U. Gali-Muhtasib, Thymoquinone: Fifty years of success in the battle against cancer models. *Drug Discov. Today* **19**, 18–30 (2014).
5. S. Ivankovic *et al.*, The antitumor activity of thymoquinone and thymohydroquinone in vitro and in vivo. *Exp. Oncol.* **28**, 220–224 (2006).
6. M. Khader, P. M. Eckl, Thymoquinone: An emerging natural drug with a wide range of medical applications. *Iran. J. Basic Med. Sci.* **17**, 950–957 (2014).

7. D.-Y. Xie, D.-M. Ma, R. Judd, A. L. Jones, Artemisinin biosynthesis in *Artemisia annua* and metabolic engineering: Questions, challenges, and perspectives. *Phytochem. Rev.* **15**, 1093–1114 (2016).
8. R. Croteau, R. E. B. Ketchum, R. M. Long, R. Kaspera, M. R. Wildung, Taxol biosynthesis and molecular genetics. *Phytochem. Rev.* **5**, 75–97 (2006).
9. U. Scheler *et al.*, Elucidation of the biosynthesis of carnosic acid and its reconstitution in yeast. *Nat. Commun.* **7**, 12942 (2016).
10. R. B. Croteau, E. M. Davis, K. L. Ringer, M. R. Wildung, (-)-Menthol biosynthesis and molecular genetics. *Naturwissenschaften* **92**, 562–577 (2005).
11. D. J. McGarvey, R. Croteau, Terpenoid metabolism. *Plant Cell* **7**, 1015–1026 (1995).
12. J. Degenhardt, T. G. Köllner, J. Gershenzon, Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. *Phytochemistry* **70**, 1621–1637 (2009).
13. A. J. Poulose, R. Croteau, γ-terpinene synthetase: A key enzyme in the biosynthesis of aromatic monoterpenes. *Arch. Biochem. Biophys.* **191**, 400–411 (1978).
14. C. Crocoll, J. Asbach, J. Novak, J. Gershenzon, J. Degenhardt, Terpene synthases of oregano (*Origanum vulgare* L.) and their roles in the pathway and regulation of terpene biosynthesis. *Plant Mol. Biol.* **73**, 587–603 (2010).
15. A. S. Lima *et al.*, Genomic characterization, molecular cloning and expression analysis of two terpene synthases from *Thymus caespititius* (Lamiaceae). *Planta* **238**, 191–204 (2013).
16. K. Rudolph *et al.*, Expression, crystallization and structure elucidation of γ-terpinene synthase from *Thymus vulgaris*. *Acta Crystallogr. F Struct. Biol. Commun.* **72**, 16–23 (2016).
17. B. Lukas, R. Samuel, J. Novak, Oregano or marjoram? The enzyme γ-terpinene synthase affects chemotype formation in the genus Origanum. *Isr. J. Plant Sci.* **58**, 211–220 (2010).
18. M. D. Mendes, J. G. Barroso, M. M. Oliveira, H. Trindade, Identification and characterization of a second isogene encoding γ-terpinene synthase in *Thymus caespititius*. *J. Plant Physiol.* **171**, 1017–1027 (2014).
19. B. Tohidi, M. Rahimmalek, A. Arzani, H. Trindade, Sequencing and variation of terpene synthase gene (TPS2) as the major gene in biosynthesis of thymol in different Thymus species. *Phytochemistry* **169**, 112126 (2020).
20. A. J. Poulose, R. Croteau, Biosynthesis of aromatic monoterpenes: Conversion of gamma-terpinene to p-cymene and thymol in *Thymus vulgaris* L. *Arch. Biochem. Biophys.* **187**, 307–314 (1978).
21. M. Schalk, R. Croteau, A single amino acid substitution (F363I) converts the regiochemistry of the spearmint (-)-limonene hydroxylase from a C6- to a C3-hydroxylase. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11948–11953 (2000).
22. C. Crocoll, "Biosynthesis of the phenolic monoterpenes, thymol and carvacrol, by terpene synthases and cytochrome P450s in oregano and thyme," PhD thesis, Friedrich-Schiller-University Jena, Jena, Germany (2011).
23. D. R. Nelson *et al.*, P450 superfamily: Update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics* **6**, 1–42 (1996).
24. D. R. Nelson, The cytochrome p450 homepage. *Hum. Genomics* **4**, 59–65 (2009).
25. K. L. Ringer, E. M. Davis, R. Croteau, Monoterpene metabolism. Cloning, expression, and characterization of (-)-isopiperitenol/(-)-carveol dehydrogenase of peppermint and spearmint. *Plant Physiol.* **137**, 863–872 (2005).
26. B. Persson *et al.*, The SDR (short-chain dehydrogenase/reductase and related enzymes) nomenclature initiative. *Chem. Biol. Interact.* **178**, 94–98 (2009).
27. Y. Kallberg, U. Oppermann, B. Persson, Classification of the short-chain dehydrogenase/reductase superfamily using hidden Markov models. *FEBS J.* **277**, 2375–2386 (2010).
28. T. N. Petersen, S. Brunak, G. von Heijne, H. Nielsen, SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786 (2011).
29. H. Bannai, Y. Tamada, O. Maruyama, K. Nakai, S. Miyano, Extensive feature detection of N-terminal protein sorting signals. *Bioinformatics* **18**, 298–305 (2002).
30. R. Schuurink, A. Tissier, Glandular trichomes: Micro-organs with model status? *New Phytol.* **225**, 2251–2266 (2020).
31. I. Botnick *et al.*, Distribution of primary and specialized metabolites in *Nigella sativa* seeds, a spice with vast traditional and historical uses. *Molecules* **17**, 10159–10177 (2012).
32. N. Bornowski *et al.*, Genome sequencing of four culinary herbs reveals terpenoid genes underlying chemodiversity in the Nepetoideae. *DNA Res.* **27**, dsaa016 (2020).
33. A. O. Latunde-Dada *et al.*, Flavonoid 6-hydroxylase from soybean (*Glycine max* L.), a novel plant P-450 monooxygenase. *J. Biol. Chem.* **276**, 1688–1695 (2001).
34. L. Ralston *et al.*, Cloning, heterologous expression, and functional characterization of 5-epi-aristolochene-1,3-dihydroxylase from tobacco (*Nicotiana tabacum*). *Arch. Biochem. Biophys.* **393**, 222–235 (2001).
35. E. Wang *et al.*, Suppression of a P450 hydroxylase gene in plant trichome glands enhances natural-product-based aphid resistance. *Nat. Biotechnol.* **19**, 371–374 (2001).
36. S. Takahashi *et al.*, Functional characterization of premnaspirodiene oxygenase, a cytochrome P450 catalyzing regio- and stereo-specific hydroxylations of diverse sesquiterpene substrates. *J. Biol. Chem.* **282**, 31744–31754 (2007).
37. S. Besseau *et al.*, A pair of tabersonine 16-hydroxylases initiates the synthesis of vindoline in an organ-dependent manner in *Catharanthus roseus*. *Plant Physiol.* **163**, 1792–1803 (2013).
38. X. Tian *et al.*, Characterization of gossypol biosynthetic pathway. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E5410–E5418 (2018).
39. J. Guo *et al.*, CYP76AH1 catalyzes turnover of miltiradiene in tanshinones biosynthesis and enables heterologous production of ferruginol in yeasts. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 12108–12113 (2013).
40. H. Jörnvall *et al.*, Short-chain dehydrogenases/reductases (SDR). *Biochemistry* **34**, 6003–6013 (1995).
41. H. Jörnvall, J.-O. Höög, B. Persson, SDR and MDR: Completed genome sequences show these protein families to be large, of old origin, and of complex nature. *FEBS Lett.* **445**, 261–264 (1999).
42. J. Munkert *et al.*, Iridoid synthase activity is common among the plant progesterone 5β-reductase family. *Mol. Plant* **8**, 136–152 (2015).
43. N. Sato-Masumoto, M. Ito, Two types of alcohol dehydrogenase from Perilla can form citral and perillaldehyde. *Phytochemistry* **104**, 12–20 (2014).
44. H. J. Bouwmeester, J. Gershenzon, M. C. Konings, R. Croteau, Biosynthesis of the monoterpenes limonene and carvone in the fruit of caraway. I. Demonstration of enzyme activities and their changes with development. *Plant Physiol.* **117**, 901–912 (1998).
45. F. Karp, J. L. Harris, R. Croteau, Metabolism of monoterpenes: Demonstration of the hydroxylation of (+)-sabinene to (+)-cis-sabinol by an enzyme preparation from sage (*Salvia officinalis*) leaves. *Arch. Biochem. Biophys.* **256**, 179–193 (1987).
46. S. S. Dehal, R. Croteau, Metabolism of monoterpenes: Specificity of the dehydrogenases responsible for the biosynthesis of camphor, 3-thujone, and 3-isothujone. *Arch. Biochem. Biophys.* **258**, 287–291 (1987).
47. A. Gesell *et al.*, The gymnosperm cytochrome P450 CYP750B1 catalyzes stereospecific monoterpene hydroxylation of (+)-sabinene in thujone biosynthesis in western redcedar. *Plant Physiol.* **168**, 94–106 (2015).
48. R. Croteau, C. L. Hooper, M. Felton, Biosynthesis of monoterpenes. Partial purification and characterization of a bicyclic monoterpenol dehydrogenase from sage (*Salvia officinalis*). *Arch. Biochem. Biophys.* **188**, 182–193 (1978).
49. R. Croteau, N. M. Felton, Substrate specificity of monoterpenol dehydrogenases from *Foeniculum vulgare* and *Tanacetum vulgare*. *Phytochemistry* **19**, 1343–1347 (1980).
50. S. Okamoto *et al.*, A short-chain dehydrogenase involved in terpene metabolism from *Zingiber zerumbet*. *FEBS J.* **278**, 2892–2900 (2011).
51. Q. Liu *et al.*, Reconstitution of the costunolide biosynthetic pathway in yeast and *Nicotiana benthamiana*. *PLoS One* **6**, e23255 (2011).
52. G. W. Turner, R. Croteau, Organization of Monoterpene Biosynthesis in Mentha. Immunocytochemical localizations of geranyl diphosphate synthase, limonene-6-hydroxylase, isopiperitenol dehydrogenase, and pulegone reductase. *Plant Physiol.* **136**, 4215–4227 (2004).
53. S. Bak *et al.*, Cytochromes P450. *Arab. Book/American Soc. Plant Biol.* **9**, tab.0144 (2011).
54. G. W. Turner, E. M. Davis, R. B. Croteau, Immunocytochemical localization of short-chain family reductases involved in menthol biosynthesis in peppermint. *Planta* **235**, 1185–1195 (2012).
55. C. Chothia, J. Janin, Principles of protein-protein recognition. *Nature* **256**, 705–708 (1975).
56. I. Massova, P. A. Kollman, Computational alanine scanning to probe protein-protein interactions: A novel approach to evaluate binding free energies. *J. Am. Chem. Soc.* **121**, 8133–8143 (1999).
57. K. Jørgensen *et al.*, Metabolon formation and metabolic channeling in the biosynthesis of plant natural products. *Curr. Opin. Plant Biol.* **8**, 280–291 (2005).