Occurrence and biosynthesis of plant sesterterpenes (C25), a new addition to terpene diversity

Qingwen Chen¹, Jianxu Li², Yihua Ma^{1,3}, Weiliang Yuan^{1,3}, Peng Zhang² and Guodong Wang^{1,3,*}

¹State Key Laboratory of Plant Genomics, Institute of Genetics and Developmental Biology, The Innovative Academy of Seed Design, Chinese Academy of Sciences, Beijing 100101, China

²National Key Laboratory of Plant Molecular Genetics, CAS Center for Excellence in Molecular Plant Sciences, Chinese Academy of Sciences, Shanghai 200032, China

³College of Advanced Agricultural Sciences, University of Chinese Academy of Sciences, Beijing, China

*Correspondence: Guodong Wang (gdwang@genetics.ac.cn)

<https://doi.org/10.1016/j.xplc.2021.100184>

ABSTRACT

Terpenes, the largest group of plant-specialized metabolites, have received considerable attention for their highly diverse biological activities. Monoterpenes (C10), sesquiterpenes (C15), diterpenes (C20), and triterpenes (C30) have been extensively investigated at both the biochemical and molecular levels over the past two decades. Sesterterpenes (C25), an understudied terpenoid group, were recently described by plant scientists at the molecular level. This review summarizes the plant species that produce sesterterpenes and describes recent developments in the field of sesterterpene biosynthesis, placing a special focus on the catalytic mechanism and evolution of geranylarnesyl diphosphate synthase and sesterterpene synthase. Finally, we propose several questions to be addressed in future studies, which may help to elucidate sesterterpene metabolism in plants.

Key words: terpene, geranylarnesyl diphosphate synthase, metabolic gene clusters, specialized metabolites, sesterterpene (C25)

Chen Q., Li J., Ma Y., Yuan W., Zhang P., and Wang G. (2021). Occurrence and biosynthesis of plant sesterterpenes (C25), a new addition to terpene diversity. *Plant Comm.* **2**, 100184.

INTRODUCTION

Terpenes are the largest and most diverse group of naturally occurring organic compounds, and more than 80 000 terpenoids have been structurally identified to date (Christianson, 2017). In the plant kingdom, terpenes function as defense molecules (also as attractants in some cases), photosynthetic pigments (e.g., chlorophylls and carotenoids), or phytohormones (e.g., brassinosteroids, gibberellic acid, abscisic acid, and strigolactone) (Cheng et al., 2007; Gershenzon and Dudareva, 2007; Santner et al., 2009; Zhou and Pichersky, 2020). Compared with well-known terpenes (mono-, sesqui-, and diterpenes), sesterterpenes (C25) are the rarest group of terpenoids discovered to date, and slightly more than 1000 sesterterpenes have been isolated from various natural sources (Shirley et al., 2018). However, sesterterpenes are widely distributed in bacteria, fungi, insects, and plants (Li and Gustafson, 2021). The first sesterterpene (ophiobolin A, Figure 1) was isolated and described in *Ophiobolus miyabeanus*, a plant pathogenic fungus (Nozoe et al., 1965). In general, sesterterpenes can be

divided into two types based on their structures: the linear type and the carbocyclic type.

Sesterterpene biosynthetic pathways have recently been investigated in fungi and plants: prenyltransferase (PT) and terpene synthase (TPS) are required for the production of sesterterpene scaffolds. Fungal sesterterpene synthases usually contain two functional domains, an N-terminal TPS domain and a C-terminal PT domain. The first chimera sesterterpene synthase, ophiobolin F synthase, was identified from *Aspergillus clavatus* (Chiba et al., 2013). Sesterfisherol synthase from *Neosartorya fischeri* (Ye et al., 2015), stellata-2,6,19-triene synthase from *Emericella variegata* (Matsuda et al., 2015), and other synthases were subsequently identified. The fungal sesterterpene synthases and their catalytic mechanism were well summarized by Minami

Published by the Plant Communications Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and CEMPS, CAS.

et al. (2018). However, the PT and TPS involved in plant sesterterpene biosynthesis are encoded by separate genes, which we will discuss in detail below. In this review, we focus on carbocyclic plant sesterterpenes and their biosynthesis.

Distribution of sesterterpenes (C25) in the plant kingdom and their bioactivities

At the time of writing, approximately 140 sesterterpenes have been identified in multiple plant families, including Asteraceae, Lamiaceae, Lobariaceae, Gentianaceae, and Pteridaceae (Table 1). As shown in Table 1, the functions of most isolated plant sesterterpenes are unknown, probably because only a small number of purified sesterterpenes have been obtained from natural sources. Sufficient amounts of leucosceptroid A (C2) and leucosceptroid B (no hydroxyl group attached at the C11 position, unlike leucosceptroid A, Figure 1) were isolated from the leaf glandular trichomes of *Leucosceptrium canum*, and both compounds were tested for potential bioactivities. Leucosceptroid A and leucosceptroid B exhibited potent antifeedant activity against two generalist insects, beet armyworm (*Spodoptera exigua*) and cotton bollworm (*Helicoverpa armigera*), with EC₅₀ (effective concentration) values ranging from 3.78 to 20.38 µg/cm². Leucosceptroid B inhibited the growth of four agricultural pathogenic fungi (*Colletotrichum musae*, *Colletotrichum gloeosporioides*, *Fusarium oxysporum* f. sp. *niveum*, and *Rhizoctonia solani*) and exhibited EC₅₀ values between 0.27 and 0.68 mM (Luo et al., 2010). Similarly, colquhounoids A to C from the pelate glandular trichomes of *Colquhounia coccinea* var. *mollis* (Lamiaceae) also showed defensive functions against generalist insects and agricultural pathogenic fungi (Li et al., 2013). Recently, leucosceptroid B was further demonstrated to reduce fat storage by downregulating the expression of 10 genes involved in lipid metabolism in *Caenorhabditis elegans* (Ling et al., 2019). Gentianelloids A and B, a pair of epimeric sesterterpenes isolated from the medicinal plant *Gentianella turkestanorum*, showed notable immunosuppressive activity, primarily by inhibiting T cell proliferation (Guo et al., 2020). These results indicated that plant sesterterpenes have a wide range of bioactivities, which clearly warrant further investigation.

With the rapid development of genome mining and biochemical methods in *Arabidopsis* and other Brassicaceae plants (*Brassica rapa*, *Brassica napus*, *Brassica oleracea*, *Camelina sativa*, and *Capsella grandiflora*), approximately 20 sesterterpene skeletons have been produced and identified by reconstruction of their biosynthetic pathway in either *Escherichia coli* or tobacco (*Nicotiana benthamiana*) (Huang et al., 2017, 2018; Shao et al., 2017; Chen et al., 2019, 2020). Notably, only four sesterterpene skeletons were detected (C6–C9, Figure 1) in *Arabidopsis*, and to the best of our knowledge, no oxidized sesterterpenes have been identified in Brassicaceae plants to date (Huang et al., 2017; Chen et al., 2019). Among these skeletons, root-specific C8 and C9 significantly affect root microbiota assembly (9.28% constrained variance) in a similar pattern because they possess similar structures (Chen et al., 2019). The effects of these microbiota assembly changes on the *Arabidopsis* plant itself have not been fully characterized to date.

Sesterterpene biosynthesis pathway in plants

In general, plant terpenoids, including sesterterpenes (C25), are synthesized *de novo* from two common C5 precursors, isopentenyl pyrophosphate (IPP) and its isomer dimethylallyl pyrophosphate (DMAPP), which are generated from two independent pathways: the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway, which operates in plastids, and the mevalonate (MVA) pathway, which operates primarily in the cytoplasm and endoplasmic reticulum (Vranova et al., 2013; Zhao et al., 2013). DMAPP and IPP are condensed to form linear prenyl pyrophosphates with various chain lengths in chain-elongation reactions catalyzed by PTs (EC 2.5.1.1), which can be further divided into *cis*-PT (CPT) and *trans*-PT (TPT; step 1 in Figure 2A). Depending on the chain length of the predominant products, the TPTs are classified as short-chain (C10–C25, SC-PTs), medium-chain (C30–C35), or long-chain (C40–C50) PTs (Vandermoten et al., 2009). Among the SC-PTs, geranyl pyrophosphate synthases (GPPSs; EC 2.5.1.1), farnesyl pyrophosphate synthases (FPPSs; EC 2.5.1.10), and geranylgeranyl pyrophosphate synthases (GGPPSs; EC 2.5.1.29) are responsible for the production of GPP (C10, monoterpene backbone), FPP (C15, sesquiterpene backbone), and GGPP (C20, diterpene backbone), respectively. Similar to other known terpene biosynthesis pathways, cyclization by TPSs (EC 4.2.3.-; step 2 in Figure 2A) and modification (such as oxidation, acylation, and glycosylation, step 3 in Figure 2A) are proposed to be included in the sesterterpene biosynthetic pathway. Two different strategies were employed to investigate sesterterpene biosynthesis in plants. Work in *Arabidopsis* (and other Brassicaceae species) began with the functional identification of GFPPS-sesterTPS-P450 gene clusters *in vitro* (Figure 2B), and sesquiterpenes were subsequently detected *in planta* after the main products of the sesterTPS reactions (compounds 6–9 in Figure 1) were identified by nuclear magnetic resonance (NMR) (Huang et al., 2017; Shao et al., 2017; Chen et al., 2019; He et al., 2019). The *LcaGFPPS1* gene was isolated and identified from the glandular trichomes of *L. canum*, which accumulate high levels of leucosceptroid A and leucosceptroid B (Luo et al., 2010; Liu et al., 2016).

As outlined in Figure 2A, the production of sesquiterpenes and triterpenes occurs in the cytosolic compartment, whereas monoterpenes and diterpenes are synthesized in the plastids of plant cells. Subcellular localization of identified enzymes involved in sesterterpene biosynthesis suggested that sesterterpenes (GFPPSs and sesterTPSs) are produced from the plastidial MEP pathway in plant cells (Figure 2A) (Beck et al., 2013; Nagel et al., 2015; Liu et al., 2016; Chen et al., 2019). This possibility was further confirmed by inhibitor treatment experiments in *L. canum*: treatment with mevinolin (an MVA-specific inhibitor that inactivates 3-hydroxy-3-methylglutaryl CoA reductase) resulted in a 1.3-fold increase in leucosceptroid B content, whereas treatment with fosmidomycin (an MEP-specific inhibitor that inactivates 1-deoxy-D-xylulose 5-phosphate reductoisomerase) led to a pronounced decrease in leucosceptroid B content (Liu et al., 2016).

Lineage-specific GFPPSs diverged from bona fide GGPPSs in plants

The first plant GFPPS was independently identified in *Arabidopsis thaliana* by two research groups (Nagel et al., 2015;

Figure 1. Chemical structures of GFPP and representative sesterterpenes.

Compound **1** (C1) is the first structurally identified sesterterpene from nature, compounds **2** and **3** were isolated from the glandular trichomes of *Lamiaceae* plants, and compounds **4–9** were identified from *Arabidopsis thaliana* (a *Brassicaceae* plant). *Arabidopsis* TPS enzymes, which are responsible for the production of the sesterterpene skeleton from GFPP, are also shown. In **C1** and **C4–C9**, the backbone carbons are numbered following the rule for GFPP, in which the first carbon is attached to the diphosphate group. The origin information for **C10–C15** is presented in Table 1.

Wang et al., 2016). Ten GGPPS-like genes (GGPPSLs), which are thought to produce GGPP for an extended period, are found in the *Arabidopsis* genome (Tholl and Lee, 2011). Both groups demonstrated that four AtGGPPSLs produce GFPP (C25) as the main product *in vitro* when using DMAPP (or GPP, FPP, and GGPP) and IPP as cosubstrates. Moreover, GGPPS activity and GFPP were detected in *Arabidopsis* root

tissue in which AtGGPPS3 (At3g29430) and AtGGPPS4 (At3g32040) were predominately expressed (Beck et al., 2013; Chen et al., 2019). Replacement with a smaller residue at the fifth position before the FARM motif (Ser in GGPPS versus Met in GGPPS) plays a key role in GFPP production. By comparing the crystal structures of AtGGPPS11 (C20, At4g36810), AtGGPPS2 (C25, At3g14550), and AtPPPS2

Source (total identified sesterterpenes, family)	Representative sesterterpenes	Bioactivities	Biosynthetic genes identified	References
<i>Arabidopsis thaliana</i> (11, Brassicaceae)	(+)-thalianatriene (C6 ^b) (–)-retigeranin B (C7)	ND ^d	<i>AtGFPPS1–4</i> , <i>AtTPS18</i> , <i>19</i>	Nagel et al. (2015); Wang et al. (2016); Shao et al. (2017)
	(–)- <i>ent</i> -quiannulatene (C8) (+)-astallatene (C9)	Root microbiota modulation	<i>AtTPS25</i> , <i>30</i>	Huang et al. (2017); Chen et al. (2019)
	(–)-arathanadiene A (C5) (–)-arathanadiene B	ND	<i>AtTPS17</i>	Huang et al. (2018)
	(2 <i>E</i>)-flocerene	ND	<i>AtTPS06</i>	Chen et al. (2020)
<i>Capsella rubella</i> (3, Brassicaceae)	(–)-caprutriene A and C (–)-caprudiene A	ND	<i>Cr237</i> , <i>Cr089</i> (sesterTPS)	Huang et al. (2017, 2018)
<i>Leucoscepttrum canum</i> (30, Lamiaceae)	Leucosceptroid A (C2) and Leucosceptroid B	Antifeedant activities and defense against pathogenic fungi	<i>GFPPS</i>	Luo et al. (2010); Liu et al. (2016)
<i>Colquhounia coccinea</i> (3, Lamiaceae)	Colquhounoid A (C3), Colquhounoids B and C	Antifeedant activities and defense against pathogenic fungi	ND	Li et al. (2013)
<i>Salvia dominica</i> (24, Lamiaceae)	Sesterterpene lactones (C10)	Tubulin tyrosine ligase inhibitors	ND	Dal Piaz et al. (2009)
<i>Solanum tuberosum</i> (1, Solanaceae)	C25-isoprenyl alcohol (Acyclic ^c)	ND	ND	Toyoda et al. (1969)
<i>Artemisia umbelliformis</i> (1, Asteraceae)	Genepolide (C11)	ND	ND	Appendino et al. (2009)
<i>Gentianella turkestanorum</i> (2, Gentianaceae)	Gentianelloid A (C12), Gentianelloid B	Immunosuppressive activity	ND	Guo et al. (2020)
<i>Triticum aestivum</i> (1, Gramineae)	Geranylarnesol isomer (Acyclic)	ND	ND	Akihisa et al. (1999)
<i>Cydonia vulgaris</i> (4, Rosaceae)	Sesterterpene esters (C13)	ND	ND	deTommasi et al. (1996)
<i>Aleuritopteris agetae</i> (5, Sinoptiridaceae)	6/6/6-tricycle sesterterpenes (C14)	ND	ND	Kamaya et al. (1996)
<i>Lobaria aretigera</i> (2, Lobariaceae)	Retigeranic acid A (C15) Retigeranic acid B	ND	ND	Kaneda et al. (1972); Sugawara et al. (1991)
<i>Haslea ostrearia</i> (15, Naviculaceae) ^a	Tetra-unsaturated sesterterpenes (Acyclic)	Inhibition of lung cancer cells	ND	Allard et al. (2001)

Table 1. Representative sesterterpenes produced by plants.^aFour *Haslea* species were investigated.^bThe chemical structure is shown in Figure 1.^cAcyclic sesterterpenes are highlighted.^dND, not determined.

(>C25, At3g20160), Wang et al. (2016) proposed a “three-floor” model (each floor is usually composed of three residues, which are located on helices D, E, and F) to explain the product chain-length determination mechanism of plant GGPPS-like proteins (Figure 3B). The side chains of these amino acid residues can either prevent or enable additional C5 condensation to occur, with larger side chains typically preventing further product elongation. In brief, floor 1 residues determine the product specificity between FPP and GGPP, floor 2 residues determine the product specificity between GGPP and GFPP, and floor 3 residues determine the product specificity between GFPP and even longer chain products (>C25). Using this “three-floor” model, a group of GFPPSs specific to Brassicaceae was discovered (Wang et al., 2016).

Notably, the biochemical data for AtGGPPS1 (encoded by *At1g49530*) differ between the two studies, perhaps because of the different tags used in the recombinant protein. A large GST (glutathione *S-transferase*) tag (26 kDa) fused at the N terminus may have altered the activity of the tested GGPPS enzymes (Wallrapp et al., 2013). Wang et al. (2016) further determined that a large side-chain amino acid (Arg¹³²) located at floor 2 in AtGGPPS1 plays a key role in the maintenance of GGPPS activity. Mutants in which Arg¹³² is replaced with small chain residues (Ala or Val) clearly show GFPPS activity. Further investigations are warranted to clarify the biochemical and physiological functions of AtGGPPS1 in the future.

LcaGFPPS1 was subsequently characterized in *L. canum* (Liu et al., 2016). Phylogenetic analysis clearly showed that the

Figure 2. Sesterterpene biosynthesis pathway in plants.

(A) Simplified terpene biosynthesis pathways in plant cells. Double arrows indicate multiple enzymatic steps. The proposed sesterterpene pathway is highlighted by a pink arrow. Abbreviations for enzymes and chemicals are as described in the text.

(B) Putative sesterterpene gene clusters, including *GFPPS*, *TPS*, and *P450* genes, in the *Arabidopsis* genome.

Brassicaceae *GFPPS* clade and the Lamiaceae *GFPPS* clade (only *LcaGFPPS1* has been characterized to date) separated well, suggesting that *GFPPS* evolved independently from *GGPPS* in these two plant families (Figure 3A). Sequence alignment, together with the “three-floor” model, further suggests the recruitment of *GFPPS* activity by different mutations in floor 1 residues. These include the replacement of Met with Ser in Brassicaceae (the fifth position before the FARM) and the replacement of Leu with Val (one carbon difference in the side chain), probably along with other residue replacements in the Lamiaceae. This proposed model for the Lamiaceae *GFPPS* clade still requires experimental validation.

Plant sesterTPSs and their catalytic mechanism

Due to the presence of *GFPPS-TPS-P450* clusters in the *Arabidopsis* genome (Figure 2B), it is reasonable to investigate whether the *TPS*s located in these biosynthetic gene clusters (BGCs) utilize *GFPP* as a substrate to produce sesterterpene backbones, even though no sesterterpene has been reported

previously in *Arabidopsis*. Using different expression systems, 18 Brassicaceae-specific sesterTPSs were recently characterized, and 20 sesterterpene products were purified and elucidated (Huang et al., 2017, 2018; Shao et al., 2017; Chen et al., 2019, 2020). The main products of AtTPS18 and AtTPS19 were first identified among these sesterterpenes (Shao et al., 2017). Phylogenetic analysis of plant *TPS* sequences clearly shows that functional sesterTPSs evolved from the *TPS*-a subfamily, the members of which always utilize *GPP* and/or *FPP* as substrates (Chen et al., 2019). This result is consistent with the expansion of the *TPS*-a subfamily in Brassicaceae, as there are 22 *TPS*-a members (32 *TPS*s in total) in *Arabidopsis* (Chen et al., 2011, 2020). Because plant sesterTPSs have been identified only from Brassicaceae, it remains unclear whether the sesterTPSs from other plant species (e.g., Lamiaceae) evolved from the *TPS*-a subfamily or from other *TPS* subfamilies.

The crystal structure of the AtTPS18–FSP (farnesyl-S-thiolodiphosphate, an inert *FPP* analog) complex was recently determined in order to explore the cyclization mechanism of plant

Figure 3. Sequence analysis of GGPPSs and GFPPSs from plants.
All Lamiaceae GGPPSs and GFPPSs were obtained from NCBI (<https://blast.ncbi.nlm.nih.gov/>) by BLAST using known plant GGPPS/GFPPS sequences as queries.

(legend continued on next page)

sesterTPSs (Figure 4A). As expected, AtTPS18 (a type-A sesterTPS; cyclization mechanisms will be discussed later) has a larger catalytic pocket than mono-, sesqui-, and di-TPSs, which enables it to accommodate a larger GFPP substrate. The larger catalytic pocket of plant sesterTPSs is mainly due to an outward shift of helix G, and the polar Asn⁴⁹³ residue (this position was occupied by Ala, Figure 4B) plays an important role in pulling helix G outward by forming hydrogen bonds. This hypothesis was further validated by mutagenesis experiments: mutation of Asn⁴⁹³ to Phe (nonpolar residue) decreased sesterTPS activity by 60%, and replacement of the kink sequence AlaGlyMet in helix G with the kink sequence ThrThrTyr resulted in an inactive enzyme (Chen et al., 2020). Analysis of AtTPS18 structure and sequence also demonstrated that the higher number of aromatic residues (five or six) in the catalytic pocket of type-A sesterTPSs compared with classic TPSs (two or three) is a unique feature of type-A sesterTPSs (Figure 4B). Among these residues, Gly³²⁸ and Phe⁴⁹⁶ (in AtTPS18) play important roles in substrate and/or product specificity. AtTPS18^{G328W}, AtTPS19^{G325W}, AtTPS25^{G328W}, and AtTPS30^{P328W} (the 328 position always occupied by an aromatic residue, Trp or Phe, in regular plant TPSs) lost their sesterTPS activity, whereas the AtTPS18^{G328W}, AtTPS19^{G325W}, and AtTPS25^{G328W} mutants utilized GPP, FPP, or GGPP as substrates (Chen et al., 2019). AtTPS18^{G328W} and AtTPS18^{F496C} produced sesterterpenes other than **C6**, the main product of wild-type AtTPS18 (Figure 4C). Therefore, these residue features (Gly³²⁸, Tyr⁴⁶² [or Phe], Tyr⁴⁶⁵, Asn⁴⁹³, and Phe⁴⁹⁶ [or Tyr]) of sesterTPSs can be used to distinguish them from regular TPS-a members.

To date, three early-stage cyclization mechanisms of GFPP in the catalytic pocket of plant sesterTPSs have been proposed (Huang et al., 2017, 2018; Shao et al., 2017; Chen et al., 2020). The proposed cyclization mechanisms (carbocation cyclization paths to the formation of **C9** and **C5**(+)-brarapadiene A / (–)-brarapadiene B / (–)-caprudine A) are partially supported by quantum chemical computations (Huang et al., 2017, 2018). In the type-A route, the 5–15 bicyclic intermediate (IM1 in Figure 4C) is generated via C₁–IV–V cyclization (between the C₁ cation, the C₁₄–C₁₅ double bond (IV), and the C₁₈–C₁₉ double bond (V)). Most of the characterized plant sesterTPSs employ type-A cyclization mechanisms (Figure 4C). Notably, one mutated plant sesterTPS (AtTPS18^{F496C}) produces a monocyclic sesterterpene via C₁–IV cyclization during which a 14 monocyclic intermediate (IM2 in Figure 4C) is formed. These results suggest that Cys⁴⁹⁶ is involved in stabilizing IM2. In the type-B cyclization pathway, the 5–11 bicyclic intermediate (IM3 in Figure 4C) forms via C₁–III–IV cyclization (between the C₁ cation, the C₁₀–C₁₁ double bond (III), and the C₁₄–C₁₅ double

bond (IV)). To date, only one plant sesterTPS (AtTPS06) has been shown to produce a sesterterpene product via type-B cyclization mechanisms. Notably, the aforementioned residue features used to distinguish type-A sesterTPSs from regular TPS-a members cannot be applied to type-B sesterTPSs. A previous attempt to interconvert type-A sesterTPS (AtTPS18) and type-B sesterTPS (AtTPS06) failed, highlighting the complexity of the cyclization mechanisms of plant sesterTPSs (Chen et al., 2020).

CONCLUDING REMARKS AND PERSPECTIVES

Plant terpene metabolism has been investigated at the molecular level for more than 30 years since the biochemical characterizations of the sesquiterpene (C15)-related enzymes were performed (Cane, 1990 and literature therein). However, the current understanding of plant sesterterpenes, at either chemical or biosynthetic levels, is very limited. To date, only approximately 140 sesterterpenes and two types of enzymes (GFPPSs and sesterTPSs) have been identified. As shown in Table 1, however, sesterterpenes are widely distributed in the plant kingdom and are observed in plants ranging from aquatic algae to flowering plants, and these sesterterpenes exert various potential bioactivities. Research on plant sesterterpene metabolism may continue to add new findings and novel insights to the terpene field. Several open questions in this field are summarized here.

1. How wide is the distribution of sesterterpenes in the plant kingdom (novel sesterterpene discovery)? To address this basic question, more chemical analysis work, which relies heavily on chromatography (preparative and analytical scale), mass spectrometry, and NMR, is warranted.
2. What oxidases (P450s or others) and other modification enzymes are involved downstream in plant sesterterpene biosynthesis? In *Arabidopsis*, the P450 genes located in the sesterterpene BGCs will be the first targets to test. Meanwhile, we cannot exclude the possibility that other genes outside the BGCs are involved in the sesterterpene biosynthetic pathway. Gene coexpression analysis will provide useful information for candidate gene screening, given that abundant transcriptome data for *A. thaliana* are publicly available.
3. How are the carbocation intermediates controlled in the catalytic pocket of sesterTPSs to produce different final products, thereby contributing to plant sesterterpene diversity? Computationally guided methods such as quantum mechanical/molecular mechanical molecular dynamics simulations have been employed to study the

(A) Phylogenetic analysis of plant GGPPSs and GFPPSs. The maximum-likelihood phylogenetic tree was constructed using MEGA X software (Kumar et al., 2018) with default parameters. Bootstrap values (based on 1000 replicates) >70% are shown. The Brassicaceae GFPPS group is highlighted in green, and the putative Lamiaceae GFPPS group is highlighted in red. All functionally identified plant GFPPSs are marked with solid green circles. The GGPPS groups from Brassicaceae and Lamiaceae are highlighted in blue.

(B) Sequence alignment of GGPPSs and GFPPSs from Brassicaceae and Lamiaceae plants (several putative Lamiaceae GFPPSs are included). Based on the crystal structure of AtGFPPS2 and AtGGPPS11, the floor 1, 2, and 3 residues (located on helices D–F) are indicated by green triangles, golden squares, and blue stars, respectively. FARM, first aspartate-rich motif. Species abbreviations: Aco, *Aquilegia coerulea* Goldsmith; Aly, *Arabidopsis lyrata*; Ag, *Abies grandis*; At, *Arabidopsis thaliana*; Atr, *Amborella trichopoda*; Cgr, *Capsella grandiflora*; Cpa, *Carica papaya*; Cre, *Chlamydomonas reinhardtii*; Cru, *Capsella rubella*; Bst, *Boechera stricta*; Bra, *Brassica rapa*; Egu, *Erythranthe guttata*; Esa, *Eutrema salsugineum*; Him, *Handroanthus impetiginosus*; Lca, *Leucoscepterum canum*; Os, *Oryza sativa*; Pa, *Picea abies*; Pdu, *Prunus dulcis*; Ppa, *Physcomitrella patens*; Sba, *Scutellaria barbata*; Sin, *Sesamum indicum*; Smo, *Selaginella moellendorffii*; Tha, *Tarenaya hassleriana*.

product distribution controlled by TPSs (Sato et al., 2018; Raz et al., 2020). Together with the available AtTPS18 structure, multiscale simulations of sesterTPS will help us to engineer sesterTPS activity more rationally.

4. How does the whole terpene network, including plastid terpenes and cytosolic terpenes (Figure 2A), change when the sesterterpene pathway is disturbed in plants? Integrated transcriptomic and metabolomic analysis of transgenic plants (gain/loss of function) may help to answer this question.
5. What are the physiological functions of sesterterpenes in *planta* and other bioactivities? In addition to testing the corresponding transgenic plants, a large number of sesterterpenes of interest are needed to screen other bioactivities *in vivo*. To date, at least two systems (*E. coli* and *N. benthamiana*) have been shown to produce sesterterpenes effectively when metabolic genes are well characterized.

FUNDING

This work was financially supported by the Key R&D Program of Shandong Province (grant no. 2019JZZY020610), the National Key R&D Program of China (grant no. 2018YFA0900600), the National Natural Science Foundation of China (grant no. 31970315), and the State Key Laboratory of Plant Genomics of China (grant no. SKLPG2016A-13).

AUTHOR CONTRIBUTIONS

G.W., Q.C., and P.Z. planned the review outline. G.W. wrote the manuscript with contributions from all authors. Q.C., J.L., Y.M., and W.Y. prepared the figures.

ACKNOWLEDGMENTS

We apologize to all colleagues whose original publications are not cited due to space limitations. No conflict of interest declared.

Received: January 22, 2020

Revised: March 3, 2021

Accepted: March 28, 2021

Published: March 30, 2021

REFERENCES

- Akihisa, T., Koike, K., Kimura, Y., Sashida, N., Matsumoto, T., Ukiya, M., and Nikaïdo, T. (1999). Acyclic and incompletely cyclized triterpene alcohols in the seed oils of *Theaceae* and *Gramineae*. *Lipids* **34**:1151–1157.
- Allard, W.G., Belt, S.T., Masse, G., Naumann, R., Robert, J.M., and Rowland, S. (2001). Tetra-unsaturated sesterterpenoids (Haslenes) from *Haslea ostrearia* and related species. *Phytochemistry* **56**:795–800.
- Appendino, G., Tagliatela-Scafati, O., Romano, A., Pollastro, F., Avonto, C., and Rubiolo, P. (2009). Genepolide, a sesterpene gamma-lactone with a novel carbon skeleton from mountain wormwood (*Artemisia umbelliformis*). *J. Nat. Prod.* **72**:340–344.
- Beck, G., Coman, D., Herren, E., Aguila Ruiz-Sola, M., Rodriguez-Concepcion, M., Grisse, W., and Vranova, E. (2013). Characterization of the GGPP synthase gene family in *Arabidopsis thaliana*. *Plant Mol. Biol.* **82**:393–416.
- Cane, D.E. (1990). Enzymatic formation of sesquiterpenes. *Chem. Rev.* **90**:1089–1103.
- Chen, F., Tholl, D., Bohlmann, J., and Pichersky, E. (2011). The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. *Plant J.* **66**:212–229.
- Chen, Q., Jiang, T., Liu, Y.X., Liu, H., Zhao, T., Liu, Z., Gan, X., Hallab, A., Wang, X., He, J., et al. (2019). Recently duplicated sesterterpene (C25) gene clusters in *Arabidopsis thaliana* modulate root microbiota. *Sci. China Life Sci.* **62**:947–958.
- Chen, Q., Li, J., Liu, Z., Mitsunashi, T., Zhang, Y., Liu, H., Ma, Y., He, J., Shinada, T., Sato, T., et al. (2020). Molecular basis for sesterterpene diversity produced by plant terpene synthases. *Plant Commun.* **1**:100051.
- Cheng, A.X., Lou, Y.G., Mao, Y.B., Lu, S., Wang, L.J., and Chen, X.Y. (2007). Plant terpenoids: biosynthesis and ecological functions. *J. Integr. Plant Biol.* **49**:179–186.
- Chiba, R., Minami, A., Gomi, K., and Oikawa, H. (2013). Identification of ophiobolin F synthase by a genome mining approach: a sesterterpene synthase from *Aspergillus clavatus*. *Org. Lett.* **15**:594–597.
- Christianson, D.W. (2017). Structural and chemical biology of terpenoid cyclases. *Chem. Rev.* **117**:11570–11648.
- Dal Piaz, F., Vassallo, A., Lepore, L., Tosco, A., Bader, A., and De Tommasi, N. (2009). Sesterterpenes as tubulin tyrosine ligase inhibitors. First insight of structure-activity relationships and discovery of new lead. *J. Med. Chem.* **52**:3814–3828.
- deTommasi, N., DeSimone, F., Pizza, C., and Mahmood, N. (1996). New tetracyclic sesterterpenes from *Cydonia vulgaris*. *J. Nat. Prod.* **59**:267–270.
- Gershenzon, J., and Dudareva, N. (2007). The function of terpene natural products in the natural world. *Nat. Chem. Biol.* **3**:408–414.
- Guo, K., Liu, X., Zhou, T.T., Liu, Y.C., Liu, Y., Shi, Q.M., Li, X.N., and Li, S.H. (2020). Gentianelloids A and B: immunosuppressive 10,11-seco-gentianellane sesterterpenoids from the traditional uighur medicine *Gentiana turkestanorum*. *J. Org. Chem.* **85**:5511–5515.
- He, J., Chen, Q., Xin, P., Yuan, J., Ma, Y., Wang, X., Xu, M., Chu, J., Peters, R.J., and Wang, G. (2019). CYP72A enzymes catalyze 13-hydroxylation of gibberellins. *Nat. Plants* **5**:1057–1065.
- Huang, A.C., Hong, Y.J., Bond, A.D., Tantillo, D.J., and Osbourn, A. (2018). Diverged plant terpene synthases reroute the carbocation cyclization path towards the formation of unprecedented 6/11/5 and 6/6/7/5 sesterterpene scaffolds. *Angew. Chem. Int. Ed.* **57**:1291–1295.
- Huang, A.C., Kautsar, S.A., Hong, Y.J., Medema, M.H., Bond, A.D., Tantillo, D.J., and Osbourn, A. (2017). Unearthing a sesterterpene biosynthetic repertoire in the Brassicaceae through genome mining reveals convergent evolution. *Proc. Natl. Acad. Sci. U S A* **114**:E6005–E6014.
- Kamaya, R., Masuda, K., Ageta, H., and Chang, H.C. (1996). Fern constituents: sesterterpenoids isolated from fronds of *Aleuritopteris agetae*. *Chem. Pharm. Bull.* **44**:695–698.
- Kaneda, M., Takahashi, R., Iitaka, Y., and Shibata, S. (1972). Retigeranic acid, a novel sesterterpene isolated from lichens of *Lobaria retigera* group. *Tetrahedron. Lett.* **13**:4609–4611.
- Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**:1547–1549.
- Li, C.H., Jing, S.X., Luo, S.H., Shi, W., Hua, J., Liu, Y., Li, X.N., Schneider, B., Gershenzon, J., and Li, S.H. (2013). Peltate glandular trichomes of *Colquhounia coccinea* var. *mollis* harbor a new class of defensive sesterterpenoids. *Org. Lett.* **15**:1694–1697.
- Li, K., and Gustafson, K.R. (2021). Sesterterpenoids: chemistry, biology, and biosynthesis. *Nat. Prod. Rep.* <https://doi.org/10.1039/D0NP00070A>.
- Ling, Y., Teng, L.L., Hua, J., Li, D.S., Luo, S.H., Liu, Y.C., Liu, Y., and Li, S.H. (2019). Leucosceptroid B from glandular trichomes of *Leucosceptum canum* reduces fat accumulation in *Caenorhabditis*

- elegans* through suppressing unsaturated fatty acid biosynthesis. *Chin. J. Nat. Med.* **17**:892–899.
- Liu, Y., Luo, S.H., Schmidt, A., Wang, G.D., Sun, G.L., Grant, M., Kuang, C., Yang, M.J., Jing, S.X., Li, C.H., et al. (2016). A geranylarnesyl diphosphate synthase provides the precursor for sesterterpenoid (C-25) formation in the glandular trichomes of the mint species *Leucosceptum canum*. *Plant Cell* **28**:804–822.
- Luo, S.H., Luo, Q., Niu, X.M., Xie, M.J., Zhao, X., Schneider, B., Gershenzon, J., and Li, S.H. (2010). Glandular trichomes of *Leucosceptum canum* harbor defensive sesterterpenoids. *Angew. Chem. Int. Ed.* **49**:4471–4475.
- Matsuda, Y., Mitsuhashi, T., Quan, Z.Y., and Abe, I. (2015). Molecular basis for stellatic acid biosynthesis: a genome mining approach for discovery of sesterterpene synthases. *Org. Lett.* **17**:4644–4647.
- Minami, A., Ozaki, T., Liu, C.W., and Oikawa, H. (2018). Cyclopentane-forming di/sesterterpene synthases: widely distributed enzymes in bacteria, fungi, and plants. *Nat. Prod. Rep.* **35**:1330–1346.
- Nagel, R., Bernholz, C., Vranova, E., Kosuth, J., Bergau, N., Ludwig, S., Wessjohann, L., Gershenzon, J., Tissier, A., and Schmidt, A. (2015). *Arabidopsis thaliana* isoprenyl diphosphate synthases produce the C25 intermediate geranylarnesyl diphosphate. *Plant J.* **84**:847–859.
- Nozoe, S., Morisaki, M., Tsuda, K., Iitaka, Y., Takahashi, N., Tamura, S., Ishibashi, K., and Shirasaka, M. (1965). The structure of ophiobolin, a C25 terpenoid having a novel skeleton. *J. Am. Chem. Soc.* **87**:4968–4970.
- Raz, K., Levi, S., Gupta, P.K., and Major, D.T. (2020). Enzymatic control of product distribution in terpene synthases: insights from multiscale simulations. *Curr. Opin. Biotech.* **65**:248–258.
- Santner, A., Calderon-Villalobos, L.I.A., and Estelle, M. (2009). Plant hormones are versatile chemical regulators of plant growth. *Nat. Chem. Biol.* **5**:301–307.
- Sato, H., Narita, K., Minami, A., Yamazaki, M., Wang, C., Suemune, H., Nagano, S., Tomita, T., Oikawa, H., and Uchiyama, M. (2018). Theoretical study of sesterfisherol biosynthesis: computational prediction of key amino acid residue in terpene synthase. *Sci. Rep.* **8**:2473.
- Shao, J., Chen, Q.W., Lv, H.J., He, J., Liu, Z.F., Lu, Y.N., Liu, H.L., Wang, G.D., and Wang, Y. (2017). (+)-Thalianatriene and (-)-retigeranin B catalyzed by sesterterpene synthases from *Arabidopsis thaliana*. *Org. Lett.* **19**:1816–1819.
- Shirley, H.J., Jamieson, M.L., Brimble, M.A., and Bray, C.D. (2018). A new family of sesterterpenoids isolated around the Pacific Rim. *Nat. Prod. Rep.* **35**:210–219.
- Sugawara, H., Kasuya, A., Iitaka, Y., and Shibata, S. (1991). Further studies on the structure of retigeranic acid. *Chem. Pharm. Bull.* **39**:3051–3054.
- Tholl, D., and Lee, S. (2011). Elucidating the metabolism of plant terpene volatiles: alternative tools for engineering plant defenses? *Recent Adv. Phytochem.* **41**:159–178.
- Toyoda, M., Asahina, M., Fukawa, H., and Shimizu, T. (1969). Isolation of new acyclic C25-isoprenyl alcohol from potato leaves. *Tetrahedron Lett.* **10**:4879–4882.
- Vandermoten, S., Haubruge, E., and Cusson, M. (2009). New insights into short-chain prenyltransferases: structural features, evolutionary history and potential for selective inhibition. *Cell. Mol. Life Sci.* **66**:3685–3695.
- Vranova, E., Coman, D., and Griessem, W. (2013). Network analysis of the MVA and MEP pathways for isoprenoid synthesis. *Annu. Rev. Plant Biol.* **64**:665–700.
- Wallrapp, F.H., Pan, J.J., Ramamoorthy, G., Almonacid, D.E., Hillerich, B.S., Seidel, R., Patskovsky, Y., Babbitt, P.C., Almo, S.C., Jacobson, M.P., et al. (2013). Prediction of function for the polyprenyl transferase subgroup in the isoprenoid synthase superfamily. *Proc. Natl. Acad. Sci. U S A* **110**:E1196–E1202.
- Wang, C.Y., Chen, Q.W., Fan, D.J., Li, J.X., Wang, G.D., and Zhang, P. (2016). Structural analyses of short-chain prenyltransferases identify an evolutionarily conserved GFPPS clade in *Brassicaceae* plants. *Mol. Plant* **9**:195–204.
- Ye, Y., Minami, A., Mandi, A., Liu, C.W., Taniguchi, T., Kuzuyama, T., Monde, K., Gomi, K., and Oikawa, H. (2015). Genome mining for sesterterpenes using bifunctional terpene synthases reveals a unified intermediate of di/sesterterpenes. *J. Am. Chem. Soc.* **137**:11846–11853.
- Zhao, L.S., Chang, W.C., Xiao, Y.L., Liu, H.W., and Liu, P.H. (2013). Methylerythritol phosphate pathway of isoprenoid biosynthesis. *Annu. Rev. Biochem.* **82**:497–530.
- Zhou, F., and Pichersky, E. (2020). More is better: the diversity of terpene metabolism in plants. *Curr. Opin. Plant Biol.* **55**:1–10.