Characterization and functional reconstruction of a highly productive germacrene A synthase from *Liriodendron chinense*

Weijia Cheng[1,2,3,4,†], Yao Zhi[1,2,5,†], Fangfang Chen[1,2,†], Xiaochun Xiao[1,2], Hui Lu[1,2], Ranjun Li[1,2], Hangzhi Zhu[1,2], Qiuxia Wang[1,2], Xueting Fang[1,2], Zhenni Xu[1,2], Zixin Deng[1,2], Tiangang Liu[1,2,5] and Li Lu[1,2,3,6,*] (iD)

[1]*Department of Urology, Zhongnan Hospital of Wuhan University, School of Pharmaceutical Sciences, Wuhan University, Wuhan, China*

[2]*Key Laboratory of Combinatorial Biosynthesis and Drug Discovery (Ministry of Education), School of Pharmaceutical Sciences, Wuhan University, Wuhan, China*

[3]*Hubei Hongshan Laboratory, Wuhan, China*

[4]*Department of Pharmacy, Renmin Hospital, Wuhan University, Wuhan, China*

[5]*Wuhan Hesheng Technology Co., Ltd., Wuhan, China*

[6]*Hubei Key Laboratory of Urological Diseases, Hubei Clinical Research Center for Laparoscopic/Endoscopic Urologic Surgery, Zhongnan Hospital of Wuhan University, Wuhan, China*

Summary

Plants produce a large array of natural products which play important roles in flavours, fragrances and medicines. However, some high-value plant intermediate metabolites cannot be directly extracted from plants. The tulip tree (*Liriodendron chinense*) in the Magnoliaceae family is rich in sesquiterpenes. Upon characterizing the functions of 11 *Liriodendron chinense* terpene synthases, we discovered that LcTPS3 could produce high yields of (+)-germacrene A, which was shown to be a central scaffold in sesquiterpene biosynthesis. This compound can be completely transformed into β-elemene at high temperature, a broad-spectrum antitumor drug widely used in clinical treatment. By expressing LcTPS3 in a precursor-providing *Saccharomyces cerevisiae* chassis and with the aid of metabolic engineering, the fermentation yield of (+)-germacrene A has been achieved at 14.71 g/L. Site-directed mutagenesis experiments and molecular dynamics simulations revealed that the A280V suppresses the cyclization of substrate by influencing the conformation of the enzyme-substrate. The Y282L facilitates secondary cyclization to produce α-guaiene by shortening the distance between the catalytic residue Y531 and the substrate. These insights underscore the high plasticity of LcTPS3 and suggest that its targeted engineering could unlock the synthesis of a wider array of valuable sesquiterpenes.

Introduction

Synthetic biology has demonstrated significant advantages in the synthesis of plant-derived compounds, such as lycopene, artemisinin and cannabidiol (Du *et al.*, 2023; Li *et al.*, 2023; Tahir *et al.*, 2021). In addition to these end products, some plant intermediate metabolites also have high value. Germacrene A was shown to be a central scaffold in sesquiterpene biosynthesis; its further skeletal rearrangements can yield a number of sesquiterpene derivatives, such as eudesmanes and guaianes (Xu and Dickschat, 2020). In addition to being an intermediate in sesquiterpene biosynthesis, germacrene A itself can be further converted into β-elemene with 100% yield through a Cope rearrangement reaction at high temperature (Kraker *et al.*, 1998). β-elemene, a sesquiterpene derived from the medicinal plant *Rhizoma zedoariae*, is widely used in clinical medicine due to its broad-spectrum antitumor activity (Li *et al.*, 2010; Wang *et al.*, 2005). However, the unsustainable plant extraction prompted the search for environmentally friendly strategies for germacrene A or β-elemene production.

Germacrene A was derived from the linear C15 precursor farnesyl diphosphate (FPP) catalysed by germacrene A synthases (GAS) (Kraker *et al.*, 1998). Most of the GAS were cloned from plants and a small number have been found in microorganisms.

Several reports have screened the activities of GAS homologues from different sources by expressing LsLTC2 (LsGAS2) from *Lactuca sativa* (Chen *et al.*, 2022; Hu *et al.*, 2017; Zhang *et al.*, 2021) or NpGAS from cyanobacteria *Nostoc punctiforme* strain PCC 7120 (Li *et al.*, 2022). De novo germacrene A synthesis platforms were successfully established in *Escherichia coli* and *Saccharomyces cerevisiae*. Site-directed mutagenesis or co-synthesis systems were subsequently performed to achieve the highest β-elemene production of 4.7 g/L (Ye *et al.*, 2023). However, the improvement of the titer of germacrene A was still needed and the ambiguous cyclization mechanism of GAS hinders enzyme engineering efforts aimed at further yield improvement.

Terpene synthase (TPS) converts prenyl diphosphates of different lengths into hundreds of often stereochemically complex mono- and polycyclic hydrocarbons, which seed the biosynthesis of thousands of derivatives through downstream metabolic pathways (Christianson, 2017). Cyclization is the major gateway to chemical diversity in terpene synthesis. Germacrene A is a central scaffold in the biosynthesis of sesquiterpenes; its synthesis includes efficient 1,10 cyclization of FPP and avoiding secondary cyclization to form other by-products (Xu and Dickschat, 2020). Identification of key substitutions that unlock

cyclization may provide fundamental structural insights and mechanistic clues about how ring formation evolved, and hence a basis to explore this phenomenon in the TPS family. Although the resolution of TPS crystal structure and quantum chemical calculations provided insights into the cyclization mechanisms of sesquiterpenes (Gennadios et al., 2009; Köksal et al., 2011; Starks et al., 1997; Wendt et al., 1997; Zhang et al., 2016), the sequence variation and the difficulty in resolving their crystal structures make it challenging. However, molecular docking and molecular dynamics simulations combined with point mutation experiments can also effectively assist us in deciphering the catalytic mechanisms (Jin et al., 2024; Srivastava et al., 2021).

Liriodendron chinense belongs to Magnoliaceae of magnoliids, with highly economically valuable ornamental tree species. Numerous terpenoid constituents, including monoterpenes, sesquiterpenes and sesquiterpene lactones, have been isolated from this species, which contributes to its fragrance accumulation and high resistance to plant diseases and insect pests (Chen et al., 2013; Dong et al., 2009; Graziose et al., 2011). Here, through functional screening of terpene synthases from *Liriodendron chinense*, we characterized a germacrene A synthase (LcTPS3), which showed a high productive activity when heterologously expressed in the *S. cerevisiae* platform. Through a series of mutagenesis analyses and molecular dynamics simulations, we identified the pivotal residue responsible for substrate cyclization. By reengineering the molecular choreography of LcTPS3, we successfully obtained new sesquiterpene products, demonstrating its high plasticity.

Results

Functional characterization of terpene synthases from *Liriodendron chinense* revealed a highly productive germacrene A synthase

The Liriodendron genus is rich in monoterpenes, sesquiterpenes and sesquiterpene lactones, including compounds such as peroxyferolide and lipiferolide, which are derivatives of germacrene A (Figure 1a) (Graziose et al., 2011). To unravel the biosynthesis genes responsible for germacrene A production in *Liriodendron chinense*, 40 candidate terpene synthase genes were identified from its genome. These include 19 TPS-a, 10 TPS-b and 3 TPS-g genes, aligning with the previous categorization of plant TPS genes (Figure 1b) (Chen et al., 2011; Jia et al., 2022). Utilizing transcriptome data from various tissues, the expression profile of the 40 terpene synthase genes was visualized through a heatmap (Figure 1b). Integrating the phylogenetic relationship and the expression pattern of these TPS genes, 11 TPS were chosen as candidates (Table S1).

The 11 TPS genes underwent codon optimization and synthesis, followed by heterologous expression in the *S. cerevisiae* JCR27 strain, which has been genetically modified to increase the activity of key enzymes in the mevalonate (MVA) pathway (Siemon et al., 2020). The fermented products were subsequently analysed using gas chromatography–mass spectrometry (GC–MS) (Table S2). When co-expressed with monoterpene precursor synthase (GPPS), LcTPS8 produced α- and β-pinene (T2, T5) (Figure 1c). LcTPS9 was responsible for the production of β-ocimene (T9). Moreover, LcTPS10 is capable of synthesizing up to 13 different monoterpene products, which greatly enrich the plant's volatile aromatic profile. When co-expressed with sesquiterpene precursor synthase (FPPS), LcTPS2 primarily produced β-caryophyllene (F5) (Figure 1d). LcTPS5 yielded an array of 9 sesquiterpenes and sesquiterpene alcohols, including β-caryophyllene (F5), α-copaene (F2) and 4-*epi*-cuebol (F9). Noteworthy, LcTPS3 is particularly efficient in producing a high quantity of β-elemene (F4).

Metabolic engineering for high-yield germacrene A production in a yeast chassis

The product of LcTPS3 (F4) was identified as β-elemene by GC–MS; however, the high injection port temperature (250 °C) in GC–MS may lead to the complete conversion of germacrene A into β-elemene. Therefore, F4 was isolated for further structure identification. A positive specific optical rotation ($[a]_D^{20} = +26.31$, c 1.0, CH_3CH_2OH) was obtained by spectropolarimeter, which confirmed F4 is indeed (+)-germacrene A (Adio et al., 2004; Faraldos et al., 2007). Subsequently, we purified the LcTPS3 protein from heterologous expression in *Escherichia coli* and determined its optimal catalytic conditions *in vitro* (Figure S1). Our results indicated that the highest catalytic activity of LcTPS3 was achieved at 35 °C and pH = 7.5 (Figure S2). Under this condition, the enzyme kinetics parameters were further determined as $K_m = 17.72 \pm 0.47$ μM and $K_{cat} = 1.90 \pm 0.33 \, s^{-1}$. Compared to the previously reported data of GAS in *Lactuca sativa* LsLTC2 (LsGAS2) (Ye et al., 2023), *Barnadesia spinosa* BsGAS1/2 (Nguyen et al., 2016) and *Artemisia annua* AaGAS (Bertea et al., 2006), LcTPS3 demonstrates the highest activity (Table 1). The catalytic efficiency of LcTPS3 was 4.1-fold higher than that of LsLTC2. This superior performance suggests that LcTPS3 holds considerable promise for its application in the field of germacrene A fermentation engineering.

Several genes that encode germacrene A synthase have been described and utilized in metabolic engineering, including LsLTC2 from *Lactuca sativa* (Ye et al., 2023), NPGAS from *Nostoc* sp. PCC 7120 (Fordjour et al., 2023) and STpGMAS from *Tanacetum parthenium* (Majdi et al., 2011). However, LcTPS3 exhibits sequence divergence from these known germacrene A synthases, with only 40–43% amino acid sequence identity (Figures 2a and S3). To compare their catalytic capabilities on a yeast platform, we individually co-expressed LcTPS3, NPGAS, LsLTC2 and STpGMAS with a yeast–derived farnesyl diphosphate synthase (EGR20) in the JCR27 yeast strain (Figure 2b). While NPGAS yielded a higher production in the yeast chassis, it also generated a range of by-products, including 4-*epi*-cubebol (F9) and germacrene-D-4-ol (F11). To evaluate the catalytic efficiency of the four germacrene A synthase genes in yeast during shake flask fermentation experiments, we integrated a plasmid containing various germacrene A synthases and *ERG20* (yeast FPP synthase) under the regulation of galactose-inducible promoters (P_{GAL1}–P_{GAL10}) into the chromosome of the *S. cerevisiae* strain JCR27 (Table S3). In a 50 mL shake flask fermentation experiment, after 72 h, strain JLcTPS301 that harbours LcTPS3 produced a significantly higher titer of germacrene A, amounting to 151.89 mg/L (Figure 2c).

We then employed metabolic engineering strategies to further enhance the germacrene A titer (Figure 2d). We first incorporated three copies of *LcTPS3* and added an extra copy of truncated *HMG1* (encodes HMG-CoA (3-hydroxy-3-methylglutaryl coenzyme A) reductase) and *ERG20* into the chromosome. To reduce the competition for the FPP substrates caused by yeast squalene synthase (ERG9), we attenuated the upstream activating sequences (UAS) of the *ERG9* promoter. In addition, the *Ura3*
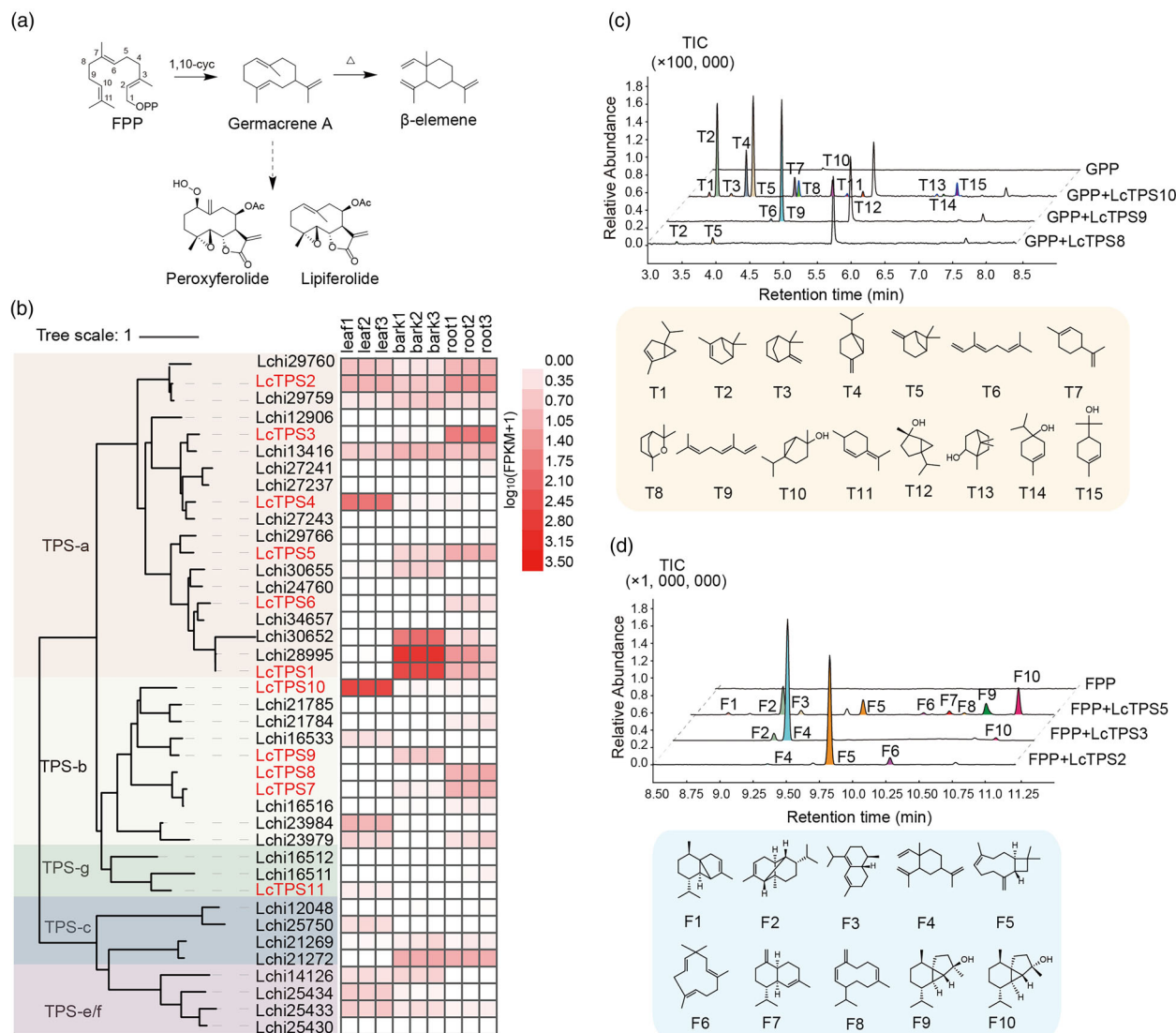
Figure 1 Functional characterization of 11 terpene synthases from *Liriodendron chinense*. (a) The components of sesquiterpenes and their derivatives in *L. chinense* (Graziose *et al.*, 2011). (b) Phylogenetic analysis and a heatmap representation of the gene expression profiles of the TPS gene family in *L. chinense*. FPKM, Fragments Per Kilobase of transcript per Million mapped reads. (c) GC–MS total ion chromatograms and chemical structures of the products of LcTPS8, LcTPS9 and LcTPS10. Different products were labelled as peaks of different colours. T1, α-thujene; T2, α-pinene; T3, camphene; T4, sabinene; T5, β-pinene; T6, *trans*-β-ocimene; T7, limonene; T8, 1,8-cineole; T9, β-ocimene; T10, sabinene hydrate; T11, isoterpinolene; T12, *cis*-sabinene hydrate; T13, borneol; T14, terpinen-4-ol; T15, α-terpineol. (d) GC–MS total ion chromatograms and chemical structures of the products of LcTPS2, LcTPS3 and LcTPS5. F1, α-cubebene; F2, α-copaene; F3, epizonarene; F4, β-elemene; F5, β-caryophyllene; F6, α-caryophyllene; F7, γ-muurolene; F8, germacrene D; F9, 4-*epi*-cubebol; F10, cubebol. Products were identified by comparison to standards or the NIST17 library.

gene was inserted into the *GAL80* locus and the resulting mutant strain with the correct disruption of the *GAL80* gene was designated as JLcTPS309; this strain was therefore induced by glucose deprivation instead of galactose, which significantly lowered the production costs (Shi *et al.*, 2019; Xie *et al.*, 2014). The germacrene A titer of the final engineered strain JLcTPS309 reached 1.11 g/L in 50 mL shake flasks (Figure 2e).

To assess the industrial feasibility of the engineered yeast strain, we performed fed-batch fermentation of the JLcTPS309 mutant in 5 L bioreactors. The fermentation process was divided into two distinct phases: strain growth and product accumulation. The glucose concentration in the medium was 40 g/L at the beginning of the fermentation, which was gradually decreased with the growth of the strain. When the glucose concentration was around 1 g/L, the supply of glucose was increased to maintain the normal growth of the strain. Then isopropyl myristate (IPM) was added after 24 h to enable an aqueous–organic two-phase fermentation, followed by the product accumulation phase, while the ethanol concentration was strictly controlled in the range of 2–5 g/L. The fermentation process ended when product accumulation no longer increased significantly and this method increased the fermentation titer to 14.71 g/L (Figure 2f), which is the highest level of microbial heterologous synthesis in *S. cerevisiae* reported to date.

Table 1 Kinetic parameters of the LcTPS3 and other germacrene A enzymes

Enzyme	Substrate	K_m (μM)	K_{cat} (s^{-1})	k_{cat}/K_m (s^{-1} μM^{-1})
LcTPS3	FPP	17.72 ± 0.47	1.90 ± 0.33	0.107
AaGAS (Bertea *et al.*, 2006)	FPP	10.8 ± 0.32	0.25 ± 0.02	0.023
LsLTC2 (LsGAS2) (Ye *et al.*, 2023)	FPP	10.9 ± 0.16	0.28 ± 0.03	0.026
BsGAS1 (Nguyen *et al.*, 2016)	FPP	7.8 ± 0.25	0.13 ± 0.01	0.017
BsGAS2 (Nguyen *et al.*, 2016)	FPP	14.8 ± 0.29	0.28 ± 0.02	0.019

LcTPS3$_{A280}$ plays a critical role in germacrene A biosynthesis

In the process of functional characterization of LcTPSs, we found LcTPS4, which showed a 78% sequence similarity with LcTPS3, did not produce any terpene products. To identify key amino acid residues in LcTPS3 essential for its function, we generated fragment substitution mutants for LcTPS3 and LcTPS4 (Figure 3a). The results revealed that the amino acids fragment 270–280 plays a critical role in LcTPS3 activity. There are two amino acid differences in this fragment and further point mutations demonstrated that the A280T mutation resulted in a complete loss of LcTPS3 activity (Figure 3a).

We then generated and assayed more LcTPS3$_{A280}$ mutants by replacing A280 with nine representative amino acid residues (Figures 3b and S4). The LcTPS3$_{A280S}$ and LcTPS3$_{A280C}$ mutants exhibited a slight decrease in activity, while LcTPS3$_{A280G}$ maintained its enzymatic activity but with a significantly reduced product yield. Mutations into L, F, K and E resulted in an almost complete loss of LcTPS3's function. Interestingly, when A280 was mutated to V or T, the cyclized germacrene A was no longer produced and a non-cyclized single product, farnesol, was observed in LcTPS3$_{A280V}$ (Figures 3b and S4).

To further investigate the enzymatic structure–function relationship, we tried protein crystallization but unfortunately failed. We then modelled the predicted LcTPS3 structure with (*E, E*)-FPP substrate equivalent to that observed in the crystal structure of 5-*epi*-aristolochene synthase with FPP (PDB:1HXA) (Starks *et al.*, 1997). In this model, residue A280 located far from the catalytic core regions may indirectly influence the formation of the product by affecting the folding of the substrate FPP (Figure 3c). To acquire comprehensive structural insights, we resorted to molecular dynamics (MD) simulation. Referring to a previous work by Srivastava *et al.* (2023), we established the most likely initial starting positions of FPP through computational simulation. Out of the eight potential starting binding models, the
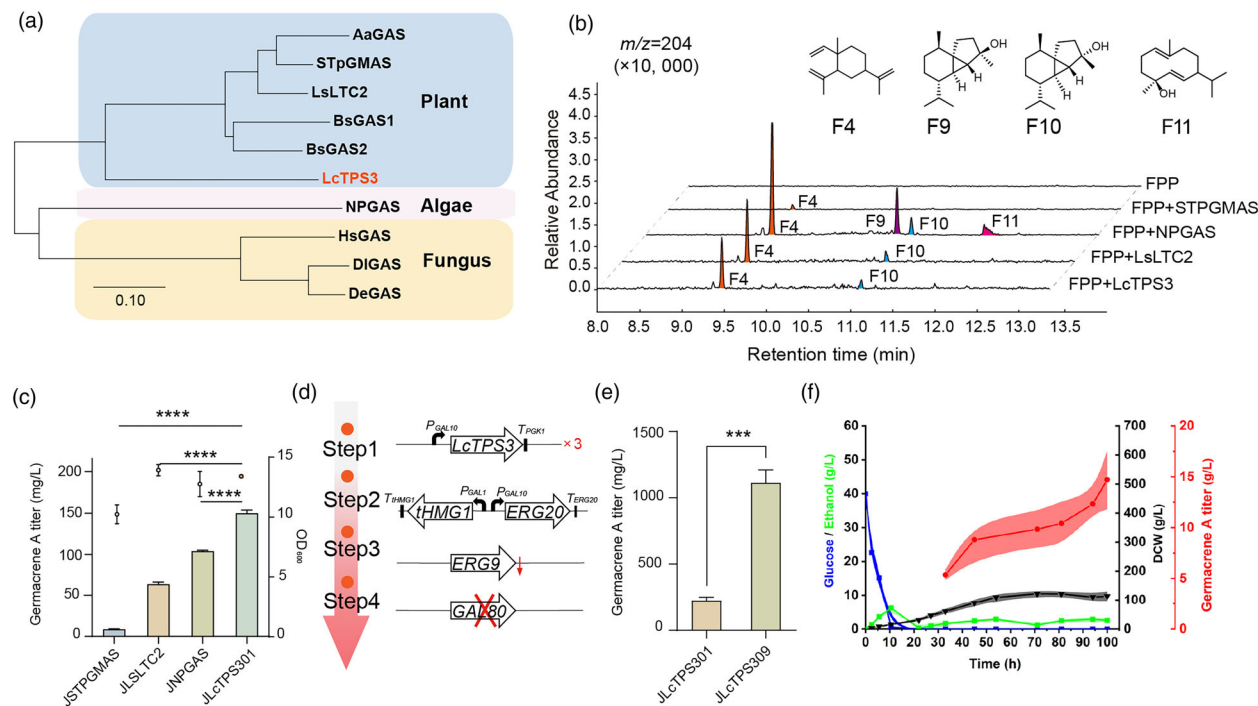

Figure 2 Metabolic engineering for high-yield germacrene A production in a yeast chassis. (a) Phylogenetic tree analysis of Germacrene A Synthase (GAS) sequences from various sources that have been reported. The GAS accession numbers are: AaGAS (DQ447636), STpGMAS (JF819848), LsLTC2 (AF489965), BsGAS1 (KM066976), BsGAS2 (KM066977), NPGAS (BAB76384), HsGAS (KZ111236), DIGAS (KZ112943), DeGAS (KAI1474031). (b) GC–MS-extracted ion chromatograms and chemical structures of sesquiterpene products. F4, β-elemene; F9, 4-*epi*-cubebol; F10, cubebol; F11, germacrene-D-4-ol. Products were identified by comparison to standards or the NIST17 library. (c) Germacrene A titer in shake flasks of strains that harbor various *GAS*. (d) Schematic depiction of the sequential engineering of mutants designated for overproduction of germacrene A within a yeast chassis. (e) Quantification of germacrene A production in engineered mutants. (f) Germacrene A titer in 5 L bioreactors. The time course of cell dry weight (CDW, black), germacrene A titer (red) and residual concentrations of glucose (blue) and ethanol (green) during fed-batch fermentation were recorded. Statistical differences were determined by ANOVA followed by the Tukey–Kramer test. $N = 3$. ***$P < 0.001$, ****$P < 0.0001$.
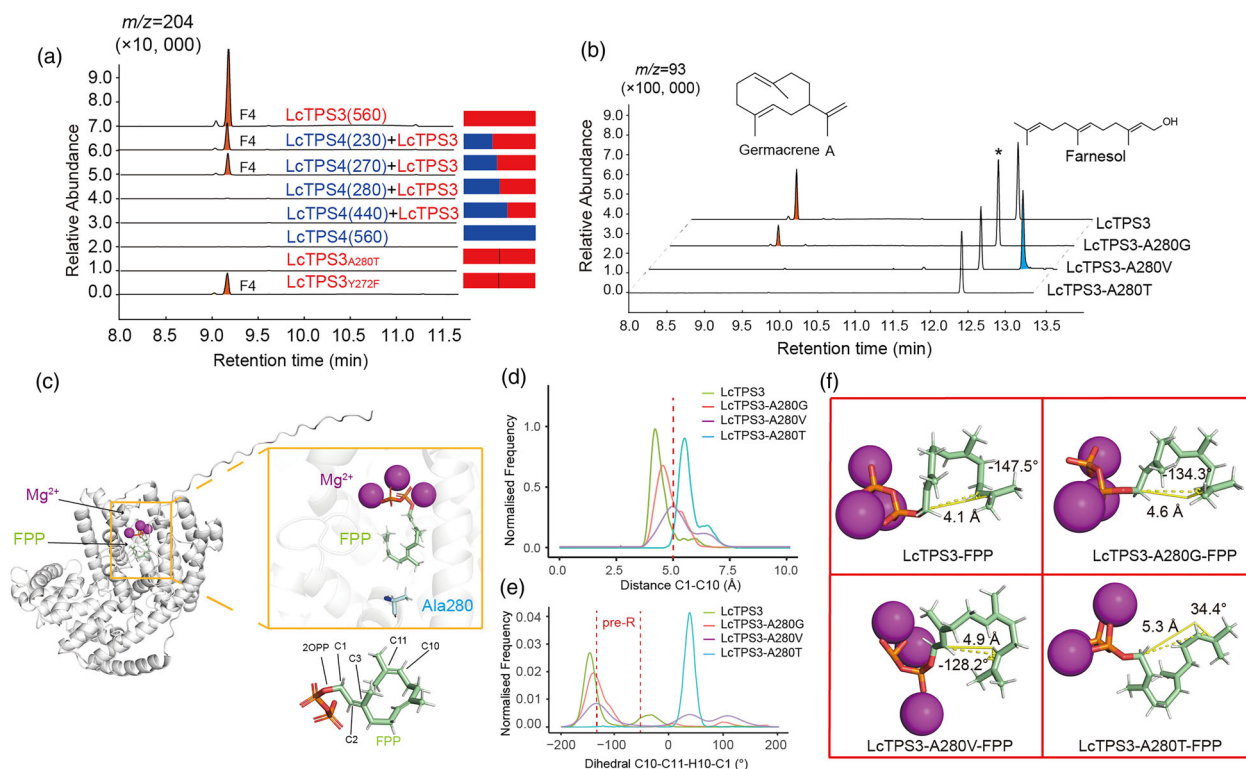
Figure 3 LcTPS3$_{A280}$ plays a critical role in germacrene A biosynthesis. (a) Schematic depiction and GC–MS-extracted ion chromatograms of the products of LcTPS3/LcTPS4 or mutants. (b) GC–MS-extracted ion chromatograms of the products of LcTPS3 and its mutants (A280G, A280V and A280T). Asterisk indicates non-specific products. (c) The protein structure and binding pocket of LcTPS3 predicted by alphafold2. The structure of LcTPS3 is represented as a cartoon in grey colour, FPP is shown as a coloured stick diagram, Mg^{2+} ions are shown as purple spheres, and atoms are coloured by element (carbon: green or blue; nitrogen: dark blue; oxygen: red; phosphorus: dark orange). (d) Histograms of the C1–C10 distance in LcTPS3 and its mutants (0.1 Å bin width). (e) Histograms of the dihedral angle C10–C11–H10–C1 of LcTPS3 and its mutants (5° bin width). (f) The binding mode observed most frequently in the MD simulations.

R-OB farnesyl starting position was the only one that achieved substantial sampling, meeting both the spatial and stereoselective requirements and was thus selected for further research (Figure S5).

Over a 100 ns simulation period of LcTPS3 in complex with the FPP substrate, the most frequently observed distance between the C1 and C10 atoms in the FPP substrate was 4.1 Å and the dihedral angle (C10–C11–H10–C1) peaked at −147.5°, suggesting the farnesyl chain could, in principle, cyclize into a germacryl cation upon ionization of FPP (Figure 3d–f). In LcTPS3 mutants complex with the substrate FPP, the distance and dihedral angle slightly increased to 4.6 Å and −134.3° in LcTPS3$_{A280G}$, aligning with the experimental results that the catalytic activity of LcTPS3$_{A280G}$ was slightly reduced. Additionally, in LcTPS3$_{A280V}$, the increase in distance to 4.9 Å hinders the formation of the germacryl cation and the dihedral angle shows a wide spread of conformations during the simulation, corresponding to its production of a linearized product. In the LcTPS3$_{A280T}$ complex with FPP, the distance continued to rise to 5.3 Å, and the dihedral angle was completely flipped to 34.4°, resulting in a complete loss of its catalytic activity (Figure 3d–f).

The results obtained from molecular dynamics (MD) simulations are consistent with the findings of the mutagenesis experiments, substantiating the significance of the A280 residue in the catalytic and cyclization processes. The LcTPS3–FPP complex, as modelled, was found to be a reasonable representation suitable for further investigation.

The functional reconstruction of LcTPS3 resulted in a secondary cyclization product α-guaiene

As germacrene A was considered a central intermediate in sesquiterpene biosynthesis (Xu and Dickschat, 2020), we next focused on exploring its functional plasticity to acquire more sesquiterpene products by redesigning its molecular choreography. Taking into account the amino acids surrounding the binding pockets of LcTPS3, along with the sequence difference among LcTPS3 and other GASs, we identified four amino acid variations, including A280I, Y282L, F304T and S308V (Figures 4a and S3). Site-directed mutagenesis showed that, while A280I, F304T and S308V led to the loss or reduction of the enzyme activity of LcTPS3, Y282L resulted in the production of a new product G2 (Figure 4b). The product G2 was then identified as α-guaiene, a secondary 1,5-cyclization product based on germacrenyl cation (Figures 4c and S6). Although Y282 is more than 10 Å away from the substrates in our docking model, additional mutations of Y282 into G, W and V also produce α-guaiene as well as a few by-products (Figure S7). This residue differs from many reports on the secondary catalytic mechanism of germacrenyl-type sesquiterpenes (Srivastava et al., 2021; Starks et al., 1997; Zhang et al., 2016).

The secondary cyclization mechanism reported for 5-epi-aristolochene synthase involves a catalytic triad (D444, Y520 and D525), in which Y520 protonates C5, creating a C1–C5 sigma bond that enables the second ring closure of germacrene A
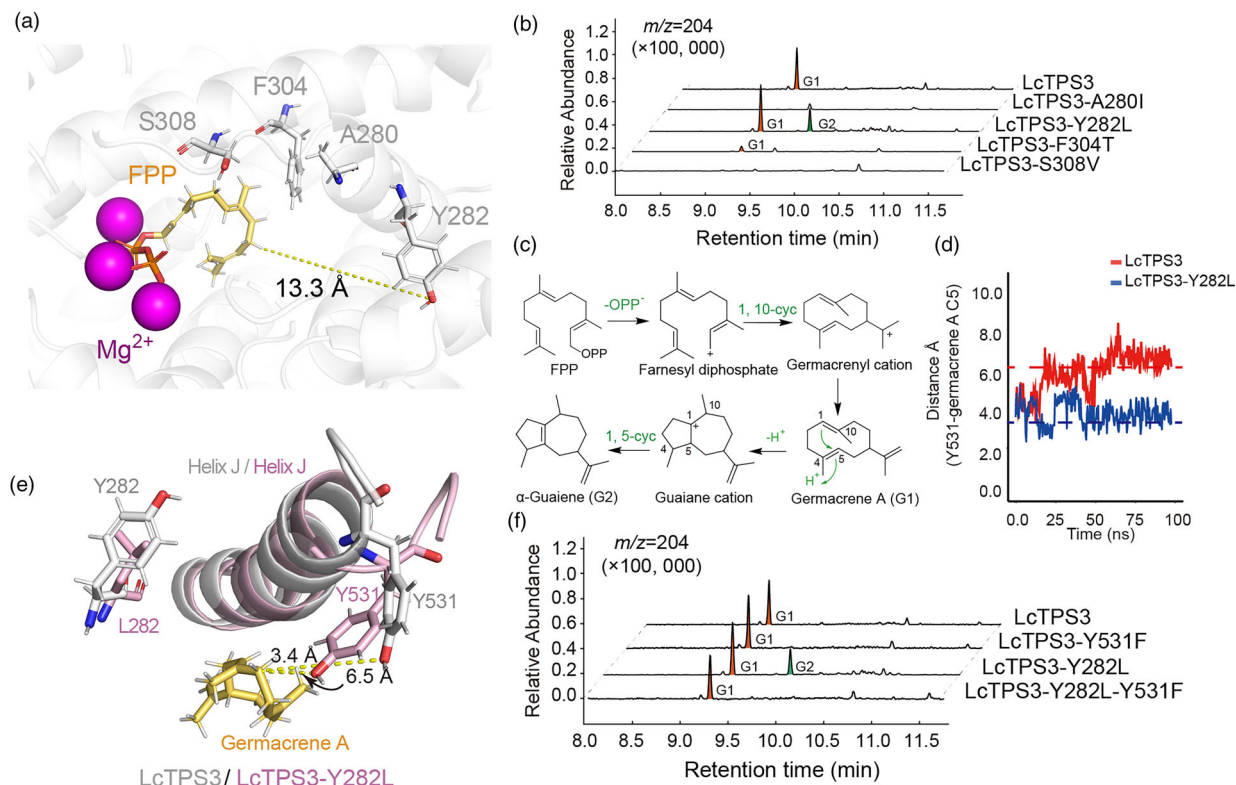
Figure 4 The functional reconstruction of LcTPS3 resulted in a secondary cyclization product α-guaiene. (a) Key amino acids around the binding pocket of LcTPS3. The structure of LcTPS3 is represented as a cartoon in grey colour, FPP is shown as a coloured stick diagram, Mg^{2+} ions are shown as purple spheres, and atoms are coloured by element (carbon: white or yellow; nitrogen: dark blue; oxygen: red; phosphorus: dark orange). The distance is represented by a yellow dashed line. (b) GC–MS-extracted ion chromatograms of products of LcTPS3 and its mutants (A280I, Y282L, F304T and S308V). G1, germacrene A; G2, α-guaiene, as determined by the standard (Figure S6). (c) Proposed biosynthetic pathway of α-guaiene. (d) Histograms of the distance between the Y531 residue and the C5 of germacrene A (0.1 Å bin width) in LcTPS3 or the $LcTPS3_{Y282L}$ mutant. (e) Distances are represented by yellow dashed lines from Y531 to the C5 of germacrene A in LcTPS3 (white) and $LcTPS3_{Y282L}$ (pink). (f) GC–MS-extracted ion chromatograms of products of LcTPS3 and LcTPS3 mutants (Y531F, Y282L and Y282L-Y531F). G1, germacrene A; G2, α-guaiene.

(Starks *et al.*, 1997). The catalytic triad is conserved in LcTPS3 (D455, Y531 and D535), even though LcTPS3 exclusively produces germacrene A (Figure S8). Based on the molecular model we generated above, we performed molecular docking and MD simulations on LcTPS3 and the $LcTPS3_{Y282L}$ mutants, with a focus on comparing the distance between Y531 and the C5 position of the germacrenyl cation. As a result, the distance between Y531 and the C5 of germacrene A has been decreased from 6.5 Å in LcTPS3 to 3.4 Å in the $LcTPS3_{Y282L}$ mutant (Figure 4d,e). This result suggests that the Y282L mutation may indirectly facilitate the production of the secondary cyclization product by shortening the distance between Y531 and the C5 position of germacrene A. To confirm the critical role of Y531 here, we created the $LcTPS3_{Y531F}$ mutant and the double mutant $LcTPS3_{Y282L/Y531F}$. While $LcTPS3_{Y531F}$ did not alter the ability of LcTPS3 to produce germacrene A, it led to an inability of $LcTPS3_{Y282L}$ to produce α-guaiene, consistent with our hypothesis (Figure 4f).

This discovery not only elucidates a novel mechanism of sesquiterpene secondary cyclization but also highlights the fact that LcTPS3 exhibits both high activity and significant plasticity. The potential to engineer the LcTPS3 enzyme may pave the way for the production of a greater variety of high-value sesquiterpenes.

Discussion

Plants possess highly complex metabolic networks, which render the composition of their metabolites particularly intricate. However, some high-value plant intermediate metabolites, such as germacrene A in *Liriodendron chinense*, cannot be directly extracted from plants. Although plants are a rich source of terpenoid skeletons, these compounds are generally present in low concentrations and exhibit spatiotemporal specificity in their distribution within plant tissues (Verma *et al.*, 2012; Wang *et al.*, 2016). The separation and identification of terpenoid skeletons necessitate a substantial plant source. In this study, we successfully biosynthesized germacrene A by employing heterologous expression of terpene synthases in a microbial host. This approach not only enhances the efficiency of terpenoid production but also provides novel strategies for further biosynthetic and metabolic engineering research.

Germacrene A undergoes the spontaneous Cope rearrangement under heat induction conditions to form β-elemene, which is recognized as a broad-spectrum antineoplastic drug (Chen *et al.*, 2012; Li *et al.*, 2010; Wang *et al.*, 2005). The biosynthesis of β-elemene via microbial chassis has been a hot topic of research over the last few years (Chen *et al.*, 2022; Hu *et al.*, 2017; Li *et al.*, 2022; Zhang *et al.*, 2021). Concurrently

with the preparation of this manuscript, one article reported that higher titers of β-elemene were achieved in the oleaginous yeast *Yarrowia lipolytica*, reaching up to 39 g/L (Liu *et al.*, 2023). However, the yield of β-elemene obtained in this study using *Saccharomyces cerevisiae* still surpasses the previously reported yield (4.7 g/L) in *Ogataea polymorpha* (Ye *et al.*, 2023). We can speculate that the expression of LcTPS3 in *Y. lipolytica* could have the potential to achieve even higher yields of β-elemene.

Although several crystal structures of terpene synthases have been revealed (Gennadios *et al.*, 2009; Köksal *et al.*, 2011; Starks *et al.*, 1997), the study of their catalytic mechanisms is still challenged due to the low sequence conservation and the complex cyclization patterns. Current research indicates that terpene synthases could catalyse a variety of complex cyclization patterns, hydrogen transfers and structural rearrangements, which could alter the form of carbocation intermediates such as the germacradienyl cation, resulting in a mixture of sesquiterpene carbocation intermediates with diverse structures (Gonzalez *et al.*, 2014; O'Brien *et al.*, 2016; Xu and Dickschat, 2020). The diversity and flexibility of this mechanism provide a wide range of chemical possibilities for the synthesis of sesquiterpenes (Chen *et al.*, 2020). However, there is scarce research on how the various structural features of the protein mediate the binding of sesquiterpene substrates and promote catalytic reactions. Protein structure prediction and molecular dynamics simulations offer alternative approaches that could significantly enhance our efficiency in predicting and designing the functionality of terpene synthases with available reference to crystal structures and catalytic mechanisms.

In addition to the *Liriodendron chinense* genome, we also investigated eight other plant genomes in the Magnoliaceae family, including *Liriodendron tulipifera* (Park *et al.*, 2019), *Magnolia biondii* (Dong *et al.*, 2021), *Magnolia kwangsiensis* (Kuang *et al.*, 2011), *Magnolia obovata* (Zhou *et al.*, 2023), *Magnolia officinalis* (Yin *et al.*, 2021), *Magnolia sieboldii* (Lu *et al.*, 2025), *Magnolia sinica* (Cai *et al.*, 2024) and *Magnolia × alba* (Hinsinger and Strijk, 2017). A total of 536 TPS genes were identified, and a phylogenetic tree was constructed (Figure S9). We examined the sequences of 38 genes in the same clade as LcTPS3, which included TPS genes from all nine species (Figure S10). The results showed that the TPS orthologs from the Magnoliaceae family only exhibit a certain degree of conservation of A280 and Y282. Magnoliaceae plants are widely recognized for their rich diversity of terpenes and terpenoids (Chen *et al.*, 2013; Dong *et al.*, 2009; Graziose *et al.*, 2011). Our analysis suggests that this richness may be attributed to the multiple copies and sequence diversity of TPS genes in the Magnoliaceae family.

As an intermediate scaffold for sesquiterpenes, germacrene A is synthesized with high catalytic efficiency by the characterized enzyme LcTPS3. Therefore, through artificial modification and design of the function of LcTPS3, its catalytic potential can be fully explored and applied. This not only helps to understand the catalytic mechanism but will further contribute to the biosynthesis of high-value sesquiterpenes. By functionally modifying LcTPS3, LcTPS3$_{Y282}$ mutants have been shown to catalyse a secondary cyclization of germacrenyl cation, resulting in the product of five to seven bicyclic skeleton of α-guaiene. α-Guaiene is a precursor to rotundone, which is the most prominent sesquiterpene identified in wine and the only known impact odorant with a peppery aroma (Lin *et al.*, 2019). The enzyme for synthesizing α-guaiene has previously been identified only in *Vitis vinifera* (Drew *et al.*, 2016) and *Stellera chamaejasme* (An *et al.*, 2020); here, we have also successfully obtained it through the resigning of LcTPS3. Thus, LcTPS3 demonstrates a significant potential for synthesizing complex sesquiterpenes, and there is potential for further in-depth exploration and functional design in the future.

Materials and methods

Chemicals and reagents

Tryptone and yeast extract were purchased from Coolaber (Beijing, China). NaCl, glucose, galactose, components of the trace metal solution and vitamin solution were purchased from Sinopharm Chemical Reagent Co., Ltd (Shanghai, China). YNB was purchased from Biotopped Co., Ltd. (Beijing, China). Amino acids, ampicillin and G418 were purchased from Shanghai yuanye Bio-Technology Co., Ltd (Shanghai, China). Germacrene A and farnesyl pyrophosphate ammonium salt were purchased from Sigma-Aldrich (St. Louis, MO, USA) for use as standards. Synthetic dropout media and YPD media were purchased from Coolaber (Beijing, China). All solvents in this study were analytical reagent grade from Sangon Biotech Co., Ltd. (Shanghai, China).

Phylogenetic and evolutionary analysis

The genomic sequence and annotation data were obtained from the BioProject PRJNA418360 (Chen *et al.*, 2019b). Transcriptomic data were sourced from the BioProject PRJNA780974 available in the NCBI database (Cheng *et al.*, 2022a). The RNA-seq reads were aligned to the reference genome using HISAT2 (v2.1.0) (Kim *et al.*, 2019). The expression levels were quantified in terms of fragments per kilobase of transcript per million mapped reads (FPKM) using StringTie (v2.0.6) (Pertea *et al.*, 2015). Putative TPS genes were identified by scanning predicted amino acid sequences against the Pfam database (v28.0) using HMMER, with an E-value cutoff of 1e-5. Sequences containing both the terpene_synth_C domain (PF03936) and the terpene_synth domain (PF01397) were classified as putative TPS genes. Phylogenetic analysis was performed with full-length amino acid sequences using MAFFT (v7.475) and FastTree (v2.1.11) with the maximum-likelihood method (Katoh and Standley, 2013; Price *et al.*, 2010). Bootstrap values were calculated with 1000 repetitions. The scale bar corresponds to 20% amino acid sequence divergence.

GC–MS analysis

The GC–MS analysis was performed as described previously (Ye *et al.*, 2022). Aliquots (1 μL) of the sample were injected into a Shimadzu Qp-2010 GC–MS system equipped with the Rxi-5Silcolumn (30 m × 0.25 mm × 0.2 μm film thickness; Rei-5Sil, PA, USA). Helium carrier gas with a column flow of 1 mL min^{-1} was used. The injector temperature was 250 °C and operated in splitless mode. The interface temperature was 280 °C. The detector was activated after a 3 min solvent delay. The oven temperature was programmed as follows: initial temperature, 50 °C for 3 min, 50–110 °C at a rate of 10 °C min^{-1}, 110–150 °C at a rate of 5 °C min^{-1} and hold for 3 min, 150–300 °C at a rate of 10 °C min^{-1} and hold for 3 min. The identification of compounds was assigned by comparison of their retention times and mass fragmentation with standards in our previously constructed terpene synthase-standard library (Zhi *et al.*, 2024) or the National Institute of Standards and Technology (NIST17) library.

Prokaryotic expression and purification of recombinant protein

Primers used in this study were synthesized by Auker Biological Engineering Co., Ltd. (Wuhan, China) and are listed in Table S4. Candidate genes were codon-optimized for *S. cerevisiae* and *E. coli*, taking codon usage bias into account and synthesized by GeneCreate Biological Engineering Co., Ltd. (Wuhan, China) (Chang *et al.*, 2024; Cheng *et al.*, 2022b; Zhi *et al.*, 2024). Polymerase chain reaction (PCR) amplification was performed using Phanta Max Super-Fidelity DNA Polymerase (Vazyme Biotech Co., Ltd., Nanjing, China) according to the manufacturer's instructions. PCR fragments were purified using the FastPure EndoFree Plasmid Mini Plus Kit (Vazyme Biotech Co., Ltd., Nanjing, China). The PCR product was digested and inserted into pET-28a (+) for heterologous expression in *Escherichia coli* strain BL21 (DE3) pLysS. Recombinant plasmid (1 μL) was added to 100 μL of chemically competent cells on ice and mixed gently. After 25 min on ice, the chemically competent cells were heated to 42 °C for 1 min and then quickly transferred to ice for 2 min. Then, 700 μL of LB medium was added and the cells were cultured at 37 °C and 200 rpm for 1 h. Plasmid-transformed *E. coli* strain BL21 (DE3) pLysS cells were inoculated in LB medium containing 50 μg/mL kanamycin, and the expression cultures were grown at 37 °C with shaking at 200 rpm. When the transformed cells were incubated to an optical density at OD_{600} of 0.6, a final concentration of 1 mm IPTG was added to induce protein expression, and the cultures were incubated at 16 °C and 150 rpm for 16 h. The cells were harvested at 6000 rpm for 10 min and then resuspended in 50 mm Tris–HCl (pH 7.5). Next, the suspension was sonicated at 150 W for 20 min (SCIENTZ-IID Sonicator; Scientz, China). The protein in the suspension was purified using Mag-Beads His-Tag (Sangon Biotech). Protein concentrations were determined using the BCA Protein Assay Kit (Solarbio Science). The purified protein was stored at −80 °C until use.

Enzymatic assays

GAS enzyme activity was determined by the malachite green assay, with some modifications. The assay was conducted in 96-well plates in a total reaction volume of 50 μL consisting of 50 mm MOPS buffer (pH 7.0), 5 mm $MgCl_2$, 5 mm KCl, 2.5 mU *Saccharomyces cerevisiae* phosphatase (Beyotime Biotech Co., Ltd., Shanghai, China), 50 μM FPP substrate and 5 μg/mL purified GAS enzyme. The mixture was incubated at 30 °C for 15 min and the reaction was terminated by adding 20 μL/well of the working reagent in the malachite green phosphate assay kit (Beyotime Biotech Co., Ltd., Shanghai, China). After incubation at 25 °C for 30 min, the absorbance was measured at 620 nm in a Multiskan GO microplate reader (Thermo Scientific, MA, USA). To establish the effects of pH on GAS enzyme activity, the pH was varied from 6.5 to 10 in the presence of 5 mm $MgCl_2$ and at 35 °C for 5 min. To test the effects of temperature on GAS enzyme activity, the temperature was varied from 20 to 45 °C in 50 mm HEPES buffer (pH 7.5) containing 5 mm $MgCl_2$. To determine the kinetic parameters, the GAS enzymes were assayed under optimal conditions using various concentrations of FPP (0 ∼ 25 μM) and 5 mm $MgCl_2$ in 100 μL of 50 mm MOPS buffer (pH 7.0) at 35 °C for 15 min. The K_m and V_{max} values were evaluated by fitting the initial velocity-versus-substrate concentration to the Michaelis–Menten kinetic equation using GRAPHPAD PRISM 5 software.

Plasmids and strains

To compare different germacrene A synthases, 80 ng of each coding sequence fragment was co-assembled with 80 ng of each respective regulatory cassette and 80 ng of the appropriate linearized expression vector pZY900 to obtain p900LcTPS01, p900LSLTC2, p900NPGAS and p900STPGMAS (Deng *et al.*, 2024). These plasmids were digested using restriction enzymes. The linearized products were loaded on an agarose gel to remove the backbone segments and purify the functional segments that were then integrated into the genome. The transformation of *S. cerevisiae* cells was performed using the LiAc/SS carrier DNA/PEG method (Siemon *et al.*, 2020) The confirmation of mutant strains was conducted through diagnostic PCR.

The *LcTPS3* gene was amplified from the p900LcTPS01 plasmid using gene-specific primers (Table S4). The amplified gene was then integrated into the plasmids pYR017 and pYR018, resulting in the plasmids p900LcTPS02 and p900LcTPS03. Various gene fragments were obtained by linearizing these plasmids. These include the *LcTPS3* gene, the *ERG20* gene, the homologous arm of *Leu* loci and the *Leu* selective gene from p900LcTPS01. Additionally, the fragment containing the *LcTPS3* gene, the *tHMG1* gene, the homologous arm of *Ura3* loci and the *His* selective gene was obtained from p900LcTPS02. Similarly, the fragment containing the *LcTPS3* gene, the homologous arm of the *YPRCdelta15* loci and the *Trp* selective gene was obtained from p900LcTPS03. A Crispr-Cas9 strategy was applied to knock down *ERG9*. *ERG9* repression was achieved by deleting a 45 bp (coordinates −220 to −175) in the UAS of the *Erg9* promoter (Chen *et al.*, 2019a). The homologous arm of the *GAL80* loci was ligated to the *Ura3* selective gene using overlap extension PCR to facilitate the knockout of *GAL80*. These fragments were subsequently utilized as templates for homologous recombination.

Fed-batch fermentation

To evaluate the yield level of germacrene A in fed-batch fermentation, The strain JLcTPS09 was fermented in 5 L bio-reactors (T&J Bio-engineering Co., LTD, Shanghai, China). Firstly, the glycerol stock was inoculated into YPD medium for activation, and then transferred to 800 mL LGM medium at a final concentration of 2% in 24 h. All seed cultures could be inoculated into a 5 L bioreactor in 16 h. The fermentation medium conditions were as follows: 40 g/L glucose, 8 g/L $KH_2PO_4 \bullet 7H_2O$, 6.2 g/L $MgSO_4 \bullet 7H_2O$ and 15 g/L $(NH_4)_2SO_4$. pH was adjusted to 5 using ammonia throughout the entire fermentation process, and the temperature was maintained at 30 °C. IPM was added at 24 h of fermentation to start the enrichment of the germacrene A. The fermentation process was based on previous studies (Huang *et al.*, 2023).

Molecular dynamics simulations

The AlphaFold2 (Tunyasuvunakool *et al.*, 2021) was used to predict the protein structure LcTPS3 models. The Mg^{2+} coordination shell (including PPi group and three Mg^{2+} ions) was rebuilt based on terpene synthases (PDB 5IK0) with a similar conserved coordination motif. The farnesyl chain was manually built in PyMOL (http://www.pymol.org/pymol), carefully avoiding clashes with the residues in the active site cavity according to the method reported before (O'Brien *et al.*, 2018). It was positioned in a cyclization-ready conformation (dC1-C10 < 5 Å). Molecular docking was performed using the software Discovery Studio (version 19.1.0). Finally, five conformations were obtained, and the structures were prepared in

the same simulation. The mutant models were constructed using the same methodology as described above.

The protein was simulated using the Amber ff99SB force field (Duan *et al.*, 2003), while water molecules were modelled using the TIP3P model (Jorgensen *et al.*, 1983). The ligand force field parameters were generated using the general AMBER force field (GAFF) (Wang *et al.*, 2004), and the partial atomic charge of the substrates was defined by the restrained electrostatic potential (RESP) (Bayly *et al.*, 1993) charge. Gaussian16_C.01 was used for RESP charge calculation, which was obtained from https://upjv.q4md-forcefieldtools.org/REDServer-Development/ (Vanquelef *et al.*, 2011). The tleap program in AMBER22 was used to generate the initial coordinates and topology files. The mutant models were constructed using the same methodology as described above.

The MD simulations were performed using the AMBER22 program, with each of the 8 FPP starting conformations being simulated independently. In summary, the preparation steps included hydrogen minimization, solvent molecule minimization and positional restraints of 5 kcal·mol^{-1}·Å^{-2} applied to C$^{\alpha}$ atoms and the Mg^{2+} ions. The positional restraints were retained during heating to 300 K (in 60 ps using NVT ensemble) and pressure equilibration (in NPT ensemble, at 300 K and 1 bar, for 100 ps). Subsequently, the positional restraints were gradually removed in a series of 50 ps NPT ensemble simulations. Following equilibration, trajectories were generated through 30 ns MD production simulations, with a target temperature of 300 K. During these simulations, the SHAKE algorithm (Ryckaert *et al.*, 1977) was employed to constrain the high-frequency stretching vibration of all bonds containing hydrogen atoms. The CPPTRAJ utility (from AmberTools23) was then used to calculate distance and dihedral angles. Figures were generated using the PyMOL Molecular Graphic System (version 2.5.2).

Author contributions

WC, YZ, FC, XX, HL, HZ, QW, XF, ZX and RL performed the experiments. WC, YZ, FC and LL conceived the project, designed the experiments and prepared the manuscript. ZD and TL discussed the results and contributed to the final manuscript. All authors read and approved the final manuscript.

Conflict of interest

The authors have applied for a patent based on this work.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. Sequences of LcTPSs in this study were provided in Table S1.

References

Adio, A.M., Paul, C., Tesso, H., Kloth, P. and König, W.A. (2004) Absolute configuration of helminthogermacrene. *Tetrahedron: Asymmetry.* **15**, 1631–1635.

An, T., Li, L., Lin, Y., Zeng, F., Lin, P. and Zi, J. (2020) Characterization of guaiene synthases from *Stellera chamaejasme* l. flowers and their application in de novo production of (-)-rotundone in yeast. *J. Agric. Food Chem.* **68**, 3214–3219.

Bayly, C.I., Cieplak, P., Cornell, W. and Kollman, P.A. (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **97**, 10269–10280.

Bertea, C.M., Voster, A., Verstappen, F.W., Maffei, M., Beekwilder, J. and Bouwmeester, H.J. (2006) Isoprenoid biosynthesis in *Artemisia annua*: cloning and heterologous expression of a germacrene A synthase from a glandular trichome cDNA library. *Arch. Biochem. Biophys.* **448**, 3–12.

Cai, L., Liu, D., Yang, F., Zhang, R., Yun, Q., Dao, Z., Ma, Y. *et al.* (2024) The chromosome-scale genome of *Magnolia sinica* (Magnoliaceae) provides insights into the conservation of plant species with extremely small populations (PSESP). *GigaScience* **13**, giad110.

Chang, X., Fang, X., Yao, Y., Xu, Z., Wu, C. and Lu, L. (2024) Identification and characterization of glycosyltransferases involved in the biosynthesis of neodiosmin. *J. Agric. Food Chem.* **72**, 4348–4357.

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, M., Zhang, J., Yu, S., Wang, S., Zhang, Z., Chen, J., Xiao, J. *et al.* (2012) Anti-cancer activity and liposome-based delivery systems of β-elemene. *Evid. Based Complement. Alternat. Med.* **2012**, 259523.

Chen, C.-Y., Wu, H.-M., Chen, H.-L., Chou, Y.-S., Juan, S.-W. and Huang, J.-C. (2013) Chemical constituents of the leaves of *Liriodendron chinense*. *Chem. Nat. Compd.* **49**, 775–776.

Chen, H., Zhu, C., Zhu, M., Xiong, J., Ma, H., Zhuo, M. and Li, S. (2019a) High production of valencene in Saccharomyces cerevisiae through metabolic engineering. *Microb. Cell Fact.* **18**, 1–14.

Chen, J., Hao, Z., Guang, X., Zhao, C., Wang, P., Xue, L., Zhu, Q. *et al.* (2019b) Liriodendron genome sheds light on angiosperm phylogeny and species-pair differentiation. *Nature Plants.* **5**, 18–25.

Chen, C.-C., Malwal, S.R., Han, X., Liu, W., Ma, L., Zhai, C., Dai, L. *et al.* (2020) Terpene cyclases and prenyltransferases: structures and mechanisms of action. *ACS Catal.* **11**, 290–303.

Chen, R., Liu, Y., Chen, S., Wang, M., Zhu, Y., Hu, T., Wei, Q. *et al.* (2022) Protein engineering of a germacrene A synthase from Lactuca sativa and its application in high productivity of germacrene A in *Escherichia coli*. *Front. Plant Sci.* **13**, 932966.

Cheng, W., Yao, Y., Wang, Q., Chang, X., Shi, Z., Fang, X., Chen, F. *et al.* (2022a) Characterization of benzylisoquinoline alkaloid methyltransferases in *Liriodendron chinense* provides insights into the phylogenic basis of angiosperm alkaloid diversity. *Plant J.* **112**, 535–548.

Christianson, D.W. (2017) Structural and chemical biology of terpenoid cyclases. *Chem. Rev.* **117**, 11570–11648.

Deng, X., Ye, Z., Duan, J., Chen, F., Zhi, Y., Huang, M., Huang, M. *et al.* (2024) Complete pathway elucidation and heterologous reconstitution of (+)-nootkatone biosynthesis from *Alpinia oxyphylla*. *New Phytol.* **241**, 779–792.

Dong, Y., Liang, D., Huang, J. and Zhang, P. (2009) Sesquiterpenes with quinone reductase-inducing activity from Liriodendron chinense. *Nat. Prod. Commun.* **4**, 1934578X0900400403.

Dong, S., Liu, M., Liu, Y., Chen, F., Yang, T., Chen, L., Zhang, X. *et al.* (2021) The genome of Magnolia biondii Pamp. provides insights into the evolution of Magnoliales and biosynthesis of terpenoids. *Hortic. Res.* **8**(1), 38.

Drew, D.P., Andersen, T.B., Sweetman, C., Møller, B.L., Ford, C. and Simonsen, H.T. (2016) Two key polymorphisms in a newly discovered allele of the Vitis vinifera TPS24 gene are responsible for the production of the rotundone precursor α-guaiene. *J. Exp. Bot.* **67**, 799–808.

Du, B., Sun, M., Hui, W., Xie, C. and Xu, X. (2023) Recent advances on key enzymes of microbial origin in the lycopene biosynthesis pathway. *J. Agric. Food Chem.* **71**, 12927–12942.

Duan, Y., Wu, C., Chowdhury, S., Lee, M.C., Xiong, G., Zhang, W., Yang, R. *et al.* (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **24**, 1999–2012.

Faraldos, J.A., Wu, S., Chappell, J. and Coates, R.M. (2007) Conformational analysis of (+)-germacrene A by variable-temperature NMR and NOE spectroscopy. *Tetrahedron* **63**, 7733–7742.

Fordjour, E., Liu, C.L., Hao, Y., Sackey, I., Yang, Y., Liu, X., Li, Y. *et al.* (2023) Engineering Escherichia coli BL21 (DE3) for high-yield production of germacrene A, a precursor of β-elemene via combinatorial metabolic engineering strategies. *Biotechnol. Bioeng.* **120**, 3039–3056.

Gennadios, H.A., Gonzalez, V., Di Costanzo, L., Li, A., Yu, F., Miller, D.J., Allemann, R.K. *et al.* (2009) Crystal structure of (+)-δ-cadinene synthase from Gossypium arboreum and evolutionary divergence of metal binding motifs for catalysis. *Biochemistry* **48**, 6175–6183.

Gonzalez, V., Touchet, S., Grundy, D.J., Faraldos, J.A. and Allemann, R.K. (2014) Evolutionary and mechanistic insights from the reconstruction of α-humulene synthases from a modern (+)-germacrene A synthase. *J. Am. Chem. Soc.* **136**, 14505–14512.

Graziose, R., Rathinasabapathy, T., Lategan, C., Poulev, A., Smith, P.J., Grace, M., Lila, M.A. *et al.* (2011) Antiplasmodial activity of aporphine alkaloids and sesquiterpene lactones from *Liriodendron tulipifera* L. *J. Ethnopharmacol.* **133**, 26–30.

Hinsinger, D.D. and Strijk, J.S. (2017) The chloroplast genome sequence of *Michelia alba* (Magnoliaceae), an ornamental tree species. *Mitochondrial DNA Part B.* **2**, 9–10.

Hu, Y., Zhou, Y.J., Bao, J., Huang, L., Nielsen, J. and Krivoruchko, A. (2017) Metabolic engineering of Saccharomyces cerevisiae for production of germacrene A, a precursor of beta-elemene. *J. Ind. Microbiol. Biotechnol.* **44**, 1065–1072.

Huang, Y., Ye, Z., Wan, X., Yao, G., Duan, J., Liu, J., Yao, M. *et al.* (2023) Systematic mining and evaluation of the sesquiterpene skeletons as high energy aviation fuel molecules. *Adv. Sci.* **10**, 2300889.

Jia, Q., Brown, R., Köllner, T.G., Fu, J., Chen, X., Wong, G.K.-S., Gershenzon, J. *et al.* (2022) Origin and early evolution of the plant terpene synthase family. *Proc. Natl. Acad. Sci.* **119**, e2100361119.

Jin, B., Xu, K., Guo, J., Ma, Y., Yang, J., Chen, N., Zeng, T. *et al.* (2024) From functional plasticity of two diterpene synthases (IrTPS2/IrKSL3a) to Enzyme Evolution. *ACS Catal.* **14**, 2959–2970.

Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W. and Klein, M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935.

Katoh, K. and Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780.

Kim, D., Paggi, J.M., Park, C., Bennett, C. and Salzberg, S.L. (2019) Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915.

Köksal, M., Jin, Y., Coates, R.M., Croteau, R. and Christianson, D.W. (2011) Taxadiene synthase structure and evolution of modular architecture in terpene biosynthesis. *Nature* **469**, 116–120.

de Kraker J.-W. Franssen M.C. de Groot A. König W.A. Bouwmeester H.J. 1998 (+)-Germacrene A biosynthesis: The committed step in the biosynthesis of bitter sesquiterpene lactones in chicory *Plant Physiol.* 117 1381 1392 4

Kuang, D.-Y., Wu, H., Wang, Y.-L., Gao, L.-M., Zhang, S.-Z. and Lu, L. (2011) Complete chloroplast genome sequence of Magnolia kwangsiensis (Magnoliaceae): implication for DNA barcoding and population genetics. *Genome* **54**, 663–673.

Li, Q.Q., Wang, G., Huang, F., Banda, M. and Reed, E. (2010) Antineoplastic effect of β-elemene on prostate cancer cells and other types of solid tumour cells. *J. Pharm. Pharmacol.* **62**, 1018–1027.

Li, M., Wen, Q., Lv, S., Yang, R., Cheng, T., Wang, Z. and Yang, J. (2022) Co-biosynthesis of germacrene A, a precursor of β-elemene, and lycopene in engineered Escherichia coli. *Appl. Microbiol. Biotechnol.* **106**, 8053–8066.

Li, Y., Qin, W., Liu, H., Chen, T., Yan, X., He, W., Peng, B. *et al.* (2023) Increased artemisinin production by promoting glandular secretory trichome formation and reconstructing the artemisinin biosynthetic pathway in Artemisia annua. *Hortic. Res.* **10**, uhad055.

Lin, J., Massonnet, M. and Cantu, D. (2019) The genetic basis of grape and wine aroma. *Hortic. Res.* **6**, 81.

Liu, Q., Zhang, G., Su, L., Liu, P., Jia, S., Wang, Q. and Dai, Z. (2023) Reprogramming the metabolism of oleaginous yeast for sustainably biosynthesizing the anticarcinogen precursor germacrene A. *Green Chem.* **25**, 7988–7997.

Lu, X., Mei, M., Liu, L., Xu, X. and Ai, W. (2025) The chromosome-scale genome of *Magnolia sieboldii* K. Koch provides insight into the evolutionary position of magnoliids and seed germination. *Mol. Ecol. Resour.* **25**, e14030.

Majdi M. Liu Q. Karimzadeh G. Malboobi M.A. Beekwilder J. Cankar K. de Vos R. Todorović S. Simonović A. Bouwmeester H. 2011 Biosynthesis and localization of parthenolide in glandular trichomes of feverfew (*Tanacetum parthenium* L. Schulz Bip.) *Phytochemistry* 72 1739 1750

Nguyen, T.-D., Faraldos, J.A., Vardakou, M., Salmon, M., O'Maille, P.E. and Ro, D.-K. (2016) Discovery of germacrene A synthases in Barnadesia spinosa: The first committed step in sesquiterpene lactone biosynthesis in the basal member of the Asteraceae. *Biochem. Biophys. Res. Commun.* **479**, 622–627.

O'Brien, T., Bertolani, S., Tantillo, D. and Siegel, J. (2016) Mechanistically informed predictions of binding modes for carbocation intermediates of a sesquiterpene synthase reaction. *Chem. Sci.* **7**, 4009–4015.

O'Brien, T.E., Bertolani, S.J., Zhang, Y., Siegel, J.B. and Tantillo, D.J. (2018) Predicting productive binding modes for substrates and carbocation intermediates in terpene synthases—bornyl diphosphate synthase as a representative case. *ACS Catal.* **8**, 3322–3330.

Park, J., Kim, Y. and Kwon, M. (2019) The complete mitochondrial genome of tulip tree, *Liriodendron tulipifera* L.(Magnoliaceae): intra-species variations on mitochondrial genome. *Mitochondrial DNA Part B.* **4**, 1308–1309.

Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.-C., Mendell, J.T. and Salzberg, S.L. (2015) StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295.

Price, M.N., Dehal, P.S. and Arkin, A.P. (2010) FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, e9490.

Ryckaert, J.-P., Ciccotti, G. and Berendsen, H.J. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341.

Shi, B., Ma, T., Ye, Z., Li, X., Huang, Y., Zhou, Z., Ding, Y. *et al.* (2019) Systematic metabolic engineering of Saccharomyces cerevisiae for lycopene overproduction. *J. Agric. Food Chem.* **67**, 11148–11157.

Siemon, T., Wang, Z., Bian, G., Seitz, T., Ye, Z., Lu, Y., Cheng, S. *et al.* (2020) Semisynthesis of plant-derived englerin A enabled by microbe engineering of guaia-6, 10 (14)-diene as building block. *J. Am. Chem. Soc.* **142**, 2760–2765.

Srivastava P.L. Escorcia A.M. Huynh F. Miller D.J. Allemann R.K. van der Kamp M.W. 2021 Redesigning the molecular choreography to prevent hydroxylation in germacradien-11-ol synthase catalysis *ACS Catal.* 11 1033 1041

Srivastava P.L. Johns S.T. Walters R. Miller D.J. Van der Kamp M.W. Allemann R.K. 2023 Active site loop engineering abolishes water capture in hydroxylating sesquiterpene synthases *ACS Catal.* 13 14199 14204

Starks, C.M., Back, K., Chappell, J. and Noel, J.P. (1997) Structural basis for cyclic terpene biosynthesis by tobacco 5-Epi-aristolochene synthase. *Science (New York, N.Y.)* **277**, 1815–1820.

Tahir, M.N., Shahbazi, F., Rondeau-Gagné, S. and Trant, J.F. (2021) The biosynthesis of the cannabinoids. *J. Cannabis Res.* **3**, 1–12.

Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Žídek, A., Bridgland, A. *et al.* (2021) Highly accurate protein structure prediction for the human proteome. *Nature* **596**, 590–596.

Vanquelef, E., Simon, S., Marquant, G., Garcia, E., Klimerak, G., Delepine, J.C., Cieplak, P. *et al.* (2011) RED Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments. *Nucleic Acids Res.* **39**, W511–W517.

Verma, P., Mathur, A.K., Srivastava, A. and Mathur, A. (2012) Emerging trends in research on spatial and temporal organization of terpenoid indole alkaloid pathway in Catharanthus roseus: a literature update. *Protoplasma* **249**, 255–268.

Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A. and Case, D.A. (2004) Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174.

Wang, G., Li, X., Huang, F., Zhao, J., Ding, H., Cunningham, C., Coad, J. *et al.* (2005) Antitumor effect of β-elemene in non-small-cell lung cancer cells is mediated via induction of cell cycle arrest and apoptotic cell death. *Cell Mol. Life Sci.* **62**, 881–893.

Wang, S., Tu, H., Wan, J., Chen, W., Liu, X., Luo, J., Xu, J. *et al*. (2016) Spatio-temporal distribution and natural variation of metabolites in citrus fruits. *Food Chem.* **199**, 8–17.

Wendt, K.U., Poralla, K. and Schulz, G.E. (1997) Structure and function of a squalene cyclase. *Science* **277**, 1811–1815.

Xie, W., Liu, M., Lv, X., Lu, W., Gu, J. and Yu, H. (2014) Construction of a controllable β-carotene biosynthetic pathway by decentralized assembly strategy in Saccharomyces cerevisiae. *Biotechnol. Bioeng.* **111**, 125–133.

Xu, H. and Dickschat, J.S. (2020) Germacrene A–A central intermediate in sesquiterpene biosynthesis. *Chemistry – A Eur. J.* **26**, 17318–17341.

Ye, Z., Huang, Y., Shi, B., Xiang, Z., Tian, Z., Huang, M., Wu, L. *et al*. (2022) Coupling cell growth and biochemical pathway induction in Saccharomyces cerevisiae for production of (+)-valencene and its chemical conversion to (+)-nootkatone. *Metab. Eng.* **72**, 107–115.

Ye, M., Gao, J. and Zhou, Y.J. (2023) Global metabolic rewiring of the nonconventional yeast *Ogataea polymorpha* for biosynthesis of the sesquiterpenoid β-elemene. *Metab. Eng.* **76**, 225–231.

Yin, Y., Peng, F., Zhou, L., Yin, X., Chen, J., Zhong, H., Hou, F. *et al*. (2021) The chromosome-scale genome of Magnolia officinalis provides insight into the evolutionary position of magnoliids. *iScience* **24**(9), 102997.

Zhang, F., Chen, N. and Wu, R. (2016) Molecular dynamics simulations elucidate conformational dynamics responsible for the cyclization reaction in TEAS. *J. Chem. Inf. Model.* **56**, 877–885.

Zhang, W., Guo, J., Wang, Z., Li, Y., Meng, X., Shen, Y. and Liu, W. (2021) Improved production of germacrene A, a direct precursor of ß-elemene, in engineered Saccharomyces cerevisiae by expressing a cyanobacterial germacrene A synthase. *Microb. Cell Fact.* **20**, 1–15.

Zhi, Y., Dai, C., Fang, X., Xiao, X., Lu, H., Chen, F., Chen, R. *et al*. (2024) Gene-directed in vitro mining uncovers the insect-repellent constituent from mugwort (*Artemisia argyi*). *J. Am. Chem. Soc.* **146**, 30883–30892.

Zhou, L., Hou, F., Wang, L., Zhang, L., Wang, Y., Yin, Y., Pei, J. *et al*. (2023) The genome of Magnolia hypoleuca provides a new insight into cold tolerance and the evolutionary position of magnoliids. *Front. Plant Sci.* **14**, 1108701.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1 Sequence information of *L. chinense* TPS candidate genes.
Table S2 Monoterpene and sesquiterpene products for LcTPSs.
Table S3 Strains and plasmids used in this study.
Table S4 Primers used to amplify *L. chinense* TPS candidate genes.
Figure S1 Prokaryotic expression and purification of LcTPS3.
Figure S2 Optimal reaction conditions for LcTPS3.
Figure S3 Multiple sequence alignment of LcTPS3 with other GASs.
Figure S4 Relative activity of LcTPS3 and its A280 mutants.
Figure S5 Structural models and molecular dynamics simulations of LcTPS3 complexed with FPP.
Figure S6 Structure determination of α-guaiene.
Figure S7 Product profiles of LcTPS3 and its Y282 mutants.
Figure S8 Superposition of the predicted LcTPS3 structure and the previously reported catalytic triad.
Figure S9 Phylogenetic analysis of the TPS gene family in Magnoliaceae plants.
Figure S10 Multiple sequence alignment of TPSs from the same clade as LcTPS3.