

Gene-Directed In Vitro Mining Uncovers the Insect-Repellent Constituent from Mugwort (*Artemisia argyi*)

Yao Zhi,[#] Chong Dai,[#] Xueting Fang, Xiaochun Xiao, Hui Lu, Fangfang Chen, Rong Chen, Weihua Ma, Zixin Deng, Li Lu,^{*} and Tiangang Liu^{*}



Cite This: *J. Am. Chem. Soc.* 2024, 146, 30883–30892



Read Online

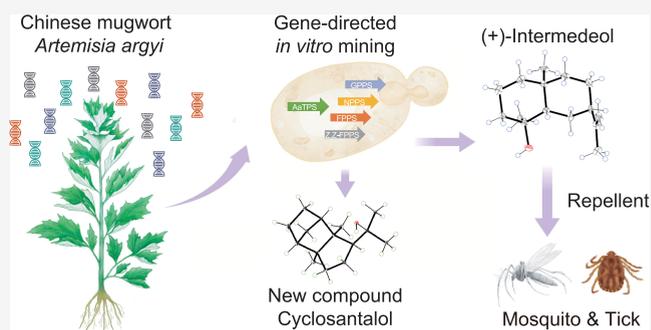
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: Plants contain a vast array of natural products yet to be discovered, particularly those minor bioactive constituents. Identification of these constituents requires a significant amount of plant material, presenting considerable technical challenges. Mugwort (*Artemisia argyi*) is a widely recognized insect repellent herb, particularly renowned for its extensive usage during the Dragon Boat Festival in China, but the specific constituent responsible for its repellent activity remains unknown. Here, we employed a gene-directed in vitro mining approach to characterize mugwort terpene synthases (TPSSs) systematically in a yeast expression system. Based on the establishment of “Terpene synthase-standard library”, we have successfully identified 54 terpene products, including a novel compound designated as cyclosantalol. Through activity screening, we have identified that (+)-intermedeol, which presents in trace amount in plants, exhibits significant repellent activity against mosquitoes and ticks. After establishing its safety and efficacy, we then achieved its biosynthetic production in a yeast chassis, with an initial yield of 2.34 g/L. The methodology employed in this study not only identified a highly effective, safe, and commercially viable insect repellent derived from mugwort but also holds promise for uncovering and producing other valuable plant natural products in future research endeavors.



INTRODUCTION

Plants harbor an untapped wealth of natural products, including a multitude of minor bioactive compounds that remain to be discovered.¹ Despite their low concentrations, these subtle elements can immensely benefit pharmaceuticals, agriculture, and are important drivers of interactions between plants and the environment.² However, isolating and identifying bioactive constituents from plants presents a formidable scientific challenge, particularly when targeting compounds that are volatile, unstable, or exist in trace quantities. An iconic example illustrating this challenge is artemisinin, an antimalarial compound derived from *Artemisia annua*. Youyou Tu's tenacity and pioneering discovery of artemisinin earned her the Nobel Prize.^{3,4}

Mugwort (*Artemisia argyi*), referred to as “Ai Hao” in Chinese, is a widely recognized traditional Chinese medicinal herb with longstanding of global usage.^{5,6} An ensuring tradition during the Dragon Boat Festival, celebrated on the fifth day of the fifth lunar month in China, entails hanging mugwort on doors or crafting sachets from its leaves. This cultural practice, which has been perpetuated for thousands of years and continues to be observed today, serves to ward off and protect against evil, disease, and bad luck.^{7,8} The rationale behind this practice could be related to pest control. As at this time of year,

coinciding with early summer, brings with it a surge in warmth and humidity, conditions conducive to insect proliferation. And the mugwort's robust, sage-like aroma could deter detrimental insects, thus potentially preventing epidemic outbreaks. Despite mugwort's significant role in traditional practices, there are surprisingly few insect repellent products on the market derived from it, as the specific bioactive components of mugwort, especially those effective against common pests like mosquitoes and ticks, remain largely unidentified.

Mosquitoes and ticks act as vectors of a wide range of infectious agents, contributing to an increasing burden worldwide with emerging diseases that affect both human and animal health. Ticks, in particular, have been associated with numerous fatalities in recent years, underscoring the urgent need for the development of new repellents.⁹ Recently, there has been heightened interest in researching plant-derived

Received: July 3, 2024

Revised: October 18, 2024

Accepted: October 21, 2024

Published: November 1, 2024



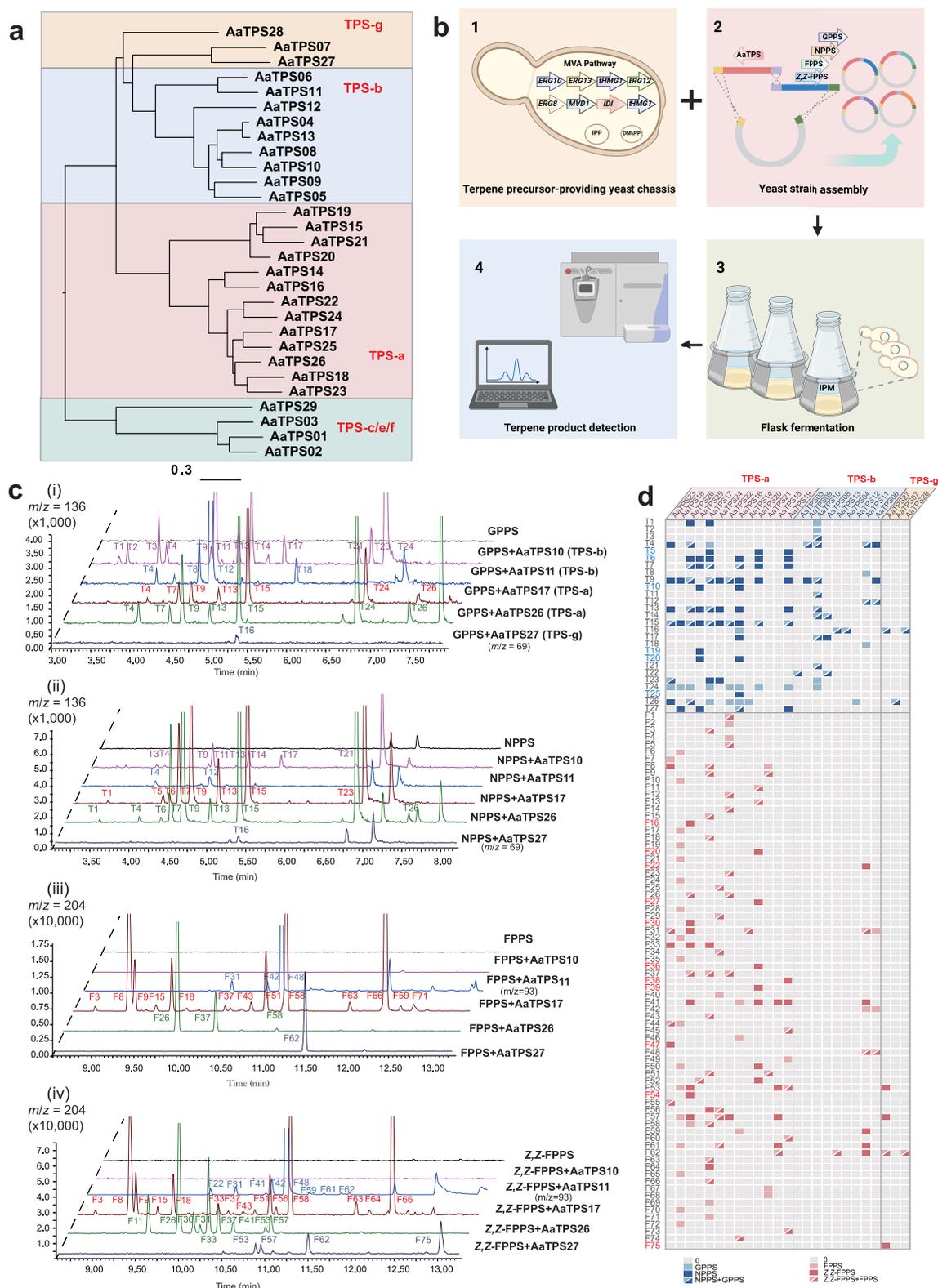


Figure 1. Systematic functional screening of AaTPSs in a yeast chassis. (a) Phylogenetic analysis showcasing the classification and relationship of terpene synthases (TPS) candidate genes from *A. argyi*. (b) Schematic depiction of the setup for the functional characterizations of AaTPSs coexpressed with GPPS, FPPS, NPPS and Z,Z-FPPS. (c) Gas-chromatography (GC) spectrum profile of products of AaTPS10, AaTPS11, AaTPS17, AaTPS26, and AaTPS27 coexpressed with GPPS (i), NPPS (ii), FPPS (iii) and Z,Z-FPPS (iv). (d) Tabulated products of AaTPS in yeast strains that coexpressed with GPPS, NPPS, FPPS, and Z,Z-FPPS. Blue and red squares represent the monoterpene and sesquiterpene products, respectively. Terpene products that are exclusively detected in AaTPS that coexpressed with NPPS or Z,Z-FPPS are highlighted in blue (monoterpene) and red (sesquiterpene) fonts.

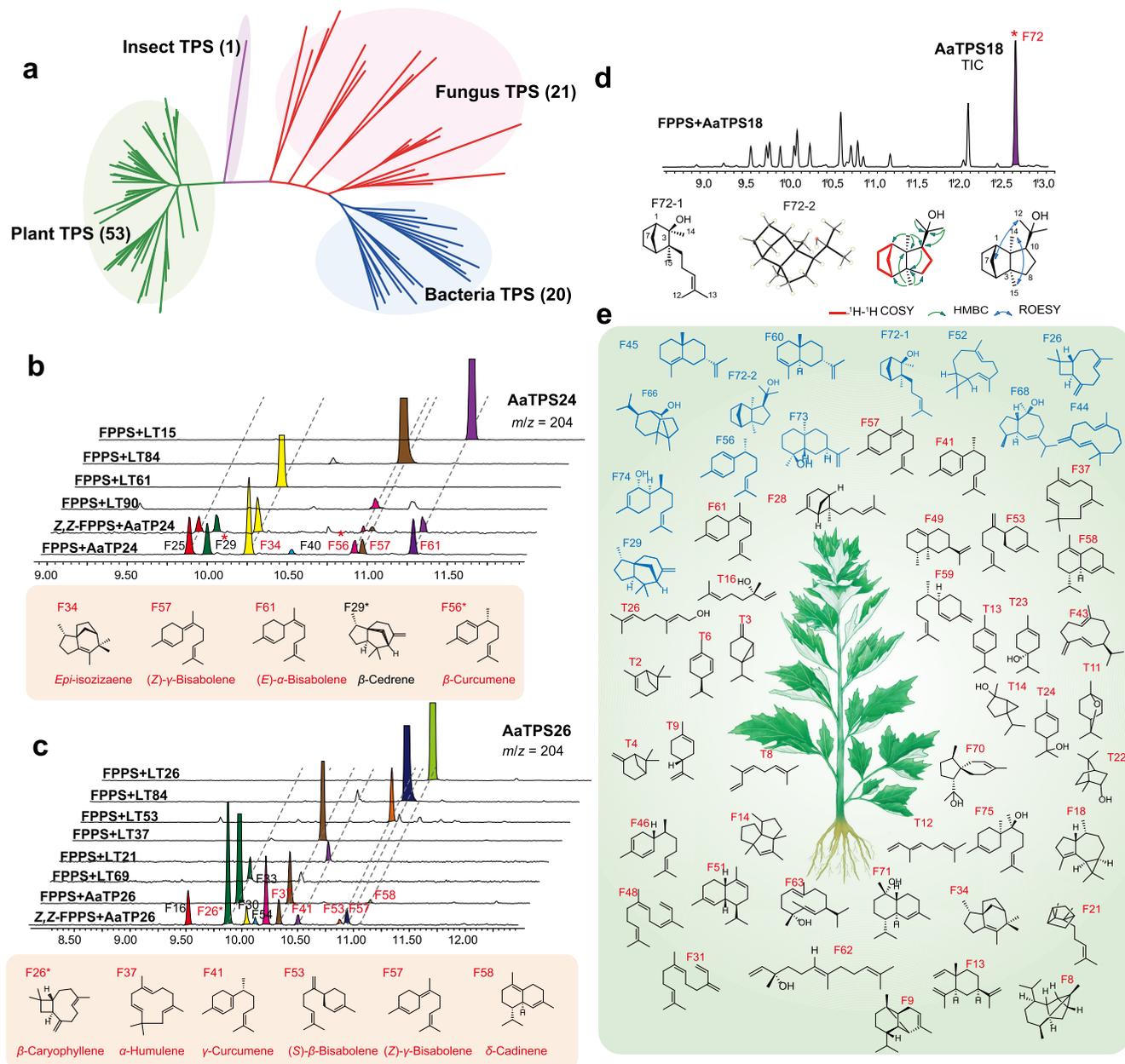


Figure 2. Terpene product determination by constructing a terpene synthase-standard library. (a) Phylogenetic analysis showed clustering of plant-derived (green), fungal-derived (red), bacterial-derived (blue), and insect-derived (purple) TPS, the brackets indicate the number of TPS genes. (b, c) Gas chromatography–mass spectrometry (GC–MS) profile ($m/z = 204$) of AaTPS24 (b) and AaTPS26 (c) products was aligned with the TPS products in the terpene synthase-standard library. The sesquiterpene structures determined by TPS-standard library (represented in red) or through nuclear magnetic resonance (NMR) (indicated by the asterisk) are presented below. (d) Total ion chromatography (TIC) of gas chromatography–mass spectrometry (GC–MS) profile of enzymatic products of AaTPS18, structure of F72–1, and X-ray Orthogonal Rank-Ordered Test Evaluation Plot (ORTEP) drawing, ^1H – ^1H COSY, HMBC, ROESY correlations of F72–2 (cyclosantalol). (e) Terpene products from yeast strains that coexpressing AaTPSs with GPPS, FPPS, NPPS and Z,Z-FPPS. Compounds were determined either via the terpene synthase-standard library (represented in black) or through NMR (represented in blue). Structure merely illustrates the relative configuration.

repellents, which are perceived as safer and more environmentally benign. Despite decades of efforts by natural product chemists to systematically isolate and characterize constituents of mugwort, with a focus on the bioactivity of its essential oil,^{8,10} recent advancements have primarily centered on sequencing its genome.^{6,11}

Terpenes constitute the largest class of volatile organic compounds emitted by plants, investigating the terpene constituents of mugwort could significantly enhance our understanding of its effectiveness in deterring mosquitoes.

But the conventional method of direct extraction and purification from plant tissues could potentially fail to capture the volatile and minor components present in the plant. Synthetic biology techniques and workflows offer significant advantages in overcoming the longstanding challenges associated with isolating and identifying terpene constituents.¹² We have previously systematically mined and characterized fungal terpene synthases (TPSs) to identify sesquiterpene products, bypassing the need to culture the fungal hosts under laboratory conditions, and novel nonsqualene-triterpene

synthases were also discovered.^{13,14} This methodology not only enables the discovery of novel compounds via the heterologous expression of TPS in microbial chassis but also provides a systematic means to enhance the production of specific chemicals, thereby laying the groundwork for subsequent bioactivity evaluations.

In this study, by using a gene-directed *in vitro* mining approach, we comprehensively examined the terpene composition of mugwort by structural elucidation of 54 terpene products. With activity screenings and safety evaluations, we have identified a superior insect-repellent constituent in mugwort that is highly effective, safe, and commercially viable.

RESULTS

Systematic Functional Screening of *Artemisia argyi* Terpene Synthase in a Yeast Chassis. An in-depth screening of the gene models from *Artemisia argyi* genome unveiled 127 protein sequences classified as AaTPS. Phylogenetic analysis further divided these AaTPS into four distinct subgroups: TPS-a, TPS-b, TPS-c/e/f, and TPS-g, based on previously established plant terpene synthase classification (Figure S1).¹⁵ As the *Artemisia argyi* genome has experienced several rounds of whole-genome duplication,¹⁶ there is a presence of highly similar TPS gene sequences (homologues). To ensure that our study included a comprehensive yet nonredundant set of TPS genes, we adopted a criterion of 90% amino acid sequence similarity to cluster these genes. From each cluster, we selected the TPS genes that showed higher expression levels based on the *Artemisia argyi* transcriptome data. Additionally, we filtered out any TPS genes that were not expressed in any of the tissues analyzed. Through this refined approach, we identified a total of 29 representative TPS genes that provide a broad coverage of the TPS gene family across the entire *Artemisia argyi* genome (Figure 1a). The 29 representative TPS genes were codon-optimized and synthesized for subsequent functional assessments. We previously engineered the yeast strain JCR27, which exhibits proficiency in synthesizing terpene precursors, particularly elevated levels of IPP (isopentenyl diphosphate) and DMAPP (dimethylallyl diphosphate), the fundamental C5 building blocks essential for terpene biosynthesis.^{17–19} These precursors, upon interaction with GPPS (geranyl diphosphate synthase) and FPPS (farnesyl diphosphate synthase), produce GPP and FPP, which function as substrates for the biosynthesis of monoterpenes and sesquiterpenes, respectively. Prior studies have demonstrated that plant TPS possess the capability to generate terpene products when supplied with NPP (nerolidyl diphosphate) or *Z,Z*-FPP (*Z,Z*-farnesyl diphosphate) as substrates.^{20–22} To fully exploit the catalytic potential of mugwort TPS, we constructed a series of yeast expression vectors by pairing each AaTPS (AaTPS1–29) with GPPS, FPPS, NPPS (nerolidyl diphosphate synthase), or *Z,Z*-FPPS (*Z,Z*-farnesyl diphosphate synthase). Harnessing the precision and efficiency of automated high-throughput biofoundry techniques, we generated 116 unique yeast strains (Figure S2). Subsequently, upon fermenting these engineered strains, the resulting compounds were analyzed using gas chromatography–mass spectrometry (GC-MS) (Figure 1b,c).

Terpene products of representative genes that coexpressed with GPPS, NPPS, FPPS, and *Z,Z*-FPPS were displayed in Figure 1c, while others are detailed in Figures S3 and S4. Among the examined AaTPS, AaTPS1, 2, 3, and 29, which were classified within the TPS-c/e/f group, showed no

response in producing any mono- or sesquiterpene products, but AaTPS4–28 yielded a total of 27 monoterpenes (T1–T27) and 75 sesquiterpenes (F1–F75) (Figure 1d). All nine AaTPSs categorized under the TPS-b group (AaTPS4–6, 8–13) synthesized monoterpene products when coexpressing GPPS or NPPS. While AaTPS10 generated a diverse mixture of 12 monoterpenes, others produced between one and four products. In contrast, AaTPSs from the TPS-a group (AaTPS14–26) exhibited significant substrate promiscuity. Out of the 13 TPSs analyzed, 11 were able to accept a variety of substrates, resulting in the production of both mono- and sesquiterpene products. Many of these enzymes generated multiple products; for instance, AaTPS17 synthesized up to 16 sesquiterpenes and 11 monoterpenes (Figure 1c). Based on their exclusive detection in yeast coexpressing AaTPS and NPPS or *Z,Z*-FPPS, monoterpenes T5, T6, T10, T19, T20, and T25, along with sesquiterpenes F16, F20, F22, F27, F30, F36, F38, F39, F47, F54, and F75 were identified (Figure 1d). Notably, upon comparing the terpene products with extracts from mugwort, we observed that 62 terpene products, including all the aforementioned products, were absent in the terpene extract derived from the leaf, stem, or root samples of the mugwort plant (Figures 1d, S5–S7).

Structure Determination of Terpene Products and Discovery of a Novel Compound Cyclosantalol. The elucidation of a compound's structure is pivotal for understanding its biological functionality. While isolating and purifying each terpene followed by structural elucidation using nuclear magnetic resonance (NMR) offers precision, it is labor-intensive, especially for minor terpene constituents found in trace amounts. In our effort to discern the structure of AaTPS terpene products, we built a local terpene synthase-standard library. This involved heterologously expressing TPSs that had been functionally characterized within our yeast chassis designated for terpene precursor provision (Figures 2a, S8, Table S1). Overall, the standard library confirmed 21 monoterpene and 106 sesquiterpene products by cross-referencing with the literature (Figure S9). Among these terpene products, 54 (42.5%) were validated through NMR, while 47 (37.0%) were authenticated against certified standards (Table S1, Figures S10 and S11). These verified terpenes were subsequently designated as benchmark standards for identifying AaTPS products.

Expression of AaTPS24 in our yeast chassis resulted in seven sesquiterpene products (Figure 2b). Notably, the main product, F34, displayed retention time and mass matching those of the sesquiterpene product from LT61, previously confirmed as *epi*-isozizaene through NMR analysis,²³ therefore, F34 was identified as *epi*-isozizaene (Figure S12). Similarly, products F56, F57, and F61 aligned in retention time and mass with those of LT90, LT84, and LT15, respectively, and were identified as β -curcumene, (*Z*)- γ -bisabolene, and (*E*)- α -bisabolene^{24–26} (Figure S12). Furthermore, F29 and F56 were isolated and purified, with their structures determined as β -cedrene and β -curcumene, respectively, through NMR analysis (Figure 2b, Tables S2–S5). Expression of AaTPS26 in our yeast chassis yielded ten distinct sesquiterpene products (Figure 2c). Among these, F26, F37, F41, F53, F57, and F58 exhibited identical retention times and masses as products derived from LT69, LT37, LT21, LT53, LT84, and LT26, respectively (Figure S13). Based on previous structural determinations of these gene products, they were identified as β -caryophyllene, α -humulene, γ -curcumene, (*S*)- β -bisabo-

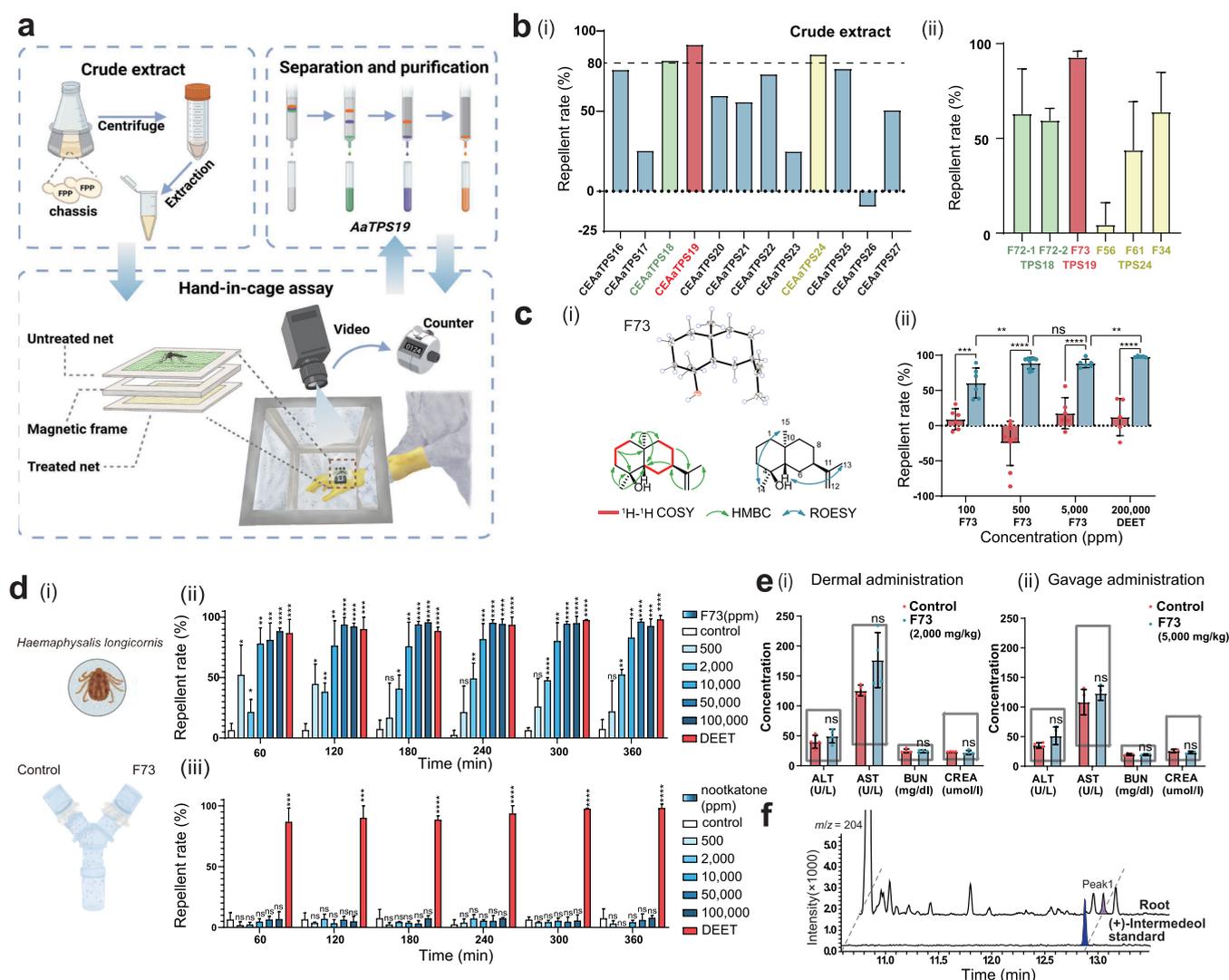


Figure 3. Identification of (+)-intermedeol as a novel insect repellent. (a) Schematic depiction of the setup for the hand-in-cage assay employed for initial repellent activity screening of yeast crude extracts and the primary metabolite of AaTPS19. (b) Initial activity screening of yeast crude extract of AaTPS16–27 that coexpresses FPPS. (i). Activity screening of the sesquiterpene products of AaTPS18, AaTPS19, and AaTPS24 against mosquitoes at the concentrations of 500 ppm ($N = 3$) (ii). (c) X-ray Orthogonal Rank-Ordered Test Evaluation Plot (ORTEP) drawing, ^1H – ^1H COSY, HMBC, ROESY correlations of (+)-intermedeol (i). Assessment of the repellent rate of F73 against mosquitoes at various dosages. DEET (200,000 ppm) was used as a positive control (ii). (d) Schematic depiction of a Y-tube olfactometer setup for dual-choice assay of ticks nymphs (*Haemaphysalis longicornis*; $N = 3$) (i). Assessment of the repellent rate of F73 and nootkatone against ticks nymphs at various dosages. DEET (20%) was used as a positive control ($N = 3$) (ii). (e) Biochemical blood analysis of mice under dermal administration (i) and gavage administration (ii). Concentrations of alanine aminotransferase (ALT), aspartate aminotransferase (AST), blood urea (BUN), and creatinine (CREA) were compared ($N = 4$). Gray box indicates the normal range. (f) GC-MS analysis for mugwort root extracts and the standard of (+)-intermedeol. For the bar plot, each dot indicates one replicate. Bars and error bars indicate means \pm s.d. Student's t test was employed, ns, not significant; *, $P < 0.1$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$.

lene, (Z)- γ -bisabolene, and δ -cadinene, respectively (Figure 2c, Tables S2–S5).^{18,19,25,27–30}

An unknown compound designated as F72 was identified as the predominant product of AaTPS18 (Figure 2d). Intriguingly, our analytical investigations revealed that F72 is not a singular entity but rather a complex mixture of two distinct compounds (Figure S14). Notably, these two constituents exhibit such similar properties that they have proven refractory to separation by conventional high-performance liquid chromatography (HPLC) and gas chromatography–mass spectrometry (GCMS) techniques. Through successive rounds of preparative thin-layer chromatography (TLC), we achieved the isolation and purification of two distinct compounds,

designated as F72–1 and F72–2. F72–1 was determined as compound *rel*-(1*R*,2*S*,3*R*,4*S*)-2,3-Dimethyl-3-(4-methyl-3-penten-1-yl) bicyclo[2.2.1]heptan-2-ol by analysis of NMR spectroscopy³¹ (Figure 2d). F72–2 has a molecular formula of $\text{C}_{15}\text{H}_{26}\text{O}$ as revealed by the high-resolution electrospray ionization mass spectrometry (HRESIMS), the ^1H and ^{13}C NMR data resembled those of epicyclosantalol,³² the HMBC correlations from Me-12 and Me-13 to C-10 and C-11 indicated that the side chain at C-10 is replaced by isopropanol (Figures S15–S18). Furthermore, a single-crystal X-ray diffraction analysis was performed for F72–2 by using Cu $K\alpha$ radiation, which not only confirmed the undescribed planar structure of F72–2, but also determined the absolute

configuration of F72–2 to be 1S,2S,3S,4R,10S (Figure 2d). Through a literature review, we identified F72–2 as a previously undescribed compound, which we have designated as cyclosantalol. The biosynthetic mechanism for cyclosantalol was proposed through a 1,6-cyclization of nerolidyl pyrophosphate, the isomer of FPP (Figure S19). Notably, neither F72–1 nor F72–2 have been previously reported as constituents of the mugwort plant.

The same analysis has been conducted for the other AaTPS products (Figure S13). In summary, utilizing our local terpene synthase-standard library, we confidently identified a total of 15 monoterpene and 31 sesquiterpene products derived from AaTPSs (Table S2). Additionally, we isolated and purified ten sesquiterpenes, including F74, F66, F45, F60, F52, F29, F56, F44, F26, and F68. Their structures were determined through NMR analysis as (1R,6R,7S)-2,10-bisaboladien-1-ol, copaborneol, 7-*epi*-selina-4,11-diene, 7-*epi*- α -selinene, (+)-bicyclogermacrene, β -cedrene, β -curcumene, γ -humulene, β -caryophyllene, and (1S,4S,5R,10S)-guai-6-en-10-ol, respectively (Figure 2e, Tables S3–S5).

Identification of (+)-Intermedeol as an Insect-Repellent Constituent. To investigate the potential insect-repellent properties of the metabolites synthesized by AaTPS, we initially screened repellent activity using yeast crude extracts derived in strains that express AaTPSs. Following fermentation, these yeast crude extracts were extracted and adjusted to a concentration of 10,000 ppm using acetone as the solvent. Their repellent activity against mosquitoes (*Aedes aegypti*) was then assessed using a hand-in-cage experiment (Figure 3a). Of the tested cell extracts, the cell extracts of AaTPS18, AaTPS19, and AaTPS24 that coexpressing with FPPS exhibited robust repellent activities, deterring mosquitoes at a rate of more than 80% (Figure 3b). The products of AaTPS18, identified as F72–1 and F72–2, were determined as described above. The AaTPS19 predominantly synthesized compound F73, which was identified as intermedeol corroborated by our terpene synthase-standard library (Figure S20). Meanwhile, F34, F56, and F61, the products of AaTPS24, were determined as *epi*-isozizaene, β -curcumene and α -bisabolene (Figure S20, Tables S3–S5). Subsequently, we conducted large-scale fermentation to extract and isolate these six compounds, their mosquito-repellent activities were further assessed at a concentration of 500 ppm. The results indicated that while F56 exhibits a low repellent rate, F72–1, F72–2, and F34 showed repellent activities above 50%, compound F73 exhibited the most remarkable repellent activity at 93% (Figure 3b). Our subsequent research has been focused on the compound F73.

F73 was extracted and isolated as colorless oil, its molecular formula was determined to be C₁₅H₂₆O based on HRESIMS data. Detailed comparison of ¹H NMR and ¹³C NMR spectra of F73 with those of (+)-intermedeol revealed that they share the same planar structure.³³ Additionally, a well-formed crystal of F73 was successfully grown by slow evaporation from a solution of petroleum ether. Crystallographic analysis confirmed the structure of F73 as (+)-intermedeol, marking the first report of the single-crystal structure for this compound (Figures 3c, S21–S24). We further evaluated its repellent properties by subjecting (+)-intermedeol in various dosages, and 20% DEET (*N,N*-Diethyl-*meta*-toluamide), a concentration considered highly effective for widespread use as insect repellent, was included as positive control.³⁴ At the concentration of 100 ppm, (+)-intermedeol has a repellent

rate of 60.5%. This rate notably rises to 88.7% when the concentration was increased to 500 ppm (Figure 3c, Table S6). However, no further improvement in repellency when the concentration is increased to 5000 ppm, with the rate remaining at 88.4%. This suggests that the optimal repellent effect of (+)-intermedeol is achieved at a concentration of 500 ppm (Figure 3c, Table S6). Although the 0.05% (500 ppm) concentration of (+)-intermedeol has a slightly lower repellent rate than 20% DEET, it exhibits high activity at this low concentration, offering better safety and environmental benefits due to its reduced dosage. Through a preliminary transcriptome study of mosquitoes, we identified 101 differentially expressed genes (DEGs) when mosquitoes were treated by (+)-intermedeol. Notably, odorant binding proteins (OBP55 and OBP27) exhibited significant down-regulation subject to (+)-intermedeol exposure, suggesting their potential involvement in chemosensory response to (+)-intermedeol (Figure S25).

Ticks are able to transmit various pathogen species and represent a growing threat to public health and agricultural systems worldwide. We extended our investigation of (+)-intermedeol's effectiveness in repelling ticks by comparing it with nootkatone, which is the active ingredient in the commercially available tick repellent, NootkaShield, and 20% DEET was also included as a positive control. Using dual-choice tests with nymphs of the Asian longhorned tick (*Haemaphysalis longicornis*), we found that (+)-intermedeol exhibited a dose-dependent repellency. Specifically, repellent rates of 81.66 and 95.20% were observed at concentrations of 10,000 ppm and 50,000 ppm, respectively. And the repellent effect of (+)-intermedeol persisted for at least 6 h at various concentrations (Figure 3d). In contrast, nootkatone showed no repellent effect on this tick species (Figures 3d, S26, Table S7). Similar results were observed in tests involving the brown planthopper (*Nilaparvata lugens*) and *Adelphocoris suturalis*, primary pests affecting rice (*Oryza sativa*) and cotton (*Gossypium hirsutum*), respectively.^{35,36} Only 16.79% of the brown planthopper nymphs and 18.67% of *A. suturalis* gravitated toward plants treated with 500 ppm (+)-intermedeol, significantly less than the solvent control (72.87 and 69.33%; *P* < 0.0001) (Figure S27, Table S6). Thus, (+)-intermedeol exhibits substantial repellency against mosquitoes, ticks, and agricultural pests.

To assess its prospective application, we conducted a biosafety evaluation of (+)-intermedeol. Using the highest recommended dosage according to GB 15670–1995 guidelines, no toxic effects were observed in mice following acute dermal and oral treatments (Figure S28). Hematological and serum biochemical parameters of (+)-intermedeol-treated groups were similar to the untreated group, and histopathological examination of vital organs showed no major changes (Figures 3e, S29). These findings confirm the nontoxic nature of (+)-intermedeol, supporting its potential as a safe and effective insect repellent.

Align with our findings, (+)-intermedeol, previously isolated and identified from leaves of American (*Callicarpa americana*) and Japanese (*Callicarpa japonica*) beautyberry (Verbenaceae), has also been demonstrated to possess mosquito repellent properties.³⁷ Interestingly, as described by the author, the residents in northeast Mississippi have traditionally used fresh crushed leaves of beautyberry as a topical treatment for draft animals to repel flies and other biting insects for several decades. This practice strikingly parallels the traditional use of

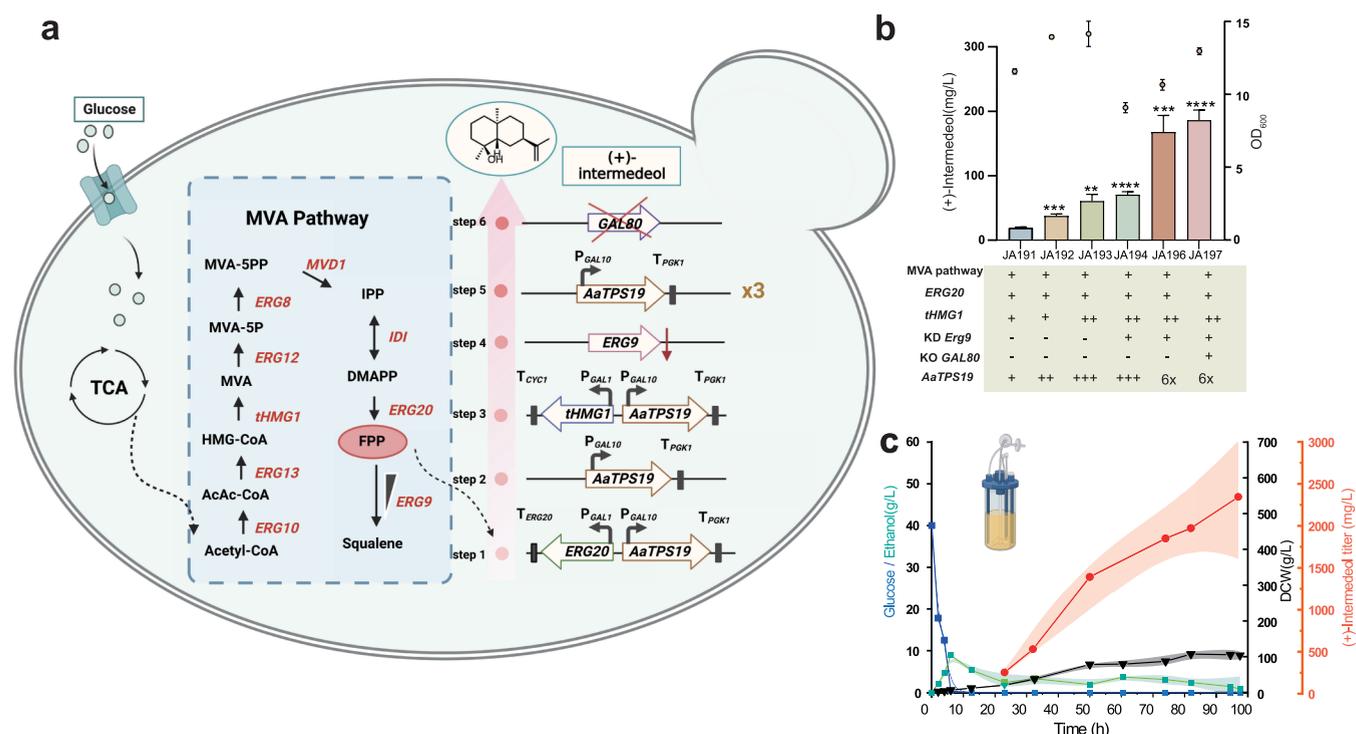


Figure 4. Titer of (+)-intermedeol in iterative mutant yeast strains cultivated in shake flask conditions and fed-batch fermentation. (a) Schematic depiction of the sequential engineering of mutants designated for overproduction of (+)-intermedeol within a yeast chassis featuring an overexpressed mevalonate (MVA) pathway. (b) Quantification of (+)-intermedeol production in engineered mutants. Bars and error bars indicate means \pm s.d. of three replicates. Student's *t*-test was employed to compare the titers of mutant strains with the initial (+)-intermedeol-producing mutant JA191 (** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$). Cell density was measured as the optical density at 600 nm (OD_{600}). Dots with bars represent means \pm s.d. of three replicates. Each “+” indicates one gene copy integrating into the yeast chromosomes. (c) Fed-batch fermentation result of JA197 in a 5 L parallel bioreactor. Changes in glucose (dark blue, g/L), ethanol (dark green, g/L), dry cell weight (DCW, black, g/L), and (+)-intermedeol titer (red, mg/L) were shown in a time-course manner, s.d. were shown as shaded areas ($N = 3$).

mugwort in China. However, despite (+)-intermedeol being recognized as a mosquito repellent as early as 2005, there has been a paucity of subsequent research in this area. The compound's scarcity in plants may have impeded comprehensive studies on its efficacy, as well as the ability to conduct thorough safety assessments, and potentially hindered its broader application and commercialization.

(+)-Intermedeol has been reported as a constituent in the essential oils of certain plants in Asteraceae, Verbenaceae, or Lamiaceae,^{38,39} but to the best of our knowledge, the presence of (+)-intermedeol within the chemical constituents of mugwort has never been documented. We found (+)-intermedeol exists in notably minute quantities in mugwort roots (273.6 ng g⁻¹ of fresh weight) (Figures 3f and S30). This aligns with our findings demonstrating its potent repellent properties even at a minimal concentration of 500 ppm. However, these trace amounts pose significant challenges for isolation and utility using conventional phytochemical methodologies, thus the gene-directed in vitro mining approach we employed here showed significant advances. Additionally, we synthesized (+)-intermedeol by transiently expressing *AaTPS19* in *Nicotiana benthamiana* leaves (Figure S31), indicating the feasibility of biosynthesizing (+)-intermedeol by *AaTPS19* in planta.

Overproduction of (+)-Intermedeol by Metabolic Engineering in Yeast Chassis. To enable large-scale production of (+)-intermedeol, essential for its application as an insect repellent, we employed metabolic engineering in yeast chassis (Figure 4a). Beginning with the JCR27 strain, we

integrated *AaTPS19* and *ERG20* (*FPPS*) into the *leu2* locus. Additional modifications included incorporating extra copies of *AaTPS19* along with the *tHMG1* gene, a key enzyme in the MVA pathway. To alleviate competition for *FPP* substrates triggered by *ERG9*, we attenuated the Upstream Activating Sequences (UAS) of the *ERG9* promoter. Subsequently, utilizing Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) technology, we introduced three additional copies of *AaTPS19* into the yeast chromosome, resulting in a titer increase up to 167.98 mg L⁻¹ in the JA196 mutant strain, approximately 8.85-fold higher than the original JA191 strain cultivated in shake-flask conditions. Furthermore, we identified a mutant strain with disrupted *Gal80* gene, allowing expression induction by glucose instead of galactose, thus reducing production costs. These metabolic engineering strategies collectively elevated (+)-intermedeol titer to 186.16 mg L⁻¹ (Figure 4b, Table S8), indicating its commercial viability.

To evaluate industrial viability, we subjected the JA197 strain to fed-batch fermentation in a 5 L bioreactor (Figure 4c). The fermentation process comprised two phases: growth and product accumulation. The growth phase initiated with a glucose concentration of 40 g L⁻¹, gradually reduced to approximately 1 g L⁻¹ to optimize strain proliferation. Subsequently, the product accumulation phase began, overlaying isopropyl myristate (IPM) onto the medium at the 24-h mark, while managing ethanol concentrations within the range of 2–5 g L⁻¹ to enhance gene expression. The fermentation process concluded when no further increments in product accumulation were observed. Remarkably, this approach

elevated fermentation titers to 2.34 g L⁻¹, demonstrating significant potential for commercial-scale production.

DISCUSSION

Here, (+)-intermedeol demonstrates superior repellent activity against mosquitoes at an exceptionally low concentration of 0.05%, outperforming other commercial repellents.^{40–42} Notably, (+)-intermedeol has demonstrated exceptional repellent activity against ticks, which have been associated with a significant number of fatalities in recent years. Given the scarcity of commercial tick repellent products, this finding underscores the pressing need for the development of novel and effective tick repellents. Additionally, (+)-intermedeol is biological safety and odorless, making it highly suitable for commercial applications. Through a series of metabolic engineering strategies, we scaled up its biosynthetic production in yeast chassis to 2.34 g L⁻¹, future efforts may focus on enhancing productivity through additional metabolic engineering strategies.

Also, we have identified 54 terpene products, including the previously unreported compound cyclosantalol and the bioactive constituent (+)-intermedeol. To the best of our knowledge, neither of these compounds has been previously documented as constituents of mugwort. This discovery implies that there may be a multitude of compounds within the plant kingdom that remain unidentified, including bioactive substances present in low concentrations and intermediates in the complex metabolic pathways of plants. The gene-directed in vitro mining approach that we conducted in this study, allows scalable production and facile purification of target compounds without relying on direct plant material extraction. This approach is particularly valuable for species facing endangerment or legal restrictions, offering an alternative pathway for successful research, development, and commercialization of valuable natural products.

The precise elucidation of a compound's structure is pivotal for understanding its biological functionality. However, the fragment ions of terpenes exhibit striking similarities, distinct mass spectrometers may yield disparate fragment ion profiles for an identical compound, rendering the exclusive reliance on its mass spectrum for structural interpretation problematic. Additionally, the repetitive identification of known terpene structures by various research groups contributes to the inefficiency in discovering novel compounds. Here, we have developed a terpene synthase-standard (TPS-Standard) library by compiling well-characterized terpene synthases. In contrast to traditional compound-based standard libraries, our gene-based library offers reproducibility and ease of expansion, making it readily shareable among the scientific community. Most notably, this library enables the efficient filtering of known TPS products, thereby significantly enhancing the rate of novel compound discovery, including new skeletal structures.

CONCLUSIONS

In summary, our investigation has identified (+)-intermedeol as an insect-repellent constituent in mugwort, a herb that has been traditionally used as a natural insect repellent. By employing a gene-directed in vitro mining approach, and establishing of "Terpene synthase-standard library", we identified 54 terpene products from mugwort, including a novel compound cyclosantalol. After establishing the safety

and efficacy of (+)-intermedeol, we achieved its biosynthetic production in a yeast chassis. Our study not only validates the efficacy of (+)-intermedeol from mugwort but also establishes a promising workflow for discovering and commercially developing additional high-value plant natural products in future research.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/jacs.4c08857>.

Detailed description of material and methods: phylogenetic analysis of TPS genes; plasmids and strains; extraction, isolation, structural analysis of terpenoids; hand-in-cage assay of *Aedes aegypti*; tick repellency bioassay; yeast metabolic engineering for the large-scale production. Supplementary figures and tables including primers/plasmids/strains used in this study; GC spectrum of monoterpenes/sesquiterpenes in yeast strains; ¹H and ¹³C NMR spectra of compounds structures; and information on the terpene synthase-standard library (PDF)

Accession Codes

Deposition Numbers 2282322 and 2366970 contain the supplementary crystallographic data for this paper. These data can be obtained free of charge via the joint Cambridge Crystallographic Data Centre (CCDC) and Fachinformationszentrum Karlsruhe Access Structures service.

AUTHOR INFORMATION

Corresponding Authors

Li Lu – Department of Urology, Zhongnan Hospital of Wuhan University, School of Pharmaceutical Sciences, Wuhan University, Wuhan 430071, China; Hubei Hongshan Laboratory, Wuhan 430071, China; orcid.org/0000-0003-0490-7163; Email: luliwhu@whu.edu.cn

Tiangang Liu – Department of Urology, Zhongnan Hospital of Wuhan University, School of Pharmaceutical Sciences, Wuhan University, Wuhan 430071, China; Wuhan Hesheng Technology Co., Ltd., Wuhan 430074, China; State Key Laboratory of Microbial Metabolism, Joint International Research Laboratory of Metabolic & Developmental Sciences, and School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200030, China; orcid.org/0000-0001-8087-0345; Email: liutg@whu.edu.cn

Authors

Yao Zhi – Department of Urology, Zhongnan Hospital of Wuhan University, School of Pharmaceutical Sciences, Wuhan University, Wuhan 430071, China; Wuhan Hesheng Technology Co., Ltd., Wuhan 430074, China

Chong Dai – Department of Urology, Zhongnan Hospital of Wuhan University, School of Pharmaceutical Sciences, Wuhan University, Wuhan 430071, China

Xueting Fang – Department of Urology, Zhongnan Hospital of Wuhan University, School of Pharmaceutical Sciences, Wuhan University, Wuhan 430071, China

Xiaochun Xiao – Department of Urology, Zhongnan Hospital of Wuhan University, School of Pharmaceutical Sciences, Wuhan University, Wuhan 430071, China

Hui Lu – Department of Urology, Zhongnan Hospital of Wuhan University, School of Pharmaceutical Sciences, Wuhan University, Wuhan 430071, China

Fangfang Chen – Department of Urology, Zhongnan Hospital of Wuhan University, School of Pharmaceutical Sciences, Wuhan University, Wuhan 430071, China

Rong Chen – Department of Urology, Zhongnan Hospital of Wuhan University, School of Pharmaceutical Sciences, Wuhan University, Wuhan 430071, China

Weihua Ma – Hubei Hongshan Laboratory, Wuhan 430071, China; Hubei Insect Resources Utilization and Sustainable Pest Management Key Laboratory, College of Plant Science and Technology, Huazhong Agricultural University, Wuhan 430070, China; orcid.org/0000-0001-9521-2038

Zixin Deng – Department of Urology, Zhongnan Hospital of Wuhan University, School of Pharmaceutical Sciences, Wuhan University, Wuhan 430071, China

Complete contact information is available at:

<https://pubs.acs.org/10.1021/jacs.4c08857>

Author Contributions

[#]Y.Z. and C.D. contributed equally to this work.

Notes

The authors declare the following competing financial interest(s): The authors have applied for a patent based on this work.

ACKNOWLEDGMENTS

We would like to thank Prof. Qing Liu for his critical and instructive comments and Chongyang Chen for the illustration drawing on the manuscript. Cartoons in Figures ¹b, ³a, ⁴a, ⁴c, S2, S8, and S28a were created in BioRender, with permission. We would like to express our gratitude to Anwei Hou, Chenxi Huang, Ziling Ye, Shuai Fu, Weijia Cheng, Yan Yao, Qixia Wang, Jing Yang, Haomin Chi, and Botian Song for their assistance and advice on the experiments. This work is supported by grants from Hesheng Tech Co. Ltd., the National Key Research and Development Program (2022YFA0912100), the Natural Science Foundation of China (32100413, 32200040), and the Hubei Provincial Research Center of Basic Discipline of Biology.

REFERENCES

- (1) Scherlach, K.; Hertweck, C. Mining and unearthing hidden biosynthetic potential. *Nat. Commun.* **2021**, *12* (1), 3864.
- (2) Baldwin, I. T.; Halitschke, R.; Paschold, A.; von Dahl, C. C.; Preston, C. A. Volatile signaling in plant-plant interactions: "talking trees" in the genomics era. *Science* **2006**, *311* (5762), 5812–5815.
- (3) Tu, Y. Artemisinin—A Gift from Traditional Chinese Medicine to the World (Nobel Lecture). *Angew. Chem., Int. Ed.* **2016**, *55* (35), 10210–10226.
- (4) Wang, J.; Xu, C.; Wong, Y. K.; Li, Y.; Liao, F.; Jiang, T.; Tu, Y. Artemisinin, the magic drug discovered from traditional Chinese medicine. *Engineering* **2019**, *5* (1), 32–39.
- (5) Liu, Y.; He, Y.; Wang, F.; Xu, R.; Yang, M.; Ci, Z.; Wu, Z.; Zhang, D.; Lin, J. From longevity grass to contemporary soft gold: Explore the chemical constituents, pharmacology, and toxicology of *Artemisia argyi* H. Lévl. & *variot essential oil*. *Journal of ethnopharmacology* **2021**, *279*, No. 114404.
- (6) Chen, H.; Guo, M.; Dong, S.; Wu, X.; Zhang, G.; He, L.; Jiao, Y.; Chen, S.; Li, L.; Luo, H. A chromosome-scale genome assembly of *Artemisia argyi* reveals unbiased subgenome evolution and key contributions of gene duplication to volatile terpenoid diversity. *Plant Communications* **2023**, *4* (3), No. 100516.

- (7) Zhang, W.-J.; You, C.-X.; Yang, K.; Chen, R.; Wang, Y.; Wu, Y.; Geng, Z.-F.; Chen, H.-P.; Jiang, H.-Y.; Su, Y. Bioactivity of essential oil of *Artemisia argyi* Lévl. et Van. and its main compounds against *Lasioderma serricorne*. *J. Oleo Sci.* **2014**, *63* (8), 829–837.

- (8) Ahmed, M.; Peiwen, Q.; Gu, Z.; Liu, Y.; Sikandar, A.; Hussain, D.; Javeed, A.; Shafi, J.; Iqbal, M. F.; An, R. Insecticidal activity and biochemical composition of *Citrullus colocynthis*, *Cannabis indica* and *Artemisia argyi* extracts against cabbage aphid (*Brevicoryne brassicae* L.). *Sci. Rep.* **2020**, *10* (1), 522.

- (9) de la Fuente, J.; Estrada-Peña, A.; Rafael, M.; Almazán, C.; Bermúdez, S.; Abdelbaset, A. E.; Kasajia, P. D.; Kabi, F.; Akande, F. A.; Ajagbe, D. O.; Bamgbose, T.; Ghosh, S.; Palavesam, A.; Hamid, P. H.; Oskam, C. L.; Egan, S. L.; Duarte-Barbosa, A.; Hekimoğlu, O.; Szabó, M. P. J.; Labruna, M. B.; Dahal, A. Perception of Ticks and Tick-Borne Diseases Worldwide. *Pathogens* **2023**, *12* (10), 1258.

- (10) Luo, D.-Y.; Yan, Z.-T.; Che, L.-R.; Zhu, J. J.; Chen, B. Repellency and insecticidal activity of seven Mugwort (*Artemisia argyi*) essential oils against the malaria vector *Anopheles sinensis*. *Sci. Rep.* **2022**, *12* (1), 5337.

- (11) Miao, Y.; Luo, D.; Zhao, T.; Du, H.; Liu, Z.; Xu, Z.; Guo, L.; Chen, C.; Peng, S.; Li, J. X. Genome sequencing reveals chromosome fusion and extensive expansion of genes related to secondary metabolism in *Artemisia argyi*. *Plant Biotechnol. J.* **2022**, *20* (10), 1902–1915.

- (12) Bai, Y.; Liu, X.; Baldwin, I. T. Using Synthetic Biology to Understand the Function of Plant Specialized Metabolites. *Annu. Rev. Plant Biol.* **2024**, *75* (1), 629–653.

- (13) Tao, H.; Lauterbach, L.; Bian, G.; Chen, R.; Hou, A.; Mori, T.; Cheng, S.; Hu, B.; Lu, L.; Mu, X.; Li, M.; Adachi, N.; Kawasaki, M.; Moriya, T.; Senda, T.; Wang, X.; Deng, Z.; Abe, I.; Dickschat, J. S.; Liu, T. Discovery of non-squalene triterpenes. *Nature* **2022**, *606* (7913), 414–419.

- (14) Chen, R.; Jia, Q.; Mu, X.; Hu, B.; Sun, X.; Deng, Z.; Chen, F.; Bian, G.; Liu, T. Systematic mining of fungal chimeric terpene synthases using an efficient precursor-providing yeast chassis. *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118* (29), No. e2023247118.

- (15) Chen, F.; Tholl, D.; Bohlmann, J.; Pichersky, E. 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.* **2011**, *66* (1), 212–229.

- (16) Miao, Y.; Luo, D.; Zhao, T.; Du, H.; Liu, Z.; Xu, Z.; Guo, L.; Chen, C.; Peng, S.; Li, J. X.; Ma, L.; Ning, G.; Liu, D.; Huang, L. Genome sequencing reveals chromosome fusion and extensive expansion of genes related to secondary metabolism in *Artemisia argyi*. *Plant Biotechnol. J.* **2022**, *20* (10), 1902–1915.

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

- (18) Ye, Z.; Shi, B.; Huang, Y.; Ma, T.; Xiang, Z.; Hu, B.; Kuang, Z.; Huang, M.; Lin, X.; Tian, Z. Revolution of vitamin E production by starting from microbial fermented farnesene to isophytol. *Innovation* **2022**, *3* (3), No. 100228.

- (19) Deng, X.; Shi, B.; Ye, Z.; Huang, M.; Chen, R.; Cai, Y.; Kuang, Z.; Sun, X.; Bian, G.; Deng, Z. Systematic identification of *Ocimum sanctum* sesquiterpenoid synthases and (–)-eremophilene overproduction in engineered yeast. *Metab. Eng.* **2022**, *69*, 122–133.

- (20) Zhou, F.; Pichersky, E. The complete functional characterisation of the terpene synthase family in tomato. *New Phytologist* **2020**, *226* (5), 1341–1360.

- (21) Akhtar, T. A.; Matsuba, Y.; Schauvinhold, I.; Yu, G.; Lees, H. A.; Klein, S. E.; Pichersky, E. The tomato cis-prenyltransferase gene family. *Plant Journal* **2013**, *73* (4), 640–652.

- (22) Schillmiller, A. L.; Schauvinhold, I.; Larson, M.; Xu, R.; Charbonneau, A. L.; Schmidt, A.; Wilkerson, C.; Last, R. L.; Pichersky, E. Monoterpenes in the glandular trichomes of tomato are synthesized from a neryl diphosphate precursor rather than geranyl diphosphate. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (26), 10865–10870.

- (23) Lin, X.; Hopson, R.; Cane, D. E. Genome Mining in *Streptomyces coelicolor*: Molecular Cloning and Characterization of a New Sesquiterpene Synthase. *J. Am. Chem. Soc.* **2006**, *128* (18), 6022–6023.
- (24) Martin, D. M.; Aubourg, S.; Schouwey, M. B.; Daviet, L.; Schalk, M.; Toub, O.; Lund, S. T.; Bohlmann, J. Functional annotation, genome organization and phylogeny of the grapevine (*Vitis vinifera*) terpene synthase gene family based on genome assembly, FLcDNA cloning, and enzyme assays. *BMC Plant Biol.* **2010**, *10* (1), 266.
- (25) Rinkel, J.; Dickschat, J. S. Stereochemical investigations on the biosynthesis of achiral (*Z*)- γ -bisabolene in *Cryptosporangium arum*. *Beilstein Journal of Organic Chemistry* **2019**, *15* (1), 789–794.
- (26) Bohlmann, J.; Crock, J.; Jetter, R.; Croteau, R. Terpenoid-based defenses in conifers: cDNA cloning, characterization, and functional expression of wound-inducible (*E*)- α -bisabolene synthase from grand fir (*Abies grandis*). *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (12), 6756–6761.
- (27) Yu, F.; Okamoto, S.; Nakasone, K.; Adachi, K.; Matsuda, S.; Harada, H.; Misawa, N.; Utsumi, R. Molecular cloning and functional characterization of α -humulene synthase, a possible key enzyme of zerumbone biosynthesis in shampoo ginger (*Zingiber zerumbet* Smith). *Planta* **2008**, *227*, 1291–1299.
- (28) Deguerry, F.; Pastore, L.; Wu, S.; Clark, A.; Chappell, J.; Schalk, M. The diverse sesquiterpene profile of patchouli, *Pogostemon cablin*, is correlated with a limited number of sesquiterpene synthases. *Archives of Biochemistry and Biophysics* **2006**, *454* (2), 123–136.
- (29) Kollner, T. G.; Schnee, C.; Gershenzon, J.; Degenhardt, J. The variability of sesquiterpenes emitted from two *Zea mays* cultivars is controlled by allelic variation of two terpene synthase genes encoding stereoselective multiple product enzymes. *Plant Cell* **2004**, *16* (5), 1115–1131.
- (30) Gennadios, H. A.; Gonzalez, V.; Di Costanzo, L.; Li, A.; Yu, F.; Miller, D. J.; Allemann, R. K.; Christianson, D. W. Crystal structure of (+)- δ -cadinene synthase from *Gossypium arboreum* and evolutionary divergence of metal binding motifs for catalysis. *Biochemistry* **2009**, *48* (26), 6175–6183.
- (31) Russell, G. B.; Hunt, M. B.; Bowers, W. S.; Blunt, J. W. A sesquiterpenoid ant repellent from *Dysoxylum spectabile*. *Phytochemistry* **1994**, *35* (6), 1455–1456.
- (32) Brunke, E. J.; Vollhardt, J.; Schmaus, G. Cyclosantalal and epicyclosantalal—new sesquiterpene aldehydes from east Indian sandalwood oil. *Flavour and Fragrance Journal* **1995**, *10* (3), 211–219.
- (33) Kesselmans, R. P. W.; Wijnberg, J. B.; Groot, A. D.; Vries, N. K. D. Synthesis of all stereoisomers of eudesm-11-en-4-ol. 1. Stereospecific synthesis of the trans- and cis-fused octahydro-8-hydroxy-4,8-dimethyl-2(1H)-naphthalenones. Conformational analysis of the cis-fused compounds. *J. Org. Chem.* **1991**, *56*, 7232–7236.
- (34) Zhang, J.-Y.; Xia, C.; Wang, H.-F.; Tang, C. Recent advances in electrocatalytic oxygen reduction for on-site hydrogen peroxide synthesis in acidic media. *Journal of Energy Chemistry* **2022**, *67*, 432–450.
- (35) Tang, B.; Xu, K.; Liu, Y.; Zhou, Z.; Karthi, S.; Yang, H.; Li, C. A review of physiological resistance to insecticide stress in *Nilaparvata lugens*. *3 Biotech* **2022**, *12* (3), 84.
- (36) Tian, C.; Tek Tay, W.; Feng, H.; Wang, Y.; Hu, Y.; Li, G. Characterization of *Adelphocoris suturalis* (Hemiptera: Miridae) transcriptome from different developmental stages. *Sci. Rep.* **2015**, *5* (1), No. 11042.
- (37) Cantrell, C. L.; Klun, J. A.; Bryson, C. T.; Kobaisy, M.; Duke, S. O. Isolation and identification of mosquito bite deterrent terpenoids from leaves of American (*Callicarpa americana*) and Japanese (*Callicarpa japonica*) beautyberry. *J. Agric. Food Chem.* **2005**, *53* (15), 5948–5953.
- (38) Jeong, S.-H.; Koo, S.-J.; Choi, J.-H.; Park, J.-H.; Ha, J.; Park, H.-J.; Lee, K.-T. Intermedeol isolated from the leaves of *Ligularia fischeri* var. *spiciformis* induces the differentiation of human acute promyelocytic leukemia HL-60 cells. *Planta Medica* **2002**, *68* (10), 881–885.
- (39) Cantrell, C.; Klun, J.; Bryson, C.; Kobaisy, M.; Duke, S. Isolation and identification of mosquito bite deterrent terpenoids from leaves of American (*Callicarpa americana*) and Japanese (*Callicarpa japonica*) beautyberry. *J. Agric. Food Chem.* **2005**, *53* (15), 5948–5953.
- (40) Kain, P.; Boyle, S. M.; Tharadra, S. K.; Guda, T.; Pham, C.; Dahanukar, A.; Ray, A. Odour receptors and neurons for DEET and new insect repellents. *Nature* **2013**, *502* (7472), 507–512.
- (41) Liu, F.; Wang, Q.; Xu, P.; Andrezza, F.; Valbon, W. R.; Bandason, E.; Chen, M.; Yan, R.; Feng, B.; Smith, L. B. A dual-target molecular mechanism of pyrethrum repellency against mosquitoes. *Nat. Commun.* **2021**, *12* (1), 2553.
- (42) Clarkson, T. C.; Janich, A. J.; Sanchez-Vargas, I.; Markle, E. D.; Gray, M.; Foster, J. R.; Black, W. C., IV; Foy, B. D.; Olson, K. E. Nootkatone is an effective repellent against *Aedes aegypti* and *Aedes albopictus*. *Insects* **2021**, *12* (5), 386.