

NPtagM: A Tailoring Enzyme Genome Mining Toolkit and Its Application in Terpenoid P450s from Phytopathogenic Fungi

Zhanren Cong, Njeru Joe Mukoma, Qiang Yin, Bin Zhu, Lingwei She, Tom Hsiang, Lixin Zhang, Lan Jiang,* and Xueting Liu*



Cite This: *J. Agric. Food Chem.* 2024, 72, 27225–27234



Read Online

ACCESS |



Metrics & More



Article Recommendations



Supporting Information

ABSTRACT: Terpenoids derived from phytopathogenic fungi are major participants in interactions among microorganisms, plants, and animals. The modifications catalyzed by cytochrome P450s significantly influence the structural and bioactivity diversity of the terpenoids. To conduct genome mining of P450s in pathogenic fungi, in this study, we developed a new software called Natural Products Tailoring Enzymes Genome Mining (NPtagM). By optimizing the workflow and gene prediction software, NPtagM demonstrated a 3-fold increase in the number of predicted P450s and an 8-fold reduction in runtime compared to antiSMASH. We then used it to extract 1189 dereplicated terpenoid P450s from our in-house fungal genomes. Using a sequence similarity network analysis, we identified a family that potentially produced eremophilane-type sesquiterpenoids. The heterologous expression in *Aspergillus oryzae* resulted in the production of two new and four known eremophilanes. Our results highlight the potential of NPtagM in genome mining for tailoring enzymes from phytopathogenic fungi.

KEYWORDS: NPtagM, tailoring enzyme genome mining, terpenoids, cytochrome P450, phytopathogenic fungi

INTRODUCTION

Terpenoids, with over 170,000 known varieties, represent the largest and most structurally diverse family of natural products.¹ Their diverse structures contribute to a broad range of biological activities, including antibacterial, anti-inflammatory, and cytotoxic effects,² making them significant for both plant defense and ecological interactions. Fungi are the primary microbial source of terpenoids, accounting for about 80% of microbial terpenoids,¹ with those derived from phytopathogenic fungi playing a crucial role in interactions and defense responses between plants, microorganisms, and animals. While protecting themselves, they also impact the environment, influencing plant communities and ecosystem evolution. Additionally, their economic value is increasingly apparent, offering a vast application potential. Previous studies delineated the biosynthesis of fungal terpenoids into two steps. First, terpene synthases (TPSs) catalyze the cyclization of acyclic substrates to produce structurally diverse skeletons. Subsequently, these skeletons are enzymatically decorated by tailoring enzymes such as cytochrome P450s (P450s), FAD-dependent monooxygenases, and methyltransferases.³

P450 enzymes represent one of the most important families among these tailoring enzymes associated with terpenoids. Almost all of the terpenoid biosynthetic pathways are reported to include at least one P450 enzyme.⁴ These enzymes catalyze various reactions on inert carbon–hydrogen skeletons of terpenoids, including hydroxylation, C–C bond cleavage, and oxidative cyclization, among others. These reactions typically serve as the starting point for structural rearrangement of terpenoids.⁵ Modifications catalyzed by P450 enzymes can significantly enhance the bioactivity and structural diversity of terpenoids.^{3,5,6} Given the significance of P450 enzymes in

terpenoid biosynthesis, genome mining of terpene-related P450s in fungi is a promising approach to unveiling novel terpenoids and deepen our understanding of terpenoid biosynthesis.

However, the targeted genome mining of tailoring enzymes faces some challenges. Unlike core enzymes, the tailoring enzymes are not only involved in the biosynthesis of natural products but also widely distributed in the biosynthesis of primary metabolites. This ubiquity implies that a certain portion of tailoring enzymes extracted from genomes will not contribute to the biosynthesis of natural products.⁷ Given that tailoring enzymes are typically clustered with core enzymes in microorganisms, this issue could be solved by extracting tailoring enzymes associated with biosynthetic gene clusters (BGCs).

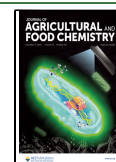
Several software programs, such as antiSMASH and PRISM,^{8,9} have been widely used to predict BGCs in fungal genomes. To apply them in the genome mining of tailoring enzymes, in this study, we developed a new software, i.e., NPtagM, that allows targeted prediction and specific extraction of natural product tailoring enzymes from large microbial genomics data sets based on antiSMASH. Using NPtagM, we extracted P450s associated with terpene biosynthesis from an in-house fungal genome library and performed clustering analysis using sequence similarity network (SSN) analysis,¹⁰

Received: August 20, 2024

Revised: November 22, 2024

Accepted: November 22, 2024

Published: December 2, 2024



which led to the discovery of a class of potential P450s that which might produce eremophilane-type sesquiterpenoids. Using the heterologous expression system in *Aspergillus oryzae*, two new eremophilane-type sesquiterpenoids were isolated and characterized.

MATERIALS AND METHODS

NPtagM Algorithm. NPtagM accepts microorganism genome files as the input data. Prodigal is utilized for gene prediction in bacteria genomes,¹¹ whereas GlimmerHMM¹² and AUGUSTUS¹³ are employed for gene prediction in fungi genomes. Then, potential tailoring enzymes were extracted from microorganism protein files using hmmersearch from HMMer¹⁴ and BLAST¹⁵ with cutoff values set at 0.01 and 10^{-5} , respectively, and CD-HIT¹⁶ was employed to dereplicate these enzymes based on sequence identity with a cutoff value set at 0.9. Cutoff values of HMMer, BLAST, CD-HIT, and gene cluster length can be adjusted according to specific requirements. Subsequently, NPtagM located the tailoring enzyme in genome using BLAST, and the gene cluster within 50 kb upstream and downstream of the tailoring enzyme was extracted and submitted to antiSMASH⁸ for identifying BGCs, excluding tailoring enzymes located distantly from the core gene. NPtagM could also automatically extract studied tailoring enzymes from MIBiG⁴ using BLAST and HMMer with the same cutoff criteria (details in the Supplementary Text in the SI). NPtagM was developed within Python and could run on a Linux Ubuntu system. The source code of NPtagM is available on GitHub at <https://github.com/ZhanRenCong/NPtagM>.

NPtagM Validation. NPtagM algorithm was validated in two ways: (1) running on 440 experimentally characterized fungi BGCs from MIBiG and (2) running on 10 bacteria genomes and 10 fungi genomes alongside antiSMASH. We downloaded the P450 hmm profile from the Pfam-A database (Pfam accession number: PF00067) and used it as the query file of NPtagM.¹⁷ GlimmerHMM and AUGUSTUS were utilized as the fungi gene prediction software in NPtagM, with *Aspergillus oryzae* as the reference organism in AUGUSTUS. Prodigal and prokka were utilized as the bacteria gene prediction software. Other parameters were the same as described above. P450s from the BGCs predicted by antiSMASH and MIBiG were extracted by HMMer with an e-value cutoff set at 0.01. The P450s predicted by NPtagM and antiSMASH are listed in Tables S1–S6.

Extraction and Bioinformatics Analysis of Terpenoid Biosynthetic P450s from the In-House Fungi Genome Library Using NPtagM. A set of 430 complete genome sequences was obtained from our in-house fungi genomic library. We downloaded the P450 hmm profile from the Pfam database (Pfam accession number: PF00067) and used it as the query file. Our in-house fungi genome files and their gene protein files were used as a database. Fungal TPSs were categorized in five families based on their sequence characteristic: sesquiterpene synthases, trichodiene synthases-like sesquiterpene synthases, bifunctional terpene synthases, CPS/KS diterpene synthases, and triterpene synthases. To extract all these fungal terpene synthases, we use four different hmm profiles as the core gene query files: Terpene syn C 2 (Pfam accession number: PF19086), SQHop cyclase C (Pfam accession number: PF13243.8), CPS/KS, and trichodiene synthases-like sesquiterpene synthases (built by sequences in Table S7). AUGUSTUS was set as the gene prediction software. The distance between the P450s and TPSs was set to 10 kb. So, the extracted sequences were dereplicated by CD-HIT with a cutoff of 0.8. Finally, 1189 dereplicated P450 terpenoid biosynthetic sequences were extracted from our in-house fungi genome library. Besides these sequences, 89 and 29 terpenoid biosynthetic P450 sequences were identified from the MIBiG database and manually collected from published papers, respectively (Tables S8 and S9). The sequence identity between these P450s was calculated by using BLAST to generate SSN profiles. SSNs were then visualized using Cytoscape with an e-value cutoff set at 10^{-120} .¹⁸

Construction of Expression Plasmids for Eremophilane-Type Sesquiterpenoid BGCs and Transformation of *A. oryzae*.

The gDNA of phytopathogenic fungi *Colletotrichum siamense* 15011 was provided by Prof. Hsiang from the University of Guelph (Ontario, Canada). *Aspergillus oryzae* NSAR1 (*niaD*[−], *sC*[−], *adeA*[−], *ΔargB*) were used as the host for gene expression. Transformants of the *A. oryzae* strain were grown in an MPY medium at 30 °C and 200 rpm for 5 days. *Escherichia coli* DH10B was used for gene cloning.

The two genes *CseA* and *CseB* that encode for STS and P450, respectively, were amplified from the gDNA of *C. siamense* 15011 with primers shown in Table S10. PCRs were performed with the Q5 High-Fidelity DNA Polymerase (New England Biolabs). The ClonExpressII One Step Cloning Kit (Vazyme Biotech) was used to construct expression plasmid pUARA4-*CseA* by inserting the *cseA* PCR product into the *KpnI* restriction sites of pUARA4. pUARA4-*CseAB* was constructed by inserting the *CseB* PCR product into the *NdeI* restriction sites of pUARA4-*CseA* utilizing the same kit. The constructed plasmids and transformants are summarized in Table S11. The STS and P450 sequences are available in the National Center for Biotechnology Information under accession numbers PP516636 and PP516637.

Transformation of *A. oryzae* NSAR1 (1.0×10^8 cells) was performed by the protoplast-polyethylene glycol method reported previously to construct the transformants AO-*CseA* (containing pUARA2-*CseA*) and AO-*CseAB* (containing pUARA2-*CseAB*).¹⁹ AO is an abbreviation of *A. oryzae*, and AO-*CseA* means a transformant harboring the *cseA* gene. AO transformants were grown on an MPY (maltose polypeptone yeast extract medium; 3% maltose, 1% polypeptone, and 0.5% yeast extract) medium containing 1% (NH₄)₂SO₄ and 0.01% adenine for 3–5 days at 30 °C.

To detect the transcription, the total RNA of AO transformants was extracted from dried mycelia using an AxyPrep Multisource Total RNA Miniprep Kit (Axygen) according to the manufacturer's instructions and then treated with DNase I (Life Technologies) for reverse transcription. Complementary DNA (cDNA) was synthesized with a PrimeScript RT Reagent Kit with a gDNA Eraser (Takara) using the oligo dT primer according to the manufacturer's instructions. RNA integrity was confirmed by electrophoresis on a TAE (Tris-base acid-EDTA) agarose gel, and the concentration was determined by using a Nanodrop instrument (Thermo Scientific). Reverse transcripts were prepared from 500 ng of the total RNA by the PrimeScript RT Reagent Kit using the gDNA Eraser (Takara) with random primers following the manufacturer's instructions. PCR was performed with the Q5 High-Fidelity DNA Polymerase (New England Biolabs) in the presence of 25 ng of reverse transcribed RNA.

Heterologous Expression of Eremophilane-Type Sesquiterpenoid BGCs and Identification of Metabolites. Mycelia of STS transformant AO-*CseA* were inoculated into the MPY medium (5 mL) containing 0.925% (NH₄)₂SO₄, 0.15% methionine, and 0.01% adenine in a 20 mL headspace screw top vial at 30 °C for 3 days. Mycelia of transformant AO-*CseAB* were inoculated into the MPY medium (50 mL) containing 0.925% (NH₄)₂SO₄ and 0.01% adenine in a 250 mL Erlenmeyer flask to prepare seed culture. Fermentation was then carried out in 10 Erlenmeyer flasks (3000 mL), each containing 1000 mL of the MPY medium at 30 °C and 220 rpm for 5 days.

The fermentation products of transformant AO-*CseA* were extracted with a solid-phase microextraction (SPME) fiber (50/30 μm DVB/CAR/PDMS; Stableflex, 24 Ga, manual holder) for 15 min at room temperature. After extraction, the SPME fiber was inserted into the injection port of a QP2010SE (Shimadzu, Kyoto, Japan) GC–MS apparatus with a DB-5 MS capillary column (0.25 mm × 30 m, 0.25 μm film thickness, SHIMADZU) in splitless mode. The column temperature was increased at 30 °C/min from 60 to 120 °C, then increased at 5 °C/min to 180 °C, and then increased at 30 °C/min to 270 °C. The flow rate of the helium carrier gas was 1.41 mL/min. The structures of compounds detected by GC–MS were identified by comparing their mass spectra with those of terpenoids in the National Institute of Standards and Technology (NIST) standard reference database.

The culture broth of transformant AO-*CseAB* was extracted exhaustively with ethyl acetate (10 L × 3). The combined organic

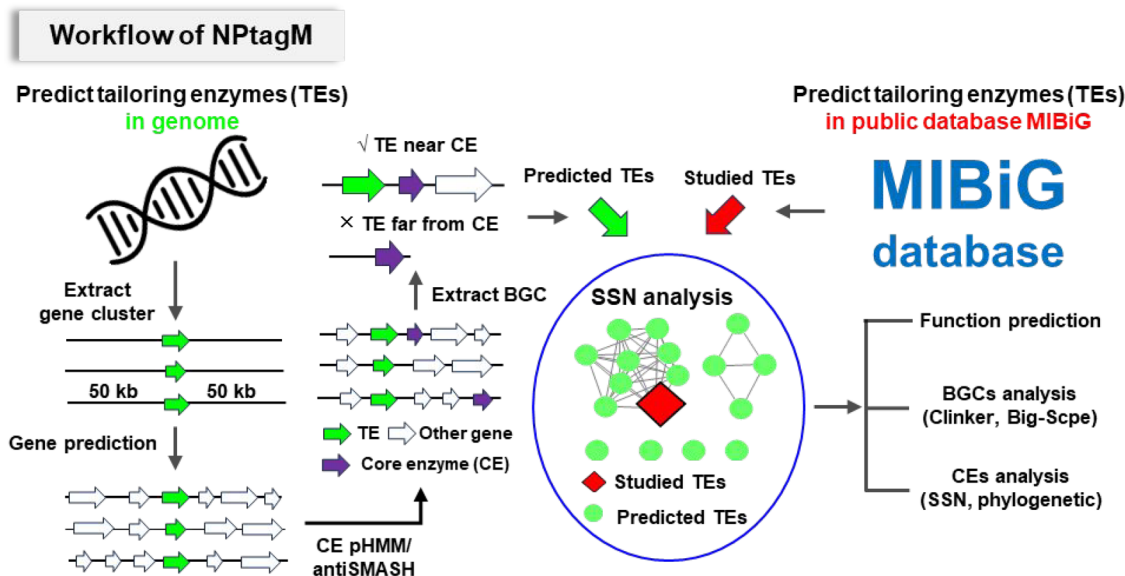


Figure 1. Workflow of NPtagM. The NPtagM workflow begins by extracting potential tailoring enzymes from microorganism genomes. Subsequently, NPtagM retrieves the gene cluster associated with the identified tailoring enzymes, subjecting it to further gene prediction utilizing various software tools. NPtagM then identifies core enzyme genes within these annotated gene clusters using antiSMASH or users' own pHMM, excluding tailoring enzymes distant from the core enzymes. Additionally, NPtagM extracts target tailoring enzymes from the MIBiG database. An SSN could be constructed to visualize the sequence relationships among reported and predicted tailoring enzymes. For interested clusters from SSN, their products could be further predicted by integrating the core enzymes and BGC analysis.

layers were concentrated to yield a dark brown oil extract (5.23 g), were subjected to the silica gel column eluted by a gradient hexane-EtOAc-methanol, and afforded fractions G1–G9. G4 (189 mg) was purified by preparative RP-HPLC using an ACE C18-PFP (10 × 250 mm) column eluting at a flow rate of 4.0 mL/min using an isocratic elution (24% CH₃CN-H₂O) to obtain **1** (8.3 mg, t_R = 28.4 min), **2** (2.4 mg, t_R = 31.5 min), **4** (2.2 mg, t_R = 34.6 min), **5** (1.6 mg, t_R = 38.2 min), and **6** (2.6 mg, t_R = 45.7 min). G3 (143 mg) was purified by preparative RP-HPLC using an ACE C18-PFP (10 × 250 mm) column eluting at a flow rate of 4.0 mL/min using a gradient elution: 0 min, 25% CH₃CN-H₂O and 35 min, 35% CH₃CN to yield **3** (2.2 mg, t_R = 30.1 min).

RESULTS

The Workflow of the NPtagM Algorithm Detects NPs-Related Tailoring Enzymes. To address the prolonged runtime and reduced accuracy in extracting NPs-related tailoring enzymes, we developed NPtagM. This algorithm significantly reduces runtime by employing a “linkage extraction” method, which focuses on searching for core enzyme genes located near tailoring enzyme genes. Additionally, to enhance the prediction accuracy, NPtagM selectively applies AUGUSTUS,¹³ which has demonstrated superior performance in predicting genes through fungal genomes compared to GlimmerHMM,¹² the tool traditionally used in antiSMASH.

The workflow of NPtagM is scanning a genome for the potentially novel NPs-related tailoring enzymes. All tailoring enzymes were first extracted from the microbial genomes using HMMer or BLAST, and redundancies were removed using CD-HIT based on sequence identity. Subsequently, gene clusters 50 kb upstream and downstream of the tailoring enzyme were extracted from the genome and subjected to additional gene prediction using GlimmerHMM or AUGUSTUS. NPtagM then employs antiSMASH or users' own core gene profiles' hidden Markov model (pHMM) to identify core enzyme genes within these annotated gene clusters. Finally,

biosynthetic gene clusters (BGCs) that contain both the core and modified enzyme genes are filtered and retained. Meanwhile, NPtagM will automatically extract specific tailoring enzymes and their gene clusters from the MIBiG with the same parameters and export them together with those retrieved from the genome, which are further characterized by SSN clustering analysis and experiments (Figure 1).

Comparison and Validation of the Efficiency of NPtagM with antiSMASH. To demonstrate the accuracy of NPtagM, we used it to predict P450s in 440 experimentally characterized fungal BGCs from MIBiG. A total of 507 P450s were manually extracted from the annotated BGCs. When antiSMASH was used to predict biosynthetic P450s from this genomic library, 250 P450s were successfully identified. In contrast, NPtagM identified 273 and 420 biosynthetic P450s utilizing GlimmerHMM and AUGUSTUS as the gene prediction software, respectively (Table 1).

Table 1. Accuracy of NPtagM Compared to antiSMASH^a

software	accuracy
antiSMASH	253/507 (49.90%)
NPtagM (GlimmerHMM)	273/507 (53.85%)
NPtagM (AUGUSTUS)	420/507 (82.84%)

^aBoth tools were tested by extracting biosynthetic P450s in 440 experimentally characterized fungal BGCs from MIBiG.

The runtime and efficiency of the NPtagM algorithm were further validated by running it on 10 fungi genomes alongside antiSMASH. Using GlimmerHMM as the gene prediction software, antiSMASH requires approximately 245 min to detect all BGCs. NPtagM identifies all P450s in BGCs in approximately 33 min, representing an 8-fold reduction in runtime compared to antiSMASH. Additionally, a similar number of biosynthetic P450s were extracted (38 for antiSMASH and 41 for GlimmerHMM). When AUGUSTUS

was used as the gene prediction software, NPtagM extracted 110 P450s, nearly 3 times the number extracted by antiSMASH. However, we also observed a significant increase in runtime, from 33 to 215 min, due to the lengthy time required for gene prediction across the entire genome. Submitting the protein sequence files along with the genome data could bypass this step, reducing the time to 53 min (Table 2). Besides P450s, we also applied NPtagM in predicting α -

Table 2. Performance of NPtagM Compared to antiSMASH^a

software	runtime	number of extracted P450
antiSMASH	245 min 36 s	38
NPtagM (GlimmerHMM)	33 min 31 s	41
NPtagM (AUGUSTUS)	215 min 16 s	110
NPtagM (AUGUSTUS) [skip gene prediction across genomes]	53 min 18 s	

^aBoth tools were tested by extracting biosynthetic P450s from a data set consisting of 10 fungal genomes.

ketoglutarate-dependent mononuclear nonheme iron enzymes, which demonstrated a 2-fold increase in the number of predicted enzymes and an 11-fold reduction in runtime (Table S12).

SSN Analysis of Terpenoid Biosynthetic P450s from the Fungi Genome Library Forms a Comprehensive Landscape of Terpenoid Biosynthetic P450s. Utilizing NPtagM, 1189 dereplicated terpenoid biosynthetic P450 sequences were successfully extracted from our in-house fungal genome library comprising 430 fungal genomes. Additionally, 89 terpene biosynthesis P450s were identified from the MIBiG database, and 29 were manually collected from published papers (Tables S8 and S9). These P450s are associated with various TPSs, i.e., 127 with triterpene synthases (TTSs), 119 with trichodiene synthase-like sesquiterpene synthases (TDTs), 464 with sesquiterpene synthases, 282 with monofunctional diterpene/sesterterpene synthases (MFTSs), and 215 with CPS/KS diterpene synthases (CPS/KSs) (Figure 2).

Subsequently, we systematically investigated the sequence relationships among terpenoid biosynthetic P450s, aiming to create a comprehensive map that could be systematically searched to identify P450s of biosynthetic or taxonomic significance. We used BLAST to compute a sequence identity distance matrix for the 1189 predicted terpenoid biosynthetic P450s from our in-house fungal genomic library, along with the 118 reported members from MIBiG and published papers. Using SSN with an e-value cutoff set at 10^{-120} , we identified 321 families of terpenoid biosynthetic P450s with distinct core genetic components (Figure 3). These families include 2 large families (families 1 and 2, with 473 and 117 P450s), 7 medium families (families 3 to 9, each containing 16–100 P450s), 26 small families (each containing 4–15 P450s), and 286 mini families (each containing 1–3 P450s). The SSN analysis revealed an unexpected robust correlation between P450 sequences and their corresponding TPSs. A total of 473 P450s from CPS/KS and MFTS BGCs, distributed within family 1, are thought to be implicated in the oxidation of diterpene and sesterterpene skeletons, and 240 P450s from STS and TDTs

BGCs are grouped into six families (families 2–7), likely responsible for modifying sesquiterpene skeletons. Additionally, 30 P450s from TTS BGCs are categorized into families 8 and 9, probably involved in the modification of triterpene skeletons.

Function Analysis of the P450 Families Revealed Unexplored P450s Responsible for Eremophilane-Type Sesquiterpenoids. In total, 54.3% of fungal P450s were clustered in nine large and medium families, and we further predicted the functions of all large and medium families containing reported P450s by integrating the TPS and BGC analysis (details in the Supporting Information, Figure S1). Through secondary SSN analysis, P450s in family 1 are implicated in modifying a spectrum of diesterpene/sesterterpene skeletons (Figure S2). P450s identified in families 2 and 3 are likely involved in the production of eremophilane-type sesquiterpenoids and heptelidic acid analogs, respectively (Figures S3–S5). P450s in family 4 may be involved in modifying tricyclic sesquiterpene skeletons to generate protoilludane-type and other fungi-specific sesquiterpenoids (Figure S6). P450s in family 6 are implicated in the biosynthesis of botryane-type sesquiterpenoids (Figures S7 and S8). Moreover, P450s in families 8 and 9 may be responsible for producing fusidane-type and hopane-type triterpenoids, respectively (Figures S9 and S10). However, there are some P450s whose functions cannot be well clustered, such as families 5 and 7. It is possible that these P450s are multifunctional and are involved in the modification of different types of terpene skeletons.

We focused on family 2, which comprises 117 P450s, including two reported P450s responsible for the biosynthesis of eremophilane-type sesquiterpenoid PR-toxin.²⁰ Eremophilane-type sesquiterpenoids exhibit a range of phytotoxic, antimicrobial, anticancer, and immunomodulatory properties (Figure 4A).²¹ Over 200 eremophilanes have been isolated from different fungi.²² However, aside from PR toxins, there have been few reports of BGCs for such compounds. Subsequent sequence similarity analysis of STSs associated with family 2 P450s revealed that over 90% of STSs in this group clustered with two STSs (prx2 and AtAS) capable of producing aristolechene, the skeleton of eremophilane (Figure 4B).^{23,24} The BGCs containing these P450s are approximately 20 kb in size and harbor a core set of genes, including STS and P450s, as well as other oxidoreductases such as short-chain dehydrogenases, α -KG-dependent oxidoreductases, FAD-dependent enzymes, and GMC oxidoreductases (Figure 4C and Table S13). These results strongly suggest that P450s from family 2 may be involved in eremophilane-type sesquiterpene biosynthesis.

The Heterologous Expression of the Selected P450 Demonstrated Its Capability to Produce Eremophilane-Type Sesquiterpenoids. One BGC from *C. siamense* 15011 (*C. siamense*eremophilane, *Cse*) that harbors the STS and P450 from family 2 was selected and identified (Figure 5A and Table S13) to investigate whether these P450s participate in the oxidative modifications of eremophilane-type sesquiterpenoids. STS encoded by *cseA* possessed a conserved motif similar with other reported STS, including N-terminal ¹¹³DDXXD/E¹¹⁷ motif and ²⁴²NSE/DTE²⁴⁸ triad in the TC domain (Figure S11). To validate the function of *cseA*, we construct an AO-*cseA* transformant by introducing *cseA* into *A. oryzae*. The fermentation products of the AO and AO-*cseA* were compared by GC–MS, and sesquiterpene skeleton aristolechene was

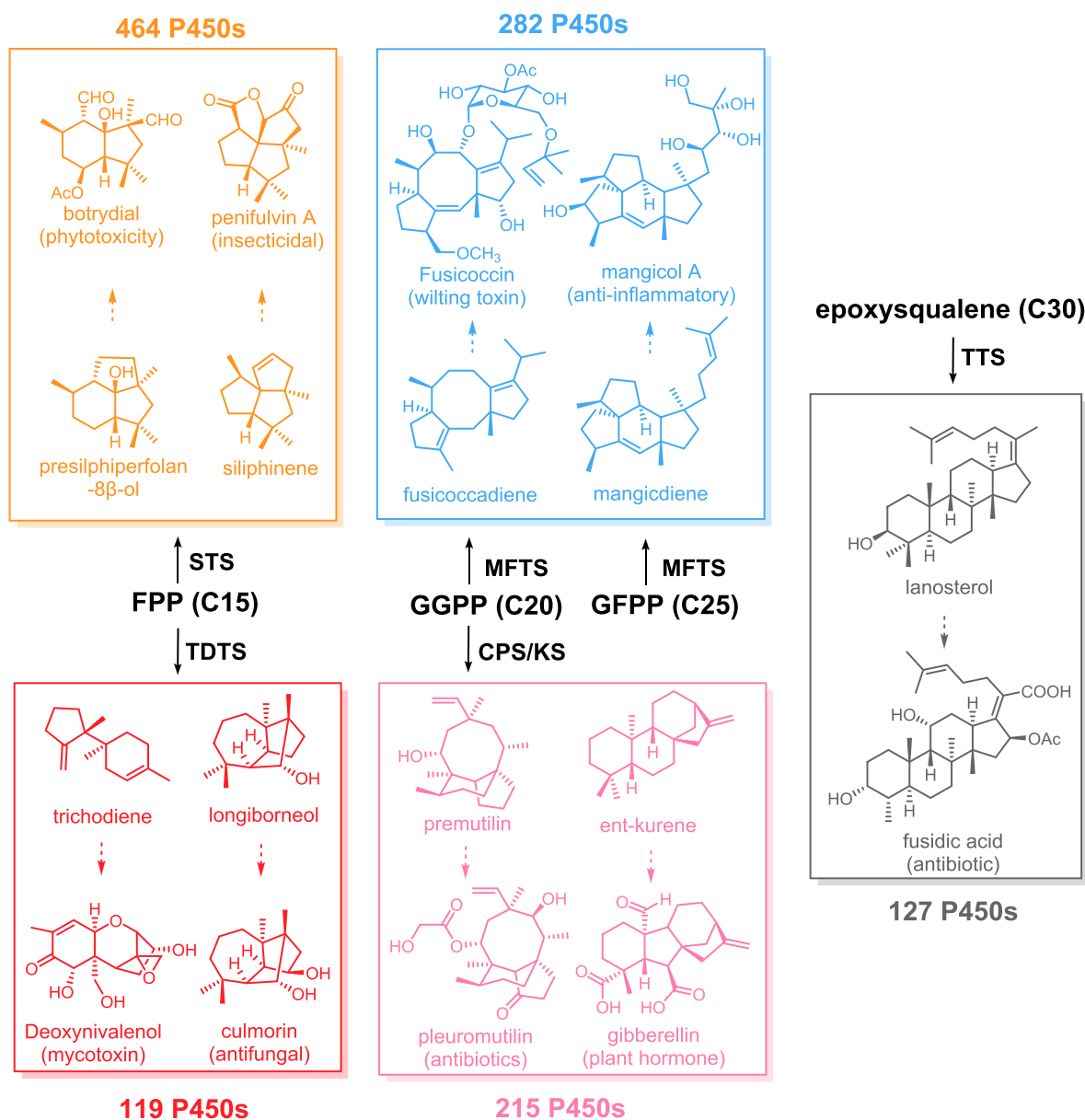


Figure 2. Terpene skeletons produced by different fungal TPS families and their modified bioactive terpenoids. The number of P450 enzymes associated with each TPS that were extracted from our in-house fungal genomes using NPtagM was indicated near each box.

detected as an additional peak in AO-*cseA*, with a molecular ion peak at m/z 204 (Figure 5B). Subsequently, P450 encoded by *CseB*, homologous to P450 ORF11 catalyzed hydroxylation in the biosynthesis of PR-toxin,²⁵ was cotransferred into *A. oryzae* to generate the transformant AO-*cseAB*. Six sesquiterpenes with different site oxidations of aristolochene were isolated and identified from AO-*CseAB*, including two new eremophilanes (1 and 2), along with four reported ones (3–6) (Figure 5C,D).

Compound 1 was isolated as a colorless oil, λ_{\max} = 200 nm. The molecular formula $C_{15}H_{25}O_2$ was deduced by LC-HRMS (m/z 219.1746 [$M-H_2O + H$]⁺, calcd. for 219.1749), indicating four units of unsaturation. The 1H NMR, ^{13}C NMR, and HSQC spectra of 1 revealed the presence of 15 carbon signals and showed the presence of two methyl groups, four methylene groups, one olefinic methylene group, one hydroxymethyl group, two methine groups, one olefinic methine carbon, one hydroxyl methine group, and three

quaternary carbons. A comparison of the 1D and 2D NMR data of 1 to paraconiothins D indicated that they shared the same eremophilane skeleton.²⁶ The hydroxylated position was assigned as C-2 based on the HMBC correlation of H-1/C-9, H-1/C-10, and H-1/C-3 and 1H – 1H COSY correlations of H-1/H-2, H-2/H-3. Another hydroxylated position was assigned as C-12 based on the HMBC correlation of H-12/C-13. Thus, the planar structure of compound 1 was fully constructed. Based on key NOESY correlations, including H-1 β /H₃-14, H-2/H-4, H-3 β /H₃-14, H-7/H₃-14, and H-3 β /H-15, the methyl group (C-14 and C-15), H-7, and the hydroxy group of C-2 were β -oriented. Thus, the relative configuration of compound 1 could be determined. In addition, a crystal of 6 was obtained using methanol: hexane (1:1) at 4 °C. The single-crystal X-ray crystallographic experiment was performed using Cu K α radiation, yielding a Flack parameter of 0.08 (11) (Figure 5E and Table S14; CCDC: 2346750), which indicated that the absolute configuration of 6 was 4*S*, 5*R*, 7*R*. Considering that 1

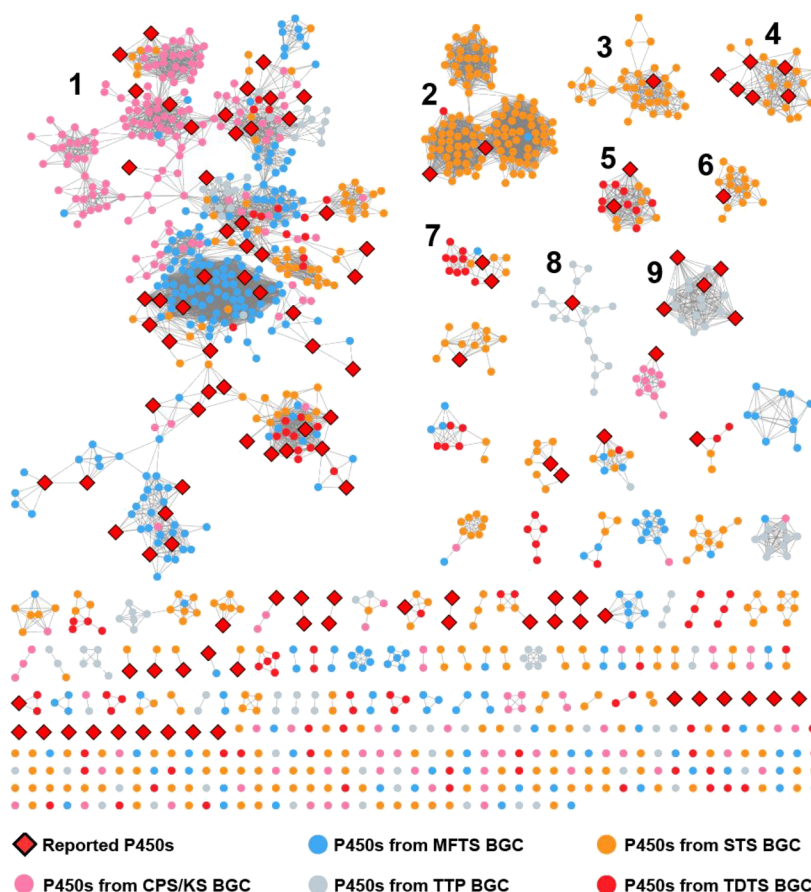


Figure 3. An SSN was constructed to depict the sequence relationships among reported fungal and extracted functionally uncharacterized terpenoid biosynthetic P450s, with an e-value cutoff set at 10^{-120} . In this network, reported P450s from both the MIBiG database and published literature are represented as red diamonds. Predicted P450s from diverse terpenoid BGCs are depicted as circles, each color-coded to represent a distinct type.

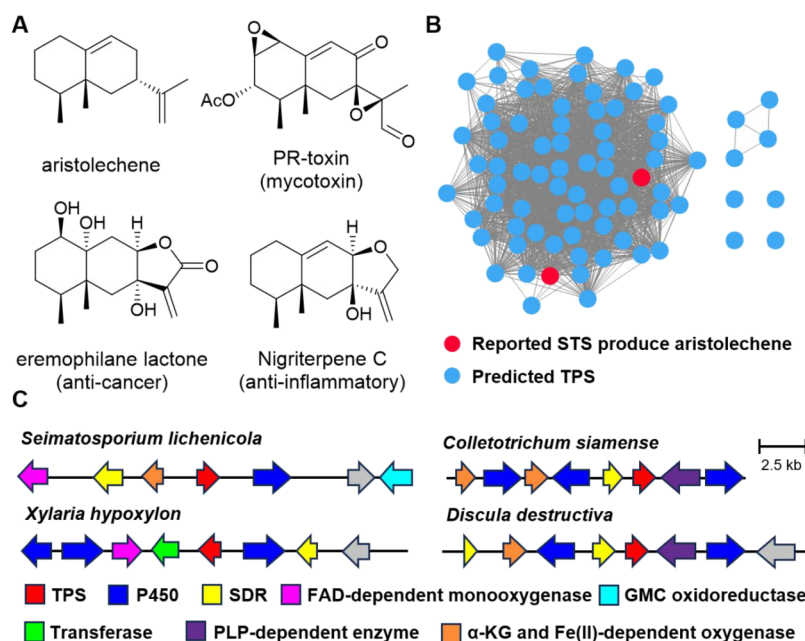


Figure 4. Genome mining to identify P450 enzymes potentially responsible for producing eremophilane-type sesquiterpenoids. (A) Structure and bioactivity of eremophilane-type sesquiterpenoids. (B) SSN analysis of TPSs from family 1, along with two previously reported STSs known to produce aristolechene using an e-value cutoff of 10^{-80} . (C) BGCs of P450s in family 2.

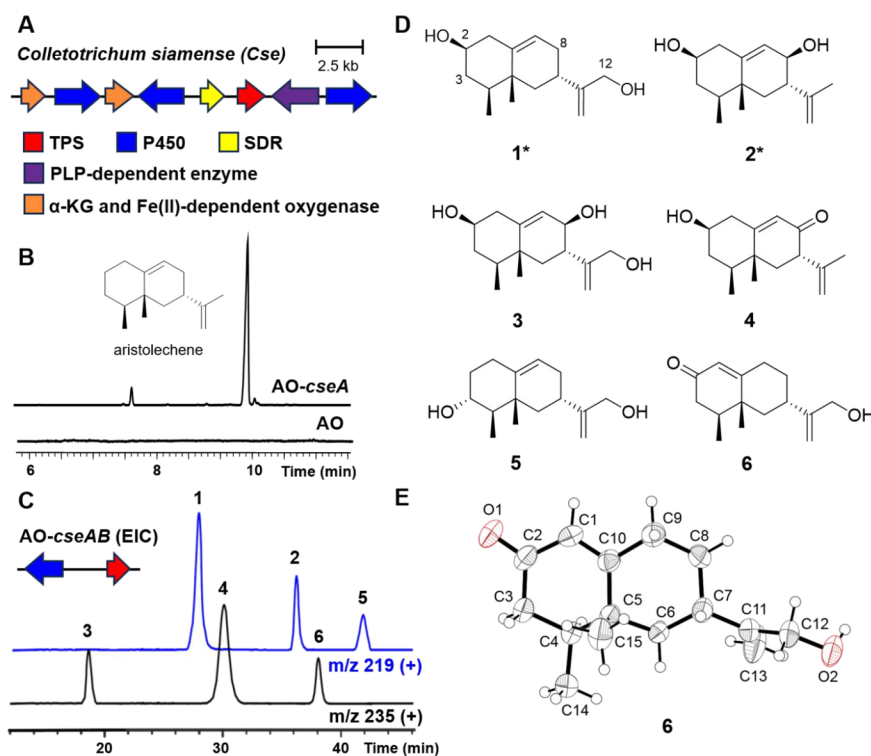


Figure 5. Heterologous expression of the selected P450 proves its capability to produce eremophilane-type sesquiterpenoids. (A) BGC in this study. (B) GC–MS profile of *A. oryzae* transformant containing STS *CseA*. (C) EIC profile of the *A. oryzae* transformant containing STS *CseA* and P450 *CseB*. (D) Structures of isolated eremophilane-type sesquiterpenoids 1–6 from AO-*CseAB*. The asterisk means the new compound. (E) X-ray crystal structure of 6.

shares the same biosynthetic pathway as 6, the absolute configuration of 1 was determined as 2*R*, 4*S*, 5*R*, 7*R* (Figures S12–S18 and S26, Table S15).

Compound 2 was isolated as a colorless oil, λ_{\max} = 198 nm. The molecular formula $C_{15}H_{25}O_2$ was deduced by LC-HRMS (m/z 219.1746 $[M-H_2O + H]^+$, calcd. for 219.1749), indicating four units of unsaturation. The 1D and 2D NMR signals of 2 closely resembled those of 1,²⁶ indicating that they share the same skeleton. The difference between these two compounds was that the hydroxy group at C-12 in compound 1 was absent and an additional hydroxy group was substituted at C-8, which could be confirmed by the 1H – 1H COSY correlations of H-8/H-9 and HMBC correlations of H₃-12/C-9, H₃-12/C-10, and H₃-12/C-11. Thus, the planar structure of 2 was fully constructed. Based on key NOESY correlations, H-8/H-13 indicated that the hydroxy group of C-2 was β -oriented. Based on the identified absolute configuration of 6, the absolute configuration of 2 was elucidated as 2*R*, 4*S*, 5*R*, 7*R*, 8*S* (Figures S19–S26, Table S16).

Compound 3 was isolated as a white powder. The molecular formula $C_{15}H_{25}O_3$ was deduced by LC-HRMS (m/z 235.1696 $[M-H_2O + H]^+$, calcd. for 235.1698), indicating four units of unsaturation and three units of hydroxylation. The NMR data are in good agreement with paraconiothins D in the previous report (Figures S27 and S28, Table S17).²⁶

Compound 4 was isolated as a white powder. HR-ESI-MS data of 4 revealed its molecular formula as $C_{15}H_{23}O_2$ (m/z 235.1696 $[M + H]^+$, calcd. for 235.1698), indicating five units of unsaturation, one unit of hydroxylation, and one carbonyl. The NMR data are in good agreement with acremeremophilanes G in the previous report (Figures S29 and S30, Table S17).²⁷

Compound 5 was isolated as a white powder. HR-ESI-MS data of 5 revealed its molecular formula as $C_{15}H_{25}O_2$ (m/z 219.1748 $[M-H_2O + H]^+$, calcd. for 219.1749), indicating four units of unsaturation, one unig hydroxylation, and one carbonyl. The NMR data are in good agreement with those of paraconiothins C in the previous report (Figures S31 and S32, Table S18).²⁶

Compound 6 was isolated as a white powder. HR-ESI-MS data of 6 revealed its molecular formula as $C_{15}H_{23}O_2$ (m/z 235.1696 $[M + H]^+$, calcd. = 235.1698), indicating five units of unsaturation and two units of hydroxylation. The NMR data are in good agreement with 12-hydroxynootkatone in the previous report (Figures S33 and S34, Table S18).²⁸

DISCUSSION

NPs are an important source for bioactive compounds, many of which have been widely applied in the pharmaceutical or agricultural industry. The biosyntheses of most NPs can be delineated into two steps. Initially, core enzymes catalyze simple building blocks to produce structurally diverse skeletons. Subsequently, these skeletons are enzymatically decorated by tailoring enzymes such as P450s, FAD-dependent monooxygenases, and methyltransferases, thereby further enriching the chemical and bioactivity diversity of NPs.²⁹ Recent advances in genome mining have revealed numerous novel terpenoids and biosynthetic enzymes, often facilitated by core enzymes. Genome mining of tailoring enzymes promises to unveil novel NPs and deepen our understanding of NP biosynthesis. While software products like antiSMASH and PRISM have been extensively utilized to predict BGCs and tailoring enzymes in fungal genomes, their use for the specific extraction of NP-related tailoring enzymes in extensive fungal

genomes is hindered by serious issues, such as lengthy processing times and low accuracy. This is because these software tools need to search BGCs across the entire genome and utilize fast, but low-accuracy, fungal gene prediction software. In this study, we introduced a new software program, NPtagM, designed for scanning NP-related tailoring enzymes from microorganism genomes. By optimizing the workflow and gene prediction software, NPtagM demonstrated a 3-fold increase in the number of predicted fungal biosynthetic P450s and an 8-fold reduction in runtime compared to antiSMASH. The development of NPtagM enables the identification of specific tailoring enzymes from extensive genomic databases, which could significantly accelerate the genome mining of tailoring enzymes. The tailoring enzymes extracted by NPtagM could be further dereplicated and analyzed based on sequence identity using SSN, phylogenetic trees, and CD-HIT. Besides tailoring enzymes, NPtagM provided their corresponding BGCs, which could be further analyzed by BiG-SCAPE, CORASON, and clinker. Based on the analysis of the sequences of tailoring enzymes and their BGCs, users can make preliminary predictions about their functions.

However, NPtagM faces two main challenges. The first is the trade-off between accuracy and speed. Currently, NPtagM utilizes AUGUSTUS or GlimmerHMM for gene prediction, but neither of these tools offers high accuracy or fast performance. Recent advancements in machine learning and deep learning have significantly improved genomics,³⁰ and we believe that new software leveraging these efficient algorithms could potentially address this issue. The second is the function prediction of extracted tailoring enzymes. Due to the sequence and function diversity as well as the large gap between functionally characterized and uncharacterized tailoring enzymes, understanding the sequence–function relationship of extensive tailoring enzymes still encounters a huge challenge. Recently, the rapid development of protein structure prediction and molecular dynamic simulation techniques has deepened our understanding of enzyme catalytic mechanisms.^{31,32} Combining these protein structure-based approaches may subsequently overcome the existing limitation.

P450s catalyze various reactions on the inert carbon–hydrogen skeleton of terpenoids, including hydroxylation, C–C bond cleavage, oxidative cyclization, etc. The modifications catalyzed by P450s significantly influence the structural and bioactivity diversity of terpenoids.^{3,6} Several papers have performed comprehensive bioinformatics analysis on bacterial terpenoid biosynthetic P450s.^{10,33} However, as the major microbial resource of terpenoids, the sequence and function diversity of fungal terpenoid biosynthetic P450s still remains untapped. In this study, utilizing NPtagM, we provide a comprehensive landscape of terpenoids' biosynthetic P450s. The SSN analysis classified all P450s into 321 families, with many families not clustered with reported ones. The sequence diversity of P450s holds great potential for the discovery of more novel terpenoids and reactions.

Eremophilane-type sesquiterpenoids constitute one of the largest fungal terpenoid families with diverse bioactivities. Some of these compounds have found applications in the pharmaceutical and food industries. For instance, PR-toxin, a highly oxygenated eremophilane-type sesquiterpenoid found in *Penicillium roqueforti*, poses significant challenges to food quality and safety due to its potent mycotoxin properties.³⁴ Another example is an eremophilane-type sesquiterpene lactone isolated from *Xylaria* sp. BCC 21097, which has

demonstrated moderate cytotoxic activities against various cancer cell lines (KB, MCF-7, and NCI-H187) with IC₅₀ values ranging from 3.8 to 21 μ M.³⁵ Additionally, nigriterpene C sourced from *Xylaria nigripes* has been shown to inhibit lipopolysaccharide-induced inducible nitric oxide production without inducing significant cytotoxic effects.³⁶ In this study, we characterized a new P450 capable of oxidizing C-2, C-3, C-8, and C-12 of the aristolechene skeleton to produce eremophilane-type sesquiterpenoids. Additionally, we identified a large P450 family, consisting of 117 members, that might produce eremophilane-type sesquiterpenoids. Further characterization of these P450s might lead to the discovery of more structurally diverse eremophilane-type sesquiterpenoids.

In summary, a new tool called NPtagM was developed, which significantly improves efficiency with a 3-fold increase in the number of predicted fungal biosynthetic P450s and an 8-fold reduction in runtime compared to existing methods. Using NPtagM, we extracted 1189 dereplicated P450s involved in terpenoid biosynthesis from fungal genomes, identifying a family potentially producing eremophilane-type sesquiterpenoids. Heterologous expression of a P450 and the corresponding sesquiterpene synthase resulted in the production of two new eremophilanes in *A. oryzae*, showcasing that NPtagM has potential for genome mining in NP biosynthesis from microorganisms.

■ ASSOCIATED CONTENT

Data Availability Statement

The source code of NPtagM is available on GitHub (<https://github.com/ZhanRenCong/NPtagM>).

Supporting Information


The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jafc.4c07653>.

Function prediction of P450s from large and medium families by integrating the TPSSs and BGCs analysis (Figures S1–S10); protein sequence alignment of CseA and other STSs (Figure S11); ¹H (600 MHz) and ¹³C (150 MHz) NMR spectra of 1–6 and ¹H–¹H COSY, HMBC, and HSQC NMR spectra of compounds 1 and 2 (Figures S12–S32); the P450s predicted by NPtagM and antiSMASH (Tables S1–S6); reported TPSSs used in the construction of pHMM (Table S7); MIBiG terpenoid BGCs containing P450s (Table S8); reported fungal terpenoid biosynthetic P450s (Table S9); primers used in this study (Table S10); plasmids and transformants constructed in this study (Table S11); performance of NPtagM compared to antiSMASH (Table S12); annotation of Cse BGC (Table S13); X-ray crystallographic data for 6 (Table S14); and NMR data of 1–6 (Tables S15–18) (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Lan Jiang – Department of Cardiothoracic Surgery, Children's Hospital of Nanjing Medical University, Nanjing 210093, China; Email: jianglan0426@163.com

Xueting Liu – State Key Laboratory of Bioreactor Engineering, East China University of Science of Technology, Shanghai 200237, China;  orcid.org/0000-0002-1322-8253; Email: liuxueting@ecust.edu.cn

Authors

Zhanren Cong – State Key Laboratory of Bioreactor Engineering, East China University of Science of Technology, Shanghai 200237, China

Njeru Joe Mukoma – State Key Laboratory of Bioreactor Engineering, East China University of Science of Technology, Shanghai 200237, China

Qiang Yin – State Key Laboratory of Bioreactor Engineering, East China University of Science of Technology, Shanghai 200237, China

Bin Zhu – State Key Laboratory of Bioreactor Engineering, East China University of Science of Technology, Shanghai 200237, China; Engineering Research Centre of Pharmaceutical Process Chemistry, Ministry of Education, and Laboratory of Pharmaceutical Crystal Engineering & Technology, School of Pharmacy, East China University of Science and Technology, Shanghai 200237, China;

orcid.org/0000-0003-1085-1231

Lingwei She – Key Laboratory of Molecular Medicine and Biotherapy, School of Life Science, Beijing Institute of Technology, 100081 Beijing, China

Tom Hsiang – School of Environmental Sciences, University of Guelph, Guelph, Ontario N1G2W1, Canada

Lixin Zhang – State Key Laboratory of Bioreactor Engineering, East China University of Science of Technology, Shanghai 200237, China

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.jafc.4c07653>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Science and Technology Commission of Shanghai Municipality (21NL2600100), National Key Research and Development Program of China (2020YFA090032 and 2019YFA0906200), the National Natural Science Foundation of China (21977029 and 21907031), and the Open Project Funding of the State Key Laboratory of Bioreactor Engineering, the 111 Project (B18022). Furthermore, we would like to acknowledge the Natural Science and Engineering Research Council of Canada for funding the genome sequencing and assembly of fungal genomes under the supervision of Prof. T. Hsiang.

REFERENCES

- (1) Zeng, T.; Liu, Z.; Zhuang, J.; Jiang, Y.; He, W.; Diao, H.; Lv, N.; Jian, Y.; Liang, D.; Qiu, Y.; Zhang, R.; Zhang, F.; Tang, X.; Wu, R. TeroKit: A Database-Driven Web Server for Terpenome Research. *J. Chem. Inf. Model.* **2020**, *60* (4), 2082–2090.
- (2) Avalos, M.; Garbeva, P.; Vader, L.; van Wezel, G. P.; Dickschat, J. S.; Ulanova, D. Biosynthesis, evolution and ecology of microbial terpenoids. *Nat. Prod. Rep.* **2022**, *39* (2), 249–272.
- (3) Yin, Z.; Dickschat, J. S. Engineering fungal terpene biosynthesis. *Nat. Prod. Rep.* **2023**, *40* (1), 28–45.
- (4) Kautsar, S. A.; Blin, K.; Shaw, S.; Navarro-Muñoz, J. C.; Terlouw, B. R.; van der Hooft, J. J. J.; van Santen, J. A.; Tracanna, V.; Suarez Duran, H. G.; Pascal Andreu, V.; Selem-Mojica, N.; Alanjary, M.; Robinson, S. L.; Lund, G.; Epstein, S. C.; Sisto, A. C.; Charkoudian, L. K.; Collemare, J.; Linington, R. G.; Weber, T.; Medema, M. H. MIBiG 2.0: a repository for biosynthetic gene clusters of known function. *Nucleic Acids Res.* **2020**, *48* (D1), D454–D458.
- (5) Zhang, X.; Guo, J.; Cheng, F.; Li, S. Cytochrome P450 enzymes in fungal natural product biosynthesis. *Nat. Prod. Rep.* **2021**, *38* (6), 1072–1099.
- (6) Zhao, Y.; Liang, F.; Xie, Y.; Duan, Y. T.; Andeadelli, A.; Pateraki, I.; Makris, A. M.; Pomorski, T. G.; Staerk, D.; Kampranis, S. C. Oxetane ring formation in taxol biosynthesis is catalyzed by a bifunctional cytochrome P450 enzyme. *J. Am. Chem. Soc.* **2024**, *146* (1), 801–810.
- (7) Rao, L.; Yuan, G.-Y.; Chen, X.-Y.; Ran, J.-L.; Zou, Y. Reshaping the Diversity of Oxidized Polyquinane Sesquiterpenoids by Cytochrome P450s. *Org. Lett.* **2023**, *25* (18), 3276–3280.
- (8) Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S. Y.; Medema, M. H.; Weber, T. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* **2019**, *47* (W1), W81–W87.
- (9) Wei, B.; Du, A. Q.; Zhou, Z. Y.; Lai, C.; Yu, W. C.; Yu, J. B.; Yu, Y. L.; Chen, J. W.; Zhang, H. W.; Xu, X. W.; Wang, H. An atlas of bacterial secondary metabolite biosynthesis gene clusters. *Environ. Microbiol.* **2021**, *23* (11), 6981–6992.
- (10) Rudolf, J. D.; Chang, C. Y.; Ma, M.; Shen, B. Cytochromes P450 for natural product biosynthesis in *Streptomyces*: sequence, structure, and function. *Natural product reports* **2017**, *34* (9), 1141–1172.
- (11) Hyatt, D.; Chen, G. L.; LoCascio, P. F.; Land, M. L.; Larimer, F. W.; Hauser, L. J. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinf.* **2010**, *11*, 119.
- (12) Majoros, W. H.; Pertea, M.; Salzberg, S. L. TigrScan and GlimmerHMM: two open source ab initio eukaryotic gene-finders. *Bioinformatics* **2004**, *20* (16), 2878–2879.
- (13) Stanke, M.; Diekhans, M.; Baertsch, R.; Haussler, D. Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics* **2008**, *24* (5), 637–644.
- (14) Potter, S. C.; Luciani, A.; Eddy, S. R.; Park, Y.; Lopez, R.; Finn, R. D. HMMER web server: 2018 update. *Nucleic Acids Res.* **2018**, *46* (W1), W200–W204.
- (15) Camacho, C.; Boratyn, G. M.; Joukov, V.; Vera Alvarez, R.; Madden, T. L. ElasticBLAST: accelerating sequence search via cloud computing. *BMC Bioinf.* **2023**, *24* (1), 117.
- (16) Fu, L.; Niu, B.; Zhu, Z.; Wu, S.; Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **2012**, *28* (23), 3150–3152.
- (17) Mistry, J.; Chuguransky, S.; Williams, L.; Qureshi, M.; Salazar, G. A.; Sonnhammer, E. L. L.; Tosatto, S. C. E.; Paladin, L.; Raj, S.; Richardson, L. J.; Finn, R. D.; Bateman, A. Pfam: The protein families database in 2021. *Nucleic Acids Res.* **2021**, *49* (D1), D412–D419.
- (18) Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13* (11), 2498–2504.
- (19) Nagamine, S.; Liu, C.; Nishishita, J.; Kozaki, T.; Sogahata, K.; Sato, Y.; Minami, A.; Ozaki, T.; Schmidt-Dannert, C.; Maruyama, J. i.; Oikawa, H.; Master, E. R. Ascomycete *Aspergillus oryzae* is an efficient expression host for production of basidiomycete terpenes by using genomic DNA sequences. *Appl. Environ. Microbiol.* **2019**, *85* (15), e00409–e00419.
- (20) Hidalgo, P. I.; Poirier, E.; Ullan, R. V.; Piqueras, J.; Meslet-Cladiere, L.; Coton, E.; Coton, M. *Penicillium roqueforti* PR toxin gene cluster characterization. *Appl. Microbiol. Biotechnol.* **2017**, *101* (5), 2043–2056.
- (21) Hou, C.; Kulka, M.; Zhang, J.; Li, Y.; Guo, F. Occurrence and biological activities of eremophilane-type sesquiterpenes. *Mini Rev. Med. Chem.* **2014**, *14*, 664–677.
- (22) Yuyama, K. T.; Fortkamp, D.; Abraham, W. R. Eremophilane-type sesquiterpenes from fungi and their medicinal potential. *Biol. Chem.* **2017**, *399* (1), 13–28.
- (23) Proctor, R. H.; Hohn, T. M. Aristolochene synthase. Isolation, characterization, and bacterial expression of a sesquiterpenoid biosynthetic gene (Ari1) from *Penicillium roqueforti*. *J. Biol. Chem.* **1993**, *268* (6), 4543–4548.

- (24) Cane, D. E.; Kang, I. Aristolochene synthase: purification, molecular cloning, high-level expression in *Escherichia coli*, and characterization of the *Aspergillus terreus* cyclase. *Arch. Biochem. Biophys.* **2000**, 376 (2), 354–364.
- (25) Hidalgo, P. I.; Ullan, R. V.; Albillos, S. M.; Montero, O.; Fernandez-Bodega, M. A.; Garcia-Estrada, C.; Fernandez-Aguado, M.; Martin, J. F. Molecular characterization of the PR-toxin gene cluster in *Penicillium roqueforti* and *Penicillium chrysogenum*: cross talk of secondary metabolite pathways. *Fungal Genet Biol.* **2014**, 62, 11–24.
- (26) Nakashima, K.-i.; Tomida, J.; Hirai, T.; Kawamura, Y.; Inoue, M. Paraconiothins A–J: Sesquiterpenoids from the Endophytic Fungus *Paraconiothyrium brasiliense* ECN258. *J. Nat. Prod.* **2019**, 82 (12), 3347–3356.
- (27) Cheng, Z.; Zhao, J.; Liu, D.; Proksch, P.; Zhao, Z.; Lin, W. Eremophilane-Type Sesquiterpenoids from an *Acremonium* sp. Fungus Isolated from Deep-Sea Sediments. *J. Nat. Prod.* **2016**, 79 (4), 1035–1047.
- (28) Xu, J.; Su, J.; Li, Y.; Tan, N. Eremophilane-type sesquiterpenes from *Alpinia oxyphylla* with inhibitory activity against nitric oxide production. *Chem. Nat. Compd.* **2013**, 49, 457–461.
- (29) Chiang, C. Y.; Ohashi, M.; Tang, Y. Deciphering chemical logic of fungal natural product biosynthesis through heterologous expression and genome mining. *Nat. Prod Rep* **2023**, 40 (1), 89–127.
- (30) Eraslan, G.; Avsec, Z.; Gagneur, J.; Theis, F. J. Deep learning: new computational modelling techniques for genomics. *Nat. Rev. Genet* **2019**, 20 (7), 389–403.
- (31) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zidek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, 596 (7873), 583–589.
- (32) Lou, T.; Li, A.; Xu, H.; Pan, J.; Xing, B.; Wu, R.; Dickschat, J. S.; Yang, D.; Ma, M. Structural insights into three sesquiterpene synthases for the biosynthesis of tricyclic sesquiterpenes and chemical space expansion by structure-based mutagenesis. *J. Am. Chem. Soc.* **2023**, 145 (15), 8474–8485.
- (33) Lin, G.-M.; Voigt, C. A. Design of a redox-proficient *Escherichia coli* for screening terpenoids and modifying cytochrome P450s. *Nature Catalysis* **2023**, 6 (11), 1016–1029.
- (34) Dubey, M. K.; Aamir, M.; Kaushik, M. S.; Khare, S.; Meena, M.; Singh, S.; Upadhyay, R. S. PR toxin - biosynthesis, genetic regulation, toxicological potential, prevention and control measures: overview and challenges. *Front. Pharmacol.* **2018**, 9, 288.
- (35) Isaka, M.; Chinthanom, P.; Boonruangprapa, T.; Rungjindamai, N.; Pinruan, U. Eremophilane-type sesquiterpenes from the fungus *Xylaria* sp. BCC 21097. *J. Nat. Prod* **2010**, 73, 683–687.
- (36) Chang, J.-C.; Hsiao, G.; Lin, R.-K.; Kuo, Y.-H.; Ju, Y.-M.; Lee, T.-H. Bioactive constituents from the termite nest-derived medicinal fungus *Xylaria nigripes*. *J. Nat. Prod* **2017**, 80 (1), 38–44.