Genome Mining of Fungal Unique Trichodiene Synthase-like Sesquiterpene Synthases

Zhanren Cong [1], Qiang Yin [1], Kunhong Tian [1], Njeru Joe Mukoma [1], Liming Ouyang [1], Tom Hsiang [2], Lixin Zhang [1], Lan Jiang [3,*] and Xueting Liu [1,*]

[1] State Key Laboratory of Bioreactor Engineering, East China University of Science of Technology, Shanghai 200237, China; joemukosh@gmail.com (N.J.M.)

[2] School of Environmental Sciences, University of Guelph, 50 Stone Road East, Guelph, ON N1G 2W1, Canada

[3] Department of Cardiothoracic Surgery, Children's Hospital of Nanjing Medical University, Nanjing 210093, China

* Correspondence: jianglan0426@163.com (L.J.); liuxueting@ecust.edu.cn (X.L.)

Abstract: Sesquiterpenoids served as an important source for natural product drug discovery. Although genome mining approaches have revealed numerous novel sesquiterpenoids and biosynthetic enzymes, the comprehensive landscape of fungal sesquiterpene synthases (STSs) remains elusive. In this study, 123 previously reported fungal STSs were subjected to phylogenetic analysis, resulting in the identification of a fungi-specific STS family known as trichodiene synthase-like sesquiterpene synthases (TDTSs). Subsequently, the application of hidden Markov models allowed the discovery of 517 TDTSs from our in-house fungi genome library of over 400 sequenced genomes, and these TDTSs were defined into 79 families based on a sequence similarity network. Based on the novelty of protein sequences and the completeness of their biosynthetic gene clusters, 23 *TDTS* genes were selected for heterologous expression in *Aspergillus oryzae*. In total, 10 TDTSs were active and collectively produced 12 mono- and sesquiterpenes, resulting in the identification of the first chamipinene synthase, as well as the first fungi-derived cedrene, sabinene, and camphene synthases. Additionally, with the guidance of functionally characterized TDTSs, we found that TDTSs in Family 1 could produce bridged-cyclic sesquiterpenes, while those in Family 2 could synthesize spiro- and bridged-cyclic sesquiterpenes. Our research presents a new avenue for the genome mining of fungal sesquiterpenoids.

Keywords: fungi; genome mining; sesquiterpenes; monoterpenes; sesquiterpene synthase; trichodiene synthase

1. Introduction

Sesquiterpenoids, with over 40,000 compounds, represent one of the largest and most structurally diverse natural product families [1,2]. Their complex and diverse backbones are derived from farnesyl diphosphate (FPP) as a precursor and catalyzed by sesquiterpene synthases (STSs) [3]. These compounds exhibit a wide spectrum of bioactivities, such as antibacterial, anti-inflammatory, and cytotoxic effects, and have a wide application in the pharmaceutical industry (such as artemisinin and bilobalide), food and flavor industry (such as nootkatone and valencene), as well as biofuels (such as farnesene), with a promising economic value [4].

Fungi represent a rich source of diverse sesquiterpenes with pharmaceutical and agricultural relevance [5], including illudin M (cytotoxic), abscisic acid (plant growth regulation), and penifulvin (insecticidal) [6]. In the past decade, over 150 fungal STSs have been functionally characterized through genome mining, resulting in the discovery of numerous novel sesquiterpenes and enzymes [7]. However, the genome mining of fungal STSs has been limited to a small number of species, such as *Coprinopsis cinerea*, *Omphalotus olearius*, and *Stereum hirsutum* [7–11]. The sequence diversity, abundance, and distribution

of fungal STSs on a global scale remain elusive. This limitation may lead to the oversight of novel sesquiterpene structures or the rediscovery of similar-function STSs.

Trichodiene synthase-like sesquiterpene synthase (TDTS) is a fungi-specific STS family that may catalyze the biosynthesis of complex and diverse sesquiterpenes through the bisabolene carbocation intermediate. For example, the representative trichodiene is post-modified to generate well-known trichothecene mycotoxins such as deoxynivalenol and T-2 toxin [12]. Furthermore, they can also catalyze the production of longiborneol, barbatene, and acorodiene with complex bridged or spiro structures. Two of them can catalyze the generation of monoterpenes, with only three fungal monoterpene synthases (MTSs) previously reported (Figure 1). Currently, there are only 28 reported TDTSs, the uniqueness of which suggests that further exploration may reveal more unexpected terpenoids.
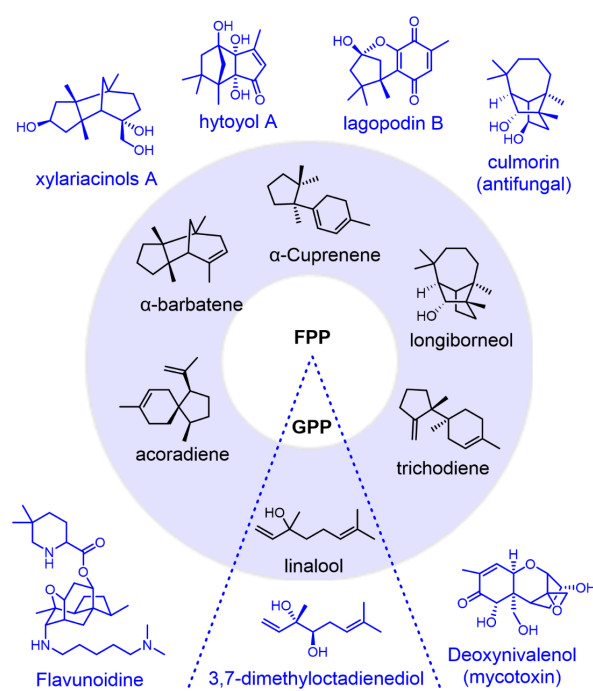

Figure 1. The terpene skeletons are produced by fungal TDTSs and their modified terpenoids.

Verification of the relationship between the sequences and functions of STSs could improve the discovery of new sesquiterpenes. In this study, the 123 previously reported fungal STSs were subjected to phylogenetic analysis, resulting in the identification of a fungus-specific STS family known as TDTSs. The application of hidden Markov models (HMMs) allowed the discovery of 517 TDTSs from our in-house fungi genome library. Subsequently, we constructed a sequence similarity network (SSN) to partially reveal the distribution of this special STS family, which allowed us to define all known TDTSs into 79 families based on sequence identity. Combined with the heterologous expression in *Aspergillus oryzae*, the STS responsible for the biosynthesis of champinene was identified for the first time, as were three key enzymes involved in the biosynthesis of cedrene, sabinene, and camphene, all characterized by fungi. Additionally, with the help of functionally characterized TDTSs, we found that TDTSs in Family 1 could produce bridged cyclic sesquiterpenes, while those in Family 2 could synthesize spiro and bridged cyclic sesquiterpenes. This study provides a foundation for the development of fungal TDTSs.

2. Materials and Methods

2.1. Strains and Media

Strains of *A. oryzae* NSAR1 (*niaD−*, *sC−*, *adeA−*, Δ*argB*) were used as the host for gene expression. The *A. oryzae* NSAR1 wildtype strain was cultivated in DPY medium for protoplast preparation (dextrose 20 g/L, polypeptone 10 g/L, yeast extract 5 g/L). *A. oryzae*

NSAR1 transformants were cultivated in MA medium for transformant selection (dextrose 20 g/L, sorbitol 218.6 g/L, NH_4Cl 2 g/L, $(NH_4)_2SO_4$ 1 g/L, KCl 0.5 g/L, NaCl 0.5 g/L, KH_2PO_4 1 g/L, $MgSO_4 \cdot 7H_2O$ 0.5 g/L, $FeSO_4 \cdot 7H_2O$ 0.02 g/L, Methionine 1.5 g/L, Adenine 0.1 g/L, agar 15 g/L, pH 5.5). *A. oryzae* transformants were also cultivated in MPY medium for fermentation (maltose 30 g/L, polypeptone 10 g/L, yeast extract 5 g/L, and $(NH_4)_2SO_4$ 0.925 g/L). *Escherichia coli* DH10b was used for gene cloning. *E. coli* transformants were cultivated in LB medium (tryptone 10 g/L, yeast extract 5 g/L, and NaCl 10 g/L).

2.2. Collection and Phylogenetic Analysis of Reported Fungal STSs

A total of 123 reported fungal STS sequences were collected from published papers and public databases (Genebank [http://www.ncbi.nlm.nih.gov/ (accessed on 8 December 2023)], UniProtKB [http://www.uniprot.org/ (accessed on 08/12/2023)], and JGI [http://www.jgi.doe.gov/ (accessed on 8 December 2023)]). The accession number, gene name, organism, products, and cyclization pathway of all reported STSs are summarized in Tables S1 and S2. Fungal STSs were labeled as germacrene cyclization intermediate derived, humulene cyclization intermediate derived, or bisabolene cyclization intermediate derived according to their products and cyclization intermediates. The cyclization intermediates for each sesquiterpene were determined using the scheme detailed in IUBMB's Enzyme Nomenclature Supplement 24 (2018) [13]. The cyclization intermediates for STSs with multiple products were determined by their major products. The sequence alignment of 123 previously reported TDTSs was made in Clustal Omega 1.2.2 with default parameters [14]. Phylogenetic analysis of fungal STSs was performed in MEGA7 [15] using the maximum-likelihood method with the Jones–Taylor–Thornton (JTT) model, and default parameters were used. The phylogenetic tree was rendered using iTOL 6.9 [16].

2.3. Extraction and SSN Analysis of TDTSs from Fungi Genome Database

A profile hidden Markov model (pHMM) was built with default parameters using 27 reported TDTSs to screen TDTSs from the in-house fungal genome database (>450 genomes) with the default parameters (Table S1), resulting in the extraction of a total of 1513 sequences [17]. To exclude the STSs belonging to the terpene synthase (TPTS) family from these sequences, we built another pHMM using 98 reported TPSs (Table S2). Each sequence extracted from the in-house genome database was scored by two HMM profiles using HM-Mer 3.0. The 965 TDTS sequences were selected with higher scores on TDTS pHMM. These 965 sequences were further de-replicated by CD-Hit v4.8.1 with a sequence identity cut-off set at 0.8, resulting in 517 TDTS sequences. TDTSs are then located in their respective genomes using BLAST [18], followed by the extraction of the biosynthetic gene clusters (BGCs) within 10 kb upstream and downstream of the identified TDTSs. The genes on BGCs were predicted by AUGUSTUS 3.5.0 [19], with *Aspergillus oryzae* as the reference organism. The sequence identity of these predicted TDTSs, together with the reported ones, was calculated using BLAST 2.2.30 to generate SSN profiles with default parameters. SSNs were then visualized using Cytoscape 3.6.0 with an e-value cut-off set at 10^{-75} [20]. The alignment of all selected TDTSs and reported TDTSs was completed using Clustal Omega 1.2.2 with default parameters and visualized with Jalview 2.11.3.2 [21]. The visualization of the HMM logos of TDTS and TPTS pHMM was finished using Skyline [22] with default parameters.

2.4. Construction of Expression Plasmids

The 8 TDTS genes were synthesized by GenScript Biotech (Shanghai, China). The other 15 TDTS genes were amplified from the gDNA of corresponding wild strains with primers shown in Tables S3 and S4. PCR reactions were performed with the Q5® High-Fidelity DNA Polymerase (NEB biotech, Ipswich, MA, USA). pUARA4 was digested utilizing *Kpn*I (NEB biotech). The ClonExpress®II One Step Cloning Kit (Vazyme Biotech, Nanjing, China) was used to construct expression plasmids by inserting the PCR product of different TDTS genes into the *Kpn*I restriction sites of pUARA4 through homologous recombination to produce heterologous expression plasmids, pUARA4-TDTS. The single clones were manually picked

out and utilized as the template for PCR using $2 \times$ Taq Plus Master Mix (Vazyme) with the primers listed in Table S4. Positive clones were sequenced by Tsingke Biotech to confirm the gene was inserted into the right restriction site with no mutations. Plasmid DNA was extracted using the Plasmid Mini Kit I (OMEGA, Norcross, GA, USA). The active *TDTS* sequences were deposited in the National Center for Biotechnology Information under accession number PP516626-35. The 13 inactive *TDTS* sequences deposited in the National Center for Biotechnology Information under accession number PP776614-26

2.5. Preparation of Protoplasts and Transformation of A. oryzae

A. oryzae was selected as the heterologous expression host, and the protoplast–polyethylene glycol method was used for the transformant's construction. Spores of *A. oryzae* NSAR1 (1.0×10^8 cells) were placed into DPY medium (50 mL) in a 250 mL flask and shaken at 220 rpm at 30 °C for 2 days. Hyphae were collected through a filter, followed by washing with a 0.8 M NaCl solution 2–3 times and drying. We then dissolve 100 mg of Yatalase (TAKARA) in 20 mL of Solution 0 (0.8 M NaCl, 10 mM KH_2PO_4) and filter the solvent to sterilize. We then dissolved 100 mg of yatalase (TAKARA, Kusatsushi, Japan) in 20 mL of Solution 0 (0.8 M NaCl, 10 mM KH_2PO_4), and the liquid was filtered through a 0.22 μm pore size filter to sterilize. Then, the hyphae were placed in Yatalase solution and incubated at 30 °C and 220 rpm for 2–3 h until most of the hyphae disintegrated. The protoplasts were collected by filtering and washing with miracloth (Merck, Rahway, NJ, USA) and washed with a 0.8 M NaCl solution three times. Finally, the concentration of protoplasts was adjusted to 1.0×10^8 cells/mL with solution 2 (Sorbitol 218.6 g/L, $CaCl_2 \cdot 2H_2O$ 7.35 g/L, NaCl 2.05 g/L, 1 M Tris-HCl 10 mL/L). Then, 200 μL prepared protoplasts and 10 μg heterologous expression plasmid were mixed, followed by incubating on ice for 20 min. Then 1 mL Solution 3 (PEG 600 g/L, $CaCl_2 \cdot 2H_2O$ 7.35 g/L, 1 M Tris-HCl 10 mL/L) was added, followed by incubation at room temperature for 20 min. Then, 10 mL Solution 2 was added, followed by centrifugation at 4 °C at $800 \times g$ for 10 min. The supernatant was discarded, and the precipitation was resuspended in 1 mL of Solution 2. The supernatant was spread onto MA medium and cultured at 30 °C for 3–5 days. After 3–5 days, single colonies were picked and transferred to new MA medium plates. After 2–3 rounds of transfer, these AO (*A. oryzae*) transformants were verified by picking up single colonies manually and suspending them in 100 μL NaOH solution (25 mM) and lysed through heating at 100 °C for 10 min. The supernatant was used as the template for PCR using $2 \times$ Taq Plus Master Mix (Vazyme) with the primers listed in Table S4. At least three positive transformants were selected for fermentation and metabolite analysis. A series of AO-*TDTS* transformants (harboring *TDTS*) were constructed utilizing this method [10]; details are summarized in Table S5.

2.6. Fermentation and Analysis of the Metabolites from A. oryzae Transformants

A. oryzae transformants were inoculated into MPY medium (5 mL) in a 20 mL headspace screw-top vial and incubated at 30 °C and 220 rpm for 3 days. The volatile compounds produced by *A. oryzae* transformants were extracted with a solid-phase microextraction (SPME) fiber (50/30 μm DVB/CAR/PDMS; Stableflex, 24Ga, 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) gas chromatography–mass spectrometry (GC–MS) apparatus with a DB5 MS capillary column (0.25 mm × 30 m, 0.25 μm film thickness, SHIMADZU) for 10 s. The SPME fiber was maintained in the injection port to avoid contamination. The temperature of the injection port was set at 250 °C, and the GC–MS program was started in splitless mode. The column temperature was increased at 30 °C/min from 60 °C to 120 °C, increased at 5 °C/min to 180 °C, then increased at 30 °C/min to 270 °C. The flow rate of the helium carrier gas was 1.41 mL/min. The *m/z* detection range was set at 50–250. The structures of compounds detected by GC–MS were identified by comparing their mass spectra with data on spectra of terpenoids at the National Institute of Standards and Technology (NIST 14) standard reference database.

3. Results

3.1. Identification of TDTSs through Phylogenetic Analysis of Fungal STSs

To provide an overview of fungal STSs, phylogenetic analysis of the 123 previously reported was performed, resulting in clustering into clades A and B, containing 96 and 27 STSs, respectively (Figure 2A). Further protein family analysis showed that STSs in Clade A had a complete domain of the terpene synthase C (PF19086), while STSs in Clade B did not contain the C domain but had a very special trichodiene synthase domain (PF06330), which has only been reported in fungi (Mistry et al., 2021) (Figure 2B). In addition to protein family differences, the phylogenetic analysis also supported the function divergencies of the STSs from two clades. STSs in Clade A catalyze cyclization through germacrene (C1–C10) and humulene (C1–C11) carbocation intermediates to produce sesquiterpenes such as cadinene, viridiflorene, and protoilludene. However, most STSs from Clade B can catalyze cyclization through a bisabolene (C1–C6) carbocation intermediate to yield sesquiterpenes such as trichodiene, acoradiene, and barbatene (Figure 2C). Due to their sequence and functional differences, we named STSs from Clade A and Clade B as terpene synthase-like sesquiterpene synthases (TPTSs) and trichodiene synthase-like sesquiterpene synthases (TDTSs), respectively.
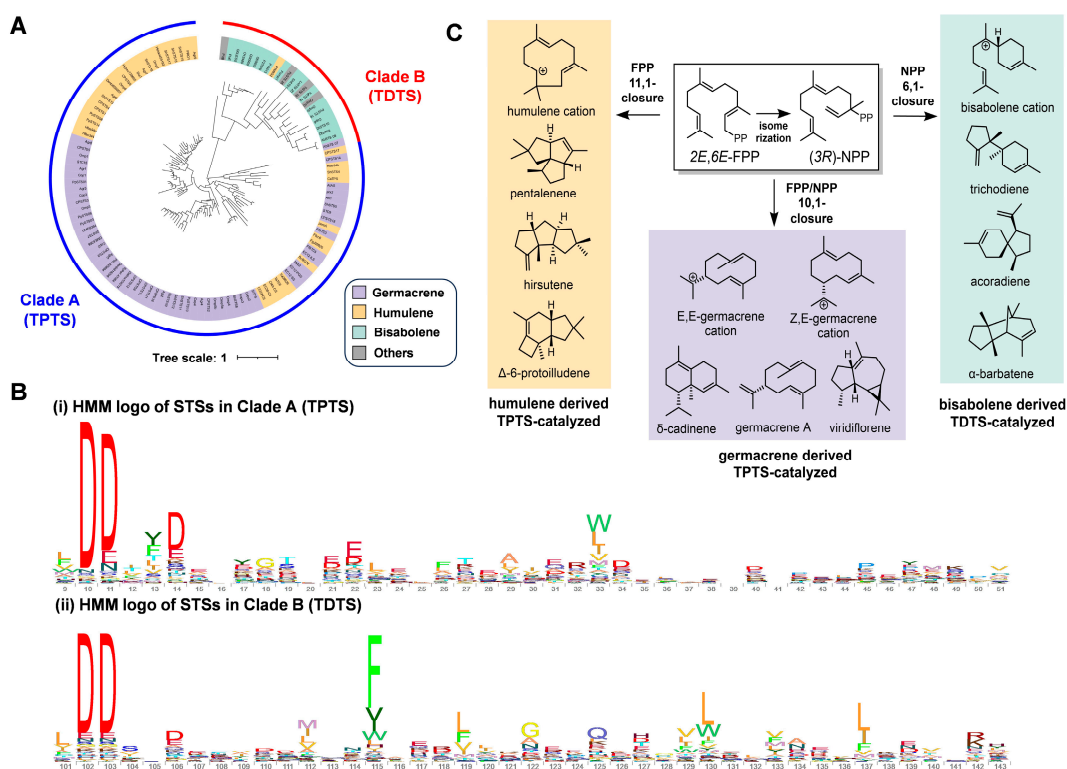

Figure 2. The phylogenetic analysis of reported fungal STSs revealed a fungi-specific TDTS family. (**A**) The phylogenetic analysis of fungal STSs classified them into two families: TPTS family and TDTS family. (**B**) HMM logos of TDTS and TPTS protein families. (**C**) Fungal STSs from two clades catalyzed cyclization through three different pathways.

3.2. Extraction and SSN Analysis of TDTSs from a Fungal Genome Database

A comprehensive map was constructed to allow focus on TDTSs with biosynthetic or taxonomic significance. From our in-house fungal genome library, which comprises 430 fungal genomes, 517 TDTS sequences were extracted through two rounds of profile Hidden Markov Model (pHMM) screening. The first round used the TDTS pHMM profile generated from 27 functionally characterized TDTSs to extract potential TDTS sequences (Table S1), while the second round utilized the TPTS pHMM profile generated from 98 functionally characterized TPTSs to exclude the TPTS sequences extracted by the

TDTS pHMM (Table S2), and then CD-HIT was used to remove sequences with sequence similarity >80%, resulting in the identification of 517 de-duplicated TDTS sequences. Analysis of genome sequence revealed that TDTS was widely distributed in approximately 54% of species in our in-house 430 members' fungal genome library. Overall, Basidiomycetes account for the majority, with 58% of TDTSs deriving from Basidiomycetes (305/517) and 42% from Ascomycetes (212/517) (Figures S1 and S2).

BLAST was used to compute a sequence identity distance matrix for the 517 predicted TDTSs along with the 27 previously reported ones. Utilizing a sequence similarity network with an e-value cut-off set at 10^{-75}, the above 545 TDTSs were further defined into 79 families based on sequence identity, giving 2 large families (Families 1 and 3, each with more than 50 TDTSs), 9 medium-sized families (Families 2 and 4–11, each containing 10–50 TDTSs), 10 small families (each with 3–10 TDTSs), and 71 mini families (each with 1–3 TDTSs) (Figure 3). The SSN analysis revealed distinct sequence differences between TDTSs from Ascomycota and Basidiomycota. Most Ascomycota-derived TDTSs (140 out of 212) were classified into Families 1 and 2, whereas most Basidiomycota-derived (228 out of 305) TDTSs were placed in 8 other families (Families 3–10).
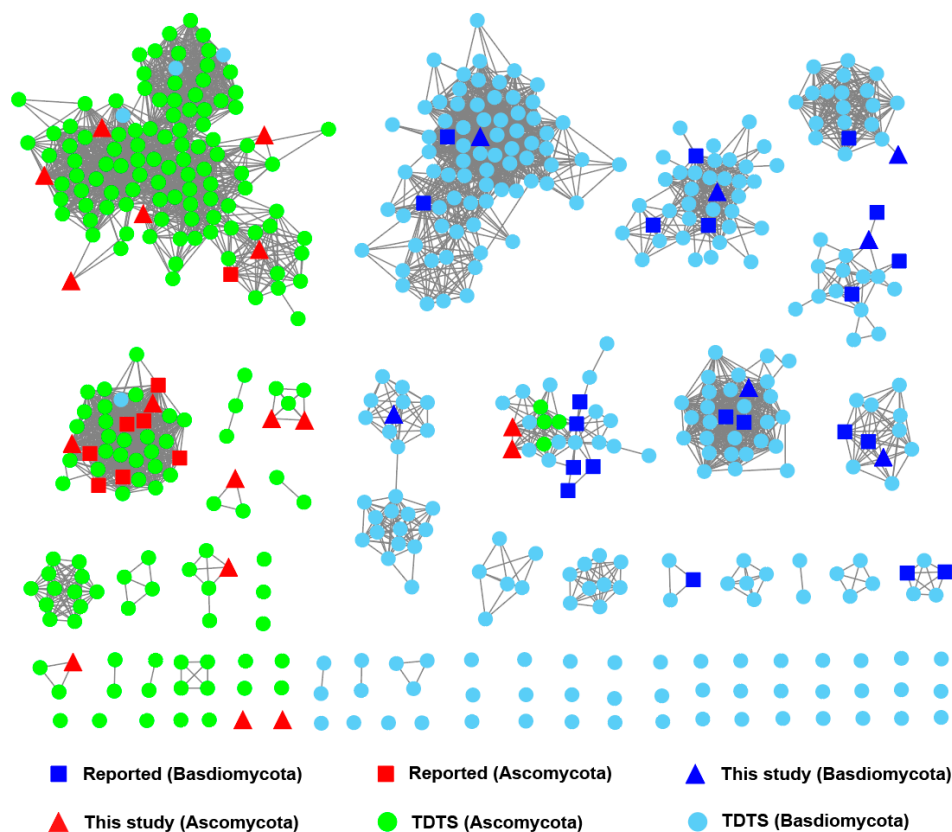

Figure 3. A sequence similarity network of previously reported and extracted fungal TDTSs was constructed with an e-value cut-off set at 10^{-75}.

3.3. Heterologous Expression of TDTSs

These criteria were established for the selection of TDTS for heterologous expression to characterize their function: (1) had a distant evolutionary relationship with the reported TDTS (identity < 60%) (Table S3); (2) contained a complete biosynthetic gene cluster (associated with tailoring enzymes such as P450 monooxygenase and other oxidation-reduction enzymes) (Figures S3 and S4); and (3) possessed complete conservative domains (DDXXD and NSE/DTE motif) (Figure S5). TDTSs were selected for heterologous expression in *A. oryzae* to characterize their function. Twenty-three TDTSs satisfied these criteria.

SPME-GCMS assays revealed that 10 of the 23 selected TDTSs were activated to produce a series of mono- and sesquiterpenes. Notably, two rare TDTSs, ni15096-TDTS

and bb12017-TDTS, produced the monoterpene skeletons sabinene (**1**) and camphene (**2**), respectively (Figure 4i,ii). Although related biosynthetic enzymes have been reported from plants before, this is the first time that such enzymes have been reported from fungi. The remaining eight TDTSs produced sesquiterpene skeletons: chamipinene (**3**), acoradiene (**4**), longifolene (**5**), longiborneol (**6**), sativene (**7**), α-muurolene (**8**), α-cuprenene (**9**), cedrene (**10**, **11**), and γ-bisabolene (**12**) (Figures 4iii–x and S6). Among them, chamipinene (**3**), a rare sesquiterpene, has only been reported from the plant *Illicium oligandrum* [23], with the related biosynthetic enzyme unknown. We report for the first time that an14053-TDTS is the enzyme responsible for catalyzing chamipinene.
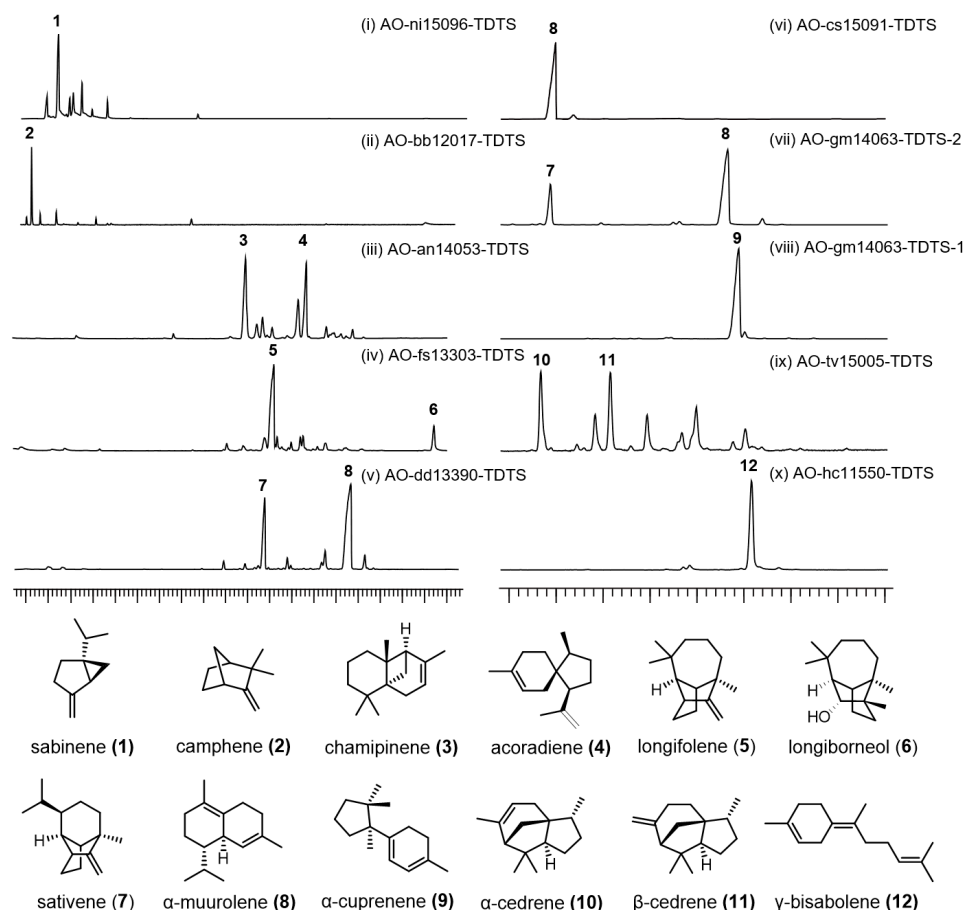
Figure 4. GC–MS profiles of sesquiterpenes produced by *A. oryzae* transformants harbor the selected *TDTSs*. Sesquiterpenes and monoterpenes in the vial headspace were detected by SPME. The structures of compounds detected by GC–MS were identified by comparing their mass spectra with the spectra of terpenoids in the National Institute of Standards and Technology (NIST 14) standard reference database.

3.4. Differences in Cyclization Paths of TDTSs in Families 1 and 2

The 27 reported TDTSs, combined with 10 newly functionally characterized ones, were utilized as a guide to investigate the cyclization pathways of each TDTS family. We found that TDTSs from Family 1 and Family 2 catalyze the cyclization of FPP through similar cyclization pathways.

Unlike most TDTSs that catalyze cyclization through bisabolene intermediates, TDTSs in Family 1 generate sesquiterpenes **5**, **6**, **7**, and **8** through germacrene and humulene carbocation intermediates. We propose the cyclization pathway for **7** and **8**. Following the diphosphate dissociation from FPP, the newly formed C1–C10 bond results in a tertiary carbocation on C11, creating the germacrene carbocation intermediate **IM0**. Carbocation intermediate **IM0** undergoes a 1,2-hydrogen transfer reaction, with subsequent deprotonation

at H7 to yield **8**. While the newly formed C3–C7 bond forms the **IM3** cation, followed by a Wagner–Meerwein rearrangement and deprotonation at C3 to produce **7**. The cyclization of **5** and **6** was similar to that of **7**, with the first cyclization occurring at C1 and C11 instead of at C1 and C10 [24] (Figure 5A).
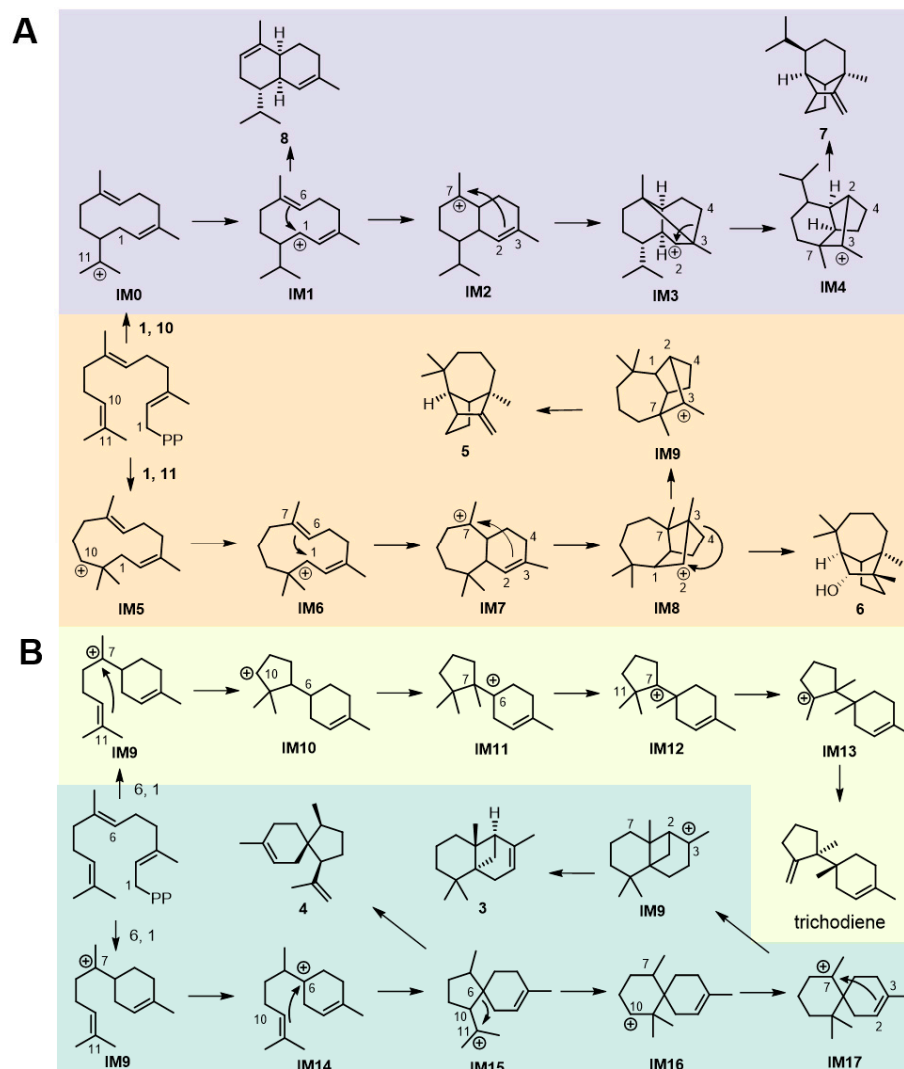

Figure 5. Proposed cyclization mechanism of TDTSs from (**A**) Family 1 and (**B**) Family 2.

TDTSs in Family 2 generated sesquiterpenes **3**, **4**, and trichodiene through the bisabolene carbocation intermediate. We propose the cyclization pathway for trichodiene. Following the diphosphate dissociation from FPP, C6 attacks the C1 carbocation to generate the bisabolene carbocation intermediate **IM9**. Intermediate **IM9** undergoes ring closure between C7 and C11, resulting in the **IM10** cation. This intermediate undergoes a 1,4-hydride shift from C10 to C6, and then two successive 1,2-C shifts lead to **IM13**, followed by deprotonation to produce trichodiene. The cyclization pathways of **3** and **4** share the same bisabolene cyclization intermediate, **IM9**, with trichodiene. The newly formed C6–C10 bond yielded the **IM15** cation, followed by deprotonation at C-11 to produce **4**. Intermediate **IM15** underwent a ring expansion and 1,4-hybrid shift to form the **IM16** cation. Subsequent cyclization at C2 and C7 forms the intermediate **IM17**, followed by deprotonation at C3 to yield **3** [25] (Figure 5B).

Within Family 2, certain TDTS synthesized structurally complex sesquiterpenes **3** and **4** via the **IM6** carbocation intermediate, while others produced the common sesquiterpene trichodiene. To distinguish between these two types of TDTS from their sequences, a

secondary SSN was generated with a different e-value cut-off of 10^{-130}, which indicates that TDTSs in Families 2–1, 2–2, and 2–3 produced trichodiene, **4** and **3** and **4** as major products, respectively (Figure 6).
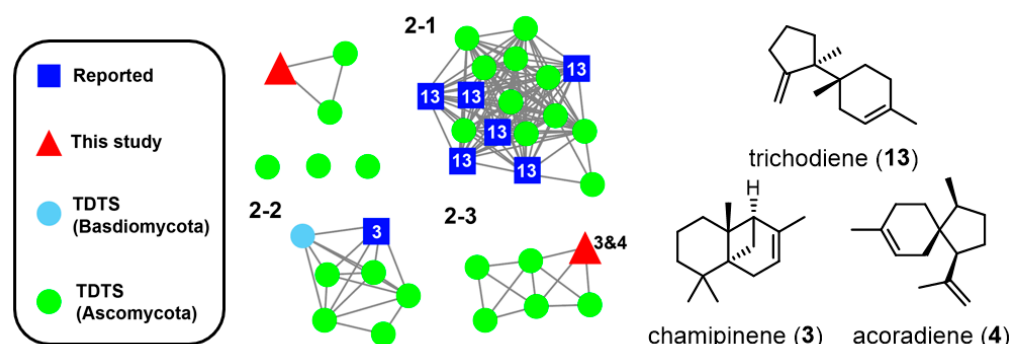

Figure 6. For a more detailed examination of TDTSs within Family 2, a secondary SSN was generated using a different e-value cut-off of 10^{-130}.

4. Discussion

Fungi offer a rich reservoir for the discovery of novel sesquiterpenoids with diverse bioactivities. Sesquiterpene synthases (STSs) catalyze the cyclization of acyclic FPP precursors to produce a variety of cyclic sesquiterpene scaffolds, which significantly enhance the bioactivity and structural diversity of sesquiterpenoids [3]. The genome mining of fungal STSs could facilitate the discovery of novel sesquiterpenes and enzymes. In this study, phylogenetic analysis of previously reported STSs resulted in the identification of a fungus-specific STS family, trichodiene synthase-like sesquiterpene synthases (TDTSs). The application of HMMs further allowed the discovery of 517 TDTSs from our in-house fungal genome library. Considering the novelty of protein sequences and the completeness of their biosynthetic gene cluster, 23 TDTS genes were selected for biochemical characterization in *Aspergillus oryzae*, leading to the identification of the first chamipinene synthase as well as the first fungal-derived cedrene, sabinene, and camphene synthases. The BGCs of all active TDTSs also contained tailoring enzymes, including cytochromes P450, transferases, and other oxidoreductases (Figure S3). The co-expression of these tailoring enzymes with TDTSs holds promise for generating novel terpenoids.

Monoterpenoids comprise a large terpenoid family widely utilized in the pharmaceutical, food, and perfume industries. While plants are the major known source of monoterpenoids, monoterpenoids from fungi are rare and poorly studied. Only around 70 monoterpenoids have been identified in fungi, accounting for approximately 1% of total fungal terpenoids [1]. In this study, we identified two fungi-derived MTSs, sabinene and camphene synthases. Although several sabinene synthases [26] and camphene synthases [27] have been identified in plants, none have been characterized in fungi prior to this study. Sabinene finds applications as flavorings, perfume additives, fine chemicals, and advanced biofuels. Similarly, camphene can be used as a food additive and fragrance component [26]. Our work not only broadens the diversity of known fungus-derived MTSs but also presents new avenues for the biological production of these valuable compounds. The further co-expression of TDTSs and modified genes in biosynthetic gene clusters should also provide insights into the biosynthesis of this valuable terpenoid family. Among the five functionally characterized MTSs, four belong to the TDTS family, suggesting the substantial potential of TDTSs to produce monoterpenes.

Trichoderma species are renowned for their ability to colonize the root surface and promote plant growth [28]. Previous research has indicated that cedrene, produced by *Trichoderma guizhouense*, can modulate *Arabidopsis* root development through the transport and signaling of auxin [29]. However, the fungal STS responsible for cedrene production remains unknown. In this study, the first fungal cedrene synthases were characterized from

Trichoderma voglmayr, offering potential opportunities for developing *Trichoderma* strains with agricultural applications.

Recent advancements in genome mining of fungal STSs have unveiled numerous novel sesquiterpenes and enzymes [7]. However, the sequence-function relationships of STSs have been barely studied. In this study, guided by 38 functionally characterized TDTSs, we explored the sequence-function relationships among different TDTS families. TDTSs in Family 1 could produce bridged cyclic sesquiterpenes such as **5**, **6**, and **7**, while those in Family 2 could synthesize spiro and bridged cyclic sesquiterpenes such as **3** and **4**. Further genome mining of these two TDTS families holds the potential to unveil additional complex polycyclic sesquiterpenes.

In summary, in this study, through the phylogenetic analysis of 123 reported fungal STSs, we identified a fungi-specific STS family, TDTS. The systematic screening of TDTSs from our in-house genome library and the further SSN analysis provided a comprehensive landscape of TDTSs. Considering the novelty of protein sequences and the completeness of their BGCs, 23 TDTS genes were selected for biochemical characterization in *A. oryzae*, resulting in the identification of the first champipinene synthase as well as the first fungal cedrene, sabinene, and camphene synthases. Additionally, with the guidance of 38 functionally characterized TDTSs, we made preliminary speculation regarding the potential sequence-function relationships of TDTSs from Ascomycota. Our research presents a new avenue for the genome mining of fungal sesquiterpenoids.

References

1. Zeng, T.; Liu, Z.; Zhuang, J.; Jiang, Y.; He, W.; Diao, H.; Lv, N.; Jian, Y.; Liang, D.; Qiu, Y.; et al. TeroKit: A database-driven web server for terpenome research. *J. Chem. Inf. Model.* **2020**, *60*, 2082–2090. [CrossRef] [PubMed]
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*, 249–272. [CrossRef] [PubMed]

3. Miller, D.J.; Allemann, R.K. Sesquiterpene synthases: Passive catalysts or active players? *Nat. Prod. Rep.* **2012**, *29*, 60–71. [CrossRef] [PubMed]

4. Bergman, M.E.; Davis, B.; Phillips, M.A. Medically Useful Plant Terpenoids: Biosynthesis, Occurrence, and Mechanism of Action. *Molecules* **2019**, *24*, 3961. [CrossRef] [PubMed]

5. Yin, Z.; Dickschat, J.S. Engineering fungal terpene biosynthesis. *Nat. Prod. Rep.* **2023**, *40*, 28–45. [CrossRef]

6. Dai, Q.; Zhang, F.-L.; Feng, T. Sesquiterpenoids Specially Produced by Fungi: Structures, Biological Activities, Chemical and Biosynthesis (2015–2020). *J. Fungi* **2021**, *7*, 1026. [CrossRef] [PubMed]

7. Wang, S.; Chen, R.; Yuan, L.; Zhang, C.; Liang, D.; Qiao, J. Molecular and functional analyses of characterized sesquiterpene synthases in mushroom-forming fungi. *J. Fungi* **2023**, *9*, 1017. [CrossRef] [PubMed]

8. Zhang, C.; Chen, X.; Orban, A.; Shukal, S.; Birk, F.; Too, H.P.; Ruhl, M. *Agrocybe aegerita* serves as a gateway for identifying sesquiterpene biosynthetic enzymes in higher fungi. *ACS Chem. Biol.* **2020**, *15*, 1268–1277. [CrossRef] [PubMed]

9. Masunaga, N.; Kitaoka, T.; Ichinose, H. Biocatalyst collection and heterologous expression of sesquiterpene synthases from basidiomycetous fungi: Discovery of a novel sesquiterpene hydrocarbon. *Microb. Biotechnol.* **2022**, *16*, 632–644. [CrossRef]

10. Nagamine, S.; Liu, C.; Nishishita, J.; Kozaki, T.; Sogahata, K.; Sato, Y.; Minami, A.; Ozaki, T.; Schmidt-Dannert, C.; Maruyama, J.; et al. Ascomycete aspergillus oryzae is an efficient expression host for production of basidiomycete terpenes by using genomic DNA sequences. *Appl. Environ. Microbiol.* **2019**, *85*, e00409–e00419.

11. Zhang, T.; Feng, J.; He, W.; Rong, X.; Lv, H.; Li, J.; Li, X.; Wang, H.; Wang, L.; Zhang, L.; et al. Genomic and transcriptomic approaches povide a pedictive framework for sesquiterpenes bosynthesis in *Desarmillaria tabescens* CPCC 401429. *J. Fungi* **2023**, *9*, 481. [CrossRef]

12. Proctor, R.H.; McCormick, S.P.; Gutierrez, S. Genetic bases for variation in structure and biological activity of trichothecene toxins produced by diverse fungi. *Appl. Microbiol. Biotechnol.* **2020**, *104*, 5185–5199. [CrossRef] [PubMed]

13. Webb, E.C. Enzyme nomenclature 1992. In *Recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the Nomenclature and Classification of Enzymes*; Academic Press: New York, NY, USA, 1992; Volume 6.

14. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Soding, J.; et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539. [CrossRef] [PubMed]

15. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [CrossRef] [PubMed]

16. Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **2021**, *49*, W293–W296. [CrossRef] [PubMed]

17. 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*, W200–W204. [CrossRef] [PubMed]

18. Camacho, C.; Boratyn, G.M.; Joukov, V.; Vera Alvarez, R.; Madden, T.L. ElasticBLAST: Accelerating sequence search via cloud computing. *BMC Bioinform.* **2023**, *24*, 117. [CrossRef] [PubMed]

19. Stanke, M.; Diekhans, M.; Baertsch, R.; Haussler, D. Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics* **2008**, *24*, 637–644. [CrossRef]

20. 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*, 2498–2504. [CrossRef]

21. Procter, J.B.; Carstairs, G.M.; Soares, B.; Mourao, K.; Ofoegbu, T.C.; Barton, D.; Lui, L.; Menard, A.; Sherstnev, N.; Roldan-Martinez, D.; et al. Alignment of Biological Sequences with Jalview. *Methods Mol. Biol.* **2021**, *2231*, 203–224. [CrossRef]

22. Wheeler, T.J.; Clements, J.; Finn, R.D. Skylign: A tool for creating informative, interactive logos representing sequence alignments and profile hidden Markov models. *BMC Bioinform.* **2014**, *15*, 7. [CrossRef]

23. Lu, H.N.; Ma, S.G.; Liu, Y.B.; Qu, J.; Li, Y.; Xu, S.; Zhu, H.; Yu, S.S. Sesquiterpenes from the roots of Illicium oligandrum. *J. Asian Nat. Prod. Res.* **2015**, *17*, 430–438. [CrossRef] [PubMed]

24. 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*, 8474–8485. [CrossRef]

25. Hong, Y.J.; Tantillo, D.J. Branching out from the bisabolyl cation. Unifying mechanistic pathways to barbatene, bazzanene, chamigrene, chamipinene, cumacrene, cuprenene, dunniene, isobazzanene, iso-gamma-bisabolene, isochamigrene, laurene, microbiotene, sesquithujene, sesquisabinene, thujopsene, trichodiene, and widdradiene sesquiterpenes. *J. Am. Chem. Soc.* **2014**, *136*, 2450–2463. [CrossRef]

26. Cao, Y.; Zhang, H.; Liu, H.; Liu, W.; Zhang, R.; Xian, M.; Liu, H. Biosynthesis and production of sabinene: Current state and perspectives. *Appl. Microbiol. Biotechnol.* **2018**, *102*, 1535–1544. [CrossRef] [PubMed]

27. Jorg, B.; Michael, P.; Vasanthi, R.; Sadanobu, K.; Croteau, R. cDNA cloning, characterization, and functional expression of four new monoterpene synthase members of the Tpsd gene family from grand fir (*Abies grandis*). *Arch. Biochem. Biophys.* **1999**, *368*, 232–243. [CrossRef]

28. Druzhinina, I.S.; Seidl-Seiboth, V.; Herrera-Estrella, A.; Horwitz, B.A.; Kenerley, C.M.; Monte, E.; Mukherjee, P.K.; Zeilinger, S.; Grigoriev, I.V.; Kubicek, C.P. *Trichoderma*: The genomics of opportunistic success. *Nat. Rev. Microbiol.* **2011**, *9*, 749–759. [CrossRef] [PubMed]

29. Li, Y.; Shao, J.; Fu, Y.; Chen, Y.; Wang, H.; Xu, Z.; Feng, H.; Xun, W.; Liu, Y.; Zhang, N.; et al. The volatile cedrene from *Trichoderma guizhouense* modulates *Arabidopsis* root development through auxin transport and signalling. *Plant Cell Environ.* **2022**, *45*, 969–984. [CrossRef]

30. Lopez-Gallego, F.; Agger, S.A.; Abate-Pella, D.; Distefano, M.D.; Schmidt-Dannert, C. Sesquiterpene synthases Cop4 and Cop6 from *Coprinus cinereus*: Catalytic promiscuity and cyclization of farnesyl pyrophosphate geometric isomers. *Chembiochem* **2010**, *11*, 1093–1106. [CrossRef]

31. Wawrzyn, G.T.; Quin, M.B.; Choudhary, S.; Lopez-Gallego, F.; Schmidt-Dannert, C. Draft genome of *Omphalotus olearius* provides a predictive framework for sesquiterpenoid natural product biosynthesis in Basidiomycota. *Chem. Biol.* **2012**, *19*, 772–783. [CrossRef]

32. Bahadoor, A.; Schneiderman, D.; Gemmill, L.; Bosnich, W.; Blackwell, B.; Melanson, J.E.; McRae, G.; Harris, L.J. Hydroxylation of longiborneol by a Clm2-encoded CYP450 monooxygenase to produce culmorin in *Fusarium graminearum*. *J. Nat. Prod.* **2016**, *79*, 81–88. [CrossRef]

33. Yee, D.A.; Kakule, T.B.; Cheng, W.; Chen, M.; Chong, C.T.Y.; Hai, Y.; Hang, L.F.; Hung, Y.S.; Liu, N.; Ohashi, M.; et al. Genome mining of alkaloidal terpenoids from a hybrid terpene and nonribosomal peptide biosynthetic pathway. *J. Am. Chem. Soc.* **2020**, *142*, 710–714. [CrossRef] [PubMed]

34. Nosenko, T.; Zimmer, I.; Ghirardo, A.; Kollner, T.G.; Weber, B.; Polle, A.; Rosenkranz, M.; Schnitzler, J.P. Predicting functions of putative fungal sesquiterpene synthase genes based on multiomics data analysis. *Fungal Genet. Biol.* **2023**, *165*, 103779. [CrossRef] [PubMed]

35. Proctor, R.H.; McCormick, S.P.; Kim, H.S.; Cardoza, R.E.; Stanley, A.M.; Lindo, L.; Kelly, A.; Brown, D.W.; Lee, T.; Vaughan, M.M.; et al. Evolution of structural diversity of trichothecenes, a family of toxins produced by plant pathogenic and entomopathogenic fungi. *PLoS Pathog.* **2018**, *14*, e1006946. [CrossRef] [PubMed]

36. Ichinose, H.; Kitaoka, T. Insight into metabolic diversity of the brown-rot basidiomycete *Postia placenta* responsible for sesquiterpene biosynthesis: Semi-comprehensive screening of cytochrome P450 monooxygenase involved in protoilludene metabolism. *Microb. Biotechnol.* **2018**, *11*, 952–965. [CrossRef] [PubMed]

37. Li, R.; Yao, B.; Zeng, H. Identification and characterization of a nerol synthase in fungi. *J. Agric. Food Chem.* **2023**, *72*, 416–423. [CrossRef] [PubMed]

38. Ichinose, H.; Ukeba, S.; Kitaoka, T. Latent potentials of the white-rot basidiomycete *Phanerochaete chrysosporium* responsible for sesquiterpene metabolism: CYP5158A1 and CYP5144C8 decorate (E)-alpha-bisabolene. *Enzyme Microb. Technol.* **2022**, *158*, 110037. [CrossRef] [PubMed]

39. Al-Salihi, S.A.A.; Dao, T.T.; Williams, K.; Bailey, A.M.; Foster, G.D. The biogenetic origin of the biologically active naematolin of Hypholoma species involves an unusual sesquiterpene synthase. *Mol. Biotechnol.* **2019**, *61*, 754–762. [CrossRef] [PubMed]

40. Burkhardt, I.; Kreuzenbeck, N.B.; Beemelmanns, C.; Dickschat, J.S. Mechanistic characterization of three sesquiterpene synthases from the termite-associated fungus *Termitomyces*. *Org. Biomol. Chem.* **2019**, *17*, 3348–3355. [CrossRef]

41. Lackner, G.; Peters, E.E.; Helfrich, E.J.; Piel, J. Insights into the lifestyle of uncultured bacterial natural product factories associated with marine sponges. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E347–E356. [CrossRef]

42. Mischko, W.; Hirte, M.; Fuchs, M.; Mehlmer, N.; Bruck, T.B. Identification of sesquiterpene synthases from the Basidiomycota *Coniophora puteana* for the efficient and highly selective beta-copaene and cubebol production in *E. coli*. *Microb. Cell Fact.* **2018**, *17*, 164. [CrossRef] [PubMed]

43. Yap, H.Y.; Muria-Gonzalez, M.J.; Kong, B.H.; Stubbs, K.A.; Tan, C.S.; Ng, S.T.; Tan, N.H.; Solomon, P.S.; Fung, S.Y.; Chooi, Y.H. Heterologous expression of cytotoxic sesquiterpenoids from the medicinal mushroom *Lignosus rhinocerotis* in yeast. *Microb. Cell Fact.* **2017**, *16*, 103. [CrossRef] [PubMed]

44. Agger, S.; Lopez-Gallego, F.; Schmidt-Dannert, C. Diversity of sesquiterpene synthases in the basidiomycete *Coprinus cinereus*. *Mol. Biotechnol.* **2009**, *72*, 1181–1195. [CrossRef] [PubMed]

45. Ntana, F.; Bhat, W.W.; Johnson, S.R.; Jorgensen, H.J.L.; Collinge, D.B.; Jensen, B.; Hamberger, B. A sesquiterpene synthase from the endophytic fungus *Serendipita indica* catalyzes formation of viridiflorol. *Biomolecules* **2021**, *11*, 898. [CrossRef] [PubMed]

46. Zeng, H.; Yin, G.; Wei, Q.; Li, D.; Wang, Y.; Hu, Y.; Hu, C.; Zou, Y. Unprecedented [5.5.5.6]dioxafenestrane ring construction in fungal insecticidal sesquiterpene biosynthesis. *Angew. Chem. Int. Ed. Engl.* **2019**, *58*, 6569–6573. [CrossRef] [PubMed]

47. Li, Z.; Wang, Y.; Jiang, B.; Li, W.; Zheng, L.; Yang, X.; Bao, Y.; Sun, L.; Huang, Y.; Li, Y. Structure, cytotoxic activity and mechanism of protoilludane sesquiterpene aryl esters from the mycelium of *Armillaria mellea*. *J. Ethnopharmacol.* **2016**, *184*, 119–127. [CrossRef] [PubMed]

48. Yuan, Y.; Litzenburger, M.; Cheng, S.; Bian, G.; Hu, B.; Yan, P.; Cai, Y.; Deng, Z.; Bernhardt, R.; Liu, T. Sesquiterpenoids produced by combining two sesquiterpene cyclases with promiscuous myxobacterial CYP260B1. *Chembiochem* **2019**, *20*, 677–682. [CrossRef] [PubMed]

49. Buchvaldt Amby, D.; Manczak, T.; Petersen, M.A.; Sundelin, T.; Weitzel, C.; Grajewski, M.; Simonsen, H.T.; Jensen, B. Role of the *Colletotrichum acutatum* sesquiterpene synthase CaTPS in the biosynthesis of sesquiterpenoids. *Microbiology* **2016**, *162*, 1773–1783. [CrossRef] [PubMed]

50. Wu, W.; Tran, W.; Taatjes, C.A.; Alonso-Gutierrez, J.; Lee, T.S.; Gladden, J.M. Rapid discovery and functional characterization of terpene synthases from four endophytic *Xylariaceae*. *PLoS ONE* **2016**, *11*, e0146983. [CrossRef]

51. 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. [CrossRef]

52. Burkhardt, I.; Siemon, T.; Henrot, M.; Studt, L.; Rçsler, S.; Tudzynski, B.; Christmann, M.; Dickschat, J.S. Mechanistic characterisation of two sesquiterpene cyclases from the plant pathogenic fungus *Fusarium fujikuroi*. *Angew. Chem. Int. Ed.* **2016**, *55*, 8748–8751. [CrossRef] [PubMed]

53. Shaw, J.J.; Berbasova, T.; Sasaki, T.; Jefferson-George, K.; Spakowicz, D.J.; Dunican, B.F.; Portero, C.E.; Narvaez-Trujillo, A.; Strobel, S.A. Identification of a fungal 1,8-cineole synthase from *Hypoxylon* sp. with specificity determinants in common with the plant synthases. *J. Biol. Chem.* **2015**, *290*, 8511–8526. [CrossRef] [PubMed]

54. De Sena Filho, J.G.; Quin, M.B.; Spakowicz, D.J.; Shaw, J.J.; Kucera, K.; Dunican, B.; Strobel, S.A.; Schmidt-Dannert, C. Genome of *Diaporthe* sp. provides insights into the potential inter-phylum transfer of a fungal sesquiterpenoid biosynthetic pathway. *Fungal Biol.* **2016**, *120*, 1050–1063. [CrossRef]

55. Sun, X.; Cai, Y.S.; Yuan, Y.; Bian, G.; Ye, Z.; Deng, Z.; Liu, T. Genome mining in *Trichoderma viride* J1-030: Discovery and identification of novel sesquiterpene synthase and its products. *Beilstein J. Org. Chem.* **2019**, *15*, 2052–2058. [CrossRef] [PubMed]

56. Brock, N.L.; Huss, K.; Tudzynski, B.; Dickschat, J.S. Genetic dissection of sesquiterpene biosynthesis by *Fusarium fujikuroi*. *Chembiochem* **2013**, *14*, 311–315. [CrossRef]

57. Moraga, J.; Dalmais, B.; Izquierdo-Bueno, I.; Aleu, J.; Hanson, J.R.; Hernandez-Galan, R.; Viaud, M.; Collado, I.G. Genetic and molecular basis of botrydial biosynthesis: Connecting cytochrome P450-encoding genes to biosynthetic intermediates. *ACS Chem. Biol.* **2016**, *11*, 2838–2846. [CrossRef] [PubMed]

58. 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*, 4543–4548. [CrossRef] [PubMed]

59. 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*, 354–364. [CrossRef]

60. Murai, K.; Lauterbach, L.; Teramoto, K.; Quan, Z.; Barra, L.; Yamamoto, T.; Nonaka, K.; Shiomi, K.; Nishiyama, M.; Kuzuyama, T.; et al. An unusual skeletal rearrangement in the biosynthesis of the sesquiterpene trichobrasilenol from *Trichoderma*. *Angew. Chem. Int. Ed. Engl.* **2019**, *58*, 15046–15050. [CrossRef]