Genome of *Diaporthe* sp. provides insights into the potential inter-phylum transfer of a fungal sesquiterpenoid biosynthetic pathway

Jose Guedes de Sena Filho[1,2,†], **Maureen B. Quin**[3,†], **Daniel J Spakowicz**[1], **Jeffrey J. Shaw**[1], **Kaury Kucera**[1], **Brian Dunican**[1], **Scott A. Strobel**[1], and **Claudia Schmidt-Dannert**[3,*]

[1]Yale University, Dept of Molecular Biophysics and Biochemistry, 800 West Campus Drive, West Haven, CT 06516, USA

[2]EMBRAPA-Empresa Brasileira de Pesquisa Agropecuária-Coastal Tablelands, Av. Beira Mar, 3250, Jardins, 49025-040 - Aracaju, SE – Brazil

[3]University of Minnesota, Dept. of Biochemistry, Molecular Biology and Biophysics, 1479 Gortner Avenue, St Paul, MN 55108, USA

Abstract

Fungi have highly active secondary metabolic pathways which enable them to produce a wealth of sesquiterpenoids that are bioactive. One example is 6-protoilludene, the precursor to the cytotoxic illudins, which are pharmaceutically relevant as anticancer therapeutics. To date, this valuable sesquiterpene has only been identified in members of the fungal division Basidiomycota. To explore the untapped potential of fungi belonging to the division Ascomycota in producing 6-protoilludene, we isolated the a fungal endophyte *Diaporthe* sp. BR109 and show that it produces a diversity of terpenoids including 6-protoilludene. Using a genome sequencing and mining approach 17 putative novel sesquiterpene synthases were identified in *Diaporthe* sp. BR109. A phylogenetic approach was used to predict which gene encodes 6-protoilludene synthase, which was then confirmed experimentally. These analyses reveal that the sesquiterpene synthase and its putative sesquiterpene scaffold modifying cytochrome P450(s) may have been acquired by inter-phylum horizontal gene transfer from Basidiomycota to Ascomycota. Bioinformatic analyses indicate that inter-phylum transfer of these minimal sequiterpenoid secondary metabolic pathways may have occurred in other fungi. This work provides insights into the evolution of fungal sesquiterpenoid secondary metabolic pathways in the production of pharmaceutically relevant bioactive natural products.

*Corresponding author: Claudia Schmidt-Dannert; schmi232@umn.edu Tel: 612-625-5782 Fax: 612-625-5780.
†These authors contributed equally to this work

AUTHOR CONTRIBUTIONS

Conceived and designed experiments: JGSF, MBQ, CS-D; conducted experiments: JGSF, MBQ, DJS, JJS, KK, BFD; analyzed data: JGSF, MBQ, CS-D; provided reagents: JGSF, SAS, CS-D; wrote the manuscript: MBQ, CS-D.

1. INTRODUCTION

Higher fungi (divisions Ascomycota and Basidiomycota) are a valuable source of natural products. Fungi produce bioactive compounds to protect their mycelia and/or fruiting bodies against threats posed by competing microorganisms, invertebrates, insects and larger animals (Alves et al., 2013; Aly et al., 2011; Gershenzon and Dudareva, 2007; Schmidt-Dannert, 2015). Some of these bioactive compounds are of value as pharmaceutical agents; as anticancer, antiviral, antimicrobial and antifungal drugs (Abraham, 2001; Elisashvili, 2012; Evidente et al., 2014; Misiek and Hoffmeister, 2007). However, chemical synthesis of these structurally complex molecules is frequently infeasible or economically impractical and as such, biosynthetic methods are required for their production. This can be achieved either by the native fungal producer or more commonly, in the case of low-yield or genetically intractable strains, by engineered, recombinant systems. Therefore, there is a need to identify and characterize the natural products biosynthetic enzymes and pathways for the development of such heterologous production systems, facilitating access to these chemically diverse scaffolds (Ajikumar et al., 2010; Boettger and Hertweck, 2013; Lin et al., 2013; Lin et al., 2014; Paddon et al., 2013; Spakowicz and Strobel, 2015; Tsunematsu et al., 2013; Wawrzyn et al., 2012a; Wiemann et al., 2013; Wiemann and Keller, 2014; Yaegashi et al., 2014).

One of the largest classes of natural products is the terpenoids, with more than 40,000 compounds isolated. Several hundred different terpene scaffolds have been described, making this a highly diverse class of chemical structures (Christianson, 2008). Among the fungi, mushroom-forming Basidiomycota are prolific producers of structurally diverse terpene natural products (Quin et al., 2014; Schmidt-Dannert, 2015). Fungal terpenes are produced from the mevalonate pathway-derived five carbon precursors dimethylallyl diphosphate (DMAPP) or isopentenyl diphosphate (IPP) (Miziorko, 2011). Condensation of monomers of DMAPP and IPP results in linear hydrocarbons of varying length, such as the C15 molecule farnesyl pyrophosphate (($2E$,$6E$)-FPP), which is the precursor to the volatile sesquiterpenes. The ionization-dependent removal of inorganic pyrophosphate (PPi) from ($2E$,$6E$)-FPP, and the initial cyclization of the reactive carbocation is catalyzed by enzymes known as sesquiterpene synthases (Christianson, 2006; Miller and Allemann, 2012). Different sesquiterpene synthases can catalyze distinct carbocation cyclization mechanisms: the enzyme can catalyze a 1,10 or 1,11 cyclization following removal of PPi from ($2E$,$6E$)-FPP (Lesburg et al., 1997), yielding a *E,E*-germacradienyl cation or *trans*-humulyl cation, respectively. Alternatively, the enzyme can cause ($2E$,$6E$)-FPP to undergo an initial isomerization yielding ($3R$)-nerolidyl pyrophosphate (NPP) (Cane et al., 1981), followed by a 1,6 or 1,10 cyclization yielding a bisabolyl cation or *Z,E*-germacradienyl cation, respectively (Vedula et al., 2007). Subsequently, the carbocation undergoes a series of further cyclizations, ring rearrangements, methyl and hydride shifts, until a final deprotonation mediated either by attack by a water molecule (Pinedo et al., 2008) or by the

PPi (Chen et al., 2013) results in the release of the sesquiterpene from the enzyme active site. The sesquiterpene can then be further modified by accessory enzymes such as cytochrome P450 monoxygenases and other types of oxidoreductases that are typically located in biosynthetic gene clusters together with the terpene synthases in fungi, yielding the final bioactive sesquiterpenoid product (Agger et al., 2009).

Of particular interest are fungal sesquiterpene/oids derived from a 1,11 cyclization of (2*E*, 6*E*)-FPP, via a *trans*-humulyl cation. Many of these chemicals are highly bioactive, and have potential as pharmaceutical agents (Abraham, 2001). A large number of *trans*-humulyl cation derived sesquiterpenoids have been isolated from *Basidiomycota*, a significant portion of which appear to be derived from modified variants of the sesquiterpene scaffold Δ6-protoilludene (Abraham, 2001; Quin et al., 2014; Schmidt-Dannert, 2015). Δ6-protoilludene derived sesquiterpenoids have potent anticancer properties, and are currently being developed as therapeutics (McMorris et al., 2003; Schobert et al., 2011a; Schobert et al., 2011b). As part of a larger effort to identify and characterize the fungal biosynthetic pathways involved in Δ6-protoilludene derived sesquiterpenoid production, we previously biochemically characterized a number of *Basidiomycota* sesquiterpene synthases, including five Δ6-protoilludene synthases (Agger et al., 2009; Quin et al., 2013a; Wawrzyn et al., 2012b). Using our knowledge of the experimentally validated cyclization mechanisms of these enzymes we mined the rapidly expanding Joint Genome Initiative (JGI) database of fungal genomes (Cuomo and Birren, 2010) to create a predictive framework (Wawrzyn et al., 2012b). Our framework enables the sequence-guided initial prediction of cyclization mechanism of putative fungal sesquiterpene synthases prior to cloning and characterization efforts, and has proven to be reliable in the directed discovery and characterization of additional sesquiterpene synthases from *Basidiomycota*, based upon cyclization mechanism of choice (Quin et al., 2013a).

In contrast to the *Basidiomycota*, relatively few studies have described sesquiterpene/oids isolated from *Ascomycota*. Of the *Ascomycota* sesquiterpenes, only a small number of those described are derived from the *trans*-humulyl cation. For example, α-humulene and koraiol (Brock et al., 2013) are produced by *Fusarium fujikorii*; and presilphiperfolan-8β-ol (Pinedo et al., 2008; Wang et al., 2009) is produced by *Botrytis cinerea*; the corresponding sesquiterpene synthases have been identified and characterized from these major plant pathogens. Our previous phylogenetic analyses showed that these two Ascomycota enzymes form a clade that is distinct from the one that contains the Basidiomycota Δ6-protoilludene synthases identified and characterized by our group, indicating a difference in sequence (Schmidt-Dannert, 2015). To the best of our knowledge, only one study has shown production of Δ6-protoilludene by the plant pathogenic Ascomycete *Ceratocystis piceae* (Hanssen et al., 1986), and the sesquiterpene synthase has not yet been identified. Bioinformatics analyses indicate that overall, Ascomycota have comparatively fewer putative sesquiterpene synthases encoded within their genomes compared to the sesquiterpene synthase rich Basidiomycota, which may explain why fewer sesquiterpene synthases have been characterized from Ascomycota (Schmidt-Dannert, 2015). Nonetheless, the fact that *trans*-humulyl cation producing sesquiterpene synthases have already been identified and characterized; and that at least one Ascomycete is known to produce Δ6-

protoilludene; indicates an underlying untapped potential of Ascomycota to produce bioactive sesquiterpenoids derived from a 6-protoilludane scaffold.

In this work we sought to identify 1,11-cyclizing sesquiterpene synthases, specifically 6-protoilludene synthases, in Ascomycota with the goal of accessing novel bioactive terpenoids. We focused our efforts on endophytic fungi which are known to be a source of pharmaceutically relevant terpenoids (Barra et al., 2014; Chandra, 2012; de Souza et al., 2011), including monoterpenes that are synthesized by a recently discovered monoterpene synthase in the endophyte *Hypoxylon* sp. (Shaw et al., 2015a). *De novo* genome sequencing of a new environmentally isolated fungal endophyte, coded here as *Diaporthe* sp. BR109, and subsequent bioinfomatics analyses, led to the discovery and characterization of the first 6-protoilludene synthase from an Ascomycete. Unexpectedly, our results also suggest a potential horizontal gene transfer of at least partial biosynthetic pathways leading to 6-protoilludene derived sesquiterpenoids, apparently across fungal divisions. These findings deepen our appreciation of the complex nature of evolution of secondary metabolic pathways in fungi.

2. MATERIALS AND METHODS

2.1 Endophytic fungal isolation, culturing and sesquiterpene extractions

Caesalpinia pyramidalis Tul (Fabaceae) plant stem and leaf material were collected in northeast Brazil. The plant material was surface sterilized and plated on isolation media comprised of 1 × potato dextrose agar (PDA) (24 g/L potato dextrose), 0.1 × PDA (2.4 g/L potato dextrose) (EMD Millipore) or water agar (15 g/L agar) for isolation of *fungi*, as described elsewhere (Shaw et al., 2015a). A fungal isolate growing from the plated plant stem material was subcultured on a fresh 1 × PDA plate and was coded BR109. For detection of sesquiterpene production, BR109 was grown for 14 and 28 days in 100 mL potato dextrose broth (PDB) liquid media at 25 °C with shaking at 125 rpm. Fungal sesquiterpenes were extracted from the culture biomass using 2×30 mL chloroform and ethyl acetate. Extractions were filtered and were dried under vacuum prior to shipping to the University of Minnesota for GC/MS analyses.

2.2 GC/MS analysis of terpene/oids

GC/MS analysis was conducted on a HP GC 7890A chromatograph coupled to an anion-trap mass spectrometer HP MSD triple axis detector (Agilent Technologies, Santa Clara, CA, USA). Compounds were separated on a HP-5MS capillary column (30 m × 0.25 mm × 1.0 mm) with an injection port temperature of 250 °C and helium as carrier gas. The oven temperature started at 60 °C and was increased at 10 °C min^{-1} to a final oven temperature of 250 °C. Mass spectra were scanned over the range 5 – 300 amu at 1 sec intervals. The retention indices and mass spectra of compound peaks were compared to reference spectra available in reference libraries in Mass Finder 4 (Joulain and Konig, 1998) and the National Institute of Standards and Technology (NIST) standard reference database.

2.3 Phylogenetic analysis of fungal isolate BR109

The internal transcribed spacer (ITS) rDNA of fungal isolate BR109 was amplified by colony PCR using a modified method (van Zeijl et al., 1997), using trichoderma lysing enzyme (Sigma-Aldrich). The primers used for amplification were ITS1 (TCCGTAGGTGAACCTGCGG) and ITS4 (TCCTCCGCTTATTGATATGC). The resulting PCR product was sequenced by the W.M. Keck Foundation. The ITS sequence was aligned with homologs identified by BLAST using MUSCLE v3.8.31 and a maximum likelihood phylogenetic tree was constructed using RAxML v7.2.8 and MrBayes v3.2.0.

2.4 Genome sequencing of fungal isolate BR109

Genomic DNA was isolated from a 7 day old culture grown in potato dextrose media using a DNeasy Plant Mini kit (Qiagen), as described elsewhere (Gianoulis et al., 2012). Two libraries were prepared (450 bp and 3 kb mate-pair) according to Illumina protocols. Libraries were sequenced on the Illumina Hiseq 2000 using paired-end 2×250 bp reads and assembled Allpaths-LG v44034 (Gnerre et al., 2011). The genome data was submitted to GenBank (Bioproject ID PRJN288728).

2.5 Sesquiterpene synthase identification

Gene predictions were conducted using AUGUSTUS (Stanke et al., 2004), using reference species *Botrytis cinerea, Fusarium graminearum, Chaetomium globosum*, and *Magnaporthe grisea*. Amino acid predictions were manually checked for accuracy by creating a sequence alignment using Clustal W Thompson et al., 1994) with previously characterized sesquiterpene synthases to check for the presence and correct alignment of conserved metal binding motifs present in all sesquiterpene synthases (Lesburg et al., 1998). For predictions of cyclization mechanism of putative sesquiterpene synthases, the curated list of amino acid sequences were aligned, and phylogenetic analyses were carried out using MEGA 6 (Tamura et al., 2013) with the default parameters for the Neighbor-Joining Saitou and Nei, 1987) method and a bootstrap test of phylogeny (500 replicates). Biosynthetic gene clusters surrounding the putative sesquiterpene synthases were manually annotated using AUGUSTUS and BLAST. Further details of bioinformatics methods, all gene sequences, and protein sequences, are provided in the Supporting Information.

2.6 Cloning and expression of sesquiterpene synthase in E. coli

Putative sesquiterpene synthase Dia1 was synthesized with codon optimization for expression in *E. coli* and cloned into plasmid backbones pUC57-Kan using restriction sites NcoI and XmaI (GENEWIZ). We subcloned Dia1 into our in-house plasmid pUCBB (Vick et al., 2011) that allows for constitutive expression using restriction sites BamHI and SpeI. Plasmid pUCBB-Dia1 was transformed into *E. coli* BL21 and single colonies were isolated on Lysogeny Broth (LB) agar supplemented with ampicillin (100 μg mL^{-1}). Single colonies were used to inoculate 2L flasks of LB broth (500 mL) supplemented with ampicillin 100 μg mL^{-1}), and the cultures were incubated overnight at 30 °C with shaking at 225 rpm. The headspace of flasks was sampled the next day for volatile sesquiterpenes for 10 minutes using a solid-phase microextraction (SPME) fiber, followed by GC-MS analysis.

3. RESULTS

3.1 Endophytic fungal isolate BR109 produces mono- and sesquiterpenes

In our search for terpenoid producing endophytic fungi that produce pharmaceutically relevant *trans*-humulyl derived sesquiterpenes, we serendipitously isolated an uncharacterized fungal endophyte (coded BR109) from a flowering tree that is endemic to northeastern Brazil. The endophyte isolate produced a range of mono- and sesquiterpenoids, including 6-protoilludene (*vide infra*) and was identified as *Diaporthe* (*Phomopsis*) sp. by ITS sequencing. Phylogenetic analyses showed that fungal isolate BR109 was most closely related to *Diaporthe anacardii* and *Diaporthe foeniculaceae* (Fig 1)(Gomes et al., 2013). Like the 6-protoilludene producing *Ceratocystis piceae* (Ascomycota, Sordariomycetes, Hypocreomycetidae) (Hanssen et al., 1986) and the recently genome sequenced *Hypoxolon* sp. (Ascomycota, Sordariomycetes, Xylariomycetidae) which encodes a monoterpene synthase (Shaw et al., 2015a), this fungus also belongs to the Sordiomycetes but is a representative of the third subclass, Sordariomycetidae, within this fungal class.

To identify the terpenoids produced by the fungal isolate BR109, metabolites produced by the fungus were solvent extracted from culture biomass. GC/MS analysis of the extracts confirmed that the isolate produced a variety of terpene/oids (Fig 2), indicating that the fungus likely expressed a diversity of terpene synthases, including the sought after 6-protoilludene synthase. Several peaks in the GC/MS chromatogram had corresponding mass spectra that could be assigned using reference libraries as C10 monoterpenes and monoterpenoids: terpinen-4-ol **1**, trans-piperitol **2**, γ-terpineol **3** and piperiton epoxide **4**. Several other peaks could be assigned as C15 sesquiterpenes and sesquiterpenoids: neotrifaradiene **5**, 6-protoilludene **6**, 3,7-di-epi-trifara-9,14-diene **7**, β-elemene (the heat-induced Cope rearrangement product of germacrene A) **8**, thujopsene **9**, nardosina-7,9,11-triene **10**, aristolene **11**, α-muurolene **12**, α-bulnesene **13**, guaia-9,11-diene **14**, norpatchulenol **15**, and maaliol **16** (Fig 2, Fig 3). Other peaks which could not be confidently assigned using reference libraries, had mass spectra with parent ions corresponding to a *m/z* 206, 216, 222, 234, 235 and 236, which are indicative of modified sesquiterpenoids.

The presence of a wide range of sesquiterpene products suggested that fungal isolate BR109 expressed a complement of terpene synthases that catalyze different types of cyclization reactions. For example, neotrifaradiene **5** and 3,7-di-epi-trifara-9,4-diene **7** share a trifarane backbone which has been proposed to be derived from nerolidol (Sonwa et al., 2001). BR109 therefore may express an acyclic sesquiterpene synthase, nerolidol synthase, which has so far only been described from plants (Green et al., 2012; Nagegowda et al., 2008), bacteria (Nakano et al., 2011) and from the endophyte Hypoxylon sp. (Shaw et al., 2015a). Furthermore, the spectrum of sesquiterpenes made by the fungus suggested that it expressed cyclic sesquiterpene synthases responsible for almost all possible initial cyclization reactions of FPP; for example, thujopsene 9 has been proposed to be derived from a 1,6 cyclization of NPP via a bisabolyl cation (Wu et al., 2005); one of the most abundant products α-muurolene 12, and the guaiane skeleton-derived products 13 and 14 have been proposed to be synthesized from a direct 1,10 cyclization of (2E,6E)-FPP via a E,E-germacradienyl

cation (de Kraker et al., 1998; Wawrzyn et al., 2012b); and 6-protoilludene 6 is the result of a 1,11 cyclization of (2E,6E)-FPP via a trans-humulyl cation (Quin et al., 2013a; Quin et al., 2015; Seemann et al., 2002; Wawrzyn et al., 2012b). None of the products identified, however, appeared to be derived from a 1,10 cyclization of NPP via a Z,E-germacradienyl cation.

3.2 Genome and complement of putative sesquiterpene synthases in BR109

Owing to the fact that fungal isolate BR109 likely expressed a diversity of sesquiterpene synthases, and produced 6-protoilludene **6**, we set out to sequence the genome of this uncharacterized endophyte in order to map its terpene synthase complement and identify the 6-protoilludene synthase. Following de novo genome sequencing of fungal isolate BR109, the final genome assembly length was estimated to be 56 Mbp in size across 145 scaffolds, with an N50 of 1.8 Mbp, which is similar to other reported endophyte genomes (Gianoulis et al., 2012; Shaw et al., 2015a; Shaw et al., 2015b). A total of 16,469 predicted genes were identified (Table S1).

To identify genes encoding putative sesquiterpene synthases in BR109 we carried out BLAST searches using the sequences of diverse fungal sesquiterpene synthases previously cloned and characterized by our group: Cop1–6 from the Basidiomycete Coprinus cinereus (Agger et al., 2009; Lopez-Gallego et al., 2010a; Lopez-Gallego et al., 2010b); Omp1–11 from the Basidiomycete Omphalotus olearius (Quin et al., 2013b; Wawrzyn et al., 2012b); and Stehi159379, Stehi128017, Stehi25180, Stehi64702 and Stehi73029 from the Basidiomycete Stereum hirsutum (Quin et al., 2013a; Quin et al., 2015). We also included sequences of sesquiterpene synthases characterized from Ascomycetes by other groups: presilphiperfolan-8β-ol synthase BcBOT2 from Botrytis cinerea (Wang et al., 2009), koraiol synthase Ffsc4 and acorenol synthase Ffsc6 from Fusarium fujikori (Brock et al., 2013), trichodiene synthase FsTS from Fusarium sporotrichioides (Hohn and Vanmiddlesworth, 1986), longiborneol synthase FgCLM1 from *Fusarium graminearium* (McCormick et al., 2010), aristolochene synthase PrAS from *Penicillium roqueforti* (Caruthers et al., 2000), as well as 11 terpene synthases Hyp1–11 recently identified in *Hypoxylon* sp. (Shaw et al., 2015a). Using our searches we identified a total of 17 genes encoding putative sesquiterpene synthases in the genome of BR109, named Dia1–17 (Fig 4, Supporting Information).

We have previously shown in Basidiomycota that phylogenetic analyses comparing putative sesquiterpene synthases to characterized enzymes are predictive of the cyclization mechanisms (Quin et al., 2013a; Wawrzyn et al., 2012b). Using this same approach, we noted that the potential sesquiterpene synthases from BR109 clustered in distinct clades with other characterized enzymes, potentially according to cyclization mechanism (Fig 4). Most of the predicted Dia sesquiterpene synthases clustered most closely with homologs from other Ascomycota. Dia6–8 clustered in a clade with Ascomycota terpene synthases Hyp1,2,5,6,7 and 8 from *Hypoxylon* sp.. Dia6 was located on a branch with Hyp7, an inactive enzyme; Dia7 was located on a branch with Hyp6, an inactive enzyme; and Dia8 was located on a branch with Hyp2, which produces δ-cadinene (Shaw et al., 2015a). It is not reliable to draw conclusions regarding putative cyclization mechanisms for this particular cluster of Dia enzymes because so few of the characterized Hyp enzymes with

which they share sequence identity were active. The next distinct clade contained putative terpene synthases Dia9–12; Dia9 and Dia10 are closely related to aristolochene synthase PrAS which catalyzes a 1,10 cyclization (Caruthers et al., 2000) and Dia12 was closely related to acorenol synthase Ffsc6 which may catalyze a 1,6 cyclization (Brock et al., 2011). Dia11 is not located on the same branch as any other sesquiterpene synthase in this cluster. Notably, this particular cluster contains previously characterized Ascomycota enzymes catalyzing apparently diverse cyclizations, raising the possibility that our predictive framework may not be as reliable for predicting the activity of Ascomycota sesquiterpene synthases as it is for Basidiomycota (Quin et al., 2013a). The largest complement of putative Dia sesquiterpene synthases clustered with Ascomycota NPP 1,6 cyclizing, bisabolyl cation producing enzymes: Dia13 and Dia14 were most closely related to trichodiene synthase FsTS (Hohn and Vanmiddlesworth, 1986), Dia15 was most closely related to longiborneol synthase FgCLM1 (McCormick et al., 2010), and Dia16 and Dia17 were located on a separate branch from Dia15.

Finally, four of the putative sesquiterpene synthases Dia2–5 clustered separately with the Ascomycota sesquiterpene synthases BcBOT2 (Wang et al., 2009) and Ffsc4 (Brock et al., 2011) that catalyze a 1,11 cyclization of 2*E*,6*E*-FPP, and which we have previously shown cluster separately from the 1,11-cyclizing Basidiomycota enzymes (Quin et al., 2014; Schmidt-Dannert, 2015). Surprisingly, one single putative sesquiterpene synthase, Dia1, clustered very closely with previously characterized 1,11 cyclizing Basidiomycota   6-protoilludene synthases (Engels et al., 2011; Quin et al., 2013a; Wawrzyn et al., 2012b), providing strong support that this particular enzyme may be responsible for the observed   6-protoilludene production by the endophyte BR109.

3.3 Evidence of horizontal gene transfer in   6-protoilludene synthases

Our phylogenetic analyses indicated that, for the most part, the cyclization mechanism of Ascomycota sesquiterpene synthases cannot be inferred by their relationship to Basidiomycota enzymes, as the Dia putative sesquiterpene synthases generally were most similar to *Ascomycota* enzymes, regardless of their cyclizaition mechanism. For example, Dia9 and Dia10 clustered with the *Ascomycete* (2*E*,6*E*)-FPP 1,10 cyclizing, (*E,E*)-germacradienyl cation producing enzyme aristolochene synthase PrAS (Caruthers et al., 2000); not with the *Basidiomycete* versions of this subset of sesquiterpene synthases, Cop1–3 and Omp1–3 (Agger et al., 2009; Wawrzyn et al., 2012b), which are located distantly on the phylogenetic tree (Fig 4). It was therefore intriguing that Dia1 was so closely related to previously characterized *Basidiomycota* 1,11 cyclizing   6-protoilludene synthases Omp6&7, Stehi25180, Stehi64702 and Stehi73029, and Armga1 (Engels et al., 2011; Quin et al., 2013a; Wawrzyn et al., 2012b), as opposed to other 1,11 cyclizing *Ascomycota* sesquiterpene synthases Ffsc4 (Brock et al., 2011) and BcBOT2 (Wang et al., 2009), with which Dia2–5 associate on a separate clade (Fig 4). This obvious difference in primary sequence led us to question whether Dia1 had been acquired via a cross20 phyla horizontal gene transfer event between *Basidiomycota* and *Ascomycota*; a phenomenon that has been described for a small number of fungal secondary metabolic genes and pathways (Slot and Hibbett, 2007; Slot and Rokas, 2011; Wisecaver et al., 2014). The gene encoding Dia1 also shares a more similar intron splicing pattern with genes encoding *Basidiomycota* 1,11

cyclizing sesquiterpene synthases than with genes encoding Ascomycota 1,11 cyclizing enzymes (Fig 5A, Supporting Information), further supporting the theory that horizontal gene transfer may have resulted in the presence of a Basidiomycete-like gene encoding a 6-protoilludene synthase in an Ascomycete genome.

In order to determine whether Dia1 encoded a predicted 6-protoilludene synthase, we synthesized a codon optimized variant of the Dia1 gene, and cloned and expressed the enzyme in *E. coli* and analyzed its activity using previously used methods (Engels et al., 2011; Quin et al., 2013a; Wawrzyn et al., 2012b). The major volatile sesquiterpene produced by *E. coli* cells overexpressing Dia1 was the *trans*-humulyl cation derived 6-protoilludene **6** (Fig 6), confirming that our prediction of cyclization mechanism and activity was correct. Further, minor peaks were also observed including pentalenene **17**, 2-sterpurene **18**, african-1-ene **19**, brasila-5(10),6-diene **20** (Fig 6), all of which are 1,11 cyclization products; as well as the 1,10 cyclization product β-elemene **8** (a heat induced rearrangment product of germacrene A), which has been observed in the headspace of other heterologously overexpressed 6-protoilludene synthases (Quin et al., 2013a).

3.4 Horizontal gene transfer of fungal sesquiterpenoid secondary metabolic pathways

With the confirmation of Dia1 as a 6-protoilludene synthase that is most closely related to *Basidiomycota* enzymes with the same activity (Fig 4 & Fig 5A), we hypothesized that Dia1 was likely acquired by BR109 or its ancestors by horizontal gene transfer from a *Basidiomycota* source. We next questioned whether such a cross-phyla transfer of a *Basidiomycota* 6-protoilludene synthase could be detected in other *Ascomycota*, or whether this was a rare case limited to BR109 and closely related endophytic *Fungi*. Homology searches of 317 *Ascomycota* genomes available at the Joint Genome Institutes MycoCosm genome database with the Dia1 protein sequence as a search template identified 20 putative sesquiterpene synthases homologs that had high sequence identity (E-values < $1E^{-60}$) to Dia1). Interestingly, these homologs were identified in the genomes of fungi belonging to two related fungal classes; the Sordariomycetes and Letiomycetes (Table S2), suggesting that the *Basidiomycota* gene may be acquired by a shared ancestor of these two fungal classes, and was maintained by some, but lost by other members of these two classes. Several of the identified Dia1 homologs share similar intron splicing patterns of their predicted gene architectures (most notably, introns 2 and 3 are separated by a small exon) suggesting a common evolutionary origin (Fig 5B).

All of the previously characterized *Basidiomycota* 6-protoilludene synthases characterized by our group are located in predicted biosynthetic gene clusters (Quin et al., 2013a; Wawrzyn et al., 2012b). The initial 6-protoilludene scaffold is modified into to a wide diversity of bioactive sesquiterpenoids (Abraham, 2001; Quin et al., 2014). These modifications are likely catalyzed by enzymes such as cytochrome P450 monooxygenases, oxidoreductases and methyltransferases. We therefore searched within the genomic regions surrounding these 20 putative sesquiterpene synthase homologs, as well as Dia1, for putative associated biosynthetic genes. Eight of the homologs, as well as Dia1, appear to be part of putative biosynthetic gene clusters that we manually annotated (Fig 7, Table S2). These clusters vary in size, containing anywhere between two and 16 genes encoding putative

transporters, transcription factors, scaffold modifying enzymes (eg. methyltransferases), P450s and oxidoreductases. The presence of genes encoding P450s and oxidoreductases surrounding 6-protoilludene synthase Dia1, as well as the other eight homologs, suggests that these biosynthetic genes likely modify the 6-protoilludene scaffold into bioactive secondary metabolites. For example, the gene cluster associated with Diaam1|5282 identified in the recently sequenced genome of the grapevine trunk pathogen *Diaporthe ampelina* may play an important role in its pathogenicity (Morales-Cruz et al., 2015).

It is notable that every biosynthetic gene cluster that we have identified as being associated with a (putative) 6-protoilludene synthase contains at least one cytochrome P450 monooxygenase (Fig 7) (Quin et al., 2013a; Wawrzyn et al., 2012b), which may indicate that P450 monooxygenases have an essential function in the biosynthesis of modified sesquiterpenoids, although this remains to be verified. All of the putative biosynthetic gene clusters identified for Dia1 and its eight homologs have a P450 enzyme located adjacent to the predicted terpene synthase (Fig 7). If a modification catalyzed by a P450 is essential for the biosynthesis of modified bioactive sesquiterpenoids derived from 6-protoilludene; the cytochrome P450 and terpene synthase may potentially be horizontally transferred together to yield a partial or complete secondary metabolic pathway that provides a fitness advantage to the host (Slot and Rokas, 2011). Our analyses of the predicted P450s located adjacent to Dia1 and the other predicted 6-protoilludene synthase homologs (Fig 7) showed that some were highly related in sequence and potentially in function (CYP family) to *Basidiomycota* P450s (Supporting Information). Putative P450s associated with (putative) sesquiterpene synthases Dia1, Diaam1, Melsp1 and Thiar1 were classified according to sequence as belonging to CYP5037, a cytochrome P450 family that is enriched in *Basidiomycota* (Syed et al., 2014). Furthermore, the closest match identified by BLAST search for these four putative P450s belongs to a *Basidiomycete*, suggesting that at least in these clusters, a terpene synthase and a P450 may have been acquired together from a *Basidiomycota* source by horizontal gene transfer. This is in contrast to the other five P450s associated with the *Ascomycota* sesquiterpene synthase homologs, which were more similar in sequence to *Ascomycota* P450s and these clusters likely recruited a cytochrome P450 from an existing complement of enzymes.

4. DISCUSSION

Sesquiterpenoids belong to a large and diverse class of natural products, many of which have potential as pharmaceutical agents due to their bioactivities (Fraga, 2013). Fungi represent one of the greatest sources of bioactive sesquiterpenoids (Abraham, 2001), therefore the discovery and characterization of fungal sesquiterpenoid biosynthetic pathways is of key interest. The goal of this work was to explore the biosynthetic potential of the relatively understudied fungal division *Ascomycota* in respect to their production of bioactive sesquiterpenoids derived from the 6-protoilludene scaffold that is typically associated with *Basidiomycota*. Here, we have conducted *de novo* genome sequencing of an environmentally isolated fungal endophyte which produces both monoterpenoids and sesquiterpenoids (Fig 2 & 3), and we used a genome mining and phylogenetic approach to identify and characterize the first 6-protoilludene synthase (Fig 6) from an *Ascomycete*.

Previously, we have shown that our predictive framework was reliable for the targeted discovery of novel sesquiterpene synthases from *Basidiomycota* based on cyclization mechanism (Quin et al., 2013a). This has the potential to streamline the discovery and characterization of new enzymes, in particular, those that catalyze a 1,11 cyclization of (2*E*, 6*E*)-FPP to produce *trans*-humulyl cation derived sesquiterpene scaffolds, such as 6-protoilludene. This study shows that predicting cyclization mechanism of putative sesquiterpene synthases from *Ascomycota* based upon phylogenetic clustering may not be as well defined as it is for *Basidiomycota* (Fig 4). This may be related to the fact that relatively fewer sesquiterpene synthases have been identified and characterized from *Ascomycota* than *Basidiomycota* (Schmidt-Dannert, 2015); or it could be due to the fact that in this study we purposefully compared *Ascomycota* and *Basidiomycota* sequences. Our results indicate that there is an underlying difference in gene architecture and amino sequence between *Ascomycota* and *Basidiomycota* sesquiterpene synthases (Fig 4 & 5), which may make phylogenetic analyses more difficult. Future studies depend upon the availability of a more diverse set of experimentally characterized *Ascomycota* sesquiterpene synthases. Coincidently, using this comparative approach led us to identify an unexpected case of horizontal gene transfer across fungal phyla.

Horizontal gene transfer from species to species across divisions and classes is a commonly recognized mechanism in fungal genomic evolution that drives metabolic diversification (Fitzpatrick, 2012; Wisecaver et al., 2014). Gene duplication and horizontal gene transfer likely accounts for the wide-spread presence of 6-protoilludene synthases and associated gene clusters (Quin et al., 2013a; Wawrzyn et al., 2012b) in different members of the division *Basidiomycota*. Similarly, a 6-protoilludene synthase and associated genes may have been transferred from a *Basidiomycete* to one or several members within the *Ascomycota* division, followed by species to species transfer within the division, for example, between the closely related *Diaporthe* sp. BR109 and *Diaporthe ampelina*. So far, our analysis has identified 20 6-protoilludene synthase homologs in the sequenced genomes of fungi from only two closely related *Ascomycota* classes: *Sordariomycetes* and *Leotiomycetes*, which are known to be more proficient in horizontal gene transfer of metabolic genes, and especially clustered genes, in comparison to other fungal divisions and classes (Wisecaver *et al.*, 2014). Eight of the putative 6-protoilludene synthase homologs are located in a biosynthetic gene cluster and in four of those fungi, the terpene synthase and an associated P450 are most closely related to *Basidiomycota* homologs, suggesting that they likely have originally been transferred as a cluster. This strongly suggests that uptake of these genes, despite the expense of alternative splicing arrangements of genes, must provide an evolutionary fitness to the host (Lawrence and Roth, 1996).

It is not known which, if any, of the fungal strains that we have identified by sequence analysis produce 6-protoilludene and/or modified sesquiterpenoids derived from this scaffold, nor is it known what bioactivities such modified sesquiterpenoids could have and how those activities could benefit the host. Nonetheless, the conserved presence of a cytochrome P450 in all putative 6-protoilludene synthase containing biosynthetic gene clusters, which is some cases contain only a sesquiterpene synthase and (a) cytochrome P450(s) (Fig 7), could indicate a putative minimal operon architecture required for function. Likewise, the apparent horizontal transfer of both a sesquiterpene synthase and one

cytochrome P450 in at least some of the *Ascomycota* species could suggest a minimal operon architecture required for evolutionary fitness. Future work will require the characterization of cytochrome P450s within these clusters, and studies of the modified sesquiterpene scaffolds that they may be involved in creating, to determine whether any of these hypotheses are correct.

In general, more natural product related studies have been carried out with *Ascomycota* than *Basidiomycota*, despite the fact that *Basidiomycota* are a rich resource for bioactive sesquiterpenoids. This may be due to several factors, including the fact that reliable laboratory cultivation techniques for natural products isolation are comparably well established for filamentous fungi and endophytes (Posch et al., 2013; Venugopalan and Srivastava, 2015); more genome sequences of *Ascomycetes* are publically available (Schmidt-Dannert, 2015); and the genetics of *Ascomycetes* are better understood, enabling the development of genetic manipulation techniques (Brakhage, 2013; Brakhage and Schroeckh, 2011; Scharf and Brakhage, 2013). *Ascomycota* could therefore serve as a useful platform for the discovery and/or production of further modified bioactive sesquiterpenoids such as those derived from the 6-protoilludene scaffold. Furthermore, fungal endophytes such as the *Ascomycete Diaporthe* (*Phomopsis*) sp. produce a diversity of interesting and bioactive sesquiterpene scaffolds, such as cadinanes (Silva et al., 2006), drimanes (Zang le et al., 2012) and meroterpenoids (Hemtasin et al., 2011; Ma et al., 2015; Tajima et al., 2004) and as we have shown here, *trans*-humulane scaffolds. *Diaporthe* (*Phomopsis*) sp. is also known to produce monoterpenoids (Sassa et al., 2003; Singh et al., 2011) as we have also shown here; as well as diterpenoids (Toyomasu et al., 2008) and triterpenoids (Zhang et al., 2013). This fungal species therefore represents a rich source of terpene synthases that could potentially be characterized. Increased availablity of *Diaporthe* (*Phomopsis*) sp. genome sequences, such as the one we present here, or other novel species (Koloniuk et al., 2014) will facilitate further studies in this area.

5. Conclusion

This work describes the first characterization of a 6-protoilludene synthase from an *Ascomycete*, and provides insights into the evolutionary pathway of individual genes and potentially minimal sesquiterpenoid secondary metabolic pathways across fungal phyla. These results offer a basis towards the discovery and characterization of further novel biotechnologically relevant enzymes involved in the biosynthesis of bioactive natural products in fungi.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

REFERENCES

Abraham WR. Bioactive sesquiterpenes produced by fungi - are they useful for humans as well? Curr. Med. Chem. 2001; 8:583–606. [PubMed: 11281843]

Agger S, Lopez-Gallego F, Schmidt-Dannert C. Diversity of sesquiterpene synthases in the basidiomycete *Coprinus cinereus*. Mol. Microbiol. 2009; 72:1181–1195. [PubMed: 19400802]

Ajikumar PK, Xiao WH, Tyo KE, Wang Y, Simeon F, Leonard E, Mucha O, Phon TH, Pfeifer B, Stephanopoulos G. Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. Science. 2010; 330:70–74. [PubMed: 20929806]

Alves MJ, Ferreira IC, Dias J, Teixeira V, Martins A, Pintado M. A review on antifungal activity of mushroom (Basidiomycetes) extracts and isolated compounds. Curr. Top. Med. Chem. 2013; 13:2648–2659. [PubMed: 24083794]

Aly AH, Debbab A, Proksch P. Fungal endophytes: unique plant inhabitants with great promises. Appl. Microbiol. Biotechnol. 2011; 90:1829–1845. [PubMed: 21523479]

Barra L, Schulz B, Dickschat JS. Pogostol biosynthesis by the endophytic fungus Geniculosporium. Chem Bio Chem. 2014; 15:2379–2383.

Boettger D, Hertweck C. Molecular diversity sculpted by fungal PKS-NRPS hybrids. Chem Bio Chem. 2013; 14:28–42.

Brakhage AA. Regulation of fungal secondary metabolism. Nat. Rev. Microbiol. 2013; 11:21–32. [PubMed: 23178386]

Brakhage AA, Schroeckh V. Fungal secondary metabolites - strategies to activate silent gene clusters. Fungal Genet. Biol. 2011; 48:15–22. [PubMed: 20433937]

Brock NL, Huss K, Tudzynski B, Dickschat JS. Genetic dissection of sesquiterpene biosynthesis by *Fusarium fujikuroi*. Chem Bio Chem. 2013; 14:311–315.

Brock NL, Tudzynski B, Dickschat JS. Biosynthesis of sesqui- and diterpenes by the gibberellin producer *Fusarium fujikuroi*. Chem Bio Chem. 2011; 12:2667–2676.

Cane DE, Swanson S, Murthy PPN. Trichodiene biosynthesis and the enzymatic cyclization of farnesyl pyrophosphate. J. Am. Chem. Soc. 1981; 103:2136–2138.

Caruthers JM, Kang I, Rynkiewicz MJ, Cane DE, Christianson DW. Crystal structure determination of aristolochene synthase from the blue cheese mold, *Penicillium roqueforti*. J. Biol. Chem. 2000; 275:25533–25539. [PubMed: 10825154]

Chandra S. Endophytic fungi: novel sources of anticancer lead molecules. Appl. Microbiol. Biotechnol. 2012; 95:47–59. [PubMed: 22622838]

Chen M, Al-lami N, Janvier M, D'Antonio EL, Faraldos JA, Cane DE, Allemann RK, Christianson DW. Mechanistic insights from the binding of substrate and carbocation intermediate analogues to aristolochene synthase. Biochemistry. 2013; 52:5441–5453. [PubMed: 23905850]

Christianson DW. Structural biology and chemistry of the terpenoid cyclases. Chemical Reviews. 2006; 106:3412–3442. [PubMed: 16895335]

Christianson DW. Unearthing the roots of the terpenome. Curr. Opin. Chem. Biol. 2008; 12:141–150. [PubMed: 18249199]

Cuomo CA, Birren BW. The Fungal Genome Initiative and Lessons Learned from Genome Sequencing. Methods Enzymol. 2010; 470:833–855. [PubMed: 20946837]

de Kraker JW, Franssen MC, de Groot A, Konig WA, Bouwmeester HJ. (+)-Germacrene A biosynthesis. The committed step in the biosynthesis of bitter sesquiterpene lactones in chicory. Plant Physiol. 1998; 117:1381–1392. [PubMed: 9701594]

de Souza JJ, Vieira IJ, Rodrigues-Filho E, Braz-Filho R. Terpenoids from endophytic fungi. Molecules. 2011; 16:10604–10618. [PubMed: 22183885]

Elisashvili V. Submerged cultivation of medicinal mushrooms: bioprocesses and products (review). Int J Med Mushrooms. 2012; 14:211–239. [PubMed: 22577974]

Engels B, Heinig U, Grothe T, Stadler M, Jennewein S. Cloning and characterization of an Armillaria gallica cDNA encoding protoilludene synthase, which catalyzes the first committed step in the synthesis of antimicrobial melleolides. J. Biol. Chem. 2011; 286:6871–6878. [PubMed: 21148562]

Evidente A, Kornienko A, Cimmino A, Andolfi A, Lefranc F, Mathieu V, Kiss R. Fungal metabolites with anticancer activity. Nat. Prod. Rep. 2014; 31:617–627. [PubMed: 24651312]

Fitzpatrick DA. Horizontal gene transfer in fungi. FEMS Microbiol. Lett. 2012; 329:1–8. [PubMed: 22112233]

Fraga BM. Natural sesquiterpenoids. Nat Prod Rep. 2013; 30:1226–1264. [PubMed: 23884176]

Gershenzon J, Dudareva N. The function of terpene natural products in the natural world. Nat. Chem. Biol. 2007; 3:408–414. [PubMed: 17576428]

Gianoulis TA, Griffin MA, Spakowicz DJ, Dunican BF, Alpha CJ, Sboner A, Sismour AM, Kodira C, Egholm M, Church GM, Gerstein MB, Strobel SA. Genomic analysis of the hydrocarbon-producing, cellulolytic, endophytic fungus *Ascocoryne sarcoides*. PLoS Genet. 2012; 8:e1002558. [PubMed: 22396667]

Gnerre S, Maccallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, Sharpe T, Hall G, Shea TP, Sykes S, Berlin AM, Aird D, Costello M, Daza R, Williams L, Nicol R, Gnirke A, Nusbaum C, Lander ES, Jaffe DB. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. Proc. Natl. Acad. Sci. U S A. 2011; 108:1513–1518. [PubMed: 21187386]

Gomes RR, Glienke C, Videira SI, Lombard L, Groenewald JZ, Crous PW. *Diaporthe*: a genus of endophytic, saprobic and plant pathogenic fungi. Persoonia. 2013; 31:1–41. [PubMed: 24761033]

Green SA, Chen X, Nieuwenhuizen NJ, Matich AJ, Wang MY, Bunn BJ, Yauk YK, Atkinson RG. Identification, functional characterization, and regulation of the enzyme responsible for floral (E)-nerolidol biosynthesis in kiwifruit (*Actinidia chinensis*). J. Exp. Bot. 2012; 63:1951–1967. [PubMed: 22162874]

Hanssen HP, Sprecher E, Abraham WR. 6-Protoilludene, the major volatile metabolite from *Ceratocystis piceae* liquid cultures. Phytochemistry. 1986; 25:1979–1980.

Hemtasin C, Kanokmedhakul S, Kanokmedhakul K, Hahnvajanawong C, Soytong K, Prabpai S, Kongsaeree P. Cytotoxic pentacyclic and tetracyclic aromatic sesquiterpenes from *Phomopsis archeri*. J. Nat. Prod. 2011; 74:609–613. [PubMed: 21341709]

Hohn TM, Vanmiddlesworth F. Purification and characterization of the sesquiterpene cyclase trichodiene synthetase from *Fusarium sporotrichioides*. Arch. Biochem. Biophys. 1986; 251:756–761. [PubMed: 3800398]

Joulain, D.; Konig, WA. The atlas of spectral data of sesquiterpene hydrocarbons. EB-Verlag, Hamburg: 1998.

Koloniuk I, El-Habbak MH, Petrzik K, Ghabrial SA. Complete genome sequence of a novel hypovirus infecting *Phomopsis longicolla*. Arch. Virol. 2014; 159:1861–1863. [PubMed: 24473711]

Lawrence JG, Roth JR. Selfish operons: horizontal transfer may drive the evolution of gene clusters. Genetics. 1996; 143:1843–1860. [PubMed: 8844169]

Lesburg CA, Caruthers JM, Paschall CM, Christianson DW. Managing and manipulating carbocations in biology: terpenoid cyclase structure and mechanism. Curr. Opin. Struct. Biol. 1998; 8:695–703. [PubMed: 9914250]

Lesburg CA, Zhai G, Cane DE, Christianson DW. Crystal structure of pentalenene synthase: mechanistic insights on terpenoid cyclization reactions in biology. Science. 1997; 277:1820–1824. [PubMed: 9295272]

Lin HC, Chooi YH, Dhingra S, Xu W, Calvo AM, Tang Y. The fumagillin biosynthetic gene cluster in *Aspergillus fumigatus* encodes a cryptic terpene cyclase involved in the formation of beta-trans-bergamotene. J. Am. Chem. Soc. 2013; 135:4616–4619. [PubMed: 23488861]

Lin HC, Tsunematsu Y, Dhingra S, Xu W, Fukutomi M, Chooi YH, Cane DE, Calvo AM, Watanabe K, Tang Y. Generation of complexity in fungal terpene biosynthesis: discovery of a multifunctional cytochrome P450 in the fumagillin pathway. J. Am. Chem. Soc. 2014; 136:4426–4436. [PubMed: 24568283]

Lopez-Gallego F, Agger SA, Abate-Pella D, Distefano MD, Schmidt-Dannert C. Sesquiterpene synthases Cop4 and Cop6 from *Coprinus cinereus*: catalytic promiscuity and cyclization of farnesyl pyrophosphate geometric isomers. Chem Bio Chem. 2010a; 11:1093–1106.

Lopez-Gallego F, Wawrzyn GT, Schmidt-Dannert C. Selectivity of fungal sesquiterpene synthases: role of the active site's H-1 alpha loop in catalysis. Appl. Environ. Microbiol. 2010b; 76:7723–7733. [PubMed: 20889795]

Ma X, Wang W, Li E, Gao F, Guo L, Pei Y. A new sesquiterpene from the entomogenous fungus Phomopsis amygdali. Nat. Prod. Res. 2015:1–5.

McCormick SP, Alexander NJ, Harris LJ. CLM1 of *Fusarium graminearum* encodes a longiborneol synthase required for culmorin production. Appl. Environ. Microbiol. 2010; 76:136–141. [PubMed: 19880637]

McMorris TC, Cong Q, Kelner MJ. Structure-activity relationship studies of illudins: analogues possessing a spiro-cyclobutane ring. The Journal of Organic Chemistry. 2003; 68:9648–9653. [PubMed: 14656090]

Miller DJ, Allemann RK. Sesquiterpene synthases: passive catalysts or active players? Nat Prod Rep. 2012; 29:60–71. [PubMed: 22068697]

Misiek M, Hoffmeister D. Fungal genetics, genomics, and secondary metabolites in pharmaceutical sciences. Planta Medica. 2007; 73:103–115. [PubMed: 17245664]

Miziorko HM. Enzymes of the mevalonate pathway of isoprenoid biosynthesis. Arch. Biochem. Biophys. 2011; 505:131–143. [PubMed: 20932952]

Morales-Cruz A, Amrine KC, Blanco-Ulate B, Lawrence DP, Travadon R, Rolshausen PE, Baumgartner K, Cantu D. Distinctive expansion of gene families associated with plant cell wall degradation, secondary metabolism, and nutrient uptake in the genomes of grapevine trunk pathogens. BMC Genomics. 2015; 16:469. [PubMed: 26084502]

Nagegowda DA, Gutensohn M, Wilkerson CG, Dudareva N. Two nearly identical terpene synthases catalyze the formation of nerolidol and linalool in snapdragon flowers. Plant J. 2008; 55:224–239. [PubMed: 18363779]

Nakano C, Kim HK, Ohnishi Y. Identification and characterization of the linalool/nerolidol synthase from *Streptomyces clavuligerus*. Chem Bio Chem. 2011; 12:2403–2407.

Paddon CJ, Westfall PJ, Pitera DJ, Benjamin K, Fisher K, McPhee D, Leavell MD, Tai A, Main A, Eng D, Polichuk DR, Teoh KH, Reed DW, Treynor T, Lenihan J, Fleck M, Bajad S, Dang G, Dengrove D, Diola D, Dorin G, Ellens KW, Fickes S, Galazzo J, Gaucher SP, Geistlinger T, Henry R, Hepp M, Horning T, Iqbal T, Jiang H, Kizer L, Lieu B, Melis D, Moss N, Regentin R, Secrest S, Tsuruta H, Vazquez R, Westblade LF, Xu L, Yu M, Zhang Y, Zhao L, Lievense J, Covello PS, Keasling JD, Reiling KK, Renninger NS, Newman JD. High-level semi-synthetic production of the potent antimalarial artemisinin. Nature. 2013; 496:528–532. [PubMed: 23575629]

Pinedo C, Wang CM, Pradier JM, Dalmais B, Choquer M, Le Pecheur P, Morgant C, Collado IG, Cane DE, Viaud M. Sesquiterpene Synthase from the Botrydial Biosynthetic Gene Cluster of the Phytopathogen *Botrytis cinerea*. ACS Chem. Biol. 2008; 3:791–891. [PubMed: 19035644]

Posch AE, Herwig C, Spadiut O. Science-based bioprocess design for filamentous fungi. Trends Biotechnol. 2013; 31:37–44. [PubMed: 23183302]

Quin MB, Flynn CM, Schmidt-Dannert C. Traversing the fungal terpenome. Nat. Prod. Rep. 2014; 31:1449–1473. [PubMed: 25171145]

Quin MB, Flynn CM, Wawrzyn GT, Choudhary S, Schmidt-Dannert C. Mushroom hunting by using bioinformatics: application of a predictive framework facilitates the selective identification of sesquiterpene synthases in basidiomycota. Chem Bio Chem. 2013a; 14:2480–2491.

Quin MB, Michel SN, Schmidt-Dannert C. Moonlighting metals: insights into regulation of cyclization pathways in fungal Delta6-protoilludene sesquiterpene synthases. Chem Bio Chem. 2015

Quin MB, Wawrzyn G, Schmidt-Dannert C. Purification, crystallization and preliminary X-ray diffraction analysis of Omp6, a protoilludene synthase from *Omphalotus olearius*. Acta Crystallogr Sect F Struct Biol Cryst Commun. 2013b; 69:574–577.

Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 1987; 4:406–425. [PubMed: 3447015]

Sassa T, Kenmoku H, Sato M, Murayama T, Kato N. (+)-Menthol and its hydroxy derivatives, novel fungal monoterpenols from the fusicoccin-producing fungi, *Phomopsis amygdali* F6a and Niigata 2. Biosci Biotechnol Biochem. 2003; 67:475–479. [PubMed: 12723593]

Scharf DH, Brakhage AA. Engineering fungal secondary metabolism: a roadmap to novel compounds. J. Biotechnol. 2013; 163:179–183. [PubMed: 22820338]

Schmidt-Dannert C. Biosynthesis of terpenoid natural products in fungi. Adv. Biochem. Eng. Biotechnol. 2015; 148:19–61. [PubMed: 25414054]

Schobert R, Knauer S, Seibt S, Biersack B. Anticancer active illudins: recent developments of a potent alkylating compound class. Curr. Med. Chem. 2011a; 18:790–807. [PubMed: 21182482]

Schobert R, Seibt S, Mahal K, Ahmad A, Biersack B, Effenberger-Neidnicht K, Padhye S, Sarkar FH, Mueller T. Cancer selective metallocenedicarboxylates of the fungal cytotoxin illudin M. J. Med. Chem. 2011b; 54:6177–6182. [PubMed: 21848340]

Seemann M, Zhai G, de Kraker JW, Paschall CM, Christianson DW, Cane DE. Pentalenene synthase. Analysis of active site residues by site-directed mutagenesis. J. Am. Chem. Soc. 2002; 124:7681–7689. [PubMed: 12083921]

Shaw JJ, Berbasova T, Sasaki T, Jefferson-George K, Spakowicz DJ, Dunican BF, Portero CE, Narvaez-Trujillo A, Strobel SA. Identification of a fungal 1,8-cineole synthase from *Hypoxylon* sp. with specificity determinants in common with the plant synthases. J. Biol. Chem. 2015a; 290:8511–8526. [PubMed: 25648891]

Shaw JJ, Spakowicz DJ, Dalal RS, Davis JH, Lehr NA, Dunican BF, Orellana EA, Narvaez-Trujillo A, Strobel SA. Biosynthesis and genomic analysis of medium-chain hydrocarbon production by the endophytic fungal isolate *Nigrograna mackinnonii* E5202H. Appl. Microbiol. Biotechnol. 2015b; 99:3715–3728. [PubMed: 25672844]

Silva GH, Teles HL, Zanardi LM, Marx Young MC, Eberlin MN, Hadad R, Pfenning LH, Costa-Neto CM, Castro-Gamboa I, da Silva Bolzani V, Araujo AR. Cadinane sesquiterpenoids of Phomopsis cassiae, an endophytic fungus associated with *Cassia spectabilis* (Leguminosae). Phytochemistry. 2006; 67:1964–1969. [PubMed: 16857221]

Singh SK, Strobel GA, Knighton B, Geary B, Sears J, Ezra D. An endophytic Phomopsis sp. possessing bioactivity and fuel potential with its volatile organic compounds. Microb. Ecol. 2011; 61:729–739. [PubMed: 21331608]

Slot JC, Hibbett DS. Horizontal transfer of a nitrate assimilation gene cluster and ecological transitions in fungi: a phylogenetic study. PLoS One. 2007; 2:e1097. [PubMed: 17971860]

Slot JC, Rokas A. Horizontal transfer of a large and highly toxic secondary metabolic gene cluster between fungi. Curr. Biol. 2011; 21:134–139. [PubMed: 21194949]

Sonwa MM, Konig WA, Lahlou EH, Asakawa Y. Sesquiterpene hydrocarbons with trifarane backbone in the liverwort *Trocholejeunea sandvicensis*. Phytochemistry. 2001; 57:499–506. [PubMed: 11394847]

Spakowicz DJ, Strobel SA. Biosynthesis of hydrocarbons and volatile organic compounds by fungi: bioengineering potential. Appl. Microbiol. Biotechnol. 2015; 99:4943–4951. [PubMed: 25957494]

Stanke M, Steinkamp R, Waack S, Morgenstern B. AUGUSTUS: a web server for gene finding in eukaryotes. Nucleic Acids Res. 2004; 32:W309–W312. [PubMed: 15215400]

Syed K, Shale K, Pagadala NS, Tuszynski J. Systematic identification and evolutionary analysis of catalytically versatile cytochrome p450 monooxygenase families enriched in model basidiomycete fungi. PLoS One. 2014; 9:e86683. [PubMed: 24466198]

Tajima N, Nukina M, Kato N, Sassa T. Novel Fusicoccins R and S, and the fusicoccin S aglycon (phomopsiol) from *Phomopsis amygdal* i niigata 2-A, and their seed germination-stimulating activity in the presence of abscisic acid. Biosci Biotechnol Biochem. 2004; 68:1125–1130. [PubMed: 15170119]

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. 2013; 30:2725–2729. [PubMed: 24132122]

Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994; 22:4673–4680. [PubMed: 7984417]

Toyomasu T, Niida R, Kenmoku H, Kanno Y, Miura S, Nakano C, Shiono Y, Mitsuhashi W, Toshima H, Oikawa H, Hoshino T, Dairi T, Kato N, Sassa T. Identification of Diterpene Biosynthetic Gene Clusters and Functional Analysis of Labdane-Related Diterpene Cyclases in *Phomopsis amygdali*. Biosci. Biotechnol. Biochem. 2008; 72:1038–1047. [PubMed: 18391465]

Tsunematsu Y, Ishiuchi K, Hotta K, Watanabe K. Yeast-based genome mining, production and mechanistic studies of the biosynthesis of fungal polyketide and peptide natural products. Nat. Prod. Rep. 2013; 30:1139–1149. [PubMed: 23824111]

van Zeijl CM, van de Kamp EH, Punt PJ, Selten GC, Hauer B, van Gorcom RF, van den Hondel CA. An improved colony-PCR method for filamentous fungi for amplification of PCR-fragments of several kilobases. J. Biotechnol. 1997; 59:221–224. [PubMed: 9519482]

Vedula LS, Zhao Y, Coates RM, Koyama T, Cane DE, Christianson DW. Exploring biosynthetic diversity with trichodiene synthase. Arch. Biochem. Biophys. 2007; 466:260–266. [PubMed: 17678871]

Venugopalan A, Srivastava S. Endophytes as in vitro production platforms of high value plant secondary metabolites. Biotechnol. Adv. 2015; 33:873–887. [PubMed: 26225453]

Vick JE, Johnson ET, Choudhary S, Bloch SE, Lopez-Gallego F, Srivastava P, Tikh IB, Wawrzyn GT, Schmidt-Dannert C. Optimized compatible set of BioBrick vectors for metabolic pathway engineering. Appl. Microbiol. Biotechnol. 2011; 92:1275–1286. [PubMed: 22033566]

Wang C-M, Hopson R, Lin X, Cane DE. Biosynthesis of the Sesquiterpene Botrydial in *Botrytis cinerea*. Mechanism and Stereochemistry of the Enzymatic Formation of Presilphiperfolan-8-ol. J. Am. Chem. Soc. 2009; 131:8360–8361. [PubMed: 19476353]

Wawrzyn GT, Bloch SE, Schmidt-Dannert C. Discovery and characterization of terpenoid biosynthetic pathways of fungi. Methods Enzymol. 2012a; 515:83–105. [PubMed: 22999171]

Wawrzyn GT, Quin MB, 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. 2012b; 19:772–783. [PubMed: 22726691]

Wiemann P, Guo CJ, Palmer JM, Sekonyela R, Wang CC, Keller NP. Prototype of an intertwined secondary-metabolite supercluster. Proc. Natl. Acad. Sci. 2013; 110:17065–17070. [PubMed: 24082142]

Wiemann P, Keller NP. Strategies for mining fungal natural products. J. Ind. Microbiol. Biotechnol. 2014; 41:301–313. [PubMed: 24146366]

Wisecaver JH, Slot JC, Rokas A. The evolution of fungal metabolic pathways. PLoS Genet. 2014; 10:e1004816. [PubMed: 25474404]

Wu S, Schoenbeck MA, Greenhagen BT, Takahashi S, Lee S, Coates RM, Chappell J. Surrogate splicing for functional analysis of sesquiterpene synthase genes. Plant Physiol. 2005; 138:1322–1333. [PubMed: 15965019]

Yaegashi J, Oakley BR, Wang CC. Recent advances in genome mining of secondary metabolite biosynthetic gene clusters and the development of heterologous expression systems in *Aspergillus nidulans*. J. Ind. Microbiol. Biotechnol. 2014; 41:433–442. [PubMed: 24342965]

Zang le Y, Wei W, Guo Y, Wang T, Jiao RH, Ng SW, Tan RX, Ge HM. Sesquiterpenoids from the mangrove-derived endophytic fungus *Diaporthe* sp. J. Nat. Prod. 2012; 75:1744–1749. [PubMed: 23006147]

Zhang Y, Hao F, Liu N, Xu Y, Jia A, Yang Z, Xia X, Liu C. Stereochemical determination of a new and cytotoxic euphane triterpenoid from the plant endophytic fungus *Phomopsis chimonanthi*. J Antibiot (Tokyo). 2013; 66:679–682. [PubMed: 23921820]
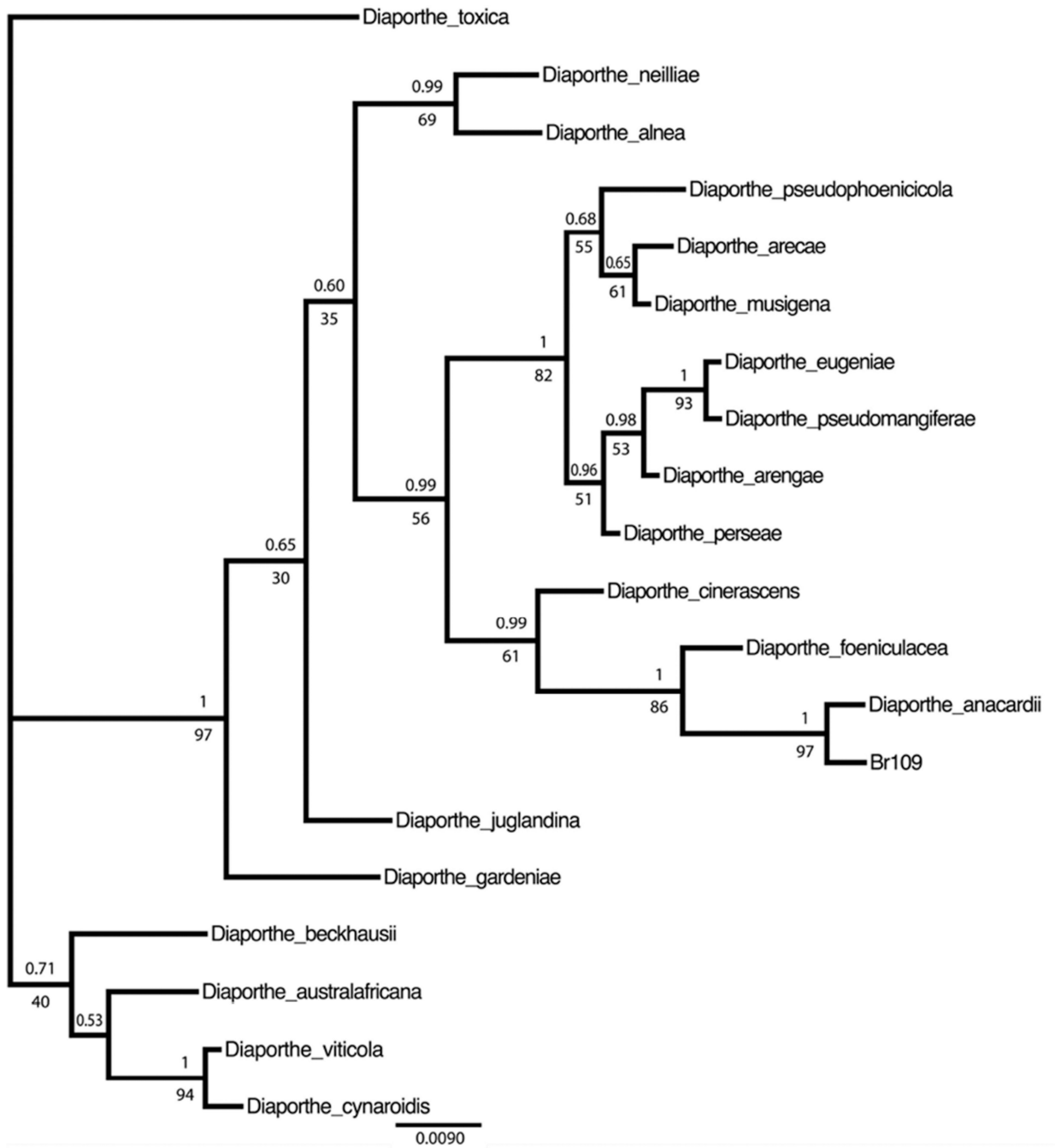
Fig 1. Phylogenetic analyses of fungal isolate BR109

ITS sequencing shows that BR109 is most closely related to *Diaporthe anacardii* and *Diaporthe foeniculacea*.
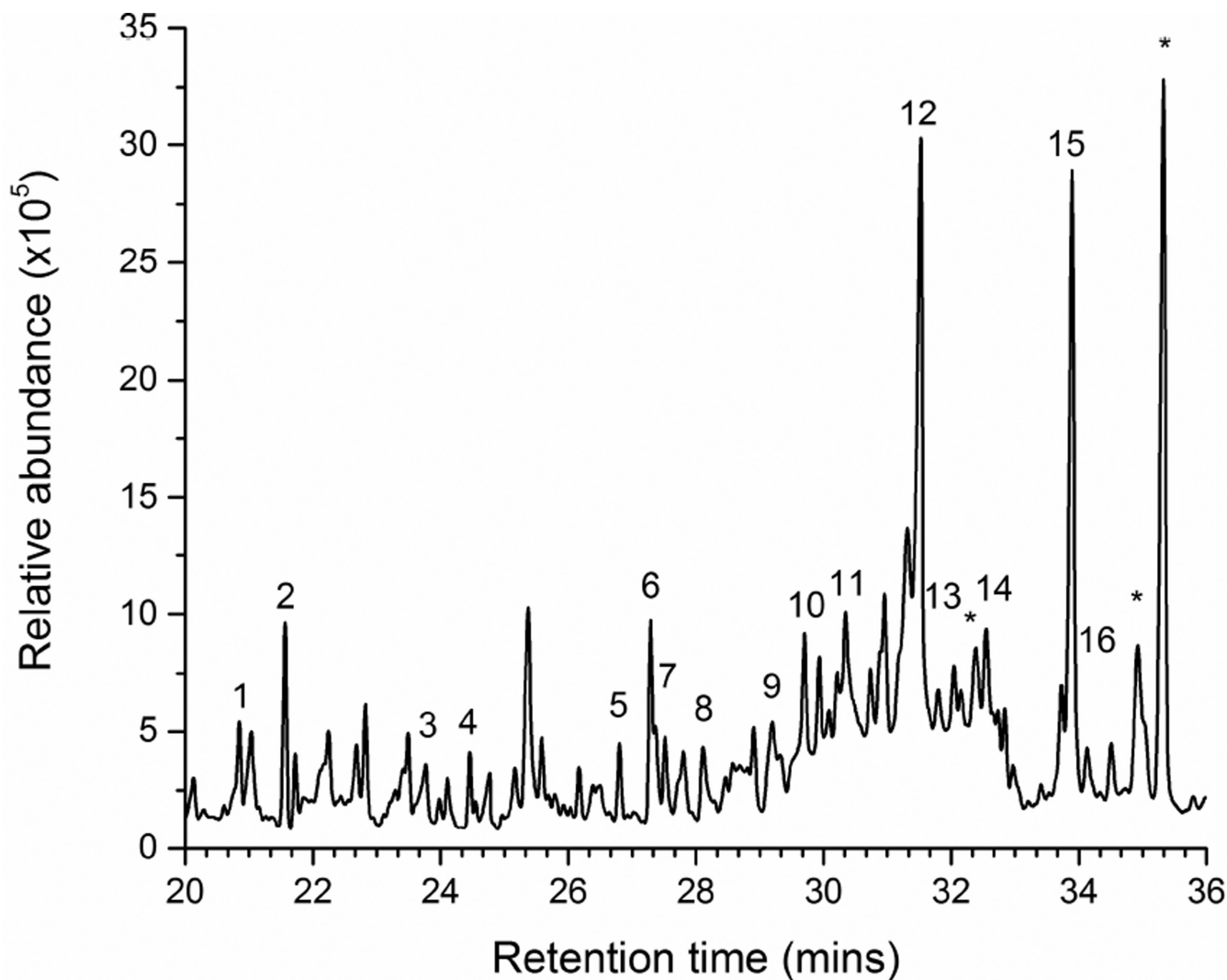
Fig 2. GC/MS analysis of the terpenes/terpenoids produced by fungal isolate BR109

Peaks that could be assigned based on comparison of mass spectra to known libraries of terpenes are labelled with numbers. Numbering of peaks on the chromatogram correspond to the structures shown in Fig 3. Peaks that could not be assigned, but had *m/z* parent ions likely related to modified sesquiterpenoids, are labelled with an asterisk *.
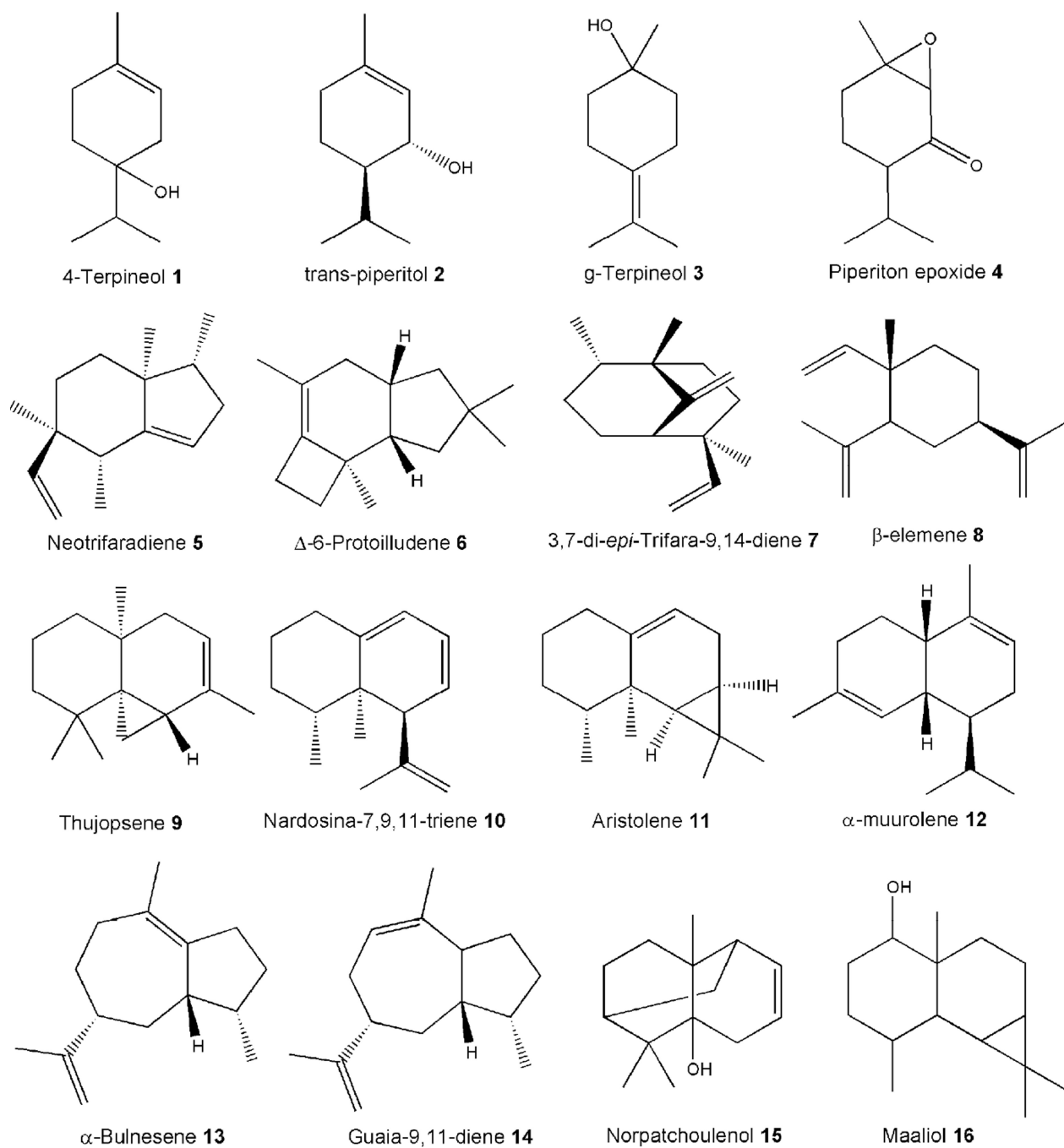
Fig 3. Chemical structures of terpenoids produced by fungal isolate BR109

Identified terpenes/oids are numbered according to the peak labelling in Fig 2. Relative stereochemistries of structures are shown.
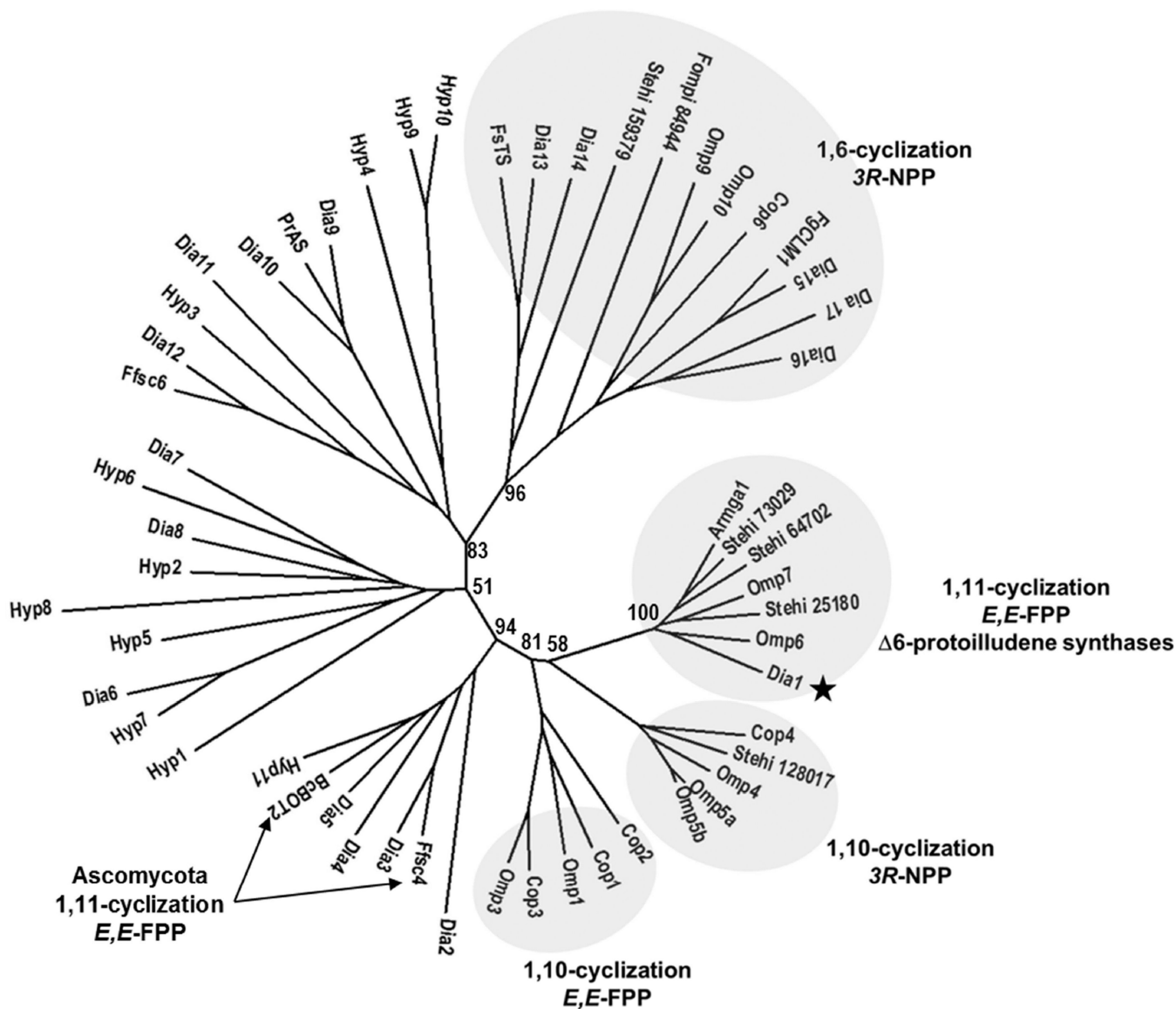
Fig 4. Phylogenetic analyses of the 17 putative sesquiterpene synthases Dia1–17 identified in the genome of fungal isolate BR109

Putative sesquiterpene synthases Dia1–17 cluster in distinct clades alongside other characterized sesquiterpene synthases from *Basidiomycota* and *Ascomycota*, apparently according to cyclization mechanism. Clusters that contain sesquiterpene synthases previously characterized by our group with known cyclization mechanisms are highlighted with a shaded oval. The sesquiterpene synthase that we characterize in this study, Dia1, is highlighted with a star. Dia1 clusters closely with other previously characterized   6-protoilludene synthases from *Basidiomycota*. Bootstrap values (1000 replicates) are shown for the first internal nodes (see Supporting Information).
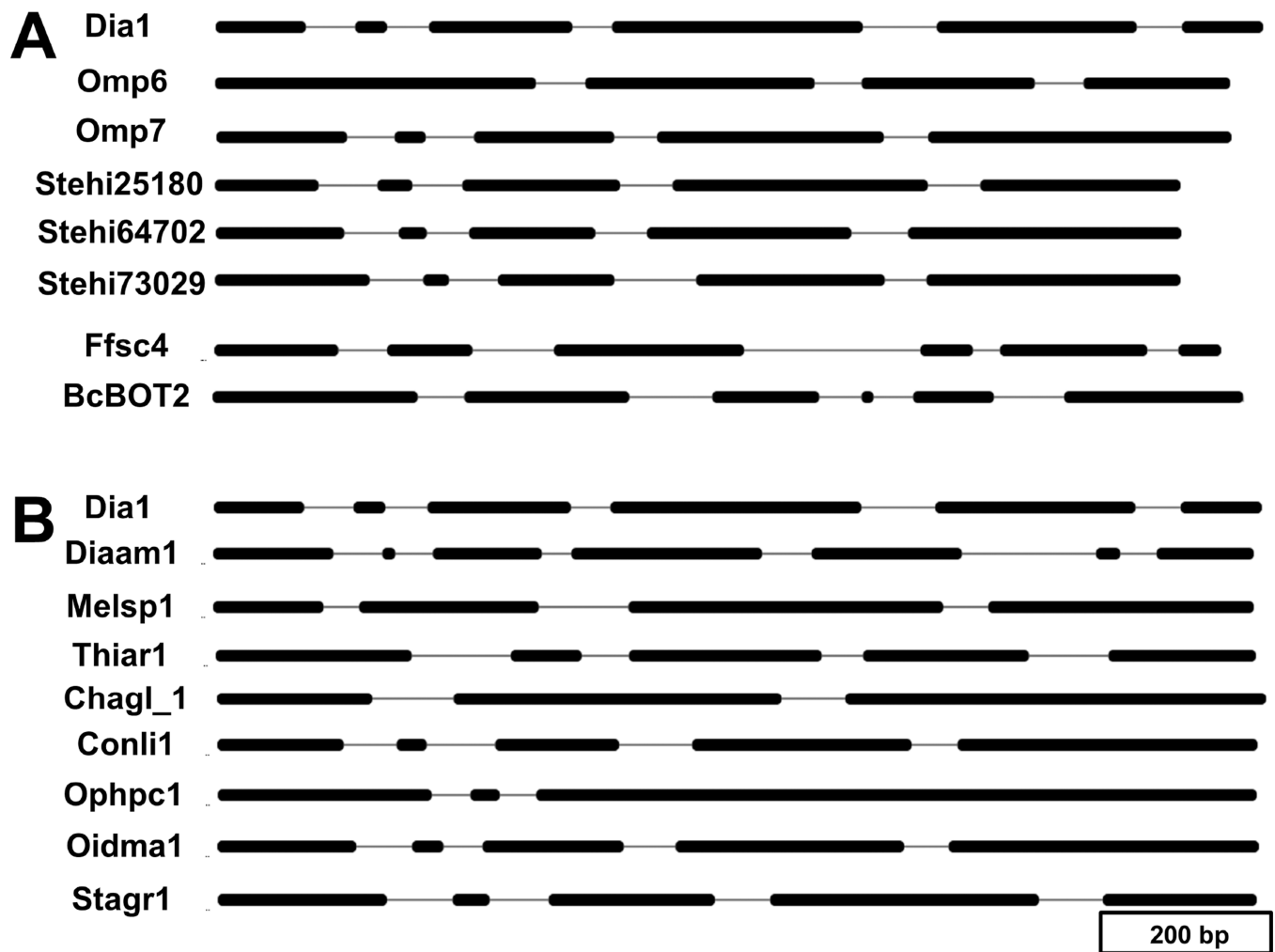
Fig 5. Intron splicing patterns of *Basidiomycota* and *Ascomycota* genes encoding 1,11 cyclizing sesquiterpene synthases differ

(A) The gene architecture of Dia1 is compared to characterized *Basidiomycota* genes encoding 1,11 cyclizing sesquiterpene synthases Omp6, Omp7, Stehi25180, Stehi64702 and Stehi73029; as well as characterized *Ascomycota* genes encoding 1,11 cyclizing sesquiterpene synthases Ffsc4 and BcBOT2. The splicing patterns between characterized *Basidiomycota* and *Ascomycota* genes are different; Dia1 most closely matches the splicing pattern of *Basidiomycota* genes. **(B)** The gene architecture of Dia1 is compared to putative *Ascomycota* genes that may encode sesquiterpene synthases as determined by sequence comparisons. Terpene synthase gene names refer to those indicated in Fig 7. Exons are indicated by black rounded rectangles, introns are indicated by grey lines. The scale bar represents 200 bp.
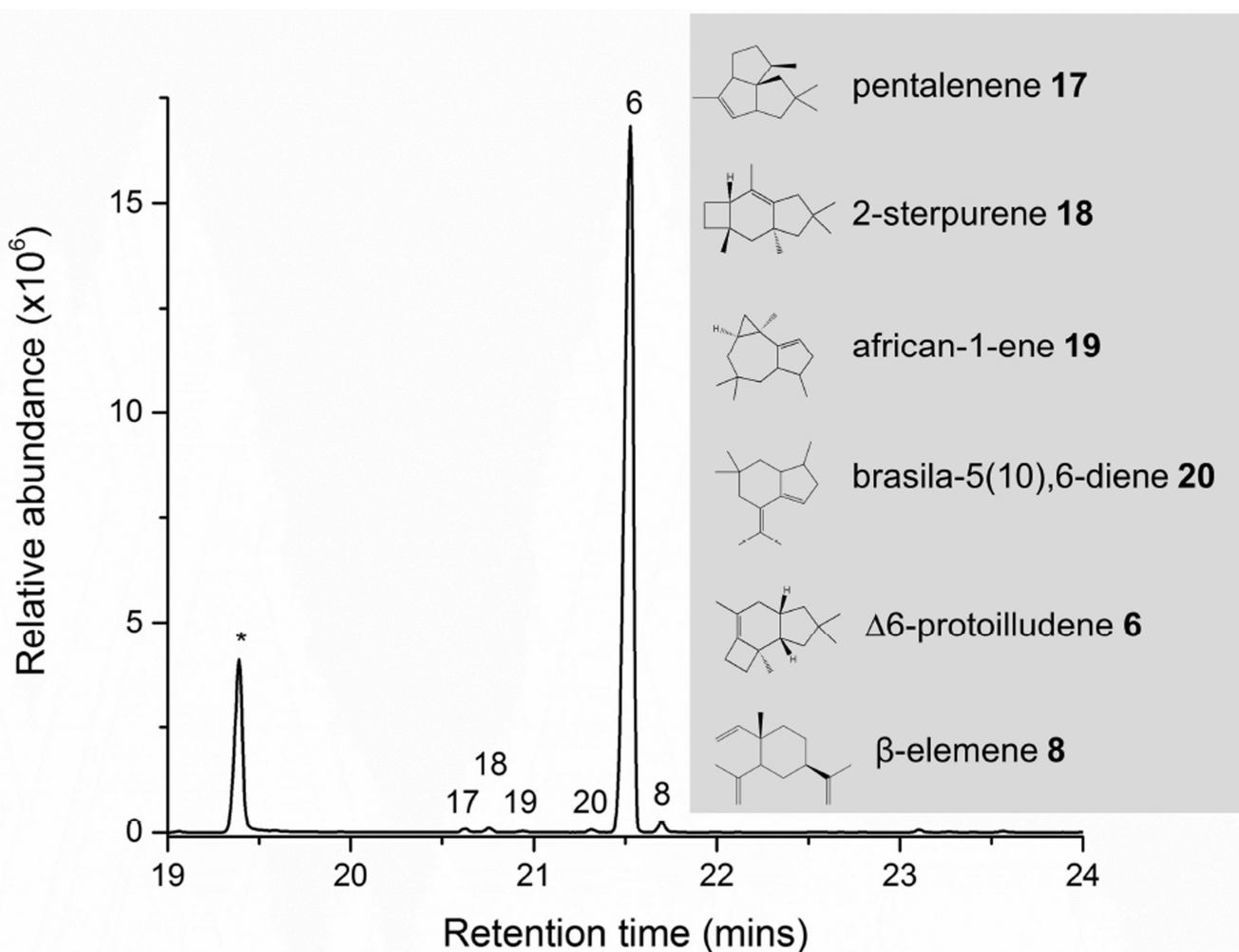
Fig 6. GC/MS analysis of the volatile sesquiterpenes produced by *E. coli* overexpressing Dia1
Headspace analysis of *E. coli* cells overexpressing Dia1 reveals   6-protoilludene **6** as a major peak on the chromatogram, as well as minor peaks corresponding to pentalenene **17**, 2-sterpurene **18**, african-1-ene **19** and brasila-5(10),6-diene **20**, which are all derived from a 1,11 cyclization of (2*E*,6*E*)-FPP. β elemene **8** is a heat induced rearrangment product of a 1,10 cyclization of (2*E*,6*E*)-FPP. Peaks that could be assigned based on comparison of mass spectra to known libraries of terpenes are labelled with numbers, and chemical structures are shown in the grey box. Indole, highlighted with an asterisk (*), is a breakdown product of tryptophan naturally produced by *E. coli*, and serves as an internal standard.
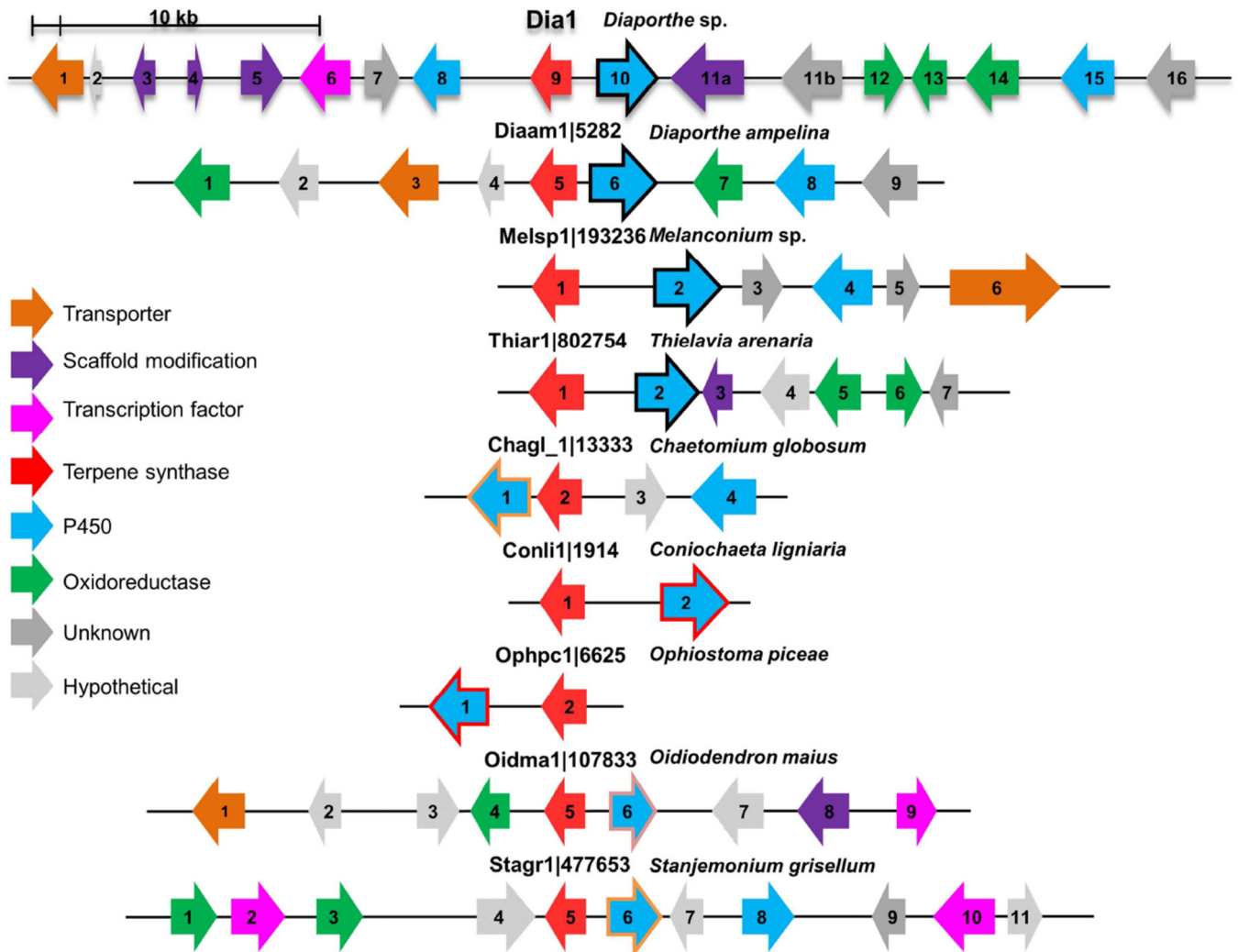
Fig 7. Putative sesquiterpene synthases and biosynthetic gene clusters associated with 6-protoilludene synthase Dia1 and its closely related *Ascomycota* homologs, and classification of the P450 most closely associated with these putative terpene synthases

Homologs of 6-protoilludene synthase Dia1 were identified by BLAST searches of *Ascomycota* genomes in JGI using a cutoff value of 1E-60. Putative terpene synthases are indicated by a red arrow, with the JGI identifier and species name provided above the arrow. Potential biosynthetic genes clustered around the terpene synthase were manually annotated using AUGUSTUS and BLAST. Identified were genes putatively encoding transporters, enzymes involved in scaffold modification, transcription factors, cytochrome P450 monooxygenases and oxidoreductases (colored arrows show their putative functions). Genes with unknown function or annotated as hypothetical are indicated by grey arrows. The numbers within arrows refer to gene model prediction from AUGUSTUS outputs. The P450 most closely associated with the putative terpene synthases (bold arrow outline) was categorized according to CYP family using Dr. Nelson's P450 database and BLAST server (http://drnelson.uthsc.edu/CytochromeP450.html). Dia_g10 (3e-167), Diaam_g6 (1e-128), Melsp_g2 (2e-109), Thiar_g2 (5e-165) belong to CYP family 5032 (black arrow outline) associated with *Basidiomycota*, while the other P450s are most closely related to

Ascomycota CYP families: Chagl_g1 (2e-143), Stagr_g6 (9e-172) belong to CYP family 65 (orange arrow outline); Conli_g2 (0), Ophpc_g1 (0) belong to CYP 582 (rose arrow outline); and Oidma_g6 (2e-52) to CYP 682 (pink arrow outline). E-values for top P450 hits are shown in parentheses as an indicator of significance of CYP family assignments (see Supporting Information for sequences).