Published in final edited form as:

Chembiochem. 2013 December 16; 14(18): . doi:10.1002/cbic.201300349.

Mushroom hunting using bioinformatics: Application of a predictive framework facilitates the selective identification of sesquiterpene synthases in Basidiomycota

Dr. Maureen B. Quin^{[a],†}, Christopher M. Flynn^{[a],†}, Grayson T. Wawrzyn^[a], Dr. Swati Choudhary^[a], and Prof. Claudia Schmidt-Dannert^{*,[a]}

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

Abstract

Basidiomycota represent a diverse source of natural products, particularly the sesquiterpenoids. Recently, the genome sequencing, mining, and subsequent discovery of a suite of sesquiterpene synthases was described in *Omphalotus olearius*. A predictive framework was developed to facilitate the discovery of sesquiterpene synthases in Basidiomycota. Phylogenetic analyses indicated a conservation of both sequence and initial cyclization mechanisms used. Here, the first robust application of this predictive framework is reported. It is used to pursue and selectively identify sesquiterpene synthases that follow a 1,6-, 1,10-, and 1,11-cyclization mechanism in the crust fungus *Stereum hirsutum*. The successful identification and characterization of a 1,6- and a 1,10-cyclizing sesquiterpene synthase, as well as three 1,11-cyclizing Δ -6 protoilludene synthases, is described. This study verifies the accuracy and utility of the predictive framework as a roadmap for the discovery of specific sesquiterpene synthases from Basidiomycota, representing an important step forward in natural product discovery.

Keywords

kinetics; anticancer agents; biosynthesis; natural products; GC/MS

Introduction

Basidiomycota are a rich source of natural products. Mushroom-forming fungi are widespread in nature; current estimates of fungal diversity exceed that of land plants by a ratio of 10:1.^[1,2] The ubiquitous nature of mushrooms can be attributed to their highly active secondary metabolic pathways. Many of the bioactive secondary metabolites produced serve to protect the mushroom from predators, enhance survival under harsh growth conditions, and regulate cell proliferation and differentiation.^[3] The isolation, chemical characterization, and chemical synthesis of bioactive compounds from mushrooms has led to the production of pharmaceuticals with antimicrobial, antifungal, and anticancer applications.^[3,4]

*Address correspondence to: Prof. Claudia Schmidt-Dannert, Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, 1479 Gortner Avenue, St. Paul, MN 55108, USA Tel. 612 625-5782; Fax. 612 625-5780; schmi232@umn.edu.

[†]These authors contributed equally to this work

Supporting information for this article is available on the WWW under <http://www.chembiochem.org> or from the author.

Sesquiterpenoids represent a major group of secondary metabolites isolated from Basidiomycota. These natural products are bioactive compounds due to their unique chemical structures, which result from complex modifications and rearrangements of the precursor sesquiterpene scaffold by accessory modifying enzymes.^[5,6] Sesquiterpenes are produced by a class of enzymes known as sesquiterpene synthases, which cyclize the 15-carbon linear molecule farnesyl pyrophosphate ((2*E*,6*E*)-FPP) via a specific ring closure, producing hundreds of different types of volatile cyclic scaffolds.^[7,8] Despite this diversity of products, all sesquiterpene synthases share a common three dimensional fold and catalytic mechanism. The cyclization cascade is initiated by a divalent metal-ion mediated dephosphorylation of (2*E*,6*E*)-FPP to produce a highly reactive carbocation, which is subjected to a series of rearrangement reactions until a final deprotonation or attack by water. The resulting volatile cyclic product is then released from the catalytic active site.^[9,10] The robust nature and cyclization products of these enzymes makes them ideal candidates for study, not only to further our understanding of carbocation chemistry, but also to develop a toolbox for the biosynthetic production of pharmaceutically relevant compounds that may not be accessible by traditional chemical syntheses.^[11,12]

Only a handful of fungal sesquiterpene synthases have been cloned, purified, and enzymatically characterized.^[13–24] Comprehensive studies are challenging, not only because it can be relatively difficult to cultivate mushrooms under laboratory conditions,^[25] but also because few systematic surveys of sequence information have been conducted. As such, the identification of novel fungal sesquiterpene synthases has been a labor-intensive and unpredictable process. However, the relative affordability of next-generation sequencing technologies^[26] has led to a large increase in the number of sequenced fungal genomes. The availability of sequence information facilitates mining of genomes for the discovery of new sesquiterpene synthases and pathways, as has been applied to the well-studied bacterial^[27–29] and plant terpene synthases.^[30] Our group recently carried out a comprehensive phylogenetic analysis of all sequenced Basidiomycota genomes. It was discovered that sesquiterpene synthases appeared to cluster, not only according to sequence conservation, but also by cyclization mechanism.^[24] If correct, these phylogenetic analyses offer a predictive framework, allowing the focused discovery of novel sesquiterpene synthases based upon cyclization mechanism of choice, streamlining the identification and cloning of novel sesquiterpene synthases which produce desirable natural products.

The goal of this work was to test the reliability of the predictive framework to selectively identify novel sesquiterpene synthases from the crust fungus *Stereum hirsutum*. This fungus was chosen because it has a sequenced genome and according to our bioinformatics analyses, it possesses several putative sesquiterpene synthases. Further, a search in SciFinder® (www.cas.org) shows that compared to other genome sequenced Basidiomycota a number of bioactive natural products have been isolated and characterized from *Stereum* strains. The majority of sesquiterpenoids previously isolated from *Stereum* strains derive from a *trans*-humulyl cation generated from (2*E*,6*E*)-FPP by an initial 1,11-cyclization reaction.^[31–38] The *trans*-humulyl cation is the first committed step in the production of a wide swathe of pharmaceutically relevant sesquiterpene scaffolds made by Basidiomycota.^[5] It was therefore hypothesized that *S. hirsutum* would also be a prolific producer of sesquiterpenes, including those derived from the *trans*-humulyl cation. This work shows that not only was this hypothesis correct, but also that by using the predictive framework, it was possible to selectively clone, heterologously express, and enzymatically characterize 1,6-, 1,10- and 1,11-cyclizing sesquiterpene synthases, including three novel Δ -6 protoilludene synthases. These data serve to validate the predictive framework as a practical and reliable guide, facilitating the directed discovery of sesquiterpene synthase activities in Basidiomycota. The predictive framework will be a valuable resource for future

natural product discovery in Basidiomycota, which will become increasingly important as additional fungal genomes are sequenced.

Results and Discussion

S. hirsutum produces a wide range of sesquiterpenes

Several sesquiterpenoids derived from the *trans*-humulyl cation have been isolated and characterized from the fruiting bodies of *Stereum* sp. [32,33,35,36] and from culture media, [37,38] indicating that members of this genus produce 1,11-cyclizing sesquiterpene synthases. Additionally, sesquiterpenoids derived from a 1,6-cyclization [34] as well as a 1,10-cyclization mechanism [31,35] have also been isolated from *Stereum* sp. (Scheme 1). No sesquiterpene synthase has yet been described or characterized from *Stereum*.

To confirm that the genome sequenced strain *S. hirsutum* is a prolific producer of sesquiterpene scaffolds, the headspace of a liquid culture was analyzed for the presence of volatile hydrocarbons by gas chromatography/mass spectrometry (GC/MS) (Figure 1, Figure S1). Volatile sesquiterpene production was apparent after 8 days, and the relative abundance of these products increased over the sampling period of 21 days. The major sesquiterpene produced was Δ -6 protoilludene **7**, a 1,11-cyclization product, and the second most abundant sesquiterpene was β -elemene **8**, a heat-induced Cope rearrangement product of the 1,10-cyclization product germacrene A. [39] Other abundant volatiles included α -humulene **12**, hirsutene **6**, and pentalenene **4**, all of which are 1,11-cyclization products. [5] Other less abundant products included the 1,10-cyclization product δ -cadinene **13**, and the 1,6-cyclization product sesquisabinene A **11**. These findings confirmed that *S. hirsutum* possesses a number of as-of-yet uncharacterized sesquiterpene synthases that follow a 1,6-, a 1,10- and a 1,11-cyclization mechanism.

Using phylogenetic analyses to predict 1,6-, 1,10-, and 1,11-cyclizing sesquiterpene synthases in *S. hirsutum*

Previously, our group carried out a comprehensive bioinformatic survey and phylogenetic analysis of 40 Basidiomycota genomes using sequence and function data for characterized sesquiterpene synthases from *C. cinereus* and *O. olearius* as a guide. [24] A total of 542 putative sesquiterpene synthases were found. The enzymes formed five distinct clades in a phylogenetic tree, apparently clustering by sequence conservation and cyclization mechanism. Clade I consisted of enzymes (Omp1–3, Cop 1–3) that utilize a 1,10-cyclization of (2*E*,6*E*)-FPP to produce sesquiterpenes derived from a *E,E*-germacradienyl cation. [13,24]

Clade II consisted of enzymes (Cop4, Omp4, Omp5a and 5b) that shared a 1,10-cyclization of (3*R*)- nerolidyl diphosphate (NPP) mechanism, [40,41] producing sesquiterpenes derived from a *Z,E*-germacradienyl cation. Clade III consisted of enzymes believed to share a common 1,11-cyclization of (2*E*,6*E*)-FPP mechanism, producing the *trans*-humulyl cation. Omp6 and Omp7, both of which are Δ -6 protoilludene synthases that were characterized from *O. olearius*, clustered in this group. Cop5, which is a putative pentalenene synthase from *C. cinereus*, also clustered in Clade III. Clade IV consisted of enzymes that shared a 1,6-cyclization of (3*R*)-NPP or (3*S*)-NPP mechanism, [40,42] producing a bisabolyl cation. The α -cupranene synthase Cop6 from *C. cinereus*, and the α/β -barbatene synthase Omp9 as well as the daucene/*trans*-dauca-4(11),8-diene synthase Omp10 from *O. olearius* clustered in this group. Finally, Clade V consisted of enzymes believed to share a 1,6-cyclization mechanism.

We set out to establish whether this apparent phylogenetic clustering according to cyclization mechanism could be used to predict enzyme function from a genomic

perspective for *S. hirsutum*. Multiple sequence alignments and phylogenetic analyses revealed a total of 18 putative sesquiterpene synthases from *S. hirsutum* located in Clades I–IV. Two of the sesquiterpene synthases were clustered in Clade I, five of the sesquiterpene synthases clustered together in Clade II, nine of the sesquiterpene synthases were clustered with sequences in Clade III, and two of the sesquiterpene synthases were located in Clade IV (Figure 2A). Note that none of the sesquiterpene synthases were located in the previously described Clade V.^[24]

The two Clade I putative sesquiterpene synthases, Steh1|45387 and Steh1|167646, clustered together with Omp3, Cop2 and Cop3, likely utilizing a 1,10-cyclization of (2*E*, 6*E*)-FPP to yield a *E,E*-germacradienyl cation-derived product, such as β -elemene **8**, which is a heat-induced Cope rearrangement product of the 1,10-cyclization product germacrene A.^[39]

Of the Clade II putative sesquiterpene synthases, Steh1|111121, Steh1|111127 and Steh1|128017 are clustered most closely with Cop4, Omp4, Omp5a and Omp5b. It was predicted that these would likely be responsible for the production of the δ -cadinene **13** detected in the headspace of *S. hirsutum* cultures, via a *Z,E*-germacradienyl cation from a postulated 1,10-cyclization of (3*R*)-NPP.^[13,24] The other predicted 1,10-cyclizing sesquiterpene synthases Steh1|155443 and Steh1|146390 are located together on a separate branch.

Half of the 18 predicted sesquiterpene synthases are located in Clade III (1,11-cyclization mechanism), substantiating the finding that *S. hirsutum* is a prolific producer of *trans*-humulyl cation derived sesquiterpenes (Figure 1). Steh1|25180, Steh1|73029, Steh1|64702 and Steh1|69906 cluster together, and are closely related to the Δ -6 protoilludene synthases Omp6 and Omp7^[24] (Figure 2B). It was therefore predicted that these putative sesquiterpene synthases would produce the Δ -6 protoilludene **7** detected in the culture headspace, utilizing a 1,11-ring closure mechanism to yield a *trans*-humulyl cation. Steh1|70268 and Steh1|50042 are more closely related to the putative pentalenene synthase Cop5, and therefore may produce pentalenene **4**. The other three putative 1,11-cyclizing sesquiterpene synthases in Clade III, Steh1|122776, Steh1|52743, and Steh1|161672, did not cluster tightly with other members, and are instead located on separate branches. One or more of these enzymes may be responsible for the production of α -humulene **12** or hirsutene **6** by *S. hirsutum*.

Finally, the two members of Clade IV, Steh1|113028 and Steh1|159379, are located on a separate branch and are not as tightly clustered as members of other clades. However, they clustered closest to Cop6, and were therefore predicted to utilize a 1,6-cyclization of (3*R*)-NPP or (3*S*)-NPP to yield a bisabolyl cation-derived product such as sesquisabinene A **11**, in a similar fashion to the α -cuprenene synthase Cop6.^[19,24]

Cloning the suite of predicted sesquiterpene synthases from *S. hirsutum* cDNA

To test the accuracy of our predictive framework, the genes encoding the putative enzymes were cloned and expressed heterologously. Gene predictions of the 18 putative sesquiterpene synthases from the *S. hirsutum* genome sequence were refined by manual reannotation. All of the potential transcripts obtained for each sesquiterpene synthase-encoding gene were first aligned against known sequences of previously isolated sesquiterpene synthases to identify the most likely transcript(s) encoding a functional sesquiterpene synthase. Using gene specific primers designed for the manually predicted gene models, PCR amplification products of several splice variants were successfully obtained from cDNA for Steh1|159379, Steh1|113028, Steh1|128017, Steh1|25180, Steh1|64702 and Steh1|73029. Incorrectly spliced isoforms were excluded from further analysis due to the presence of internal stop codons or frameshift mutations. Several

unsuccessful attempts were made to obtain a correctly spliced version of Steh1|113028, which is located adjacent to Steh1|159379, leading the authors to believe that either the gene prediction is incorrect or that it is a pseudogene. Notably, despite exhaustive efforts, PCR amplification products could not be obtained from cDNA for any of the other 12 predicted sesquiterpene synthases. These genes may not be expressed under the growth conditions used,^[43] may be pseudogenes,^[44] or may not be accurately predicted using the fungal gene prediction models available (Table S1). Furthermore, Basidiomycota genes typically have many small introns and exons that are difficult or impossible to predict using current fungal gene models, making the task of correctly predicting splice variants of genes particularly tedious and challenging,^[44] and in most cases preclude the feasibility of gene synthesis approaches.

In order to confirm that members of certain clades would follow the predicted cyclization mechanisms, the PCR products of the putative sesquiterpene synthases were cloned into our in-house expression vector pUCBB,^[45] which allows for constitutive expression of genes in the heterologous host *E. coli*. The putative sesquiterpene synthases were functionally characterized by GC/MS analysis of the volatile components of the headspace from recombinant BL21 (DE3) *E. coli* cultures.^[46,47]

Predicted 1,6- and 1,10-cyclizing sesquiterpene synthases produce bisabolyl cation and putative *Z,E*-germacradienyl cation-derived terpenes

The culture headspace of cells expressing the Clade IV member Steh1|159379 contained one major peak which was identified as β -barbatene, as well as minor peaks which were identified as α -barbatene and α -cuprenene (Figure 3A). Although these peaks were not identified in the *S. hirsutum* headspace analysis perhaps due to further modification into a non-volatile compound(s) (Figure 1), identical retention indices and mass fragmentation patterns were obtained for the major peaks from the culture headspace of *E. coli* cells expressing the previously characterized α/β -barbatene synthase Omp9 from *O. olearius*,^[24] as well as the α -cuprenene synthase Cop6 from *C. cinereus* ^[19] (Figure S2). These data confirm that Steh1|159379 utilizes a 1,6-ring closure of (3*R*)-NPP or (3*S*)-NPP to yield bisabolyl cation-derived products,^[40,42] as had been predicted using our phylogenetic analyses.

The headspace of cells expressing the Clade II member Steh1|128017 contained one major peak, δ -cadinene **13**, as well as several minor peaks including β -copaene, sativene, γ -muurolene, α -muurolene, β -cubebene, germacrene D, and sesquisabinene A **11** (Figure 3B, Figure S3). Of these, δ -cadinene **13** and sesquisabinene A **11** were identified in the headspace of *S. hirsutum* cultures (Figure 1). Retention indices and mass fragmentation patterns of the δ -cadinene **13** peak matched those obtained for the δ -cadinene synthases Omp4 ^[24] and Cop4,^[13,19] and the γ -cadinene synthase Omp5a ^[24] (Figure S3). These findings suggest that Steh1|128017 utilizes a putative 1,10-cyclization of (3*R*)-NPP to yield sesquiterpenes derived from a *Z,E*-germacradienyl cation. However, sesquisabinene A **11** is proposed to derive from a 1,6-cyclization mechanism.^[48] Also, a recent study raised into question whether δ -cadinene synthase utilized a 1,10- or a 1,6-ring closure, although conclusive evidence could not be obtained to confirm which mechanism was correct.^[40] It is therefore not clear whether Steh1|128017 follows dual cyclization mechanisms, or whether the putative 1,6-cyclization mechanism is actually a 1,10-ring closure, or *vice versa*. These unanswered questions highlight the versatile nature of the terpene synthases as an enzyme class. Taken together, these data confirm the accuracy of our predictive framework in the selective identification of sesquiterpene synthases that follow the currently accepted 1,6- and 1,10-cyclization mechanisms of (3*R*)-NPP and (3*S*)-NPP.

Predicted 1,11-cyclizing sesquiterpene synthases produce *trans*-humulyl cation derived terpenes

The *trans*-humulyl cation is the precursor to a wide range of pharmaceutically relevant sesquiterpenoids,^[5] making the directed discovery of 1,11-cyclizing sesquiterpene synthases particularly relevant. The culture headspace of *E. coli* cells expressing the Clade III members Steh1|25180, Steh1|64702 and Steh1|73029 all contained one major peak which was identified as Δ -6 protoilludene **7**, further confirming the accuracy of our predictive framework (Figure 4). Identical retention indices and mass fragmentation patterns were obtained for the major peak from the culture headspace of *E. coli* cells expressing the previously characterized Δ -6 protoilludene synthases Omp6 and Omp7 from *O. olearius*^[24] (Figure S4). Other minor peaks were also observed in the culture headspace of Steh1|64702 and Steh1|73029. One of these peaks was identified as β -elemene **8** (Figure 4), also produced by Omp3^[24] (Figure S5).

These results support our hypothesis that putative sesquiterpene synthases in Clade III would likely follow a 1,11-cyclization mechanism, producing *trans*-humulyl derived products. Quantum-chemical calculations and isotopic labeling experiments indicate that a 1,11-ring closure of (2*E*,6*E*)-FPP yields a highly reactive *trans*-humulyl carbocation. A 1,2-hydride shift followed by further ring closures results in the protoilludyl cation, which undergoes a final deprotonation, resulting in Δ -6 protoilludene **7** (Scheme 2).^[49,50] Interestingly, the presence of a small amount of β -elemene **8** in the headspace of Steh1|64702 and Steh1|73029 expressing cultures suggests that these sesquiterpene synthases are less specific than Steh1|25180, and the Δ -6 protoilludene synthases Omp6 and Omp7 from *O. olearius*, which display only one single peak when analyzed by GC/MS, both *in vivo* and *in vitro*. Likewise, heterologous expression of the Δ -6 protoilludene synthase Pro1 from *A. gallica* resulted in one peak, suggesting that it is also a highly specific enzyme, although the activity of this enzyme has not yet been fully studied *in vitro*.^[17] Steh1|25180, Steh1|64702 and Steh1|73029 sesquiterpene synthases appear to preferentially utilize a 1,11-cyclization mechanism, but Steh1|64702 and Steh1|73029 also simultaneously follow a 1,10-cyclization pathway to yield a small amount of β -elemene **8**, a germacrene A-derived product (Scheme 2), again highlighting the mechanistic versatility of these enzymes.

Kinetic characterization of Steh1|25180, Steh1|64702 and Steh1|73029 highlights differences in catalytic efficiencies

To further understand the cyclization mechanism employed by Steh1|25180, Steh1|64702 and Steh1|73029, the sesquiterpene synthases were purified and their activities were characterized *in vitro*. The enzymes were incubated with (2*E*,6*E*)-FPP under standard assay conditions and the product profiles were determined by GC/MS. Steh1|25180, Steh1|64702 and Steh1|73029 all maintained their preference for a 1,11-cyclization mechanism, producing high levels of Δ -6 protoilludene **7** as their sole sesquiterpene product (Figure 5). Therefore, the low level of the 1,10-cyclization derived product β -elemene **8**, which was produced during *in vivo* heterologous expression of Steh1|64702 and Steh1|73029 must be the result of specific, unknown intracellular conditions. Future studies to determine the conditions that regulate the cyclization mechanism of these Δ -6 protoilludene synthases would provide novel insights into the biochemical mechanisms underlying the product specificity of sesquiterpene synthases.

Additionally, the kinetic parameters for purified Steh1|25180, Steh1|64702 and Steh1|73029 were calculated with (2*E*,6*E*)-FPP as a substrate (Table 1). Catalytic turnover rates are similar between the three enzymes, however their affinities for (2*E*,6*E*)-FPP varied such that the catalytic efficiency of Steh1|73029 was two-fold greater than that of Steh1|64702, and four-fold greater than that of Steh1|25180. These findings emphasize the relative

importance of these apparently identical enzymes to the fungal host. Δ -6 protoilludene **7** is the major volatile sesquiterpene product of *S. hirsutum* (Figure 2). It could be envisioned that *in vivo* the three Δ -6 protoilludene synthases act in a cooperative fashion to maintain sufficient levels of this key intermediate sesquiterpene scaffold, which is a precursor to a wide swathe of bioactive sesquiterpenoids produced by *Stereum* sp. [32,33,35,36] (Scheme 1).

Insights into the production of *trans*-humulyl-derived sesquiterpenoids by Basidiomycota

The majority of sesquiterpenoids isolated from *Stereum* strains are derived from a modified *trans*-humulyl cation scaffold (Scheme 1),^[5] indicating that *Stereum* is not only a prolific producer of 1,11-cyclizing sesquiterpene synthases, but also of accessory scaffold modifying enzymes. *Stereum* strains also produce the stereumins, which are derived from a modified δ -cadinene scaffold,^[31,35] and methoxylaricinolic acid, which is derived from a modified drimane scaffold,^[34] suggesting that *Stereum* strains have several highly active secondary metabolic pathways. Gene clustering of biosynthetic secondary metabolic pathways is a hallmark characteristic of fungi.^[51] It was therefore predicted that the *S. hirsutum* sesquiterpene synthases would be located in gene clusters with the accessory scaffold modifying enzymes that are responsible for the production of the final bioactive sesquiterpenoids.

To identify potential sesquiterpene biosynthetic gene clusters, sequence regions 10–15 kb upstream and downstream of the genes encoding the functionally cloned and expressed sesquiterpene synthases were manually reannotated. The β -barbatene synthase Steh1|159379 and the δ -cadinene synthase Steh1|128017 are located in putative biosynthetic gene clusters, containing oxidoreductases, but lacking P450 monooxygenases typically associated with the biosynthesis of bioactive natural products (Figure S6). The Steh1|128017 cluster may be involved in the biosynthesis of the stereumins which are known to be produced by *Stereum* strains and are proposed to be derived from δ -cadinene (Scheme 1).^[31,35,52] Although no P450 monooxygenases are associated with this cluster, and are most likely required for stereumin biosynthesis, additional scaffold modifying enzymes could be located on satellite clusters known from other fungal natural product biosynthetic pathways such as for trichothecene mycotoxin biosynthesis by *Fusarium graminearum*.^[53]

The three Δ -6 protoilludene synthases, Steh1|25180, Steh1|64702 and Steh1|73029 are located in large gene clusters which encode for a range of scaffold modifying enzymes, including several P450 monooxygenases (Figure 6, Data S1). Steh1|73029 is located in the smallest biosynthetic gene cluster, with three putative P450 monooxygenases, two oxidoreductases, two aldo-keto reductases, a FAD binding protein, a hydrolase, two transferases and a transporter. Steh1|64702 biosynthetic gene cluster also contains three P450 monooxygenases, two oxidoreductases, two reductases, an O-methyltransferase, a dehydrogenase and a transporter. Steh1|25180 is located in the largest biosynthetic gene cluster, with three P450 monooxygenases, although two of these may result from a gene duplication event. Additionally, this gene cluster has fewer oxidoreductases, no transferases, and no transporter (Figure 6). To date, the Δ -6 protoilludene-derived bioactive sesquiterpenoids that have been isolated from *Stereum* strains are the stereostreins (Scheme 1).^[32,33] However, the identification of three large gene clusters associated with the three functional Δ -6 protoilludene synthases, Steh1|25180, Steh1|64702 and Steh1|73029 would suggest that *Stereum hirsutum* has the potential of producing a number of yet-to-be identified bioactive sesquiterpenoids from the Δ -6 protoilludene scaffold. Bio-synthesis of other *trans*-humulyl cation-derived bioactive sesquiterpenoids such as the sterpuranes isolated from *Stereum* sp. [54–57] and the hirsutenes from *Stereum hirsutum* [37,38] (Scheme 1) likely involve putative Clade III sesquiterpene synthase members that could not be cloned from cDNA (Figure 2).

The strained cyclobutyl ring of Δ -6 protoilludene **7** is believed to undergo different secondary rearrangements upon activation by scaffold modifying enzymes such as P450 mono-oxygenases, yielding the final bioactive sesquiterpenoid.^[5,6,13,58] Significantly, all of the *S. hirsutum* Δ -6 protoilludene synthase gene clusters share in common the presence of several P450 monooxygenases. Additionally, the Δ -6 protoilludene-derived stereoreins isolated from *Stereum* sp. have had at least one oxygen atom inserted, typically at the C-8 position of the cyclohexene ring, yielding a ketone (Scheme 1). Likewise, the previously described biosynthetic gene clusters associated with the Δ -6 protoilludene synthases Omp6 and Omp7 both contain at least one P450 monooxygenase, and many of the illudin sesquiterpenoids isolated from *O. olearius* are also oxygenated at this same C-8 position.^[24] The similarity of the modifications to the Δ -6 protoilludene scaffold produced by *O. olearius* and *Stereum* strains may represent a common functionality of their sesquiterpene synthase associated P450 monooxygenases, perhaps indicating a common evolutionary ancestor.

In order to establish whether P450 monooxygenases responsible for modifying the Δ -6 protoilludene scaffold could share a common ancestor, a comparative sequence analysis was undertaken. Phylogenetic analyses were carried out by aligning the amino acid sequences of the predicted P450 monooxygenases associated with the Δ -6 protoilludene synthases Omp6 and Omp7^[24] with those identified in the Steh1|25180, Steh1|64702 and Steh1|73029 clusters (Figure 6). Interestingly, the P450 monooxygenases associated with Steh1|64702 are all closely related to those associated with Omp6 and Omp7. Additionally, one of the P450 monooxygenases from the Steh1|73029 gene cluster (Steh1|73029h) is closely related to the Omp6 and Omp7 P450 monooxygenases, while the other two P450 monooxygenases (Steh1|73029j and Steh1|73029o) are more distantly related. Notably, the Steh1|25180 P450 monooxygenases are all located on a separate branch, and are not closely associated with other P450 monooxygenases, resulting from a divergence in sequence, and potentially in function (Figure 7). Given the degree of sequence conservation between the predicted Δ -6 protoilludene synthase associated P450 monooxygenases, as well as the three characterized Δ -6 protoilludene synthases in *S. hirsutum*, it could be hypothesized that their presence results from a process of horizontal gene transfer. However, the question as to whether the apparent manifold nature of the Δ -6 protoilludene synthases and their associated biosynthetic gene clusters indicates a difference in functionality in the production of the different bioactive compounds derived from a modified Δ -6 protoilludene scaffold, known to be produced by *Stereum* sp. (Scheme 1), remains open to speculation.

Knowledge of the sequence conservation of the P450 monooxygenases associated with Δ -6 protoilludene synthases could be used to guide biosynthetic pathway elucidation and in particular, identification of the steps leading to the rearrangement of the cyclobutyl ring of Δ -6 protoilludene that give rise to the diverse humulyl-derived bioactive sesquiterpene scaffolds made by Basidiomycota. In addition, knowledge of key P450 function and sequence conservation could open paths for the engineering of biosynthetic routes to novel bioactive sesquiterpenoids through combinatorial approaches and/or P450 protein engineering.

Conclusions

This work serves as a proof-of-concept of our predictive framework as an accurate tool for the directed discovery of sesquiterpene synthases based upon cyclization mechanism of choice. Phylogenetic analyses have uncovered several novel sesquiterpene synthases in the Basidiomycete *S. hirsutum*. These bioinformatic studies enabled the selective identification of sesquiterpene synthases which were predicted to follow a 1,6- and a 1,10-cyclization mechanism, producing bisabolyl cation and *Z,E*-germacradienyl cation derived products, respectively, as well as enzymes that utilize a 1,11-cyclization mechanism, producing *trans*-

humulyl cation derived scaffolds. Biosynthetic gene clusters encoding for the 1,11-cyclizing sesquiterpene synthases and their accessory modifying enzymes have been described. These sesquiterpene synthases have been cloned, expressed, and their activities characterized. GC/MS analysis confirmed the discovery of a β -barbatene synthase, which follows the predicted 1,6-cyclization mechanism, and a δ -cadinene synthase, which follows the predicted 1,10-cyclization mechanism. Additionally, three novel Δ -6 protoilludene synthases, which follow the predicted 1,11-cyclization mechanism, have been characterized. The kinetic parameters show these enzymes to be highly active, producing significant amounts of Δ -6 protoilludene **7**. However, Steh1|64702 and Steh1|73029 appear to be less specific than other previously described Δ -6 protoilludene synthases, and are capable of simultaneously producing a 1,10-cyclization derived scaffold, β -elemene **8** when expressed in a heterologous host. The discovery and characterization of these Δ -6 protoilludene synthases, which are located in large biosynthetic gene clusters, furthers the development of a biosynthetic toolkit. It is envisioned that this toolkit could be applied in a combinatorial approach to evolve and build robust biocatalysts for the production of specific protoilludene-derived bioactive compounds, such as new anticancer illudin derivatives.^[59] At present, many of these compounds are not accessible due to the inhibitory cost and difficulty in chemical syntheses. Producing the compounds biologically is a much more cost effective and efficient option. These results represent an important step for natural product discovery in fungi, which with the predictive framework described herein, has the potential to become a much more efficient, streamlined, and strategic enterprise.

Experimental Section

Chemicals and reagents

All chemicals and reagents were purchased from Sigma-Aldrich (St. Louis, MO), unless indicated otherwise.

Homolog identification, phylogenetic tree construction and biosynthetic cluster prediction

Sequences for the biochemically characterized sesquiterpene synthases from *C. cinereus* and *O. olearius*^[13,24] were aligned with the putative sesquiterpene synthases from *S. hirsutum* FP-91666 SS1 v1.0 using the gene models associated with this genome, provided by the Joint Genome Institute (JGI)^[60]. Alignments were computed using ClustalW^[61] and phylogenetic analyses were conducted using MEGA version 5.2^[62] using the default parameters for the Neighbor-Joining method^[63] with a bootstrap test of phylogeny (500 replicates). Alignments were manually examined for the presence and proper alignment of the conserved metal-binding motifs characteristic for sesquiterpene synthases.^[7] For the identification of putative sesquiterpene biosynthetic gene clusters, each of the scaffolds containing the 18 sesquiterpene synthases from *S. hirsutum* were analyzed for the presence of nearby enzymes implicated in secondary metabolite biosynthesis. A previously described workflow was used for gene identification and splicing predictions.^[47] For prediction of the putative function of cluster genes, the top conserved domain hit (CDD) at NCBI was used. For the comparative analysis of P450s, the sequences of Omp6-i, -j, -h, and Omp7-b^[24] were used in BLASTp searches of the *S. hirsutum* FP-91666 SS1 v1.0 genome via the JGI server, using an E value threshold of 10. The results were manually scanned 20 kb up/downstream to confirm association with putative sesquiterpene synthases using Augustus.^[64] The putative amino acids sequences were aligned and phylogenetic analyses were carried out using MEGA version 5.2^[62] using the default parameters for the Neighbor-Joining method^[63] with a bootstrap test of phylogeny (500 replicates).

Growth of *S. hirsutum* and headspace analysis of volatile compounds

Stereum hirsutum FP-91666 SS1 was obtained as a gift from Dr. Robert Blanchette (University of Minnesota, USA) and was grown on potato dextrose agar (PDA) for 2–3 weeks in the dark prior to inoculating liquid cultures with a agar plug (0.5 cm²). Liquid cultures (100 mL) were cultivated in foil-sealed flasks (500 mL) for up to 21 days in Rich Medium (20 g/L malt extract, 20 g/L glucose, 5 g/L peptone) at room temperature, in the dark. Volatile terpenoids which had accumulated in the headspace of flasks containing *S. hirsutum* growing on Rich Medium were sampled every 4–5 days for 15 minutes by inserting a 100 µm poly-dimethylsiloxane fiber (Supelco, Bellefonte, PA) solid phase microextraction fiber (SPME) through the aluminum foil seal of the culture flask, as performed previously.^[13,24] Similarly, following 21 days of growth, the culture supernatant was clarified by centrifugation at 4000 rpm for 10 minutes, and sampled for non-volatile terpenoid accumulation using a 100 µm poly-dimethylsiloxane-divinylbenzene SPME fiber. Terpenoids were separated and analyzed by GC/MS, described below.

S. hirsutum mRNA extraction and cDNA preparation

Mycelial *S. hirsutum* cultivated for 21 days in liquid Rich Medium was dried manually by compression 4–5 times on filter paper, dissected, aliquoted into eppendorf tubes, flash-frozen in liquid N₂, and stored at –80 °C. Frozen tissue was ground into a fine powder using a N₂ cooled, sterile mortar and pestle. mRNA was then extracted using TRIzol® Reagent (Life Technologies, Grand Island, NY) following the manufacturer's procedures. Briefly, powdered tissue was incubated in TRIzol Reagent for 5 minutes, separated by centrifugation from insoluble components, extracted with chloroform, and precipitated with isopropanol. The resulting precipitated nucleotides were washed with 75 % ethanol, and resuspended in DNA/RNA-free water. Single-stranded *S. hirsutum* cDNA was synthesized using SuperScript™ III First Strand Synthesis System for RT-PCR (Life Technologies, Grand Island, NY) utilizing oligo(dT)₂₀ primers for RT-PCR, followed by RNase H treatment.

Cloning of sesquiterpene synthases from *S. hirsutum* cDNA

Following initial phylogenetic analyses to predict sesquiterpene synthases in *S. hirsutum*, gene predictions were further refined by manually re-scanning the genomic sequence of each putative sesquiterpene synthase and 10–15 kb up/downstream of open reading frames using Augustus.^[64] Potential sesquiterpene synthase-encoding transcripts were then manually aligned against previously isolated sesquiterpene synthases using MEGA version 5.2 to identify the most likely transcript(s) encoding the functional sesquiterpene synthase. A comparison between initial JGI predictions and cloned ORF and protein sequences are presented in Tables S1 and S2, respectively. Finally, gene specific primers (Table S1) were used to amplify sesquiterpene synthase-encoding genes from previously prepared *S. hirsutum* cDNA. Genes were PCR amplified from *S. hirsutum* cDNA using PfuUltra High Fidelity Polymerase (Agilent Technologies, Inc., Santa Clara, CA). PCR products were cloned using the Zero Blunt® TOPO system (Life Technologies, Grand Island, NY) and were transformed into the supplied MachI™ competent cells. Positive clones were isolated by blue-white colony screening, and plasmids were isolated using the Wizard® Plus SV Miniprep kit (Promega, Madison, WI). Expected DNA sequences were confirmed by single pass DNA sequencing (ACGT Inc., Wheeling, IL). The genes were then subcloned into the pUCBB vector^[45], to allow constitutive expression of genes in the heterologous host *E. coli*, and into the pET32b(+) vector system (EMD Millipore, Billerica, MA), to allow inducible expression in frame with a thioredoxin tag to improve solubility and a His₆-tag for purification purposes.

Sampling of volatile compounds produced by *E. coli* cultures expressing putative sesquiterpene synthases

Activity of putative sesquiterpene synthases was confirmed by heterologous expression in *E. coli* BL21 (DE3) cells, which were transformed with pUCBB constructs. Single colonies were isolated on Lysogeny Broth (LB) agar plates supplemented with ampicillin (100 µg/mL), and were used to inoculate LB broth (50 mL). Cultures were incubated for 16 hours at 30 °C with shaking at 220 rpm. The headspace of the cultures was sampled for 10 minutes by SPME followed by GC/MS analysis as described below.

Gas chromatography/mass spectrometry analysis

GC/MS analysis was conducted on an HP GC 7890A coupled to an anion-trap mass spectrometer HP MSD triple axis detector (Agilent Technologies, Santa Clara, CA). Separation of compounds was performed using a HP-5MS capillary column (30 m × 0.25 mm × 1.0 µm) with an injection port temperature of 250 °C and helium as a carrier gas. The oven temperature started at 60 °C and was increased 10 °C min⁻¹ to a final oven temperature of 250 °C. Mass spectra were scanned in the range of 5–300 atomic mass units at 1 s intervals. Compounds produced were identified by first calibrating the GC/MS with a C8–C20 alkane mix. Retention indices and mass spectra of compound peaks were compared to reference spectra in MassFinder's (software version 4) terpene library [65] and in the National Institute of Technology (NIST) standard reference database as described previously.[24,47]

Expression and purification of Steh1|25180, Steh1|64702 and Steh1|73029

Rosetta (DE3) pLysS cells were transformed with pET32b (+) constructs of Steh1|25180, Steh1|64702 and Steh1|73029. Single colonies were isolated on LB agar supplemented with ampicillin (100 µg/mL) and chloramphenicol (50 µg/mL), and were used to inoculate LB starter cultures (50 mL). Large-scale expression was carried out in LB (1000 mL). Cultures were incubated at 30 °C with shaking at 220 rpm until an OD₆₀₀ 0.6 was reached. Protein expression was induced by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) (0.5 mM), and cells were incubated for a further 3 hours at 30 °C. Cells were harvested by centrifugation, lysed by sonication and soluble proteins were separated from the cell slurry by centrifugation. The His₆-tagged proteins were bound to a HisTrapTM FF column (GE Healthcare, Pittsburgh, PA) in Buffer A (50 mM Tris.HCl, 100 mM NaCl, 5 mM imidazole (pH 7.5)), and were eluted in Buffer B (50 mM Tris.HCl, 100 mM NaCl, 250 mM imidazole (pH 7.5)). Following overnight cleavage of the thioredoxin tag by enterokinase (EMD Millipore, Billerica, MA), the tags were separated from the sesquiterpene synthase by passage over a HisTrapTM FF column. The proteins were purified to homogeneity by passage over a S200 10/300 GL size exclusion column preequilibrated with Buffer C (50 mM Tris.HCl pH 7.5, 100 mM NaCl, 1 mM β-mercaptoethanol, 10 % (v/v) glycerol).

Enzyme characterization of purified Steh1|25180, Steh1|64702 and Steh1|73029

In vitro activity assays with purified sesquiterpene synthases were performed in Buffer D (20 mM Tris-HCl, 200 mM NaCl, 10 mM MgCl₂ (pH 8.0)) in a final reaction volume of 100 µL. Sesquiterpene synthases (10 µg) were incubated with (2*E*,6*E*)-FPP (2 µM) in a sealed glass vial, and reactions were carried out at 21 °C for 16 hours. The headspace of the reaction vessel was sampled for 10 min by SPME followed by GC/MS analysis as described above. The kinetic parameters of Steh1|25180, Steh1|64702 and Steh1|73029 were determined using previously published protocols.[24,66] Briefly, sesquiterpene synthases were incubated with (2*E*,6*E*)-FPP at a range of concentrations (1–100 µM). Enzyme activity was monitored using the PiPer PyrophosphateTM enzyme coupled spectrophotometric assay

(Invitrogen, Carlsbad, CA), which measures the release of pyrophosphate (PPi). Control reactions without enzyme and without substrate were also carried out, and all reactions were performed in triplicate. The kinetic parameters were determined by nonlinear regression analysis of the data in Origin 9.0.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by the National Institutes of Health Grant GM080299 (to C.S.-D.).

References

1. Blackwell M. American journal of botany. 2011; 98:426–438. [PubMed: 21613136]
2. O'Brien HE, Parrent JL, Jackson JA, Moncalvo JM, Vilgalys R. Applied and environmental microbiology. 2005; 71:5544–5550. [PubMed: 16151147]
3. Zhong JJ, Xiao JH. Adv Biochem Eng Biotechnol. 2009; 113:79–150. [PubMed: 19475376]
4. Wasser SP. Appl Microbiol Biotechnol. 2011; 89:1323–1332. [PubMed: 21190105]
5. Abraham WR. Curr Med Chem. 2001; 8:583–606. [PubMed: 11281843]
6. Fraga BM. Nat Prod Rep. 2011; 28:1580–1610. [PubMed: 21808787]
7. Christianson DW. Chem Rev. 2006; 106:3412–3442. [PubMed: 16895335]
8. Miller DJ, Allemann RK. Nat Prod Rep. 2012; 29:60–71. [PubMed: 22068697]
9. Lesburg CA, Caruthers JM, Paschall CM, Christianson DW. Current opinion in structural biology. 1998; 8:695–703. [PubMed: 9914250]
10. Yamada Y, Cane DE, Ikeda H. Methods Enzymol. 2012; 515:123–162. [PubMed: 22999173]
11. Bhatti HN, Zubair M, Rasool N, Hassan Z, Ahmad VU. Natural product communications. 2009; 4:1155–1168. [PubMed: 19769003]
12. Wilkinson B, Micklefield J. Nature chemical biology. 2007; 3:379–386.
13. Agger S, Lopez-Gallego F, Schmidt-Dannert C. Molecular microbiology. 2009; 72:1181–1195. [PubMed: 19400802]
14. Cane DE, Kang I. Archives of biochemistry and biophysics. 2000; 376:354–364. [PubMed: 10775423]
15. Cane DE, Wu Z, Proctor RH, Hohn TM. Archives of biochemistry and biophysics. 1993; 304:415–419. [PubMed: 8346917]
16. Caruthers JM, Kang I, Rynkiewicz MJ, Cane DE, Christianson DW. The Journal of biological chemistry. 2000; 275:25533–25539. [PubMed: 10825154]
17. Engels B, Heinig U, Grothe T, Stadler M, Jennewein S. The Journal of biological chemistry. 2011; 286:6871–6878. [PubMed: 21148562]
18. Faraldos JA, Gonzalez V, Allemann RK. Chemical communications. 2012; 48:3230–3232. [PubMed: 22349314]
19. Lopez-Gallego F, Agger SA, Abate-Pella D, Distefano MD, Schmidt-Dannert C. Chembiochem: a European journal of chemical biology. 2010; 11:1093–1106. [PubMed: 20419721]
20. Pinedo C, Wang CM, Pradier JM, Dalmais B, Choquer M, Le Pecqueur P, Morgant G, Collado IG, Cane DE, Viaud M. ACS chemical biology. 2008; 3:791–801. [PubMed: 19035644]
21. Proctor RH, Hohn TM. The Journal of biological chemistry. 1993; 268:4543–4548. [PubMed: 8440737]
22. Rynkiewicz MJ, Cane DE, Christianson DW. Proceedings of the National Academy of Sciences of the United States of America. 2001; 98:13543–13548. [PubMed: 11698643]
23. Shishova EY, Yu F, Miller DJ, Faraldos JA, Zhao Y, Coates RM, Allemann RK, Cane DE, Christianson DW. The Journal of biological chemistry. 2008; 283:15431–15439. [PubMed: 18385128]

24. Wawrzyn GT, Quin MB, Choudhary S, Lopez-Gallego F, Schmidt-Dannert C. *Chem Biol*. 2012; 19:772–783. [PubMed: 22726691]
25. Elisashvili V. *International journal of medicinal mushrooms*. 2012; 14:211–239. [PubMed: 22577974]
26. Shendure J, Lieberman Aiden E. *Nature biotechnology*. 2012; 30:1084–1094.
27. Cane DE, Ikeda H. *Accounts of chemical research*. 2012; 45:463–472. [PubMed: 22039990]
28. Citron CA, Gleitzmann J, Laurenzano G, Pukall R, Dickschat JS. *Chembiochem: a European journal of chemical biology*. 2012; 13:202–214. [PubMed: 22213220]
29. Keeling CI, Weisshaar S, Ralph SG, Jancsik S, Hamberger B, Dullat HK, Bohlmann J. *BMC plant biology*. 2011; 11:43. [PubMed: 21385377]
30. Chen F, Tholl D, Bohlmann J, Pichersky E. *The Plant journal: for cell and molecular biology*. 2011; 66:212–229. [PubMed: 21443633]
31. Ainsworth AM, Rayner ADM, Broxholme SJ, Beeching JR, Pryke JA, Scard PR, Berriman J, Powell KA, Floyd AJ, Branch SK. *Mycol Res*. 1990; 94:799–809.
32. Isaka M, Srisanoh U, Choowong W, Boonpratuang T. *Organic letters*. 2011; 13:4886–4889. [PubMed: 21848293]
33. Isaka M, Srisanoh U, Sappan M, Supothina S, Boonpratuang T. *Phytochemistry*. 2012; 79:116–120. [PubMed: 22595359]
34. Kim YH, Yun BS, Ryoo IJ, Kim JP, Koshino H, Yoo ID. *The Journal of antibiotics*. 2006; 59:432–434. [PubMed: 17025020]
35. Li G, Liu F, Shen L, Zhu H, Zhang K. *Journal of natural products*. 2011; 74:296–299. [PubMed: 21210711]
36. Liermann JC, Schuffler A, Wollinsky B, Birnbacher J, Kolshorn H, Anke T, Opatz T. *The Journal of organic chemistry*. 2010; 75:2955–2961. [PubMed: 20345148]
37. Yoo NH, Kim JP, Yun BS, Ryoo IJ, Lee IK, Yoon ES, Koshino H, Yoo ID. *The Journal of antibiotics*. 2006; 59:110–113. [PubMed: 16629412]
38. Yun BS, Lee IK, Cho Y, Cho SM, Yoo ID. *Journal of natural products*. 2002; 65:786–788. [PubMed: 12027770]
39. Faraldos JA, Wu S, Chappell J, Coates RM. *Tetrahedron*. 2007; 63:7733–7742. [PubMed: 20617157]
40. Faraldos JA, Miller DJ, Gonzalez V, Yoosuf-Aly Z, Cascon O, Li A, Allemann RK. *J Am Chem Soc*. 2012; 134:5900–5908. [PubMed: 22397618]
41. Gennadios HA, Gonzalez V, Di Costanzo L, Li A, Yu F, Miller DJ, Allemann RK, Christianson DW. *Biochemistry*. 2009; 48:6175–6183. [PubMed: 19489610]
42. Faraldos JA, O'Maille PE, Deltas N, Noel JP, Coates RM. *J Am Chem Soc*. 2010; 132:4281–4289. [PubMed: 20201526]
43. Hertweck C. *Nature chemical biology*. 2009; 5:450–452.
44. Martin F, Selosse MA. *New Phytol*. 2008; 180:296–310. [PubMed: 19138220]
45. Vick JE, Johnson ET, Choudhary S, Bloch SE, Lopez-Gallego F, Srivastava P, Tikh IB, Wawrzyn GT, Schmidt-Dannert C. *Appl Microbiol Biotechnol*. 2011; 92:1275–1286. [PubMed: 22033566]
46. Rabe P, Dickschat JS. *Angewandte Chemie*. 2013; 52:1810–1812. [PubMed: 23307484]
47. Wawrzyn GT, Bloch SE, Schmidt-Dannert C. *Methods Enzymol*. 2012; 515:83–105. [PubMed: 22999171]
48. Kollner TG, O'Maille PE, Gatto N, Boland W, Gershenzon J, Degenhardt J. *Archives of biochemistry and biophysics*. 2006; 448:83–92. [PubMed: 16297849]
49. Gutta P, Tantillo DJ. *J Am Chem Soc*. 2006; 128:6172–6179. [PubMed: 16669687]
50. Zu L, Xu M, Lodewyk MW, Cane DE, Peters RJ, Tantillo DJ. *J Am Chem Soc*. 2012; 134:11369–11371. [PubMed: 22738258]
51. Keller NP, Turner G, Bennett JW. *Nature reviews Microbiology*. 2005; 3:937–947.
52. Westfelt L. *Acta Chemica Scandinavica*. 1970; 24:1618–1622.
53. Desjardins AE. *Journal of agricultural and food chemistry*. 2009; 57:4478–4484. [PubMed: 19385595]

54. Ayer WA, Nakashima TT, Saeedighomi MH. *Can J Chem.* 1984; 62:531–533.
55. Ayer WA, Saeedighomi MH. *Can J Chem.* 1981; 59:2536–2538.
56. Ayer WA, Saeedighomi MH. *Tetrahedron Lett.* 1981; 22:2071–2074.
57. Opatz T, Kolshorn H, Anke H. *The Journal of antibiotics.* 2008; 61:563–567. [PubMed: 19160524]
58. Tokai T, Koshino H, Takahashi-Ando N, Sato M, Fujimura M, Kimura M. *Biochemical and biophysical research communications.* 2007; 353:412–417. [PubMed: 17188234]
59. Tanasova M, Sturla SJ. *Chem Rev.* 2012; 112:3578–3610. [PubMed: 22482429]
60. Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martinez AT, Otilar R, Spatafora JW, Yadav JS, Aerts A, Benoit I, Boyd A, Carlson A, Copeland A, Coutinho PM, de Vries RP, Ferreira P, Findley K, Foster B, Gaskell J, Glotzer D, Gorecki P, Heitman J, Hesse C, Hori C, Igarashi K, Jurgens JA, Kallen N, Kersten P, Kohler A, Kues U, Kumar TK, Kuo A, LaButti K, Larrondo LF, Lindquist E, Ling A, Lombard V, Lucas S, Lundell T, Martin R, McLaughlin DJ, Morgenstern I, Morin E, Murat C, Nagy LG, Nolan M, Ohm RA, Patyshakuliyeva A, Rokas A, Ruiz-Duenas FJ, Sabat G, Salamov A, Samejima M, Schmutz J, Slot JC, St John F, Stenlid J, Sun H, Sun S, Syed K, Tsang A, Wiebenga A, Young D, Pisabarro A, Eastwood DC, Martin F, Cullen D, Grigoriev IV, Hibbett DS. *Science.* 2012; 336:1715–1719. [PubMed: 22745431]
61. Thompson JD, Gibson TJ, Higgins DG. *Current protocols in bioinformatics/editorial board, Andreas D. Baxevanis ...[et al.].* 2002; Chapter 2(Unit 2 3)
62. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. *Molecular biology and evolution.* 2011; 28:2731–2739. [PubMed: 21546353]
63. Saitou N, Nei M. *Molecular biology and evolution.* 1987; 4:406–425. [PubMed: 3447015]
64. Stanke M, Steinkamp R, Waack S, Morgenstern B. *Nucleic acids research.* 2004; 32:W309–312. [PubMed: 15215400]
65. Joulain, D.; Konig, WA. Hamburg. Germany: E. B. Verlag; 1998.
66. Lopez-Gallego F, Wawrzyn GT, Schmidt-Dannert C. *Applied and environmental microbiology.* 2010; 76:7723–7733. [PubMed: 20889795]
67. Abell C, Leech AP. *Tetrahedron Lett.* 1988; 29:4337–4340.
68. Garms S, Kollner TG, Boland W. *The Journal of organic chemistry.* 2010; 75:5590–5600. [PubMed: 20704432]
69. Jansen BJM, de Groot Ae. *Nat Prod Rep.* 2004; 21:449–477. [PubMed: 15282630]

Figure 1. Volatile sesquiterpene production by *S. hirsutum*

The headspace of a liquid culture of *S. hirsutum* was sampled and analyzed for the production of volatile sesquiterpenes by GC/MS over a period of 21 days. Identified sesquiterpene products are numbered and their relative stereochemical structures are shown. One peak (labeled with a question mark ?) with a m/z 204 could not be identified. Peaks labeled with an asterisk * are likely modified terpenoids, each with a m/z 218–222.

Figure 2. Phylogenetic analysis of sesquiterpene synthase homologs

Sesquiterpene synthase homologs from *Coprinus cinereus* (Cop), *Omphalotus olearius* (Omp) and *Stereum hirsutum* (Steh1) are labeled. A) Sesquiterpene synthases form distinct clades in an unrooted neighbor-joining phylogram. Clade III, described previously,^[24] is highlighted with a box. Gene accession numbers highlighted with a circle indicate cloned and characterized sesquiterpene synthases. B) An unrooted neighbor-joining tree of members of Clade III. Branches are labeled with their bootstrap values. Gene accession numbers highlighted in bold indicate cloned and characterized sesquiterpene synthases.

Figure 3. Volatile sesquiterpene production by *E. coli* cultures expressing Stehi1|159379 and Stehi1|128017

The headspace of cultures was sampled and analyzed for production of sesquiterpenes by GC/MS. *E. coli* cells constitutively expressing A) Stehi1|159379 produced primarily the following identified products: β -barbatene **B**, α -barbatene **A**, and small quantities of α -cuprenene **C**. All of these sesquiterpenes are products of a 1,6-cyclization of (3*R*)-NPP. (Note that peaks identified in the headspace of *E. coli* cells expressing Stehi1|159379 were not detected in the culture headspace of *S. hirsutum* and peak labeling follows instead that of Figure S2 showing Cop6 and Omp9 products). *E. coli* cells constitutively expressing B) Stehi1|128017 produced as major products that were identified in the *S. hirsutum* culture headspace (Figures 2 and S1): δ -cadinene **13**, a putative 1,10-cyclization of (3*R*)-NPP product, as well as sesquisabinene **11**, which is believed to result from a 1,6-cyclization of (3*R/S*)-NPP. A number of other sesquiterpenoid products are also produced and are identified and labeled in Figure S3. Indole (labeled with an asterisk *), a breakdown product of tryptophan, is naturally produced by *E. coli* and serves as an internal standard.

Figure 4. Volatile sesquiterpene production by *E. coli* cultures expressing Stehi1|25180, Stehi1|64702 and Stehi1|73029

The headspace of *E. coli* cultures expressing A) Stehi1|25180, B) Stehi1|64702 and C) Stehi1|73029 was sampled and analyzed for production of sesquiterpenes by GC/MS. The major product of cells expressing Stehi1|25180, Stehi1|64702 and Stehi1|73029 is Δ-6 protoilludene **7**, a 1,11-cyclization of (2*E*,6*E*)-FPP. The minor peak observed for Stehi1|64702 and Stehi1|73029 was identified as β-elemene **8**, which results from a 1,10-cyclization of (2*E*,6*E*)-FPP. Indole (labeled with an asterisk *), a breakdown product of tryptophan, is naturally produced by *E. coli* and serves as an internal standard.

Figure 5. *In vitro* activities of purified Stehi1|25180, Stehi1|64702 and Stehi1|73029

Purified enzymes were incubated with (2*E*,6*E*)-FPP and the headspace of vials was sampled and analyzed for production of sesquiterpenes by GC/MS. A single peak, Δ -6 protoilludene **7** was observed for A) Stehi1|25180, B) Stehi1|64702, and C) Stehi1|73029.

Figure 6. Biosynthetic gene clusters of 1,11-cyclizing sesquiterpene synthases
 Stehi1|25180, Stehi1|64702 and Stehi1|73029 are located in putative biosynthetic gene clusters, which are manually annotated. Predicted open reading frames are colored according to function. Sesquiterpene synthases are represented in red, P450 monooxygenases are blue, hydrocarbon scaffold modifying enzymes are green, and transporters are orange. The peptide reference numbers are included for each open reading frame (Data S1). Also included in parentheses is the top conserved domain hit at NCBI. Stehi1|25180 is located in a large gene cluster with three putative P450 monooxygenases, one reductase, two HMG-CoA synthases, and two Cu-oxidases. Note that Stehi1|25180 –i and –j appear to have resulted from a gene duplication event (Data S1), and are identical. Stehi1|64702 is also located in a large gene cluster with three putative P450 monooxygenases, five reductases, an aldehyde dehydrogenase and a multidrug efflux transporter. Stehi1|73029 is located in a gene cluster with three putative P450 monooxygenases, four reductases, a FAD binding protein, two transferases, a hydrolase and a multiple drug transporter.

Figure 7. Phylogenetic analysis of putative Δ -6 protoilludene-associated P450 monooxygenase homologs

An unrooted neighbor-joining tree of P450 monooxygenase homologs from the *Omphalotus olearius* (Omp) and *Stereum hirsutum* (Stehi1) Δ -6 protoilludene synthase associated biosynthetic gene clusters. Branches are labeled with their bootstrap values. P450 labeling is consistent with that of Figure 6 and previously described biosynthetic gene clusters.^[24] (Note that Stehi|25180 -i and -j appear to result from a gene duplication event (Figure 6, Data S1)).

Scheme 1. Representative examples of modified sesquiterpenoids previously isolated from *Stereum* sp

Most sesquiterpenoids isolated from *Stereum* sp. are derived from the *trans*-humulyl cation, a product of the (2*E*,6*E*)-FPP 1,11-cyclization mechanism. Atom numbering is based upon designations in the FPP precursor. Sterpuric acid, sterpolide A, sterpurene-3,12,14-triol, and the sterelactones result from the modification of the sterpurene scaffold.^[54–57,67]

Modification of the hirsutene scaffold yields hirsutenols A-F,^[37,38] as well as complicatic acid and chlorostereone.^[36] The sterostreins^[32,33] are modified products of the Δ-6 protoilludene scaffold. Additionally, the stereumins are derived from the (3*R*)-NPP 1,10-cyclization product δ-cadinene,^[31,35] (+)-torreyol is a product of attack by water on either the (2*E*,6*E*)-FPP 1,10-cyclization intermediate muurolyl cation or the (3*R*)-NPP 1,10-cyclization intermediate cadinyl-cation.^[68] Finally, methoxylaricinolic acid results from a putative 2,7-cyclization product drimane^[34,69] (shown in parentheses). Relative stereochemistries of sesquiterpenoids are shown.

Scheme 2. Cyclization pathways for the production of Δ -6 protoilludene, germacrene A and β -elemene

A metal ion mediated dephosphorylation of (2*E*,6*E*)-FPP yields a reactive carbocation, followed by a specific ring closure. A 1,11-cyclization mechanism produces a *trans*-humulyl cation, a precursor to a wide range of sesquiterpenes. The *trans*-humulyl cation can undergo a 1,2 hydride shift followed by two cyclizations and loss of a proton, resulting in Δ -6 protoilludene **7**. A 1,10-cyclization mechanism yields a *E,E*-germacradienyl cation, and loss of a proton leads to germacrene A. β -elemene **8** is the heat induced Cope rearrangement product of germacrene A. Relative stereochemistries of sesquiterpenes are shown.

Table 1
Kinetic parameters determined for Stehi1|25180, Stehi1|64702 and Stehi1|73029

Data were collected using (2*E*,6*E*)-FPP as a substrate in a coupled spectrophotometric assay.

	<i>K_m</i> [M]	<i>k_{cat}</i> [s ⁻¹]	<i>k_{cat}/K_m</i> [M ⁻¹ s ⁻¹]
Stehi1 25180	$(5.02 \pm 0.9) \times 10^{-6}$	$(4.4 \pm 0.8) \times 10^{-3}$	$(8.89 \pm 0.7) \times 10^2$
Stehi1 64702	$(1.91 \pm 0.3) \times 10^{-6}$	$(3.7 \pm 0.6) \times 10^{-3}$	$(19.51 \pm 1.5) \times 10^2$
Stehi1 73029	$(1.52 \pm 0.2) \times 10^{-6}$	$(6.3 \pm 0.7) \times 10^{-3}$	$(41.84 \pm 5.1) \times 10^2$