



Title	Ascomycete <i>Aspergillus oryzae</i> Is an Efficient Expression Host for Production of Basidiomycete Terpenes by Using Genomic DNA Sequences
Author(s)	Nagamine, Shota; Liu, Chengwei; Nishishita, Jumpei; Kozaki, Takuto; Sogahata, Kaho; Sato, Yoshiro; Minami, Atsushi; Ozaki, Taro; Schmidt-Dannert, Claudia; Maruyama, Jun-ichi; Oikawa, Hideaki
Citation	Applied and Environmental Microbiology, 85(15), e00409-19 <a href="https://doi.org/10.1128/AEM.00409-19">https://doi.org/10.1128/AEM.00409-19</a>
Issue Date	2019-08
Doc URL	<a href="http://hdl.handle.net/2115/76550">http://hdl.handle.net/2115/76550</a>
Type	article (author version)
File Information	Appl. Environ. Microbiol..pdf



[Instructions for use](#)

## **Ascomycota *Aspergillus oryzae* is an efficient expression host for production of Basidiomycota terpenes using genomic DNA sequences**

**Shota Nagamine<sup>1</sup>, Chengwei Liu<sup>1</sup>, Jumpei Nishishita<sup>1</sup>, Takuto Kozaki<sup>1</sup>, Kaho Sogahata<sup>1</sup>, Yoshiro Sato<sup>1</sup>, Atsushi Minami<sup>1,\*</sup>, Taro Ozaki<sup>1</sup>, Claudia Schmidt-Dannert<sup>2</sup>, Jun-ichi Maruyama,<sup>3</sup> and Hideaki Oikawa<sup>1,4,\*</sup>**

1. Department of Chemistry, Faculty of Science, Hokkaido University, Kita-ku Kita 10 Jo Nishi 8 Chome, Sapporo 060-0810, Japan.
2. Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, St. Paul, Minnesota, USA.
3. Department of Biotechnology, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan.
4. Lead Contact

### **Running title (not to exceed 54 characters) 53 characters**

Expression of Basidiomycota gene using genomic DNA sequences

### **Keywords**

Basidiomycota, sesquiterpene synthase, functional analysis, *Aspergillus oryzae*

### **Corresponding authors**

Hideaki Oikawa and Atsushi Minami

### **E-mail addresses**

hoik@sci.hokudai.ac.jp, aminami@sci.hokudai.ac.jp

1 **Abstract**

2 Basidiomycota fungi are an attractive resource for biologically active natural products for use in  
3 pharmaceutically relevant compounds. Recently, genome projects of mushroom fungi have  
4 provided a great deal of biosynthetic gene cluster information. However, functional analyses of the  
5 gene clusters for natural products were largely unexplored because of the difficulty of cDNA  
6 preparation and lack of gene manipulation tools for Basidiomycota fungi. To develop a versatile  
7 host for Basidiomycota genes, we examined gene expression using genomic DNA sequences in the  
8 robust Ascomycota host *Aspergillus oryzae*, which is frequently used for the production of  
9 metabolites from filamentous fungi. Exhaustive expression of 30 terpene synthase genes from  
10 Basidiomycota *Clitopilus pseudo-pinsitus* and *Stereum hirsutum* showed two splicing patterns,  
11 completely spliced cDNAs giving terpenes (15 cases) and mostly spliced cDNAs, indicating that *A.*  
12 *oryzae* correctly spliced most of introns at the predicted positions and lengths. The mostly spliced  
13 cDNAs were expressed after PCR-based removal of introns, resulting in the successful production  
14 of terpenes (14 cases). During this study, we observed relatively frequent mispredictions in the  
15 automated program. Hence, the complimentary use of AO-expression and automated prediction will  
16 be a powerful tool for genome mining.

17

18 **Importance**

19 The recent large influx of genome sequences from Basidiomycota, which are prolific producers  
20 of bioactive natural products may provide opportunities to develop novel drug candidates. The  
21 development of a reliable expression system is essential for the genome mining of natural products  
22 because of the lack of a tractable host for heterologous expression of Basidiomycota genes. For this  
23 purpose, we applied the Ascomycota *Aspergillus oryzae* system for the direct expression of fungal  
24 natural product biosynthetic genes from genomic DNA. Using this system, 29 sesquiterpene  
25 synthase genes and diterpene biosynthetic genes for bioactive pleuromutilin were successfully  
26 expressed. Together with the use of computational tools for intron-prediction, this *Aspergillus*  
27 *oryzae* system represents a practical method for the production of Basidiomycota natural products.

## 1 **Introduction**

2 Mushroom-forming Basidiomycota fungi are known to be prolific producers of structurally  
3 diverse, bioactive natural products (NPs) and have been used since ancient times in traditional  
4 medicine (1). This group of fungi is particularly known for their terpenoid NPs, which appear to be  
5 the major class of NPs produced by them (1, 2). Well-known examples of pharmaceutically  
6 relevant terpenoid-derived compounds include the widely used livestock antibiotic pleuromutilin  
7 (3), the anticancer compound illudin S (4), the melleolide antibiotics (5), and the nerve growth  
8 factor-synthesis promoting erinacines (6) (Figure 1). The 1000 fungal genomes project performed  
9 by the Joint Genome Institute (JGI) (7) has resulted in a large influx of Basidiomycota genome  
10 sequences and genomic data from hundreds of Basidiomycota taxa that can now be searched *in*  
11 *silico* for NP biosynthetic gene clusters.

12  
13 However, despite their incredible potential for NP discovery, Basidiomycota fungi are a largely  
14 unexplored territory for drug discovery compared to Ascomycota fungi. While an abundance of  
15 genetic tools and techniques for transformation and genetic manipulation are available for yeast and  
16 filamentous fungi, such methods are mostly lacking for Basidiomycota fungi (2, 8). In addition,  
17 many Basidiomycota strains are difficult to grow under laboratory conditions, requiring the  
18 determination of suitable growth conditions and often long fermentation times. Characterization of  
19 their biosynthetic genes therefore typically requires heterologous expression in a more genetically  
20 tractable host. Unfortunately, Basidiomycota have very intron-rich genomes and genes that contain  
21 very small and unpredictable exons, especially in cytochrome P450 monooxygenase genes, that  
22 currently necessitate the amplification of genes from cDNA. Even if a gene is expressed, alternative  
23 splice variants may be produced that are not functional (2). All pioneering works to access the large  
24 diversity of terpenoid NPs made by Basidiomycota have therefore relied on the amplification of  
25 functional terpene synthase (TS) genes from cDNA for expression and characterization in *E. coli* or  
26 yeasts, limiting the complete functional analysis of all computationally identified TSs in  
27 Basidiomycota genomes (9-13). Being able to directly express Basidiomycota biosynthetic genes  
28 from genomic DNA in a suitable fungal surrogate host would therefore greatly accelerate the  
29 functional characterization of NP pathways.

30  
31 We reasoned that a genetically tractable Ascomycota could potentially be used as a heterologous  
32 host for the functional expression of Basidiomycota NP biosynthetic genes. Unlike yeasts, which  
33 have relatively few genes with introns, a splicing machinery of filamentous fungi such as  
34 *Aspergillus* is more similar to that of Basidiomycota (14). Heterologous expression of biosynthetic  
35 genes from filamentous fungi directly from genomic DNA has become a powerful approach to  
36 elucidate biosynthetic pathways. We have used this strategy with the Ascomycota *Aspergillus*  
37 *oryzae* (AO) for the heterologous production of several NPs and their pathway intermediates  
38 (15-18) and to access difficult to obtain pathway intermediates for the study of intriguing enzyme  
39 reactions (19-21). Recently, we successfully produced in AO the Basidiomycota-derived diterpene  
40 pleuromutilin from cDNA amplified biosynthetic genes. Like others, we found the cloning of  
41 functional genes from Basidiomycota cDNA to be difficult and tedious. We and others also showed

1 that the native AO NADPH cytochrome P450 reductase supports the activity of the Basidiomycota  
2 P450 monooxygenases pathway (22, 23). Given our success with using AO for the direct expression  
3 of fungal NP biosynthetic genes from genomic DNA and the ability of AO to functionally express a  
4 Basidiomycota terpenoid NP biosynthetic enzymes, including its P450 monooxygenase, we  
5 therefore hypothesized that this Ascomycota would make an excellent platform strain for the  
6 characterization and production of Basidiomycota NP pathways, in particular its terpenoid  
7 pathways.

8  
9 Here, we describe a new strategy using AO as an expression host for the heterologous  
10 production of Basidiomycota NP biosynthetic genes. We selected two different terpenoid classes  
11 (pleuromutilin diterpene and sesquiterpenoids) from two different Basidiomycota and performed a  
12 systematic analysis of the predicted and experimentally obtained splicing patterns in AO. Our  
13 results indicated that a significant fraction of genes was correctly spliced by AO, while many  
14 partially spliced genes could be corrected based on gene model predictions and functionally  
15 expressed in AO.

## 16 17 **Results and Discussion**

### 18 ***Expression of pleuromutilin biosynthetic genes from genomic DNA in *Aspergillus oryzae****

19 We previously achieved the heterologous production of the Basidiomycota diterpene  
20 pleuromutilin using an *A. oryzae* expression system (22). Therefore, we initially examined the  
21 heterologous expression of pleuromutilin biosynthetic genes using genomic DNA sequences in this  
22 Ascomycota host. When we tested the gene expression of GGPP synthase *ple4* and terpene synthase  
23 *ple3* in *A. oryzae*, the resultant transformant (TF) successfully produced premutilin (**1**) (Figure 2).  
24 Sequencing of cDNA sequences recovered from the TF AO-*ple3/4* showed that all introns (*ple3/4*:  
25 3/4) were correctly spliced (Figure 2). We next examined the expression of intron rich P450 genes  
26 (*ple1/5/6*: 10/13/11). cDNA sequence analysis showed that all 11 introns of *ple6* were successfully  
27 spliced while *ple1/5* each contained one unspliced introns (*ple1*: intron-6 out of 10; *ple5*: intron-4  
28 out of 13), which can be readily removed by PCR-based methods. Based on these promising results,  
29 we speculated that gene splicing in Basidiomycota resembles that of Ascomycota fungi (14, 24).

### 30 31 ***Functional gene expression analysis of STS genes***

32 Given that almost all introns of *ple* genes were successfully spliced in AO, we next performed a  
33 comprehensive analysis of Basidiomycota gene expression from genomic DNA in this host. To  
34 establish the versatility of Basidiomycota heterologous expression using AO, we chose to test the  
35 expression of sesquiterpene synthases (STSs) because they are prolific in Basidiomycota genomes  
36 and are small enzymes that convert abundant farnesyl pyrophosphate (FPP) in host cells into  
37 various cyclic, volatile products that can be readily identified by GC-MS analysis of cultures (25).  
38 In this work, we selected 31 STS genes for characterization from two Basidiomycota fungi, our  
39 pleuromutilin producer *C. pseudo-pinsitus* and for comparison, *Stereum hirsutum* which was  
40 previously investigated for sesquiterpenoid production (11) (Figure 3, Table S1).

41

## 1 ***1. STSs from *Clitopilus pseudo-pinsitus****

2 Our Local-Blast search using aristolochene synthase as a query sequence identified 18 STS  
3 candidate genes (CpSTS1-18) in *C. pseudo-pinsitus* (Figure 3, Table S1). Open reading frames  
4 (ORFs) were manually predicted by comparison with functionally characterized fungal STS genes.  
5 Except for CpSTS10, which lacked the conserved motifs (DDXXD; NSE) of STSs, 17 ORFs were  
6 selected for expression in AO. To ensure the expression of target genes, we used the recently  
7 developed fungal CRISPR-Cas9 system specifically optimized for AO (26). In the presence of the  
8 CRISPR/Cas9 plasmid with the protospacer sequence for the SC103 locus, which was found in a  
9 strain that produced betaenone in good yield (27), *ligD*-deficient *A. oryzae* NSPID1 (28) was  
10 transformed using a donor DNA plasmid carrying a targeted STS gene. Usually, the knock-in rate  
11 ranged from 80 to 90%, much higher than in normal random integrations. GC-MS analysis of  
12 headspace volatiles from these TFs revealed that nine AO TFs expressed functional STS that  
13 cyclized FPP into the cyclic products (Figures 3, S1, and S2). Sesquiterpenes were identified by  
14 retention indices (RI) and comparison of MS fragmentation patterns with those reported in  
15 databases or the literature (Figure S1 and Table S2).

16  
17 To find out why the remaining eight TFs did not produce sesquiterpenoids, cDNA of STS were  
18 recovered and sequenced. The results showed that although AO correctly spliced the vast majority  
19 of introns that were computationally predicted by Augustus (29), some introns were (partially)  
20 skipped, thereby generating multiple splicing patterns leading to incomplete protein (ex. one intron;  
21 CpSTS5/7/13/14/16, two introns; CpSTS8/17) (Figures 3 and S3, and Table S3). Following the  
22 removal of nonspliced introns in amplified cDNA by PCR-based techniques,  
23 CpSTS5/7/8/13/14/16/17 were expressed by *E. coli* TFs and yielded terpenes (Figures 3, S1, and S2,  
24 and Table S2). Only CpSTS15 did not give any product although all introns were likely spliced  
25 correctly.

26 Sesquiterpenes produced by CpSTSs include 1) aromadendrene type sesquiterpenes featuring  
27 the 5-7-3 tricyclic skeleton, such as ledene (**2**) and virifloridol (**3**) (C1,C10-cyclization products)  
28 (Figure 3b), 2) structurally unique sesquiterpenes such as 6-protoilludene (**7**), sterpurene (**8**),  
29 pentalenene (**9**), and  $\beta$ -caryophyllene (**10**) (C1,C11-cyclization products) (Figure 3c), 3) cadinene  
30 type sesquiterpenes possessing the 6-6 bicyclic structure, such as  $\delta$ -cadinene (**12**) and  $\gamma$ -muurolene,  
31 and biosynthetically related ones such as  $\alpha$ -cubebene (**13**) (C1,C10-cyclization products) (Figures  
32 3e and S2), and 4) linear and monocyclic sesquiterpenes such as  $\alpha$ -farnesene (**11**) and  $\beta$ -elemene  
33 (Figures 3c and S2). Among them, **7** and cadinenes are frequently isolated from fungi and the  
34 corresponding STSs have been functionally characterized (2). By contrast, **8** produced by CpSTS1  
35 is known as a fungi specific sesquiterpene and a sterpurene synthase has not been identified.  
36 Successful functional characterization of CpSTS1 enabled us to search putative sterpurene  
37 synthases in public database and identify homologs in *Psilocybe cyanescens* (72%, PPQ78014.1),  
38 *Gymnopilus dilepis* (72%, PPQ64797.1), and *Fibularhizoctonia sp.* CBS 109695 (66%,  
39 KZP33092.1) although **8** and the related metabolites have not been isolated from them.

1 Based on a previous phylogenetic analysis (10), 69 Basidiomycota STSs are divided into clades  
2 I-IV (Figure 4). The majority of characterized CpSTS (11/17) in this work groups with fungal STS  
3 in clades II and III that are proposed to catalyze the C1,C10- and C1,C11-cyclization of FPP,  
4 respectively (Figures 3, 4, and S2). The remaining STS group with clade I (2 STSs), clade IV (2  
5 STSs), and unclassified clade (1 STS). We used this phylogenetic framework as a guide to  
6 investigate the relationship of homologous STSs in the context of sequence homology, gene  
7 structures, and chemical structures/reaction mechanisms. The species-specific STSs  
8 CpSTS8/9/11/12/13/16 (sequence identity: >40%) belonging to clade II have similar gene structures  
9 (Figure 3a, Table S3) (30), possibly arising by gene duplication of ancestral STS, a well known  
10 evolutionary mechanism in this class of fungi (31). These STSs produced the aromadendrene family  
11 members and aristolene (**6**), which are frequently found as plant volatiles (Figures 3b and S4), but  
12 only one corresponding STS from *Citrus* (CsSesquiTPS5) has been characterized (32). All of them  
13 are most likely biosynthesized via bicyclogermacrene (**IM1**) (30, 32-34). These results suggested a  
14 strong relationship of the gene structure and the cyclization mechanism. It should be emphasized  
15 that comprehensive analysis of 17 STSs provides direct evidence regarding the gene structure and  
16 the enzyme function.

17

## 18 **2. STSs from *S. hirsutum***

19 Previously, 18 STS genes were bioinformatically predicted in the mushroom *S. hirsutum* and  
20 five STS genes were successfully amplified from cDNA for functional expression in *E. coli* (11, 35).  
21 The remaining STS were not studied, partially because of cDNA availability limitations. We sought  
22 this suite of STS from *S. hirsutum* would be another good test system for our direct gene expression  
23 strategy. Of the 18 STS genes (ShSTS1-18), we excluded five genes from our study to avoid the  
24 expression of ShSTS2/9/14/15/6 that showed >60% homology to ShSTS1/8/13/16 and CpSTS3  
25 (Table S4). GC-MS analysis of culture headspace showed that seven AO TFs expressing  
26 (ShSTS-1/4/7/8/11/16/17) produced sesquiterpenes such as **7**, **12**, **15**,  $\beta$ -barbatene, and hirsutene,  
27 (Figures 3, S1, S2 and Table S2). As in the case of CpSTSs, sequencing of cDNA showed intron  
28 skipping at one (ShSTS5/12/13/18) or two positions (ShSTS3/10) (Figure 3d). Removal of  
29 predicted introns in ShSTS3/5/10/12/13/18 gave functional STSs in *E. coli* transformants that  
30 synthesized terpenes such as **10**, **11**, **13**, **14**, **16**, and  $\gamma$ -cadinene (Figures 3, S1, S2 and Table S2).  
31 The metabolite profiles revealed that clade III ShSTSs as well as those of CpSTSs produce  
32 structurally unique sesquiterpene such as **7-10** as a single product while other STSs appear to give  
33 multiple biosynthetically related products.

34

35 A strong correlation between gene structure and function was also found in *Stereum* TSs; clade  
36 II ShSTS8/9/10/11/12 revealed a relatively high sequence identity (>48%) and the gene structures  
37 of ShSTS10/11/12 are nearly identical, whereas that of the ShSTS8 gene is likely modified by the  
38 insertion of an extra intron (Figures 3d and 4). These STSs generated cadinene-type sesquiterpenes  
39 biosynthesized via *Z,E*-germacrenedieryl cation (**IM2**) (Figure 3e). In addition, a conservation of  
40 sequence identity (>54%) and gene structure is observed for clade III ShSTSs, ShSTS15/16/17/18  
41 (Figures 3c and 4), which produce **7** as a characteristic and major sesquiterpene made by

1 mushrooms. The gene structure of ShSTS13 resembles with those of Clade III ShSTSs except that  
2 an extra *N*-terminal intron is inserted (Figure 3d). The structural difference between **7**  
3 (ShSTS15/16/17/18) and **10** (ShSTS13) might reflect to alter the transition states derived from  
4 common intermediate **IM3**. When we searched for putative biosynthetic gene clusters surrounding  
5 clade II and III Cp- and Sh-STS paralogs/orthologs, we found quite divergent arrangements of  
6 putative biosynthetic genes for modifying terpene scaffold (Figure S5).

### 7 8 ***Splicing of Basidiomycota genes in AO.***

9 To gain insights into the splicing patterns of Basidiomycota genes in AO, we analyzed the  
10 sequences of spliced and nonspliced introns of genes investigated in this study (Table S3). Overall,  
11 CpSTS have a slightly smaller mean intron length (56 nt, size range: 47-69 nt) and average number  
12 of introns (4 introns) compared to ShSTS with an intron length of 64 nt (size range: 48-122 nt) and  
13 5 introns on average. The 5'- and 3'-splice site consensus sequences (GUxxGU and YAG,  
14 respectively) matched those from *C. neoformans* (Basidiomycota fungi) and *Fusarium* sp.  
15 (Ascomycete) (14, 24). Analysis of branch sites revealed that the majority introns (82%, 108/132)  
16 has a CTNAN motif, while 13% (17/132) contain a TTNAN motif and for the remaining (7/132) no  
17 motif could be determined. Branch site motif appears to be the most important predictor for splicing  
18 in AO, as 96% of introns with a CTNAN are spliced in AO, while only 65% of introns with a  
19 TTNAN or 43% of discernable motif are spliced. Only one spliced intron (CpSTS8-intron1) out of  
20 all introns spliced in AO (132 number) was mis-spliced in AO (Figure S3), demonstrating that  
21 introns in Basidiomycota genes that are recognized for splicing are faithfully and reliably spliced by  
22 AO. Intron skipping tends to occur at one or two positions in STS genes. A similar pattern was  
23 observed for pleuromutilin genes investigated in this work (Figure 2).

### 24 ***Estimation of reliability for computational prediction of introns and ORFs***

25 Comparison of the cDNA sequences of the large number of 30 functional STS sequences  
26 characterized in this work with computationally predicted gene structures provides an excellent  
27 assessment of the reliability of bioinformatics tools such as Augustus (29) for Basidiomycota gene  
28 prediction (Tables S3 and S5). In Augustus, the accuracy of gene structure predictions depended on  
29 the gene model used; predictions with the mushroom *Coprinus cinereus* model instead of AO  
30 (expression host) were much more accurate. Even with the mushroom gene model, Augustus  
31 mispredicted 11 out 30 STS gene structures (Figures 3 and S3, Table S5). Misprediction included i)  
32 insertion of an extra intron in *CpSTS16*; ii) skipping of an essential intron in *ShSTS5*; iii) prediction  
33 of shorter/longer introns in *CpSTS18* and *ShSTS12/16*; iv) misprediction of introns closely located  
34 in the 3'-end of *CpSTS4/13*, *ShSTS17* (extra intron), and *CpSTS9* (intron skipping) and the 5'-end of  
35 CpSTS8 and ShSTS11 (Figure 5). Most errors concerning ORFs and introns located close to the  
36 3'-terminal end of the genes (example iv) can be avoided by manual prediction by considering the  
37 amino acid sequences of the known STSs and the gene structures of paralogous STS genes (Figure  
38 3). Therefore, automated prediction may be useful when these measures are taken. Actually, we  
39 used Augustus prediction when we needed to remove nonspliced intron(s) in the AO expression. We  
40 successfully used this strategy to identify and subsequently remove nonspliced introns by AO in



1 recovered cDNA for expression of active STS by EC-TFs (F1). Consequently, direct expression of  
2 Basidiomycota genes from genomic DNA in AO in conjunction with bioinformatics proved to be an  
3 effective and reliable strategy to quickly obtain functional NP biosynthetic genes.  
4

## 5 **Concluding remarks**

6 In this work we developed an efficient approach for heterologous expression of Basidiomycota  
7 NP biosynthetic genes directly from genomic DNA in the genetically tractable and well-studied  
8 Ascomycota host *A. oryzae*. We show that direct expression of unspliced genes from genomic  
9 DNA from Basidiomycota in AO results in almost correctly spliced introns and that a combination of  
10 computational and manual gene structure prediction can be used to identify and correct nonspliced  
11 introns in recovered cDNA by a quick PCR step. Out of 30 STS expressed in AO using this method,  
12 half of them were functionally expressed and a simple PCR step yielded functional cDNA for the  
13 remaining STS. Comprehensive analysis of gene structures indicated a strong correlation between  
14 gene structure preservation and evolution of enzyme function. Comparison of intron sequences  
15 revealed intron features such as branch sites that are less likely to be spliced by an Ascomycota such  
16 as AO. This information will aid in the prediction of Basidiomycota STS activities, provide insights  
17 in the diversification of NP cluster in this class of fungi and improve computational gene prediction  
18 methods. Further, by knowing which introns are less likely to be processed, genomic sequences can  
19 be preemptively, spot corrected prior to introduction into a versatile, heterologous Ascomycota  
20 production host – thereby opening the door for the accessing the largely undiscovered NP diversity  
21 of Basidiomycota mushrooms.  
22

## 23 **Materials and Methods**

24 **General.** All reagents commercially supplied were used as received. Column chromatography was  
25 carried out on 60N silica gel (Kanto Chemicals). Optical rotations were recorded on JASCO P-2200  
26 digital polarimeter. <sup>1</sup>H-NMR spectra were recorded on Bruker DRX-500 or Bruker AMX-500  
27 spectrometer (500 MHz for <sup>1</sup>H-NMR). NMR spectra were recorded in CDCl<sub>3</sub> (99.8 atom% enriched,  
28 Kanto). <sup>1</sup>H chemical shifts were reported in δ value based on residual chloroform (7.26 ppm) as a  
29 reference. Data are reported as follows: chemical shift, multiplicity (s = singlet, d = doublet, t =  
30 triplet, q = quartet, m = multiplet, br = broad), coupling constant (Hz), and integration. GC-MS  
31 analyses were conducted with MS-2010 (Shimadzu). Mass spectra were obtained with a Waters  
32 ACQUITY QDa (ESI mode).

33 Oligonucleotides for polymerase chain reactions (PCRs) were purchased from Hokkaido System  
34 Science Co., Ltd. PCRs were performed with a BioRad S1000 thermal cycler. All PCR reactions  
35 were performed with the KOD-Plus-Neo (TOYOBO). The assembly of DNA fragments was  
36 performed by using either In-Fusion Advantage PCR cloning kit (Clontech Laboratories) or HiFi  
37 DNA Assembly Master Mix (New England Biolabs).  
38

39 **Strains and culture conditions.** *Escherichia coli* HST08 was used for cloning and following  
40 standard recombinant DNA techniques. *E. coli* BL21-Gold(DE3) was used for protein expression. A  
41 fungal host strain used in this study was *A. oryzae* NSAR1 (36), a quadruple auxotrophic mutant

1 (*niaD*<sup>-</sup>, *sC*<sup>-</sup>, *ΔargB*, *adeA*<sup>-</sup>), and *A. oryzae* NSPID1 (28), a strain with a highly efficient  
2 gene-targeting background (*niaD*<sup>-</sup>, *sC*<sup>-</sup>, *ΔpyrG*, *ΔligD*) for expression. AO-*bet123* (27), a betaenone  
3 B highly producing transformant, was used for genome sequencing.

4  
5 **Genome Sequencing and Analysis.** Genome sequencing of AO-*bet123* was performed by  
6 Hokkaido System Science Co., Ltd. (Hokkaido, Japan) with an Illumina HiSeq 2000 system. Read  
7 mapping was performed with the following programs, BWA, Samtools, GATK, and Picard. The  
8 results showed a clear gap between AO090103000023 and AO090103000025 on the chromosome 8.  
9 This site is tentatively defined as SC103 region in this study.

10  
11 **Construction of a Cas9 plasmid pC9SC103.** The Cas9 plasmid, pC9SC103, harbouring guide  
12 RNA sequences for SC103 was constructed as follows; The gRNA fragments (G-Fr1: U6P-SC103  
13 and G-Fr2: SC103-U6T) were amplified from ppAsACas9gwA (26) with primer set as shown in  
14 Supplemental Table S6. These primary products were used as a template for the fusion PCR to  
15 afford guide RNA sequences flanked by U6 promoter/terminator set. The PCR product (U6P-guide  
16 RNA sequence for SC103-U6T) was then inserted into *Sma*I-digested ppAsACas9 to construct  
17 pC9SC103.

18  
19 **Construction of donor vectors, pDP1031 and pDP1032.** The donor vectors, pDP1031 and  
20 pDP1032, harbouring homologous arms of SC103 was constructed as follows; Two DNA  
21 fragments, SC103-up and SC103-down, were amplified from genomic DNA of *Aspergillus oryzae*  
22 NSAR1 with primer sets as shown in Supplemental Table S6. Each PCR product was inserted into  
23 *Hind*III/*Eco*RI-digested pUC19 to afford pUC19-SC103. The DNA fragment harbouring  
24 primer-terminator sets for overexpression in *A. oryzae* were amplified from either pTAex3 or  
25 pUARA2 with primer set as shown in Supplemental Table S6. Each PCR product was inserted into  
26 *Bam*HI-digested pUC19-SC103 to afford pDP1031 (single primer/terminator set) and pDP1032  
27 (tandem primer/terminator sets).

28  
29 **Bioinformatics analysis.** STS gene annotation, alignment, and gene structure prediction were  
30 performed as described previously (11). Briefly, gene predictions were made in Augustus (29) with  
31 different fungal gene models using the genomic region 10-15 kb flanking identified STS ORFs.  
32 STS gene predictions were manually aligned with functionally characterized fungal STS. All  
33 sequence alignments and phylogenetic analysis were performed in MEGA7 (37) using MUSCLE  
34 (38) for protein alignments. Phylogenetic analysis of fungal STSs (**Figure 2B**) was done using the  
35 Neighbor-joining method using the Poisson correction methods with 500 bootstrap replications (39,  
36 40).

37 Accession numbers and references protein sequences are as follows:

38 **Basidiomycota STS:** *Coprinus cinereus* (Cop1-4, Cop6) [XP 001832573, XP 001836556, XP  
39 01832925, XP 01836356, XP 01832548], *Omphalotus olearius* (Omp1-10), *Fomitopsis pinicola*  
40 [FomPi84944], *Stereum hirsutum* [Stehi1|159379, 128017, 25180, 64702, 73029], *Armillaria*  
41 *gallica* (ArmGa1) [P0DL13], and *Postia placenta* [PpSTS01 [BBD74517.1], 02 [BBD74518.1], 03

1 [BBD74519.1], 06 [BBD74520.1], 07 [BBD74521.1], 08 [BBD74522.1], 09 [BBD74523.1], 10  
2 [BBD74524.1], 13 [BBD74525.1], 14 [BBD74526.1], 16 [BBD74527.1], 18 [BBD74528.1], 22  
3 [BBD74529.1], 24 [BBD74530.1], 25 [BBD74531.1], 29 [BBD74532.1].

4 **Ascomycota STS:** *Fusarium fujikori* (Ffsc4) [HF563560.1] and Ffsc6 [HF563561.1], *Fusarium*  
5 *gramineareum* (FgCLM1) [GU123140]), *Aspergillus terreus* (atAS) [Q9UR08], *Penicillium*  
6 *roqueforti* (prAS) [W6Q4Q9], and *Botrytis cinerea* (BcBOT2) [AAQ16575.1].

7  
8 **Accession number:** The sequences of sesquiterpene synthase genes found in *C. pseudo-pinsitus*  
9 ATCC20527 have been deposited in the DNA Data Bank of Japan (DDBJ). The accession numbers  
10 are summarized in Table S1.

11  
12 **Construction of *A. oryzae* expression plasmids of *ple* and sesquiterpene synthase genes.**  
13 Pleuromutin biosynthetic genes (*ple13456*) were amplified from the genomic DNA of *C.*  
14 *pseudo-pinsitus* ATCC20527 (22) with primer set as shown in Supplemental Table S6. Each PCR  
15 product was inserted into appropriate restriction site (site 1 and/or site 2) of pDP1031 or pDP1032  
16 to construct the following expression plasmids; pDP1031-*ple5* (*KpnI* site), pDP1032-*ple6* (*NheI*  
17 site), pDP1032-*ple3* (*KpnI*)/4 (*SpeI*), and pDP1032-*ple7* (*KpnI*)-*ple1* (*SpeI*).

18 Sesquiterpene synthase genes were amplified from the genomic DNA of *C. pseudo-pinsitus*  
19 ATCC20527 and *S. hirsutum* FP-91666 S1 (11) with primer set as shown in Supplemental Table S6.  
20 Each PCR product was then inserted into the appropriate restriction sites of expression vectors to  
21 construct the following expression plasmids; pTAex3-*CpSTS1*, pUARA2-*CpSTS2*,  
22 pUARA2-*CpSTS3*, pTAex3-*CpSTS4*, pDP1031-*CpSTS5*, pDP1031-*CpSTS6*, pDP1031-*CpSTS7*,  
23 pDP1031-*CpSTS8*, pDP1031-*CpSTS9*, pDP1031-*CpSTS10*, pDP1031-*CpSTS11*,  
24 pDP1031-*CpSTS12*, pUSA2-*CpSTS13*, pDP1031-*CpSTS14*, pDP1031-*CpSTS15*, pUSA2-*CpSTS16*,  
25 pDP1031-*CpSTS17*, pDP1031-*CpSTS18*, pDP1032-*ShSTS1/18*, pDP1031-*ShSTS3*,  
26 pDP1031-*ShSTS4*, pDP1031-*ShSTS5*, pDP1031-*ShSTS7*, pDP1031-*ShSTS8*, pDP1031-*ShSTS10*,  
27 pDP1032-*ShSTS11/16*, pDP1031-*ShSTS12*, pDP1031-*ShSTS13*, and pDP1031-*ShSTS17*.

28  
29 **Transformation of *Aspergillus oryzae* (genome-editing method).** Most transformants used in this  
30 study was constructed by genome-editing method; a spore suspension of *A. oryzae* NSPID1 (1.0 x  
31 10<sup>8</sup> cells) were inoculated into CD (0.3 % of NaNO<sub>3</sub>, 0.2 % of KCl, 0.1 % of K<sub>2</sub>HPO<sub>4</sub>, 0.05 % of  
32 MgSO<sub>4</sub>·7H<sub>2</sub>O, 2 % of dextrin, 0.002 % of FeSO<sub>4</sub>·7H<sub>2</sub>O, 0.15 % of methionine, 0.488 % of uracil,  
33 0.2 % of uridine, 100 mL, pH 5.5) medium supplemented with appropriate nutrients. After 3 days  
34 incubation at 30 °C (200 rpm), mycelia was collected by filtration and washed with water.  
35 Protoplasting was performed using Yatalase (Takara; 5.0 mg mL<sup>-1</sup>) in Solution 1 (0.8 mM of NaCl,  
36 10 mM of NaH<sub>2</sub>PO<sub>4</sub>, pH 6.0) at 30 °C for 2 h. Protoplasts were centrifuged at 2,000 rpm (Beckman  
37 JLA10.500) for 5 min and washed with 0.8 M of NaCl solution. Then, protoplasts were adjusted to  
38 2.0 x 10<sup>8</sup> cells/mL by adding Solution 2 (0.8 M of NaCl, 10 mM of CaCl<sub>2</sub>, 10 mM of Tris-HCl, pH  
39 8.0) and Solution 3 (40 % (w/v) of PEG4000, 50 mM of CaCl<sub>2</sub>, 50 mM of Tris-HCl, pH 8.0) in 4/1  
40 volume ratio. To the protoplast solution (200  $\mu$ L) was added a cas9 plasmid (2  $\mu$ g) and a donor  
41 plasmid (5  $\mu$ g). The aliquot was incubated on ice for 20 min and then Solution 3 (1 mL) added to

1 the aliquot. After 20 min incubation at room temperature, Solution 2 (10 mL) added to the mixtures  
2 and the mixture was centrifuged at 2,000 rpm (Beckman JLA10.500) for 5 min. After decantation,  
3 the residue was diluted with Solution 2 (500  $\mu$ L) and the mixture (100  $\mu$ L) was poured onto the  
4 CD agar plate (1.5 %) supplemented with 4.7% of NaCl and then overlaid with the soft-top CD agar  
5 (0.6 %) containing 21.8 % of sorbitol. The plates were incubated at 30 °C for 3-7 days.

6  
7 **Analysis of the metabolites from AO-transformants harboring a STS gene.** Mycelia of each  
8 transformant was inoculated into a MPY medium (1 mL) containing 20 mM of uridine and 0.2% of  
9 uracil in 10 mL test tube. Each culture was incubated at 30 °C for 4 days. The volatile organic  
10 compounds were extracted by a SPME fiber (50/30  $\mu$ m DVB/CAR/PDMS, Stableflex, 24Ga,  
11 Manual Holder), which was conditioned by inserting it into the GC injector to prevent contamination,  
12 during 30 min at room temperature. After extraction, the fiber was pulled into the needle sheath and  
13 the SPME device was removed from the vial and then inserted into the injection port of a GC-MS  
14 QP2010 apparatus (Shimadzu, Kyoto, Japan) with a HP-5 capillary column (0.32 mm  $\times$  30 m, 0.25  
15  $\mu$ m film thickness; J&W Scientific, Folsom, CA). Each sample was injected onto the column at  
16 60 °C in the splitless mode. The column temperature was increased by 4 °C min<sup>-1</sup> to 180 °C. The  
17 flow rate of the helium carrier gas was 0.66 mLmin<sup>-1</sup> (**method A**).

#### 18 19 **Isolation of sesquiterpenes.**

20 **Sterpurene;** Mycelia of AO-*CpSTS1* were inoculated into a solid medium containing polished rice  
21 (100 g) and adenine (10 mg) in 500 mL Erlenmeyer flasks. Each culture was incubated at 30 °C for  
22 12 days. After extraction with ethyl acetate, the extract was concentrated in vacuo to afford crude  
23 extracts. The crude extracts were purified with silica gel column chromatography (hexane) to give  
24 sterpurene (20.0 mg from 1.3 kg of rice medium from AO-*CpSTS1*). The NMR data are in good  
25 agreement with the reported data (Figure S6, 41).

26 **9-Alloaromadendrene;** Mycelia of AO-*CpSTS11* was inoculated into a MPY medium (30 mL)  
27 containing 20 mM of uridine and 0.2 % of uracil in 200 mL Erlenmeyer flasks. Each culture was  
28 incubated at 30 °C for 5 days. The crude extracts were purified with silica gel column  
29 chromatography (hexane) to give 9-alloaromadendrene (1.0 mg from 150 mL of MPY medium  
30 containing uridine and uracil). The NMR data are in good agreement with the reported data (Figure  
31 S6, 42).

32  
33 **cDNA preparation.** Each transformant was grown on MPY medium containing 20 mM of uridine  
34 and 0.2 % of uracil for 3-5 days at 30 °C. Total RNA was extracted from each dried mycelia using  
35 TRIzol® Reagent (Invitrogen) according to the manufacturer's instructions and then treated with  
36 DNase I (Life Technologies) for reverse transcription. Complementary DNA (cDNA) was  
37 synthesized with PrimeScript™ II 1st strand cDNA synthesis kit (Takara) using the oligo dT  
38 primer according to the manufacturer's instructions. The cDNA was used as a template of the PCR  
39 reactions for direct sequencing and subcloning of each STS gene into the pColdI vector.

40  
41 **Estimation of sequences of functionally active STSs.** cDNA sequences derived from AO  
42 transformant were aligned with the predicted sequence by Augustus for identifying the nonspliced

1 introns in cDNA sequences. Alignments were computed by using ClustalW. The resultant  
2 transcripts of putative intronless STSs were then manually examined for the presence of and proper  
3 alignment of the conserved motifs characteristic for sesquiterpene synthases.

4  
5 **Construction of *E. coli* expression plasmids.** Intronless DNA clone was prepared from PCR-based  
6 removal of the nonspliced introns (Figures 3 and 4) as follows; two DNA fragments (upstream- and  
7 downstream regions of the nonspliced intron) were separately amplified from cDNA with primer set  
8 as shown in Supplemental Table S6. Each PCR product was inserted into *Nde*I-, *Eco*RI-, or  
9 *Kpn*I-digested pColdI to construct the following expression plasmids; pColdI-*CpSTS5*,  
10 pColdI-*CpSTS7*, pColdI-*CpSTS8*, pColdI-*CpSTS13*, pColdI-*CpSTS14*, pColdI-*CpSTS16*,  
11 pColdI-*CpSTS17*, pColdI-*ShSTS3*, pColdI-*ShSTS5*, pColdI-*ShSTS10*, pColdI-*ShSTS12*,  
12 pColdI-*ShSTS13*, and pColdI-*ShSTS18*.

13  
14 **Analysis of the metabolites from EC-transformants.** The constructed plasmids were separately  
15 introduced into *E. coli* BL21-Gold(DE3) for overexpression. The transformant was grown at 37 °C  
16 at an OD<sub>600</sub> of ~0.6 in 500 mL flask. After cooling at 4 °C, isopropyl β-D-thiogalactopyranoside  
17 (0.1 mM) was added to the culture. After incubation at 16 °C for 17 h, the volatile organic  
18 compounds were extracted by a SPME fiber (50/30um DVB/CAR/PDMS. Stableflex, 24Ga,  
19 Manual Holder) during 30 min at room temperature. The extracts were then analyzed by a GC-MS  
20 apparatus according to **method A**.

## 21 **Supplemental Material**

22 Supplemental material for this article may be found at xxx.

## 23 **Acknowledgments**

24 This work was financially supported by Grants-in-Aid for Japan Society for the Promotion of  
25 Science Grants JP15H01835 to H.O., JP16H03277 and JP16H06446 to A.M., and JP18K14342 to  
26 L.C.

## 27 **References**

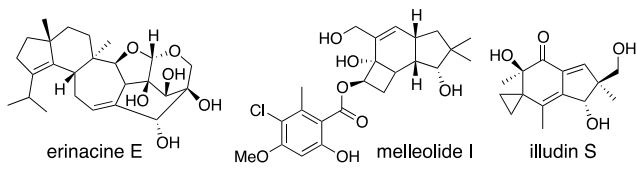
- 28 1. Lorenzen K, Anke T. 1998. Basidiomycetes as a source for new bioactive natural products. *Curr*  
29 *Org Chem* 2;329-364.
- 30 2. Quin MB, Flynn CM, Schmidt-Dannert C. 2014. Traversing the fungal terpenome. *Nat Prod*  
31 *Rep* 31;1449-1473.
- 32 3. Kavanagh F, Hervey A, Robbins WJ. 1951. Antibiotic substances from Basidiomycetes: VIII.  
33 *Pleurotus multilus* (Fr.) Sacc. and *Pleurotus passeckerianus* Pilat. *Proc Natl Acad Sci USA*  
34 37;570-574.<sup>[1]</sup><sup>[SEP]</sup>
- 35 4. Mcmorris TC, Anchel M. 1963. Structures of Basidiomycete metabolites illudin S and illudin M.  
36 *J Am Chem Soc* 85;831-832.

- 1 5. Midland SL, Izac RR, Wing RM., Zaki AI, Munnecke DE, Sims JJ. 1982. Melleolide, a new  
2 antibiotic from *Armillaria mellea*. Tetrahedron Lett 23;2515-2518.
- 3 6. Kawagishi H, Shimada A, Hosokawa S, Mori H, Sakamoto H, Ishiguro Y, Sakemi S, Bordner J,  
4 Kojima N, Furukawa S. 1996. Erinacines E, F, and G, stimulators of nerve growth factor  
5 (NGF)-synthesis, from the mycelia of *Hericiium erinaceum*. Tetrahedron Lett 37;7399-7402.
- 6 7. Martin F, Cullen D, Hibbett D, Pisabarro A, Spatafora JW, Baker SE, Grigoriev IV. 2011.  
7 Sequencing the fungal tree of life. New Phytol 190;818-821.
- 8 8. Stadler M, Hoffmeister D. 2015. Fungal natural products-the mushroom perspective. Front  
9 Microbiol 6;127.
- 10 9. Agger S, Lopez-Gallego F, Schmidt-Dannert C. 2009. Diversity of sesquiterpene synthases in  
11 the basidiomycete *Coprinus cinereus*. Mol Microbiol 72;1181-1195.
- 12 10. Wawrzyn GT, Quin MB, Choudhary S, Lopez-Gallego F, Schmidt-Dannert C. 2012. Draft  
13 genome of *Omphalotus olearius* provides a predictive framework for sesquiterpenoid natural  
14 product biosynthesis in Basidiomycota. Chem Biol 19;772-783.
- 15 11. Quin MB, Flynn CM, Wawrzyn GT, Choudhary S, Schmidt-Dannert C. 2013. Mushroom  
16 hunting by using bioinformatics: Application of a predictive framework facilitates the selective  
17 identification of sesquiterpene synthases in Basidiomycota. ChemBioChem 14;2480-2491.
- 18 12. Ichinose H, Kitaoka T. 2018. Insight into metabolic diversity of the brown-rot basidiomycete  
19 *Postia placenta* responsible for sesquiterpene biosynthesis: semi-comprehensive screening of  
20 cytochrome P450 monooxygenase involved in protoilludene metabolism. Microb Biotechnol  
21 11;952-965.
- 22 13. Burkhardt I, Kreuzenbeck N, Beemelmans C, Dickschat JS. 2019. Mechanistic  
23 characterization of three sesquiterpene synthases from the termite-associated fungus  
24 Termitomyces. Org Biomol Chem 17;3348-3355.
- 25 14. Kupfer DM, Drabenstot SD, Buchanan KL, Lai HS, Zhu H, Dyer DW, Roe BA, Murphy JW.  
26 2004. Introns and splicing elements of five diverse fungi. Eukaryot Cell 3;1088-1100.
- 27 15. Fujii R, Minami A, Tsukagoshi T, Sato N, Sahara T, Ohgiya S, Gomi K, Oikawa H. 2011. Total  
28 biosynthesis of diterpene aphidicolin, a specific inhibitor of DNA polymerase alpha:  
29 Heterologous expression of four biosynthetic genes in *Aspergillus oryzae*. Biosci Biotechnol  
30 Biochem 75;1813-1817.
- 31 16. Tagami K, Liu CW, Minami A, Noike M, Isaka T, Fueki S, Shichijo Y, Toshima H, Gomi K,  
32 Dairi T, Oikawa H. 2013. Reconstitution of biosynthetic machinery for indole-diterpene  
33 paxilline in *Aspergillus oryzae*. J Am Chem Soc 135;1260-1263.
- 34 17. Ye Y, Minami A, Mandi A, Liu C, Taniguchi T, Kuzuyama T, Monde K, Gomi K, Oikawa H.  
35 2015. Genome mining for sesterterpenes using bifunctional terpene synthases reveals a unified  
36 intermediate of di/sesterterpenes. J Am Chem Soc 137;11846-11853.

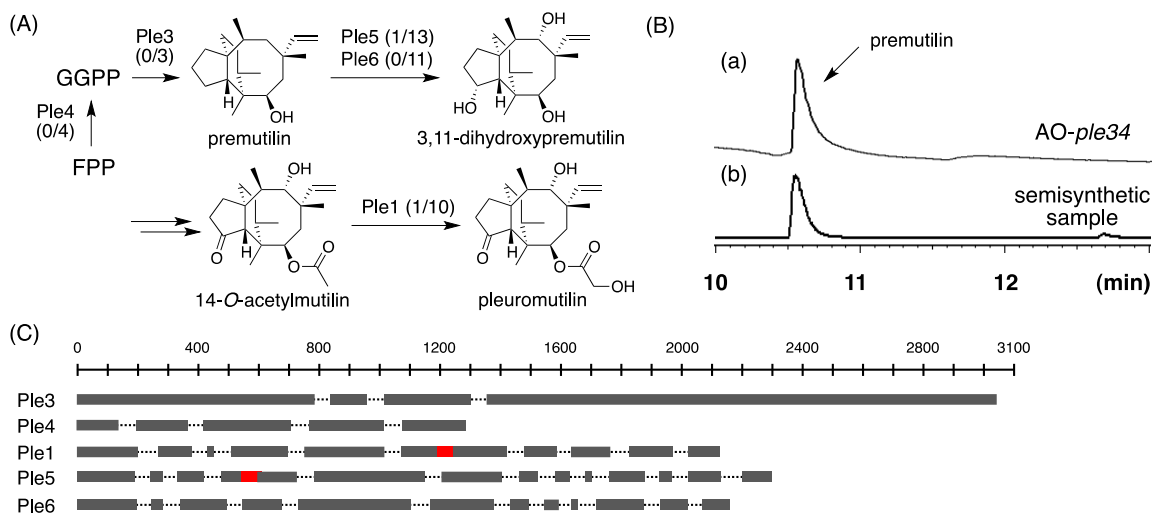
- 1 18. Heneghan MN, Yakasai AA, Halo LM, Song Z, Bailey AM, Simpson TJ, Cox RJ, Lazarus CM.  
2 2010. First heterologous reconstruction of a complete functional fungal biosynthetic multigene  
3 cluster. *ChemBioChem* 11;1508-1512.
- 4 19. Liu C, Tagami K, Minami A, Matsumoto T, Frisvad JC, Suzuki H, Ishikawa J, Gomi K, Oikawa  
5 H. 2015. Reconstitution of biosynthetic machinery for the synthesis of the highly elaborated  
6 indole diterpene penitrem. *Angew Chem Int Ed* 54;5748-5752.
- 7 20. Ye Y, Minami A, Igarashi Y, Izumikawa M, Umemura M, Nagano N, Machida M, Kawahara T,  
8 Shin-ya K, Gomi K, Oikawa H. 2016. Unveiling the biosynthetic pathway of the ribosomally  
9 synthesized and post-translationally modified peptide ustiloxin B in filamentous fungi. *Angew*  
10 *Chem Int Ed* 55;8072-8075.
- 11 21. Matsuda Y, Iwabuchi T, Wakimoto T, Awakawa T, Abe I. 2015. Uncovering the unusual D-ring  
12 construction in terretonin biosynthesis by collaboration of a multifunctional cytochrome P450  
13 and a unique isomerase. *J Am Chem Soc* 137;3393-3401.
- 14 22. Yamane M, Minami A, Liu C, Ozaki T, Takeuchi I, Tsukagoshi T, Tokiwano T, Gomi K, Oikawa  
15 H. 2017. Biosynthetic machinery of diterpene pleuromutilin isolated from Basidiomycete fungi.  
16 *ChemBioChem* 18;2317-2322.
- 17 23. Bailey AM, Alberti F, Kilaru S, Collins CM, de Mattos-Shiple K, Hartley AJ, Hayes P, Griffin  
18 A, Lazarus CM, Cox RJ, Willis CL, O'Dwyer K, Spence DW, Foster GD. 2016. Identification  
19 and manipulation of the pleuromutilin gene cluster from *Clitopilus passeckerianus* for increased  
20 rapid antibiotic production. *Sci Rep* 6;25202.
- 21 24. Phasha MM, Wingfield BD, Coetzee MPA, Santana QC, Fourie G, Steenkamp ET. 2017.  
22 Architecture and distribution of introns in core genes of four *Fusarium* species. *G3-Genes*  
23 *Genom Genet* 7;3809-3820.
- 24 25. Dickschat JS. 2014. Capturing volatile natural products by mass spectrometry. *Nat Prod Rep*  
25 31;838-861.
- 26 26. Katayama T, Tanaka Y, Okabe T, Nakamura H, Fujii W, Kitamoto K, Maruyama J. 2016.  
27 Development of a genome editing technique using the CRISPR/Cas9 system in the industrial  
28 filamentous fungus *Aspergillus oryzae*. *Biotechnol Lett* 38;637-642.1
- 29 27. Ugai T, Minami A, Fujii R, Tanaka M, Oguri H, Gomi K, Oikawa H. 2015. Heterologous  
30 expression of highly reducing polyketide synthase involved in betaenone biosynthesis. *Chem*  
31 *Commun* 51;1878-1881.
- 32 28. Maruyama J, Kitamoto K. 2008. Multiple gene disruptions by marker recycling with highly  
33 efficient gene-targeting background (Delta ligD) in *Aspergillus oryzae*. *Biotechnol Lett*  
34 30;1811-1817.
- 35 29. Hoff K J, Stanke M. 2013. WebAUGUSTUS-a web service for training AUGUSTUS and  
36 predicting genes in eukaryotes. *Nucleic Acids Res* 41;W123-W128.
- 37 30. Miller DJ, Allemann RK. 2012. Sesquiterpene synthases: passive catalysts or active players?  
38 *Nat Prod Rep* 29;60-71.

- 1 31. Syed K, Shale K, Pagadala NS, Tuszynski J. 2014. Systematic identification and evolutionary  
2 analysis of catalytically versatile cytochrome P450 monooxygenase families enriched in model  
3 Basidiomycete fungi. *Plos One* 9;e86683.
- 4 32. Alquezar B, Rodriguez A, de la Pena M, Pena L. 2017. Genomic analysis of terpene synthase  
5 family and functional characterization of seven sesquiterpene synthases from *Citrus sinensis*.  
6 *Front Plant Sci* 8;1481.
- 7 33. Chou WKW, Fanizza I, Uchiyama T, Komatsu M, Ikeda H, Cane DE. 2010. Genome mining in  
8 *Streptomyces avermitilis*: Cloning and characterization of SAV\_76, the synthase for a new  
9 sesquiterpene, avermitilol. *J Am Chem Soc* 132;8850-8851.
- 10 34. Dickschat JS, Martens T, Brinkhoff T, Simon M, Schulz S. 2005. Volatiles released by a  
11 *Streptomyces* species isolated from the North Sea. *Chem Biodivers* 2;837-865.
- 12 35. Flynn CM, Schmidt-Dannert C. 2018. Sesquiterpene synthase-3-hydroxy-3-methylglutaryl  
13 coenzyme A synthase fusion protein responsible for hirsutene biosynthesis in *Stereum hirsutum*.  
14 *Appl Environ Microb* 84;e00036-18.
- 15 36. Jin FJ, Maruyama J, Juvvadi PR, Arioka M, Kitamoto K. 2004. Development of a novel  
16 quadruple auxotrophic host transformation system by argB gene disruption using adeA gene and  
17 exploiting adenine auxotrophy in *Aspergillus oryzae*. *FEMS Microbiol Lett* 239;79-85.
- 18 37. Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular evolutionary genetics analysis  
19 version 7.0 for bigger datasets. *Mol Biol Evol* 33;1870-1874.
- 20 38. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high  
21 throughput. *Nucleic Acids Res* 32;1792-1797.
- 22 39. Jones DT, Taylor WR, Thornton JM. 1992. The rapid generation of mutation data matrices from  
23 protein sequences. *Comput Appl Biosci* 8;275-282.
- 24 40. Felsenstein J. 1985. Confidence limits on phylogenies: An approach using the bootstrap.  
25 *Evolution* 39;783-791.
- 26 41. Ayer WA, Saedi-Ghomi MH. 1981. 1-sterpurene-3,12,14-triol and 1-sterpurene, metabolites of  
27 silver-leaf disease fungus *Stereum purpureum*. *Can J Chem* 59;2536-2538.
- 28 42. Friedel HD, Matusch R. 1987. Neue aromadendran-derivate aus tolu-balsam. *Helv Chim Acta*  
29 70;1753-1759.
- 30

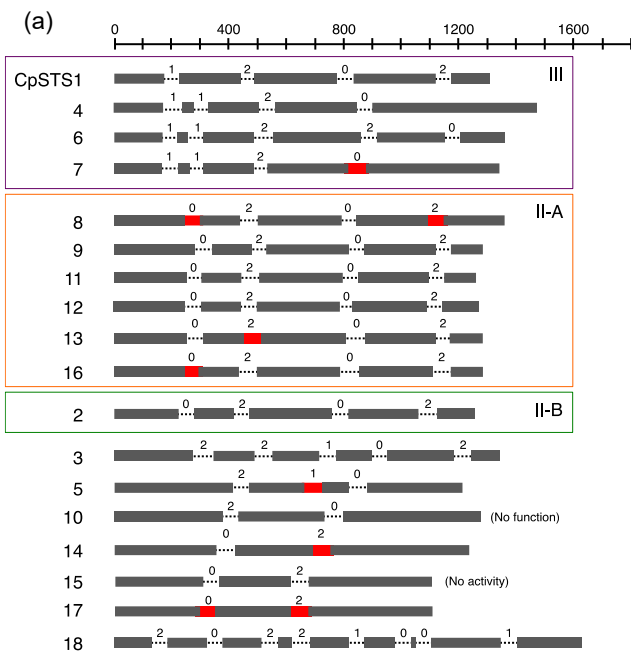




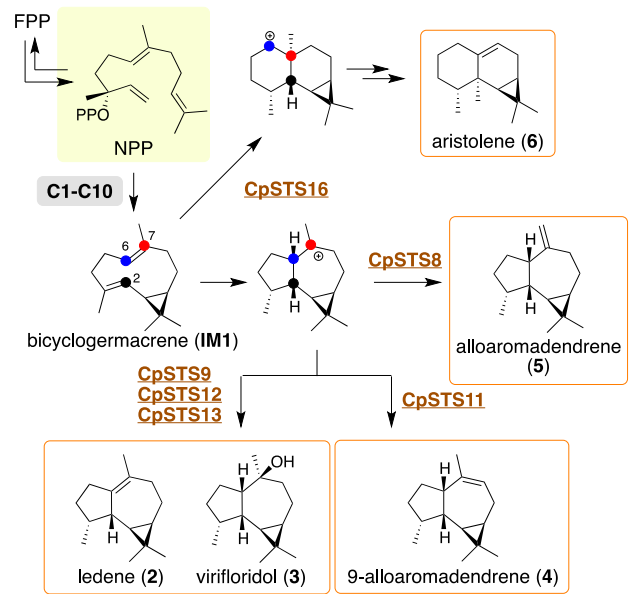
**Figure 1.** Chemical structures of biologically active terpenoids produced by Basidiomycota fungi.



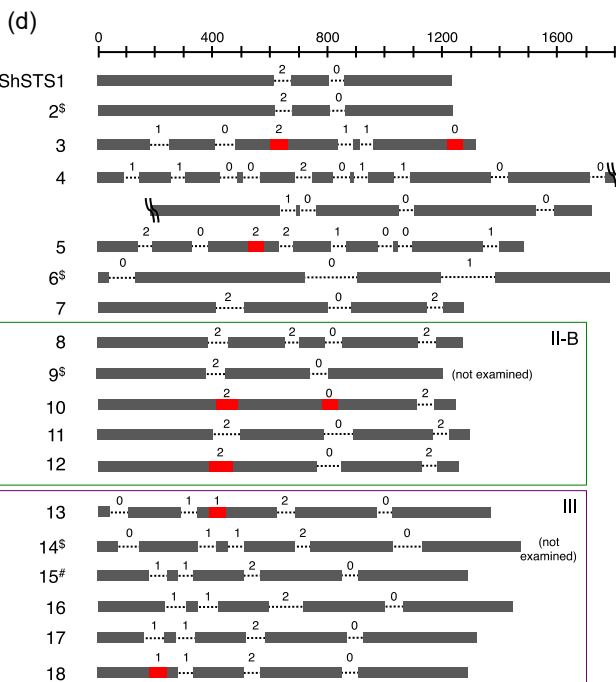
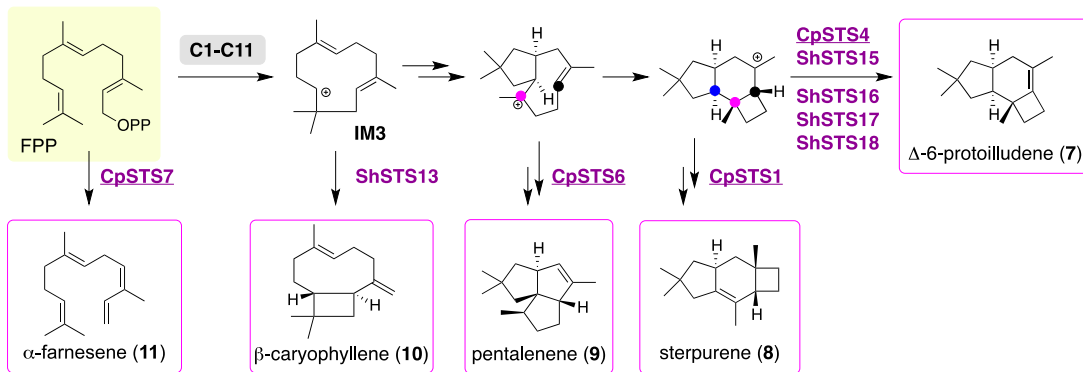
**Figure 2.** (A) Biosynthetic pathway of pleuromutilin. The numbers in parentheses are the nonspliced introns in the cDNA recovered from AO-TF and the total introns of the corresponding gene. (B) GC-MS profiles of the metabolites from (a) AO-*ple34* and (b) synthetic sample. (C) Schematic view of the gene structures of *ple3*, *ple4*, *ple1*, *ple5*, and *ple6*. Exon regions are shown in shaded boxes. Non-spliced introns are shown in red boxes.



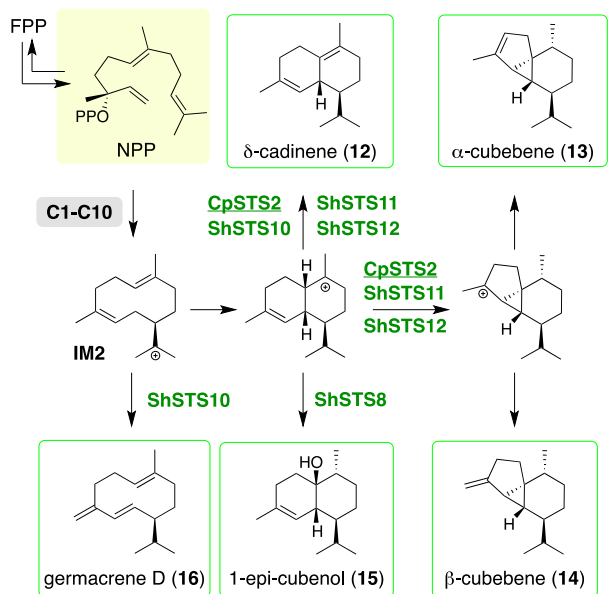
(b) C1-C10 cyclization catalyzed by clade II\_A STSs



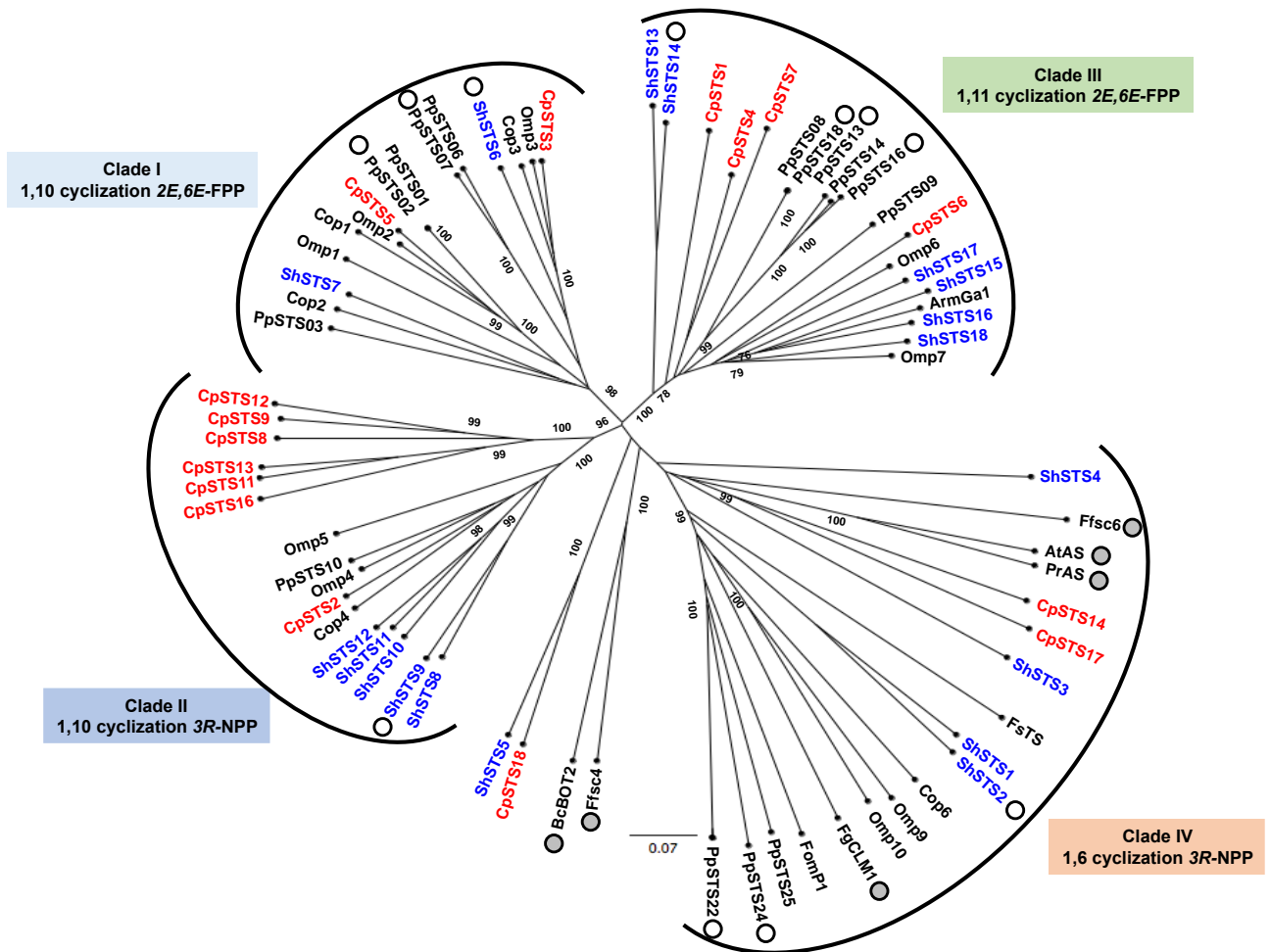
(c) C1-C11 cyclization catalyzed by clade III STSs



(e) C1-C10 cyclization catalyzed by clade II\_B STSs



**Figure 3.** (a) Schematic view of the gene structures of CpSTSs. The dark bars indicate exons and the dotted lines do introns along with intron phases. The red bars represent the computationally predicted introns that were not spliced (skipped) in cDNA from AO-TFs, of which removal resulted in functional STS expression in *E. coli* TF. Genes with the colored boxes are proposed paralogs or orthologs in the *C. pseudo-pinsitus* and *S. hirsutum* that belong to cyclization clade III (purple) or clade II\_A (orange) or II\_B (green). Non-functional refers to CpSTS10 lacking conserved active site motifs. (b) Proposed cyclization mechanism of terpene products catalyzed by clade II\_A STSs. (c) Proposed cyclization mechanism of terpene products catalyzed by clade III STSs. (d) Schematic view of the gene structures of ShSTSs. A gene structure of functional ShSTS15 with # was reported previously (31). Genes with \$ denote the computationally predicted gene structure of STS not investigated in this work. STSs from *C. pseudo-pinsitus* are underlined in (b), (c), and (e). Stereochemistries of putative carbocation intermediates are predicted by those of cyclized products.



**Figure 4.** Phylogenetic tree of functionally characterized STSs except four ShSTSs (white circles) and seven PpSTSs (white circles) derived from Basidiomycetes fungi. STSs derived from *C. pseudo-pinsitus* and *S. hirsutum* are described in red and blue colors, respectively. STSs derived from Ascomycetes fungi are indicated by grey circle.



**Figure 5.** Schematic view of the gene structures of active and mispredicted STSs. The blue bars indicate the mispredicted sites in computationally predicted (Augustus (26)) gene structures compared to the functional gene structure shown directly above.