*Article*

# Genome Assembly and Genetic Traits of the Pleuromutilin-Producer *Clitopilus passeckerianus* DSM1602

Thomas Schafhauser [1,*], Daniel Wibberg [2,3], Antonia Binder [1,4], Christian Rückert [2], Tobias Busche [2], Wolfgang Wohlleben [1,5,6] and Jörn Kalinowski [2,*]

1 Mikrobiologie und Biotechnologie, Interfakultäres Institut für Mikrobiologie und Infektionsmedizin, Eberhard Karls Universität Tübingen, Auf der Morgenstelle 28, 72076 Tuebingen, Germany
2 Centrum für Biotechnologie, CeBiTec, Universität Bielefeld, Universitätsstr. 27, 33615 Bielefeld, Germany
3 Institute of Bio- and Geosciences IBG-5, Computational Metagenomics, Forschungszentrum Jülich GmbH, 52425 Juelich, Germany
4 Institut für Mikrobiologie, Technische Universität Dresden, Zellescher Weg 20b, 01062 Dresden, Germany
5 Cluster of Excellence EXC 2124—Controlling Microbes to Fight Infections, 72076 Tuebingen, Germany
6 German Center for Infection Research (DZIF), Partner Site Tübingen, 72076 Tuebingen, Germany
* Correspondence: thomas.schafhauser@biotech.uni-tuebingen.de (T.S.); joern@cebitec.uni-bielefeld.de (J.K.)

**Abstract:** The gilled mushroom *Clitopilus passeckerianus* (Entolomataceae, Agaricales, Basidiomycota) is well known to produce the terpenoid pleuromutilin, which is the biotechnological basis for medically important antibiotics such as lefamulin and retapamulin. Their unique mode of action and good tolerance entails an increasing demand of pleuromutilin-derived antibiotics in veterinary and human health care. Surprisingly, despite their pharmaceutical importance, no genome sequence is available of any pleuromutilin-producing fungus. Here, we present the high-quality draft genome sequence of the pleuromutilin-producer *C. passeckerianus* DSM1602 including functional genome annotation. More precisely, we employed a hybrid assembly strategy combining Illumina sequencing and Nanopore sequencing to assemble the mitochondrial genome as well as the nuclear genome. In accordance with the dikaryotic state of the fungus, the nuclear genome has a diploid character. Interestingly, the mitochondrial genome appears duplicated. Bioinformatic analysis revealed a versatile secondary metabolism with an emphasis on terpenoid biosynthetic enzymes in *C. passeckerianus* and also in related strains. Two alleles of biosynthetic gene clusters for pleuromutilin were found in the genome of *C. passeckerianus*. The pleuromutilin genes were reassembled with yeast-specific elements for heterologous expression in *Saccharomyces cerevisiae*. Our work lays the foundation for metabolic strain engineering towards higher yields of the valuable compound pleuromutilin.

## 1. Introduction

The basidiomycete *Clitopilus passeckerianus* is a saprotrophic fungus belonging to the family of Entolomataceae within the fungal order Agaricales, also known as gilled mushrooms. *Clitopilus* species are especially endemic in northern temperate areas [1]. *C. passeckerianus* has been named after its discoverer Friedrich Passecker, an Austrian agronomist, who encountered it as a contamination in agricultural mushroom farms [2]. The fungus occasionally also grows on wood and old horse dung [3], and has even been reported from deep mining galleries [2].

Morphologic and genetic observations [2,4–7] reveal that *C. passeckerianus* undergoes the typical life cycle of Agaricales. Its main form of appearance is the dikaryotic mycelium composed of microscopic hyphae harboring two distinct haploid nuclei [4–7]. The dikaryon is able to differentiate into macroscopic fruiting bodies [2,7], which are characterized by a

tiny stipe, a small white cap, and amber-to-pinkish gills [2]. The gill tissue generates slightly pink-toned [2], haploid basidiospores upon meiosis [7]. The spores germinate on a suitable substrate and produce a haploid mycelium—the monokaryon [7]. When two compatible monokaryons meet and fuse, the life cycle is completed as the dikaryotic mycelium—populated by haploid nuclei of both parental genotypes—is rebuilt [7]. As eukaryotes, fungal cells contain a mitogenome besides the nuclear genome. The mitochondrial DNA (mtDNA) in Agaricales is usually inherited uniparentally (i.e., from one of the parental monokaryons) to give rise to a homoplasmic dikaryon containing only a single type of mtDNA throughout the whole mycelium [8–10]. The mtDNA and its inheritance in *C. passeckerianus* has not yet been investigated.

*Clitopilus passeckerianus* are famous for the production of pleuromutilin, a tricyclic diterpene antibiotic of great medicinal importance, which interferes with protein synthesis of mainly Gram-positive bacteria [11]. Pleuromutilin was first isolated 70 years ago in the golden age of antibiotic discovery from *C. passeckerianus* and *C. scyphoides* [11]—at that time both assigned to the genus *Pleurotus*, hence the compound's name, pleuromutilin. Since then, several pleuromutilin-producing strains have become known, which form a tight monophyletic clade within the Entolomataceae, suggesting that pleuromutilin production has only evolved once [5]. Although it has been used in veterinary medicine for decades, pleuromutilin has just recently been applied in human medicine in the form of the semi-synthetic derivatives retapamulin (for topic application) and lefamulin (for systemic application). Despite the longtime usage in livestock, resistance against pleuromutilin is very rare and develops rather slowly [12,13]. Moreover, the unique binding site of pleuromutilins at the 23 S rRNA of bacterial ribosome 50 S subunit drastically reduces the chance of cross-resistance to other antibiotics [14,15]. These favorable features together with their good tolerance make pleuromutilins precious drugs in human health care.

The biosynthetic pathway of pleuromutilin has been elucidated simultaneously in *C. passeckerianus* and the related strain *C. pseudo-pinsitus* by two research groups [4,16,17] via the consecutive expression of the involved genes in *Aspergillus oryzae*. According to that, seven *ple* genes encode (i) a geranyl-geranyl diphosphate (GGPP) synthase responsible for combining four C-5 isoprene units to the linear C-20 GGPP, (ii) a terpene synthase catalyzing the key reaction and committed step from GGPP to the tricyclic compound pre-mutilin, and (iii) finally tailoring enzymes such as P450 monooxygenases, a dehydrogenase, and an acetyl-transferase altogether generating the final product pleuromutilin. In both the genomes of *C. passeckerianus* and *C. pseudo-pinsitus*, the *ple* genes are organized in a biosynthetic gene cluster (BGC) with identical architecture [4,17].

The fungus *C. passeckerianus* is genetically tractable using polyethylene glycol-mediated or *Agrobacterium*-mediated transformation [6]. Moreover, molecular tools for gene manipulation such as dominant selection markers and a gene silencing system have been developed [6]. However, all attempts aimed at increasing pleuromutilin production—e.g., by inserting additional biosynthetic genes—were unsuccessful, supposedly due to problems related to the nuclear redundancy [7] and a phenomenon called sense suppression [4]. In contrast, heterologous expression of the pleuromutilin BGC in the ascomycete host *Aspergillus oryzae* led to a stable and increased pleuromutilin production [4,17]. These studies constitute the first gene clusters from a basidiomycete to be successfully expressed in an ascomycete. Besides *A. oryzae*, no other host strain has been tested for heterologous pleuromutilin production.

The availability of complete fungal genomes is expanding rapidly; however, basidiomycetes still are represented only moderately—the Joint Genome Institute [18] currently lists 624 basidiomycete genomes, among them 183 from Agaricales, in contrast to 1448 ascomycete genomes [19]. Surprisingly, despite the pharmaceutical importance of pleuromutilin-producing species, no genome sequence is available from any such fungi. Indeed, the root endophyte and non-producer *C. hobsonii* is the only member of the family Entolomataceae whose genome has been sequenced. To facilitate genome assembly

though, the DNA has been extracted from the haploid monokaryon instead of the dikaryotic mycelium of *C. hobsonii* [20].

In this study, we established high-quality nuclear and mitochondrial genome sequences of the dikaryotic fungus *C. passeckerianus* DSM1602 including functional genome annotation. We performed phylogenetic and comparative genome analysis with related gilled mushrooms to estimate genomic similarity of the analyzed fungi. Furthermore, our study provides insights into a pronounced secondary metabolism of *C. passeckerianus* and related strains with an emphasis on terpenoid BGCs. The high-quality *C. passeckerianus* genome will help to investigate fundamental traits in Agaricales genetics and it paves the way for biotechnological approaches that aim at metabolic pathway engineering or complete BGC refactoring of enzymatic pathways such as the pleuromutilin biosynthesis.

## 2. Materials and Methods

### 2.1. Clitopilus passeckerianus Growth and DNA Isolation

The strain *C. passeckerianus* DSM1602 (=ATCC, NRRL 3100), obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ), was routinely grown in 50 mL potato dextrose broth (PDB; Karl Roth GmbH, Karlsruhe, Germany) in a 100 mL baffled flask at 180 rpm and 25 °C. Whole genome DNA was isolated and purified with the 'NucleoBond HMW DNA' kit from Macherey-Nagel (Düren, Germany) according to the manufacturer's instructions. The isolated DNA was used for sequencing when the required spectroscopic criteria were met (260/280 ratio value: ~1.8; 260/230 ration value: 2.0–2.2) and when a 0.5% agarose control gel confirmed the presence of large DNA fragments (>>10 kbp) and the absence of visible amounts of RNA.

### 2.2. Intron/Exon Analysis of Ple Genes in C. passeckerianus

For intron/exon analysis of the pleuromutilin biosynthesis genes *ple*, RNA was isolated from cultures grown for approx. 7 days in PDB using the 'Quick-RNA Fungal/Bacterial Kit' (cat no R2014; Zymo Research, Freiburg, Germany). Isolated RNA was treated with DNase (cat no AM1906, Thermo Fisher Scientific) and reverse transcribed to cDNA using the 'RevertAid RT Kit' (cat no K1691, Thermo Fisher Scientific, Waltham, MA, USA) with anchored oligo(dT)18-primers. Gene transcripts were PCR-amplified with the Q5 HF DNA polymerase (cat no M0491, New England Biolabs, Ipswich, MA, USA) using primers listed in Tables S1 and S2, sub-cloned using the 'CloneJet PCR Cloning Kit' (cat no K1231, Thermo Fisher Scientific), and Sanger-sequenced at Eurofins Genomics (Ebersberg, Germany).

### 2.3. Nanopore Library Preparation and GridION Sequencing

A sequencing library with genomic DNA of *C. passeckerianus* DSM1602 was prepared as recently described for other species [21,22]. In brief, the Nanopore DNA Sequencing kit (SQK-LSK09, Oxford Nanopore Technologies [ONT], Oxford, UK) was used according to the manufacturer's instructions. Sequencing was performed on an ONT GridION Mk1 sequencer using an R9.4.1 flow cell according to the manufacturer's instructions.

### 2.4. Illumina Library Preparation and MiSeq Sequencing

Whole-genome shotgun PCR-free libraries of *C. passeckerianus* DSM1602 for Illumina sequencing were constructed as described before for other species [22]. Quality of the resulting libraries was controlled by using an Agilent 2000 Bioanalyzer with an Agilent High Sensitivity DNA Kit (Agilent Technologies, Santa Clara, CA, USA) for fragment sizes of 500—1000 bp. Paired-end sequencing was performed on the Illumina MiSeq platform ($2 \times 300$ bp, v3 chemistry). Adapters and low-quality reads were removed by an in-house software pipeline prior to polishing, as recently described [23].

### 2.5. Base Calling, Reads Processing, and Assembly

The software MinKNOW 3.6.5. (ONT) was used to control the run using the 48 h sequencing run protocol. Base calling was performed offline using guppy 3.2.10. The

assembly was performed using Canu v2.1.1 [24]. The resulting contigs were polished with the Illumina reads using Pilon [25], which was run for ten iterative cycles. BWA-MEM [26] was used for read mapping in the first five iterations and Bowtie2 v2.3.2 [27] was used in the remaining five iterations.

### 2.6. Gene Prediction and Genome Annotation

Gene prediction was performed by applying GeneMark-ES (v4.6.2) [28] using default settings and the fungi mode. Predicted genes were functionally annotated using a modified version of the genome annotation platform GenDB 2.0 [29] for eukaryotic genomes as previously described [30]. For automatic annotation within the platform, similarity searches against different databases, including COG [31], KEGG [32], and SWISS-PROT [33], were performed. In addition to genes, putative tRNA genes were identified with tRNAscan-SE [34]. Completeness, contamination, and strain heterogeneity were estimated with BUSCO (v3.0.2) [35], using the fungi-specific single-copy marker genes database (odb9).

### 2.7. Comparative Genome Analyses and Phylogenetic Analysis

Genome comparisons and investigations of phylogeny were accomplished using a modified version of the comparative genomics program EDGAR designed to handle eukaryotic genomes and their multi-exon genes [36], as described recently [21,22]. Average nucleotide identity analysis (ANI), average amino acid identity analysis (AAI), and percentage of conserved proteins (POCP) analysis were performed, as described previously [21,22,37,38].

### 2.8. Identification and Analysis of the Mitogenome

The mitogenome was identified on the basis of two particular criteria. In basidiomycetes, the mitochondrial DNA (mtDNA) generally features a lower GC content than the nuclear genome [39]. Therefore, on the one hand, contigs were identified whose GC content clearly deviated from the average GC content. On the other hand, contigs that were over-represented were considered to belong to the mtDNA since eukaryotes usually contain multiple mitochondria per cell [40]. This led to a single contig that fulfilled both criteria. This contig represents the mitogenome. Subsequently, automatic gene prediction and annotation were performed using the genome annotation systems GenDB 2.0 [29] and Prokka [41]. Intergenic regions were analyzed by BLAST programs [42] for any additional coding sequences that might have been missed in preceding analyses. Sequence interpretation was refined manually by means of splice-site analyses, as described before [43].

### 2.9. Construction of the Yeast-PleBGCs

The yeast-pleBGCs consist of 16–22 discrete DNA elements that were PCR-amplified (Q5 DNA polymerase, New England Biolabs, Ipswich, MA, USA) using either cDNA of *C. passeckerianus* or DNA molecules synthesized by Biocat (Heidelberg, Germany) as a template. The DNA fragments were then step-wise assembled in vitro (see Figure S2 for details) and propagated in *Escherichia coli* 'NovaBlue Singles Cells Novagen' (Merck, Darmstadt, Germany). First, in the case of yeast-pleBGC-C, the fragments were fused pairwise by fusion PCR (overlap extension PCR, see Table S3). Next, between three and five DNA fragments were assembled by Gibson assembly (NEBuilder HiFi DNA assembly mix, cat no E2621, New England Biolabs) and introduced at the same time in the vector pBluescript, which was linearized with EcoRV. Finally, the assembled fragments were excised from pBluescript using type IIS restriction enzymes to generate DNA fragments (four per assembly) without remnants of the restriction site. This allowed seamless Gibson assembly of the excised large fragments with an EcoRI/Kpn2I-fragment of pBR322. The EcoRI/Kpn2I-fragment of pBR322 contains an ampicillin resistance cassette and an origin of replication for selection and propagation in *E. coli* (see Figure S3). The sequences of the yeast-pleBGCs are available as genbank-files in the Supplementary Materials.

## 3. Results and Discussion

### 3.1. Generation of a High-Qualitiy Draft Genome of Clitopilus passeckerianus DSM1602

In order to decipher the complete genome sequence of *Clitopilus passeckerianus* DSM1602, third-generation Nanopore sequencing together with second-generation Illumina sequencing were applied. Nanopore sequencing resulted in 22.81 Gb, which were assembled into 166 high-quality contigs with a total size of 65.6 Mb and an average GC content of 49.4%. For the final assembly, the contigs obtained from Nanopore sequencing were polished with raw data from a whole-genome shotgun Illumina paired-end sequencing approach. In total, 27,568 SNPs, 30 ambiguous bases, 88,514 small insertions, and 8213 small deletions were identified and corrected by this polishing approach. The combination of long-read (Nanopore) and short-read (Illumina) sequencing improved base accuracy and significantly reduced error rates in the final genome.

The obtained genome sequence was applied to automatic annotation as described in the Section 2 which resulted in 23,566 protein coding sequences (CDS), 422 tRNA genes, and 126 rRNA genes (Table 1).

**Table 1.** Sequencing and assembly statistics of the *C. passeckerianus* genome.

| Feature | Value |
| --- | --- |
| Total number of raw bases (ONT) | 22.81 Gb |
| Number of reads (ONT) | 4.59 million |
| Total number of raw bases (Illumina) | 4.70 Gb |
| Number of reads (Illumina) | 15.7 million |
| Number of contigs | 167 |
| Largest contig | 3,100,797 bp |
| Total length | 65,616,532 bp |
| GC content | 49.43% |
| N50 (shortest contig length at ½ genome length) [1] | 673,524 bp |
| Number of predicted genes | 23,566 |

[1] N50 can be described as a weighted median statistic such that 50% of the entire assembly is contained in contigs equal to or larger than this value [44].

A BUSCO analysis was performed to estimate the genomic data quality. This analysis revealed that 93.4% (284 of 303) of eukaryotic core genes—defined by the BUSCO dataset [35]—were completely detected, 2.0% partially, and only 4.0% were missing. According to our experience, fungal genomes can be regarded as fully covered when the value is above 90%. Therefore, the assembled genome sequence of *C. passeckerianus* is assumed to be complete.

In summary, the combination of Illumina and Nanopore sequencing resulted in a high-quality and state-of-the-art fungal draft genome sequence highly suitable for further genome analysis.

### 3.2. The C. passeckerianus Genome Has a Diploid Character

The DNA used for sequencing was isolated from untreated dikaryotic mycelium that contained a set of two distinct haploid nuclei. Therefore, genes were expected to be present in allelic variants and the sequence dataset should have had a diploid character. To verify this, several analyses were performed. In a first analysis, all contigs were compared with each other by BLAST in order to detect local similarities. It was observed that a large number of contigs share similarity with (at least) one other contig at regions of more than 1000 bp of lengths. The largest matching region is longer than 100 kb, featuring a sequence identity of 99%. The two corresponding contigs likely display (parts of) homologous chromosomes of the two haploid nuclei. Apparently, this contig pair contains allelic variants of genes that are significantly different to each other, which is why they were not assembled into one contig. Moreover, the occurrence of sequence differences is not evenly distributed over the aligned contig pair, indicating that some regions are more conserved than others.

In a second analysis, the BUSCO results (see above) were investigated for duplications. It turned out that among the identified eukaryotic core genes approx. 50% appeared twice—which strongly points to the presence of homologous chromosomes. The reason that the duplication of core genes is not closer to 100% can be explained by the fact that the other half of the core genes is obviously very highly conserved so that no allelic variants can be discriminated and hence they appear only once in the screening.

In a third analysis, a contig-length vs. read-count plot was calculated based on an Illumina-only assembly. The plot illustrates the coverage of individual contigs compared to the average coverage. As we could show in previous studies for other fungi [23,37,38,45] most contigs are generally covered 0.5–1.5×. However, in diploid genomes, an additional group of contigs with coverage rates of 1.5–3.0× appears. This is also the case here, indicated as the groups II and III in Figure 1. Contigs of these two groups encode typical housekeeping genes and they almost represent the entire nuclear genome of *C. passeckerianus*. The two groups were interpreted as follows: group II comprises sequences of homologous chromosomes that are sufficiently different to be assembled as separate contigs, whereas group III consists of contigs with identical or nearly identical allelic variants of homologous chromosomes, which thus are covered by twice as many reads as group II contigs. In other words, the two haploid nuclei of *C. passeckerianus* are in parts so similar that they are represented on identical contigs (group III) in the dataset.
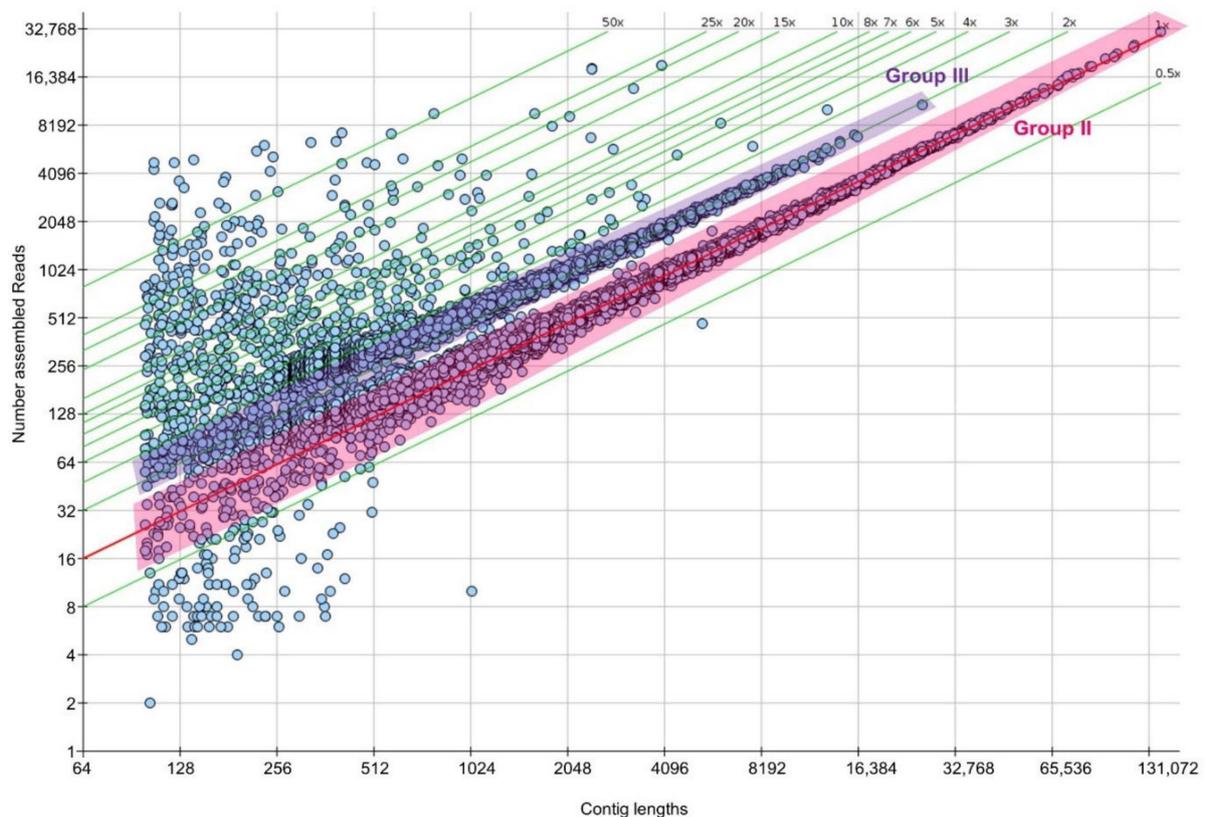


**Figure 1.** Contig-length vs. read-count plot. Dots represent the length of a contig (*x*-axis, in bp) plotted against the number of reads (*y*-axis) assembled into that contig. The red line displays the calculated coverage of unique contigs (1×), green lines display multiples of that coverage rate. Several contig groups can be discriminated: (I) underrepresented contigs with a coverage of below 0.5 (usually caused by inevitable contamination), (II) contigs with a coverage between 0.5× and 1.5× (chromosomal contigs, highlighted), (III) contigs with a coverage between 1.5× and 3× (chromosomal contigs, highlighted), (IV) contigs with a coverage between 3× and 50× (mtDNA), and (V) contigs with a coverage above 50× (ribosomal DNA).

Taken together, the three analyses clearly demonstrate a diploid character of the sequence dataset. This reflects the dikaryotic state of the mycelium of *C. passeckerianus* and is in accordance to previous observations [4–7].

### 3.3. The Mitogenome Appears in Duplicate

One contig of the *C. passeckerianus* genome dataset was identified as the mitogenome, as described in the Material and Methods section. The GC content of the mitochondrial DNA (mtDNA) is only 27.1% and clearly differs from that of the nuclear genome (49.4%). Interestingly, this contig is unusually long (87 kb) and harbors two copies of the mitochondrial genes. Because such a duplication is very uncommon and could be an artifact of the assembly process, we manually bisected the contig to obtain two distinct circular mtDNA chromosomes (mtDNA-1 and mtDNA-2). They have the same coverage rate of sequence reads and slightly differ in size and in the sequence (Tables 2 and 3). In both mtDNAs, 32 protein coding sequences and 26 sequences encoding tRNAs for all 20 amino acids were identified. Moreover, small and large rRNA loci are present in the mtDNAs. The small subunit ribosomal RNA genes (16 S rRNA gene = rns) are identical in both mtDNAs. Based on sequence comparisons with bacterial 16 S rRNA genes, the mitogenome of *C. passeckerianus* might originate from a bacterium related to ancestors of the family Peptoniphilaceae. The mitochondrial 16 S rRNA shows the highest degree of similarity (68%) to *Lagierella massiliensis* strain SIT14.

**Table 2.** Features of the two mtDNA chromosomes.

| Feature | MtDNA-1 | MtDNA-2 |
|---|---|---|
| Length | 40.291 bp | 40.382 bp |
| Number of CDS | 32 | 32 |
| Number of encoded tRNA | 26 | 26 |
| Number of encoded rRNA clusters | 3 | 3 |
| GC content | 27.1% | 27.1% |

**Table 3.** Similarity of the two mtDNA chromosomes.

| | |
|---|---|
| **Sequence Identity** | 99.63% |
| **Alignment Length** | 40,405 bp |
| **Identical Positions** | 40,259 bp |
| **Gaps** | 137 |

The presence of two sets of mitochondrial genes is exceptional and leaves room for three interpretations (see Figure 2): (i) the mitochondrial chromosome consists of two consecutive repeats of mitochondrial genes making it twice as long as normal—as was the original outcome of the Nanopore sequence assembly, (ii) the mycelium is heteroplasmic containing two sets of different mitochondria per cell with 'normal-sized' chromosomes, or (iii) the mycelium is chimeric or mosaic with respect to the mitochondrion and comprises two sets of homoplasmic hyphae, each with a distinct 'normal-sized' mtDNA.
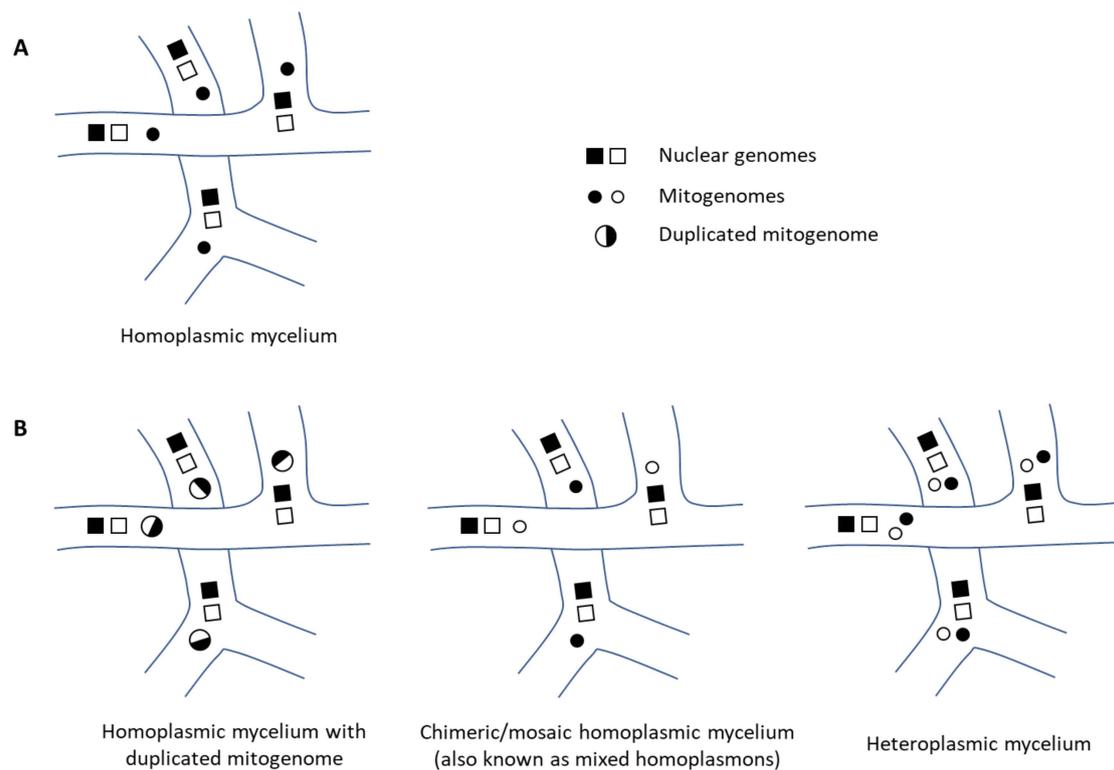
**Figure 2.** Distribution of mitogenomes and nuclear genomes in fungal mycelium. (**A**) Common scenario in Agaricales. (**B**) Possible scenario in *C. passeckerianus*.

To ensure cellular integrity, mitochondria need to be inherited from the parental hyphae to the daughter hyphae. In Agaricales, the inheritance of mitochondria is best studied in laboratory mating experiments of *Agaricus* spp. [46–48], *Coprinus cinereus* [49,50], *Agarocybe aegerita* [51], and *Armillaria* spp. [8–10], from which the following general picture emerges. When two monokaryons with distinct nuclei and distinct mitochondria mate and fuse during anastomosis, the resulting hyphal compartment is dikaryotic and heteroplasmic [10,49]. However, this state is only transitional as the different mitochondria quickly segregate into subsequent hyphae that recover from the junction [46] and upon only a few events of cell division the proceeding dikaryotic mycelium is homoplasmic again [47,48], containing either one of the two mitochondria. In parallel, both of the nuclei usually spread reciprocally from the anastomosed compartment to the resident mycelium of the other mating type, respectively, while the mitochondria do not migrate [46,50]. As a consequence of both events—mitochondrial segregation in the novel hyphae and nuclear migration in the resident hyphae—the mycelium as a whole has become a chimeric or mosaic mycelium with respect to the mitochondria [9,49–51]. In contrast, the nuclear genome is identical within the—now completely dikaryotic—mycelium. Since both mating types can contribute their mitochondria to the homoplasmic progeny, this pattern of inheritance is probably best referred to as doubly uniparental [52]. Several experiments indicate that the chimeric state of the mycelium, however, is not permanent and one homoplasmon eventually outcompetes the other, which results in the dominant presence of one mtDNA [9,51,53]. This is in accordance with the observation that the vast majority of natural samples from Agaricales are predominantly homoplasmic [8–10,54,55]. With that in mind, and the fact that the DNA of *C. passeckerianus* had not been extracted from recently mated mycelium, the observed duplicated state of the mtDNA in *C. passeckerianus*—including all proposed interpretations—certainly displays a peculiarity and is a matter for further investigations.

*3.4. Phylogeny and Comparative Genome Analysis Confirms Genetic Proximity of Clitopilus spp.*

In order to verify the taxonomic placement of *C. passeckerianus* within the basidiomycete order Agaricales, the comparative genomics platform EDGAR 3.0 [36] was applied. Eleven genomes of sequenced Agaricales members served as a comparison. The bioinformatic program calculated 1303 core genes for the selected species, on the basis of which a phylogenetic tree was computed. According to that, four groups can roughly be discriminated (Figure 3), at which *C. passeckerianus* clearly clusters with *C. hobsonii*, another representative of the species *Clitopilus* and the only one with an available genome sequence.
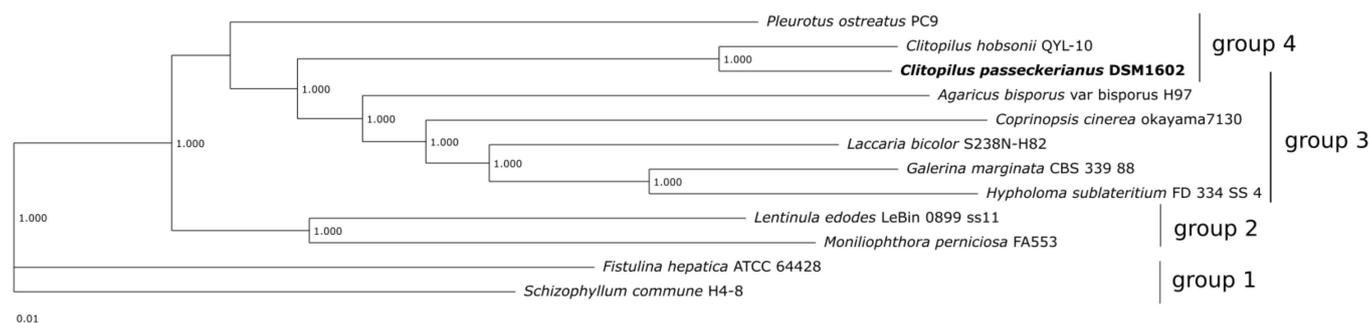


**Figure 3.** Phylogenetic tree of twelve selected Agaricales. The phylogenetic tree was build using SH-like local support values computed with 'FASTtree 2.1′ within the comparative genomics tool 'EDGAR 3.0′ on the bases of 1303 core genes. Branches show a perfect support value of 1.000.

The other selected Agaricales are only distantly related to the genus *Clitopilus* and two fungi, *Fistulina hepatica* and *Schizophyllum commune*, are genetically so different that they appear as outgroups in the tree. The phylogenetic results are also reflected by calculations of genome-wide pairwise similarity, the Average Nucleotide Identity (ANI), the Average Amino Acid Identity (AAI), and the Percentage of Conserved Proteins (POCP) (see Tables S4–S6). These calculations disclose high values only for the closely related *Clitopilus* species. For example, *C. passeckerianus* shares 83.7% of its encoded proteins with *C. hobsonii* (based on POCP) and the ANI value of both genomes is 78.1%. As we could recently demonstrate, fungal strains or isolates from the same species have ANI values above 97% [22,38,56]. Likewise, POCP analysis were shown to be suitable to define genus boundaries. A threshold value of 70% is regarded appropriate to separate fungal genera from fungal species [21,22,38]. Accordingly, the investigations on genomic similarities clearly show that *C. passeckerianus* belongs to the same genus as *C. hobsonii* and both fungi constitute distinct species within this genus.

Similarities in the genomes of both *Clitopilus* strains were investigated further. To determine the conservation of gene blocks in the same relative position of both genomes, a macrosynteny analysis was carried out. This revealed a high degree of genomic synteny showing conserved chromosomes with only a few inversions, rearrangements of small areas, and some duplicated regions (see Figure S1). This makes a similar number of (non-homologous) chromosomes in both genomes likely, which in *C. hobsonii* has been indicated with 10 [20]. In order to compare functional genes of both strains, we needed to conduct a bioinformatic gene annotation also for *C. hobsonii*. In total, 12,973 genes were predicted. This is approx. half the number of genes predicted for *C. passeckerianus* (23,566, see above), which is due to the fact that *C. hobsonii* was sequenced as an artificial monokaryon instead of the natural dikaryon [20]. Based on the comparative genomics tool EDGAR [36], the two *Clitopilus* strains share 8654 core-genes, which represents approx. 75% of all *C. hobsonii* genes. These core-genes encode to a large extent basic cellular processes and play roles in general metabolism according to analyses with respective databases (such as COG, KEGG, and SWISS-PROT). However, both genomes also contain a noticeable number of genes with predicted functions in plant cell wall degradation, as judged by a dbCAN-CAZy [57]

analysis (see Table S7). Besides the core-genes, *C. hobsonii* and *C. passeckerianus* harbor 4832 and 14,339 singletons (paralogous genes are not counted), respectively, and thus the pan-genome consists of 27,825 genes (4832 + 14,339 + 8654). Although some singletons could be assigned to KEGG pathways, most of them are annotated as hypothetical genes.

In summary, the comparative genome analysis illustrates that both species share a significant amount of genetic content. Simultaneously, a sufficient number of individual genes remain that potentially are associated with the different live styles of the fungi: in contrast to the saprotrophic *C. passeckerianus*, *C. hobsonii* has been observed to establish biotrophic symbiosis with plants [58].

### 3.5. Clitopilus passeckerianus Harbors a Huge Genetic Potential for Secondary Metabolites

In order to detect biosynthetic gene clusters (BGCs), coding for terpenes, and other secondary metabolites, an analysis with the fungal version of the bioinformatics tool antiSMASH [59] was performed together with a BLAST search using the terpene synthase of the pleuromutilin BGC [17] from *C. pseudo-pinsitus* (accession number LC314149) as the query. This resulted in the large number of 76 BGCs present in the genome of *C. passeckerianus*, which displays a huge genetic capacity of the fungus to produce secondary metabolites. Especially outstanding is the high number of 55 terpene BGCs, of which two were assigned to the pleuromutilin biosynthesis (more details in the next chapter). In addition to *C. passeckerianus*, we investigated the distribution of BGCs in 10 further Agaricales genomes with the result that terpene BGCs generally were very abundant, particularly in *Galerina marginata* (29 clusters) and in *C. hobsonii* (22 clusters). However, none of them comprised a pleuromutilin BGC. Other frequently found BGCs were those coding for non-ribosomal peptides (NRP) and NRP-like metabolites, whereas BGCs for indoles, fungal-RiPPs, siderophores, type I polyketides, or hybrids thereof, were represented in the low single-digit level (Figure 4).
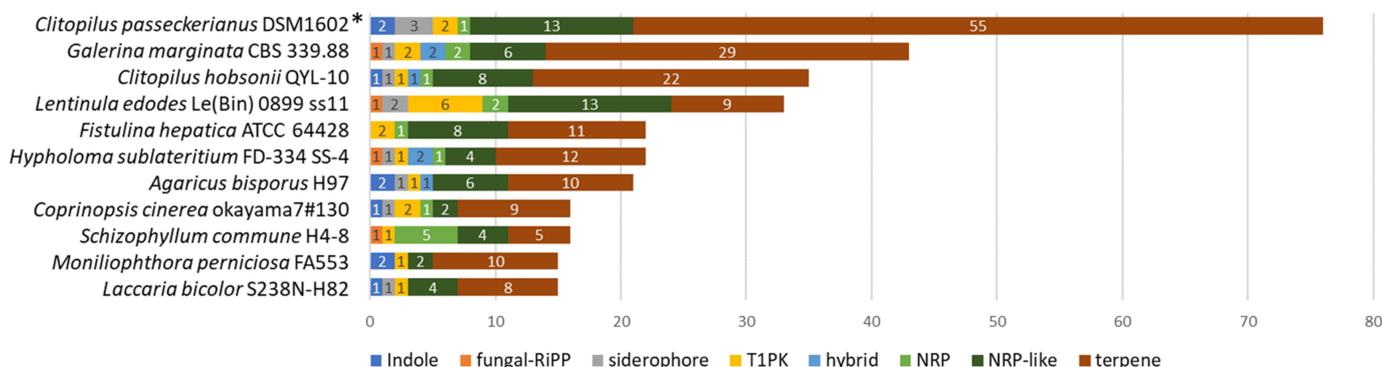


**Figure 4.** Distribution of BGCs in genomes of selected Agaricales. * Note that the genome of *C. passeckerianus* derives from the dikaryon and the actual number of BGCs presumably is even higher (see main text for details).

It needs to be considered that—in contrast to the other genomes investigated here, which are likely haploid monokaryons, as judged from their genomic size—the genome of *C. passeckerianus* derives from the dikaryon. Since the two haploid nuclei of the dikaryon are in parts so similar that they are represented on identical contigs in the dataset (see Section 3.2), the BGCs located on these contigs were identified only once by antiSMASH, although they would appear twice in the cell. Therefore, the total number of BGCs present in the dikaryotic genome of *C. passeckerianus* certainly is significantly higher than 76 and—consequently—the number of unique, non-homologous BGCs certainly is higher than half that value.

The terpene BGCs in *C. passeckerianus* contain a total of 63 predicted terpene synthase genes. Terpene synthases catalyze dephosphorylation and coupling/cyclization reactions of linear isoprenoids of varying length such as GGPP. Two distinct classes of terpene synthases

are distinguished by their substrate activation mechanism, which is either ionization-dependent (class I) or protonation-dependent (class II) [60,61]. In addition, bifunctional enzymes exist. Conserved motifs in the enzymes give hints to the reaction mechanism. The predicted terpene synthases of *C. passeckerianus* were screened for these motifs (see Figure 5). According to that, all but six contain at least one of the characteristic motifs and thus likely display functional enzymes. Moreover, the majority belongs to class I terpene synthases.
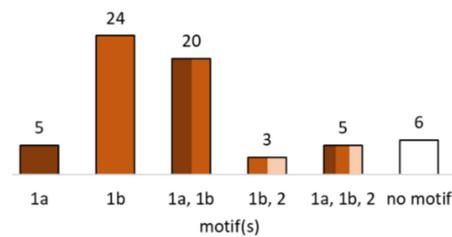


**Figure 5.** Number of predicted terpene synthases in *C. passeckerianus* with conserved motifs. 1a: class I motif DDXX(D/E); 1b: class I motif (N/D)DXX(S/T)XXXE; 2: class II motif DXDD.

Terpenoid production is widespread in fungi [62,63]. Basidiomycetes in particular are known prolific producers [64] and an increasing number of basidiomycete terpene synthases are characterized [64,65]. Nevertheless, our swift analysis suggests that the actual potential to produce terpenoids by members of the Agaricales could even be higher than previously thought. In particular, the numerous terpene synthases in *C. passeckerianus* indicate a pronounced and versatile terpenoid metabolism. Considering that even very similar enzymes are able to produce a manifold terpenoid chemistry [60], it is very likely that terpenoids other than pleuromutilin will be isolated from this fungus in prospective studies.

*3.6. Two Pleuromutilin Biosynthesis Gene Clusters Present in C. passeckerianus*

Two of the terpene BGCs were identified as pleuromutilin BGCs based on their similarity to the pleuromutilin BGC (*ple*BGC) of *C. pseudo-pinsitus* (Figure 6). Cluster architectures coincide and global nucleotide sequence identities are very high. The two *ple*BGCs of *C. passeckerianus* are localized on different contigs with a high degree of sequence identity and gene synteny, which is why they very likely represent allelic variants on homologous chromosomes of the dikaryotic fungus. Introns of the *ple* genes were verified by PCR amplification, cloning, and sequencing of *ple* transcripts, revealing intron numbers and positions that are in accordance with the findings of other groups [4,17] (see Figure 6 and Table 4). The function of *ple* genes in *C. passeckerianus* (Figure 7) has been proofed by Alberti et al. via heterologous expression of one of the clusters in *A. oryzae* [16]. Amino acid sequences of the *ple* gene products from both alleles are nearly identical, suggesting that either allelic cluster is able to autonomously give rise to pleuromutilin biosynthesis. This is in accordance with the fact that pleuromutilin is detected from monokaryotic cultures of *C. passeckerianus* [7].
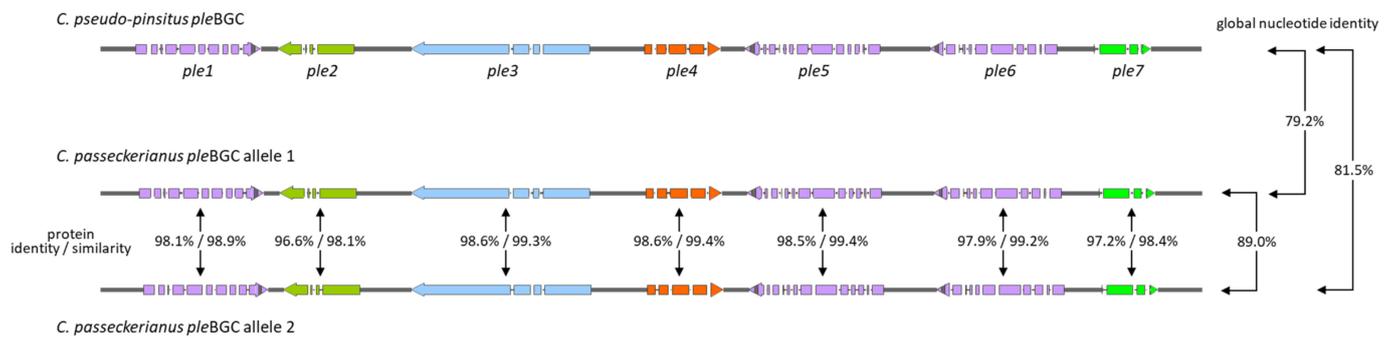
**Figure 6.** Pleuromutilin biosynthetic gene clusters (*ple*BGCs). Introns are illustrated as interruptions in the genes. Nucleotide identities and protein similarities/identities are indicated.

**Table 4.** The pleuromutilin biosynthesis genes in *C. passeckerianus*.

| Gene Name | Alternative Name | Function | *ple*BGC | Length (bp) | Number of Introns | Protein Length (aa) | Nucleotide Identity | Amino Acid Identity (Similarity) |
|---|---|---|---|---|---|---|---|---|
| *ple1* | Pl-p450-3 | Cytochrome P450 | 1 | 2107 | 10 | 523 | 93.9% | 98.1% (98.9%) |
| | | | 2 | 2108 | 10 | 522 | | |
| *ple2* | Pl-atf | Acetyltransferase | 1 | 1301 | 3 | 377 | 93.7% | 96.6% (98.1%) |
| | | | 2 | 1304 | 3 | 377 | | |
| *ple3* | Pl-cyc | Terpene synthase | 1 | 3041 | 3 | 959 | 96.7% | 98.6% (99.3%) |
| | | | 2 | 3040 | 3 | 959 | | |
| *ple4* | Pl-gps | GGPP synthase | 1 | 1291 | 4 | 350 | 95.0% | 98.6% (99.4%) |
| | | | 2 | 1290 | 4 | 350 | | |
| *ple5* | Pl-p450-1 | Cytochrome P450 | 1 | 2286 | 13 | 523 | 96.3% | 98.5% (99.4%) |
| | | | 2 | 2284 | 13 | 523 | | |
| *ple6* | Pl-p450-2 | Cytochrome P450 | 1 | 2166 | 11 | 525 | 90.4% | 97.9% (99.2%) |
| | | | 2 | 2166 | 11 | 525 | | |
| *ple7* | Pl-sdr | Dehydrogenase | 1 | 945 | 3 | 253 | 89.7% | 97.2% (98.4%) |
| | | | 2 | 945 | 3 | 253 | | |
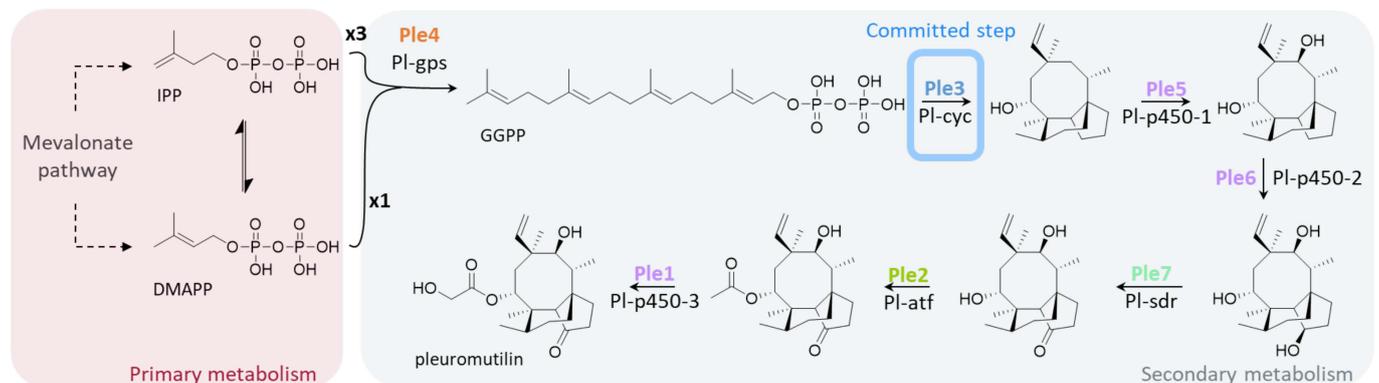


**Figure 7.** Pleuromutilin biosynthesis according to [16,17]. IPP: isopentenyl pyrophosphate; DMAPP: dimethylallyl diphosphate.

### 3.7. Refactoring the pleBGC for Heterolgous Expression in Saccharomyces cerevisiae

Motivated by the success of heterologous pleuromutilin production in the Ascomycete host *Aspergillus oryzae* [16,17] and the recent advances of heterologous expression of basidiomycete terpene biosynthesis genes in yeast [66,67], we envisaged to reconstitute the *ple*BGC for heterologous expression in the fast-growing, well-studied and biotechnologically relevant yeast *Saccharomyces cerevisiae*.

For this purpose, three composite pleuromutilin gene clusters consisting of various yeast-specific *ple* expression cassettes were conceptualized. The clusters were referred to as yeast-*ple*BGCs-A, -B, and -C, and contained all seven *ple* genes controlled by a different set of promoters, respectively (see Figure 8 and Table 5). Every *ple* gene should be flanked by

its own promoter and terminator to avoid homologous sequences in the yeast-*ple*BGCs that may give rise to undesired recombination events in the final yeast host. Promoter sequences were taken from Maury et al., who identified *S. cerevisiae* promoters that were active at late stages during fermentation [68]. Usage of such promoters would allow the decoupling of cell growth from product formation, which is advantageous in industrial fermentation processes as the yeast metabolism is automatically channeled into the desired direction—i.e., biomass or pleuromutilin—during the individual fermentation steps. Given the anticipated functionality of both alleles of the *ple*BGCs in *C. passeckerianus*, one of the clusters was arbitrarily chosen as a basis for the yeast-specific *ple* expression cassettes. Furthermore, a set of seven high-capacity terminators was selected. Sequences of these terminators were taken from Curran et al., who identified terminators that increased mRNA half-life and improved gene expression in *S. cerevisiae* [69]. Besides the *ple* expression cassettes, the conceptualized yeast-*ple*BGCs comprised the hygromycin resistance cassette *hph*MX6 for future selection of transformants. Moreover, the clusters were flanked by delta elements of the yeast retrotransposon Ty1 in order to allow for stable (and multi-copy) chromosomal integration [70] of the yeast-*ple*BGCs in prototrophic *S. cerevisiae* wild-type strains.
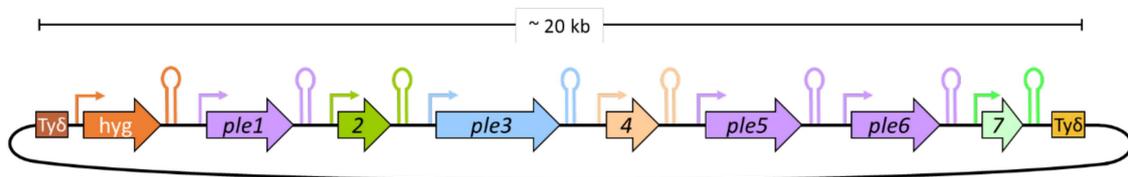


**Figure 8.** Schematic representation of a yeast-*ple*BGC. Tyδ: delta elements of the yeast retrotransposon Ty1; large arrows: genes; small arrows: yeast-specific promoters; hairpin structures: yeast-specific, high-capacity terminators. Expression cassettes are composed of a promoter, gen, and a terminator.

**Table 5.** Composition of the yeast-specific *ple* expression cassettes.

| Promoters [1] | | | Genes [2] | | Terminators [3] |
|---|---|---|---|---|---|
| **Yeast-*ple*BGC-A** | **Yeast-*ple*BGC-B** | **Yeast-*ple*BGC-C** | **All Yeast-*ple*BGCs** | | **All Yeast-*ple*BGCs** |
| pSPG4 | pYIG1 | pPGK1 | *ple1* | (P450-3) | tCYC* |
| pCTA1 | pYEF3 | pTEF1 | *ple2* | (acetyltransferase) | tVPS13 |
| pACS1 | pTDH3 | pADH2 | *ple3* | (terpene synthase) | tCPS1 |
| pICL1 | pADH1 | pADY2 | *ple4* | (GGPP synthase) | tPRM9 |
| pADH2 | pPGK1 | pFBA1 | *ple5* | (P450-1) | tHIS5 |
| pATO1 | pTEF1 | pATO1 | *ple6* | (P450-2) | tSPG5 |
| pADY2 | pFBA1 | pCTA1 | *ple7* | (dehydrogenase) | tUBX6 |

[1] According to [68]. [2] Amplified from cDNA of *C. passeckerianus ple*BGC-1. [3] According to [69].

In order to construct the designed yeast-*ple*BGCs, all the individual elements were PCR-amplified. In the case of the *ple* genes, the cDNA of *C. passeckerianus* served as a template to avoid the missplicing of the intron-dense *ple* genes. All other DNA elements were amplified from synthesized DNA molecules. In a next step, the discrete DNA elements (up to 22 per yeast-*ple*BGC) were step-wise assembled using a combination of fusion-PCR, type IIS restriction enzymes for seamless cleavage of intermediate products, and two consecutive rounds of Gibson assembly (see Figure S2 for more details). The final vectors carrying correctly assembled yeast-*ple*BGCs were identified by restriction digestion and the final sequence was verified by Sanger-sequencing of the entire composite gene clusters.

In summary, the pleuromutilin biosynthesis genes were successfully reassembled to yeast-specific gene clusters that are now ready for usage in future experiments. The step-wise Gibson assembly in combination with type IIS restriction enzyme digestion has proved to be a strong tool to put together 22 discrete DNAs to one large molecule. A high-quality genome sequence that includes information on the pleuromutilin biosynthesis genes—which we provided in the course of this study—has been the essential basis for

that approach. In prospective studies, we will transfer the refactored yeast-*ple*BGCs into appropriate *S. cerevisiae* strains and evaluate heterologous pleuromutilin production.

## 4. Conclusions

In this study, a high-quality annotated genome of the pleuromutilin-producing fungus *C. passeckerianus* was established, providing a valuable resource for future transcriptomic, proteomic, and comparative genomic studies, which will help to understand the fundamental biological traits of this important pharmaceutical species. A multi-locus phylogenetic tree was prepared and enabled precise ranking within other sequenced genomes of Agaricales. The findings on duplicated mitochondrial genomes in *C. passeckerianus* sheds new lights on mitochondrial heritage in basidiomycetes. It remains to be seen whether similar mitochondrial features are going to be discovered in further dikaryotic genomes. Possibly, the heteroplasmic state is more widespread in basidiomycete fungi than previously assumed.

In addition, our study paves the way for studies aimed at interfering with the cellular metabolism to boost the production of pleuromutilin. Given the fact that protocols for genetic manipulation have been successfully developed, rational metabolic engineering approaches are now possible. In particular, the abundant terpene synthase genes are attractive targets as they display potential competitors for common precursors in the terpenoid metabolism. Therefore, silencing or deletion of (non-essential) terpene synthase genes should increase the substrate pool for pleuromutilin biosynthesis. Transcription analysis will help to identify the most competing pathways. Moreover, pathway genes of the primary metabolism such as the mevalonate pathway are now accessible and allow for a targeted manipulation of metabolic fluxes towards terpenoid biosynthesis [71,72].

Furthermore, in light of the plethora of BGCs that have been detected, our study revealed an extremely pronounced secondary metabolism of *C. passeckerianus*. Particularly, the large number of terpene synthase genes present in *C. passeckerianus* and related species suggests that much chemical and enzymatic diversity regarding terpenoids and their biosynthesis remains to be discovered. With respect to the fact that pleuromutilin is the only secondary metabolite known from *C. passeckerianus*, the majority of BGCs appear to be silent or only weakly expressed. Therefore, it seems obvious to assume that many more secondary metabolites await discovery using measures to activate these clusters (e.g., alternative cultivation conditions) or by applying sensitive analytical screening methods.

**Conflicts of Interest:** The funders had no role in: the design of the study; in the collection, analyses, or interpretation of data; or in the writing of the manuscript. The funders approved the publishing of the results.

# References

1. Kumla, J.; Suwannarach, N.; Sungpalee, W.; Sri-Ngernyuang, K.; Lumyong, S. Clitopilus lampangensis (Agaricales, Entolomataceae), a new species from northern Thailand. *MycoKeys* **2019**, *58*, 69–82. [CrossRef] [PubMed]
2. Passecker, F. Ein neuer Unkrautpilz auf Champignonbeeten. (*Pleurotus passeckerianus* Pilát). *Z. Pflanzenkrankh. (Pflanzenpathol.) Pflanzenschutz* **1936**, *46*, 271–277.
3. Lohrmann, F. Katzenohrräsling—*Clitopilus passeckerianus*. In Pilzbestimmer. Available online: https://pilzbestimmer.de/Detailed/10833.html (accessed on 14 April 2022).
4. Bailey, A.M.; Alberti, F.; Kilaru, S.; Collins, C.M.; de Mattos-Shipley, K.; Hartley, A.J.; Hayes, P.; Griffin, A.; Lazarus, C.M.; Cox, R.J.; et al. Identification and manipulation of the pleuromutilin gene cluster from *Clitopilus passeckerianus* for increased rapid antibiotic production. *Sci. Rep.* **2016**, *6*, 25202. [CrossRef] [PubMed]
5. Hartley, A.J.; de Mattos-Shipley, K.; Collins, C.M.; Kilaru, S.; Foster, G.D.; Bailey, A.M. Investigating pleuromutilin-producing *Clitopilus* species and related basidiomycetes. *FEMS Microbiol. Lett.* **2009**, *297*, 24–30. [CrossRef]
6. Kilaru, S.; Collins, C.M.; Hartley, A.J.; Bailey, A.M.; Foster, G.D. Establishing Molecular Tools for Genetic Manipulation of the Pleuromutilin-Producing Fungus *Clitopilus passeckerianus*. *Appl. Environ. Microbiol.* **2009**, *75*, 7196–7204. [CrossRef]
7. de Mattos-Shipley, K.M.J.; Foster, G.D.; Bailey, A.M. Insights into the Classical Genetics of *Clitopilus passeckerianus*—The Pleuromutilin Producing Mushroom. *Front. Microbiol.* **2017**, *8*, 1056. [CrossRef]
8. Smith, M.L.; Anderson, J.B. Mitochondrial DNAs of the fungus *Armillaria ostoyae*: Restriction map and length variation. *Curr. Genet.* **1994**, *25*, 545–553. [CrossRef]
9. Smith, M.L.; Duchesne, L.C.; Bruhn, J.N.; Anderson, J.B. Mitochondrial genetics in a natural population of the plant pathogen armillaria. *Genetics* **1990**, *126*, 575–582. [CrossRef]
10. Saville, B.J.; Yoell, H.; Anderson, J.B. Genetic exchange and recombination in populations of the root-infecting fungus Armillaria gallica. *Mol. Ecol.* **1996**, *5*, 485–497. [CrossRef]
11. Kavanagh, F.; Hervey, A.; Robbins, W.J. Antibiotic Substances from Basidiomycetes. *Proc. Natl. Acad. Sci. USA* **1951**, *37*, 570–574. [CrossRef]
12. Paukner, S.; Riedl, R. Pleuromutilins: Potent Drugs for Resistant Bugs—Mode of Action and Resistance. *Cold Spring Harb. Perspect. Med.* **2016**, *7*, a027110. [CrossRef]
13. Van Duijkeren, E.; Greko, C.; Pringle, M.; Baptiste, K.E.; Catry, B.; Jukes, H.; Moreno, M.; Pomba, M.C.M.F.; Pyörälä, S.; Rantala, M.; et al. Pleuromutilins: Use in food-producing animals in the European Union, development of resistance and impact on human and animal health. *J. Antimicrob. Chemother.* **2014**, *69*, 2022–2031. [CrossRef]
14. Poulsen, S.M.; Karlsson, M.; Johansson, L.B.; Vester, B. The pleuromutilin drugs tiamulin and valnemulin bind to the RNA at the peptidyl transferase centre on the ribosome. *Mol. Microbiol.* **2008**, *41*, 1091–1099. [CrossRef]
15. Yang, L.P.H.; Keam, S.J. Retapamulin. *Drugs* **2008**, *68*, 855–873. [CrossRef]
16. Alberti, F.; Khairudin, K.; Venegas, E.R.; Davies, J.; Hayes, P.M.; Willis, C.L.; Bailey, A.M.; Foster, G.D. Heterologous expression reveals the biosynthesis of the antibiotic pleuromutilin and generates bioactive semi-synthetic derivatives. *Nat. Commun.* **2017**, *8*, 1–9. [CrossRef]
17. Yamane, M.; Minami, A.; Liu, C.; Ozaki, T.; Takeuchi, I.; Tsukagoshi, T.; Tokiwano, T.; Gomi, K.; Oikawa, H. Biosynthetic Machinery of Diterpene Pleuromutilin Isolated from Basidiomycete Fungi. *ChemBioChem* **2017**, *18*, 2317–2322. [CrossRef]
18. Nordberg, H.; Cantor, M.; Dusheyko, S.; Hua, S.; Poliakov, A.; Shabalov, I.; Smirnova, T.; Grigoriev, I.; Dubchak, I. The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. *Nucleic Acids Res.* **2013**, *42*, D26–D31. [CrossRef]
19. Grigoriev, I.V.; Nikitin, R.; Haridas, S.; Kuo, A.; Ohm, R.A.; Otillar, R.; Riley, R.; Salamov, A.A.; Zhao, X.; Korzeniewski, F.; et al. MycoCosm portal: Gearing up for 1000 fungal genomes. *Nucleic Acids Res.* **2013**, *42*, D699–D704. [CrossRef]
20. Peng, L.; Shan, X.; Wang, Y.; Martin, F.; Vilgalys, R.; Yuan, Z. Hybrid Genome Assembly and Gene Repertoire of the Root Endophyte *Clitopilus hobsonii* QYL-10 (Entolomataceae, Agaricales, Basidiomycetes). *Mol. Plant-Microbe Interactions* **2021**, *34*, 711–714. [CrossRef]
21. Ullmann, L.; Wibberg, D.; Busche, T.; Rückert, C.; Müsgens, A.; Kalinowski, J.; Blank, L.M. Seventeen Ustilaginaceae High-Quality Genome Sequences Allow Phylogenomic Analysis and Provide Insights into Secondary Metabolite Synthesis. *J. Fungi* **2022**, *8*, 269. [CrossRef]
22. Wibberg, D.; Stadler, M.; Lambert, C.; Bunk, B.; Spröer, C.; Rückert, C.; Kalinowski, J.; Cox, R.J.; Kuhnert, E. High quality genome sequences of thirteen Hypoxylaceae (Ascomycota) strengthen the phylogenetic family backbone and enable the discovery of new taxa. *Fungal Divers.* **2020**, *106*, 7–28. [CrossRef]
23. Wibberg, D.; Andersson, L.; Tzelepis, G.; Rupp, O.; Blom, J.; Jelonek, L.; Pühler, A.; Fogelqvist, J.; Varrelmann, M.; Schlüter, A.; et al. Genome analysis of the sugar beet pathogen *Rhizoctonia solani* AG2-2IIIB revealed high numbers in secreted proteins and cell wall degrading enzymes. *BMC Genom.* **2016**, *17*, 245. [CrossRef] [PubMed]
24. Koren, S.; Walenz, B.P.; Berlin, K.; Miller, J.R.; Bergman, N.H.; Phillippy, A.M. Canu: Scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. *Genome Res.* **2017**, *27*, 722–736. [CrossRef] [PubMed]

25. Walker, B.J.; Abeel, T.; Shea, T.; Priest, M.; Abouelliel, A.; Sakthikumar, S.; Cuomo, C.A.; Zeng, Q.; Wortman, J.; Young, S.K.; et al. Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. *PLoS ONE* **2014**, *9*, e112963. [CrossRef]

26. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *Genomics* **2013**, arXiv:1303.3997. [CrossRef]

27. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [CrossRef]

28. Borodovsky, M.; Lomsadze, A. Eukaryotic Gene Prediction Using GeneMark.hmm-E and GeneMark-ES. *Curr. Protoc. Bioinform.* **2011**, *35*, 4.6.1–4.6.10. [CrossRef]

29. Meyer, F.; Goesmann, A.; McHardy, A.C.; Bartels, D.; Bekel, T.; Clausen, J.; Kalinowski, J.; Linke, B.; Rupp, O.; Giegerich, R.; et al. GenDB—An open source genome annotation system for prokaryote genomes. *Nucleic Acids Res.* **2003**, *31*, 2187–2195. [CrossRef]

30. Rupp, O.; Becker, J.; Brinkrolf, K.; Timmermann, C.; Borth, N.; Pühler, A.; Noll, T.; Goesmann, A. Construction of a Public CHO Cell Line Transcript Database Using Versatile Bioinformatics Analysis Pipelines. *PLoS ONE* **2014**, *9*, e85568. [CrossRef]

31. Tatusov, R.L.; Fedorova, N.D.; Jackson, J.D.; Jacobs, A.R.; Kiryutin, B.; Koonin, E.V.; Krylov, D.M.; Mazumder, R.; Mekhedov, S.L.; Nikolskaya, A.N.; et al. The COG database: An updated version includes eukaryotes. *BMC Bioinform.* **2003**, *4*, 41. [CrossRef]

32. Kanehisa, M.; Furumichi, M.; Tanabe, M.; Sato, Y.; Morishima, K. KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **2017**, *45*, D353–D361. [CrossRef]

33. Bairoch, A. The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res.* **2000**, *28*, 45–48. [CrossRef]

34. Chan, P.P.; Lin, B.Y.; Mak, A.J.; Lowe, T.M. tRNAscan-SE 2.0: Improved detection and functional classification of transfer RNA genes. *Nucleic Acids Res.* **2021**, *49*, 9077–9096. [CrossRef]

35. Manni, M.; Berkeley, M.R.; Seppey, M.; Simão, F.A.; Zdobnov, E.M. BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. *Mol. Biol. Evol.* **2021**, *38*, 4647–4654. [CrossRef]

36. Dieckmann, M.A.; Beyvers, S.; Nkouamedjo-Fankep, R.C.; Hanel, P.H.G.; Jelonek, L.; Blom, J.; Goesmann, A. EDGAR3.0: Comparative genomics and phylogenomics on a scalable infrastructure. *Nucleic Acids Res.* **2021**, *49*, W185–W192. [CrossRef]

37. Wibberg, D.; Genzel, F.; Verwaaijen, B.; Blom, J.; Rupp, O.; Goesmann, A.; Zrenner, R.; Grosch, R.; Pühler, A.; Schlüter, A. Draft genome sequence of the potato pathogen Rhizoctonia solani AG3-PT isolate Ben3. *Arch. Microbiol.* **2017**, *199*, 1065–1068. [CrossRef]

38. Wibberg, D.; Genzel, F.; Verwaaijen, B.; Blom, J.; Rupp, O.; Goesmann, A.; Zrenner, R.; Grosch, R.; Pühler, A.; Schlüter, A. Genome Analyses of the Less Aggressive *Rhizoctonia solani* AG1-IB Isolates 1/2/21 and O8/2 Compared to the Reference AG1-IB Isolate 7/3/14. *J. Fungi* **2021**, *7*, 832. [CrossRef]

39. Villa, V.D.; Storck, R. Nucleotide Composition of Nuclear and Mitochondrial Deoxyribonucleic Acid of Fungi. *J. Bacteriol.* **1968**, *96*, 184–190. [CrossRef]

40. Ballard, J.W.O.; Whitlock, M.C. The incomplete natural history of mitochondria. *Mol. Ecol.* **2004**, *13*, 729–744. [CrossRef]

41. Seemann, T. Prokka: Rapid Prokaryotic Genome Annotation. *Bioinformatics* **2014**, *30*, 2068–2069. [CrossRef]

42. Altschul, S.F.; Madden, T.L.; Schäffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402. [CrossRef] [PubMed]

43. Eikmeyer, F.; Hadiati, A.; Szczepanowski, R.; Wibberg, D.; Schneiker-Bekel, S.; Rogers, L.M.; Brown, C.J.; Top, E.M.; Pühler, A.; Schlüter, A. The complete genome sequences of four new IncN plasmids from wastewater treatment plant effluent provide new insights into IncN plasmid diversity and evolution. *Plasmid* **2012**, *68*, 13–24. [CrossRef]

44. n.d. N50, L50, and Related Statistics. In Wikipedia. Available online: https://en.wikipedia.org/wiki/N50,_L50,_and_related_statistics (accessed on 30 May 2022).

45. Wibberg, D.; Jelonek, L.; Rupp, O.; Hennig, M.; Eikmeyer, F.; Goesmann, A.; Hartmann, A.; Borriss, R.; Grosch, R.; Pühler, A.; et al. Establishment and interpretation of the genome sequence of the phytopathogenic fungus *Rhizoctonia solani* AG1-IB isolate 7/3/14. *J. Biotechnol.* **2013**, *167*, 142–155. [CrossRef] [PubMed]

46. Hintz, W.; Anderson, J.B.; Horgen, P.A. Nuclear migration and mitochondrial inheritance in the mushroom *Agaricus bitorquis*. *Genetics* **1988**, *119*, 35–41. [CrossRef]

47. Jin, T.; Sonnenberg, A.S.M.; Van Griensven, L.J.L.D.; Horgen, P.A. Investigation of Mitochondrial Transmission in Selected Matings between Homokaryons from Commercial and Wild-Collected Isolates of *Agaricus bisporus* (= *Agaricus brunnescens*). *Appl. Environ. Microbiol.* **1992**, *58*, 3553–3560. [CrossRef] [PubMed]

48. Jin, T.; Horgen, P.A. Uniparental Mitochondrial Transmission in the Cultivated Button Mushroom, *Agaricus bisporus*. *Appl. Environ. Microbiol.* **1994**, *60*, 4456–4460. [CrossRef] [PubMed]

49. Baptista-Ferreira, J.L.C.; Economou, A.; Casselton, L.A. Mitochondrial genetics of Coprinus: Recombination of mitochondrial genomes. *Curr. Genet.* **1983**, *7*, 405–407. [CrossRef] [PubMed]

50. May, G.; Taylor, J.W. Patterns of mating and mitochondrial DNA inheritance in the agaric Basidiomycete *Coprinus cinereus*. *Genetics* **1988**, *118*, 213–220. [CrossRef]

51. Barroso, G.; Labarère, J. Genetic evidence for nonrandom sorting of mitochondria in the basidiomycete *Agrocybe aegerita*. *Appl. Environ. Microbiol.* **1997**, *63*, 4686–4691. [CrossRef]

52. Aanen, D.K.; Kuyper, T.W.; Debets, A.J.M.; Hoekstra, R.F. The evolution of non–reciprocal nuclear exchange in mushrooms as a consequence of genomic conflict. *Proc. R. Soc. B Boil. Sci.* **2004**, *271*, 1235–1241. [CrossRef]

53. Ye, L.-Y.; Deng, Y.-J.; Mukhtar, I.; Meng, G.-L.; Song, Y.-J.; Cheng, B.; Hao, J.-B.; Wu, X.-P. Mitochondrial genome and diverse inheritance patterns in *Pleurotus pulmonarius*. *J. Microbiol.* **2020**, *58*, 142–152. [CrossRef]

54. Wang, P.; Sha, T.; Zhang, Y.; Cao, Y.; Mi, F.; Liu, C.; Yang, D.; Tang, X.; He, X.; Dong, J.; et al. Frequent heteroplasmy and recombination in the mitochondrial genomes of the basidiomycete mushroom *Thelephora ganbajun*. *Sci. Rep.* **2017**, *7*, 1626. [CrossRef]

55. Choi, J.-H.; Suzuki, T.; Ono, A.; Kotajima, M.; Tanaka, Y.; Suzuki, T.; Kawagishi, H.; Dohra, H. The complete mitochondrial genome sequence of the fairy ring-forming fungus *Lepista sordida*. *Mitochondrial DNA Part B* **2022**, *7*, 712–714. [CrossRef]

56. Wibberg, D.; Rupp, O.; Blom, J.; Jelonek, L.; Kröber, M.; Verwaaijen, B.; Goesmann, A.; Albaum, S.; Grosch, R.; Pühler, A.; et al. Development of a *Rhizoctonia solani* AG1-IB Specific Gene Model Enables Comparative Genome Analyses between Phytopathogenic *R. solani* AG1-IA, AG1-IB, AG3 and AG8 Isolates. *PLoS ONE* **2015**, *10*, e0144769. [CrossRef]

57. Yin, Y.; Mao, X.; Yang, J.; Chen, X.; Mao, F.; Xu, Y. dbCAN: A web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **2012**, *40*, W445–W451. [CrossRef]

58. Peng, L.; Shan, X.; Yang, Y.; Wang, Y.; Druzhinina, I.S.; Pan, X.; Jin, W.; He, X.; Wang, X.; Zhang, X.; et al. Facultative symbiosis with a saprotrophic soil fungus promotes potassium uptake in American sweetgum trees. *Plant Cell Environ.* **2021**, *44*, 2793–2809. [CrossRef]

59. Blin, K.; Shaw, S.; Kloosterman, A.M.; Charlop-Powers, Z.; van Wezel, G.P.; Medema, M.H.; Weber, T. antiSMASH 6.0: Improving cluster detection and comparison capabilities. *Nucleic Acids Res.* **2021**, *49*, W29–W35. [CrossRef]

60. Wendt, K.U.; Schulz, G.E. Isoprenoid biosynthesis: Manifold chemistry catalyzed by similar enzymes. *Structure* **1998**, *6*, 127–133. [CrossRef]

61. Dickschat, J.S. Isoprenoids in three-dimensional space: The stereochemistry of terpene biosynthesis. *Nat. Prod. Rep.* **2011**, *28*, 1917–1936. [CrossRef]

62. Chen, H.-P.; Liu, J.-K. Secondary Metabolites from Higher Fungi. In *Progress in the Chemistry of Organic Natural Products*; Springer: Cham, Switzerland, 2017; Volume 106, pp. 1–201. [CrossRef]

63. Koczyk, G.; Pawłowska, J.; Muszewska, A. Terpenoid Biosynthesis Dominates among Secondary Metabolite Clusters in *Mucoromycotina* Genomes. *J. Fungi* **2021**, *7*, 285. [CrossRef]

64. Gressler, M.; Löhr, N.A.; Schäfer, T.; Lawrinowitz, S.; Seibold, P.S.; Hoffmeister, D. Mind the mushroom: Natural product biosynthetic genes and enzymes of Basidiomycota. *Nat. Prod. Rep.* **2021**, *38*, 702–722. [CrossRef]

65. Schmidt-Dannert, C. Biosynthesis of Terpenoid Natural Products in Fungi. *Adv. Biochem. Eng. Biotechnol.* **2015**, *148*, 19–61. [CrossRef]

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

67. Harvey, C.J.B.; Tang, M.; Schlecht, U.; Horecka, J.; Fischer, C.R.; Lin, H.-C.; Li, J.; Naughton, B.; Cherry, J.; Miranda, M.; et al. HEx: A heterologous expression platform for the discovery of fungal natural products. *Sci. Adv.* **2018**, *4*, eaar5459. [CrossRef]

68. Maury, J.; Kannan, S.; Jensen, N.B.; Öberg, F.K.; Kildegaard, K.R.; Forster, J.; Nielsen, J.; Workman, C.T.; Borodina, I. Glucose-Dependent Promoters for Dynamic Regulation of Metabolic Pathways. *Front. Bioeng. Biotechnol.* **2018**, *6*, 63. [CrossRef]

69. Curran, K.A.; Karim, A.S.; Gupta, A.; Alper, H.S. Use of expression-enhancing terminators in Saccharomyces cerevisiae to increase mRNA half-life and improve gene expression control for metabolic engineering applications. *Metab. Eng.* **2013**, *19*, 88–97. [CrossRef]

70. Sakai, A.; Shimizu, Y.; Hishinuma, F. Integration of heterologous genes into the chromosome of Saccharomyces cerevisiae using a delta sequence of yeast retrotransposon Ty. *Appl. Microbiol. Biotechnol.* **1990**, *33*, 302–306. [CrossRef]

71. Botas, A.; Eitel, M.; Schwarz, P.N.; Buchmann, A.; Costales, P.; Núñez, L.E.; Cortés, J.; Morís, F.; Krawiec, M.; Wolański, M.; et al. Genetic Engineering in Combination with Semi-Synthesis Leads to a New Route for Gram-Scale Production of the Immunosuppressive Natural Product Brasilicardin A. *Angew. Chem. Int. Ed.* **2021**, *60*, 13536–13541. [CrossRef]

72. Schwarz, P.N.; Roller, L.; Kulik, A.; Wohlleben, W.; Stegmann, E. Engineering metabolic pathways in *Amycolatopsis japonicum* for the optimization of the precursor supply for heterologous brasilicardin congeners production. *Synth. Syst. Biotechnol.* **2018**, *3*, 56–63. [CrossRef]