

## Article

# Chromosome-Level Genome Assembly and Annotation of the Highly Heterozygous *Phallus echinovolvatus* Provide New Insights into Its Genetics

Mengya An <sup>1,2</sup>, Ruoxi Liang <sup>3</sup>, Yanliu Chen <sup>1</sup>, Jinhua Zhang <sup>1</sup>, Xiuqing Wang <sup>1</sup>, Xing Li <sup>1</sup>, Guohua Qu <sup>1</sup> and Junfeng Liang <sup>1,\*</sup>

<sup>1</sup> Research Institute of Tropical Forestry, Chinese Academy of Forestry, Guangzhou 510520, China; anmy0505@163.com (M.A.); xiaoliuchen0316@126.com (Y.C.); jhzhzhang0111@163.com (J.Z.); clearmoon2008@126.com (X.W.); lixing@scbg.ac.cn (X.L.); qgh0107@163.com (G.Q.)

<sup>2</sup> College of Forestry, Nanjing Forestry University, Nanjing 210037, China

<sup>3</sup> Honors College, Northwestern Polytechnical University, Xi'an 710129, China; liangruoxi@mail.nwpu.edu.cn

\* Correspondence: jfliang2000@163.com

**Abstract:** *Phallus echinovolvatus* is a well-known edible and medicinal fungus with significant economic value. However, the available whole-genome information is lacking for this species. The chromosome-scale reference genome (Monop) and two haploid genomes (Hap1 and Hap2) of *P. echinovolvatus*, each assembled into 11 pseudochromosomes, were constructed using Illumina, PacBio-HiFi long-read sequencing, and Hi-C technology. The Monop had a size of 36.54 Mb, with 10,251 predicted protein-coding genes and including 433 carbohydrate-active enzyme genes, 385 cytochrome P450 enzyme genes, and 42 gene clusters related to secondary metabolite synthesis. Phylogenetic and collinearity analysis revealed a close evolutionary relationship between *P. echinovolvatus* and *Clathrus columnatus* in the core Phallales clade. Hap1 and Hap2 had sizes of 35.46 Mb and 36.11 Mb, respectively. Collinear relationships were not observed for 15.38% of the genes in the two haplotypes. Hap1 had 256 unique genes, and Hap2 had 370 unique genes. Our analysis of the *P. echinovolvatus* genome provides insights into the genetic basis of the mechanisms underlying the metabolic effects of bioactive substances and will aid ongoing breeding efforts and studies of genetic mechanisms.

**Keywords:** *Phallus echinovolvatus*; whole genome; dikaryon; comparative genomics; CAZymes; secondary metabolism; terpene



Academic Editors: Quanju Xiang and Yunfu Gu

Received: 30 November 2024

Revised: 8 January 2025

Accepted: 13 January 2025

Published: 15 January 2025

**Citation:** An, M.; Liang, R.; Chen, Y.; Zhang, J.; Wang, X.; Li, X.; Qu, G.; Liang, J. Chromosome-Level Genome Assembly and Annotation of the Highly Heterozygous *Phallus echinovolvatus* Provide New Insights into Its Genetics. *J. Fungi* **2025**, *11*, 62. <https://doi.org/10.3390/jof11010062>

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

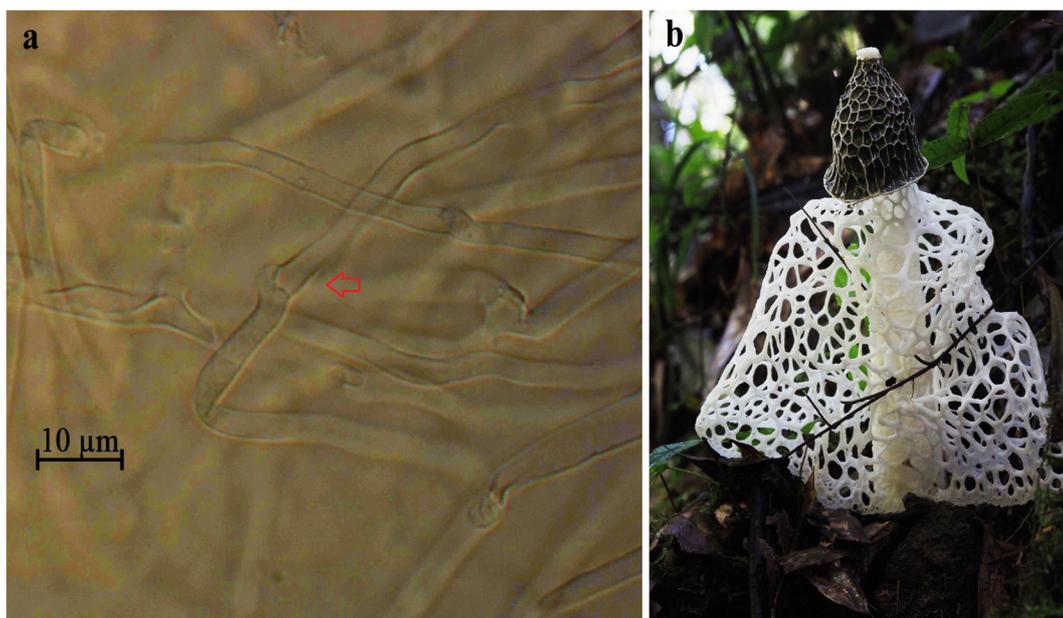
## 1. Introduction

*Phallus echinovolvatus* (M. Zang, D.R. Zheng & Z.X. Hu) Kreisel, originally discovered in Hunan Province, China, in 1988, was initially named *Dictyophora echinovolvata* M. Zang, D.R. Zheng & Z.X. Hu [1]. This fungus is widely distributed in regions south of the Yangtze River in China and in Southeast Asia [2,3]. The genus *Dictyophora* was initially established to classify a group of fungi within the family Phallaceae, characterized by a distinctive net-like “skirt” hanging from the cap. In the early 19th century, these fungi were grouped under the genus *Dictyophora* Desvaux. However, in 1996, Kreisel reduced *Dictyophora* as a section within the genus *Phallus* (P. sect. *Dictyophora* (Desvaux) Kreisel) [4], and the species name was revised to *Phallus echinovolvatus* (M. Zang, D.R. Zheng & Z.X. Hu) Kreisel. Recent molecular phylogenetic studies have shown that the presence or absence of the “skirt” is insufficient to justify the independent taxonomic status of *Dictyophora* [5,6]. These

findings provide strong evidence for the taxonomic adjustment, which has since been widely accepted by mycologists [7,8]. Nevertheless, some studies continue to use the genus name *Dictyophora*.

*P. echinovolvatius* is rich in various nutrients and is widely consumed as a functional food in daily life in countries such as China, Japan, Germany, and North America [9,10]. It also has significant biomedical effects. Previous studies have shown that it plays an important role in the immune system, including in tumor cell inhibition and antibiosis [11–14]. The genome is essential for molecular and genetic research in macrofungi. Few genomic and genetic studies of *P. echinovolvatius* have been conducted compared with other edible mushrooms, such as *Lentinula edodes* [15], *Flammulina velutipes* [16], *Auricularia heimuer* [17], *Morchella importuna* [18], and *Agaricus bisporus* [19]. This research gap has greatly impeded downstream investigations and the utilization of this mushroom. Although the genomes of *Phallus indusiatus* [20] and *Phallus rubrovolvatius* [21,22] have been published, the genome of *P. echinovolvatius* has not yet been sequenced. The molecular basis and evolution of the component biosynthesis in *P. echinovolvatius* are rarely reported due to the lack of a high-quality reference genome.

Microscopic observations have revealed distinctive clamp connections, indicating that the strain (RITF7875) is a dikaryon (Figure 1a). In macrofungi, the presence of two haploid nuclei within a single cell is a widespread phenomenon. This specialized dikaryotic structure poses major challenges to the assembly of macrofungal genomes. Following the emergence of high-fidelity (HiFi) and chromosome conformation capture (Hi-C) sequencing technologies, assembly tools designed for HiFi data, such as Hifiasm and HiCanu, have been used for the assembly of the two haploid genomes of heterozygous organisms [23]. The use of HiFi and Hi-C sequencing strategies has facilitated the phasing and assembly of chromosome-level genomes for several eukaryotic organisms, such as humans [24], *Takifugu ocellatus* [25], diploid *Suaeda glauca* [26], walnut [27], apple [28], *Puccinia triticina* [29], and *Puccinia polysora* [30].



**Figure 1.** Dikaryotic mycelia and fruiting body of *P. echinovolvatius*. (a) Microscopic observation of dikaryotic mycelia of the RITF7875 strain where the clamp connections are indicated by arrows. (b) Cultivated *P. echinovolvatius* under trees in Guangdong Province, China.

We generated a chromosome-level reference genome and two haploid genomes of *P. echinovolvatius* by integrating PacBio, Illumina, and Hi-C sequencing data. The total

length of the reference genome was 43.85 Mb, with a contig N50 of 1.20 Mb. The assembled sequences were anchored to 11 pseudochromosomes with an integration efficiency of 83.58%. The total lengths of Hap1 and Hap2 were 37.68 Mb and 38.46 Mb, with contig N50 lengths of 0.65 Mb and 0.73 Mb, respectively. Approximately 35.46 Mb (94.16%) of Hap1 and 36.11 Mb (93.98%) of Hap2 were assigned to 11 pseudochromosomes. A total of 10,251, 9316, and 9328 protein-coding gene models were predicted for the Monop, Hap1, and Hap2, respectively. In summary, we present novel genomic information about *P. echinovolvatus* through gene annotation and comparative genomic analysis. The genome serves as a direct and comprehensive representation of a species' genetic information, offering insights unclouded by the effects of convergent evolution or morphological plasticity. This clarity makes genomic analysis an invaluable tool for exploring evolutionary relationships, genetic diversity, and functional traits, bypassing the confounding influences of superficial similarities or adaptive morphological changes. The genetic data generated in this study will significantly contribute to the study of the taxonomy and evolutionary relationships within the Phallaceae family, providing a crucial theoretical foundation for future genetic breeding and the development of active compounds in *P. echinovolvatus*.

## 2. Materials and Methods

### 2.1. Fungal Strain and Genome Sequencing

The dikaryotic *P. echinovolvatus* strain (RITF7875) used in this study was derived from the Forest Resources and Protection Laboratory of the Research Institute of Tropical Forestry, Chinese Academy of Forestry (RITF) (Figure 1). The isolated and purified vegetative mycelia were cultured in a liquid PD medium (30% potato, 2% glucose) for 14 days. Mycelium collected through centrifugation was flash-frozen in liquid nitrogen and stored in a refrigerator at  $-80^{\circ}\text{C}$ .

The high-quality genomic DNA from *P. echinovolvatus* (RITF7875) was extracted using an Omega Bio-Tek Fungal DNA extraction kit (E.Z.N.A.<sup>®</sup> Fungal DNA Kit, Omega Bio-Tek, Norcross, GA, USA) according to the manufacturer's instructions. The library was constructed using the Illumina TruSeq<sup>™</sup> Nano DNA Sample Prep Kit (Illumina, San Diego, CA, USA) method (with an insert size of 450 bp). After library construction, paired-end sequencing was performed using the Illumina NovaSeq 6000 platform (Shanghai BIOZERON Co., Ltd., Shanghai, China). To enhance the accuracy of the subsequent assembly, adapter sequences were removed from the raw data using Trimmomatic v0.39 software [31] after quality control by FastQC v0.11.9 [32].

The high-molecular-weight genomic DNA was then sheared to a target size of 15–20 kb, and a SMRTbell library was constructed. The genomic library was sequenced in circular consensus sequencing (CCS) mode on the PacBio Sequel II platform. To obtain a more accurate assembly, the original sequencing data were processed as follows: (1) polymerase reads with a length of less than 200 bp were removed; (2) polymerase reads with a quality score of less than 0.80 were removed; (3) CCS reads were extracted from polymerase reads, and adapter sequences were removed; and (4) CCS reads with a length of less than 200 bp were removed.

For Hi-C sequencing, cells were treated with formaldehyde to crosslink the DNA with the proteins, which preserves the conformation of the DNA. Following cell lysis, the crosslinked DNA was treated with a restriction enzyme to generate sticky ends. After end-polishing and repair, biotin was introduced to label the oligonucleotide ends. Adjacent DNA fragments were ligated using DNA ligase. Protein digestion was performed to release the DNA from the crosslinked state, followed by DNA purification and fragmentation into 500–700 bp fragments. The biotin-labeled DNA was captured using streptavidin magnetic

beads for library construction. After library preparation, Hi-C sequencing was performed using the Illumina NovaSeq 6000 platform.

## 2.2. Genome Assembly and Assessment

Before assembly, k-mer analysis was used to estimate various genome characteristics. The Illumina sequencing data were analyzed using 21-mers to estimate the genome size, heterozygosity, and repeatability.

The software Hifiasm (<https://github.com/chhylp123/hifiasm>, accessed on 5 November 2023) was used to assemble the HiFi reads [33]. The ALLHiC tool (<https://github.com/tanghaibao/allhic>, accessed on 10 November 2023) was used to link the genome contigs or scaffolds [34]. Hi-C-assisted assembly is based on the principle that cis-interactions are stronger than trans-interactions, and the strength of cis-interactions increases at closer distances. Therefore, the contigs or scaffolds were clustered, sorted, and oriented to construct the Hi-C-based chromosome-level assembly, followed by heatmap visualization to display the chromosome contacts [35,36]. BUSCO v4.1.2 [37] was used to assess the completeness and accuracy of the *P. echinovolvatius* genome.

## 2.3. Component Prediction

We used AUGUSTUS v3.2.3 [38] for the de novo gene prediction, and the alignments were performed using GeneWise v2.4 [39] to identify gene-coding regions and intron regions. Subsequently, EVIDENCEModeler v1.1.1 [40] software was used to integrate the aforementioned results and predict all the protein-coding genes. RepeatMask (<http://www.repeatmasker.org>, accessed on 5 January 2024) [41] was used to identify interspersed repeats (IRs) by aligning the sequences against a known repetitive sequence database. TRNAscan-SE v2.0.7 [42] was used to predict the transfer RNA (tRNA) regions and the secondary structures of tRNAs. Ribosomal RNA (rRNA) predictions were conducted using rRNAmmer v1.2 [43]. The Rfam v13.0 [44] was used to make comparisons against the RNA families (Rfam) database and acquire annotations, and its integrated search tool (with the default parameters) was used to identify small RNAs (sRNAs), small nuclear RNAs (snRNAs), and microRNAs (miRNAs).

## 2.4. Gene Annotation

The predicted genes were subjected to BLAST alignment against the Non-Redundant (NR), the Kyoto Encyclopedia of Genes and Genomes (KEGG), the Evolutionary Genealogy of Genes: Non-supervised Orthologous Groups (eggNOG), the Gene Ontology (GO), and the Swiss-Prot databases to predict gene functions. PfamScan v1.6, provided by the Pfam database, was used to identify structural domains. TMHMM v2.0 was used to predict transmembrane domain structures. To classify members of the cytochrome P450 gene family in *P. echinovolvatius*, BLAST v 2.10.1 was used to align all the protein sequences of *P. echinovolvatius* against the Fungal Cytochrome P450 Database [45]. The predicted results were named according to the Cytochrome P450 Database (<https://cyped.biocatnet.de/>, accessed on 5 June 2024). The e-value cutoff was set to be less than or equal to  $1 \times 10^{-5}$ .

## 2.5. Comparative Genomics Analysis

The genome and protein sequences of other macrofungi were downloaded from the National Center for Biotechnology Information (<https://www.ncbi.nlm.nih.gov/>, accessed on 4 June 2024) (Table S1). Carbohydrate-active enzymes (CAZymes) were predicted using dbCAN3 [46]. The online tool antiSMASH7.0 [47] was used to predict genes potentially related to secondary metabolites. OrthoFinder (v2.5.4) [48] software was used to construct the orthologous groups of 19 macrofungi. Single-copy orthologous genes were obtained through screening and used to construct the phylogenetic tree to reveal the phylogenetic

relationships among *P. echinovolvatus* and other related taxa. The single-copy orthologs were further aligned using MUSCLE (v3.8.31) [49], and then the conserved region was extracted with the aid of Gblocks 0.91b [50] to obtain the codon-type nucleic acid comparison, which was concatenated for phylogenetic tree construction. Maximum likelihood (ML) estimation of the model parameters was performed using jModeltest v2.1.10 software [51] and GTR+I+G was the optimal model. PhyMLv3.0 [52] was used to construct the ML phylogenetic tree. Collinearity analysis was performed using TBtools-II [53]. The results were visualized using the ChiPlot Cloud Platform (<https://www.chiplot.online/>, accessed on 5 June 2024). The gene sequences of Hap1 and Hap2 were extracted from the Hap1 and Hap2 genome sequences and compared with each other using blastn (e-value  $< 1 \times 10^{-5}$ ). Genes without aligned partners between Hap1 and Hap2 were considered unique to either Hap1 or Hap2.

### 3. Results and Discussions

#### 3.1. Genome Sequencing and Assembly

We used the Illumina and PacBio sequencing platforms to sequence the genome of *P. echinovolvatus*. Illumina sequencing generated 7.04 Gb of clean data. Hi-C sequencing yielded 6.91 Gb of clean reads. Additionally, PacBio sequencing provided 24.37 Gb of HiFi reads (Table S2).

The 7.04 Gb of clean data obtained through Illumina sequencing were used for the k-mer analysis to generate a histogram of the depth distribution of the sequencing data ( $k = 21$ ) (Figure S1). The estimated genome size of *P. echinovolvatus* was 37.8 Mb, and the estimated heterozygosity was approximately 1.69%. Notably, two major peaks were observed in the graph, which were positioned between 50 and 150, and the first peak was higher than the second. This demonstrates that RITF7875 is a dikaryotic strain.

Based on the HiFi reads, we initially assembled two haploid genomes, Hap1 and Hap2, along with the reference genome, Monop. The Hap1, Hap2, and Monop genomes comprised 183, 122, and 76 contigs, and their sizes were 37.68 Mb, 38.46 Mb, and 43.85 Mb, respectively. The N50 was 0.65 Mb, 0.73 Mb, and 1.20 Mb for the Hap1, Hap2, and Monop genomes, and the guanine-cytosine (GC) content percentages were 43.96%, 43.97%, and 43.98%, respectively (Table 1). The contigs or scaffolds were clustered, sorted, and oriented using ALLHiC software to construct the chromosome-level genome, and interaction mapping was used to correct inconsistencies. Finally, three genomes with 11 chromosomes each were generated, with sizes of 35.46 Mb (Hap1), 36.11 Mb (Hap2), and 36.54 Mb (Monop), and scaffold N50 lengths of 3.31 Mb, 3.33 Mb, and 3.43 Mb, respectively (Table 1). Based on the interaction maps of the Hap1, Hap2, and Monop genomes, 11 distinct blocks were observed, which corresponded to 11 chromosomes. The strong correlation between the chromosome-level genome assembly and the Hi-C interaction maps indicates that this assembly was highly reliable (Figure S2).

Compared to the other two species in the *Phallus* genus, the genomes of the three species show significant differences. *P. indusiatus* has the largest genome size (67.32 Mb), while *P. rubrovolvatus* has the smallest (32.89 Mb). The proportion of the coding gene length to the total genome length is highest in *P. rubrovolvatus* (63.1%) and lowest in *P. echinovolvatus* (37.09%). The average coding gene length in *P. rubrovolvatus* is much greater than that in *P. echinovolvatus* and *P. indusiatus*, whereas *P. indusiatus* has a much higher gene count than *P. rubrovolvatus* and *P. echinovolvatus* (Table 1). These differences may be due to species-specific characteristics or variations in sequencing technologies; further investigation is needed to clarify the underlying causes.

**Table 1.** Genome assembly features of *P. echinovolvatus*, *P. rubrovolvatus* and *P. indusiatus*.

Contents	Monop	Hap1	Hap2	<i>P. rubrovolvatus</i>	<i>P. indusiatus</i>
Sequencing Technology	Illumina + Hi-C + PacBio HiFi			Illumine + ONT	
Number of Scaffolds	76	183	122	132	216
Total Length (Mb)	43.85	37.68	38.46	32.89	67.32
Scaffold N50 (Mp)	1.20	0.65	0.73	2.7	0.79
Guanine-Cytosine Content (%)	43.98	43.96	43.97	45.16	44.05
Gene Number	10,251	9316	9328	9725	19,909
Gene Total Length (Mb)	16.21	14.83	14.78	20.76	31.87
Gene Average Length (bp)	1582	1592	1585	2135.07	1601
Gene Length/Genome (%)	37.09	39.38	38.47	63.1	47.25
Pseudo-Chromosome	11	11	11	11	–
Pseudo-Chromosome Total Length (Mb)	36.54	36.11	35.46	28.24	–
Pseudo-Chromosome N50 Length (Mb)	3.42	3.38	3.36	2.75	–
Chromosome Anchoring Rate for Contigs (%)	83.58	94.16	93.98	85.57	–

Furthermore, the Benchmarking Universal Single-Copy Orthologs (BUSCOs, basidiomycota\_odb11) were used to assess the assembly quality. A total of 1764 BUSCOs were identified in the genome assembly, and the completeness rates of the Hap1, Hap2 and Monop genomes were 95.6%, 95.8% and 96.7%, respectively (Table 2). These results indicate that the *P. echinovolvatus* genome sequence had high completeness and contiguity.

**Table 2.** Statistics from the BUSCO analysis of the *P. echinovolvatus* genome.

Term	Hap1		Hap2		Monop	
	BUSCO Number	Percentage	BUSCO Number	Percentage	BUSCO Number	Percentage
Complete BUSCOs	1687	95.6	1691	95.8	1706	96.7
Complete and single-copy BUSCOs	1553	88	1579	89.5	1460	82.8
Complete and duplicated BUSCOs	134	7.6	112	6.3	246	13.9
Fragmented BUSCOs	13	0.7	12	0.7	13	0.7
Missing BUSCOs	64	3.7	61	3.5	45	2.6
Total BUSCO groups searched	1764	1.0	1764	1.0	1764	1.0

### 3.2. Genome Component of *P. echinovolvatus*

#### 3.2.1. Gene Prediction

In the reference genome (Monop), a total of 10,251 coding genes with an average length of 1582 bp were predicted. The cumulative length of these genes amounted to 16.21 Mb, accounting for 37.09% of the genome. The average numbers of exons and introns per gene were 7.51 and 6.51, respectively. The two haploid genomes predicted 9316 and 9328 coding genes, which represented 39.38% and 38.47% of the genome, respectively. The average counts of exons and introns per gene were 7.54 and 6.54, respectively (Table S3).

#### 3.2.2. Repeat Sequences Prediction

Repetitive sequences are widespread in eukaryotic genomes and are considered ancient components of the genome. They can be divided into two categories: interspersed repeats (IRs) and tandem repeats (TRs) [54]. TRs include minisatellite DNA, microsatellite DNA, and satellite DNA. Discrete repeats include retrotransposons and DNA transposons [55]. Repeat sequences of *P. echinovolvatus* were identified using both homology-based and de novo strategies. Repeat sequences accounted for 17.73%, 14.45%, and 15.84% of the Monop, Hap1, and Hap2 genomes, respectively. Most IRs were long terminal repeats, long interspersed nuclear elements, and DNA transposable elements. Most TRs

comprised minisatellite DNA and microsatellite DNA (Table S4). Repetitive sequences play an important role in maintaining the spatial structure of chromosomes, gene expression, genetic regulation, and biological evolution. Accurately understanding the structure of the repetitive sequences in the genome contributes to genome research [56].

### 3.2.3. Non-Coding RNA Prediction

The predictions of non-coding RNAs in the *P. echinovolvatus* genome are shown in Table S5. A total of 107 tRNAs, five rRNAs, and 14 snRNAs were predicted in the reference genome (Monop). Among the tRNAs, 41 were unidentified anticodons, and the remaining anticodon tRNAs comprised 17 common amino acid codons. In the Hap1 and Hap2 genomes, 103 and 94 tRNAs, four and five rRNAs, and 12 and 14 snRNAs were predicted, respectively.

### 3.3. Annotation of Gene Function

Functional analysis of the predicted gene sequences was conducted using multiple public databases (NR, GO, KOG, KEGG, P450, Swiss-Prot, HMHMM, Pfam). This analysis resulted in the identification of 8851, 8106, and 8080 annotated genes for the Monop, Hap1, and Hap2 genomes, respectively (Table S6). Based on the gene annotation results, we conducted further analysis of the reference genome (Monop).

#### 3.3.1. GO Annotation

A total of 3402 genes were annotated in the GO database. The GO terms in the Biological Process category were the most common and included “cellular process” (3195 genes), “metabolic process” (2516 genes), and “biological regulation” (1427 genes); within the Cellular Component category, the GO terms “cellular anatomical entity” (3191 genes), “intracellular” (3135 genes), and “protein-containing complex” (1613 genes) were detected; and within the Molecular Function category, the GO terms “catalytic activity” (1659 genes) and “binding” (1576 genes) were detected (Figure 2).

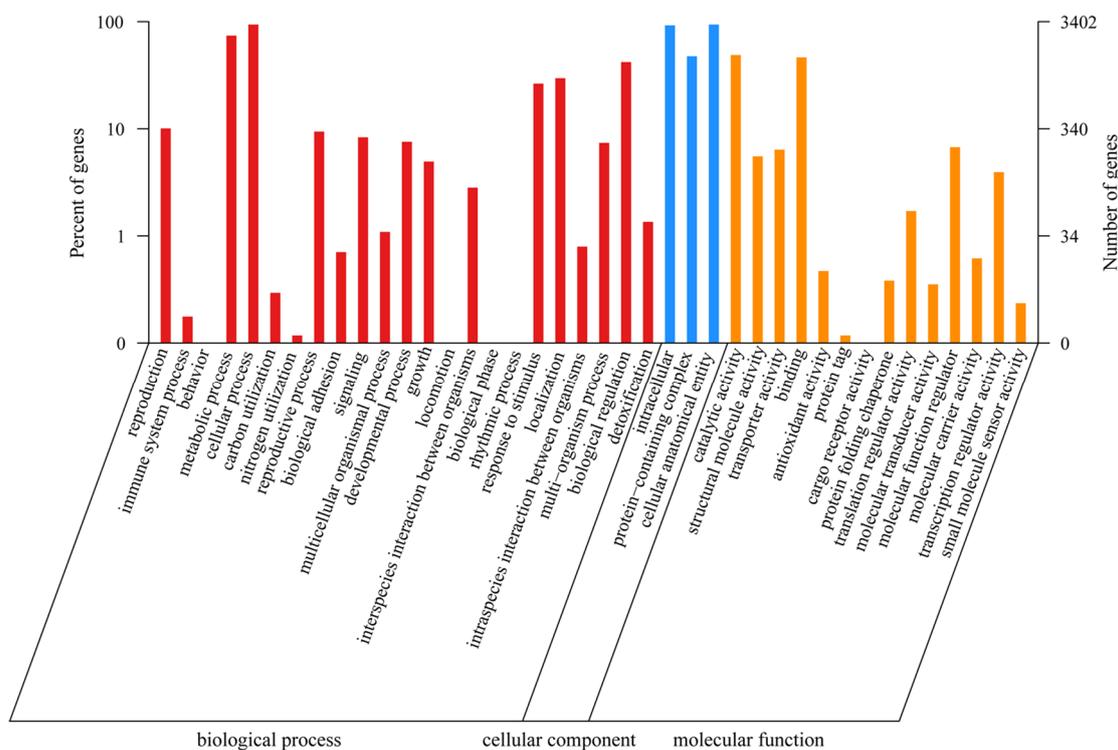
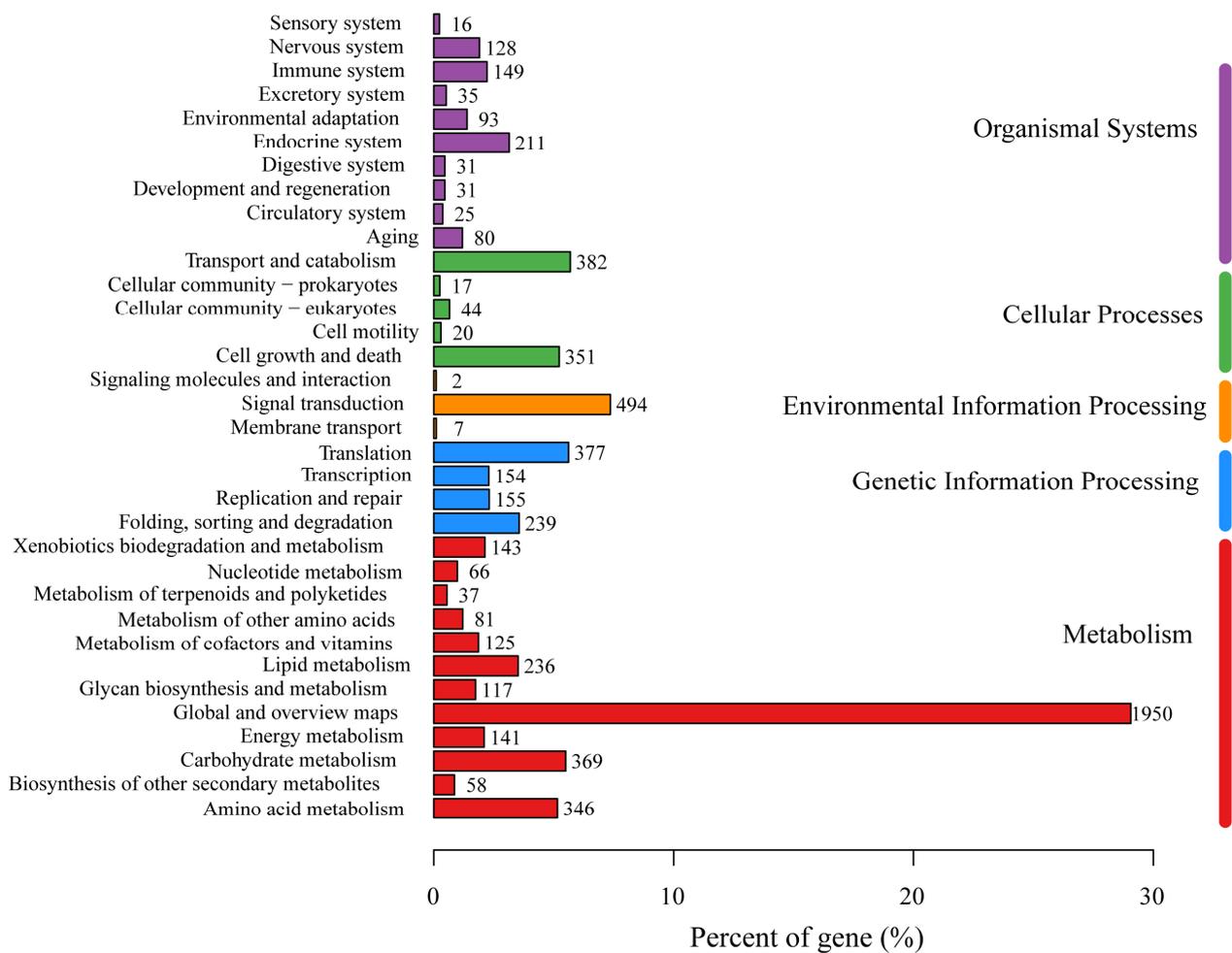


Figure 2. The GO function annotation of *P. echinovolvatus*.

### 3.3.2. KEGG Annotation

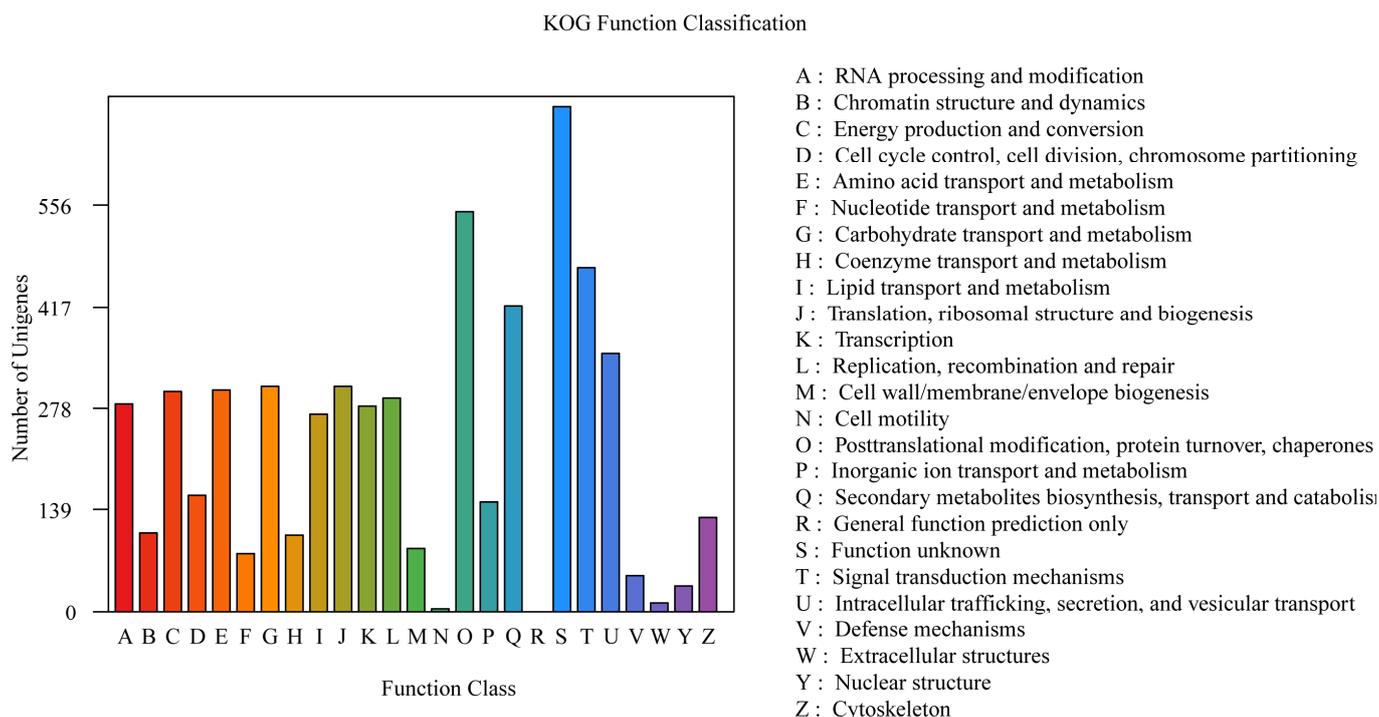
Annotations from the KEGG database were obtained for 3566 genes, which accounted for 34.79% of all the genes. These genes were classified into six major categories: Metabolism (11 branches, 1719 genes), Genetic Information Processing (4 branches, 925 genes), Environmental Information Processing (3 branches, 503 genes), Cellular Processes (5 branches, 814 genes), Organismal Systems (10 branches, 799 genes), and Human Diseases (12 branches, 1276 genes). Within the Metabolism category, the 1719 genes were further classified into 11 subcategories, primarily “Carbohydrate metabolism” (369 genes), “Amino acid metabolism” (346 genes), “Lipid metabolism” (236 genes), “Energy metabolism” (141 genes), “Xenobiotics biodegradation and metabolism” (143 genes), “Metabolism of cofactors and vitamins” (125 genes), and “Glycan biosynthesis and metabolism” (117 genes) (Figure 3).



**Figure 3.** The KEGG function annotation of *P. echinovolvatius*.

### 3.3.3. KOG Annotation

The annotations derived from the KOG database are shown in Figure 3. A total of 5657 genes were annotated in the KOG database, which accounted for 55.18% of all the genes. The results revealed that the genes were enriched in various metabolic processes, including “Posttranslational modification, protein turnover, chaperones” (547 genes), “Signal transduction mechanisms” (471 genes), “Secondary metabolite biosynthesis, transport and catabolism” (419 genes), “Intracellular trafficking, secretion, and vesicular transport” (353 genes), “Carbohydrate transport and metabolism” (308 genes), “Translation, ribosomal structure and biogenesis” (308 genes), “Amino acid transport and metabolism” (303 genes), and “Energy production and conversion” (301 genes) (Figure 4).



**Figure 4.** The KOG function annotation of *P. echinovolvatius*.

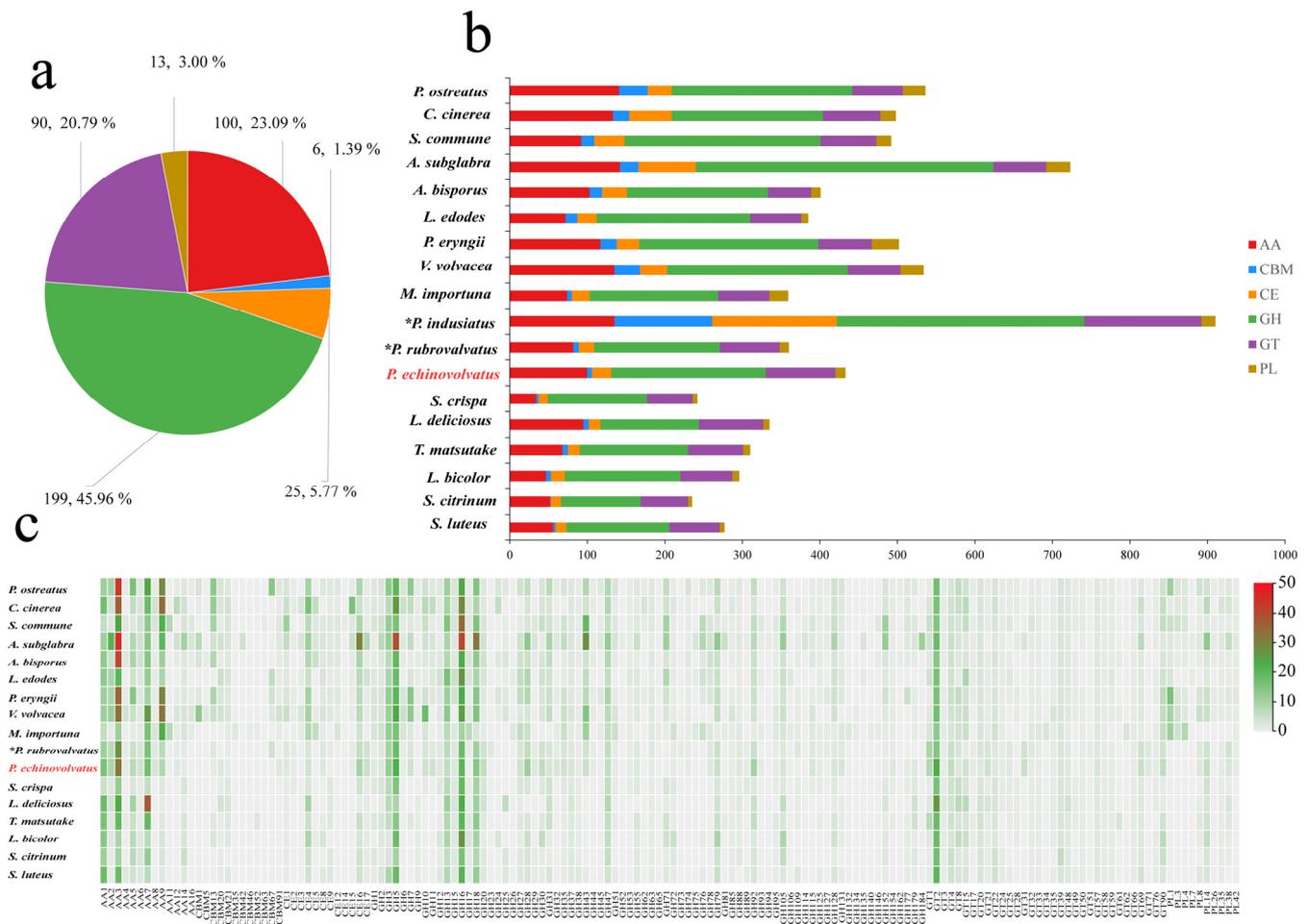
### 3.3.4. The Cytochromes P450 (CYPs) Family

Cytochrome P450 enzymes (CYPs) are terminal oxidases and the major catalysts involved in the metabolism of drugs involved in detoxification, the degradation of xenobiotics, and the biosynthesis of secondary metabolites [57,58]. *P. echinovolvatius* had a total of 385 CYP genes in 30 families according to Nelson's nomenclature [59]. The CYP51 family contained the greatest number of genes (116 genes), followed by the CYP53 family (42 genes), CYP620 family (28 genes), CYP504 family (28 genes), CYP512 family (23 genes), CYP125 family (22 genes), CYP505 family (20 genes), and CYP78 family (18 genes) (Table S7). *P. echinovolvatius* had a few P450 genes involved in KEGG pathways. The results of the gene function annotation revealed that 10 and 11 genes were involved in "Metabolism of xenobiotics by cytochrome P450" and "Drug metabolism cytochrome P450", respectively.

## 3.4. Comparative Genomics Analysis

### 3.4.1. CAZymes

Complex carbohydrates are widely distributed in nature, and they play numerous biological roles within organisms, including serving as structural molecules, sources of energy, and mediators of cell recognition within the same organism or between different organisms. Carbohydrate-active enzymes are involved in the assembly and breakdown of these complex carbohydrates [60]. A total of 433 CAZyme genes were identified from the genome of *P. echinovolvatius*, and these were distributed across 103 CAZyme families, including 199 glycoside hydrolases (GHs, 49 families, 45.96%), 90 glycosyl transferases (GTs, 27 families, 20.79%), 13 polysaccharide lyases (PLs, 5 families, 3.00%), 25 carbohydrate esterases (CEs, 7 families, 5.77%), 100 auxiliary activity enzymes (AAs, 11 families, 23.09%), and six carbohydrate-binding modules (CBMs, 4 families, 1.39%) (Figure 5a and Table S8).



**Figure 5.** CAZymes in *P. echinovolvatus* and other fungi. (a) The distribution of CAZymes in *P. echinovolvatus*. (b) The distribution of CAZymes in another 23 fungi. (c) Heatmap representing the CAZyme families distributed in *P. echinovolvatus* and other fungi. GH, glycoside hydrolase; GT, glycosyltransferase; PL, polysaccharide lyase; CE, carbohydrate esterase; CBM, carbohydrate-binding module; AA, auxiliary activity. \* Referenced [20,21]. The red font is the target species sequenced in this paper.

Fungi play a key role in the degradation of plant biomass. Symbiotic fungi can obtain the nutrients needed for growth and development by establishing a symbiotic relationship with the host, and saprophytic fungi mainly obtain nutrients by decomposing humus. Under natural conditions, *P. echinovolvatus* usually grows in the decayed litter layer of bamboo forest. The number of genes encoding carbohydrate enzymes predicted in the genome of *P. echinovolvatus* was more than that in symbiotic fungi but less than that in most saprophytic fungi. The genome of *P. echinovolvatus* contained genes that encode enzymes required for cellulose degradation, such as GH1, GH3, GH5, GH6, GH7, GH12, GH74, GH92, and AA9, as well as hemicellulases, including GH10, GH11, GH16, GH30, GH115, CE1, and CE2; ligninases, such as AA1, AA2, AA3, AA7, and AA5; and pectinases, such as GH2, GH28, GH35, GH43, GH51, GH53, GH78, GH88, GH93, GH105, and CE8. Additionally, in the *P. echinovolvatus* genome, there were six genes related to CBMs, including four families; most were related to the CBM91 and CBM20 families. Many biomass-degrading fungi commonly employ CBMs for plant cell wall degradation [61]. In the case of *P. echinovolvatus*, CBMs form chimeric enzymes with CAZymes from several families to enhance the hydrolysis of carbohydrates, as well as insoluble substrates such as cellulose, chitin, and starch. We compared the numbers of enzymes in different families

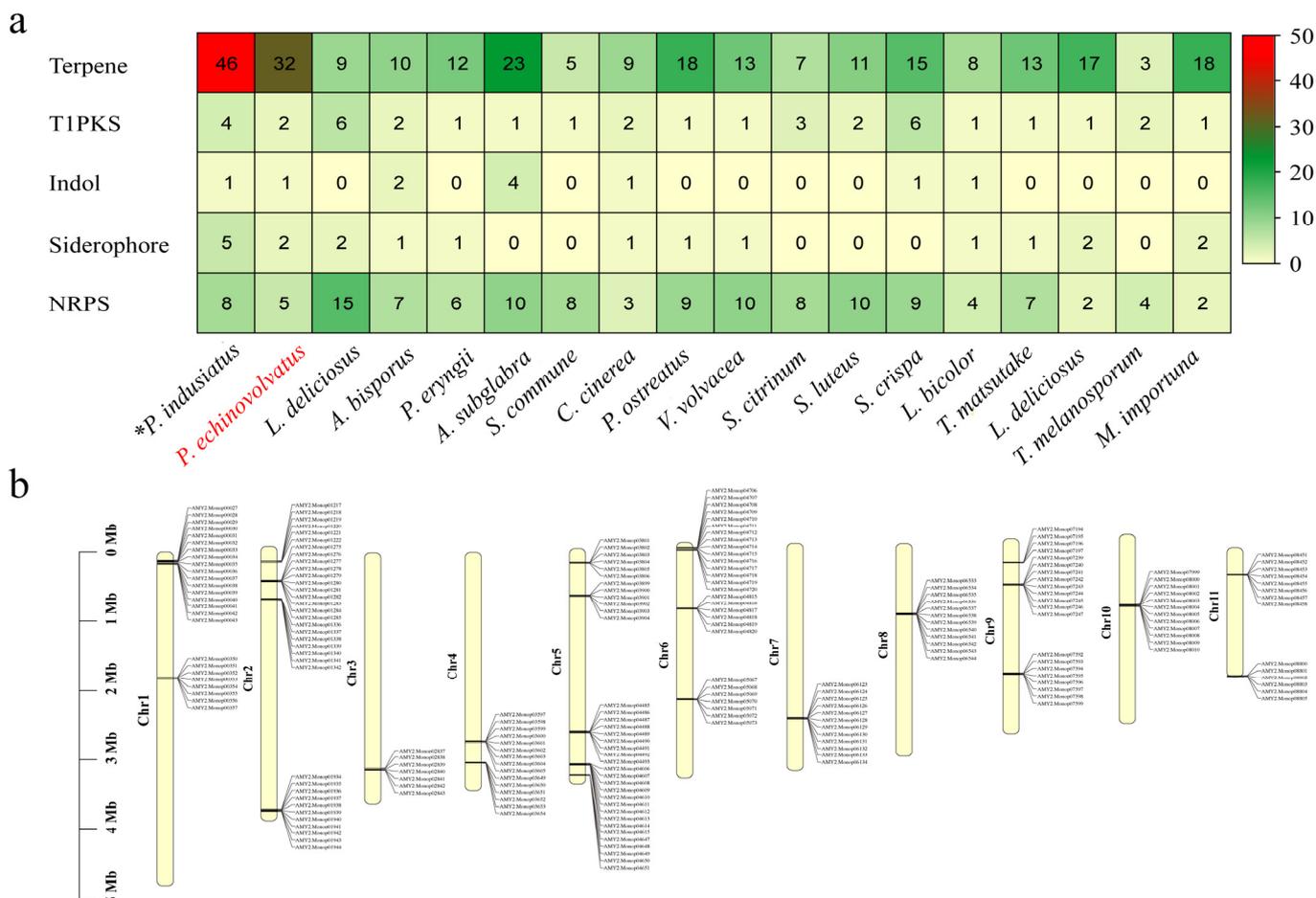
in *P. echinovolvatatus* with those in other mushrooms and found that the numbers of AA9, GH5, GH43, CBM13, and CE4 were significantly lower in *P. echinovolvatatus* than in most saprophytic fungi. Both *P. echinovolvatatus* and *P. rubrovolvatatus* lacked genes encoding PL1, PL3, and PL4 for pectin degradation. These results suggest that *P. echinovolvatatus* might have relatively weak plant cell-wall-degrading abilities, which might contribute to its longer growth cycle (Figure 5 and Table S8).

Fungi rely on CAZymes to degrade complex polysaccharides such as cellulose, hemicellulose, and lignin, which allows them to access and utilize diverse carbon sources present in their environment. The CAZyme repertoire of a fungus reflects its ability to colonize specific substrates. Interestingly, *P. indusiatus* has a significantly higher number of CAZyme genes compared to *P. echinovolvatatus* and *P. rubrovolvatatus*—2.10 and 2.53 times higher, respectively, highlighting the significant expansion of the CAZyme repertoire in *P. indusiatus* compared to the other two species. This variation in the CAZyme gene numbers appears to be associated with differences in the genome size, suggesting that genome expansion may play a role in the diversification of CAZyme repertoires among these species. The extensive CAZyme repertoire of *P. indusiatus* suggests an ecological adaptation to niches that demand a broader range of enzymatic activities, potentially targeting complex substrates like bamboo or other lignocellulosic materials. In contrast, *P. echinovolvatatus* and *P. rubrovolvatatus* may specialize in simpler substrates, necessitating fewer CAZymes. The larger number of CAZyme genes likely provides *P. indusiatus* with greater metabolic flexibility, allowing it to adapt to diverse or challenging environments. While *P. indusiatus* stands out for its extensive CAZyme gene set, the specific evolutionary and ecological drivers of this disparity remain unclear and warrant further investigation, particularly in the context of substrate availability and habitat specialization.

#### 3.4.2. Secondary Metabolites

During the growth of edible fungi, a wide variety of secondary metabolites are produced, and these serve as important sources of bioactive substances in edible fungi. These bioactive substances include amino acids, polysaccharides, terpenes, vitamins, and other active compounds. The secondary metabolites often exhibit unique biological activities. For example, they have various functions, such as reducing blood lipid levels, inhibiting tumor growth, enhancing immune responses, and regulating metabolism. Consequently, these secondary metabolites have significant implications for modern pharmaceutical development [62,63]. The gene clusters involved in the biosynthesis of secondary metabolites in the *P. echinovolvatatus* genome were predicted using antiSMASH. As shown in Figure 6a and Table S9, a total of 42 predicted secondary metabolite gene clusters were identified, including 32 terpene gene clusters, two T1PKS gene clusters, two siderophore gene clusters, one indole gene cluster, and five NRPS gene clusters. T1PKSs are involved in the biosynthesis of complex polyketides, which include many clinically important compounds, such as antibiotics, antifungals, and anticancer agents [64]. The relatively low number of T1PKS clusters in *P. echinovolvatatus* compared to other fungi suggests limited production of complex polyketides, but the presence of these clusters indicates potential for unique bioactive compounds. Similarly, indole clusters are sparse, with most species having zero to four clusters. *P. echinovolvatatus* has one cluster, indicating limited indole metabolite biosynthesis compared to other fungi. Siderophores are iron-chelating compounds that facilitate iron uptake in iron-limited environments. Siderophore clusters are fairly evenly distributed, ranging from zero to five clusters [65]. *P. echinovolvatatus* has two clusters, indicating its potential for iron-chelating activities, which are vital for fungal growth and survival. Additionally, non-ribosomal peptide synthetases (NRPSs) are involved in the biosynthesis of peptides with various biological activities, including antibiotics, immuno-

suppressants, and siderophores [66,67]. The number of NRPS gene clusters is often closely related to the ecological adaptability and diversity of secondary metabolites. A total of five NRPS gene clusters were predicted in the genome of *P. echinovolvat*, indicating a lower count compared to certain other fungi. This suggests that *P. echinovolvat* may rely more on other secondary metabolic pathways, such as terpenoid biosynthesis, for environmental adaptation.



**Figure 6.** Secondary metabolite prediction results for *P. echinovolvat*. (a) Comparison of the secondary metabolites of *P. echinovolvat* with those of other fungi. (b) The distribution of gene clusters encoding sesquiterpenes on chromosomes. NRPS, non-ribosomal peptides; PKS, polyketides. \* Referenced [20]. The red font is the target species sequenced in this paper.

It is worth noting that terpene-related gene clusters account for the largest proportion of these clusters, which is lower than that of *P. indusiatus* (46 clusters) but significantly higher than in other macrofungi, indicating a much greater potential for synthesizing terpene compounds compared to other macrofungi. Mapping the terpene gene clusters to the chromosomes revealed their distribution across all the chromosomes (Figure 6b and Table S9).

Terpenoid compounds are a class of natural organic compounds formed by the polymerization of isoprene units (C<sub>5</sub>H<sub>8</sub>). They are important components of secondary metabolites, widely found in microorganisms. They have important physiological activities and biological functions, and they are particularly important in drug development and growth regulation. For example, guanacastane-type diterpenoids with antitumor activity have been isolated from *Coprinus radians* [68]. The triterpene acid extract from *Ganoderma lucidum* can inhibit the activity of human hepatoma SMMC-7721 and human colon carcinoma SW620

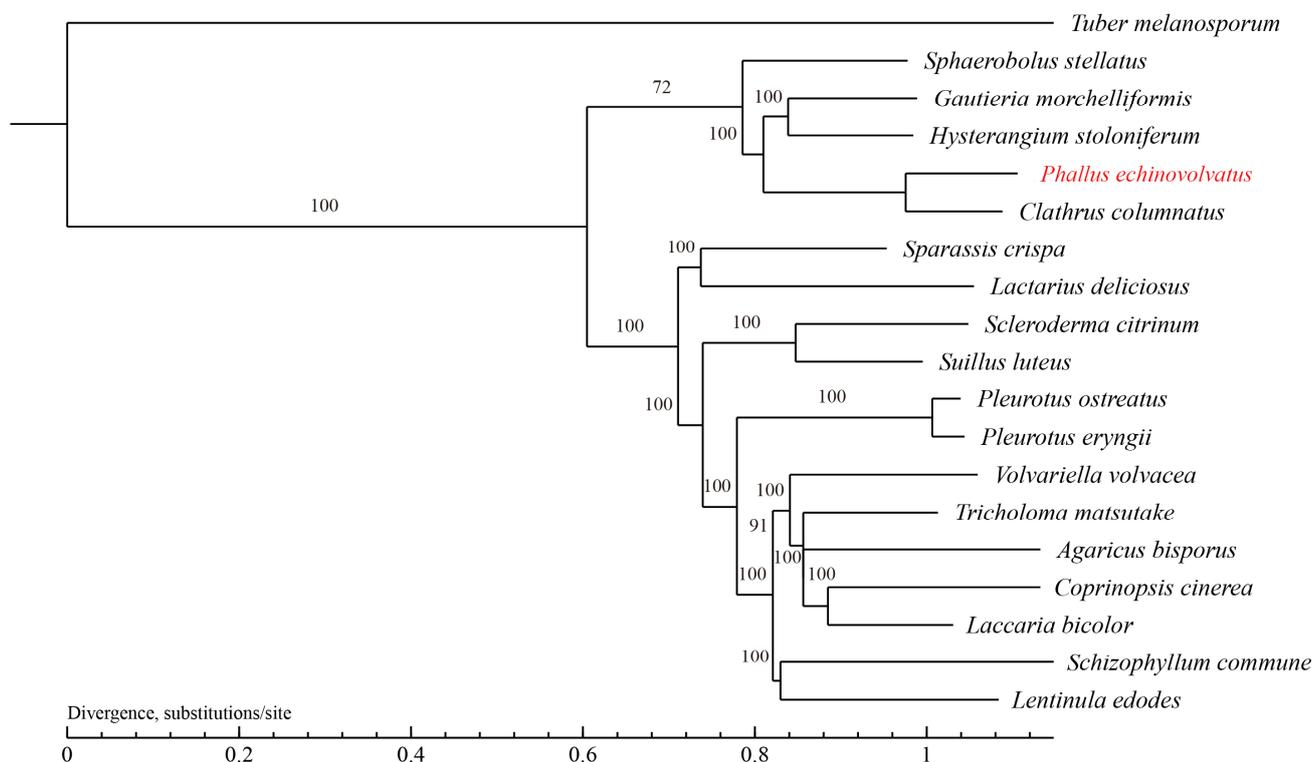
cells [69]. Terpenoid compounds perform various crucial functions within organisms. They act as volatile substances to deter pests and pathogens, attract pollinators or seed dispersers, and help microorganisms adapt to environmental stresses [70–72]. Previous studies have indicated that one of the primary constituents of the volatile aroma of bamboo fungi is volatile terpenoids, such as geranylacetone,  $\beta$ -patchoulene, limonene, alpha-chamigrene, cedrene,  $\gamma$ -selinene, and cedrol [73,74]. Plants utilize scent to attract animals to aid in seed dispersal and reproduction. Possibly similar to the reproductive strategy of plants, when *P. echinovolvat* matures, the surface of the cap secretes mucilage containing numerous sexual spores and emits a strong odor, which might attract more insects to facilitate spore dissemination. The genomes of both *P. indusiatus* and *P. echinovolvat* are significantly enriched with gene clusters associated with terpene biosynthesis; this likely stems from long-term adaptation to the environment. Isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) are common precursors for the biosynthesis of terpenes. IPP and DMAPP are primarily derived from the mevalonate (MVA) and methylerythritol phosphate (MEP) pathways [75,76]. Based on the KEGG annotation results, there were 15 enzymes involved in “terpenoid backbone biosynthesis (map00900)” in *P. echinovolvat*, where acetyl-CoA C-acetyltransferase, isopentenyl-diphosphate delta-isomerase, and phosphomevalonate kinase were encoded by two genes each, and the remaining 12 enzymes were encoded by a single gene. Similar to other basidiomycetes, the MVA pathway was detected in *P. echinovolvat*; however, the MEP/DOXP pathway was not detected (Figure S3 and Table S10). There were two enzymes involved in “Sesquiterpenoid and triterpenoid biosynthesis (map00909)”, where farnesyl-diphosphate farnesyltransferase was encoded by three genes and squalene monooxygenase was encoded by one gene. Furthermore, we identified two genes encoding lanosterol synthase (LSS) (AMY2.Monop07530.1, AMY2.Monop07562.1 [EC:5.4.99.7]); one enzyme was involved in “Diterpenoid biosynthesis (map00904)”, six enzymes were involved in “Ubiquinone and other terpenoid-quinone biosynthesis (map00130)”, NAD(P)H dehydrogenase (quinone) was encoded by four genes, and the remaining five enzymes were encoded by a single gene. However, no enzymes associated with the “monoterpenoid biosynthesis (map00902)” pathway were identified.

The diversity of the secondary metabolite gene clusters underscores the variability in the biosynthesis of secondary metabolites across different species. A comparison with other macrofungi revealed that *P. echinovolvat* exhibits significant potential for terpene biosynthesis, making it a promising source of diverse terpenoids that are not only ecologically significant but also possess substantial biological and pharmaceutical value. Its rich terpenoid diversity positions it as a valuable candidate for natural product discovery, particularly in drug development and other biotechnological applications. To fully realize this potential, future studies should focus on elucidating the terpene biosynthetic pathways of *P. echinovolvat* through transcriptomic, proteomic, and metabolomic, combined with bioinformatics analysis and experimental validation. This approach will enable a deeper understanding of the specific functions of these gene clusters and exploration of their application potential.

### 3.4.3. Phylogenetic Analysis

The evolutionary relationships between *P. echinovolvat* and 18 other fungi (*T. melanosporum* from Ascomycota as the outgroup) were investigated. A total of 629 single-copy orthologous gene families were found and used to construct the phylogenetic tree, and 844 genes were unique to *P. echinovolvat* (Table S11). As shown in Figure 7, *P. echinovolvat* was nested within a large clade formed by five species under Phallales and clustered with *C. columnatus*, indicating that *P. echinovolvat* is closely related to *C. columnatus*. *P. echinovolvat* and *C. columnatus* both belong to the family Phallaceae, and the

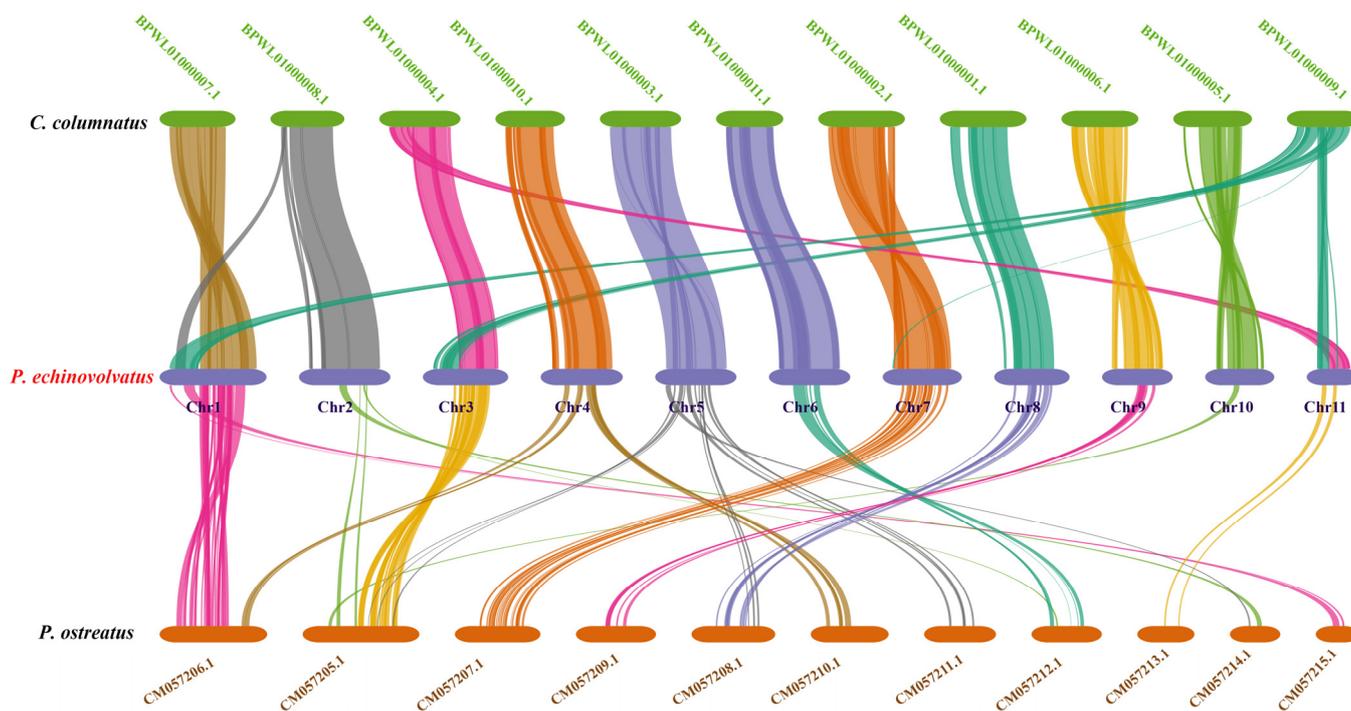
results of the phylogenetic analysis were consistent with the classification results based on the morphological and molecular traits.



**Figure 7.** Phylogenetic tree of 19 fungi based on single-copy orthologous genes. *T. melanosporum* served as the outgroup. The red font is the target species sequenced in this paper.

#### 3.4.4. Collinearity Analysis

Based on the phylogenetic analysis, we conducted a genome collinearity analysis using *C. columnatus*, which is closely related to *P. echinovolvatius* (its genome comprises 11 long contigs and eight short contigs), and *P. ostreatus*, which is more distantly related (its genome comprises 11 chromosomes). We observed pronounced collinearity between *P. echinovolvatius* and *C. columnatus*, and the collinearity between *P. echinovolvatius* and *P. ostreatus* was not significant. The results indicated the high reliability of the *P. echinovolvatius* genome assembly. Previous studies have found that structural genomic variants play an important role in the evolution of species [77]. According to the results of the collinearity analysis, genomic structural variations were observed both within and between chromosomes. Chromosomal segment heterotopy was observed in chromosomes 5, 7, 9, and 10 in *P. echinovolvatius* and *C. columnatus*. Major chromosomal rupture and fusion events occurred between chromosomes 2, 3, and 11. Obvious rearrangements, ruptures, and fusion events were detected in chromosome 1 (Figure 8).



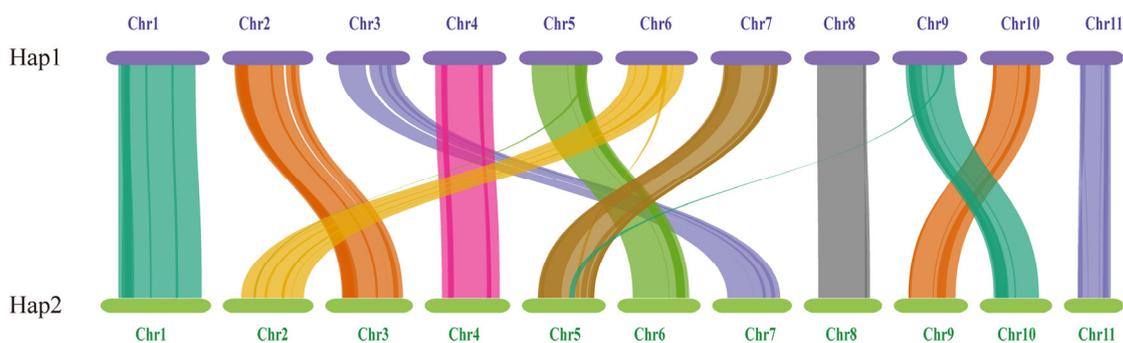
**Figure 8.** The genome collinearity among *P. echinovolatus*, *C. columnatus*, and *P. ostreatus*. The red font is the target species sequenced in this paper.

### 3.4.5. Comparative Analysis of Hap1 and Hap2

The length and gene count of each chromosome in Hap1 and Hap2 are shown in Table 3. A total of 156 pairs of one-to-one matching gene blocks were detected, which accounted for 84.62% of all the genes. Strong collinearity was detected between Hap1 and Hap2; chromosomes 2, 3, 5, 6, 7, 9, and 10 in Hap1 corresponded to chromosomes 3, 7, 6, 2, 5, 10, and 9 in Hap2, respectively. Chromosomes 5, 6, and 9 in Hap1 and chromosomes 2 and 5 in Hap2 experienced minor fragmentation and fusion events. These results suggested that the two haploid nuclei might have been derived from different parents (Figure 9).

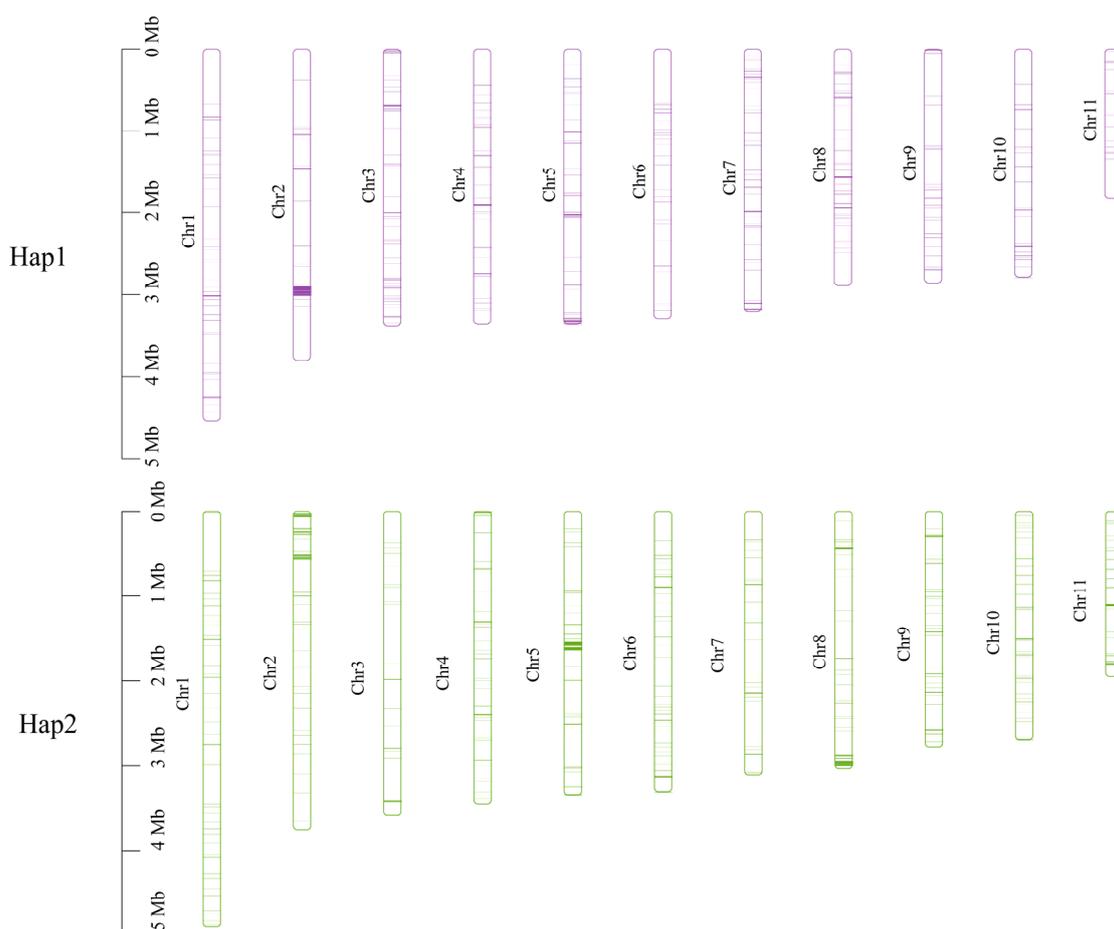
**Table 3.** The length and gene count of each chromosome.

ID	Hap1		Hap2	
	Length (bp)	Gene Number	Length (bp)	Gene Number
Chr1	4,557,232	1177	4,928,922	1169
Chr2	3,821,430	812	3,777,475	980
Chr3	3,400,287	844	3,601,952	766
Chr4	3,382,853	923	3,473,486	949
Chr5	3,377,001	1052	3,363,986	614
Chr6	3,305,313	867	3,326,919	1032
Chr7	3,217,999	586	3,117,068	719
Chr8	2,906,280	763	3,049,163	783
Chr9	2,867,821	579	2,808,455	679
Chr10	2,792,080	687	2,697,050	574
Chr11	1,831,909	454	1,961,784	465
Total	35,460,205	8744	36,106,260	8730



**Figure 9.** The genome collinearity among Hap1 and Hap2.

Subsequently, non-collinear gene sequences were extracted from the genome and subjected to blastn alignment for the two haplotypes ( $e\text{-value} < 1 \times 10^{-5}$ ). There were 256 unique genes in Hap1 and 370 unique genes in Hap2. These unique genes were widely distributed across each chromosome, with a predominant location on chromosome 2 in Hap1 and on chromosomes 5 and 8 in Hap2 (Figure 10).



**Figure 10.** Location of the unique genes found within each of the two monokaryons on each chromosome.

Functional enrichment analysis was conducted for unique genes, and the KEGG annotation results indicated that the unique genes in Hap1 were primarily enriched in pathways related to “Biosynthesis of secondary metabolites”, “Metabolic pathways”, and “Ubiquitin-mediated proteolysis”. The unique genes in Hap2 were significantly enriched in “Metabolic pathways” and “Oxidative phosphorylation” (Table S12). The results of the

GO analysis indicated that the unique genes in Hap1 were mainly enriched in processes related to “cellular protein modification process”, “macromolecule modification”, “organelle membrane”, and “protein modification process”. The unique genes in Hap2 were primarily enriched in processes related to “envelope”, “organelle envelope”, and “oxidoreductase activity” (Table S13).

The genome of the dikaryotic strain of *P. echinovolvatatus* is highly heterozygous, and the results from the phased assembly and annotation indicate genetic differences between the two nuclei. Given the limitations of current sequencing technologies and analytical tools, future work involving the mononucleation of dikaryotic mycelium, followed by separate sequencing of the resulting mononuclear strains, could yield more accurate genomic information. This approach would also allow for a deeper investigation into the roles of each nucleus in the growth and reproduction of *P. echinovolvatatus*.

#### 4. Conclusions

In this study, we present the genomic information about *P. echinovolvatatus* obtained by integrating different sequencing technologies. Functional annotations of the genomes were obtained using multiple public databases. This whole-genome assembly, along with the associated annotation data, represents the first chromosome-level genome assembly for *P. echinovolvatatus*. These new data will aid future studies on the evolution of species and phylogenetic analyses based on genomic data. Furthermore, these data will be useful for breeding programs, as well as for studies of developmental mechanisms and the pathways underlying the biosynthesis of bioactive compounds in *P. echinovolvatatus*. Our findings provide genetic and molecular insights into the evolutionary history of *P. echinovolvatatus* and offer genomic resources to further facilitate gene editing to enhance desirable traits.

**Supplementary Materials:** The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/jof11010062/s1>, Figure S1: Histogram of the 21-mer depth distribution of the Illumina sequencing reads of *P. echinovolvatatus*; Figure S2: Whole-genome Hi-C interaction map at 50 k resolution; Figure S3: “Terpenoid backbone biosynthesis” pathway of *P. echinovolvatatus*; Table S1: Macrofungi used in this study; Table S2: Statistics for the sequencing data of the *P. echinovolvatatus* genome; Table S3. Characteristics of the gene prediction of *P. echinovolvatatus*; Table S4: Statistics of the *P. echinovolvatatus* repetitive sequence prediction results; Table S5: Statistical results of the non-coding RNAs in *P. echinovolvatatus*; Table S6: Statistics of *P. echinovolvatatus* protein-coding gene annotation; Table S7: Statistics of the *P. echinovolvatatus* P450 annotation (Monop); Table S8: Carbohydrate-active enzyme annotation results (Monop); Table S9: Statistics of the antiSMASH database annotation (Monop); Table S10: Putative genes involved in terpenoid backbone biosynthesis (Monop); Table S11: Statistics of the clustered results (Monop); Table S12: The KEGG function annotation of unique genes; Table S13: The GO function annotation of unique genes.

**Author Contributions:** J.L. conceptualized the study; M.A., R.L. and J.Z. participated in the study design; M.A., Y.C. and G.Q. conducted the experiments and analyzed the data; M.A. wrote the original draft of the manuscript; J.L., X.W. and X.L. revised and improved the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China (No. 32370018 and 31770657).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The genome sequencing of *P. echinovolvatatus* RITF7875 generated for this study has been submitted to the NCBI (BioProject: PRJNA 1192246).

**Acknowledgments:** We acknowledge everyone who contributed to this article.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Zhang, M.; Zheng, D.; Hu, Z. A new species of the Genus *Dictyophora* from China. *Mycotaxon* **1988**, *33*, 145–148.
2. Melanda, G.C.S.; Silva-Filho, A.G.S.; Lenz, A.R.; Menolli, N.; de Lima, A.A., Jr.; Ferreira, R.J.; de Assis, N.M.; Cabral, T.S.; Martín, M.P.; Baseia, I.G. An overview of 24 years of molecular phylogenetic studies in *Phallales* (Basidiomycota) with notes on systematics, geographic distribution, lifestyle, and edibility. *Front. Microbiol.* **2021**, *12*, 689374. [[CrossRef](#)] [[PubMed](#)]
3. Sommai, S.; Khamsuntorn, P.; Somrithipol, S.; Luangsa-Ard, J.J.; Pinruan, U. *Phallus Chiangmaiensis* sp. Nov. and a record of *P. merulinus* in Thailand. *Mycobiology* **2021**, *49*, 439–453. [[CrossRef](#)] [[PubMed](#)]
4. Kreisel, H. A preliminary survey of the genus *Phallus* sensu lato. *Czech Mycol.* **1996**, *48*, 273–281. [[CrossRef](#)]
5. Cabral, T.S.; Marinho, P.; Goto, B.T.; Baseia, I.G. *Abrachium*, a new genus in the *Clathraceae*, and *Itajahya* reassessed. *Mycotaxon* **2012**, *119*, 419–429. [[CrossRef](#)]
6. Moreno, G.; Khalid, A.N.; Alvarado, P.; Kreisel, H. *Phallus hadriani* and *P. roseus* from Pakistan. *Mycotaxon* **2013**, *125*, 45–51. [[CrossRef](#)]
7. Li, H.; MORTIMER, E.E.; Karunaratna, S.C.; Xu, J.; Hyde, K.D. New species of *Phallus* from a subtropical forest in Xishuangbanna, China. *Phytotaxa* **2014**, *163*, 91–103. [[CrossRef](#)]
8. Song, B.; Li, T.; Li, T.H.; Huang, Q.J.; Deng, W.Q. *Phallus fuscoechinovolvatus* (Phallaceae, Basidiomycota), a new species with a dark spinose volva from southern China. *Phytotaxa* **2018**, *334*, 19–27. [[CrossRef](#)]
9. Læssøe, T.; Spooner, B. The uses of Gasteromycetes. *Mycologist* **1994**, *8*, 154–159. [[CrossRef](#)]
10. Deng, C.; Shang, J.; Fu, H.; Chen, J.; Liu, H.; Chen, J. Mechanism of the immunostimulatory activity by a polysaccharide from *Dictyophora indusiata*. *Int. J. Biol. Macromol.* **2016**, *91*, 752–759. [[CrossRef](#)]
11. Cheng, C.; Liu, F.; Wang, B.; Qu, P.; Liu, J.; Zhang, Y.; Liu, W.; Tong, Z.; Deng, G. Influences of *Serendipita indica* and *Dictyophora echinovolvata* on the growth and fusarium wilt disease resistance of banana. *Biology* **2022**, *11*, 393. [[CrossRef](#)] [[PubMed](#)]
12. Huang, M.; Tian, H.; Sun, B.; Liu, Y. Isolation and identification of oxacyclopentadecan-2-one from the dried fruiting body of *Dictyophora echinovolvata* Zang, Zheng et Hu. *Flavour. Fragr. J.* **2012**, *27*, 75–76. [[CrossRef](#)]
13. Lin, C.; Zhang, H.; Chen, L.; Fang, Y.; Chen, J. Immunoregulatory function of *Dictyophora echinovolvata* spore polysaccharides in immunocompromised mice induced by cyclophosphamide. *Open Life Sci.* **2021**, *16*, 620–629. [[CrossRef](#)] [[PubMed](#)]
14. Yu, W.X.; Lin, C.Q.; Zhao, Q.; Lin, X.J.; Dong, X.L. Neuroprotection against hydrogen peroxide-induced toxicity by *Dictyophora echinovolvata* polysaccharide via inhibiting the mitochondria-dependent apoptotic pathway. *Biomed. Pharmacother.* **2017**, *88*, 569–573. [[CrossRef](#)]
15. Gao, Q.; Yan, D.; Song, S.; Fan, Y.; Wang, S.; Liu, Y.; Huang, Y.; Rong, C.; Guo, Y.; Zhao, S.; et al. Haplotype-resolved genome analyses reveal genetically distinct nuclei within a commercial cultivar of *Lentinula edodes*. *J. Fungi* **2022**, *8*, 167. [[CrossRef](#)]
16. Park, Y.J.; Baek, J.H.; Lee, S.; Kim, C.; Rhee, H.; Kim, H.; Seo, J.S.; Park, H.R.; Yoon, D.E.; Nam, J.Y.; et al. Whole genome and global gene expression analyses of the model Mushroom *Flammulina velutipes* reveal a high capacity for lignocellulose degradation. *PLoS ONE* **2014**, *9*, e93560.
17. Fang, M.; Wang, X.; Chen, Y.; Wang, P.; Lu, L.; Lu, J.; Yao, F.; Zhang, Y. Genome sequence analysis of *Auricularia heimuer* combined with genetic Linkage Map. *J. Fungi* **2020**, *6*, 37. [[CrossRef](#)]
18. Liu, W.; Chen, L.; Cai, Y.; Zhang, Q.; Bian, Y. Opposite polarity monospore genome de novo sequencing and comparative analysis reveal the possible heterothallic life cycle of *Morchella importuna*. *Int. J. Mol. Sci.* **2018**, *19*, 2525. [[CrossRef](#)]
19. Morin, E.; Kohler, A.; Baker, A.R.; Foulongne-Oriol, M.; Lombard, V.; Nagy, L.G.; Ohm, R.A.; Patyshakuliyeva, A.; Brun, A.; Aerts, A.L.; et al. Genome sequence of the button mushroom *Agaricus bisporus* reveals mechanisms governing adaptation to a humic-rich ecological niche. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 17501–17506. [[CrossRef](#)]
20. Duan, M.; Long, S.; Wu, X.; Feng, B.; Qin, S.; Li, Y.; Li, X.; Li, C.; Zhao, C.; Wang, L.; et al. Genome, transcriptome, and metabolome analyses provide new insights into the resource development in an edible fungus *Dictyophora indusiata*. *Front. Microbiol.* **2023**, *14*, 1137159. [[CrossRef](#)]
21. Ma, L.; Yang, C.; Xiao, D.; Liu, X.; Jiang, X.; Lin, H.; Ying, Z.; Lin, Y. Chromosome-level assembly of *Dictyophora rubrovolvata* genome using third-generation DNA sequencing and Hi-C analysis. *G3* **2023**, *13*, jkad102. [[CrossRef](#)] [[PubMed](#)]
22. Wang, J.; Tian, H.; Xu, Y.; Peng, X.; Zhang, C.; Lu, M.; Lu, Z.; Xu, S.; Wen, T.; Yang, C. Comprehensive analysis of the metabolome and the whole genome transcriptome to reveal the regulatory mechanism the flavor of *Phallus rubrovolvatus*. *Food Biosci.* **2024**, *60*, 104483. [[CrossRef](#)]
23. Cheng, H.; Jarvis, E.D.; Fedrigo, O.; Koepfli, K.-P.; Urban, L.; Gemmill, N.J.; Li, H. Haplotype-resolved assembly of diploid genomes without parental data. *Nat. Biotechnol.* **2022**, *40*, 1332–1335. [[CrossRef](#)] [[PubMed](#)]

24. Vollger, M.R.; Logsdon, G.A.; Audano, P.A.; Sulovari, A.; Porubsky, D.; Peluso, P.; Wenger, A.M.; Concepcion, G.T.; Kronenberg, Z.N.; Munson, K.M.; et al. Improved assembly and variant detection of a haploid human genome using single-molecule, high-fidelity long reads. *Ann. Hum. Genet.* **2020**, *84*, 125–140. [[CrossRef](#)] [[PubMed](#)]
25. Zeng, Q.; Zhou, Z.; He, Q.; Li, L.; Pu, F.; Yan, M.; Xu, P. Chromosome-level haplotype-resolved genome assembly for *Takifugu ocellatus* using PacBio and Hi-C technologies. *Sci. Data* **2023**, *10*, 22. [[CrossRef](#)] [[PubMed](#)]
26. Yi, L.; Sa, R.; Zhao, S.; Zhang, X.; Lu, X.; Mu, Y.; Bateer, S.; Su, S.; Wang, S.; Li, Z.; et al. Chromosome-scale, haplotype-resolved genome assembly of *Suaeda glauca*. *Front. Genet.* **2022**, *13*, 884081. [[CrossRef](#)]
27. Han, L.; Luo, X.; Zhao, Y.; Li, N.; Xu, Y.; Ma, K. A haplotype-resolved genome provides insight into allele-specific expression in wild walnut (*Juglans regia* L.). *Sci. Data* **2024**, *11*, 278. [[CrossRef](#)]
28. Sun, X.; Jiao, C.; Schwaninger, H.; Chao, C.T.; Ma, Y.; Duan, N.; Khan, A.; Ban, S.; Xu, K.; Cheng, L.; et al. Phased diploid genome assemblies and pan-genomes provide insights into the genetic history of apple domestication. *Nat. Genet.* **2020**, *52*, 1423–1432. [[CrossRef](#)]
29. Duan, H.; Jones, A.W.; Hewitt, T.; Mackenzie, A.; Hu, Y.; Sharp, A.; Lewis, D.; Mago, R.; Upadhyaya, N.M.; Rathjen, J.P.; et al. Identification and correction of phase switches with Hi-C data in the Nanopore and HiFi chromosome-scale assemblies of the dikaryotic leaf rust fungus *Puccinia triticina*. *bioRxiv* **2021**. [[CrossRef](#)]
30. Liang, J.; Li, Y.; Dodds, P.N.; Figueroa, M.; Sperschneider, J.; Han, S.; Tsui, C.K.M.; Zhang, K.; Li, L.; Ma, Z.; et al. Haplotype-phased and chromosome-level genome assembly of *Puccinia polysora*, a giga-scale fungal pathogen causing southern corn rust. *Mol. Ecol. Resour.* **2023**, *23*, 601–620. [[CrossRef](#)]
31. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)] [[PubMed](#)]
32. Wingett, S.W.; Andrews, S. FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Research* **2018**, *7*, 1338. [[CrossRef](#)] [[PubMed](#)]
33. Cheng, H.; Concepcion, G.T.; Feng, X.; Zhang, H.; Li, H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat. Methods* **2021**, *18*, 170–175. [[CrossRef](#)] [[PubMed](#)]
34. Zhang, X.; Zhang, S.; Zhao, Q.; Ming, R.; Tang, H. Assembly of allele-aware, chromosomal-scale autopolyploid genomes based on Hi-C data. *Nat. Plants* **2019**, *5*, 833–845. [[CrossRef](#)]
35. Burton, J.N.; Adey, A.; Patwardhan, R.P.; Qiu, R.; Kitzman, J.O.; Shendure, J. Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. *Nat. Biotechnol.* **2013**, *31*, 1119–1125. [[CrossRef](#)]
36. Lieberman-Aiden, E.; van Berkum, N.L.; Williams, L.; Imakaev, M.; Ragozcy, T.; Telling, A.; Amit, I.; Lajoie, B.R.; Sabo, P.J.; Dorschner, M.O.; et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **2009**, *326*, 289–293. [[CrossRef](#)]
37. Simao, F.A.; Waterhouse, R.M.; Ioannidis, P.; Kriventseva, E.V.; Zdobnov, E.M. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **2015**, *31*, 3210–3212. [[CrossRef](#)]
38. Hoff, K.J.; Stanke, M. WebAUGUSTUS—a web service for training AUGUSTUS and predicting genes in eukaryotes. *Nucleic Acids Res.* **2013**, *41*, W123–W128. [[CrossRef](#)]
39. Birney, E.; Clamp, M.; Durbin, R. GeneWise and Genomewise. *Genome Res.* **2004**, *14*, 988–995. [[CrossRef](#)]
40. Haas, B.J.; Salzberg, S.L.; Zhu, W.; Pertea, M.; Allen, J.E.; Orvis, J.; White, O.; Buell, C.R.; Wortman, J.R. Automated eukaryotic gene structure annotation using EVIDENCEModeler and the program to assemble spliced alignments. *Genome Biol.* **2008**, *9*, r7. [[CrossRef](#)]
41. Tarailo-Graovac, M.; Chen, N. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr. Protoc. Bioinform.* **2009**, *4*, 4.10.11–14.10.14. [[CrossRef](#)] [[PubMed](#)]
42. Lowe, T.M.; Eddy, S.R. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **1997**, *25*, 955–964. [[CrossRef](#)] [[PubMed](#)]
43. Lagesen, K.; Hallin, P.; Rodland, E.A.; Stærfeldt, H.-H.; Rognes, T.; Ussery, D.W. RNAMmer: Consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* **2007**, *35*, 3100–3108. [[CrossRef](#)] [[PubMed](#)]
44. Griffiths-Jones, S.; Moxon, S.; Marshall, M.; Khanna, A.; Eddy, S.R.; Bateman, A. Rfam: Annotating non-coding RNAs in complete genomes. *Nucleic Acids Res.* **2005**, *33*, D121–D124. [[CrossRef](#)]
45. Chen, W.; Lee, M.-K.; Jefcoate, C.; Kim, S.-C.; Chen, F.; Yu, J.-H. Fungal Cytochrome P450 Monooxygenases: Their distribution, structure, functions, family expansion, and evolutionary origin. *Genome Biol. Evol.* **2014**, *6*, 1620–1634. [[CrossRef](#)]
46. Zheng, J.; Ge, Q.; Yan, Y.; Zhang, X.; Huang, L.; Yin, Y. dbCAN3: Automated carbohydrate-active enzyme and substrate annotation. *Nucleic Acids Res.* **2023**, *51*, W115–W121. [[CrossRef](#)]
47. Blin, K.; Shaw, S.; Augustijn, H.E.; Reitz, Z.L.; Biermann, F.; Alanjary, M.; Fetter, A.; Terlouw, B.R.; Metcalf, W.W.; Helfrich, E.J.N.; et al. antiSMASH 7.0: New and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Res.* **2023**, *51*, W46–W50. [[CrossRef](#)]

48. Emms, D.M.; Kelly, S. OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.* **2019**, *20*, 238. [[CrossRef](#)]
49. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [[CrossRef](#)]
50. Castresana, J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* **2000**, *17*, 540–552. [[CrossRef](#)]
51. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. jModelTest 2: More models, new heuristics and parallel computing. *Nat. Methods* **2012**, *9*, 772. [[CrossRef](#)] [[PubMed](#)]
52. Guindon, S.; Dufayard, J.-F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321. [[CrossRef](#)] [[PubMed](#)]
53. Chen, C.; Wu, Y.; Li, J.; Wang, X.; Zeng, Z.; Xu, J.; Liu, Y.; Feng, J.; Chen, H.; He, Y.; et al. TBtools-II: A “one for all, all for one” bioinformatics platform for biological big-data mining. *Mol. Plant.* **2023**, *16*, 1733–1742. [[CrossRef](#)]
54. Wöstemeyer, J.; Kreibich, A. Repetitive DNA elements in fungi (Mycota): Impact on genomic architecture and evolution. *Curr. Genet.* **2002**, *41*, 189–198. [[CrossRef](#)]
55. Biscotti, M.A.; Olmo, E.; Heslop-Harrison, J.S. Repetitive DNA in eukaryotic genomes. *Chromosome Res.* **2015**, *23*, 415–420. [[CrossRef](#)]
56. Plohl, M.; Luchetti, A.; Mestrovic, N.; Mantovani, B. Satellite DNAs between selfishness and functionality: Structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. *Gene* **2008**, *409*, 72–82. [[CrossRef](#)]
57. Cresnar, B.; Petric, S. Cytochrome P450 enzymes in the fungal kingdom. *Biochim. Biophys. Acta* **2011**, *1814*, 29–35. [[CrossRef](#)]
58. Guengerich, F.P. Cytochrome P450s and other enzymes in drug metabolism and toxicity. *AAPS J.* **2006**, *8*, E101–E111. [[CrossRef](#)]
59. Nelson, D.R. The Cytochrome p450 homepage. *Hum. Genom.* **2009**, *4*, 59–65. [[CrossRef](#)]
60. Lombard, V.; Ramulu, H.G.; Drula, E.; Coutinho, P.M.; Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **2014**, *42*, D490–D495. [[CrossRef](#)]
61. Payne, C.M.; Knott, B.C.; Mayes, H.B.; Hansson, H.; Himmel, M.E.; Sandgren, M.; Ståhlberg, J.; Beckham, G.T. Fungal cellulases. *Chem. Rev.* **2015**, *115*, 1308–1448. [[CrossRef](#)] [[PubMed](#)]
62. Ha, J.W.; Kim, J.; Kim, H.; Jang, W.; Kim, K.H. Mushrooms: An important source of natural bioactive compounds. *Nat. Prod. Sci.* **2020**, *26*, 118–131. [[CrossRef](#)]
63. Taofiq, O.; Martins, A.; Barreiro, M.F.; Ferreira, I.C.F.R. Anti-inflammatory potential of mushroom extracts and isolated metabolites. *Trends Food Sci. Technol.* **2016**, *50*, 193–210. [[CrossRef](#)]
64. Robertsen, H.L.; Musiol-Kroll, E.M. Actinomycete-derived polyketides as a source of antibiotics and lead structures for the development of new antimicrobial drugs. *Antibiotics* **2019**, *8*, 157. [[CrossRef](#)]
65. Roskova, Z.; Skarohlid, R.; McGachy, L. Siderophores: An alternative bioremediation strategy? *Sci. Total Environ.* **2022**, *819*, 153144. [[CrossRef](#)]
66. Bozhüyük, K.A.J.; Präve, L.; Kegler, C.; Schenk, L.; Kaiser, S.; Schelhas, C.; Shi, Y.N.; Kutenlochner, W.; Schreiber, M.; Kandler, J.; et al. Evolution-inspired engineering of nonribosomal peptide synthetases. *Science* **2024**, *383*, eadg4320. [[CrossRef](#)]
67. Hai, Y.; Jenner, M.; Tang, Y. Fungal siderophore biosynthesis catalysed by an iterative nonribosomal peptide synthetase. *Chem. Sci.* **2020**, *11*, 11525–11530. [[CrossRef](#)]
68. Ou, Y.X.; Li, Y.Y.; Qian, X.M.; Shen, Y.M. Guanacastane-type diterpenoids from *Coprinus radians*. *Phytochemistry* **2012**, *78*, 190–196. [[CrossRef](#)]
69. Wang, X.L.; Ding, Z.Y.; Liu, G.Q.; Yang, H.; Zhou, G.Y. Improved production and antitumor properties of triterpene acids from submerged culture of *Ganoderma lingzhi*. *Molecules* **2016**, *21*, 1395. [[CrossRef](#)]
70. Plett, J.M.; Wojtalewicz, D.; Plett, K.L.; Collin, S.; Kohler, A.; Jacob, C.; Martin, F. Sesquiterpenes of the ectomycorrhizal fungus *Pisolithus microcarpus* alter root growth and promote host colonization. *Mycorrhiza* **2024**, *34*, 69–84. [[CrossRef](#)]
71. Li, H.P.; Yang, W.J.; Qu, S.X.; Pei, F.; Luo, X.; Mariga, A.M.; Ma, L. Variation of volatile terpenes in the edible fungi mycelia *Flammulina velutipes* and communications in fungus-mite interactions. *Food Res. Int.* **2018**, *103*, 150–155. [[CrossRef](#)] [[PubMed](#)]
72. Sidorova, D.E.; Plyuta, V.A.; Padiy, D.A.; Kupriyanova, E.V.; Roshina, N.V.; Koksharova, O.A.; Khmel, I.A. The effect of volatile organic compounds on different organisms: Agrobacteria, plants and insects. *Microorganisms* **2021**, *10*, 69. [[CrossRef](#)] [[PubMed](#)]
73. Sun, Y.; Xiang, Z.; Luo, Q.; Li, L.; Liu, N.; Tang, J.; Zhang, H.; Pan, G.; Long, H.; Zou, F. Analysis of volatile components and flavor characteristics about different temperature varieties of *Dictyophora duplicata*. *Biotechnology* **2019**, *29*, 586–592.
74. Zheng, j.; Li, T.T.; Song, J.X.; Kan, J.Q. Analysis of Volatile Compounds in Embryo of *Dictyophora indusiata* (Vent. Pers.) Fisch Using Headspace-Solid Phase Microextraction Combined with GC-MS. *Food Sci.* **2014**, *35*, 125–128.
75. Pu, X.; Dong, X.; Li, Q.; Chen, Z.; Liu, L. An update on the function and regulation of methylerythritol phosphate and mevalonate pathways and their evolutionary dynamics. *J. Integr. Plant Biol.* **2021**, *63*, 1211–1226. [[CrossRef](#)]

- 
76. Shi, L.; Ren, A.; Mu, D.; Zhao, M. Current progress in the study on biosynthesis and regulation of ganoderic acids. *Appl. Microbiol. Biotechnol.* **2010**, *88*, 1243–1251. [[CrossRef](#)]
  77. Wellenreuther, M.; Merot, C.; Berdan, E.; Bernatchez, L. Going beyond SNPs: The role of structural genomic variants in adaptive evolution and species diversification. *Mol. Ecol.* **2019**, *28*, 1203–1209. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.