



Article

Chromosome-Scale Genome Assembly Provides Insights into Fresh Pine Wood Decay Strategies of the *Wolfiporia hoelen*Chi Yang ¹, Donglai Xiao ¹, Xiaoling Jiang ¹, Yaru Li ¹, Xiaoyu Liu ¹, Hui Lin ¹, Chuansen Liu ² and Lu Ma ^{1,*} 

¹ Institute of Edible Mushroom, National and Local Joint Engineering Research Center for Breeding & Cultivation of Featured Edible Mushroom, Fujian Academy of Agricultural Sciences, Fuzhou 350014, China; yangchi@faas.cn (C.Y.); xiaodonglai@faas.cn (D.X.); jiangxiaoling@faas.cn (X.J.); 13625083980@163.com (Y.L.); liuxiaoyu@faas.cn (X.L.); linhui-syjs@faas.cn (H.L.)

² Agriculture and Rural Bureau of Datan County, Sanming 366100, China; syjls@163.com

* Correspondence: malu@faas.cn

Abstract: The sclerotia of *Wolfiporia hoelen* (Fr.) Y.C. Dai & V. Papp is an important traditional Chinese medicine with diverse pharmacological properties. This study utilized a combination of PacBio Long-Read Sequencing, Illumina Short-Read Sequencing, and Hi-C Sequencing to generate a high-quality chromosome-level genome assembly of a *W. hoelen* strain Minling A5. There were 112 contigs in the genome, with 62.95 Mb in total length and 4.21 Mb in length for the contig N50. The average GC content was 51.89%. Based on Hi-C data, we corrected the CCS data and scaffolded them into 14 pseudo-chromosomes. The genome contained 44.37% repetitive sequences and 12,670 protein-coding genes, 86.53% (10,963) of which could be functionally annotated in at least one of the KOG, GO, Pfam, Swissprot, TrEMBL, NR, and KEGG databases. In addition, 240 transfer RNAs, 97 ribosomal RNAs, and 103 other non-coding RNAs were identified in the *W. hoelen* genome. A total of 755 pseudogenes were also identified, with an average length of 2665.51 bp. Further, there were 398, 100, 2837, 519, and 2068 genes annotated by CAZymes, TCDB, PHI, P450, and DFVF databases, respectively. One notable attribute of *W. hoelen* is its capacity to thrive in a substrate of fresh pine sawdust. Through an analysis of the growth on various pure wood sawdust culture media, we found that the growth of *W. hoelen* and *Sparassis latifolia* on pine sawdust was similar to that on broad-leaved wood sawdust, while the growth of *Pleurotus ostreatus*, *P. eryngii*, and *Cyclocybe aegerita* was slower than that on broad-leaved wood sawdust. By the functional annotation analysis of orthogroups in these five mushroom-forming fungi, it was determined that 645 orthogroups were specifically common in *W. hoelen* and *S. latifolia*. The genes in these specific orthogroups were significantly enriched in 12 pathways, including steroid biosynthesis, biosynthesis of antibiotics, and tyrosine metabolism. The high-quality genome and comparative genome analysis results significantly contribute to advancing our foundational knowledge of *W. hoelen* biology, while also offering valuable insights for the development of innovative biotechnological approaches aimed at enhancing the efficient and sustainable utilization of *Pinus*.

Keywords: *Wolfiporia hoelen*; genome; chromosome; pine sawdust decay; comparative genomics



Citation: Yang, C.; Xiao, D.; Jiang, X.; Li, Y.; Liu, X.; Lin, H.; Liu, C.; Ma, L. Chromosome-Scale Genome Assembly Provides Insights into Fresh Pine Wood Decay Strategies of the *Wolfiporia hoelen*. *Horticulturae* **2024**, *10*, 703. <https://doi.org/10.3390/horticulturae10070703>

Academic Editors: Chenyang Huang and Mingwen Zhao

Received: 26 April 2024

Revised: 21 June 2024

Accepted: 27 June 2024

Published: 3 July 2024



Copyright: © 2024 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

Pinus is an economically and ecologically important soft conifer species in China, which has been seriously infected by pine wilt disease (PWD) since the beginning of the 20th century [1]. PWD was caused by a migratory plant-parasitic nematode, *Bursaphelenchus xylophilus*, which resulted in increasingly serious economic losses and ecological damage, especially in China and Japan [2,3]. As per the declaration made by the China National Forestry and Grassland Administration, there are a total of 19 provincial epidemic areas of PWD in mainland China in the year 2022 [4]. PWD rapidly spreads, leading to a significantly high mortality rate of the affected plants and more than one billion dollars of mean annual

economic losses in China [5]. The management of pine wilt disease prevention and control in this region is of utmost concern due to the severity of the situation. Therefore, exploring the treatment and resource utilization of epidemic trees has become particularly important.

The decomposition of wood has important implications for the biogeochemical cycling of carbon (C) and nitrogen (N) in forest soils and litter. Mushroom-forming fungi play important roles in decomposing wood in the global carbon cycle. However, sawdust derived from pine trees is infrequently utilized in cultivating mushrooms due to the inhibitory effects of the resin content on the growth of most fungi directly on coniferous wood [6]. There are several wood-decaying fungi reported to colonize coniferous wood directly [6], and our previous study revealed that *Sparassis latifolia* Y.C. Dai & Zheng Wang can grow on fresh pine wood sawdust substrate [7,8]. The molecular mechanism of *S. latifolia*'s degradation of fresh pine sawdust was investigated through the application of comparative transcriptomics techniques and the genes enriched in nitronate monooxygenase activity, dioxygenase activity, and oxidation-reduction process GO terms and peroxisome KEGG pathway were found to play key roles [9]. However, the current industrialized cultivation mode of *S. latifolia* requires transporting pine trees to the factory, which increases the risk of pine wood nematode spread.

The taxonomic designation *Wolfiporia hoelen* (Fr.) Y.C. Dai & V. Papp, previously classified as *Poria cocos* or *Wolfiporia cocos*, denotes the Chinese fungal species commonly recognized as “Fuling” [10]. This fungus is known to grow on pine roots throughout the year and is prevalent in East Asia. The main active components of *W. hoelen* are polysaccharides with biological activities, such as anti-oxidation [11], cancer immunotherapy [12], and modulating gastrointestinal function [13]. The triterpenoids in *W. cocos* also have important phytochemistry and pharmacological activities [14–17]. However, the commercial production of *W. cocos* is limited because of the severe destruction of *W. cocos* habitat and shortages in pinewood resources [18]. There is a conflict between the demand for large-scale development of *W. hoelen* and the protection of pine resources, which prevents *W. hoelen* from achieving sustainable development [19].

In order to meet the market demand for *W. hoelen* and fully utilize the epidemic pine wood, using epidemic wood and stumps to cultivate *W. hoelen* in situ is ideal. We have been using the strain of Minling A5 in recent years in southeast China, especially in Fujian province [20]. Minling A5 is an excellent variety of *W. hoelen* bred by the Institute of Edible Mushroom, Fujian Academy of Agricultural Sciences. It was recognized as a new variety by the Fujian Provincial Crop Variety Approval Committee in April 2013. But until now, there has been no report on the genetic and genome information of this variety. Therefore, it is very urgent to study the molecular mechanism of Minling A5 degrading pine wood.

W. hoelen is an aerobic saprophytic fungus that exhibits parasitic behavior in the roots of pine trees, specifically *Pinus densiflora* and *P. massoniana* Lamb. Previous study examined the transcriptome and secretome of *W. cocos* when cultivated on different carbon sources, including lodgepole pine (*Pinus contorta*) [21]. When the brown rot fungus *W. cocos* was cultivated on lodgepole pine as the exclusive carbon source compared to glucose, a total of 27 genes encoding carbohydrate-active enzymes, eight genes associated with the production of extracellular oxidants, 10 cytochrome P450 genes, and 35 transporter genes were found to be significantly upregulated greater than 4-fold, with a statistical significance of $p < 0.01$. However, the mechanism of pine wood decay by *W. hoelen* is still limited.

A high-level genome sequence will contribute to an understanding of the biological characteristics of *W. hoelen*. Efforts were undertaken to sequence the *Wolfiporia* genome, resulting in the publication of several reference genomes [22–25]. The initial strain utilized for genome assembly within the *Wolfiporia* genus was the single spore isolate MD-104 SS10 from America [26]. Then, the genome of a Chinese strain CGMCC5.78 of *W. cocos* was reported in 2020, which also conducted a comparative analysis of transcriptomes from mycelial and sclerotial tissues to identify differentially expressed genes (DEGs) associated with sclerotial development [25]. The genome sequence of *W. cocos* KMCC03342, the strain registered and maintained by the Korea Seed and Variety Service for commercial

applications, has been published [22]. Another study introduced the initial high-quality homokaryotic genome of *W. hoelen*, consisting of 14 chromosomes [24]. Huazhong Agricultural University recently published the genome sequence of an additional strain of *W. cocos* (https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_034769205.1/, accessed on 5 June 2024). However, there is still no genomic information about *W. hoelen* strain Minling A5.

This study aimed to assemble a high-quality chromosome-scale reference genome of *W. hoelen* strain Minling A5 using Illumina, PacBio, and Hi-C. The high-quality reference genome will facilitate the study of the molecular mechanisms underlying the pine wood decay by *W. hoelen*.

2. Materials and Methods

2.1. Strains and Culture Conditions

The *W. hoelen* strain Minling A5 was preserved at the Institute of Edible Mushroom, Fujian Academy of Agricultural Sciences (Fuzhou, China). The strain was inoculated on potato dextrose agar (PDA) plate in the dark at 24 °C for 7 d and then the mycelium was collected. Genomic DNA was isolated from the mycelium of *W. hoelen* strain by the cetyl-trimethyl ammonium bromide (CTAB) method [27]. Agar gel electrophoresis was used to check the integrity of the genomic DNA molecules.

2.2. Genome Features Estimation from K-mer Method

According to Illumina's instructions, the genomic DNA fragments were sonicated to a size of 350 bp and paired-end genomic libraries were made using the fragments (Illumina, San Diego, CA, USA). Using an Agilent Bioanalyser 2100 and qPCR, the library was quality checked, and sequenced on an Illumina NovaSeq 6000 sequencing platform with paired-end 150 bp reads. K-mer ($k = 19$) analysis was conducted on high-quality clean reads for genome size estimation [28].

2.3. Sequencing Libraries

The construction of sequencing libraries was performed according to previous reports [29]. For long-read sequencing, the SMRTbell® Express Template Prep Kit (version 2.0) was used to create a 15-kb SMRTbell library [30]. After shearing into a large fragment by g-TUBE (Covaris), the genomic DNA was purified and recovered by AMPure PB magnetic beads. The size and quality of the library were assessed after DNA damage repair, end repair, ligation with T-overhang, exonuclease digestion size selection, and library purification. Then, the library was sequenced using the PacBio Sequel II platform (PacBio Biosciences, Menlo Park, CA, USA) at the Beijing Biomarker Technologies Co., Ltd. (Beijing, China).

For Hi-C sequencing, libraries were prepared as previously reported [31]. To preserve DNA 3D structure in cells, formaldehyde was applied to fix the samples, and DPN II restriction enzyme was applied to digest the DNA, followed by repairing 5' overhangs with biotinylated residues. Then the Hi-C libraries were quantified and sequenced using the Illumina NovaSeq 6000 instrument (Illumina).

To predict genes, RNA-seq data were generated from the same sample according to the protocol described previously [9]. The total RNA was extracted using TRIzol reagent (Invitrogen, Gaithersburg, MD, USA). Agilent 2100 Bioanalyzer or SMA3000 was used to measure RNA purity, concentration, and RNA integrity. A library was constructed on the Illumina HiSeq 6000 platform (San Diego, CA, USA) and sequenced according to the Illumina standard protocol with read length of PE150.

2.4. Genome Assembly

To obtain a chromosome-level whole genome assembly for *W. hoelen*, a combined approach of Illumina, PacBio, and Hi-C technology was used.

The *W. hoelen* genome was analyzed using the K-mer_freq_stat script from Biomarker Technologies in Beijing, China, to estimate its dimensions, genetic diversity, and repetitive

sequences based on the distribution of K-mer frequencies in Illumina paired-end reads. In order to analyze K-mer frequencies, SOAPec (v. 2.01) and GenomeScope (v. 2.0) software were used.

The PacBio raw reads were filtered in advance of genome assembly, and hifiiasm (version 0.12) [32] software was then used to assemble high-accuracy CCS data. To confirm the quality of the assembled genome, the BUSCO v5 assessment tool was used to assess single-copy orthologous genes [33]. The specific BUSCO gene set was fungi_odb10, which contains 290 conserved core genes of fungi.

The raw Hi-C sequence data underwent quality control procedures to ensure the acquisition of clean data, which involved the removal of adapters, non-AGCT bases at the 5' end and low-quality reads. For chromosome-level assembly, clean reads were first mapped to the *W. hoelen* genome using BWA's default parameters [34]. After mapping paired-end reads to the genome, dangling ends, self-annealing sequences, and dumped pairs were removed [35]. Afterwards, HiC-Pro v2.10.0 was used to filter invalid read pairs [36]. LACHESIS was utilized to organize scaffolds/contigs into super scaffolds [37], with the following parameters: -enzyme ^GATC -CLUSTER_MAX_LINK_DENSITY 2 -CLUSTER_MIN_RE_SITES 5 -ORDER_MIN_N_RES_IN_TRUNK 5 -ORDER_MIN_N_RES_IN_SHREDS 5 -CLUSTER_NONINFORMATIVE_RATIO.

2.5. Genome Annotation

The repeat sequence database was constructed by MITE-Hunter [38], LTR_FINDER v1.05 [39], RepeatScout v1.0.5 [40], and PILER-DF v2.4 [41]. For the final repeat sequence, PASTECClassifier [42] was used to classify the database, which was then merged with Repbase's database [43]. RepeatMasker (version 4.1.0) was used to identify and classify repeat sequences in the genome using homology searches.

A combination method was used to annotate protein-coding genes, which included homology search, de novo prediction, and transcript-based assembly. For de novo prediction, Augustus v2.4 [44], Genscan [45], GeneID v1.4 [46], GlimmerHMM v3.0.4 [47], and SNAP (version 2006-07-28) [48] were employed using its default parameters. Homology-based prediction was carried out with GeMoMa v1.3.1 [49,50]. Hisat2 v2.0.4 [51] and Stringtie v1.2.3 [51] were employed for RNA-seq based prediction. The Unigene sequences were predicted through transcriptome assembly utilizing TransDecoder v2.0 (<https://github.com/TransDecoder/TransDecoder/wiki>, accessed on 22 September 2022; The Broad Institute, Cambridge, MA, USA) and PASA v2.0.2 [52]. These prediction results were finally combined using EVM software (version 1.1.1) [53].

The proteins were subsequently annotated by Gene Ontology (GO, <http://www.geneontology.org>, accessed on 1 August 2022) [54], Kyoto Encyclopedia of Genes and Genomes (KEGG) (<http://www.genome.jp/kegg/>, accessed on 1 August 2022) [55], InterPro (<https://www.ebi.ac.uk/interpro/>, accessed on 1 August 2022) [56], Swiss-Prot (<http://www.uniprot.org>, accessed on 1 August 2022) [57], and TrEMBL (<http://www.uniprot.org/>, accessed on 1 August 2022) [57]. In order to detect reliable tRNA positions, tRNAscan-SE (version 2.0.3) [58] was used. In addition, we retrieved the other ncRNAs from the Rfam database (<http://eggnogdb.embl.de/>, accessed on 1 August 2022). Furthermore, the pathogenicity were analyzed by blast against CAZy (<https://bcbl.unl.edu/dbCAN2>, accessed on 1 August 2022), TCDB (<http://www.tcdb.org/>, accessed on 1 August 2022), PHI (<http://www.phi-base.org/index.jsp>, accessed on 1 August 2022), CYPED (<https://cyped.biocatnet.de/>, accessed on 1 August 2022), DFVF (<http://sysbio.unl.edu/DFVF/>, accessed on 1 August 2022), and EffectorP (<http://effectorp.csiro.au>, accessed on 1 August 2022) [59] databases.

2.6. Growth Test of Edible Mushroom-Forming Fungi on Pine Wood Sawdust

All the strains used for growth tests were preserved at the Institute of Edible Mushroom, Fujian Academy of Agricultural Sciences (Fuzhou, China). After the strains of *P. ostreatus* (Heiping No. 1), *P. eryngii* (Xin 23), *C. aegerita* (Gucha No. 1), *S. latifolia* (SP-C),

and *W. hoelen* (Mingling A5) were recovered, it was inoculated into freshly prepared PDA culture medium and covered with a piece of glass paper. The mycelia with the same diameters were cultivated on two types of wood substrates (pine wood and hardwood) under dark conditions at 24 °C for 1 week. Following the pulverization of the sawdust, it should be sieved through a 1 mm mesh, moistened to achieve a moisture content of approximately 60%, transferred into a test tube, and subjected to sterilization at 121 °C for a duration of 2 h. After incubating for one week at 24 °C in the absence of light, the growth of the fungi strains in test tubes was observed.

2.7. Comparative Genomics Analysis

The protein sequences of 4 related species (*P. ostreatus*, *P. eryngii*, *C. aegerita*, *S. latifolia*) were downloaded from NCBI. The comparative genomics analysis was according to our previous study [31]. In brief, the OrthoMCL (<https://orthomcl.org/>) [60] software was used to classify the protein sequences and analyze the gene families. The classification process entailed conducting a statistical analysis of the distinct gene families present in each strain, the gene families that were common among the strains, and the single-copy gene families specific to each strain. The gene families were annotated in the Pfam database (accessed on 31 October 2022) [61] and a Venn diagram was created using their statistical results. KEGG enrichment analysis for the genes was implemented by KOBAS software (version: 2.1.1, accessed on 6 February 2024) [62] with a corrected *p*-value cutoff of 0.05 to judge statistically significant enrichment. Evolutionary relationships between species were analyzed through the construction of evolutionary trees utilizing single-copy orthologous gene families and the phyML 3.0 software [63].

3. Results

3.1. Sequencing and Assembly of the Genome

The sequenced strain Minling A5 has been widely culturing in south China, especially in Fujian province. The sclerotia production of strain Minling A5 is shown in Figure 1A. By sequencing, we obtained 4.15 Gb Illumina short reads, 84.40 Gb PacBio SMRT reads, and 8.47 Gb Hi-C data (Table 1).

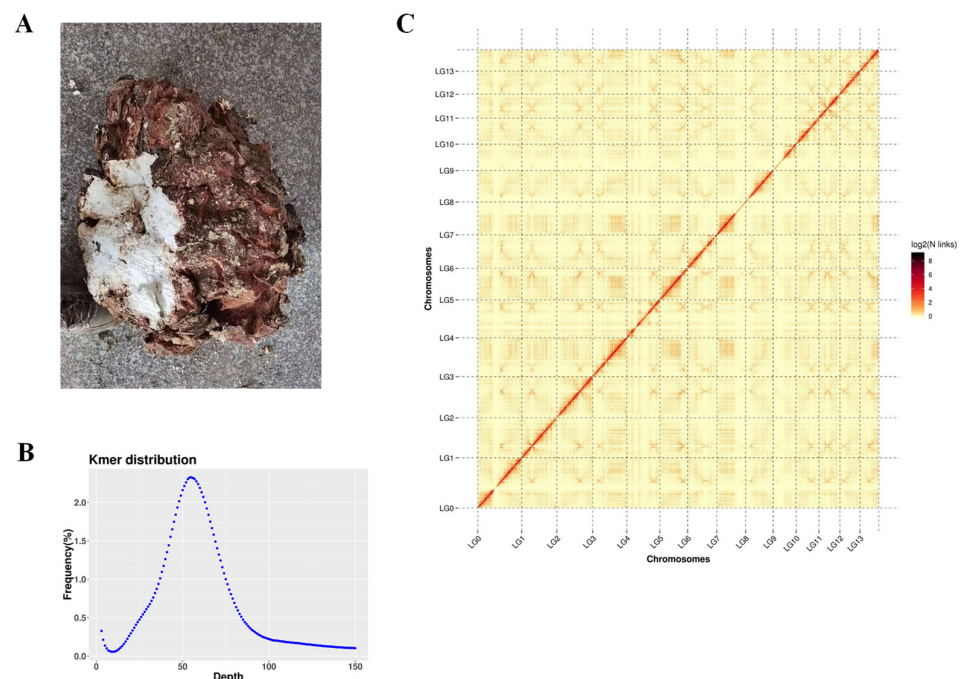


Figure 1. The genome characteristics of *W. hoelen* Minling A5. (A) The sclerotia production of strain Minling A5. (B) K-mer distribution of *W. hoelen* genome sequencing reads. (C) Hi-C interaction heatmap of all bins. The numbers on the left represent the chromosomes.

Table 1. Summary statistics of the sequencing and assembly of the *W. hoelen* strain Minling A5 genome.

Library Type	Sequencing Mode	Clean Data (Gb)	Application
Illumina	Pair end 150 bp	4.15	Genome survey and correction
PacBio	Sequel II HiFi	84.4	Genome assembly
Hi-C	Pair end 150 bp	8.47	Assisted assembly at the chromosomal level
Transcriptome	Pair end 150 bp	7.69	gene annotation
Genome assembly and scaffolding at chromosomal level			
Contig number		112	
Contig length (bp)		62,945,509	
Contig N50 (bp)		4,211,296	
Contig N90 (bp)		2,386,851	
Scaffold number		96	
Scaffold length (bp)		62,947,109	
Scaffold N50 (bp)		4,456,852	
Scaffold N90 (bp)		3,064,705	
GC content (%)		51.89	
Anchored chromosomes size (bp)		61,059,735	

In the 19-K-mer analysis, 63.81 was calculated as the dominant peak depth and 65.04 Mb was estimated as the genome size and there was approximately 0.45 heterozygosity and 47.66% repetitive sequence content (Supplementary Table S1). The genome size is slightly bigger than that of the *Wolfiporia* genus collected in NCBI (<https://www.ncbi.nlm.nih.gov/datasets/genome/?taxon=81056>, accessed on 5 June 2024) (43.77–64.44 Mb). *W. hoelen* genome showed some degree of heterozygosity and duplication, as indicated by one main peak in the 17-mer frequency distribution analysis (Figure 1B).

Based on CCS (Circular Consensus Sequencing) data, 112 contigs with 62.95 Mb total length and 4.21 Mb contig N50 length were assembled (Table 1). The corrected genome had an average GC content of 51.89%. Subsequently, as the result of the Hi-C clean data, we corrected the CCS data and scaffolded them into 14 pseudo-chromosomes (Figure 1C). The sequences mapped to the chromosomes were 61,059,735 bp long, accounting for 100% of the total length of the sequence (Table S2). All the 30 corresponding sequences were found to map to the chromosome, with a detailed distribution outlined in Table S3. Telomeres were identified as being present in the majority of chromosomes, with a predominant localization at the terminal regions of the chromosomes (Table S4).

3.2. Assessment of Genomic Integrity

By evaluating the accuracy of the initial assembly, there was 99.91% coverage of the assembled genome with 98.59% of the Illumina reads (Supplementary Table S5). According to BUSCO assay results, 94.83% of BUSCOs were complete (Table S6), indicating that the assembly integrity was adequate.

3.3. Genome Annotation

The genome contained 27.93 Mb repetitive sequences, which accounted for 44.37% of the genome (Table S7). Five major types of repeats were detected, including class I (31.74%), class II (4.29%), potential host gene (0.48%), SSR (0.02%), and unknown duplications (9.68%).

A total of 12,670 protein-coding genes were successfully yielded (Table S8), which had an average length of 2062.38 bp. An average of 6.48 exons per gene measured 242.97 bp in length, while an average of 5.48 introns measured 89.03 bp in length (Table S9). Among

these genes, 86.53% (10,963) could be functionally annotated by at least one of the KOG, GO, Pfam, Swissprot, TrEMBL, NR, and KEGG databases (Table S10). There were 240 transfer RNAs, 97 ribosomal RNAs, and 103 other non-coding RNAs identified in the *W. hoelen* genome (Table S11). In addition, 755 pseudogenes were also identified with an average length of 2665.51 bp (Table S12).

There were 398 CAZymes identified in *W. hoelen* (Table S14), which was similar to the species *W. cocos* (205) and *W. hoelen* (214) in the same genus [24]. These CAZymes included 167 glycoside hydrolases (GHs), 76 glycosyl transferases (GTs), three polysaccharide lyases (PLs), 74 carbohydrate esterases (CEs), 34 carbohydrate-binding modules (CBMs), and 44 auxiliary activities (AAs). Additionally, there were 100, 2837, 519, and 2068 genes annotated by TCDB, PHI, P450, and DFVF databases, respectively (Table S13).

3.4. The Growth of Five Edible Mushroom-Forming Fungi on Pine Wood Sawdust

One of the characteristics of *W. hoelen* is that it can grow in association with pine trees. To understand the mechanism underlying this characteristic, we first inoculated five different edible mushroom-forming fungi (*P. ostreatus*, *P. eryngii*, *C. aegerita*, *S. latifolia*, and *W. hoelen*) into pine sawdust and broad-leaved wood sawdust culture media. The genome information of these strains is shown in Table S15. As shown in Figure 2, the growth of *S. latifolia* and *W. hoelen* on pine sawdust was similar to that on broad-leaved wood sawdust, while the growth of *P. ostreatus*, *P. eryngii*, and *C. aegerita* was slower than that on broad-leaved wood sawdust. Meanwhile, the growth rate of *W. hoelen* was five to six times that of *S. latifolia*, instead of two to three times as previously reported [24].

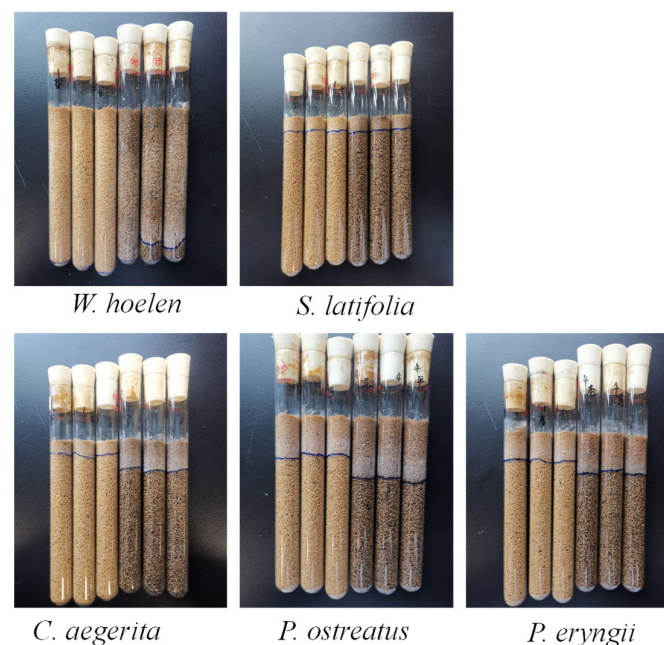


Figure 2. The growth of five different edible mushroom-forming fungi (*P. ostreatus*, *P. eryngii*, *C. aegerita*, *S. latifolia*, and *W. hoelen*) on pine sawdust and broad-leaved wood sawdust culture media. The mycelium of uniform diameter was cultivated on two distinct wood substrates (pine wood and hardwood) under dark conditions at a temperature of 24 °C for 1 week.

3.5. Employing Comparative Genomics Analysis to Identify Genes Associated with Pine Utilization

S. latifolia and *W. hoelen* grew similarly on pine sawdust and broad-leaved wood sawdust, while *P. ostreatus*, *P. eryngii*, and *C. aegerita* grew more slowly on pine sawdust compared to broad-leaved wood sawdust. Thus, we hypothesised that the genes unique to *S. latifolia* and *W. hoelen* play a significant role in the ability to thrive on fresh pine wood. Analyzing the structural and functional properties of orthogroups, we found 4694 common orthogroups in the genome of selected five species, while 2517 orthogroups were specific

to *W. hoelen* (Figure 3A and Table 2). Importantly, there were 645 orthogroups specifically common in *W. hoelen* and *S. latifolia* (Table S16), which were significantly enriched in 12 pathways (corrected p value < 0.05), including steroid biosynthesis, biosynthesis of antibiotics, and tyrosine metabolism (Figure 3B and Table S17). A maximum likelihood (ML) phylogenetic analysis was conducted for five mushroom-forming fungal species using shared single-copy orthologous genes. The results revealed that *W. hoelen* was more closely related to *S. latifolia* than to the other three species (Figure 3C).

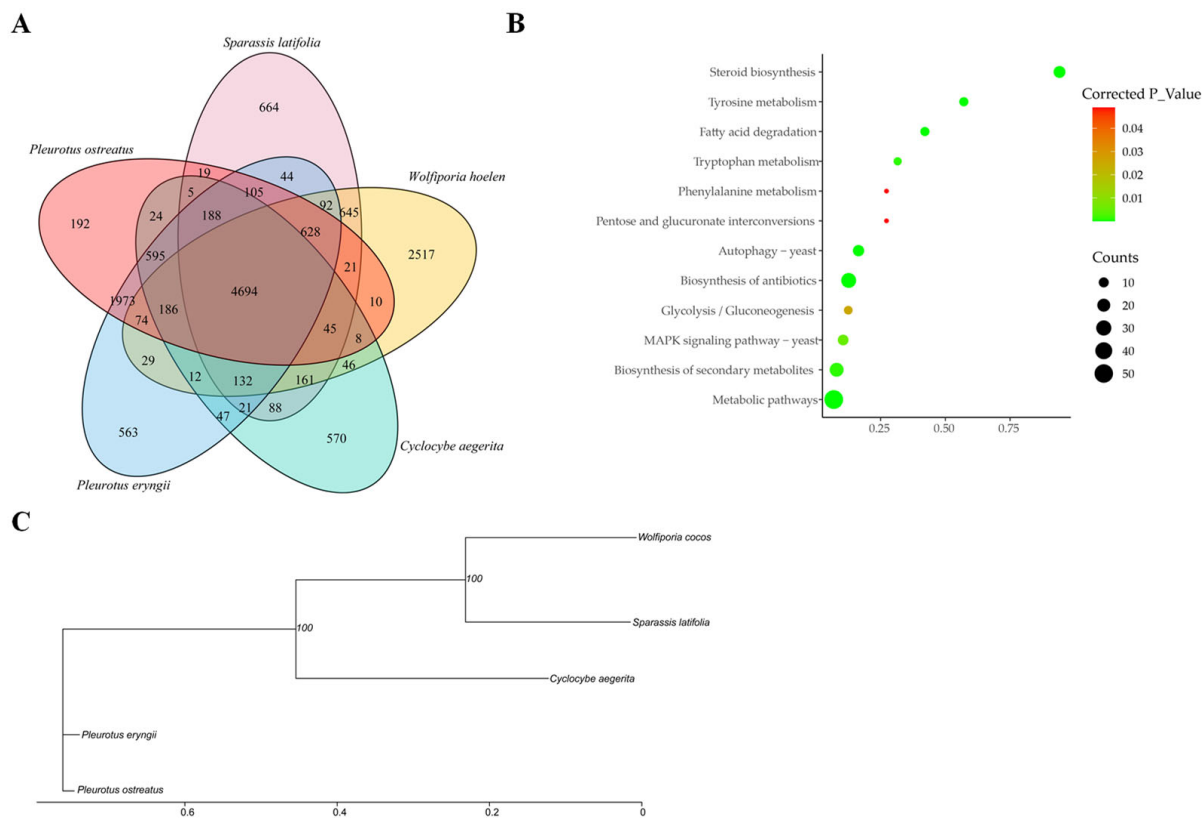


Figure 3. Comparison of the genomes of the five mushroom-forming fungi. (A) Gene family annotation venn diagram. (B) KEGG enrichment of the genes in orthogroups specifically common in *W. hoelen* and *S. latifolia*. The enrichment plot of significantly enriched pathways (corrected p value < 0.05) was drawn by an online website [64] (<https://www.bic.ac.cn/ImageGP/>, accessed on 3 April 2024). (C) Phylogenetic tree of the five mushroom-forming fungi. The single-copy orthologous genes from OrthoMCL were used to construct a phylogenetic tree.

Table 2. The gene family statistics of the five selected mushroom-forming fungi. The OrthoMCL software was used to classify the protein sequences and analyze the gene families.

Species	Total Gene Number	Cluster Gene Number	Total Gene Family Number	Unique Gene Family Number
<i>C. aegerita</i>	13,341	10,574	6822	570
<i>P. eryngii</i>	15,954	12,632	9383	563
<i>P. ostreatus</i>	11,718	10,504	8769	192
<i>S. latifolia</i>	15,016	12,972	7552	664
<i>W. hoelen</i>	12,670	11,763	9300	2517

4. Discussion

As the world’s largest conifer genus, *Pinus* is arguably the most important tree species [65], as well as in China. Pine wilt disease (PWD) is a migratory endoparasitic

nematode that infects the stem of conifer trees and affects mainly species of the genus *Pinus*, which is caused by the pine wood nematode (PWN) *Bursaphelenchus xylophilus*. In Eastern Asia and Western Europe countries, several pine species have shown high susceptibility to PWD [66]. PWD has been causing huge losses to the Chinese ecological environment, natural landscape, and social economy [67]. To limit the spreading of PWD, several measures were implemented, including detecting symptomatic trees, eliminating them, and burning or fumigating the wood. This has led to a large number of infected trees and tree stumps that need to be treated.

Basidiomycetous fungi are largely responsible for the biodegradation of woody biomass [68]. However, there are only a few species of fungi can directly grow on fresh pine sawdust, including *W. hoelen* [24], Schizophyllaceae [69], *Rhodonia placenta* [70], and *S. latifolia* [9]. *W. hoelen* and *S. latifolia* were cultivated on a large scale in China with pine sawdust substrates. Our previous study screened out some important genes in the fresh pine wood decay by *S. latifolia* through RNA-seq combined weighted gene co-expression analysis (WGCNA) [9]. However, the current industrialized cultivation mode of *S. latifolia* requires transporting pine trees to the factory, leading to a high risk of pine wood nematode spread. Additionally, the growth of *W. hoelen* on pine sawdust substrates was much faster than that of *S. latifolia* (Figure 2). Therefore, cultivating *W. hoelen* is a more effective way to utilize epidemic wood.

“Fuling” is an important crude drug in traditional Chinese medicine with activities of anti-tumor, anti-oxidant, anti-bacterial, immunomodulation, anti-inflammatory, and liver and kidney protection [71]. The market demand for “Fuling” keeps growing [72]. Additionally, cultivating *W. hoelen* on nematode-affected pine wood can kill all nematodes and *Monochamus* in 60 days, which was a workable way of utilizing nematode-affected pine wood on the mountain and harmless utilization [73]. So, cultivating *W. hoelen* is the best way to utilize nematode-affected pine wood. In recent years, *W. hoelen* strain Minling A5 has been widely culturing in south China, especially in Fujian province, and the primary raw materials utilized are pine stumps remaining post-logging [20]. This illustrates the significant contribution of this strain to the utilization of resources in pine trees affected by epidemics and pine stumps.

Information about an organism’s genomic sequence can contribute to understanding some of its biological aspects. However, no genetic information of *W. hoelen* strain Minling A5 has been reported until now. In this study, we assembled the first high-quality genome of *W. hoelen* strain Minling A5. There are several genome sequencing reports about the *Wolfiporia* genus [22–25], the genome characteristics of *Wolfiporia* genus are shown in Table 3. The genome size of *W. hoelen* strain Minling A5 was 62.95 Mb, which is very close to the genome of a strain of WCLT (62 Mb). But it is bigger than the strain of CGMCC 5.78 (50.6 Mb), KMCC03342 (55.5 Mb), and GDMCC 5.219 (60.2 Mb), while smaller than the strain of SS20 (64.44 Mb). The scaffolds number of Minling A5 (96) was close to the report of strain SS20 (78), while much less than the other reports. Another significant indicator, N50 of scaffolds, our result was much longer than the others. The number of protein-coding genes in Minling A5 (12,670) was also more than that of other reports, except for the KMCC03342 strain. Additionally, we obtained a chromosome-scale assembly of the *W. hoelen* strain Minling A5 consisting of 14 chromosomes (Figure 1C and Table 3). There were two reports on chromosome level assemble in *Wolfiporia* genus [23,24]. One is a 64.44 Mb homokaryotic genome of *W. hoelen* with 14 chromosomes and assembled with a Contig N50 length of 3.76 Mb [24]. Another one is 62 Mb haploid genome of a cultivated *W. cocos* strain (WCLT) in China and also showed 14 chromosomes under microscopy and Hi-C analysis. The number of chromosomes was consistent with our result (Figure 1). These findings indicate that the genome assembly we have achieved exhibits a high level of quality.

Table 3. Genome characteristics of *Wolfiporia* genus. The blank columns were due to the lack of corresponding data provided in the references.

Strains	Sequencing Strategy	Genome Size (Mb)	Number of Scaffolds	N50 of Scaffolds (kb)	Anchored to Chromosome (Mb)	Number of Protein-Coding Genes	Average Gene Length (bp)	Percentage of Repeat Sequences (%)	Transposable Elements (%)	GC Content (%)	Reference
CGMCC 5.78	HiSeq 2000 Illumina and a fosmid-to-fosmid strategy	50.6	351	835		10,908	1829	-	33.5	51.7	[23]
WCLT	HiSeq2500 Illumina and SMRT technology on the PacBio	62	145	1599.1	61.127	11,906	1332.76	46.6	-	51.86	[21]
SS20	Novaseq6000 Illumina and SMRT technology on the PacBio	64.44	78	3760	58.26	10,567	2004	48.56	46.26	50.15	[22]
KMCC03342	Oxford Nanopore and SMRT technology on the PacBio	55.5				14,296				52.2	[20]
GDMCC 5.219	PacBio Sequel II	60.2	183	1300						52	https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_034769205.1/ , accessed on 5 June 2024
Minling A5	Novaseq6000 Illumina and SMRT technology on the PacBio	62.95	96	4456.8	61.06	12,670	2062.38	44.37		51.89	This study

Wolfiporia is a fungus that typically colonizes the roots of pine trees [24]. The wood decay mechanism of *W. cocos* was examined by transcriptome and secretome, which found that the differentially expressed genes were involved in iron homeostasis, iron reduction, and extracellular peroxide generation [21]. In lodgepole pine, 207 genes demonstrated a 4-fold increase in transcript levels compared to those observed with glucose. However, the presence of a large number of differentially expressed genes poses a challenge in identifying the key genes of interest. Comparative genomics plays a crucial role in elucidating evolutionary dynamics across species, facilitating the discovery of novel genes and serving as a foundation for subsequent functional genomics investigations [74,75].

Sawdust from fresh pine trees is not commonly utilized to cultivate edible mushrooms [6]. However, *W. hoelen* and *S. latifolia* can directly grow on fresh pine wood sawdust substrate [7–9,21,76]. Therefore, we hypothesized that the genes specific to *S. latifolia* and *W. hoelen* are important for growing on fresh pine wood. The integration of phenotype analysis and comparative genomics analysis can facilitate the identification of crucial genes associated with pine wood decay. Through the structural and functional annotation analysis of orthogroups, we found 645 specific common in *W. hoelen* and *S. latifolia* (Table S16) and the genes in these orthogroups were significantly enriched in 12 KEGG pathways, including steroid biosynthesis, biosynthesis of antibiotics, and tyrosine metabolism (Figure 3B and Table S17). In our results of comparative genomics analysis, the gene family of cytochrome P450 was specifically common in *W. hoelen* and *S. latifolia* (Table S16), while there were also differentially expressed cytochrome P450 genes in previous reports on the pine wood decay mechanism of *W. hoelen* [21] and *S. latifolia* [9]. This suggests that cytochrome P450 is a significant factor in the degradation mechanism of pine wood. Compared to our previous report on the mechanisms of *S. latifolia* decay fresh pine wood sawdust substrate [9], there were several common pathways, including biosynthesis of antibiotics, tyrosine metabolism, MAPK signaling pathway–yeast, and tryptophan metabolism.

Compounds derived from the Pinus species have demonstrated antimicrobial properties [77,78]. A study found that rosin from pine trees makes it difficult for fungi to grow on coniferous wood [6], which has been shown to have antifungal activities [79,80]. Resin acids, characterized by the chemical formula $C_{19}H_{29}COOH$, comprise a significant portion, up to 95 wt.% of rosin [81]. In the results of comparative genomics analysis, the genes in orthogroups specific to *W. hoelen* and *S. latifolia* were enriched in the KEGG pathways related to fatty acid metabolism, including steroid biosynthesis, fatty acid degradation, and glycolysis/gluconeogenesis (Table S16). Our previous study also found that lipase genes were differentiated and expressed in *S. latifolia* induced by fresh pine sawdust [9]. So, we think that if a fungus demonstrates the ability to metabolize rosin or exhibit resistance to its antifungal properties, there is a high probability that it will thrive on fresh pine sawdust.

In conclusion, a combination of Illumina, PacBio Sequel II, and Hi-C sequencing technologies was used to sequence and assemble the genome of *W. hoelen* strain Minling A5 and provided a 62.95 Mb genome sequence consisting of 112 contigs with an N50 length of 4.21 Mb. The capacity of *W. hoelen* to utilize fresh pine wood was investigated through comparative genomics analysis, revealing that genes within specific orthogroups enriched in 12 pathways, such as steroid biosynthesis, biosynthesis of antibiotics, fatty acid degradation, and glycolysis/gluconeogenesis, and tyrosine metabolism, played a significant role (Figure 4). The genome sequencing and comparative genomics analysis of *W. hoelen* conducted in this study mark a notable advancement in elucidating the pine wood degradation mechanism of this fungus. The findings of this research hold promise for making substantial contributions to the advancement of sustainable bioenergy production and environmental protection.

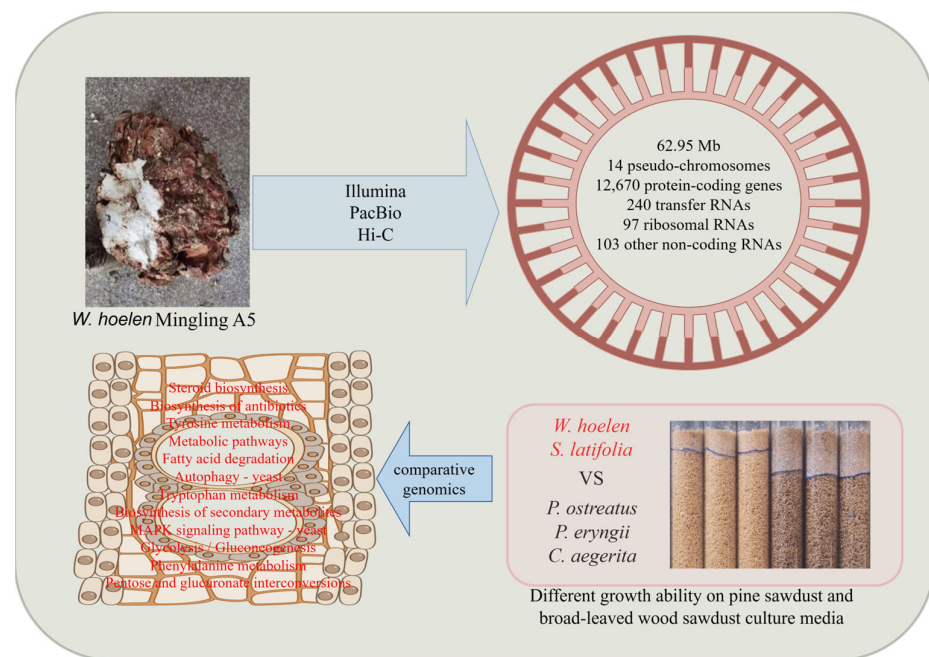


Figure 4. Proposed model of pine wood decay mechanism of *W. hoelen*. This model was drawn by Figdraw.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/horticulturae10070703/s1>, Table S1. The data statistics of 17-mer analysis and heterozygosity of genome. Table S2. The sequence distribution of each chromosome using Hi-C technology. Table S3. The parameters of 3D-DNA analysis are based on Hi-C reads. Table S4. The telomere distributions of *W. hoelen* genome. Table S5. The consistency assessment of sequencing reads to assembled genomes. Table S6. BUSCO analysis of the assembled genome. Table S7. Statistics of transposable elements in *W. hoelen* genome. Table S8. Statistics of predicted protein-coding gene number in *W. hoelen* genome. Table S9. Analysis of predicted protein-coding genes in *W. hoelen* genome. Table S10. Statistics of annotated gene number in *W. hoelen* genome by different databases and methods. Table S11. Statistics of non-coding RNAs in *W. hoelen* genome. Table S12. Statistics of pseudogenes in *W. hoelen* genome. Table S13. Statistics of specific databases annotation in *W. hoelen* genome. Table S14. Statistics of CAZyme databases annotation in *W. hoelen* genome. Table S15. Websites of published genomes using comparative genomic analysis. Table S16. The orthogroups are specifically common in *W. hoelen* and *S. latifolia*. Table S17. KEGG enrichment of the genes in orthogroups specific common in *W. hoelen* and *S. latifolia*.

Author Contributions: Conceptualization, C.Y. and L.M.; Formal analysis, X.J., H.L. and L.M.; Funding acquisition, C.Y. and L.M.; Investigation, X.L. and C.L.; Methodology, D.X. and X.J.; Project administration, X.J. and Y.L.; Software, C.Y.; Supervision, L.M.; Validation, D.X. and X.J.; Writing—original draft, C.Y.; Writing—review & editing, C.Y. and L.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the 5511 Collaborative Innovation Project of Fujian Province (XTCXGC2021007), the project from Fujian Academy of Agricultural Sciences (CXTD2021016-2), the Special Fund for Scientific Research in the Public Interest of Fujian Province (2021R1035003), and Seed Industry Innovation and Industrialization Project of Fujian Province (zycxny2021011).

Data Availability Statement: The sequencing data (Full-length transcriptome, Hi-C, Illumina and PacBio) have been deposited in SRA (Sequence Read Archive) database under BioProject PR-JNA1066959.

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

References

- Li, M.; Li, H.; Sheng, R.C.; Sun, H.; Sun, S.H.; Chen, F.M. The First Record of *Monochamus saltuarius* (Coleoptera; Cerambycidae) as Vector of *Bursaphelenchus xylophilus* and Its New Potential Hosts in China. *Insects* **2020**, *11*, 636. [\[CrossRef\]](#) [\[PubMed\]](#)
- Hu, L.J.; Wu, X.Q.; Li, H.Y.; Wang, Y.C.; Huang, X.; Wang, Y.; Li, Y. BxCDP1 from the pine wood nematode *Bursaphelenchus xylophilus* is recognized as a novel molecular pattern. *Mol. Plant Pathol.* **2020**, *21*, 923–935. [\[CrossRef\]](#) [\[PubMed\]](#)
- Zhou, L.F.; Chen, F.M.; Xie, L.Y.; Pan, H.Y.; Ye, J.R. Genetic diversity of pine-parasitic nematodes *Bursaphelenchus xylophilus* and *Bursaphelenchus mucronatus* in China. *Forest Pathol.* **2017**, *47*, e12334. [\[CrossRef\]](#)
- Xu, Q.; Zhang, X.; Li, J.; Ren, J.; Ren, L.; Luo, Y. Pine Wilt Disease in Northeast and Northwest China: A Comprehensive Risk Review. *Forests* **2023**, *14*, 174. [\[CrossRef\]](#)
- Zhao, J.; Huang, J.; Yan, J.; Fang, G. Economic Loss of Pine Wood Nematode Disease in Mainland China from 1998 to 2017. *Forests* **2020**, *11*, 1042. [\[CrossRef\]](#)
- Croan, S.C. Conversion of conifer wastes into edible and medicinal mushrooms. *For. Prod. J.* **2004**, *54*, 68–76.
- Ma, L.; Lin, Y.Q.; Yang, C.; Ying, Z.H.; Jiang, X.L. Production of liquid spawn of an edible mushroom, *Sparassis latifolia* by submerged fermentation and mycelial growth on pine wood sawdust. *Sci. Hortic.* **2016**, *209*, 22–30. [\[CrossRef\]](#)
- Xiao, D.L.; Ma, L.; Yang, C.; Lin, Y.Q. Transcriptome analysis of *Sparassis latifolia* cultivated with different carbon sources. *Microbiol. China* **2019**, *046*, 1654–1661.
- Yang, C.; Ma, L.; Xiao, D.; Liu, X.; Jiang, X.; Lin, Y. Comparative transcriptomics reveals unique pine wood decay strategies in the *Sparassis latifolia*. *Sci. Rep.* **2022**, *12*, 19875. [\[CrossRef\]](#)
- Stalpers, J.A.; Redhead, S.A.; May, T.W.; Rossman, A.Y.; Crouch, J.A.; Cubeta, M.A.; Dai, Y.C.; Kirschner, R.; Langer, G.J.; Larsson, K.H.; et al. Competing sexual-asexual generic names in *Agaricomycotina* (Basidiomycota) with recommendations for use. *IMA Fungus* **2021**, *12*, 22. [\[CrossRef\]](#)
- Liu, J.; Yu, J.; Peng, X. *Poria cocos* Polysaccharides Alleviates Chronic Nonbacterial Prostatitis by Preventing Oxidative Stress, Regulating Hormone Production, Modifying Gut Microbiota, and Remodeling the DNA Methylome. *J. Agric. Food Chem.* **2020**, *68*, 12661–12670. [\[CrossRef\]](#) [\[PubMed\]](#)
- Yang, J.; Dong, X.; Li, B.; Chen, T.; Yu, B.; Wang, X.; Dou, X.; Peng, B.; Hu, Q. *Poria cocos* polysaccharide-functionalized graphene oxide nanosheet induces efficient cancer immunotherapy in mice. *Front. Bioeng. Biotechnol.* **2022**, *10*, 1050077. [\[CrossRef\]](#) [\[PubMed\]](#)
- Duan, Y.; Huang, J.; Sun, M.; Jiang, Y.; Wang, S.; Wang, L.; Yu, N.; Peng, D.; Wang, Y.; Chen, W.; et al. *Poria cocos* polysaccharide improves intestinal barrier function and maintains intestinal homeostasis in mice. *Int. J. Biol. Macromol.* **2023**, *249*, 125953. [\[CrossRef\]](#)
- Lee, S.R.; Lee, S.; Moon, E.; Park, H.J.; Park, H.B.; Kim, K.H. Bioactivity-guided isolation of anti-inflammatory triterpenoids from the sclerotia of *Poria cocos* using LPS-stimulated Raw264.7 cells. *Bioorg. Chem.* **2017**, *70*, 94–99. [\[CrossRef\]](#) [\[PubMed\]](#)
- Ding, X.; Li, S.; Tian, M.; Yang, P.; Ding, Y.; Wang, Y.; Duan, G.; Zhang, D.; Chen, B.; Tan, Q. Facile preparation of a novel nanoemulsion based hyaluronic acid hydrogel loading with *Poria cocos* triterpenoids extract for wound dressing. *Int. J. Biol. Macromol.* **2023**, *226*, 1490–1499. [\[CrossRef\]](#) [\[PubMed\]](#)
- Chao, C.L.; Kuo, H.P.; Huang, H.W.; Cheng, M.Y.; Chao, H.F.; Lu, S.M.; Lin, H.C.; Wang, C.J.; Chang, T.C.; Wu, C.R. *Poria cocos* Lanostane Triterpenoids Extract Promotes Collagen and Hyaluronic Acid Production in D-Galactose-Induced Aging Rats. *Life* **2023**, *13*, 2130. [\[CrossRef\]](#) [\[PubMed\]](#)
- Chao, C.L.; Huang, H.W.; Su, M.H.; Lin, H.C.; Wu, W.M. The Lanostane Triterpenoids in *Poria cocos* Play Beneficial Roles in Immunoregulatory Activity. *Life* **2021**, *11*, 111. [\[CrossRef\]](#)
- Zhu, W.; Liu, Y.; Tang, J.; Liu, H.; Jing, N.; Li, F.; Xu, R.; Shu, S. Functional Analysis of Sterol O-Acyltransferase Involved in the Biosynthetic Pathway of Pachymic Acid in *Wolfiporia cocos*. *Molecules* **2021**, *27*, 143. [\[CrossRef\]](#)
- Zeng, G.; Li, Z.; Zhao, Z. Metabolome analysis of key genes for synthesis and accumulation of triterpenoids in *Wolfiporia cocos*. *Sci. Rep.* **2022**, *12*, 1574. [\[CrossRef\]](#)
- Cai, D.F.; Chen, D.H.; Zheng, P.W.; Yang, J.; Huang, X.; Lin, Y.Q. High yield cultivation and management techniques of Minling A5 on pine stump. *Edible Fungi China* **2016**, *35*, 74–77. [\[CrossRef\]](#)
- Gaskell, J.; Blanchette, R.A.; Stewart, P.E.; BonDurant, S.S.; Adams, M.; Sabat, G.; Kersten, P.; Cullen, D. Transcriptome and Secretome Analyses of the Wood Decay Fungus *Wolfiporia cocos* Support Alternative Mechanisms of Lignocellulose Conversion. *Appl. Environ. Microbiol.* **2016**, *82*, 3979–3987. [\[CrossRef\]](#) [\[PubMed\]](#)
- Kim, B.; Min, B.; Han, J.G.; Park, H.; Baek, S.; Jeong, S.; Choi, I.G. Draft Genome Sequence of the Reference Strain of the Korean Medicinal Mushroom *Wolfiporia cocos* KMCC03342. *Mycobiology* **2022**, *50*, 254–257. [\[CrossRef\]](#) [\[PubMed\]](#)
- Cao, S.; Yang, Y.; Bi, G.; Nelson, D.; Hu, S.; Makunga, N.P.; Yu, B.; Liu, X.; Li, X.; Hu, X. Genomic and Transcriptomic Insight of Giant Sclerotium Formation of Wood-Decay Fungi. *Front. Microbiol.* **2021**, *12*, 746121. [\[CrossRef\]](#) [\[PubMed\]](#)
- Li, S.; Meng, G.; Dong, C. Homokaryotic High-Quality Genome Assembly of Medicinal Fungi *Wolfiporia hoelen* Reveals Auto-Regulation and High-Temperature Adaption of Probable Two-Speed Genome. *Int. J. Mol. Sci.* **2022**, *23*, 10484. [\[CrossRef\]](#)
- Luo, H.; Qian, J.; Xu, Z.; Liu, W.; Xu, L.; Li, Y.; Xu, J.; Zhang, J.; Xu, X.; Liu, C.; et al. The *Wolfiporia cocos* Genome and Transcriptome Shed Light on the Formation of Its Edible and Medicinal Sclerotium. *Genom. Proteom. Bioinform.* **2020**, *18*, 455–467. [\[CrossRef\]](#) [\[PubMed\]](#)

26. Floudas, D.; Binder, M.; Riley, R.; Barry, K.; Blanchette, R.A.; Henrissat, B.; Martínez, A.T.; Otillar, R.; Spatafora, J.W.; Yadav, J.S.; et al. The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science* **2012**, *336*, 1715–1719. [[CrossRef](#)] [[PubMed](#)]
27. Biel, S.W.; Parrish, F.W. Isolation of DNA from fungal mycelia and sclerotia without use of density gradient ultracentrifugation. *Anal. Biochem.* **1986**, *154*, 21–25. [[CrossRef](#)] [[PubMed](#)]
28. Ranallo-Benavidez, T.R.; Jaron, K.S.; Schatz, M.C. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nat. Commun.* **2020**, *11*, 1432. [[CrossRef](#)] [[PubMed](#)]
29. He, K.; Zhao, L.; Yuan, Z.; Canario, A.; Liu, Q.; Chen, S.; Guo, J.; Luo, W.; Yan, H.; Zhang, D.; et al. Chromosome-level genome assembly of largemouth bass (*Micropterus salmoides*) using PacBio and Hi-C technologies. *Sci. Data* **2022**, *9*, 482. [[CrossRef](#)]
30. Korlach, J.; Bjornson, K.P.; Chaudhuri, B.P.; Cicero, R.L.; Flusberg, B.A.; Gray, J.J.; Holden, D.; Saxena, R.; Wegener, J.; Turner, S.W. Real-time DNA sequencing from single polymerase molecules. *Methods Enzymol.* **2010**, *472*, 431–455. [[CrossRef](#)]
31. Yang, C.; Ma, L.; Xiao, D.L.; Liu, X.Y.; Jiang, X.L.; Ying, Z.H.; Lin, Y.Q. Chromosome-scale assembly of the *Sparassis latifolia* genome obtained using long-read and Hi-C sequencing. *G3 Genes | Genomes | Genet.* **2021**, *11*, jkab173. [[CrossRef](#)] [[PubMed](#)]
32. 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](#)]
33. Manni, M.; Berkeley, M.R.; Seppey, M.; Zdobnov, E.M. BUSCO: Assessing Genomic Data Quality and Beyond. *Curr. Protoc.* **2021**, *1*, e323. [[CrossRef](#)]
34. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [[CrossRef](#)]
35. Roach, M.J.; Schmidt, S.A.; Borneman, A.R. Purge Haplotigs: Allelic contig reassignment for third-gen diploid genome assemblies. *BMC Bioinform.* **2018**, *19*, 460. [[CrossRef](#)]
36. Servant, N.; Varoquaux, N.; Lajoie, B.R.; Viara, E.; Chen, C.J.; Vert, J.P.; Heard, E.; Dekker, J.; Barillot, E. HiC-Pro: An optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* **2015**, *16*, 259. [[CrossRef](#)]
37. 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](#)]
38. Han, Y.; Wessler, S.R. MITE-Hunter: A program for discovering miniature inverted-repeat transposable elements from genomic sequences. *Nucleic Acids Res.* **2010**, *38*, e199. [[CrossRef](#)] [[PubMed](#)]
39. Xu, Z.; Wang, H. LTR_FINDER: An efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.* **2007**, *35*, W265–W268. [[CrossRef](#)]
40. Price, A.L.; Jones, N.C.; Pevzner, P.A. De novo identification of repeat families in large genomes. *Bioinformatics* **2005**, *21* (Suppl. S1), i351–i358. [[CrossRef](#)]
41. Edgar, R.C.; Myers, E.W. PILER: Identification and classification of genomic repeats. *Bioinformatics* **2005**, *21* (Suppl. S1), i152–i158. [[CrossRef](#)] [[PubMed](#)]
42. Wicker, T.; Sabot, F.; Hua-Van, A.; Bennetzen, J.L.; Capy, P.; Chalhoub, B.; Flavell, A.; Leroy, P.; Morgante, M.; Panaud, O.; et al. A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* **2007**, *8*, 973–982. [[CrossRef](#)] [[PubMed](#)]
43. Jurka, J.; Kapitonov, V.V.; Pavlicek, A.; Klonowski, P.; Kohany, O.; Walichiewicz, J. Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet. Genome Res.* **2005**, *110*, 462–467. [[CrossRef](#)] [[PubMed](#)]
44. Stanke, M.; Waack, S. Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics* **2003**, *19* (Suppl. S2), ii215–ii225. [[CrossRef](#)] [[PubMed](#)]
45. Burge, C.; Karlin, S. Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* **1997**, *268*, 78–94. [[CrossRef](#)] [[PubMed](#)]
46. Alioto, T.; Blanco, E.; Parra, G.; Guigó, R. Using geneid to Identify Genes. *Curr. Protoc. Bioinform.* **2018**, *64*, e56. [[CrossRef](#)] [[PubMed](#)]
47. Majoros, W.H.; Pertea, M.; Salzberg, S.L. TigrScan and GlimmerHMM: Two open source ab initio eukaryotic gene-finders. *Bioinformatics* **2004**, *20*, 2878–2879. [[CrossRef](#)] [[PubMed](#)]
48. Korf, I. Gene finding in novel genomes. *BMC Bioinform.* **2004**, *5*, 59. [[CrossRef](#)] [[PubMed](#)]
49. Keilwagen, J.; Wenk, M.; Erickson, J.L.; Schattat, M.H.; Grau, J.; Hartung, F. Using intron position conservation for homology-based gene prediction. *Nucleic Acids Res.* **2016**, *44*, e89. [[CrossRef](#)]
50. Keilwagen, J.; Hartung, F.; Grau, J. GeMoMa: Homology-Based Gene Prediction Utilizing Intron Position Conservation and RNA-seq Data. *Methods Mol. Biol.* **2019**, *1962*, 161–177. [[CrossRef](#)]
51. Pertea, M.; Kim, D.; Pertea, G.M.; Leek, J.T.; Salzberg, S.L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* **2016**, *11*, 1650–1667. [[CrossRef](#)] [[PubMed](#)]
52. Campbell, M.A.; Haas, B.J.; Hamilton, J.P.; Mount, S.M.; Buell, C.R. Comprehensive analysis of alternative splicing in rice and comparative analyses with Arabidopsis. *BMC Genomics* **2006**, *7*, 327. [[CrossRef](#)] [[PubMed](#)]
53. 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](#)] [[PubMed](#)]

54. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T.; et al. Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **2000**, *25*, 25–29. [\[CrossRef\]](#) [\[PubMed\]](#)
55. Ogata, H.; Goto, S.; Sato, K.; Fujibuchi, W.; Bono, H.; Kanehisa, M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **1999**, *27*, 29–34. [\[CrossRef\]](#) [\[PubMed\]](#)
56. Jones, P.; Binns, D.; Chang, H.Y.; Fraser, M.; Li, W.; McAnulla, C.; McWilliam, H.; Maslen, J.; Mitchell, A.; Nuka, G.; et al. InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **2014**, *30*, 1236–1240. [\[CrossRef\]](#) [\[PubMed\]](#)
57. Boeckmann, B.; Bairoch, A.; Apweiler, R.; Blatter, M.C.; Estreicher, A.; Gasteiger, E.; Martin, M.J.; Michoud, K.; O'Donovan, C.; Phan, I.; et al. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res.* **2003**, *31*, 365–370. [\[CrossRef\]](#) [\[PubMed\]](#)
58. 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. [\[CrossRef\]](#) [\[PubMed\]](#)
59. Sperschneider, J.; Gardiner, D.M.; Dodds, P.N.; Tini, F.; Covarelli, L.; Singh, K.B.; Manners, J.M.; Taylor, J.M. EffectorP: Predicting fungal effector proteins from secretomes using machine learning. *New Phytol.* **2016**, *210*, 743–761. [\[CrossRef\]](#)
60. Li, L.; Stoeckert, C.J., Jr.; Roos, D.S. OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Res.* **2003**, *13*, 2178–2189. [\[CrossRef\]](#)
61. Mistry, J.; Chuguransky, S.; Williams, L.; Qureshi, M.; Salazar, G.A.; Sonnhammer, E.L.L.; Tosatto, S.C.E.; Paladin, L.; Raj, S.; Richardson, L.J.; et al. Pfam: The protein families database in 2021. *Nucleic Acids Res.* **2021**, *49*, D412–D419. [\[CrossRef\]](#) [\[PubMed\]](#)
62. Bu, D.; Luo, H.; Huo, P.; Wang, Z.; Zhang, S.; He, Z.; Wu, Y.; Zhao, L.; Liu, J.; Guo, J.; et al. KOBAS-i: Intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. *Nucleic Acids Res.* **2021**, *49*, W317–W325. [\[CrossRef\]](#) [\[PubMed\]](#)
63. 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. [\[CrossRef\]](#) [\[PubMed\]](#)
64. Chen, T.; Liu, Y.-X.; Huang, L. ImageGP: An easy-to-use data visualization web server for scientific researchers. *iMeta* **2022**, *1*, e5. [\[CrossRef\]](#) [\[PubMed\]](#)
65. Farjon, A.; Filer, D. *An Atlas of the World's Conifers: An Analysis of Their Distribution, Biogeography, Diversity and Conservation Status*; Brill: Leiden, The Netherlands, 2013.
66. Webster, J.; Mota, M. Pine wilt disease: Global issues, trade and economic impact. In *Pine Wilt Disease: A Worldwide Threat to Forest Ecosystems*; Springer: Dordrecht, The Netherlands, 2008.
67. Hao, Z.; Huang, J.; Li, X.; Sun, H.; Fang, G. A multi-point aggregation trend of the outbreak of pine wilt disease in China over the past 20 years. *Forest Ecol. Manag.* **2022**, *505*, 119890. [\[CrossRef\]](#)
68. Goodell, B.; Nicholas, D.D.; Schultz, T.P. (Eds.) *Wood Deterioration and Preservation: Advances in Our Changing World*; American Chemical Society: Washington, DC, USA, 2003.
69. Almási, É.; Sahu, N.; Krizsán, K.; Bálint, B.; Kovács, G.M.; Kiss, B.; Cseklye, J.; Drula, E.; Henrissat, B.; Nagy, I.; et al. Comparative genomics reveals unique wood-decay strategies and fruiting body development in the Schizophyllaceae. *New Phytol.* **2019**, *224*, 902–915. [\[CrossRef\]](#) [\[PubMed\]](#)
70. Belt, T.; Altgen, M.; Makela, M.; Hanninen, T.; Rautkari, L. Cellular level chemical changes in Scots pine heartwood during incipient brown rot decay. *Sci. Rep.* **2019**, *9*, 5188. [\[CrossRef\]](#) [\[PubMed\]](#)
71. Nie, A.; Chao, Y.; Zhang, X.; Jia, W.; Zhou, Z.; Zhu, C. Phytochemistry and Pharmacological Activities of *Wolfiporia cocos* (F.A. Wolf) Ryvarden & Gilb. *Front. Pharmacol.* **2020**, *11*, 505249. [\[CrossRef\]](#)
72. Hu, G.S.; Huang, C.G.; Zhang, Y.; Xiao, W.; Jia, J.M. Accumulation of biomass and four triterpenoids in two-stage cultured *Poria cocos* mycelia and diuretic activity in rats. *Chin. J. Nat. Med.* **2017**, *15*, 265–270. [\[CrossRef\]](#)
73. Wu, Y. A Preliminary Study on Cultivation of *Poria cocos* on Nematode affected Pine Wood. *J. Fujian For. Sci. Technol.* **2013**, *40*, 51–55.
74. Raza, Q.; Rashid, M.A.R.; Waqas, M.; Ali, Z.; Rana, I.A.; Khan, S.H.; Khan, I.A.; Atif, R.M. Genomic diversity of aquaporins across genus *Oryza* provides a rich genetic resource for development of climate resilient rice cultivars. *BMC Plant Biol.* **2023**, *23*, 172. [\[CrossRef\]](#) [\[PubMed\]](#)
75. Adomako, M.; Ernst, D.; Simkovsky, R.; Chao, Y.Y.; Wang, J.; Fang, M.; Bouchier, C.; Lopez-Igual, R.; Mazel, D.; Gugger, M.; et al. Comparative Genomics of *Synechococcus elongatus* Explains the Phenotypic Diversity of the Strains. *mBio* **2022**, *13*, e0086222. [\[CrossRef\]](#) [\[PubMed\]](#)
76. Yang, L.; Tang, J.; Chen, J.J.; Peng, A.Y.; Wang, Q.M.; Rao, L.Q.; Yang, H.; Zhang, X.W.; Yang, H.Z.; Zhang, C.; et al. Transcriptome analysis of three cultivars of *Poria cocos* reveals genes related to the biosynthesis of polysaccharides. *J. Asian Nat. Prod. Res.* **2019**, *21*, 462–475. [\[CrossRef\]](#) [\[PubMed\]](#)
77. Elkady, W.M.; Gonaïd, M.H.; Yousif, M.F.; El-Sayed, M.; Omar, H.A.N. Impact of Altitudinal Variation on the Phytochemical Profile, Anthelmintic and Antimicrobial Activity of Two *Pinus* Species. *Molecules* **2021**, *26*, 3170. [\[CrossRef\]](#) [\[PubMed\]](#)
78. Gad, H.; Al-Sayed, E.; Ayoub, I. Phytochemical discrimination of *Pinus* species based on GC-MS and ATR-IR analyses and their impact on *Helicobacter pylori*. *Phytochem. Anal.* **2021**, *32*, 820–835. [\[CrossRef\]](#) [\[PubMed\]](#)

79. Gao, Y.; Xu, R.; Gu, S.; Chen, K.; Li, J.; He, X.; Shang, S.; Song, Z.; Song, J. Discovery of Natural Rosin Derivatives Containing Oxime Ester Moieties as Potential Antifungal Agents to Control Tomato Gray Mold Caused by *Botrytis cinerea*. *J. Agric. Food Chem.* **2022**, *70*, 5551–5560. [[CrossRef](#)] [[PubMed](#)]
80. Tao, P.; Wu, C.; Hao, J.; Gao, Y.; He, X.; Li, J.; Shang, S.; Song, Z.; Song, J. Antifungal Application of Rosin Derivatives from Renewable Pine Resin in Crop Protection. *J. Agric. Food Chem.* **2020**, *68*, 4144–4154. [[CrossRef](#)]
81. Kugler, S.; Ossowicz, P.; Malarczyk-Matusiak, K.; Wierzbicka, E. Advances in Rosin-Based Chemicals: The Latest Recipes, Applications and Future Trends. *Molecules* **2019**, *24*, 1651. [[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.