

# Genome Sequencing of *Inonotus Obliquus* Reveals Insights Into Candidate Genes Involved In Secondary Metabolite Biosynthesis

**Yingce Duan**

Northeast Forestry University

**Haiyan Han**

Northeast Forestry University

**Jianzhao Qi**

Northwest A&F University

**Jin-ming Gao**

Northwest A&F University

**Zhichao Xu**

Northeast Forestry University

**Pengchao Wang**

Northeast Forestry University

**Jie Zhang**

Northeast Forestry University

**Chengwei Liu** (✉ [liuchw@nefu.edu.cn](mailto:liuchw@nefu.edu.cn))

Northeast Forestry University

---

## Research Article

**Keywords:** *Inonotus obliquus*, Genome sequencing, CAZymes, P450, secondary metabolite

**Posted Date:** October 21st, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-963591/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at BMC Genomics on April 20th, 2022. See the published version at <https://doi.org/10.1186/s12864-022-08511-x>.

**Title: Genome sequencing of *Inonotus obliquus* reveals insights into candidate genes involved in secondary metabolite biosynthesis**

**Author names and affiliations.**

Yingce Duan<sup>1</sup>, Haiyan Han<sup>1</sup>, Jianzhao Qi<sup>2</sup>, Jin-min Gao<sup>2</sup>, Zhichao Xu<sup>1</sup>, Pengchao Wang<sup>1</sup>, Jie Zhang<sup>1</sup>, Chengwei Liu<sup>1\*</sup>

1. School of life science, Northeast Forestry University, Harbin 150040, Heilongjiang, China

2. Shaanxi Key Laboratory of Natural Products & Chemical Biology, College of Chemistry & Pharmacy, Northwest A&F University, Yangling 712100, Shaanxi,

\* Corresponding author: Chengwei Liu: No. 26 Hexing Road, Harbin, 150000, China,

E-mail: liuchw@nefu.edu.cn.

**Abstract**

**Background:** *Inonotus obliquus* is an important edible and medicinal mushroom that was shown to have many pharmacological activities in preclinical trials, including anti-inflammatory, antitumor, immunomodulatory, and antioxidant effects. However, the biosynthesis of these pharmacological components has rarely been reported. The reason for this is that there have been no relevant reports on its genome information.

**Results:** we report the genome of *I. obliquus* using a combined high-throughput Illumina NovaSeq with Oxford Nanopore PromethION sequencing platform. The de novo assembled 38.18 Mb *I. obliquus* genome was determined to harbor 12525 putative genes, with 81.83% of them having detectable sequence similarities to others available in public databases. Phylogenetic analysis revealed a close evolutionary relationship between *I. obliquus* and *Fomitiporia mediterranea* and *Sanghuangporus baumii* in the core Hymenochaetales clade. The *I. obliquus* genome was found to encode a repertoire of enzymes involved in carbohydrate metabolism, along with 135 cytochrome P450 proteins. Other annotated genes included those encoding key enzymes for secondary metabolite biosynthesis, such as those from polysaccharide, melanin, and triterpenoid pathways. Among them, the *I. obliquus* genome was particularly enriched in sesquiterpenoid biosynthesis genes and gene clusters.

**Conclusions:** This study presents the first genome analysis of an important medical mushroom, *I. obliquus*, which can provide insights into the usefulness of this organism and its secondary metabolites in medicine.

**Key words:** *Inonotus obliquus*, Genome sequencing, CAZymes, P450, secondary metabolite

**Background**

*Inonotus obliquus* (Ach. ex Pers.) Pilát belongs to Basidiomycota, Agaricomycetes, Hymenochaetales, Hymenochaetaceae, and *Inonotus* taxa, and are mainly distributed in northeastern Asia, including the Siberia region of Russia, northeastern China, and the Hokkaido region of Japan [1]. This medicinal fungus is enriched in many active chemical components and is widely used for disease treatment in Russia and northeastern China as a folk remedy. *I. obliquus* is a fungus that grows on birch trees, can be parasitic or saprophytic, and is a white rot fungus. Its growth temperature ranges from 25 to 30 °C, with pH 6 [2], but it always grows in high latitude areas where living conditions in the wild are harsh. Accordingly, the formation of the fruit body and sclerotium is slower and the entire growth cycle is longer, limiting its extensive development and utilization.

Many important secondary metabolites can be derived from the mycelia, fruit body, and sclerotium,

such as polysaccharides [3], melanin [4], phenols [5], and terpenoids [6]. These compounds present in *I. obliquus* have significant pharmaceutical value, such as anti-tumor [7], anti-inflammatory [8], anti-microbial [9], anti-oxidant [10], and anti-neuroinflammatory [11]. Research has shown that 150 µg/ml polysaccharide from *I. obliquus*, which can restrain the growth of a hepatoma cell line, exhibits an inhibitory rate similar to that of mitomycin at a dose of 5 µg/ml. Mycelial polysaccharides also show stronger inhibitory activity against lipid peroxidation, at 300 µg/ml, reaching a peak inhibitory rate of 69.74% [3]. Melanin of *I. obliquus* facilitates an increase in the growth of *Bifidobacterium bifidum* 1 by 1.4-fold in comparison to that with ascorbic acid, as a control in the trials, after 24 h of cultivation [12]. The contents and species of triterpenoids from *I. obliquus* are abundant and complex and include trametenolic acid, inotodiol, and betulinic acid. These compounds reduce the viability of human cancer cell lines (IC<sub>50</sub> value < 5 µM) and have anti-proliferative properties [13].

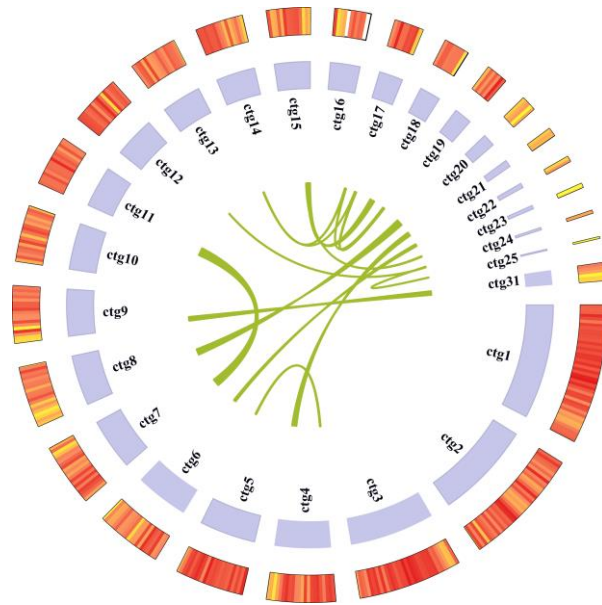
In recent years, rapid advancements in technology have gradually led to the analysis of genomes of many macro medicinal mushrooms, like *Ganoderma lucidum* [14], *Antrodia cinnamomea* [15], *Hericium erinaceus* [16], *Sanghuangporus baumii* [17], and *Wolfiporia cocos* [18]. Despite increasing interest in the active components of *I. obliquus*, very little is known about the molecular and genetic basis of the biosynthetic pathways yielding these components. Here, we report the genome sequence of *I. obliquus* based on single-molecule real-time reads from the Nanopore platform and the combination of an Illumina sequencing strategy. First, characterization analysis of the *I. obliquus* genome included gene content and genome structure. Second, we identified functional genes and gene clusters involved in secondary metabolism and biosynthesis, such as that related to polysaccharides, melanin, and terpenoids. Third, we performed classification analysis of the P450 gene family involved in secondary metabolism and biosynthesis.

## Results

### Genome sequence assembly and annotation

The genome sequence of *I. obliquus* was analyzed using the Oxford Nanopore PromethION sequencing platform, utilizing NECAT software to perform genome error correction and splicing, and finally obtaining initial joint results. Racon (version:1.4.11) software was used twice based on joint results for error correction. Finally, Pilon (version:1.23) software was used twice for error correction after purging the haplotigs to obtain the final assembled results. In total, 37,218,262 clean reads were generated, and the total number of bases was 558,2739,300 (Table S1-2).

The genome size was 38.18 Mb. This consisted of 31 contigs with an N50 of 1.88 Mb and 47.56% GC content (Fig 1). Our results indicated that genome assembly was of high quality (Table S3-4). We predicted 12,525 protein-coding genes, and 80% of the genes were annotated. Average CDS sequence length was 1296, and the longest contig length was 4.18 Mb (Table 1). The majority of the repeats were copia/gypsy (9.57% of the genome). For non-coding RNA, 88 tRNAs, 78 rRNAs, 14 snRNAs, and one sRNA were predicted. The number of long terminal repeats was 1286, occupying 3.9% of the whole genome, the number of DNA transposons was 732, occupying 0.9% of the whole genome, the number of simple repetitions was 5401, occupying 0.59% of the whole genome, and the number of satellite repetitions was 11 (Table S5-7).



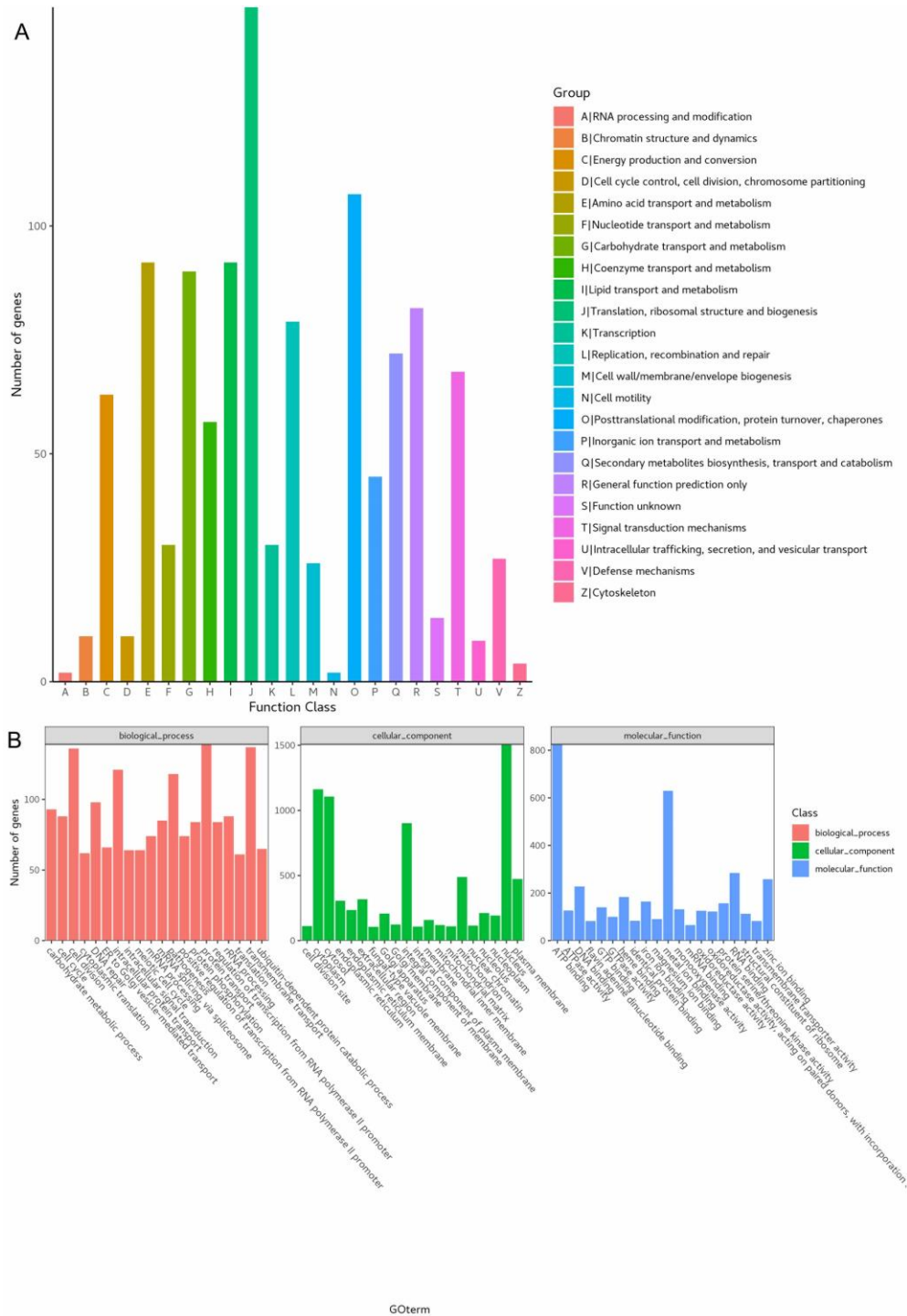
**Fig. 1** Genome diagram of *Inonotus obliquus* genome. The outermost circle is gene density, and the innermost circle is collinearity analysis. The middle circle is the size of each contig. Remaining contigs are not displayed because there were no gene annotations and their gene sizes were small.

**Table1** *De novo* genomic assembly and functional annotation of *Inonotus obliquus* genome

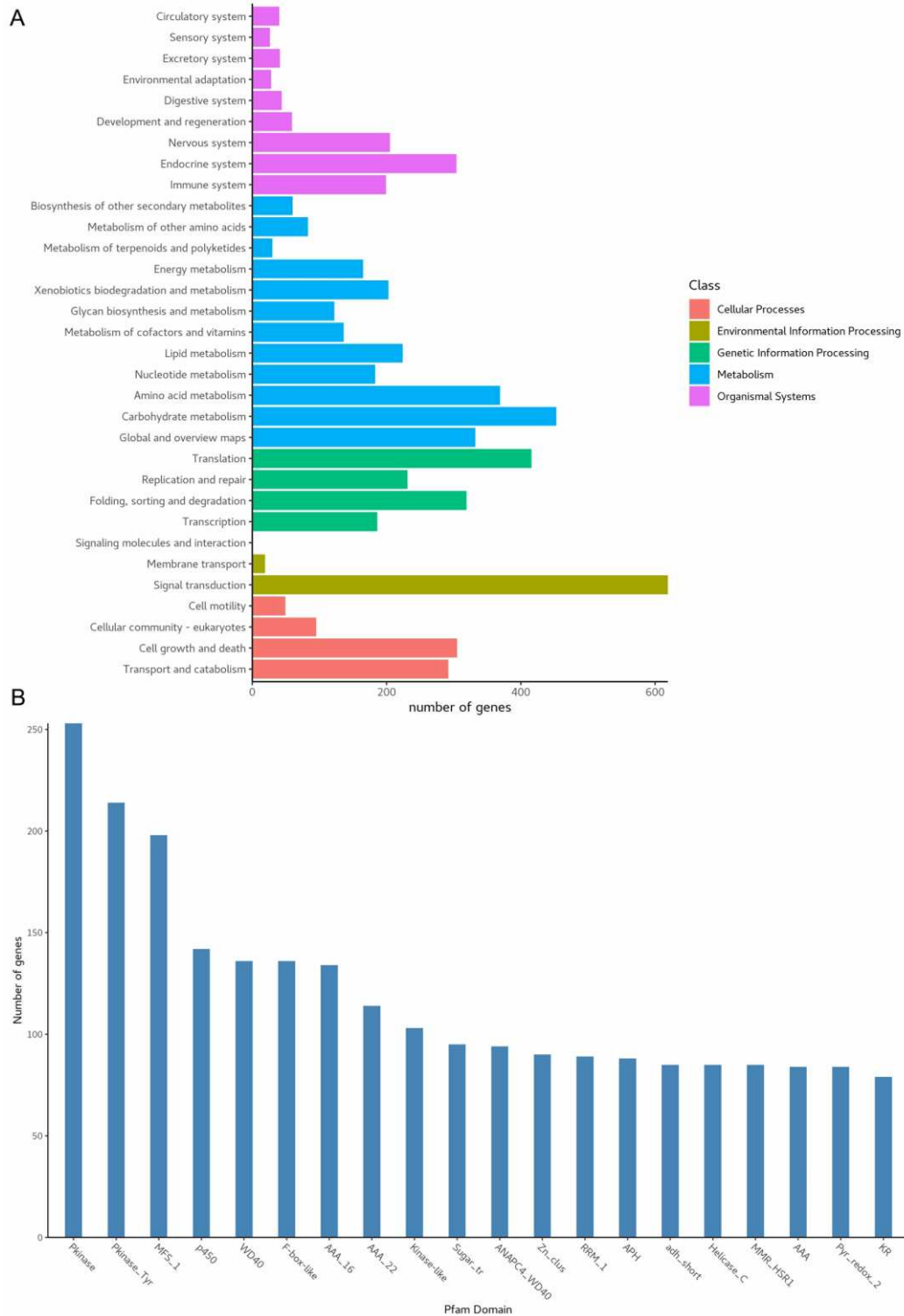
Item	Value	Item	Count	Percentage
Total_length(bp)	38181337	All	12,525	100%
Contigs	31	Nr	10,249	81.83%
GC_content(%)	47.56	Pfam	7,956	63.52%
N50(bp)	1971511	Interproscan	7,924	63.27%
N90(bp)	915121	Uniprot	5,648	45.09%
Average(bp)	1193166.78	GO	5,602	44.73%
Median(bp)	1027062.5	KEGG	4,121	32.90%
Min(bp)	2774	Refseq	3,945	31.50%
Max(bp)	4380421	Pathway	2,509	20.03%
the total number of gene	12525	COG	1,112	8.88%
Sequencing method	Nanopore; Illumina			

To predict protein sequences, 12,525 non-redundant genes were subjected to similarity analysis based on nine public databases. Most genes were mapped using the Nr database, specifically 10,249 genes/81.83%, followed by Pfam (7956 genes/63.52%), Interproscan (7924 genes/63.27%), Uniprot (5468 genes/45.09%), Gene ontology (GO; 5602 genes/44.73%), Kyoto Encyclopedia of Genes and Genomes (KEGG; 4,121 genes/32.90%), Refseq (3945 genes/31.50%), Pathway (2509 genes/20.03%), and Clusters of Orthologous Groups (COG; 1112 genes/8.88%) (Table 1). According to the COG database, “translation, ribosomal structure and biogenesis” was associated with the most genes (148). This was followed by “posttranslational modification, protein turnover, chaperones”, “amino acid transport and metabolism”, and “lipid transport and metabolism” as the most gene-rich classes in the COG groupings (Fig 2A). These findings suggest the presence of an enriched and varied array of protein and lipid metabolism functions that enable higher energy conversion efficiency. GO annotation resulted in the nucleus (1507), cytoplasm (1167), and cytosol (1105) from the cellular component category, protein transport (139) from biological processes, and ATP binding (825) and metal ion binding (630) from molecular functions (Fig 2B). These results show that the most abundant genes in the genome are

the metabolism of genetic material and energy. The KEGG functional classification showed signal transduction (619), carbohydrate metabolism (453), and translation (416) (Fig 3A). *I. obliquus* is a wild strain, in which many metabolic genes are involved in signal transduction, indicating a high degree of adaptability to the environment. We identified 253 Pkinase, 198 MFS, and 136 WD40 genes in the pfam domain of the *I. obliquus* genome (Fig 3B).



**Fig. 2** Genomic functional annotation of *Inonotus obliquus*. **a** Cluster of Orthologous Groups of proteins (COG) analysis of *Inonotus obliquus*. **b** Gene Ontology (GO) analysis of *Inonotus obliquus*.

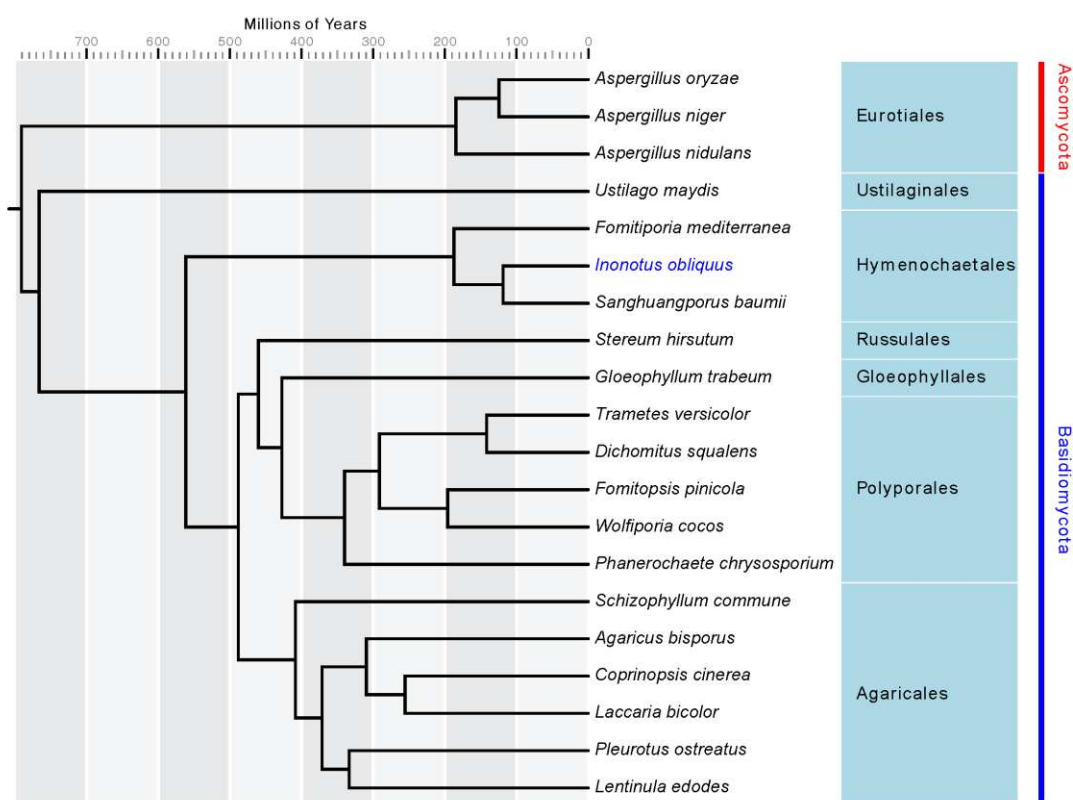


**Fig. 3** Genomic functional annotation of *Inonotus obliquus*. **a** Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of *Inonotus obliquus*. **b** Pfam domain analysis of *Inonotus obliquus*.

### Phylogenetic analysis of other fungal genomes

To investigate the evolutionary history and classification status of the *I. obliquus* genome, we identified a total of 18571 homologous gene families, among which *I. obliquus* had 7154 families. The *I. obliquus* genome was found to contain 407 specific families. In total, 699 single-copy orthologous

genes were used for phylogenetic tree construction. We found that these 20 fungal species were distributed on two branches, Basidiomycetes and Ascomycetes ( Table S8 ). The branch of Basidiomycetes was further divided into six subgroups, which corresponded to six orders, Agaricales, Polyporales, Gloeophyllales, Russulales, Hymenochaetales, and Ustilaginales. The phylogenetic tree showed that the estimated divergence time between the *I. obliquus* lineage and *F. mediterranea* lineage was approximately 195 million years ago (Mya) and that from the *S. baumii* lineage was approximately 100 Mya. These fungi belong to Hymenochaetales. The relationship was distant between *I. obliquus* and *W. cocos*, *D. squalens*, and *T. versicolor*, species from the Polyporales order. *I. obliquus*, *S. hirsutum*, and *S. commune* were determined to have the same ancestor, and the estimated divergence time was approximately 570 Mya (Fig 4).



**Fig. 4** Phylogenetic analysis of 20 fungal species based on 699 single-copy orthologous genes.

### Carbohydrate genes

This analysis revealed 380 genes encoding carbohydrate-active enzymes (CAZymes) that were distributed. These included 192 glycoside hydrolases (GHs), 75 auxiliary activities (AAs), 66 glycosyltransferases (GTs), 24 carbohydrate esterases (CEs), 11 carbohydrate-binding modules (CBMs), and 13 polysaccharide lyases (PLs) (Table 2). As white rot fungi, because of strong lignocellulose-degradation activity, carbohydrate genes of *I. obliquus* exceeded those of *G. trabeum*, *F. pinicola*, and *W. cocos*, brown rot fungi, as well as the symbiotic fungus *L. bicolor*. Straw rot *A. bisporus* and *C. cinereus* were determined to have similar or even higher quantities of the *I. obliquus* carbohydrate gene. Compared with *I. obliquus*, the white rot fungi *S. hirsutum*, *T. versicolor*, *P. ostreatus*, and *L. edodes* had more carbohydrate genes. PLs were mainly distributed in five families, including PL1, PL8, PL14, PL35, and PL38. AAs mainly included AA1-3, AA5-AA9, AA14 nine families. GTs contained 29 families, including eight chitin synthetases belonging to the GT2 family. CEs were classified as CE1, CE4, CE8-

9, CE12, CE15-17 eight families. GHs were distributed across 46 families. Cellulose and hemicellulose-degrading enzymes mostly belonged to GH1, GH3, and GH6 families; lignin-degrading enzymes mainly belonged to the AA1 and AA2 families in the genome.

**Table 2** Gene distribution of different fungi based on the six major modules of CAZymes

CAZy	Total	GH	AA	GT	CBM	CE	PL
<i>Inonotus obliquus</i>	380	199	76	66	11	24	13
<i>Sanghuangporus baumii</i>	329	175	63	61	5	18	9
<i>Fomitiporia mediterranea</i>	387	197	91	66	5	22	8
<i>Coprinopsis cinerea</i>	469	188	129	72	21	47	18
<i>Schizophyllum commune</i>	469	247	85	71	11	37	19
<i>Gloeophyllum trabeum</i>	350	199	57	62	5	19	11
<i>Dichomitus squalens</i>	433	218	102	65	9	28	15
<i>Phanerochaete chrysosporium</i>	383	183	102	66	10	21	8
<i>Laccaria bicolor</i>	311	151	53	73	9	9	7
<i>Agaricus bisporus</i>	370	172	94	52	11	33	12
<i>Pleurotus ostreatus</i>	505	226	139	65	29	27	26
<i>Stereum hirsutum</i>	518	274	127	68	7	29	19
<i>Trametes versicolor</i>	441	223	106	78	6	19	13
<i>Wolfiporia cocos</i>	269	146	42	63	2	13	4
<i>Lentinula edodes</i>	449	241	86	71	14	32	11
<i>Fomitopsis pinicola</i>	350	201	56	65	5	20	5
<i>Ustilago maydis</i>	215	106	26	63	1	17	2
<i>Aspergillus nidulans</i>	486	259	89	81	20	32	23
<i>Aspergillus niger</i>	503	253	105	102	17	30	10
<i>Aspergillus oryzae</i>	538	289	96	93	17	31	26

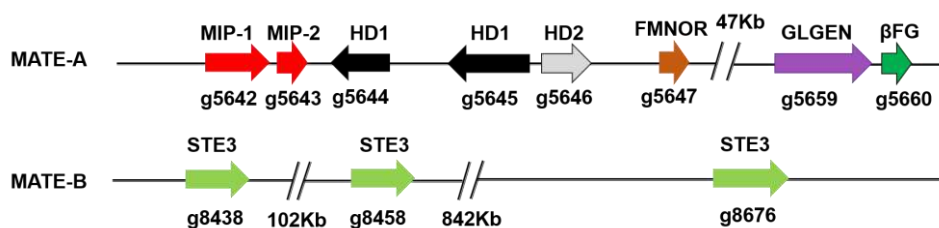
## Mating genes

Mating type recognition plays a role in the genetics and breeding of edible and medicinal mushrooms, determining the propagating system, fruiting body, and gamete quality. Homothallism is selfing fertility and heterothallism is a hybrid conception. Homothallism can be divided into biopolar and tetrapolar. *L. edodes* and *G. lucidum* are tetrapolar, whereas *Cordyceps militaris* and *W. cocos* are biopolar [19]. The biopolar homothallism sex system comprises a single factor; every monokaryon only has one mating type, and only different types of monokaryons can mate with a heterokaryon at the time of sexual reproduction. In a tetrapolar mating system, genes in mate A encode homeodomain (HD) transcription factors, and those of mate B encode the pheromone receptor and pheromone precursor genes. Mate A controls hook cell formation, nucleus pairing, hyphal cell fusion, and lock-cell formation. Mate B controls septal dissolution, nuclear migration, and lock-cell fusion with the subapical cells. Only when A and B mating types are different between the two monokaryons can mating be successful.

The mating type of *I. obliquus* has not been reported to date. In the sequencing analysis, we found HD-encoding protein type-related genes in the genome; HD1 was g5645, HD2 was g5644, and MIP genes were g5642 and g5643. The positional continuity of these four genes was located in contig 1. Pheromone receptors included g8676, g8458, and g8438, located in contig 15 (Fig 5 and Table S9). The B mating type locus was not found to be linked with homeodomain transcription factors (A mating type locus) in the tetrapolar region. According to the genome information, two mating types were not in the same contig, and we predicted that the mate type of *I. obliquus* is tetrapolar, but determination of the real



mating type requires further testing and verification.



**Fig. 5** Distribution of mating type genes in *Inonotus obliquus*. MIP, mitochondrial intermediate peptidase; HD1, homeodomain protein 1; HD2, homeodomain protein 2; FMNOR, FMN-linked oxidoreductase; GLGEN, glycosyltransferase family 8 protein; betaFG, beta-flanking gene; STE, STE3-like pheromone receptor.

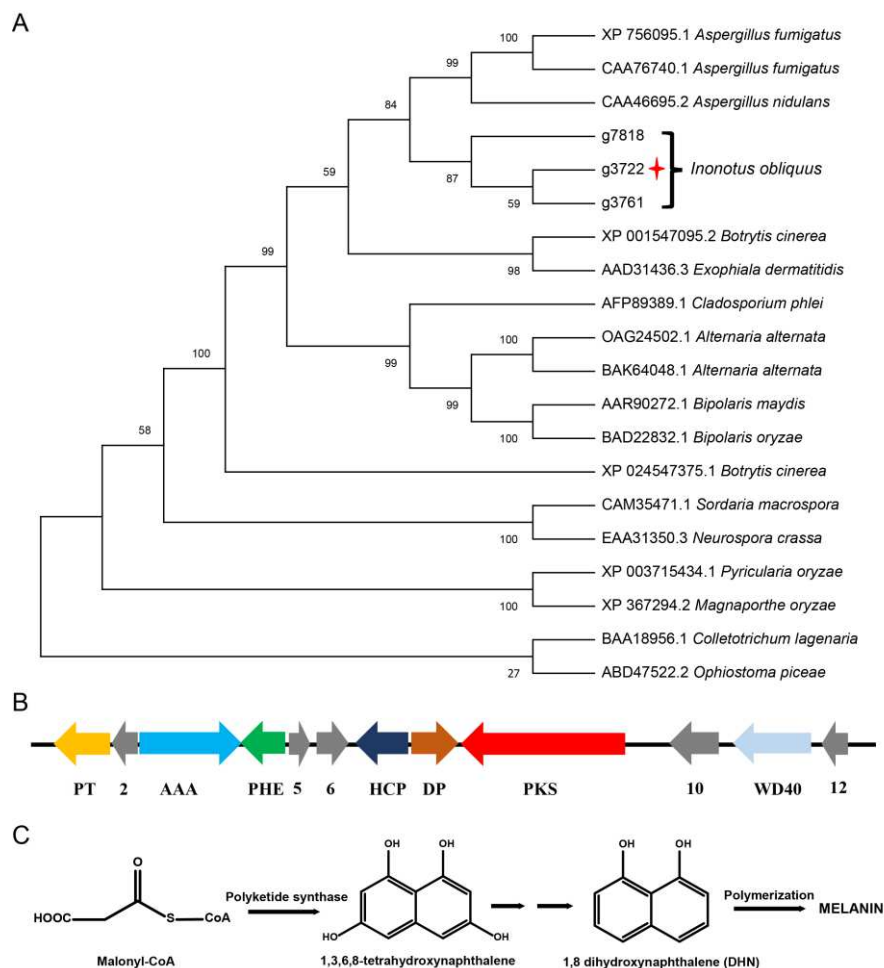
### Polysaccharide biosynthesis

*I. obliquus* polysaccharides have anti-oxidant, anti-tumor, anti-inflammatory, and anti-diabetic biological activities, with monose components mainly comprising mannose, glucose, galactose, xylose, rhamnose, and arabinose [20]. As *G. lucidum* polysaccharides involved in biosynthetic pathway correlational studies, we found 18 genes involved in polysaccharide biosynthesis, including one phosphoglucomutase, one UDP-glucose-6-dehydrogenase, and two glucan synthases (Table S10). The number of genes related to polysaccharide synthesis in *I. obliquus* was similar to that in *S. baumii*. Overexpressing phosphoglucomutase can increase the polysaccharide content in *G. lucidum* [21]. In *C. militaris*, co-expressed phosphoglucomutase and UDP-glucose 6-dehydrogenase can improve the whole content of intracellular and extracellular polysaccharides, increasing polysaccharide content by 78.13% compared with that of the wild-type strain [22]. Exogenous siRNAs were also previously applied to target beta-1,3-glucan synthase, negatively affecting the growth of the fungus *Macrophomina phaseolina*. Fungal cell walls are composed of chitin and glucan; therefore, polysaccharide synthesis is strongly correlated with regular hyphal growth [23]. The metabolism of polysaccharides found in *I. obliquus* is similar to that of other medicinal fungi.

### Melanin biosynthesis involved in polyketide synthase

Melanin is a secondary metabolite that is composed of complex heterogeneous polymers. The complete absence of nitrogen in *I. obliquus* melanin samples suggests that *I. obliquus* contains 1,8-dihydroxynaphthalene (DHN)-type melanin [13, 24]. In combined pathogenic fungus DHN-type melanin biosynthesis, the formation of melanin requires polyketide synthase (PKS), reductase, and hydrolase. There was no hydrolase found in the *I. obliquus* genome, and we only found polyketide synthase and reductase. 1,8-Dihydroxynaphthalene polymerizes melanin, which requires laccase participation, and we used sequence alignment to find 17 presumptive genes involved in melanin biosynthesis in *I. obliquus*. Phylogenetic tree analysis of three PKSs and 17 fungal melanin genes involved in PKS relationships indicated that g3722 was the closest other fungal PKS. This gene was located in contig 4 with nearby genes to form a cluster called water-soluble melanin, g3714-g3725, and all 12 genes were identified through system analysis, with a total length of 34.168 Kb. The specific functions of these genes are as follows: g3722 is a core enzyme PKS, and phylogenetic tree analyses indicated that it belongs to the type I PKS class [25, 26]. g3714 is a tRNA isopentenyl transferase, g3716 is an AAA, g3717 is a phosphodiesterase, g3720 is a HCP, g3721 is a dipeptidase, g3724 is a WD, and other genes encoded hypothetical proteins (Fig 6). Two pathways of melanin synthesis were found in the fungi. Many pathogenic fungi synthesize melanin via the DHN pathway. The second melanin pathway is named the L-DOPA pathway after L-3,4-dihydroxyphenylalanine [27]. There is more DOPA-type melanin in edible

mushrooms, such as *G. lucidum* [28], *Termitomyces albuminosus* [29], *A. bisporus* [30], and *Auricularia auricula* [31]. We found only one polyphenol oxidase gene in *I. obliquus*, but *A. bisporus* had four genes (Table S11). There might be different types of melanin in *I. obliquus*, but the specific composition is not clear. The formation of melanin mediated by the polyphenol oxidase gene in *I. obliquus* remains to be explored.



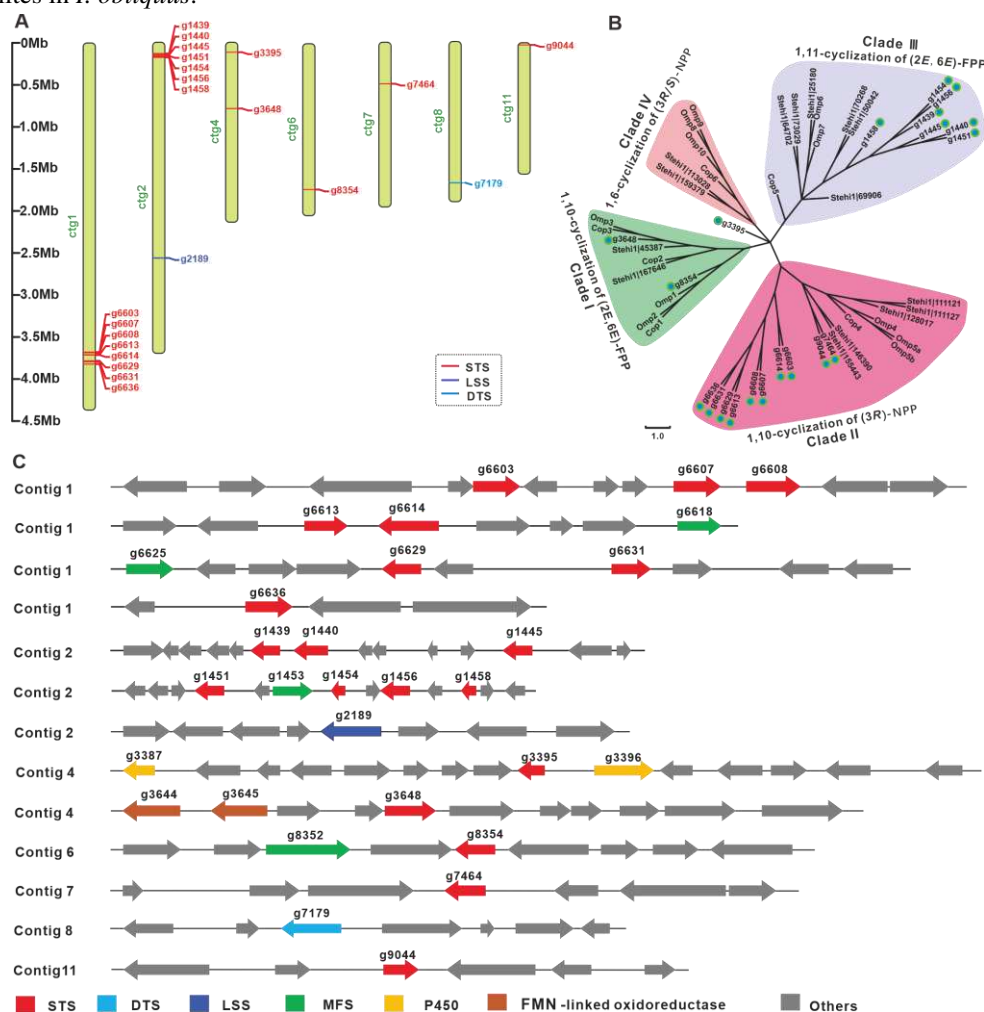
**Fig. 6** Prediction of the gene cluster structure of water-soluble melanin in *Inonotus obliquus*. **a** is the PKS phylogenetic tree; **b** is melanin gene clusters; PT, tRNA isopentenyl transferase; AAA, ATPase family associated with various cellular activities; PHEs, phosphodiesterases; HCP, hemolysin co-regulated protein; DP, dipeptidase; PKS, putative water-soluble melanin synthase; WD40, WD40 repeat protein; gray segments 2, 5, 6, 10, and 12 represent hypothetical proteins; **c** is hypothesized biosynthetic pathways of water-soluble melanin.

### Terpenoid biosynthesis

Terpenoids are one of the main secondary metabolites in *I. obliquus*, and 86 bioactive metabolites have been reported to date [32]. Most of these major terpenoid synthesis sources are mainly derived from the mevalonate pathway, which consists of 11 enzymes and 16 genes, including acetyl-CoA acyltransferase, based on four coding genes, mevalonate kinase and farnesyl diphosphate (FPP), which are encoded by two genes, and other enzymes encoded by a single-copy gene (Table S12). We found altogether 19 gene clusters of secondary metabolites by AntiSMASH fungal 6.0.0, distributed different contigs and 13 terpenoids synthesis-related gene clusters. There were 22 genes related to terpenoid

synthase in the *I. obliquus* genome, including 20 sesquiterpene synthases (STSs), one lansterol synthase, and one geranylgeranyl diphosphate synthase.

According to the same conserved domain, 20 genes were determined to be probably involved in sesquiterpene synthase. We used other known sesquiterpene synthases as identification criteria to identify the composition of sesquiterpene synthases in *I. obliquus*, such as those of *Omphalotus olearius* [33], *S. hirsutum* [34], and *C. cinereus* [35]. The 20 sesquiterpene synthases were divided into three clades. There were 10 sesquiterpene synthases belonging to Clade II, seven sesquiterpene synthases belonging to Clade III, two sesquiterpene synthases belonging to Clade I, and one was not assigned (Fig 7B). Clade II consisted of enzymes that shared a 1,10-cyclization of (3R)-nerolidyl diphosphate mechanism, producing sesquiterpenes derived from a *Z*, *E*-germacradienyl cation. Clade III consists of enzymes believed to share a common 1,11-cyclization of the (2E,6E)-FPP mechanism, producing the trans-humulyl cation. Clade I consisted of enzymes that utilize a 1,10-cyclization of (2E,6E)-FPP to produce sesquiterpenes derived from a *E,E*-germacradienyl cation [35]. It can be seen from the position of the contigs in which the genes were located that sesquiterpene synthases were almost all concentrated at the two ends of the contig (Fig 7A). This distribution phenomenon was consistent with the distribution of genes related to secondary metabolism. In the genes surrounding sesquiterpene synthase, we found that membrane transporters and cytochrome P450 were related to sesquiterpene synthesis (Fig 7C). According to relevant literature, eight sesquiterpenoids have been found in *I. obliquus* [36, 37]. The discovery of these genes will help to study the biosynthetic pathways of sesquiterpenoid secondary metabolites in *I. obliquus*.



**Fig. 7** Terpene synthases in *Inonotus obliquus*. **a** Distribution of terpene synthases on the different contigs; **b** phylogenetic analysis of sesquiterpene synthase (STS) homologues; **c** gene clusters in *I. obliquus*. DTS is diterpene synthase, and LSS is lanosterol synthase.

### Cytochrome P450 monooxygenase (CYP) family analysis

P450 plays an important role in the biosynthesis of secondary metabolites in mushrooms and participates in lignocellulose degradation, cell wall and membrane biosynthesis, glycogen synthesis, detoxification, and hazardous substance degradation. P450 has been reported in filamentous fungi, such as *A. oryzae* (152 P450s) [38], *A. nidulans* (119 P450s which contained eight pseudogenes and divided into 89 CYP families) [39]. Regarding P450, in the medicinal and edible mushroom *G. lucidum*, 219 CYP genes (197 functional genes and 22 pseudogenes) were found, divided into 42 families [14]. *A. cinnamomea* harbors 119 CYP genes [15], *Hypsizygus marmoreus* has 132 CYP genes [40], and *H. erinaceus* contains 137 CYP genes [41]. According to domain and pfam prediction, 135 P450 genes were screened in the *I. obliquus* genome. Based on family cluster analysis of 135 CYPs, it was found that 106 genes could be clustered and divided into 19 families, with the largest quantity of CYP620, CYP512, CYP5150, CYP5141, CYP5144, CYP5037, and CYP5035 genes. In total, 87 genes were identified (Fig 8). In *G. lucidum*, CYP512 family proteins might be involved in triterpenoid biosynthesis [14]. CYP5150A2 from the white-rot basidiomycete *P. chrysosporium* is capable of hydroxylating 4-propylbenzoic acid with NADPH-dependent cytochrome P450 oxidoreductase as a single redox partner [42]. Functional screening showed that CYP5035 assists in the fungal detoxification mechanism in Polyporales [43]. These cytochrome P450 subfamilies could be closely related to the formation of secondary metabolites in *I. obliquus*.

There are major bioactive compounds in *I. obliquus*, including inotodiol, botulin, and betulinic acid [2, 44]. These compounds comprise two different types of triterpenes. The synthesis of lanosterol and lupeol is catalyzed by lanosterol synthase and lupeol synthase enzymes, with 2,3-oxidosqualene as a precursor. Lanosterol produces inotodiol via the action of cytochrome P450 hydroxylation, and lupeol produces betulin and betulinic acid through the combined action of cytochrome P450 oxidase and reductase; however, cytochrome P450 and lupeol synthase have not been reported in *I. obliquus* and other fungi. We only chose the sequences in plants based on the synthesis of the same or similar substances according to the reference. Specifically, Yang et al. reported that CYP89S1, CYP97B62, and CYP86A182 have C-28 oxidation functions and catalyze the conversion of lupeol to betulinic acid in birch [45]; further, CYP90B and CYP724B have C-22 hydroxylation functions and catalyze the formation of steroids in plants such as *Arabidopsis thaliana* and *Solanum tuberosum* [46-49]. Owing to the structural similarity of triterpenes and steroids, inotodiol synthesis includes lanosterol C22 hydroxylation. Therefore, according to BLASTP screening, similar related P450 sequences in the genome were hypothesized (Table S13), comparing sequences of all P450s, which provides a foundation for further experimental verification in a later stage. Based on different gene families from phylogenetic tree analysis, fungal CYPs showed highly conserved characteristic motifs but very low overall sequence similarities [50]. Betulinic acid biosynthesis related to P450, g5553 and g3231 were determined to belong to the same family, CYP63, and g7106, g6587, g8846 were respectively CYP5032, CYP5148, CYP5037 families. Inotodiol biosynthesis is related to P450, mainly in the CYP51 and CYP512 subfamilies (Table S13). Zhang et al. reported that CYP51 belongs to the CYP superfamily and is a crucial step in the synthesis of ergosterol, which is a fungal-specific sterol. CYP51 has strong specificity and only catalyzes the demethylation of a very narrow range of substrates, including lanosterol [51].

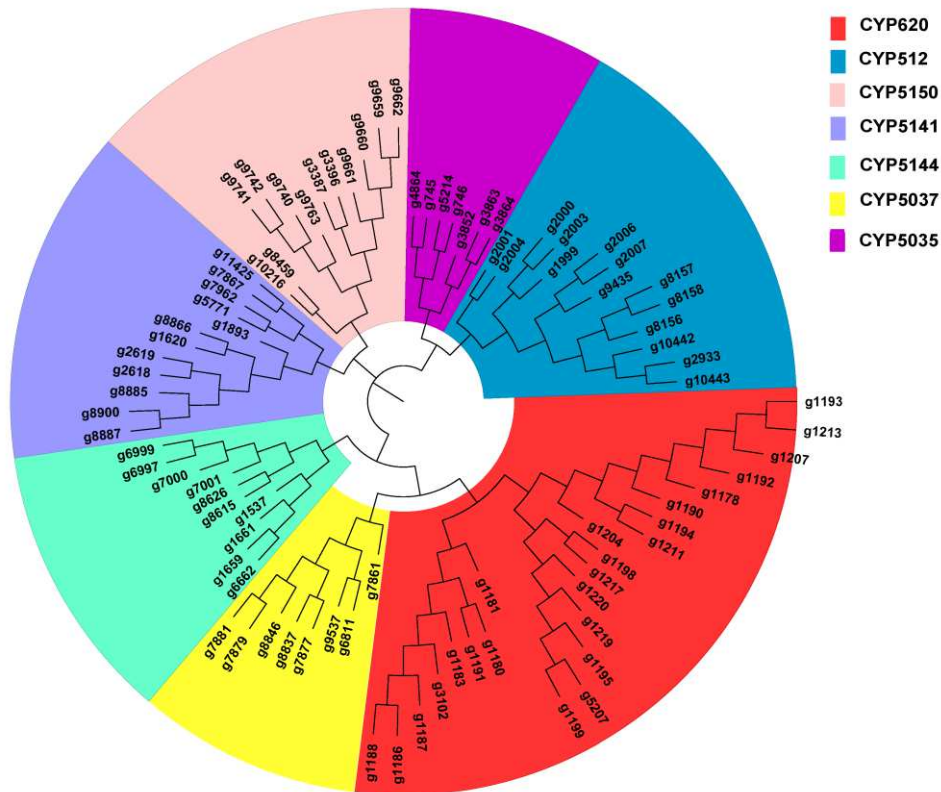


Fig. 8 Phylogenetic tree composed of the main P450 families of *Inonotus obliquus*.

## Discussion

*I. obliquus* is a typical wood-rot fungus that produces abundant enzymes related to lignocellulose degradation. We found 380 carbohydrate-related genes in the genome of this fungus. The different polarity is an important mode of reproduction that affects the transmission of genetic material. Although the polarity of many species from Agaricales and Polyporus has been analyzed [19], sufficient reports on polarity of Hymenochaetales are currently lacking. It is possible that their species are relatively evolutionarily close, as compared to those of *Sanghuangporus*, including *S. baumii*. *I. obliquus* is quite different from *S. baumii* in morphology and molecular identification. It has been reported that *F. mediterranea* is biopolar [52]. The polarity of *I. obliquus* is based on the distribution of its reproduction-related genes on the contigs, and we infer that it may be tetrapolar, but the specific polarity needs to be further tested and verified.

*I. obliquus*, the earliest folk herbal medicine widely used in Russia, has garnered attention due to its significant anti-cancer activity. In this study, we excavated and annotated important genes related to its secondary metabolism. We predicted a PKS gene cluster for melanin synthesis and carried out its annotation. The morphology also shows that *I. obliquus* is rich in melanin. The terpenoids inotodiol and betulinic acid were found the more typical ones. Betulinic acid is mostly found in plants, such as birch [45] and mulberry [53]. However, it is rarely reported in fungi, except for *S. baumii* [54], *Trametes versicolor* [55], and *I. obliquus*. We need to identify genes related to its biosynthesis to establish a consensus. In this article, a total of 20 genes related to sesquiterpene synthase have been discovered while

the surrounding genes have been annotated. However, the number of sesquiterpenes reported so far is much smaller than the original number of the encoding genes. We, therefore, speculate that many genes may be in the silent stage. Although there are probably 400 biologically active constituents in *G. lucidum* [56] and more than 300 compounds in *Phellinus* spp. [54], the number of compounds presently known in *I. obliquus* is less. P450 plays roles in hydroxylation, carbonylation, carboxylation, and ketonation, and is involved in triterpene synthesis. The analysis of P450 will be useful for finding pharmaceutical targets. We found a total 135 P450 genes in *I. obliquus* (Table S14) and analyzed candidate P450 proteins related to betulinic acid and inotodiol synthesis. Synthesis-related genes need transcriptome analysis and other follow-up experiments for further function genes verified.

## Conclusion

In this study, we presented the first genome analysis of an important medical mushroom, *I. obliquus*. The de novo sequenced and annotated genome, assembled using Oxford Nanopore PromethION sequencing platform, detailed functional annotations were made regarding the genome of *I. obliquus* using major databases. The information on the *I. obliquus* genome could provide a clear genetic background for the study of secondary metabolism and its medicinal applications. We analyzed the secondary metabolite biosynthesis genes in the *I. obliquus* genome, such as key genes related to polysaccharides, melanin and terpenoid. Additionally, we identified some candidate P450 proteins related to betulinic acid and inotodiol biosynthesis.

## Methods

### Collection of strains and culture conditions

The *I. obliquus* strain was obtained from the Microbiology Laboratory, College of Life Science, Northeast Forestry University. The fruit body was collected from the Greater Khingan Mountains area and named CT5, which was identified based on internal transcribed spacer sequences after tissue separation. The strain was grown at 30 °C for 5 days on potato dextrose medium. The *I. obliquus* genomic DNA was extracted from mycelia using the Tiangen plant DNA kit DP350, according to the manufacturer's instructions.

### Genome sequencing and assembly

After the library was built, an effective concentration and volume of the DNA library was added to the flow cell, and the flow cell was transferred to the Oxford Nanopore PromethION sequencer with Illumina NovaSeq [57] for real-time single-molecule sequencing (NCBI SRA database accession number SRR15674625). The genome size of *I. obliquus* was estimated by the k-mer method using sequencing data from the DNA library. NECAT software was used to perform genome error correction and splicing was performed to obtain the initial splicing result; then, Racon (version: 1.4.11) software [58] was used to perform two rounds of error correction on the splicing result based on the third-generation sequencing data, and finally, two rounds of Pilon were performed (version: 1.23). Error correction was performed [59], and after removing heterozygosity, the final assembly result was obtained. BUSCO software (version: 4.1.4) was used to evaluate the integrity of the predicted genes based on the fungal database (fungi\_odb10) [60].

### Gene prediction and annotation

Gene prediction was performed mainly using BRAKER software (version: 2.1.4); first, GeneMark-EX was used to train the model, and then, AUGUSTUS was called for prediction [61]. Gene function annotation referred to the annotation of gene functions and metabolic pathways based on existing

databases, including predictions of information such as motifs, structural domains, protein functions, and metabolic pathways. To obtain comprehensive gene function information, gene function annotations of nine major databases were carried out, including Nr, Pfam [62], COG [63], Uniprot [64], KEGG [65], GO [66], Pathway, Refseq [67], and Interproscan [68].

### Phylogenetic location

Together with *I. obliquus*, and other 19 species (*Gloeophyllum trabeum*, *Fomitopsis pinicola*, *Lentinula edodes*, *Pleurotus ostreatus*, *S. baumii*, *Fomitiporia mediterranea*, *W. cocos*, *Dichomitus squalens*, *Coprinopsis cinerea*, *Schizophyllum commune*, *Phanerochaete chrysosporium*, *Agaricus bisporus*, *Ustilago maydis*, *Stereum hirsutum*, *Trametes versicolor*, *Laccaria bicolor*, *Aspergillus oryzae*, *A. nidulans*, and *A. niger*), homologous gene identification and phylogenetic analysis were performed. Single-copy homologous genes were identified using OrthoFinder [69]. STAG was used to build a phylogenetic tree, and then, MCMCtree software was utilized to predict divergence time. Two groups of recent ancestor divergence times were queried as calibrated points in timetree.org (*A. niger* vs. *A. bisporus* 626-806 MYA and *A. bisporus* vs. *U. maydis* 415-482 MYA).

### Identification of *matA* and *matB* genes

Using tetrapolarity *S. commune* *MAT-A* genes as a reference with pfam domain to predict and identify conserved domains [70], we identified *MAT-A* genes in the genome. The mitochondrial intermediate peptidase gene (*mip*) was identified in the same manner. *MAT-B* genes include pheromone and pheromone precursors. The sequence length of the pheromone precursor was too short to align it for prediction. We used an annotation file to find the *MAT-A*- and *MAT-B*-specific locations.

### CAZy and CYP family in *I. obliquus*

Carbohydrates play an important role in many biological processes. A large amount of meaningful biological information can be obtained by studying carbohydrate-related enzymes. CAZy data focus on analyzing the genome, structure, and biochemical information of carbohydrate enzymes. HMMER (version: 3.2.1, filter parameter E-value <  $1e^{-18}$ ; coverage > 0.35) [71] was used to annotate protein sequences based on the CAZy database (<http://bcb.unl.edu/dbCAN2/>) [72].

Cytochrome P450 is a large family of proteins, with heme as a prosthetic group. They can catalyze the oxidation reactions of many types of substrates, and they participate in the metabolism of endogenous and exogenous substances, including drugs and environmental compounds. Diamond blastp (version > 2.9.0; parameter: -evalue  $1e^{-5}$ ) was used to annotate the target protein sequence based on the Fungal cytochrome P450 database. The reference CYP sequences were downloaded from the web (<http://p450.riceblast.snu.ac.kr/index.php?a=view>) [73].

### Secondary metabolite gene clusters in *I. obliquus*

Secondary metabolite gene clusters were determined using AntiSMASH fungal 6.0.0 [74], a web-based analysis platform. The default parameter settings were used. To verify the predicted results, the obtained gene clusters were manually checked. Blastp analysis and gene annotation were performed using the NCBI genome portal software platform. We searched all hypothetical gene models in the database using blastp and tblastn algorithms.

### Abbreviations

**AAs**, auxiliary activities; **CAZymes**, carbohydrate-active enzymes; **CE**, carbohydrate esterase; **COG**, Clusters of Orthologous Groups; **CYP**, cytochrome P450 monooxygenase; **DHN**, 1,8-dihydroxynaphthalene; **FPP**, farnesyl diphosphate; **GH**, glycoside hydrolase; **GO**, gene ontology; **GT**, glycosyltransferase; **HD**, homeodomain; **KEGG**, Kyoto Encyclopedia of Genes and Genomes; **Mya**,

million years ago; **PL**, polysaccharide lyase; **STS**, sesquiterpene synthases

### **Acknowledgments**

We would like to thank Editage (www.editage.cn) for English language editing.

### **Declaration of Competing Interest**

The authors declare that the research was conducted in the absence of any commercial or financial relationships, and there are no conflicts of interest.

### **Authors' contributions**

YD, HH, JQ, ZX and JZ collected and analyzed data. YD, CL wrote manuscript. ZX, JMG and PW interpreted the data and reviewed manuscript. CL edited the manuscript. All authors read and approved the final manuscript.

### **Funding**

This work was supported by the Fundamental Research Funds for the Central Universities (2572020DP07) and the National Natural Science Foundation of China (Project No. 31900064 and 31800031).

### **Availability of data and materials**

The *Inonotus obliquus* genomic data have been deposited under accession JAIHLT000000000 in GenBank. The version described in this paper is version JAIHLT010000000. The genome raw sequencing data and the reported assembly are associated with NCBI BioProject: PRJNA754990 and BioSample: SAMN20834359 within GenBank.

### **Declarations**

#### **Ethics approval and consent to participate**

Not applicable.

#### **Consent for publication**

Not applicable.

#### **Competing interests**

All authors declare that they have no competing interests.

### **Author details**

<sup>1</sup>School of life science, Northeast Forestry University, Harbin 150040, Heilongjiang, China

<sup>2</sup>Shaanxi Key Laboratory of Natural Products & Chemical Biology, College of Chemistry & Pharmacy, Northwest A&F University, Yangling 712100, Shaanxi, China



## References

1. Lee MW, Hur H, Chang KC, Lee TS, Jankovsky L. Introduction to distribution and ecology of sterile conks of *Inonotus obliquus*. *Mycobiology*. 2008;36(4):199-202.
2. Bai YH, Feng YQ, Mao D, Xu CP. Optimization for betulin production from mycelial culture of *Inonotus obliquus* by orthogonal design and evaluation of its antioxidant activity. *Journal of the Taiwan Institute of Chemical Engineers*. 2012;43(5):663-669.
3. Yana, Song, Jing, Hui, Wei, Kou, Ru, Xin, Fei, Jia. Identification of *Inonotus obliquus* and analysis of antioxidation and antitumor activities of polysaccharides. *Curr Microbiol*. 2008;57(5):454-62.
4. Lee J, Hyun C. Insulin-sensitizing and beneficial lipid-metabolic effects of the water-soluble melanin complex extracted from *Inonotus obliquus*. *Phytother Res*. 2014;28(9):1320-8.
5. Nakajima Y, Nishida H, Matsugo S, Konishi T. Cancer Cell Cytotoxicity of Extracts and Small Phenolic Compounds from Chaga [ *Inonotus obliquus* (persoon) Pilat]. *Journal of medicinal food*. 2009;12(3):501-507.
6. Zhao F, Mai Q, Ma J, Xu M, Wang X, Cui T, Qiu F, Han G. Triterpenoids from *Inonotus obliquus* and their antitumor activities. *Fitoterapia*. 2015;101:34-40.
7. Taji S, Yamada T, Wada SI, Tokuda H, Sakuma K, Tanaka R. Lanostane-type triterpenoids from the sclerotia of *Inonotus obliquus* possessing anti-tumor promoting activity. *Eur J Med Chem*. 2008;43(11):2373-9.
8. Taofiq O, Calhella RC, Heleno S, Barros L, Martins A, Santos-Buelga C, Queiroz MJRP, Ferreira ICFR. The contribution of phenolic acids to the anti-inflammatory activity of mushrooms: Screening in phenolic extracts, individual parent molecules and synthesized glucuronated and methylated derivatives. *Food Research International*. 2015;76:821-827.
9. Glamoclija J, Ciric A, Nikolic M, Fernandes A, Barros L, Calhella R, Ferreira I, Soković M, Van Griensven L. Chemical characterization and biological activity of Chaga (*Inonotus obliquus*), a medicinal “mushroom”. *J Ethnopharmacol*. 2015;162:323-32.
10. Lee IK, Kim YS, Jang YW, Jung JY, Yun BS. New antioxidant polyphenols from the medicinal mushroom *Inonotus obliquus*. *Bioorg Med Chem Lett*. 2007;17(24):6678-81.
11. Kou RW, Han R, Gao YQ, Li D, Gao JM: Anti-neuroinflammatory polyoxygenated lanostanoids from Chaga mushroom *Inonotus obliquus*. *Phytochemistry*. 2021;184:112647.
12. Burmasova MA, Utebaeva AA, Sysoeva EV, Sysoeva MA. Melanins of *Inonotus Obliquus*: Bifidogenic and antioxidant properties. *Biomolecules*. 2019;9(6):248.
13. Wold CW, Gerwick WH, Wangenstein H, Innngjerdigen KT. Bioactive triterpenoids and water-soluble melanin from *Inonotus obliquus* (Chaga) with immunomodulatory activity. *Journal of Functional Foods*. 2020;71:104025.
14. Chen S, Xu J, Liu C, Zhu Y, Nelson DR, Zhou S, Li C, Wang L, Guo X, Sun Y et al. Genome sequence of the model medicinal mushroom *Ganoderma lucidum*. *Nat Commun*. 2012;3:913.
15. Lu M-YJ, Fan W-L, Wang W-F, Chen T, Tang Y-C, Chu F-H, Chang T-T, Wang S-Y, Li M-y, Chen Y-H et al. Genomic and transcriptomic analyses of the medicinal fungus *Antrodia cinnamomea* for its metabolite biosynthesis and sexual development. *Proc Natl Acad Sci U S A*. 2014;111(44):E4743-52.
16. Gong W, Wang Y, Xie C, Zhou Y, Peng Y. Whole genome sequence of an edible and medicinal mushroom, *Hericium erinaceus* (Basidiomycota, Fungi). *Genomics*. 2020;112(3):2393-2399.
17. Shao Y, Guo H, Zhang J, Liu H, Wang K, Zuo S, Xu P, Xia Z, Zhou Q, Zhang H et al. The Genome of the medicinal macrofungus *Sanghuang* provides insights into the synthesis of diverse secondary metabolites. *Front Microbiol*. 2020;10:3035.

18. Luo H, Qian J, Xu Z, Liu W, Xu L, Li Y, Xu J, Zhang J, Xu X, Liu C. The *Wolfiporia cocos* genome and transcriptome shed light on the formation of its edible and medicinal sclerotium. *Genomics Proteomics Bioinformatics*. 2020;18(4):455-467.
19. James TY, Sun S, Li W, Heitman J, Kuo HC, Lee YH, Asiegbu FO, Olson A. Polyporales genomes reveal the genetic architecture underlying tetrapolar and bipolar mating systems. *Mycologia*. 2013;105(6):1374-1390.
20. Wold CW, Kjeldsen C, Corthay A, Rise F, Inngjerdingen KT. Structural characterization of bioactive heteropolysaccharides from the medicinal fungus *Inonotus obliquus* (Chaga). *Carbohydr Polym*. 2018;185:27-40.
21. Xu JW, Ji SL, Li HJ, Zhou JS, Duan YQ, Dang LZ, Mo MH. Increased polysaccharide production and biosynthetic gene expressions in a submerged culture of *Ganoderma lucidum* by the overexpression of the homologous  $\alpha$ -phosphoglucomutase gene. *Bioprocess Biosyst Eng*. 2015;38(2):399-405.
22. Wang Y, Yang X, Chen P, Yang S, Zhang H. Homologous overexpression of genes in *Cordyceps militaris* improves the production of polysaccharides. *Food Res Int*. 2021;147:110452.
23. Forster H, Shuai B. RNAi-mediated knockdown of  $\beta$ -1,3-glucan synthase suppresses growth of the phytopathogenic fungus *Macrophomina phaseolina*. *Physiological and Molecular Plant Pathology*. 2020;110:101486.
24. Zheng W, Zhang M, Zhao Y, Wang Y, Miao K, Wei Z. Accumulation of antioxidant phenolic constituents in submerged cultures of *Inonotus obliquus*. *Bioresour Technol*. 2009;100(3):1327-35.
25. Shen B. Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Curr Opin Chem Biol*. 2003;7(2):285-95.
26. Ebert MK, Spanner RE, de Jonge R, Smith DJ, Holthusen J, Secor GA, Thomma BPHJ, Bolton MD. Gene cluster conservation identifies melanin and perylenequinone biosynthesis pathways in multiple plant pathogenic fungi. *Environ Microbiol*. 2019;21(3):913-927.
27. Eisenman HC, Casadevall A. Synthesis and assembly of fungal melanin. *Appl Microbiol Biotechnol*. 2012;93(3):931-40.
28. Xu L, Li J, Chang M, Cheng Y, Geng X, Meng J, Zhu M. Comparison of physicochemical and biochemical properties of natural and arginine-modified melanin from medicinal mushroom *Ganoderma lucidum*. *J Basic Microbiol*. 2020;60(11-12):1014-1028.
29. Agnes D, Mukund KN, Nadkarni VS. Purification and characterisation of a sulphur rich melanin from edible mushroom *Termitomyces albuminosus* Heim. *Mycology*. 2018;9(4):296-306.
30. Weijn A, Bastiaan-Net S, Wichers HJ, Mes JJ. Melanin biosynthesis pathway in *Agaricus bisporus* mushrooms. *Fungal Genet Biol*. 2013;55:42-53.
31. Li J, Li Z, Zhao T, Yan X, Pang Q. Proteomic analysis of *Auricularia auricula-judae* under freezing treatment revealed proteins and pathways associated with melanin reduction. *Front Microbiol*. 2021;11:610173.
32. Zhao Y, Zheng W. Deciphering the antitumoral potential of the bioactive metabolites from medicinal mushroom *Inonotus obliquus*. *J Ethnopharmacol*. 2021;265:113321.
33. Wawrzyn G, Quin M, Choudhary S, López-Gallego F, Schmidt-Dannert C. Draft Genome of *Omphalotus olearius* provides a predictive framework for sesquiterpenoid natural product biosynthesis in basidiomycota. *Chem Biol*. 2012;19(6):772-83.
34. Flynn CM, Schmidt-Dannert AC. Sesquiterpene Synthase-3-Hydroxy-3-Methylglutaryl coenzyme a synthase fusion protein responsible for hirsutene biosynthesis in *Stereum hirsutum*. *Appl Environ Microbiol*. 2018;84(11):e00036-18.

35. Quin MB, Flynn CM, Wawrzyn GT, Choudhary S, Schmidt-Dannert C. Mushroom hunting using bioinformatics: Application of a predictive framework facilitates the selective identification of sesquiterpene synthases in Basidiomycota. *Chembiochem*. 2013;14(18):2480-91.
36. Zou CX, Wang XB, Lv TM, Hou ZL, Song SJ. Flavan derivative enantiomers and drimane sesquiterpene lactones from the *Inonotus obliquus* with neuroprotective effects. *Bioorg Chem*. 2020;96:103588.
37. Ying YM, Zhang LY, Zhang X, Bai HB, Liang DE, Ma LF, Shan WG, Zhan ZJ. Terpenoids with alpha-glucosidase inhibitory activity from the submerged culture of *Inonotus obliquus*. *Phytochemistry*. 2014;108:171-76.
38. Machida M, Asai K, Sano M, Tanaka T, Kumagai T, Terai G, Kusumoto KI, Arima T, Akita O, Kashiwagi Y. Genome sequencing and analysis of *Aspergillus oryzae*. *Nature*. 2005;438(7071):1157-61.
39. Kelly DE, Krasevec N, Mullins J, Nelson DR: The CYPome (Cytochrome P450 complement) of *Aspergillus nidulans*. *Fungal Genet Biol*. 2009;46:S53-61.
40. Min B, Kim S, Oh YL, Kong WS, Park H, Cho H, Jang KY, Kim JG, Choi IG. Genomic discovery of the hypsin gene and biosynthetic pathways for terpenoids in *Hypsizygus marmoreus*. *BMC Genomics*. 2018;19(1):789.
41. Chen J, Zeng X, Yang YL, Xing YM, Zhang Q, Li JM, Ma K, Liu HW, Guo SX. Genomic and transcriptomic analyses reveal differential regulation of diverse terpenoid and polyketides secondary metabolites in *Hericium erinaceus*. *Sci Rep*. 2017;7(1):10151.
42. Ichinose H, Wariishi H. Heterologous expression and mechanistic investigation of a fungal cytochrome P450 (CYP5150A2): Involvement of alternative redox partners. *Arch Biochem Biophys*. 2012;518(1):8-15.
43. Fessner ND, Nelson DR, Glieder A. Evolution and enrichment of CYP5035 in Polyporales: functionality of an understudied P450 family. *Appl Microbiol Biotechnol*. 2021;105(18):6779-6792.
44. Du D, Feng Z, Chen X, Ju X, Jiang J. Rapid isolation and purification of inotodiol and trametenolic acid from *Inonotus obliquus* by high-speed counter-current chromatography with evaporative light scattering detection. *Phytochem Anal*. 2011;22(5):419-423.
45. Yang J, Li Y, Zhang Y, Jia L, Sun L, Wang S, Xiao J, Zhan Y, Yin J. Functional identification of five CYP450 genes from birch responding to MeJA and SA in the synthesis of betulinic acid from lupitol. *Industrial Crops and Products*. 2021;167:113513.
46. Fujita S, Ohnishi T, Watanabe B, Yokota T, Takatsuto S, Fujioka S, Yoshida S, Sakata K, Mizutani M. Arabidopsis CYP90B1 catalyses the early C-22 hydroxylation of C27, C28 and C29 sterols. *Plant J*. 2006;45(5):765-74.
47. Yin Y, Gao L, Zhang X, Gao W. A cytochrome P450 monooxygenase responsible for the C-22 hydroxylation step in the *Paris polyphylla* steroidal saponin biosynthesis pathway. *Phytochemistry*. 2018;156:116-123.
48. Fujimoto, Yoshinori, Muranaka, Toshiya, Tsukagoshi, Yuki, Akashi, Tomoyoshi, Seki, Hikaru. Functional characterization of CYP71D443, a cytochrome P450 catalyzing C-22 hydroxylation in the 20-hydroxyecdysone biosynthesis of *Ajuga hairy* roots. *Phytochemistry*. 2016;127:23-8.
49. Ohnishi T, Watanabe B, Sakata K, Mizutani M. CYP724B2 and CYP90B3 function in the early C-22 hydroxylation steps of brassinosteroid biosynthetic pathway in tomato. *Biosci Biotechnol Biochem*. 2006;70(9):2071-80.
50. Chen W, Mi-Kyung L, Colin J, Sun-Chang K, Chen F, Jae-Hyuk Y. Fungal Cytochrome P450 Monooxygenases: Their distribution, structure, functions, family expansion, and evolutionary origin.

Genome Biol Evol. 2014;6(7):1620-34.

51. Zhang J, Li L, Lv Q, Yan L, Wang Y, Jiang Y. The Fungal CYP51s: Their functions, structures, related drug resistance, and inhibitors. *Front Microbiol.* 2019;10:691.

52. Fischer M. A new wood-decaying basidiomycete species associated with esca of grapevine: *Fomitiporia mediterranea* (Hymenochaetales). *Mycological Progress.* 2002;1(3):315-324.

53. Zhao S, Chang HP, Li X, Kim YB, Sang UP. Accumulation of rutin and betulinic acid and expression of phenylpropanoid and triterpenoid biosynthetic genes in mulberry (*Morus alba* L.). *J Agric Food Chem.* 2015;63(38):8622-30.

54. He P, Zhang Y, Li N. The phytochemistry and pharmacology of medicinal fungi of the genus *Phellinus*: a review. *Food Funct.* 2021;12(5):1856-1881.

55. Jin M, Zhou W, Jin C, Jiang Z, Diao S, Jin Z, Li G. Anti-inflammatory activities of the chemical constituents isolated from *Trametes versicolor*. *Nat Prod Res.* 2019;33(16):2422-2425.

56. Faruque AM. *Ganoderma lucidum*: Persuasive biologically active constituents and their health endorsement. *Biomed Pharmacother.* 2018;107:507-519.

57. Senol Cali D, Kim JS, Ghose S, Alkan C, Mutlu O. Nanopore sequencing technology and tools for genome assembly: computational analysis of the current state, bottlenecks and future directions. *Brief Bioinform.* 2019;20(4):1542-1559.

58. Vaser R, Sović I, Nagarajan N, Šikić M. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res.* 2017;27(5):737-746.

59. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One.* 2014;9(11):e112963.

60. Simão F, Waterhouse RM, Panagiotis I, Kriventseva EV, Zdobnov EM: BUSCO. assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics.* 2015;31(19):3210-3212.

61. Hoff KJ, Lange S, Lomsadze A, Borodovsky M, Stanke M. BRAKER1. Unsupervised RNA-Seq-Based Genome Annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics.* 2016;32(5):767-769.

62. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A et al. The Pfam protein families database in 2019. *Nucleic Acids Res.* 2019;47(D1):D427-432.

63. Galperin MY, Wolf YI, Makarova KS, Vera Alvarez R, Landsman D, Koonin EV. COG database update: focus on microbial diversity, model organisms, and widespread pathogens. *Nucleic Acids Res.* 2021;49(D1):D274-281.

64. Pundir S, Martin MJ, O'Donovan C. UniProt Tools. *Curr Protoc Bioinformatics.* 2016;53:1.29.1-1.29.15.

65. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 2017;45(D1):D353-361.

66. Attrill H, Gaudet P, Huntley RP, Lovering RC, Engel SR, Poux S, Van Auken KM, Georghiou G, Chibucos MC, Berardini TZ et al. Annotation of gene product function from high-throughput studies using the Gene Ontology. *Database (Oxford).* 2019;2019:baz007.

67. Haft DH, DiCuccio M, Badretdin A, Brover V, Chetvermin V, O'Neill K, Li W, Chitsaz F, Derbyshire MK, Gonzales NR et al. RefSeq: an update on prokaryotic genome annotation and curation. *Nucleic Acids Res.* 2018;46(D1):D851-860.

68. Jones P, Binns D, Chang HY, 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(9):1236-40.
69. Emms, David, M., Kelly, Steven. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol*. 2015;16(1):157.
70. Fowler TJ, Mitton MF, Vaillancourt LJ, Raper CA. Changes in mate recognition through alterations of pheromones and receptors in the multisexual mushroom fungus *Schizophyllum commune*. *Genetics*. 2001;158(4):1491-1503.
71. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. HMMER web server: 2018 update. *Nucleic Acids Res*. 2018;46(W1):W200-204.
72. Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, Busk PK, Xu Y, Yin Y. dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res*. 2018;46(W1):W95-101.
73. Park J, Lee S, Choi J, Ahn K, Park B, Park J, Kang S, Lee YH. Fungal cytochrome P450 database. *BMC Genomics*. 2008;9:402.
74. Medema MH, Blin K, Cimermancic P, de Jager V, Zakrzewski P, Fischbach MA, Weber T, Takano E, Breitling R. AntiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res*. 2011;39:W339-346.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile1.docx](#)
- [Additionalfile2.xlsx](#)
- [Additionalfile3.docx](#)
- [Additionalfile4.xlsx](#)