

Genome-wide analysis and prediction of genes involved in the biosynthesis of polysaccharides and bioactive secondary metabolites in high-temperature-tolerant of wild *Flammulina filiformis*

Juan Chen (✉ kibchenjuan@126.com)

Chinese Academy of Medical Sciences & Peking Union Medical College Institute of Medicinal Plant Development <https://orcid.org/0000-0002-5407-8239>

Jia-Mei Li

Chinese Academy of Medical Sciences & Peking Union Medical College Institute of Medicinal Plant Development

Yan-Jing Tang

Institute of Medicinal Plant Development, Chinese Academy of Medical Science

Ke Ma

Institute of Microbiology Chinese Academy of Sciences

Bing Li

Chinese Academy of Medical Sciences & Peking Union Medical College Institute of Medicinal Plant Development

Xu Zeng

Chinese Academy of Medical Sciences & Peking Union Medical College Institute of Medicinal Plant Development

Xiao-Bin Liu

Kunming Institute of Botany Chinese Academy of Sciences

Yang Li

Chinese Academy of Medical Sciences & Peking Union Medical College Institute of Medicinal Plant Development

Zhu-Liang Yang

Kunming Institute of Botany Chinese Academy of Sciences

Wei-Nan Xu

Fujian Agriculture and Forestry University

Bao-Gui Xie

Fujian Agriculture and Forestry University

Shun-Xing Guo

Chinese Academy of Medical Sciences & Peking Union Medical College Institute of Medicinal Plant Development

Hong-Wei Liu

Institute of Microbiology Chinese Academy of Sciences

Research article

Keywords: Edible mushroom, Gene cluster, Polysaccharide, Secondary metabolites, Gene expression, High-temperature-tolerance

Posted Date: December 11th, 2019

DOI: <https://doi.org/10.21203/rs.2.13424/v2>

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 October 17th, 2020. See the published version at <https://doi.org/10.1186/s12864-020-07108-6>.

Abstract

Background: *Flammulina filiformis* (=Asian “*F. velutipes*”) is a popular commercial edible mushroom. Many bioactive compounds with medicinal effects, such as polysaccharides and sesquiterpenoids, have been isolated and identified from *F. filiformis*, but their biosynthesis and regulation at the molecular level remains unclear. In this study, we sequenced the genome of the wild strain *F. filiformis* Liu355, predicated its the biosynthetic gene clusters (BGCs) and profiled the expression of these genes in wild and cultivar strains and in different developmental stages of the wild *F. filiformis* strain by a comparative transcriptomic analysis.

Results: We found that the genome of the *F. filiformis* was 35.01 M bp in length and harbored 10396 gene models. Thirteen putative terpenoid gene clusters were predicted and 12 sesquiterpene synthase genes belonging to four different groups and two type I polyketide synthase gene clusters were identified in the *F. filiformis* genome. The number of genes related to terpenoid biosynthesis was higher in the wild strain (119 genes) than in the cultivar strain (81 genes). Most terpenoid biosynthesis genes were upregulated in the primordium and fruiting body of the wild strain, while the polyketide synthase genes were generally upregulated in the mycelium of the wild strain. Moreover, genes encoding UDP-glucose pyrophosphorylase and UDP-glucose dehydrogenase, which are involved in polysaccharide biosynthesis, had relatively high transcript levels both in the mycelium and fruiting body of the wild *F. filiformis* strain.

Conclusions: *F. filiformis* is enriched in a number of gene clusters involved in the biosynthesis of polysaccharides and terpenoid bioactive compounds and these genes usually display differential expression between wild and cultivar strains, even in different developmental stages. This study expands our knowledge of the biology of *F. filiformis* and provides valuable data for elucidating the regulation of secondary metabolites in this unique *F. filiformis* strain.

Background

Flammulina filiformis (= Asian *F. velutipes*), a species endemic to the Asian region, also known as enokitake, winter mushroom or golden needling mushroom and belongs to the family Physalacriaceae, Agaricales [1]. Recent new phylogenetic results based on multi-gene markers and morphological comparisons demonstrated that “*F. velutipes*” in eastern Asia is not identical to the European winter mushroom *F. velutipes* and should be treated as a separate species, namely *F. filiformis*, which includes all cultivated enokitake strains in East Asia and those from South Korea and Japan with genome sequences [2]. Thus, we used the name “*F. filiformis*” instead of the Asian *F. velutipes* in our study.

F. filiformis is one of the most important and popular edible mushrooms available commercially in China. It is widely cultivated and consumed in Asian countries due to its high nutritional value and desirable taste. It has been reported that China is currently the largest producer of *F. filiformis*, with an annual production of 2.4 million tons [3]. *F. filiformis* also possesses tremendous pharmaceutical values, and many bioactive constituents have been identified such as polysaccharides [4–6], flavonoids [7],

sesquiterpenes, glycosides, proteins, and phenols [8–10]. These compounds have been shown to exhibit antitumour, anticancer, anti-atherosclerotic thrombosis inhibition, anti-aging and antioxidant effects [11, 12]. In addition, as a typical white-rot fungus, *F. filiformis* can effectively degrade lignin and produce alcohol dehydrogenase, and thus exhibiting potential for application in bioethanol production [13].

In recent decades, research has mainly focused on the phylogenetic taxonomy [1, 14], genetic diversity [15, 16], nutritional and chemical constituents [17–19], pharmacological bioactivity [20, 21] and artificial cultivation of *Flammulina* spp. [22–24]. Most studies have shown that *F. velutipes* possesses relatively high carbohydrate, protein and amino acids contents and low fat or lipid contents; thus, it is generally was recognized as a low energy delicacy [25]. In addition, bioactive polysaccharides (e.g., glucans and heteropolysaccharides), immunomodulatory proteins (e.g., FIP-fve) and multiple bioactive sesquiterpenes were also isolated and identified from the extracts, mycelia and fruiting bodies of *F. velutipes* [26]. Tang et al. [12] reviewed the compounds derived from the *F. velutipes* and their diverse biological activities. Increasing studies on the chemical compounds and biological activities of this mushroom have supported that *F. velutipes* should be exploited as a valuable resource for the development of functional foods, nutraceuticals and even pharmaceutical drugs [27].

The development of genomic and transcriptomic sequencing technologies has provided the powerful tools to understand the biology of edible mushrooms, including the effective utilization of cultivation substrates (lignocellulose) [28, 29], the mechanism of fruiting body formation and development and adaption to adverse environments, such as high temperature environments or cold-stress conditions [30–32]. For example, genome sequencing of the cultivars of *F. velutipes* from Korea and Japan revealed their high capacity for lignocellulose degradation [28, 33]. Transcriptomic and proteomic analyses of *F. velutipes* revealed key genes associated with cold- and light-stress fruiting body morphogenesis [34]. These studies provided important information for the breeding and commercial cultivation of *F. velutipes*.

Recent advances in genome sequencing have revealed that millions of putative biosynthetic gene clusters (BGCs) are hidden in fungal genomes [35, 36]. Genome mining efforts have also allowed us to understand the silencing or activation of biosynthetic pathways in microbes with the development of bioinformatics software, such as antiSMASH, SMURF and PRISM [37]. For instances, the genome-wide investigation of 66 cosmopolitan strains of *Aspergillus fumigatus* revealed 5 general types of variation in secondary metabolic gene clusters [38]. The identification of the tricyclic diterpene antibiotic pleuromutilin gene clusters on the genome-scale increased antibiotic production in *Clitopilus passeckerianus* [39]; the prediction of gene clusters involved in the biosynthesis of terperoid/ polyketide synthase (PKS) in the medicinal fungus *Hericium erinaceus* by genome and transcriptome sequencing discovered a new family of diterpene cyclases in fungi [40, 41], and the identification of the candidate cytochromes P450 genes cluster possibly related to triterpenoid biosynthesis in the medicinal mushroom *Ganoderma lucidum* by genome sequencing improved the production of effective medicinal compounds [42, 43].

However, as a popular edible mushroom that has a wide spectrum of interesting biological activities, little is known about the synthesis and regulation of bioactive secondary metabolites of *F. filiformis*. In previous experiments, we collected the wild strain of *F. filiformis* Liu355 from Longling, Yunnan and demonstrated that it could tolerate relatively high temperatures during fruiting body formation (at 18 °C–22 °C) in the laboratory and that its temperature tolerance was superior to that of the commercial strains of *F. filiformis* (Asian commercial *F. velutipes* usually produce fruiting body at low temperatures ≤ 15 °C) [16]. Thus, the wild strain is a potential and an important genetic material for future breeding or engineering of new *F. filiformis* strains because increasing the temperature tolerance can save substantial amount energy. Most interestingly, the chemical composition of the wild strain was different from that of other commercially cultivated strains of *F. filiformis*, harbouring more unique chemical compounds. A total of 13 new sesquiterpenes with noreudesmane, spiroaxane, cadinane, and cuparane skeletons were isolated and identified from the wild strain Liu355 [9].

Thus, the aims of our study are to explore the genetic features of this interesting wild strain of *F. filiformis* on a genomic scale, to identify gene clusters associated with the biosynthesis of secondary metabolites and to profile the expression differences in these candidate genes during the development of *F. filiformis*. This research will facilitate our understanding of the biology of the wild strain, provide useful datasets for molecular breeding, improving compound production and improve the production of novel compounds by heterologous pathway and metabolic engineering in the future.

Results

General features of the F. filiformis genome

Prior to our study, three “genomes” classified as *F. velutipes* were available in public databases; the relatively complete genome of the KACC42780 from Korea, L11 from China and a draft genome of TR 19 from Japan. In this study, we sequenced the genome of a wild strain of *F. filiformis* by small fragment library construction and performed a comparative genomic analysis of secondary metabolite gene clusters. The assembled genome of wild *F. filiformis* was 35.01 Mbp with approximately 118-fold genome coverage. A total of 10396 gene models were predicted, with an average sequence length of 1445 bp. The genome size and the number of predicted protein-encoding genes were very similar to the public reference genome of *F. filiformis* (= Asian *F. velutipes*) (Table 1). Functional annotation of the predicted genes showed that more than half of the predicted genes were annotated in NR database (6383 genes), and 1972, 2582, 837, and 5794 genes were annotated in the SwissProt, KEGG, COG and GO databases, respectively. We identified 107 cytochrome P450 family genes, 674 genes encoding secretory proteins and 287 genes in the CAZymes database.

There were 17293 pan-genes among the four strains of *F. filiformis*, and the pan-genome core comprised 4074 genes (on average 23.5% of each genome) (Fig. 1A). The proportion (23.5%) of core genes in the pan-genomic analysis of *F. filiformis* was similar to that in the pan-genomic analysis of 23 *Corallococcus* spp. [44]. However, the number was potentially lower than the actual number because these genomes

were not sequenced to completion. A total of 3104 orthologous genes were annotated in the KEGG database, 2722 genes were annotated in the GO database and 1055 genes were specific to the wild strain Liu355.

Functional characteristics of the predicted genes of F. filiformis

A KEGG enrichment analysis was performed to determine the functions of the predicted genes of *F. filiformis*. The results showed that the abundance of *F. filiformis* genes involved in genetic information processing and translation (253 genes) was the highest, followed by metabolism (carbohydrate metabolism with 243 genes). Twenty-one genes were involved in terpenoid and polyketide biosynthesis (Additional file 1).

Transcriptomic analysis and gene expression

We studied the gene expression differences across different developmental stages, the monokaryotic (MK), dikaryotic mycelium (DK), primordium (PD) and fruiting body (FB) stage, of the wild strain *F. filiformis* Liu355 at the transcriptomic level. Moreover, the DK of the cultivar strain of *F. filiformis* (CGMCC 5.642) was also subjected to transcriptome sequencing (Fig. 1B). Three biological replicates were designed for each sample. The average clean data for each sample was 8.07–9.32 G. We mapped the clean reads to the genome of *F. filiformis* Liu 355 using HISAT software and obtained a relatively high total mapping rate (92.63%). In addition, the expression variation between samples was the smallest between the MK and DK stages and was the greatest between the MK and FB stages of the wild *F. filiformis* strain (Additional file 2).

Among the 10396 gene models of *F. filiformis*, 9931 gene models were expressed (FPKM>5) across the four different tissues (MK, DK, PD and FB) of the wild strain and the mycelium of a cultivar strain of *F. filiformis*. A total of 6577 genes were commonly expressed in all tissues. 151 genes were specifically expressed in the cultivar strain, and 152, 116, 46, 199 genes were specifically expressed in MK, DK, PD and FB of the wild strain of *F. filiformis*, respectively (Fig. 1C). The tissue-specific and high expression transcripts in *F. filiformis* Liu355 are listed in Additional file 3. Two new genes encoding ornithine decarboxylase (involved in polyamine synthesis) were highly expressed in the mycelium of the cultivar strain (Novel01369, Novel01744), and the genes encoding oxidoreductase also had the highest expression level (gene 830, FPKM>1000). The genes encoding agroclavine dehydrogenase, acetylxylan exterase, beta-glucan synthesis-associated protein and arabinogalactan endo-1,4- β -galactosidase protein were significantly highly expressed in the FB of the wild *F. filiformis* strain, with a more than 20–100-fold change compared to their expression in the mycelium. Agroclavine dehydrogenase is involved in the biosynthesis of the fungal ergot alkaloid ergovaline [45] and β -glucan synthesis-associated protein is likely linked to the biosynthesis of fungal cell wall polysaccharide. The high expression of these genes indicates that they probably play an important role in fruiting body development and compound enrichment.

A total of 5131 genes (51.67%) were up or downregulated in at least in one stage of transition, such as from mycelium to primordium (PD vs DK, 3889 genes) and from primordium to fruiting body (FB vs PD, 3308 genes) (Fig. 1D). During primordial formation, 1780 genes are upregulated, and most of the genes were annotated as oxidoreductase activity (Go:0016491), hydrolase activity (Go:0004553) and carbohydrate metabolism (Go:0005975). The downregulated genes were mainly enriched in transmembrane transport (Go:0055085). During fruiting body development, genes related to the fungal-type cell wall (Go:0009277) and the structural constituent of the cell wall (Go:0005199) were upregulated, reflecting the dramatic changes in cell wall structure during the developmental process. In addition, Go term enrichment of differentially expressed genes (DEGs) between the wild strain Liu355 and cultivar strain CGMCC 5.642 showed that most genes displayed a similar expression profile, but peptide biosynthetic and metabolic process (Go:0006518; Go:0043043), amide biosynthetic process (Go:0043604) and ribonucleoprotein complex (Go:1901566) were upregulated in the cultivar strain of CGMCC 5.642.

KEGG enrichment analysis showed that DEGs involved in glutathione metabolism were significantly enriched in DK of the wild strain Liu 355 compared to the cultivar strain (Fig. 2). Thirty-three DEGs, including genes encoding glutathione S-transferase, ribonucleoside-diphosphate reductase, 6-phosphogluconate dehydrogenase, cytosolic non-specific dipeptidase, gamma-glutamyltranspeptidase, and glutathione peroxidase, participated in this pathway. A study on the glutathione metabolism in the filamentous fungus *Aspergillus nidulans* indicated that glutathione itself and glutathione metabolic enzymes play crucial roles in the germination of conidiospores and markedly contribute to the general stress tolerance of the fungus [46]. The high expression of genes related to glutathione metabolism in the wild strain of *F. filiformis* implied that the strain probably had strong environmental adaptation and was a potential better breeding resource. In addition, during the primordial and fruiting body development stages, the MAPK signaling pathway (45 DEGs) and starch and sucrose metabolism (26 DEGs) were significantly enriched. Tyrosine metabolism, biosynthesis of secondary metabolites and glycosphingolipid biosynthesis were also significantly enriched in the fruiting body formation stage.

Genes involved in polysaccharide biosynthesis in F. filiformis

Polysaccharides (PSs) are important and bioactive components of *F. filiformis* and other edible and medicinal mushrooms [47]. Glucans, such as lentinan, ganoderan, grifolan, and schizophyllan are well-studied PSs that can be produced by different fungi [48]. Previous studies have shown that *F. velutipes* polysaccharides (FVPs) have good immunomodulatory and anti-proliferation activities and improve learning and memory impairment in rats [4, 5]. Recently, an interesting result showed that FVPs could even be used as a material of edible film for packaging food and vegetables because of their excellent film-forming features and edible safety [49].

The biosynthetic pathways of polysaccharides (PSs) involve the synthesis of nucleotide sugar precursors, the assembly of repeating monosaccharide units, and the polymerization process [43]. Although the exact biosynthetic pathways of mushroom PSs remain unclear, polysaccharide synthesis is

not a template-driven process; instead the structures are determined primarily by the complement of polysaccharide-modifying enzymes present in any organism [50]. Phosphoglucomutase (PGM) and UDP-glucose pyrophosphorylase (UGP) are important enzymes in the biosynthetic pathway of nucleotide sugar precursors [47]. In addition, based on identification of intermediate compounds and the activities of synthesis-related enzymes, researchers constructed a simplified biosynthetic pathway of mushroom PSs and inferred several enzymes, such as glucokinase (GK), phosphoglucose isomerase (PGI), fructose-1,6-biphosphatase (FBP) and UGP that are most likely involved in the biosynthetic pathways of PSs of in edible mushrooms [43, 47]. These enzymes play key roles in glycolysis and gluconeogenesis pathways and in starch and carbohydrate metabolism.

We identified a total of 80 genes related to PS biosynthesis involved glycolysis and gluconeogenesis in the KEGG pathway analysis (KEGG map 00010) [51] at the genomic level, including glucose-6-phosphate isomerase (GPI), FBP, and mannose-6-phosphate isomerase (MPI). The expression profiles of these genes showed that genes encoding Zinc-type alcohol dehydrogenase were upregulated in both the mycelium of the wild strain compared to the cultivar strain and in the fruiting body compared to the mycelium of the wild of *F. filiformis* strain (Additional file 4 and Additional file 5). The genes encoding glycerol 2-dehydrogenase (gene9557, gene2028), 7-bisphosphatase (gene 2929), alcohol dehydrogenase (gene7891-D2, gene 9773-D2) and aryl-alcohol dehydrogenase (gene 4871, gene 612) were upregulated in mycelium of the wild strain. The expression level of the gene encoding mannose-1-phosphate guanylyltransferase (GDP) (gene 11132-D3) was the highest in the mycelium of the wild strain, with a more than 200-fold change compared to that in the mycelium of the cultivar strain. The genes encoding glycerol 2-dehydrogenase (gene 894) and sugar phosphatase (gene 11052-D2) were upregulated in the fruiting body stage of the wild strain.

Moreover, based on predicted metabolic enzymes related to PS biosynthesis in *Ganoderma lucidum* [47], we identified 21 putative essential enzymes involved in PS biosynthesis in *F. filiformis*, including GPI, MPI, UGD, UGP, hexokinase, galactokinase and transketolase, using a protein sequencing homology search method (Table 2). Among them, genes encoding UGP, UGD and fructose-bisphosphate aldolase (FDA) had relatively high transcript levels in all samples analysed (FPKM>100). These candidate genes will be functionally verified in the future by comparing the quantity of PS in different development stages and different strains of *F. filiformis*.

Carbohydrate active enzymes (CAZymes)

Secreted carbohydrate-degrading enzymes are crucial enzymes for fungal biology, both for fungal cell wall turnover and for the degradation of external carbohydrate sources. They are also responsible for the biosynthesis, degradation and modification of oligo- and polysaccharides and glycoconjugates [52]. We performed a sequencing similarity search against the CAZymes database using Blastx ($E < 10^{-5}$) and identified 287 carbohydrate-active related genes in the wild *F. filiformis* strain genome, including 149 glycoside hydrolases (GHs), 48 auxiliary activities (AAs) containing redox enzymes, 36 glycosyltransferases (GTs), 29 carbohydrate-binding modules (CBMs), 13 carbohydrate esters (CEs) and

12 polysaccharide lyases (PLs). The number of secreted carbohydrate-degrading enzymes (mostly GHs) was nearly four times that of polysaccharide biosynthetic enzymes (GTs), indicating their dependence on external carbohydrate sources for growth and development as wood-rot fungi. The genes encoding beta-1, 3-glucan-binding protein (GH16) were most abundant, with 23 members in the *F. filiformis* genome. Previous studies have shown that GH16 enzymes are involved in the formation of the chitin- β -1,3-glucan complex, suggesting a role in the processing of fungal cell wall polysaccharides [52]. The expression profiles of the genes related to CAZymes were diverse (Fig. 3). Seven genes of the 23 GH16 family genes were upregulated in the DK stage compared to the MK stage (three genes annotated as candidate glycosidase and four genes as β -glucan synthesis-associated protein), four genes of the GH16 family were upregulated in the fruiting body compared to the primordium of the wild strain Liu355.

In addition to GH16, seven out of the 14 GH5 members (four genes annotated as glucan 1,3- β -glucosidase, two genes annotated as endoglucanases and another annotated as a mannan endo-1, 4- β -mannosidase) were differentially upregulated in the DK mycelium of the wild strain compared to MK mycelium, and five genes belonging to the GH5 family were upregulated in the primordium compared to the mycelium. Eight members out of 11 GH43 genes were upregulated in the primordium compared to the mycelium of the wild strain.

Predicted bioactive secondary metabolite gene clusters of F. filiformis

In addition to macronutrients and micronutrients present in the *F. filiformis*, a large number of structurally diverse bioactive secondary metabolites, especially various novel sesquiterpenes and norsesquiterpenes, were identified from the mycelia and fruiting bodies of *F. filiformis* [9]. In this study, the gene clusters encoding terpenoid and PKS were examined using the genome and transcriptome resources in the wild strain Liu355. In total, 13 gene clusters related to terpenoid biosynthesis and two gene clusters for polyketide biosynthesis were predicted in the wild strain (Fig. 4 and Additional file 6). The numbers of gene clusters involved in terpene, PKS and NRPS biosynthesis were different in the wild strain Liu355 compared with the other three cultivar strains (KACC42780, TR19 and L11 with genome sequencing) and the gene number related to terpene synthesis was higher in the wild strain Liu355 (119 genes) than in the cultivar strain L11 (81 genes) (Table 3).

Putative genes for terpenoid biosynthesis in F. filiformis

A total of 119 genes of 13 terpenoid clusters were divided into 10 clades according to their expression levels in different developmental stages of the wild strain or cultivar strains (Additional file 7). Most genes in clade II, including encoding 4,5-DOPA dioxygenase extradiol-like protein (gene3103) and squalene synthase (gene3428), which is involved in the biosynthesis of squalene, a precursor of terpenoid compounds, were upregulated in the primordium compared to the mycelium of wild strain Liu355. The genes in clades VIII-X, including key enzymes involved in terpenoid biosynthesis, such as protoilludene synthase, candidate peroxisomal acyl-coenzyme A oxidase, L-amino acid amidase and cytochrome P450, were significantly differently expressed in the fruiting body compared with the mycelium of the wild strain Liu355 and in the mycelium of the wild strain Liu355 compared to the cultivar

strain CGMCC5.642. For example, the putative terpene synthase (gene 9115) were highly expressed and significantly upregulated in the mycelium of the wild strain (1.4-fold change) compared to the mycelium of the cultivar strain (CGMCC 5.642) and in the fruiting bodies compared with the primordium in the wild strain Liu355 (19.4-fold change). Two putative cytochrome P450 genes (gene 9114 and gene 7212) were also significantly upregulated in the dikaryotic mycelium compared to the monokaryotic mycelium of the wild strain (7.61- and 7.50- fold change, respectively) and the cultivar strain mycelium (2.1- and 1.7- fold change, respectively). This result likely explained the greater diversity of bioactive compounds in the wild strain than in the cultivar strain in a previous study.

Putative genes for sesquiterpenoid biosynthesis in F. filiformis

Sesquiterpene compounds are mainly bioactive secondary metabolites in *Flammulina*. The chemistry investigation of six strains (one wild strain and five other cultivar strains) of *F. filiformis* in a previous report revealed that the wild strain Liu355 contained many new sesquiterpenes with various skeletons, including cuparane-type and sterpurane-type sesquiterpenes [9]. However, little is known about sesquiterpene synthases (STSs) in Basidiomycota, although these fungi can produce diverse bioactive sesquiterpenes. The identification of sesquiterpene synthases from *Coprinus cinereus* and *Omphalotus olearius* provided useful guidance for the subsequent development of *in silico* approaches for the directed discovery of new sesquiterpene synthases and their associated biosynthetic genes [53]. In a previous study, 12 putative sesquiterpene synthase genes (Fla 1–12) were predicted from the genomic sequence of the cultivar strain of *F. filiformis* L11. In our study, we performed a genome-scale homologous search with sesquiterpene synthases of *O. olearius*, *C. cinereus* and *H. erinaceus* based on the genomic data of the wild strain Liu355. Similarly, 12 homologous sequences with considerable similarity (e-value < 10⁻⁵) to the known biochemically characterized sesquiterpene synthases were identified in the genome of the wild strain *F. filiformis* Liu355. Twelve STS genes of the wild *F. filiformis* strain included five genes encoding delta (6)-protoilludene synthase (gene1663-D2, gene9115, gene2784 and gene9115-D2, gene6325-D2), two genes encoding trichodiene synthase (gene1140, gene2254), two genes encoding alpha-muurolene synthase (gene1358-D2, gene1358), and one gene encoding a hypothetical protein (gene3100). The phylogenetic analysis showed that these genes consisted of four clades (Additional file 8). Five genes from the wild strain Liu355 (gene1140, gene10498-D2, gene4450, gene2785, gene2254) and two genes from the cultivar strain L11 (Fla10, Fla11) of *F. filiformis* were clustered together with the cuprenene synthases Cop6 and Omp 8-Omp10 in the clade 1. These genes are likely responsible for the 1,6- or 1,7-cyclization of 3R-nerolidyl diphosphate (NPP) (Cop6) or involved in the biosynthesis of α - and β -barbatene (Omp9), compounds known to be produced by fungi and plants and carotane sesquiterpenes (Omp10). Gene 1358-D2 (from the wild strain) and Fla6 (from the cultivar strain) in clade 2 grouped with Omp1 and Omp2 and were speculated to catalyse a 1,10-cyclization of E, E-farnesyl diphosphate (FPP) to yield cadinane, a precursor of sesquiterpenoids with noreudesmane, spiroaxane, cadinane and seco-cuparane. In a previous study, noreudesmane, spiroaxane, cadinane and seco-cuparane sesquiterpenoids were first identified as new compounds in the mycelium of the wild *F. filiformis* strain [9]. The enzymes from the wild strain (gene6325-D2, gene1663-D2, gene9115-D2 and

gene9115) and from the cultivar strain of *F. filiformis* (Fla2, Fla4, Fla7 and Fla12) that clustered with Omp6 and Omp7 in clade 3 may be capable of catalysing a 1,11-cyclization of (E, E)-FPP leading to major groups of bioactive sesquiterpenes in Basidiomycota [53]. Gene1358 and Fla9 of clade 4 clustered with Cop4, Omp4, Omp5a and Omp5b and may synthesize major compounds that require the 1, 10-cyclization of (3R)-nerolidyl diphosphate (NPP). From the previous study, we knew that the new sesquiterpenes from the wild strain of *F. filiformis* were mainly derived from 1,10-cyclization mechanisms [9]. Thus, the enzymes encoded by gene1358-D2 and gene1358 are probably responsible for the production of sesquiterpene with specific skeleton of the *F. filiformis*. The expression of gene1358-D2 and gene1358 in mycelium of the wild *F. filiformis* strain is significantly upregulated compared to cultivar strain based on our transcriptomic data (1.3- and 2.7- fold change, respectively) (Table 4). However, the differences in the expression of gene1358-D2 and gene1358 will be verified in different strains of *F. filiformis* by quantitative real-time PCR, in addition to the yield of these sesquiterpene compounds, and the reason for the wild strain possessing more new sesquiterpene compounds will be further explained.

Putative genes for polyketide biosynthesis in F. filiformis

Polyketides are a major group of secondary metabolites isolated from bacteria and fungi and fungi in particular have been important sources of natural product polyketide pharmaceuticals. The structural complexity of these polyketides makes them interesting and useful bioactive compounds and these same features also make them difficult and expensive to prepare and scale-up using synthetic methods [54]. In the past, understanding the biosynthesis and regulation of polyketides in fungi was extremely difficult because of the lack of established genetic tools. Recent studies have illustrated that *Saccharomyces cerevisiae* has potential as a tool for mining, studying and engineering fungal polyketide synthases due to its many advantages such as being a unicellular organism well suited for large-scale fermentation, having limited native secondary metabolism and having a number of developed genetic tools for protein expression and pathway constriction [54].

The diverse structures of polyketides are biosynthesized from short-chain carboxylic acid units by polyketide synthases (PKSs), PKSs have been classified into type I, type II and type III based on their product profiles and catalytic domain architecture [55]. By gene cluster prediction using antiSMASH, we found 30 genes in two gene clusters annotated as type I PKSs, and they were mainly located in a single scaffold, 24 and 78, in the *F. filiformis* genome, respectively (Fig. 4 and Additional file 6). The two gene clusters both included core genes encoding polyketide synthases (gene8217 and gene1373). The genes located on scaffold 78, including putative polyketide synthase (gene1373, gene1374) and benzoate 4-monooxygenase (gene 1372), were most upregulated in the wild strain mycelium compared with the cultivar strain mycelium, indicating that polypeptide compounds are probably abundant in the mycelia of this mushroom and especially in the wild strain.

Cytochrome P450s in the *F. filiformis* genome

It has been reported that a more divergent cytochrome P450 oxidase could be involved in secondary biosynthesis [56]. Therefore, we searched the genome of *F. filiformis* for proteins with a P450 conserved

domain using the NCBI CDD tool and BLASTp [40]. We identified 107 genes in the cytochrome P450 family, including nine putative trichodiene oxygenases, 31 O-methylsterigmatocystin oxidoreductases, five benzoate 4-monooxygenases, two linoleate 10R-lipoxygenases, two ent-kaurene oxidases, lanosterol 14- α and flavonoid hydroxylases and other candidate cytochrome P450s. Of these, 102 genes had diverse expression profiles across different tissues of *F. filiformis* (Additional file 7). Twenty-six CYP450 genes were upregulated in the mycelium of the wild strain compared to the cultivar strain and the gene54280-D3 had the highest expression level, with more than a 500-fold change. Twenty-one CYP450 genes were upregulated in the fruiting body stage compared to the mycelium stage, and genes encoding benzoate 4-monooxygenases had the highest transcript level, with 15-fold change. In the primordium stage, the gene encoding docosahexaenoic acid omega-hydroxylase was the highest differentially expressed gene.

Heat-shock proteins correspond to temperature changes in F. filiformis

In addition to unique compounds, the wild strain *F. filiformis* Liu355 can produce fruiting bodies at relatively higher temperatures than the current commercial strains, indicating that it is a potential excellent breeding resource. A temperature downshift (cold stimulation) is considered to be one of the most important and essential environmental factors for the fruiting initiation and fruiting body formation of *F. velutipes* [34]. In general, the mycelia of *F. velutipes* can grow vegetatively at 20–24 °C and fruit at an optimum temperature of 12–15 °C [32]. In our study, the wild strain of Liu355 could grow fruiting bodies at 18 °C–22 °C in the laboratory. Therefore, it is a potential excellent genetic material for *F. velutipes* breeding.

The heat-shock protein (HSP) family is known to be positively correlated with organism thermotolerance [57]. In this study, 28 genes annotated as HSPs were identified in the *F. filiformis* genome (Fig. 5). Among them, six genes were significantly upregulated in the wild strain Liu355 compared to the cultivar strain, including HSP12, HSP4, HSP104, LHS1 and GRP78, respectively. HSP12 is part of a group of small HSPs that function as chaperone proteins and are ubiquitously involved in nascent protein folding by protecting proteins from misfolding and are partially characterized as a stress response; the expression of HSP12 protein was observed in response to cold stress [58]. In *S. cerevisiae* and *C. albicans*, HSP104, in association with HSP40 and HSP70, helps in the reactivation and aggregation of denatured protein by providing disaggregated protein to HSP40 and HSP70 as a substrate [59]. The expression of HSP104 and HSP70 is regulated by the Hsf (heat-shock factor) interaction, which can be stimulated by heat stress in yeast [57].

In addition, HSP70 chaperone and two putative HSP were also found were upregulated at only primordium or young fruiting body of cultivar *F. velutipes* using the iTRAQ labeling technique [60]. However, the exact molecular function of HSPs in the high-temperature tolerance of wild *F. filiformis* and its adaptive mechanisms for relatively high temperatures need further study.

Conclusion

In our study, genome and transcriptome sequencing and the assembly and annotation of the high-temperature-tolerant wild *F. filiformis* strain were carried out, and the gene clusters associated with polysaccharides, terpenoid and polyketide biosynthesis were predicted. Comparative genomic analysis with three other Asian cultivar strains of *F. velutipes* (= *F. filiformis*) revealed that the wild strain has a similar genome size and relatively more putative gene numbers related to secondary metabolite biosynthesis. Most genes related to terpenoid biosynthesis were upregulated in the primordium and fruiting body of the wild strain, while PKS genes were generally upregulated in the mycelium of the wild strain; however, the specific regulatory pathways involved such synthesis pathways remain unresolved in this study.

Six genes belonging to the HSP family, including HSP12, HSPC4, HSP104, LHS1 and GRP78, were significantly upregulated in the wild strain Liu355 compared to the cultivar strain and may be responsible for the development of fruiting bodies at relatively high temperatures in the high-temperature-tolerant wild *F. filiformis* strain. However, the expression of these genes in other strains of *F. filiformis*, especially in strains under low-temperature developmental conditions, requires verification in future studies.

Our study provides an important genetic dataset for *F. filiformis* as a potential breeding material, and provides a foundation for enhancing the understanding of the biology of *F. filiformis*.

Materials And Methods

Fungal strains and strain culture

The wild strain Liu355 used for genomic sequencing was kindly provided by Prof. H. W. Liu (State Key Laboratory of Mycology, Institute of Microbiology, Chinese Academy of Sciences) and was first isolated from the fruiting body of *F. filiformis* collected from Longling, Yunnan Province, southwestern China. The voucher specimens of the *F. filiformis* fruiting bodies were deposited in the Cryptogamic Herbarium, Kunming Institute of Botany, Chinese Academy of Sciences (HKAS85819). Its sequence was listed under GenBank accession number KP867925 [9]. The haploid monokaryotic strain *F. filiformis* Liu355 (deposited in Mycological Laboratory of the Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences) was prepared by the protoplast mononuclear method and was grown on potato dextrose agar (PDA) at room temperature for 2–3 weeks in the dark. The fruiting bodies were obtained in sterile plastic bottles containing on growth substrate (cotton seed hulls, 78%; wheat bran, 20%; KH₂PO₄, 0.1%; MgSO₄, 0.1%; sucrose, 1%; and ground limestone, 1%; with a moisture content of 70%) at 25 °C for 30 d, followed by cold stimulation at 18 °C and 90% humidity until primordial development occurred. Cultures were maintained at low temperature (18 °C and 75% humidity) to allow full fruiting body development [61]. In addition, the genomic data of two cultivar strains from Korea (KACC42780, Bioproject PRJNA191921) and Japan (TR19, Bioproject PRJDB4587) were available from the NCBI public database, and the genomic sequence of strain L11 (Bioproject PRJNA191865) was kindly provided by the Mycological Research Center, College of Life Sciences, Fujian Agriculture and Forestry University [62]. The

cultivar dikaryotic strain (CGMCC 5.642) was obtained from the China General Microbiological Culture Collection Center (Beijing, China, <http://www.cgmcc.net/>) and stored in our laboratory.

Genome and transcriptome sequencing and analysis

Total genomic DNA of *F. filiformis* was extracted from the mononuclear mycelia in PDA medium using the Omega E. Z. N. A. fungal DNA midi kit (Omega, USA) according to the manufacturer's instructions. Total DNA was evaluated by agarose gel electrophoresis and quantified by Qubit. The *F. filiformis* strain was sequenced using 350 bp paired-end reads on an Illumina HiSeq 4000 platform via the PE150 method. Library construction and sequencing was performed at the Beijing Novogene Bioinformatics Technology Co. Ltd. (China). After removing adapters and low-quality sequences using FastQC, the high-quality reads were mapped to the reference genome sequence of *F. filiformis* L11 (Bioproject PRJNA191865) using BWA software. Functional annotation of the predicted genes was performed using BLAST against Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), SwissProt and NCBI non-redundant (NR) protein databases [40]. Pan-genomic analysis was carried out using the standalone CD-HIT tool to cluster orthologous proteins [63].

For transcriptomic sequencing, total RNA was extracted using the RNeasy Plant Mini kit (Qiagen) according to the manufacturer's protocols. Five samples were prepared; the monokaryotic mycelium (MK), dikaryotic mycelium (DK), primordium (PD) and fruiting bodies (FB) of the wild strains Liu355 and the dikaryotic mycelium of the cultivar strain CGMCC 5.462. Each sample had three biological replicates. All samples were subjected to RNA-Seq on the Illumina HiSeq2000 platform (Illumina, San Diego, CA, USA). Bioinformatics analysis was performed based on clean data with high quality, and the RNA-seq reads were mapped to the *F. filiformis* genome (Liu355) using TopHat v2.0.1253 [64]. HTSeq v0.6.1 software was used to count the read numbers mapped to each gene [65]. The FPKM value was used to calculate gene expression, and the upper-quartile algorithm was used to correct the gene expression. Gene differential expression analysis was performed using the DESeq R package (1.18.0) using a corrected p-value [66]. Genes with an adjusted P-value < 0.05 were considered differentially expressed. Hierarchical clustering of gene expression was conducted using Genesis 1.7.7 [67].

Prediction of gene clusters

The biosynthetic gene clusters were predicted using antiSMASH 3.0 software [68]. AntiSMASH currently offers a broad collection of tools and databases for automated genome mining and comparative genomics for a wide variety of different classes of secondary metabolites [69]. In addition, the homology sequence search method (Blast) was also used to identify genes related to terpenoid biosynthesis. The sesquiterpene synthases were identified based on multiple sequence alignments and phylogenetic analyses developed by the Schmidt-Dannert group [53].

Abbreviations

AA: Auxiliary Activities; *BGCs*: Biosynthetic gene clusters; *CAZymes*: Carbohydrate-Active Enzymes; *CBM*: Carbohydrate-binding module; *CE*: Carbohydrate Esterases; *CDD*: Conserved Domain Database; *COG*: Clusters Orthologous Groups; *CYP*: Cytochrome P450 proteins; *DEGs*: Differentially expressed genes; *DK*: dikaryotic mycelium; *FB*: Fruiting body; *FBP*: Fructose-1,6-bisphosphatase; *FDA*: Fructose-bisphosphate aldolase; *FPKM*: Fragments Per Kilobase of transcript per Million fragments mapped; *FVPs*: *F. velutipes* polysaccharides; *GDP*: mannose-1-phosphate guanylyltransferase; *GH*: Glycoside-Hydrolases; *GK*: glucokinase; *GO*: Gene Ontology; *GPI*: Glucose-6-phosphate isomerase; *GT*: glycosyltransferases; *HSP*: The heat shock protein; *Hsf*: heat-shock factor; *KEGG*: Kyoto Encyclopedia of Genes and Genomes; *MAPK*: mitogen-activated protein kinase; *MK*: monokaryotic; *MPI*: Mannose-6-phosphate isomerase; *PD*: primordium; *PDA*: potato dextrose agar; *PGI*: phosphoglucose isomerase; *PGM*: Phosphoglucomutase; *PKS*: Polyketide synthase; *PSs*: Polysaccharides; *UGP*: UDP-glucose pyrophosphorylase; *PL*: Polysaccharide Lyases; *STSs*: sesquiterpene synthases; *ITS*: internal transcribed spacer;

Declarations

Acknowledgments

The authors are grateful to Prof. Francis Martin for critical discussion and suggestion.

Funding

This study was supported by National Key R & D Program of China (2017YFC1701900), the Natural Science Foundation of China (81573527; 81973423), Peking Union Medical College Discipline Construction Project (201920100901) and National Basic Research Program of China (2014CB138304).

Author's contributions

JC, HWL and SXG designed the experiments. JML, YJT, BL, XL, YL and WNX performed the experiment. XBL, ZLY and HWL collected the sample and provided the strain. WNX and BGX sequenced and provided the reference genome of *F. velutipes* L11. JC, JML, KM and XZ analyzed the data, JC contributed to text writing and literature search and SXG reviewed and revised the manuscript.

Availability of data and materials

The datasets supporting the results of this article are included with in the article and additional files. The genomic and transcriptomic data have been deposited to GenBank database with the dataset identifier PRJNA531555 for genome and PRJNA530834 for transcriptome.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Ge ZW, Liu XB, Zhao K, Yang ZL. Species diversity of *Flammulina* in China: new varieties and a new record. *Mycosystema*. 2015;34:589–603.
2. Wang PM, Liu XB, Dai YC, Horak E, Steffen K, Yang ZL. Phylogeny and species delimitation of *Flammulina*: taxonomic status of winter mushroom in East Asia and a new European species identified using an integrated approach. *Mycological Progress*. 2018;17(9):1013–1030.
3. Li X, Li Y. Quality comparison and analysis on white *Flammulina velutipes* grown with bottle lines in China. *Edible Fungi China*. 2014;33:20–24 (In Chinese).
4. Lin L, Cui F, Zhang J, Gao X, Zhou M, Xu N, Zhao H, Liu M, Zhang C, Jia L. Antioxidative and renoprotective effects of residue polysaccharides from *Flammulina velutipes*. *Carbohydr Polym*. 2016;46:388–395.
5. Su A, Yang W, Zhao L, Pei F, Yuan B, Zhong L, Ma G, Hu Q. *Flammulina velutipes* polysaccharides improve scopolamine-induced learning and memory impairment in mice by modulating gut microbiota composition. *Food Funct*. 2018;9(3):1424–1432.
6. Zhang T, Ye J, Xue C, Wang Y, Liao W, Mao L, Yuan M, Lian S. Structural characteristics and bioactive properties of a novel polysaccharide from *Flammulina velutipes*. *Carbohydr Polym*. 2018;197:147–156.
7. Hu Q, Yu J, Yang W, Kimatu BM, Fang Y, Ma N, Pei F. Identification of flavonoids from *Flammulina velutipes* and its neuroprotective effect on pheochromocytoma–12 cells. *Food Chem*. 2016;204:274–282.
8. Wang Y, Bao L, Yang X, Li L, Li S, Gao H, Yao XS, Wen H, Liu HW. Bioactive sesquiterpenoids from the solid culture of the edible mushroom *Flammulina velutipes* growing on cooked rice. *Food Chem*. 2012;132(3):1346–1353.
9. Tao Q, Ma K, Yang Y, Wang K, Chen B, Huang Y, Han J, Bao L, Liu XB, Yang Z, Yin WB, Liu H. Bioactive sesquiterpenes from the edible mushroom *Flammulina velutipes* and their biosynthetic pathway confirmed by genome analysis and chemical evidence. *J Org Chem*. 2016;21:81(20): 9867–9877.
10. Li HP, Yang WJ, Qu SX, Pei F, Luo X, Mariga AM, Ma L. Variation of volatile terpenes in the edible fungi mycelia *Flammulina velutipes* and communications in fungus-mite interactions. *Food Res Int*. 2018;103:150–155.
11. Rahman MA, Abdullah N, Aminudin N. Antioxidative effects and inhibition of human low density lipoprotein oxidation in vitro of polyphenolic compounds in *Flammulina velutipes* (Golden Needle Mushroom). *Oxid Med Cell Longev*. 2015;403023.

12. Tang C, Hoo PC, Tan LT, Pusparajah P, Khan TM, Lee LH, Goh BH, Chan KG. Golden needle mushroom: A culinary medicine with evidenced-based biological activities and health promoting properties. *Front Pharmacol.* 2016;7:474.
13. Kasprzycka A, Lalak-Kańczugowska J, Tys J. *Flammulina velutipes* treatment of non-sterile tall wheat grass for enhancing biodegradability and methane production. *Bioresour Technol.* 2018;263:660–664.
14. Avin FA, Bhassu S, Shin TY, Sabaratnam V. Molecular classification and phylogenetic relationships of selected edible Basidiomycetes species. *Mol Biol Rep.* 2012;39(7):7355–7364.
15. Liu XB, Feng B, Li J, Yan C, Yang ZL. Genetic diversity and breeding history of winter mushroom (*Flammulina velutipes*) in China uncovered by genomic SSR markers. *Gene.* 2016;591:227–235.
16. Wang Q, Zhang J, Li C, Wang B, Nong W, Bian Y, Xiao Y. Phenotypic and genetic diversity of the culinary-medicinal winter mushroom *Flammulina velutipes* (Agaricomycetes) in China. *Int J Med Mushrooms.* 2018;20(6):517–536.
17. Wang Y, Bao L, Liu D, Yang X, Li S, Gao H, Yao X, Wen H, Liu H. Two new sesquiterpenes and six norsesquiterpenes from the solid culture of the edible mushroom *Flammulina velutipes*. *Tetrahedron.* 2012;68(14):3012–3018.
18. Reis FS, Barros L, Martins A, Ferreira IC. Chemical composition and nutritional value of the most widely appreciated cultivated mushrooms: an inter-species comparative study. *Food Chem Toxicol.* 2012;50(2):191–197.
19. Tsai SY, Huang EW, Lin CP. Compositional differences of the winter culinary-medicinal mushroom, *Flammulina velutipes* (Agaricomycetes), under three types of light conditions. *Int J Med Mushrooms.* 2017;19(3): 267–276.
20. Miyazawa N, Yoshimoto H, Kurihara S, Hamaya T, Eguchi F. Improvement of diet-induced obesity by ingestion of mushroom chitosan prepared from *Flammulina velutipes*. *J Oleo Sci.* 2018;67(2):245–254.
21. Wu M, Luo X, Xu X, Wei W, Yu M, Jiang N, Ye L, Yang Z, Fei X. Antioxidant and immunomodulatory activities of a polysaccharide from *Flammulina velutipes*. *J Tradit Chin Med.* 2014;34(6):733–740 (In Chinese).
22. Huang Q, Jia Y, Wan Y, Li H, Jiang R. Market survey and risk assessment for trace metals in edible fungi and the substrate role in accumulation of heavy metals. *J Food Sci.* 2015;80(7):H1612–1618.
23. Rugolo M, Levin L, Lechner BE. *Flammulina velutipes*: An option for “alperujo” use. *Rev Iberoam Micol.* 2016;33(4):242–247.
24. Xie C, Gong W, Yan L, Zhu Z, Hu Z, Peng Y. Biodegradation of ramie stalk by *Flammulina velutipes*: mushroom production and substrate utilization. *AMB Express.* 2017;7: 171.
25. Kalač P. A review of chemical composition and nutritional value of wild-growing and cultivated mushrooms. *J Sci Food Agric.* 2013;93(2):209–218.
26. Wang YQ, Bao L, Yang XL, Guo H, Dai HQ, Guo H, Yao XS, Zhang LX, Liu HW. Four new cuparene-type sesquiterpenes from *Flammulina velutipes*. *Helvetica Chimica Acta.* 2012;95:261–267.

27. Tung CH, Lin CC, Tung CC, Chen SF, Sheu F, Lu TJ. Combination of on-line desalting and HPLC-UV-ESI-MS for simultaneous detection and identification of FIP-fve and flammutoxin in *Flammulina velutipes*. *J Food Drug Anal.* 2018;26(3):1045–1053.
28. Park YJ, Baek JH, Lee S, Kim C, Rhee H, Kim H, Seo JS, Park HR, Yoon DE, Nam JY, et al. Whole genome and global gene expression analyses of the model mushroom *Flammulina velutipes* reveal a high capacity for lignocellulose degradation. *PLoS ONE.* 2014; 9(4): e93560.
29. Xie C, Yan L, Gong W, Zhu Z, Tan S, Chen D, Hu Z, Peng Y. Effects of different substrates on lignocellulosic enzyme expression, enzyme activity, substrate utilization and biological efficiency of *Pleurotus eryngii*. *Cell Physiol Biochem.* 2016;39(4):1479–1494.
30. Song HY, Kim DH, Kim JM. Comparative transcriptome analysis of dikaryotic mycelia and mature fruiting bodies in the edible mushroom *Lentinula edodes*. *Sci Rep.* 2018;8(1): 8983.
31. Wu TH, Ye ZW, Guo LQ, Yang XQ, Lin JF. De novo transcriptome sequencing of *Flammulina velutipes* uncover candidate genes associated with cold-induced fruiting. *J Basic Microbiol.* 2018;58:698–703.
32. Liu JY, Meng JL, Chang MC, Feng CP, Yuan LG. iTRAQ-based quantitative proteome revealed metabolic changes of *Flammulina velutipes* mycelia in response to cold stress. *J Proteomics.* 2017;156:75–84.
33. Kurata A, Fukuta Y, Mori M, Kishimoto N, Shirasaka N. Draft genome sequence of the basidiomycetous fungus *Flammulina velutipes* TR19. *Genome Announc.* 2016; 4(3):e00505–16.
34. Liu JY, Chang MC, Meng JL, Feng CP, Wang Y. A Comparative proteome approach reveals metabolic changes associated with *Flammulina velutipes* mycelia in response to cold and light stress. *J Agric Food Chem.* 2018; 66(14):3716–3725.
35. Doroghazi JR, Albright J, Goering AW, Ju KS, Haines RR, Tchalukov KA, Labeda DP, Kelleher NL, Metcalf WW. A road map for natural product discovery based on large-scale genomics and metabolomics. *Nat Chem Biol.* 2014;10(11):963–968.
36. 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.
37. Baral B, Akhgari A, Metsä-Ketelä M. Activation of microbial secondary metabolic pathways: Avenues and challenges. *Synth Syst Biotechnol.* 2018;3(3):163–178.
38. Lind AL, Wisecaver JH, Lameiras C, Wiemann P, Palmer JM, Keller NP, Rodrigues F, Goldman GH, Rokas A. Drivers of genetic diversity in secondary metabolic gene clusters within a fungal species. *PLoS Biol.* 2017;15(11):e2003583.
39. Bailey AM, Alberti F, Kilaru S, Collins CM, de Mattos-Shiple K, Hartley AJ, Hayes P, Griffin A, Lazarus CM, Cox RJ, et al. Identification and manipulation of the pleuromutilin gene cluster from *Clitopilus passeckerianus* for increased rapid antibiotic production. *Sci Rep.* 2016;6:25202.
40. 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 *Herichium erinaceus*. *Sci Rep.* 2017;7:10151.

41. Yang YL, Zhang S, Ma K, Xu Y, Tao Q, Chen Y, Chen J, Guo S, Ren J, Wang W, et al. Discovery and characterization of a new family of diterpene cyclases in bacteria and fungi. *Angew Chem Int Ed*. 2017;56:4749–4752.
42. 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.
43. Ma Z, Ye C, Deng W, Xu M, Wang Q, Liu G, Wang F, Liu L, Xu Z, Shi G, et al. Reconstruction and analysis of a genome-scale metabolic model of *Ganoderma lucidum* for improved extracellular polysaccharide production. *Front Microbiol*. 2018;9:3076.
44. Livingstone PG, Morphew RM, Whitworth DE. Genome sequencing and pan-Genome analysis of 23 *Corallocooccus* spp. strains reveal unexpected diversity, with particular plasticity of predatory gene sets. *Front Microbiol*. 2018; 199:3187.
45. Lorenz N, Wilson EV, Machado C, Schardl CL, Tudzynski P. Comparison of ergot alkaloid biosynthesis gene clusters in *Claviceps* species indicates loss of late pathway steps in evolution of *C. fusiformis*. *Appl Environ Microbiol*. 2007;73:7185–7191.
46. Bakti F, Király A, Orosz E, Miskei M., Emri T, Leiter É, Pócsi I. Study on the glutathione metabolism of the filamentous fungus *Aspergillus nidulans*. *Acta Microbiol Immunol Hung*. 2017;64(3):255–272.
47. Ruthes AC, Smiderle FR, Iacomini M. Mushroom heteropolysaccharides: A review on their sources, structure and biological effects. *Carbohydr Polym*. 2016;136:358–375.
48. Wang Q, Wang F, Xu, Z, Ding Z. Bioactive mushroom polysaccharides: a review on monosaccharide composition, biosynthesis and regulation. *Molecules*. 2017;22(6):pii: E955.
49. Du H, Hu Q, Yang W, Pei F, Kimatu BM, Ma N, Fang Y, Cao C, Zhao L. Development, physiochemical characterization and forming mechanism of *Flammulina velutipes* polysaccharide-based edible films. *Carbohydrate Polymers*. 2016;152:214–221.
50. Zeidan AA, Poulsen VK, Janzen T, Buldo P, Derkx PMF, Øregaard G, Neves AR. Polysaccharide production by lactic acid bacteria: from genes to industrial applications. *FEMS Microbiol Rev*. 2017;41(Supp_1): S168–S200.
51. Kanehisa M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res*. 2017;45:D353–D361.
52. Veneault-Fourrey C, Commun C, Kohler A, Morin E, Balestrini R, Plett J, Danchin E, Coutinho P, Wiebenga A, de Vries RP. et al. Genomic and transcriptomic analysis of *Laccaria bicolor* CAZome reveals insights into polysaccharides remodeling during symbiosis establishment. *Fungal Genet Biol*. 2014;72:168–181.
53. Wawrzyn GT, Quin MB, 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;19:772–783.
54. Bond C, Tang Y, Li L. *Saccharomyces cerevisiae* as a tool for mining, studying and engineering fungal polyketide synthases. *Fungal Genet Biol*. 2016; 89:52–61.

55. Shen B. Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Curr Opin Chem Biol.* 2003;7:285–295.
56. Shin J, Kim JE, Lee YW, Son H. Fungal Cytochrome P450s and the P450 Complement (CYPome) of *Fusarium graminearum*. *Toxins (Basel).* 2018;10(3):pii: E112.
57. Shui W, Xiong Y, Xiao W, Qi X, Zhang Y, Lin Y, Guo Y, Zhang Z, Wang Q, Ma Y. Understanding the mechanism of thermotolerance distinct from heat shock response through proteomic analysis of industrial strains of *Saccharomyces cerevisiae*. *Mol Cell Proteomics.* 2015;14:A779–A780.
58. Tiwari S, Thakur R, Shankar J. Role of heat-shock proteins in cellular function and in the biology of fungi. *Biotechnol. Res Int.* 2015;132635.
59. Glover JR, Lindquist S. Hsp104, Hsp70, and Hsp40: a novel chaperone system that rescues previously aggregated proteins. *Cell.* 1998; 94(1):73–82.
60. Liu J, Chang M, Meng J, Feng C, Zhao H, Zhang M. Comparative proteome reveals metabolic changes during the fruiting process in *Flammulina velutipes*. *J Agric Food Chem.* 2017; 65(24):5091–5100.
61. Luo R, Guo L, Lin J, Han F, Li Q, Kang L. A Novel high-temperature-tolerant Strain of *Flammulina velutipes* by mutagenesis. *Edible Fungi of China.* 2016;35(4):18–23(In Chinese).
62. Liu F. Preliminary study of *Flammulina velutipes* genome and transcriptome. PhD Dissertation, Fujian Agriculture and Forestry University, Fuzhou. 2014;1–117 (In Chinese).
63. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics.* 2012;28:3150–3152.
64. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg S L. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 2013;14(4):R36.
65. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics.* 2015;31(2):166–169.
66. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet.* 2009;14(1):57–63.
67. Sturn A, Quackenbush J, Trajanoski Z. Genesis: cluster analysis of microarray data. *Bioinformatics.* 2002;18:207–208.
68. Weber T, Blin K, Duddela S, Krug D, Kim HU, Brucoleri R, et al. antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* 2015;43:W237–243.
69. Blin K, Kim HU, Medema MH, Webe T. Recent development of antiSMASH and other computational approaches to mine secondary metabolite biosynthetic gene clusters. *Brief Bioinform.* 2019;20(4):1103–1113.

Tables

Due to technical limitations, Tables 1 - 4 are only available for download from the Supplementary Files section.

Supplementary Files Legend

*Additional file 1: A KEGG functional annotation of the predicted genes of *F. filiformis*.* The highest number of genes related to metabolism process and carbohydrate metabolism except for genetic information processing.

*Additional file 2: Relationships among five transcriptomes samples of *F. filiformis*.* Pairwise correlation of normalized FPKMs between RNA samples. The Pearson correlation coefficient ranges from no correlation (white) to perfect correlation (dark blue). Each sample has three biological replicates. The abbreviation: MK, monokaryotic mycelium; DK, Dikaryotic mycelium; PD, primordium; FB, Fruiting body.

*Additional file 3: Tissue-specific expression transcript in four different tissues of *F. filiformis*.*

*Additional file 4: KEGG mapping (map00010) of glycolysis/gluconeogenesis pathway [51] identified in *F. filiformis* and the putative gene expression level on different tissue of *F. filiformis*.* Red stars indicate the hits of differentially expressed genes in this map. The expression level of mapped genes (EC 5.4.2.2, EC 5.3.1.9, EC 5.3.1.1, EC 4.2.1.11, EC 1.2.4.1, EC 1.1.1.1, EC1.1.1.2) in different tissues was displayed in map. Abbreviation: F1, dikaryotic mycelium of cultivar strain CGMCC 5.642; F2, dikaryotic mycelium of wild strain Liu355; F3, fruiting body of wild strain Liu355; F4, primordium of wild strain Liu355; F5, monokaryotic mycelium of wild strain Liu355. The red and green colors indicate up-and down-regulation; black represents no significant expression change. Detail information of about the gene can be found in Additional file 5 (Note: obtained appropriate copyright permission to use the map from KEGG).

*Additional file 5: The expression of genes involved in glycolysis and gluconeogenesis pathway in *F. filiformis* genome.*

*Additional file 6: Predicted biosynthetic gene clusters involved in terpene, PKS, NRPS and siderophore in *F. filiformis* using antiSMASH tool.*

*Additional file 7: Hierarchical clustering analysis of 119 putative genes related to terpenoid biosynthesis in *F. filiformis* genome.* Expression ratios were plotted in a heatmap on a log₂ scale. The red and green colors indicate up- and down-regulation, black represents no significant expression change and grey represents missing data. The abbreviation: MK, monokaryotic mycelium; DK, Dikaryotic mycelium; PD, primordium; FB, Fruiting body. Detail information of about the gene annotation can be found in Additional file 6.

*Additional file 8: Neighbor-Joining phylogram of putative sesquiterpene synthases (STS) of *F. filiformis* were constructed based homologous protein sequences.* The number along branch represent the bootstrap value above 50%. The gene encoding sesquiterpene synthases with red dot were identified in this study. Detail information of the sequences used in phylogram can be found in reference [40, 53].

Labeled in “Cop” from the fungus *Coprinopsis cinereus*; “Omp” from fungus *Omphalotus olearius* and “Sh” from fungus *Stereum hirsutum*.

Figures

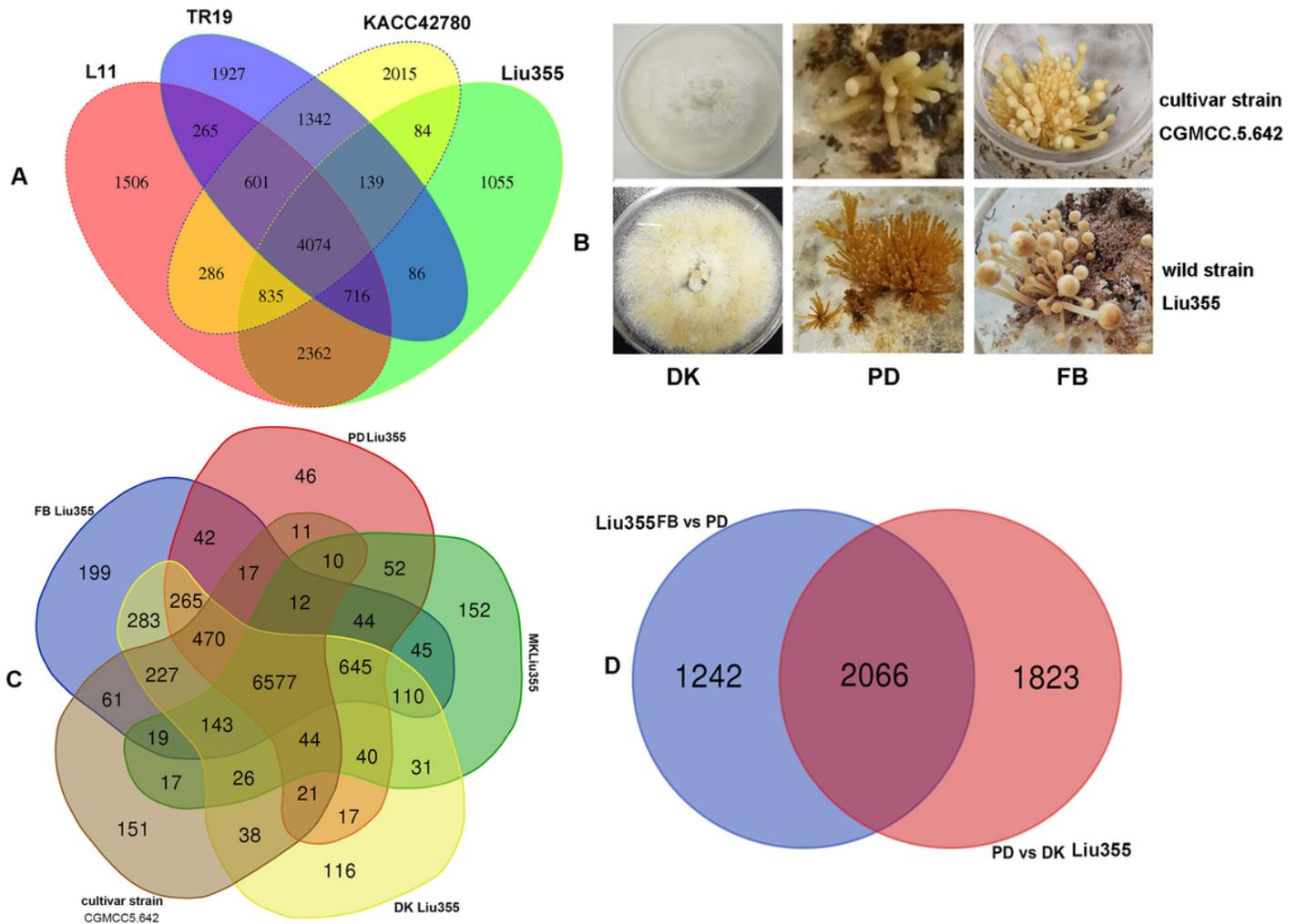


Figure 1

Samples information and venn diagram showing the numbers of orthologue genes or differentially expressed genes. (A) The numbers of orthologue genes between four strains of *F. filiformis*. L11(China) in red, TR19 (Japan) in purple, KACC42780 (Korea) in yellow and Liu355 (China) in green. (B) The samples of wild and cultivar strains of *F. filiformis*. Up-line: cultivar strain; down-line: wild strain, from left to right: Dikaryotic mycelium (DK); Primordium (PD) and Fruiting bodies (FB). (C) The numbers of differentially expressed genes (DEGs) in various comparative groups of *F. filiformis*. Fruiting body of wild strain (FB) in blue, Primordium of wild strain in red, Monokaryotic mycelium(MK) in green, Dikaryotic mycelium (DK) in yellow and cultivar strain of *F. filiformis* in brown. (C). Venn diagram showing the numbers of DEGs at adjacent development stage of *F. filiformis*. Blue color represented the number of DEGs of fruiting body

(FB) versus primordium (PD) and red color represented primordium (PD) versus dikaryotic mycelium (DK) of the wild *F. filiformis* strain. Abbreviation: MK: monokaryotic mycelium; DK: dikaryotic mycelium; PD: primordium; FB: fruiting body.

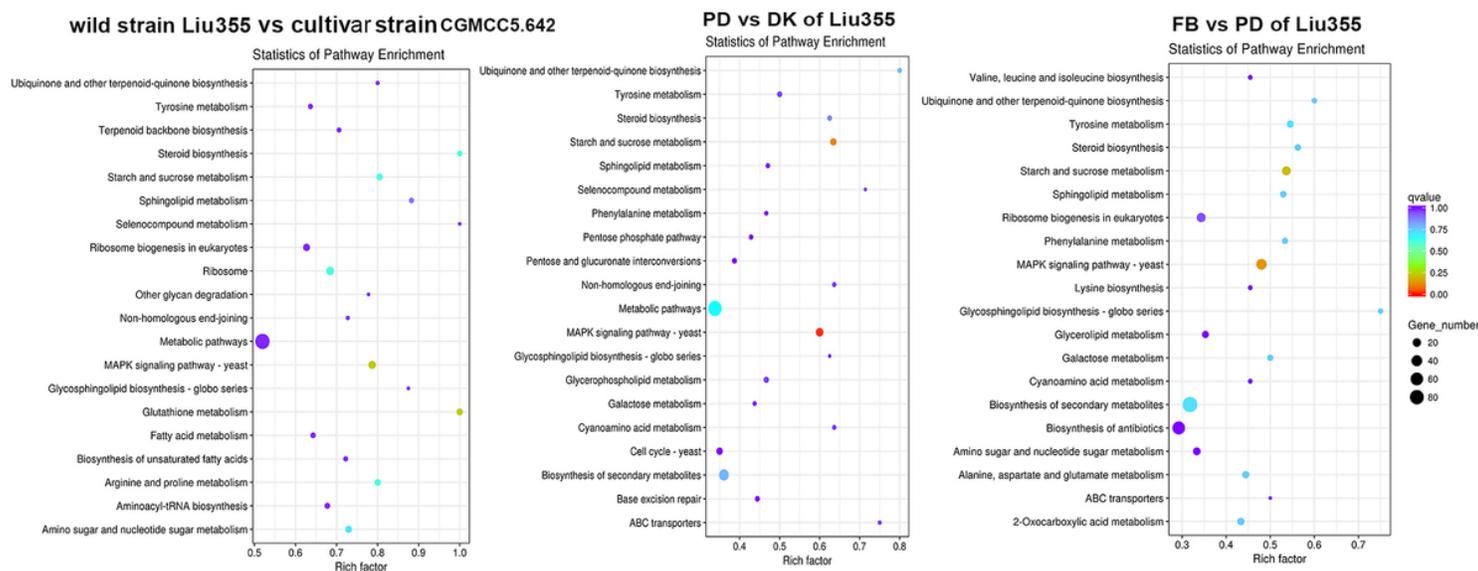


Figure 2

KEGG pathway enrichment analysis of differentially expressed genes (DEGs) during *F. filiformis* development. Left columns: pathway enrichment at mycelium stage of wild strain Liu355 compared to cultivar strain CGMCC 5.642; Middle columns: pathway enrichment at primordium stage compared to mycelium stage of wild strain Liu355; Right columns: pathway enrichment at fruiting body stage compared to primordium stage. Abbreviation: MK: monokaryotic mycelium; DK: dikaryotic mycelium; PD: primordium; FB: fruiting body.

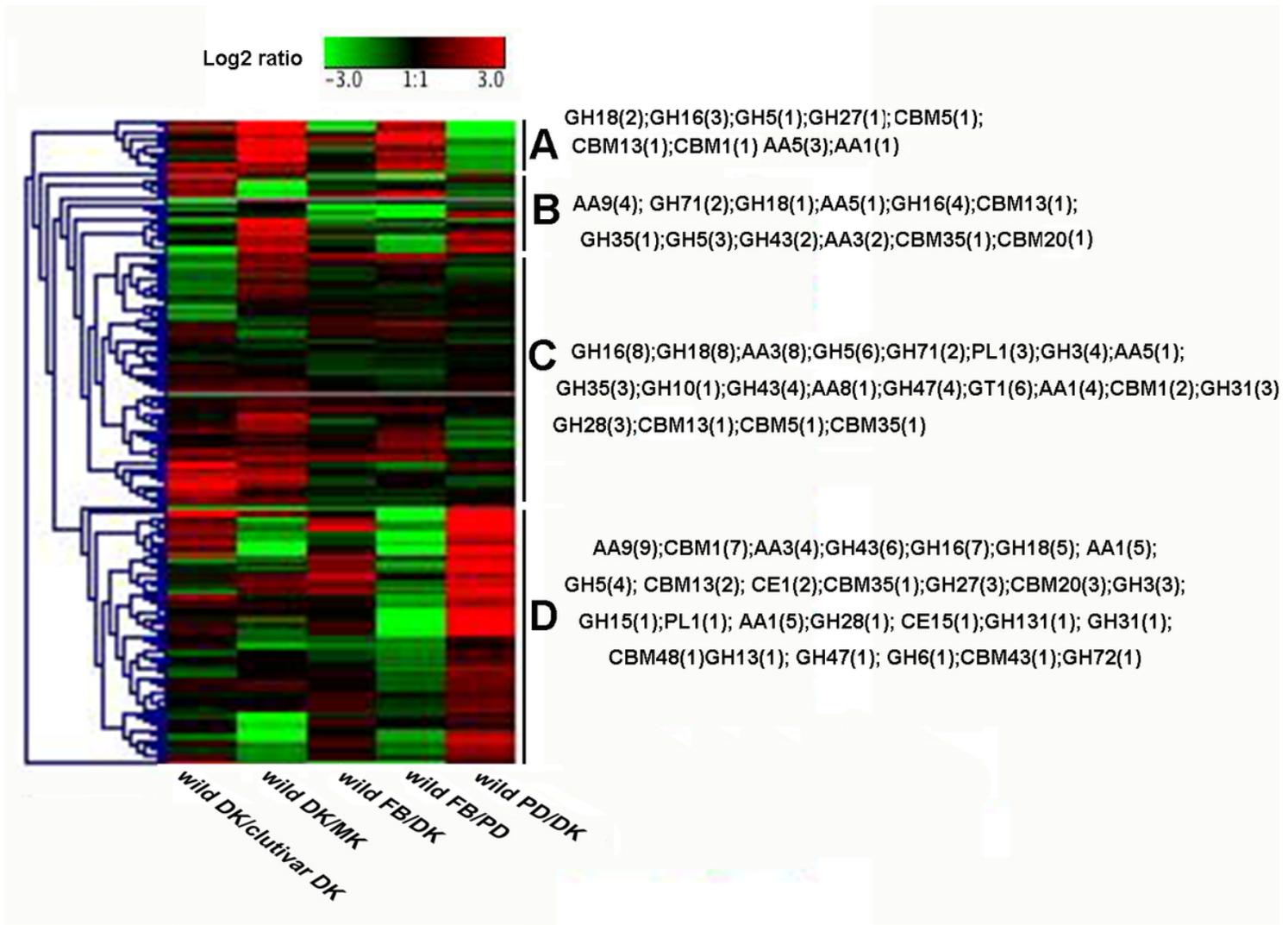


Figure 3

Hierarchical clustering analysis of 287 genes encoding CAZymes of *F. filiformis* between wild strain Liu355 and cultivar strain CGMCC 5.642 and among four development stages of wild strain Liu355. Expression ratios were plotted in a heatmap on a log2 scale. The red and green colors indicate up- and down-regulation, black represents no significant expression change and grey represents missing data. The abbreviation: MK, monokaryotic mycelium; DK, Dikaryotic mycelium; PD, primordium; FB, Fruiting body. GH, Glycoside-Hydrolases; AA, Auxiliary Activities containing redox enzymes; GT, glycosyltransferases; CBM Carbohydrate-binding module; CE, Carbohydrate Esterases (CE); PL, Polysaccharide Lyases.

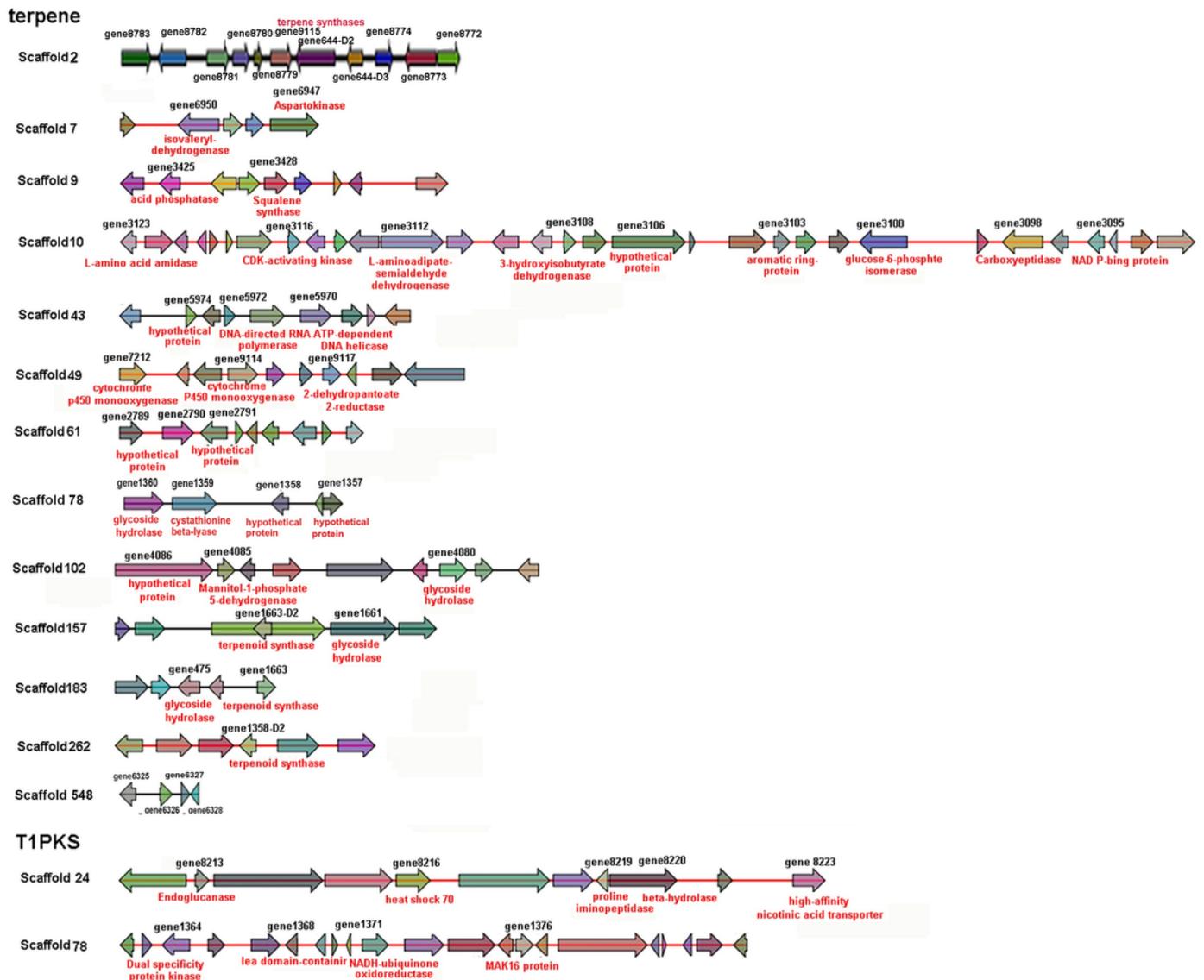


Figure 4

Identification of the 13 putative gene clusters for terpene and two polyketides gene clusters (PKS) in *F. filiformis* genome by antiSMASH software. Genes with SwissProt functional annotation were marked in red color.

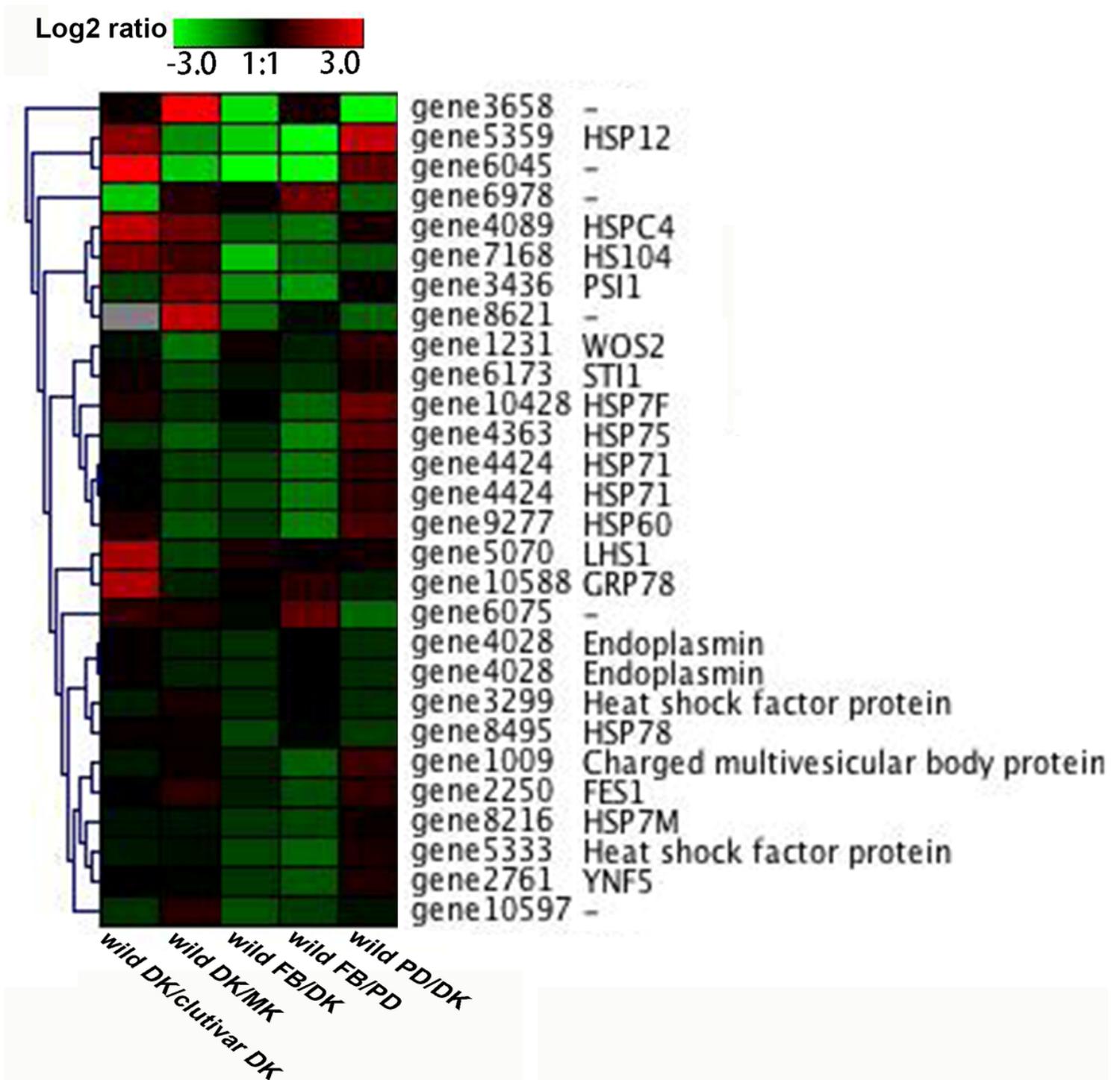


Figure 5

Hierarchical clustering analysis of 28 putative heat-shock protein encoding genes in *F. filiformis* genome between wild strain Liu355 and cultivar strain CGMCC5.642 and among four development stages of wild strain Liu355. Expression ratios were plotted in a heatmap on a log2 scale. The red and green colors indicate up- and down-regulation, black represents no significant expression change and grey represents missing. The abbreviation: MK, monokaryotic mycelium; DK, Dikaryotic mycelium; PD, primordium; FB, Fruiting body.

Supplementary Files

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

- [Table3.pdf](#)
- [Table2.pdf](#)
- [Table1.pdf](#)
- [Table4.pdf](#)