Transcriptome and proteome analyses reveal the regulatory networks and metabolite biosynthesis pathways during the development of *Tolypocladium guangdongense*

Gangzheng Wang, Min Li, Chenghua Zhang, Huijiao Cheng, Yu Gao, Wangqiu Deng, Taihui Li

**A B S T R A C T**

*Tolypocladium guangdongense* has a similar metabolite profile to *Ophiocordyceps sinensis*, a highly regarded fungus used for traditional Chinese medicine with high nutritional and medicinal value. Although the genome sequence of *T. guangdongense* has been reported, relatively little is known about the regulatory networks for fruiting body development and about the metabolite biosynthesis pathways. In order to address this, an analysis of transcriptome and proteome at differential developmental stages of *T. guangdongense* was performed. In total, 9076 genes were found to be expressed and 2040 proteins were identified. There were a large number of genes that were significantly differentially expressed between the mycelial stage and the stages. Interestingly, the correlation between the transcriptomic and proteomic data was low, suggesting the importance of the post-transcriptional processes in the growth and development of *T. guangdongense*. Among the genes/proteins that were both differentially expressed during the developmental process, there were numerous heat shock proteins and transcription factors. In addition, there were numerous proteins involved in terpenoid, ergosterol, adenosine and polysaccharide biosynthesis that also showed significant downregulation in their expression levels during the developmental process. Furthermore, both tryptophan and tryptamine were present at higher levels in the primordium stage. However, indole-3-acetic acid (IAA) levels continuously decreased as development proceeded, and the enzymes involved in IAA biosynthesis were also clearly differentially downregulated. These data could be meaningful in studying the molecular mechanisms of fungal development, and for the industrial and medicinal application of macro-fungi.

© 2020 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

In China, the fruiting bodies of some *Cordyceps* s.l. spp. are well-known traditional medicines because of their anticancer, immunomodulation, and anti-impotence effects [1–3]. Up to now, only four species of *Cordyceps* s.l., namely *C. militaris*, *C. cica-dae*, *Ophiocordyceps sinensis*, and *Tolypocladium guangdongense* previously named as *C. guangdongensis*, have been approved by the Chinese government as medicines or edible fungi, although hundreds of species have been found. *T. guangdongense* grows on the hypogeous tubers of *Elaphomyces* fungi, and is different from other *Cordyceps* s.l. species, which grow on insects [4]. To date, all four of these species have been successfully cultivated, and *C. militaris* has the largest cultivated production, and the rest have only relatively very low cultivated yields [5–9].

The fruiting body of *T. guangdongense*, which was approved as a novel food source by the Chinese government in 2013, contains numerous active and nutritional compounds such as high levels of cordycepic acid, adenosine, polysaccharides, micro-elements, and vitamins [10–11]. Based on animal tests, it has been shown that its fruiting body had a clear positive effect in improving the symptoms of several diseases, such as infection with the influenza virus H9N2 and chronic renal failure [12–13]. In addition, based on
2. Materials and methods

2.1. Sample collection at differential developmental stages

T. guangdongense strain CCTCCM206051 was used as the test strain, and was deposited in the China Center for Type Culture Collection. T. guangdongense was cultivated as described in a previous study [5]. After growth for 20 d on PDA medium at 23°C, mycelia from T. guangdongense were transferred into rice medium and incubated for 6 days in the dark at 23°C and 60–70% relative humidity. For primordium formation and fruiting body growth, the cultivation room was ventilated daily for 30 min and illuminated (500 lx) for 10 h. Using a previously described method [30], samples of the mycelium (My), primordium (P), young fruiting body (YFB) and mature fruiting body (MBF) were collected. In addition, a sample of the stroma with length 2.5–4 cm was designated as the developed fruiting body (DFB). Three replicate samples were prepared for each developmental stage. After collection, all samples representing the five differentially developmental stages were immediately frozen in liquid nitrogen and then stored at −80°C prior to RNA and protein isolation.

2.2. RNA isolation and sequencing

Total RNA was extracted from all samples using a TRIzol Kit (Invitrogen, Dalian, China) according to the manufacturer’s instruction. After qualitative and quantitative assessment of the RNA samples using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA), the mRNAs were enriched using oligo-dT magnetic beads, and then fragmented and transcribed into cDNA using random oligonucleotides. The five cDNA libraries were sequenced on a BGISEQ-500 platform at Beijing Genomics Institute (BGI)-Shenzhen, Wuhan, China.

2.3. Transcriptome data processing

After removing reads containing poly-N, sequencing adapters and low-quality reads from the raw data, the clean reads were mapped to the T. guangdongense genome using HISAT2 and Bowtie2 [31–32]. The RSEM package was used for calculating the expression levels of genes, and the differential expression analysis was performed using DESeq2 with a Q value < 0.05 [33–34]. Genes with a false discovery rate (FDR) of ≤ 0.001 and an absolute value of log2Ratio ≥ 1 were designated as differentially expressed genes (DEGs). The analysis of gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEEG) pathway enrichment was performed using Phyper based on the Hypergeometric test [35]. In addition, a weighted correlation network analysis (WGCNA) was performed to analyze the gene co-expression network and identify hub genes [36].

2.4. Protein extraction

Mycelia, primordia and developmental fruiting bodies (0.1 g) from each group (three replicates for each group) were homogenized in 0.5 mL of extraction buffer (20 mM Tris–HCl pH = 7.4, 100 mM EDTA, 2% β-mercaptoethanol, 1 mM DTT and 1% Triton X-100). Following this, 1 mL of saturated phenolic Tresturttated Phenol was added, and the mixture was shaken well for 15 min, followed by centrifugation at 25000 × g for 15 min at 4°C to obtain the supernatant (phenol phase). Then, 5 volumes of pre-cooled 0.1 M ammonium acetate in methanol supplemented with 10 mM dithiothreitol (DTT) was added to precipitate the proteins at −20°C for 2 h. The pelleted proteins were washed two times with pre-cooled acetone. After air-drying of the precipitate, a 1 × Cocktail (Sigma, St Louis, MO, USA) was added and the samples
was incubated in an ice bath for 5 min, and then centrifuged at 25000 × g for 15 min at 4 °C. Subsequently, 10 mM DTT was added to the precipitate which was then incubated in a water bath at 56 °C for 1 h. The sample was then incubated with 55 mM iodoacetamide for 45 min in the dark. The supernatant containing proteins was quantified using the Bradford method and SDS-PAGE was used to assess the quality of the protein extract [18].

2.5. Proteomic data processing

The final protein solution was diluted with 0.5 M TEAB and then digested with trypsin (Promega, Madison, WI, USA) (enzyme: protein ratio of 1:20) at 37 °C for 2 h. Peptides were labelled using the IBT method as previously described [37]. A mass spectrometric analysis of soluble intracellular proteins was undertaken, as previously described [38]. Raw MS/MS data was queried against the T. guangdongense genome using the Mascot search engine MaxQuant (version 2.3.02), and was deposited in the ProteomeXchange Consortium (identifier number PXD020391). The peptides labeled with IBTs were quantitatively analyzed using the iQuant software [38]. Proteins with a P < 0.05 and a fold change > 1.50 or < 0.67 were defined as differentially expressed proteins (DEPs). These proteins were classified and grouped using the GO and KEGG databases, and a hypergeometric test was used to analyze the GO and KEGG pathway annotations. The interaction network and subcellular location were predicted using the STRING protein interaction database and WoLF PSORT software [39–40].

2.6. Quantification of tryptophan and IAA

Samples (0.5 g) of mycelia, primordia and fruiting bodies were ground and then shaken for 15 min in 1 mL of buffer containing 50% methanol and 0.1% formic acid. The mixture was subjected to ultrasonication for 30 min in an ice bath to ensure complete homogenization of the sample. The samples were kept at 4 °C for 30 min, and centrifuged at 10000 × g for 20 min, and the supernatant was collected. The samples were filtered using a 0.3 μm filter plate and were then analyzed by LC-MS/MS (Waters Iclass-AB Sciex 6500) at the Beijing Genomics Institute (BGI)-Shenzhen, Wuhan, China. MultiQuant was used for chromatographic peak extraction and quantitative analysis of tryptophan, tryptamine, and IAA. Standards of these three metabolites were purchased from Yuanye (Shanghai, China).

2.7. Quantitative real time polymerase chain reaction (qRT-PCR) verification of gene expression

The mRNA prepared for transcriptome sequencing was also used for qRT-PCR analysis. Sixteen DEGs were selected to verify the reliability of the transcriptome. The gene-specific primers used for gene quantification were designed using Primer Primer5.0. All primers are listed in Table S1. Histone H4 was used as the internal control [41]. The qRT-PCR reaction was performed using Applied Biosystems ABI 7500 (ABI, Foster City, CA, USA) with three biological and technical repeats. The PCR reaction was performed with an initial denaturation for 2 min at 95 °C; 40 cycles of 20 s at 95 °C and annealing at 60 °C for 20 s. The relative expression levels of target genes were calculated using the 2^{-ΔΔCt} (Ct, cycle threshold) method as described by Livak [42].

3. Results

3.1. Transcriptome sequencing analysis

To investigate the gene expression changes in T. guangdongense during the developmental process, triplicate samples from the five developmental stages were used to construct a total of 15 cDNA libraries after which transcriptome sequencing was performed using BGISEQ-500 sequencing, yielding a total of 334.3 million raw reads. After removing reads containing adapter sequences, poly-N, and the low-quality reads, 330.29 million clean reads were obtained, with a mean of 22.02 million clean reads for each replicate (Table S2). Approximately 80% of the reads from the mycelia could be mapped to the T. guangdongense genome, whereas the mapping ratios of the reads from the other samples were 97.46% (Table S2). The Q30 percentages for all the clean reads in the 15 libraries were >93%. Gene expression levels between replicates for each sample exhibited a high Pearson’s correlation coefficient value (Fig. S1), indicating good repeatability between replicates. (See Fig. 1)

more than 99% of genes with FPKM values (>0) were detected in the 15 samples (Fig. 1A and Table S3). Approximately 86.31%, 98%, 98%, 98.13% and 98.37% of the genes were expressed in the My, P, YFB, DFB and MFB stages, respectively (Fig. 1A). A total of 8817 genes were expressed across the five differential developmental stages (Fig. 1B). Based on the gene FPKM values (Table S3), the expression levels of all genes were classified into four categories. During the developmental process, the majority of genes with FPKM values (10 > FPKM > 100) were considered to be moderately expressed. Approximately 17% of genes with FPKM value > 100 were considered to be highly expressed, and approximately 25% of genes with FPKM values (0 < FPKM < 10) were considered to have low levels of expression. The number of non-expressed genes (FPKM = 0 or undetected) accounted for around 2% of the total genes at each developmental stage.

3.2. Identification and functional classification of differentially expressed genes across the five developmental stages

Using a principal component analysis (Fig. S2), it was noted that the samples from the five developmental stages, each with three biological replicates, could be clearly divided into two groups, with the mycelial stage separately grouped into one cluster. As shown in Fig. 2A and Table S4, the number of DEGs between the My and P stages was the largest (3673), accounting for 40.14% of the T. guangdongense genes, followed by the DFB vs. YFB stages (400). Fewer DEGs were found in comparisons between the MFB and DFB stages (257) and between the YFB and P stages (189). The greatest number of unique DEGs (3157) was found between the MY and P stages, whereas only 25 unique DEGs were found between the YFB and P stages (Fig. 2B). In addition, only 12 DEGs were shared among the four comparisons between adjacent development stages (Fig. 2B). These results suggest that the gene expression profile at the mycelial stage is unique compared to the other developmental stages, whereas the gene expression profiles across the four other developmental stages are very similar to one another with only minor differences being found.

Classification of DEGs using GO annotation was performed to evaluate the potential functions of these DEGs across the five developmental stages. As shown in Fig. 3A and Table S5, for biological processed (BP) a total of 15 categories were found, including metabolic process (1043), cellular process (994), localization (319), biological regulation (224), cellular component organization or biogenesis, (188) and response to stimulus (183). For cellular component (CC), the largest of the 13 categories was related to membrane (1124). Other categories were related to cell (982), organelle (690) and macromolecular complex (332). For molecular function (MF), catalytic activity (1838) and binding activity (1374) were the largest categories. In comparisons between the My and P stages and between the DFB and MFB stages, the most significantly enriched GO assignment was a CC related to membrane (Fig. 3B and Table S5). The most significantly enriched GO
terms for the P vs. YFB stages and the YFB vs. DFB stages were in the MF category involved in catalytic activity. In addition, oxidoreductase and monoxygenase activities were found in a comparison of the P vs. YFB stages.

A KEGG pathway analysis was performed to further assess the functions of the DEGs across the five developmental stages (Table S6). In all four of the comparisons of two adjacent development stages, the two pathways with the greatest number of DEGs were related to MAPK signaling and amino sugar and nucleotide sugar metabolism. This was followed by endocytosis and RNA transport in three of the four comparisons (My vs P, YFB vs DFB and DFB vs MFB), whereas biosynthesis of antibiotics and mitophagy were found in the comparison of the P vs. YFB stages. The KEGG pathway enrichment also showed that the diterpenoid pathway was enriched in the comparison of P vs. YFB (P value of 0.002287), and the endocytosis pathway was enriched in the comparison of the YFB vs. DFB stages (P value of 0.0008063).

3.3. Gene co-expression network analysis

A WGCNA analysis was performed to gain a comprehensive view of the gene expression patterns during the growth and development process. A total of 1386 hub genes identified from the transcriptome data were divided into 26 modules based on the similarity in their expression patterns (Fig. 4A). As shown in Fig. 4B, the significantly enriched pathways involved ABC transporters (membrane transport), staurosporine biosynthesis (biosynthesis of other secondary metabolites), atrazine degradation (xenobiotics biodegradation and metabolism), diterpenoid biosynthesis (metabolism of terpenoids and polyketides) and MAPK signaling pathway-yeast (signal transduction). After removing genes with FPKM values < 10 in all the five stages, 1135 genes were obtained (Table S7). Among these 12 genes were differentially expressed in the four comparisons of two adjacent stages, namely Hsp30, cytochrome P450, exo-beta-1,3-glucanase and nine hypothetical proteins. A total of 961 genes showed higher expression levels at the mycelial stage, whereas only 29 genes were upregulated at the primordial stage compared to the mycelial stage. MAPK signaling pathway-yeast and amino sugar and nucleotide sugar metabolism pathway were significantly enriched pathways at the P vs. My stages (Fig. S3). The downregulated genes included three DnaJ genes, Hsp30 and 25 TFs (transcription factors), including 16 C2H2, 5 bHLH, 2 bZIP, 1 GATA and 1 HSF class TFs (Fig. 4C). In the comparison between the YFB and P stages, 31 and 30 DEGs,
including one C2H2 transcription factor and Hsp30, were downregulated and upregulated respectively at the YFB stage. In a comparison of the YFB vs. P stages, diterpenoid biosynthesis and fatty acid degradation pathways were obviously enriched, and the MAPK signaling pathway—yeast had the most DEGs (Fig. S3). A total of 89 DEGs (30 downregulated and 59 upregulated) were identified in the comparison of the DFB vs. YFB stages with the endocytosis pathway being significantly enriched (Fig. S3). During the process of fruiting body maturation (MFB vs. DFB), 62 genes were clearly upregulated in the mature fruiting body, and two pathways involved in aflatoxin biosynthesis and staurosporine biosynthesis were also enriched (Fig. S3).

### 3.4. Proteomics overview

As shown in Fig. 5A, proteome sequencing yielded a total of 711,861 spectra, of which 28,413 unique spectra were obtained. After removing the peptides with a FDR (false discovery rate) value > 1%, a total of 8430 peptides and 2040 proteins were identified and used to perform the global protein-expression profile analysis across the differential developmental stages. The theoretical molecular weight of the identified proteins varied from 7.76 kDa to 1284.32 kDa, with the majority of them (75.05%) being from 10 kDa to 80 kDa (Fig. 5B and Table S8). Based on the GO analysis (Fig. S4), for BP most of the T. guangdongense proteins identi-
fied were found to be involved in metabolic process, cellular process, localization and biological regulation. For CC, the majority of proteins were related to the biosynthesis of cell, membrane, and organelle. For MF, numerous proteins were classified into the catalytic and binding activity categories, followed by structural molecule and transporter activity categories. Based on the KOG annotation (Fig. 5C), most proteins were found in the metabolism category including amino acid transport and metabolism, energy production and conversion and lipid transport and metabolism, followed by the cellular processes and signaling categories involved in translational modification, protein turnover, chaperones and signal transduction mechanisms. In addition, 149 proteins were grouped into translation, ribosomal structure and biogenesis. The KEGG pathway analysis showed that most of the proteins were involved in carbohydrate metabolism, amino acid metabolism, translation and protein folding, sorting and degradation pathways (Table S9). It is worth pointing out that a number of proteins that remained unannotated in these databases need further study to explore their function.  

3.5. Identification and functional annotation of differentially expressed proteins

Compared to the My stage, there were 818 and 1016 DEPs identified in comparisons with the P and DFB stages, respectively (Fig. 6A and Table S10). Among these proteins, 222 and 170 proteins were clearly upregulated at the P and DFB stages, respectively, whereas 596 and 846 of these DEPs were significantly downregulated at the P and DFB stages, respectively (Fig. 6B). Comparing the P and DFB stages, 355 proteins were significantly differentially expressed, of which 305 were upregulated and 50 were downregulated. A total of 95 proteins were clearly differentially expressed during the growth and development process of T. guangdongense (Fig. 6A). To be more precise, these 95 proteins were clearly differentially expressed in all three comparisons (P vs. My, DFB vs. My and P vs. DFB), suggesting an important role in the process of primordium formation and fruiting body development. As shown in Fig. 6C, 49 of those 95 proteins were clearly downregulated at the P and DFB stages, whereas 19 proteins were...
significantly upregulated. In addition, 17 of those 95 proteins were upregulated at the P stage but downregulated at the DFB stage, whereas 9 proteins were upregulated at both the P and DFB stages. Only one protein showed an obvious downregulation expression at the P and DFB stages, but was upregulated at the DFB stage compared to the P stage.

An analysis of the GO annotation and KEGG pathways was performed to investigate the function of these DEPs across the differential developmental stages of *T. guangdongense*. The GO annotation results for these DEPs were similar to those of the DEGs in the transcriptome (Table S11). Based on the number of DEPs classified into each functional category, the three largest categories were as follows: cell, cell part and membrane for CC; metabolic process, cellular process and localization for BP; and catalytic activity, binding, and transporter activity for MF. The KEGG pathway analysis revealed that compared to the My stage, the largest number of DEPs in the P and DFB stages were related to the metabolic pathway category, followed by biosynthesis of secondary metabolites in primordia and carbon metabolism (Fig. S5). However, more DEPs were classified into the peroxisome and spliceosome pathways in the comparison of P vs. DFB. Based on the result of the KEGG pathway enrichment analysis, the most significantly enriched pathways (P < 0.05) were as follows: four pathways in the comparison of P vs. My stages, including oxidative phosphorylation, meiosis-yeast, mRNA surveillance pathway and synthesis and degradation of ketone bodies; three pathways in the comparison between the P and DFB stages, namely, biosynthesis of unsaturated fatty acids, autophagy-other and alpha-linolenic acid metabolism; and only one pathway (mRNA surveillance) from the comparison of DFB vs. My. In addition, it was found that the majority of DEPs were located to the cytosol, mitochondria, extracellular, nucleus, and plasma membrane (Fig. S6).

### 3.6. Correlation analysis between differentially expressed genes and proteins

A correlation analysis between the transcriptome and proteome was performed to better understand the molecular mechanisms involved in *T. guangdongense* development. A correlation of the 2022, 2021 and 2004 transcripts and proteins from three comparisons were identified (Table 1). A low degree of correlation between protein level changes and transcript level changes was found for all proteins. Among these, the number of differentially expressed transcripts and proteins correlated were as follows: 367 for P vs. My and 33 for DFB vs. P (Fig. 7A). In the comparison between the P and My stages, the correlated DEGs included 74
upregulated and 198 downregulated genes. Nevertheless, 7 upregulated genes and 1 downregulated gene were found in the genes correlated between the DFB and P stages (Fig. 7B and Table S12). Compared to the differentially expression levels of the transcripts correlated, approximately half of the quantified proteins in the comparison of P vs My showed the opposite changes, whereas proteins with the opposite trend only accounted for 23.20% in the comparison of DFB vs P (Fig. 7B). Among all the quantified proteins, only one protein (CTABGD1-15_A0198), annotated as an integral membrane protein, showed a similar trend for protein and transcript changes in the P vs. My and DFB vs P comparisons (Table S12). This difference between mRNA transcript changes and protein changes indicates that changes in mRNA expression levels provide only a limited viewpoint into protein expression changes, and that the developmental process of T. guangdongense may be controlled post-transcriptionally through changes in the efficiency of protein translation, protein modifications, and protein interactions.

Heat shock proteins and transcription factors play an important role in regulating the development, metabolism and resistance of an organism to abiotic and biotic stressors. From the correlation analysis between the transcriptome and the proteome, many heat shock proteins and transcription factors were significantly differentially expressed across the developmental stages of T. guangdongense (Fig. 7C). Among them, GATA and Hsp30 were clearly downregulated at the P stage at both the transcriptional and translational levels, indicating that they may be involved in primordial germination. Myb (GTABGD1-15_A08102), Hsp60 and Hsp70 (GTABGD1-15_A3890 and GTABGD1-15_A4502) only showed a downregulation of protein levels at the P vs. My stages, but no obvious change found at the mRNA transcriptional level. In contrast, bZIP and DnaJ (GTABGD1-15_A0241) were only downregulated at the mRNA transcriptional level, and Hsp90 showed a clear opposite change at the protein and mRNA transcript levels. In the comparison of DFB vs P, one DnaJ protein (GTABGD1-15_A2220) was significantly upregulated at the protein level. These

Table 1
Correlation analysis of differentially expressed proteins and genes identified from proteomic and transcriptomic analyses of different developmental stages.

| Comparison  | Gene number | Protein number | Correlation number |
|-------------|-------------|----------------|-------------------|
| P vs. My    | 8952        | 2040           | 2022              |
| DFB vs. P   | 8952        | 2040           | 2004              |

Note: My, Mycelial stage; P, Primordial stage; DFB, Developmental fruiting body stage; DE, Differential expression.
results suggest that heat shock proteins and transcription factors have a complex expression pattern and may be involved in growth and development at multiple levels.

3.7. Correlation analysis of tryptophan and IAA biosynthesis in T. Guangdongense

Based on an untargeted-metabolomic analysis of mycelia and fruiting bodies of T. guangdongense (unpublished data), metabolites related to tryptophan and IAA biosynthesis were identified (Fig. 8B). From the correlation analysis between the transcriptome and the proteome, trpF, Tam-1 (CTABGD1-15_A2976), MAO and YUCCA (CTABGD1-15_A2573) were clearly downregulated at the protein level in the comparison of P vs. My, whereas trpB (tryptophan synthase alpha chain) showed the opposite expression changes at the protein level (downregulation) and the mRNA level (upregulation). In addition, only one aldehyde dehydrogenase protein (CTABGD1-15_A3525) was significantly upregulated in the primordium and fruiting body (Fig. 8A). As shown in Fig. 8C, tryptophan and tryptamine showed a similar trend in changes, and reached a peak at the P stage, indicating a positive effect on the process of primordium formation. Nevertheless, IAA showed a continuous decrease across the developmental process of T. guangdongense. These results suggest that IAA can be synthesized via the tryptamine pathway and the indolepyruvate pathway, and that metabolites related to IAA biosynthesis may be involved in the development of T. guangdongense.

3.8. Correlation analysis of bioactive compounds in T. Guangdongense

T. guangdongense is able to produce different types of bioactive and nutritional compounds such as polysaccharides, adenosine and terpenoids. These compounds are known to be antioxidants, and possess anti-fatigue, anti-virus, and antitumor activities.

3.8.1. Triterpenoid and ergosterol biosynthesis

A total of 15 key genes involved in terpenoid and ergosterol biosynthesis were identified in the genome and transcriptome of T. guangdongense, and 12 enzymes were found in the proteome (Fig. 9A and Table S13). With respect to terpenoid biosynthesis,

----

Fig. 8. Analysis of tryptophan and IAA biosynthesis on the basis of multi-omics. A, Heatmap showing the expression change of the enzyme proteins related to tryptophan and IAA biosynthesis in the comparisons (P vs. My and DFB vs. P). Gray, Not detected in proteome analysis; Blue, Downregulation; vs, Versus; Yellow, Upregulation; G, Transcript level; Pr, Protein level. B, Overview of the IAA biosynthesis pathway based on the transcriptomic, proteomic and metabolomic data. B and C, Enol-1-(O-carboxyphenylamino)-1-deoxy ribulose phosphate and Indole-3-glycerol phosphate. C, Contents of tryptophan, tryptamine and IAA in differentially developmental stages of T. guangdongense. My, Mycelial stage; P, Primordial stage; DFB, Developmental fruiting body stage. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
the protein levels of squalene synthase (FDFT1) were significantly
downregulated in the primordia and fruiting bodies compared to
the mycelia, but there were no obvious changes found at the tran-
scriptional level (Fig. 9B). With respect to ergosterol biosynthesis,
the protein levels of three enzymes, including TM7SF2, ERG26,
and CYP61, showed an obvious downregulation in the primordia
and fruiting bodies compared to the mycelia, but TM7SF2 was
downregulated in the primordia (Fig. 9B and Table S13). Compared
to the protein expression levels in the primordia, TM7SF2 was
upregulated in fruiting bodies, whereas LSS and ERG27 were signif-
icantly downregulated. Two genes encoding lanosterol synthase
(LSS) and sterol 24-C-methyltransferase (SMT1) were clearly
upregulated in the primordia and fruiting bodies compared to
the mycelia (Fig. 9B). These results indicate that macro-fungi
maybe produce more ergosterol and terpenoids at the My stage
than in the FB stages.

Fig. 9. Analysis of terpenoid and ergosterol biosynthesis based on multi-omics. A, Ergosterol biosynthesis pathway based on the transcriptomic and genomic data. B, Heatmap showing the expression change of the enzymes related to terpenoid and ergosterol biosynthesis in the comparisons (P vs. My and DFB vs. P). Gray, Not detected in the proteome analysis; Green, Downregulation at the protein level; Red, Upregulation at the transcript level; Black, No significant change at the protein or transcript levels; Yellow, Complex changes across the different developmental stages; My, Mycelial stage; P, Primordial stage; DFB, Developmental fruiting body stage; G, Transcript level; Pr, Protein level; vs, Versus. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 10. Analysis of adenosine biosynthesis based on multi-omics. A, Adenosine biosynthesis pathway according to the transcriptome and genome data. B, Heatmap showing the expression changes of the enzymes related to adenosine biosynthesis in different comparisons (P vs. My and DFB vs. P). Gray, Not detected in the proteome analysis. My, Mycelial stage; P, Primordial stage; DFB, Developmental fruiting body stage; G, Transcript level; Pr, Protein level; vs, Versus; B, P vs My; DFB vs P.
3.8.2. Biosynthesis of polysaccharide, adenosine, and other bioactive metabolites

Two genes encoding 1,3-β-glucan synthase and four genes involved in the biosynthesis of 1,6-β-glucans were found on the basis of the functional annotation of the genome (Table S13). Among them, one protein (1,3-β-glucan synthase) showed higher expression level in primordia, but there was an obvious downregulation for UTP-glucose-1-phosphate uridylyltransferase protein, which is important in the biosynthesis of 1,6-β-glucans. With respect to adenosine synthesis (Fig. 10A), the protein levels of five enzymes (NDK, PK, AK, APRT and NrdJ) were all found to be significantly downregulated in primordia, whereas two enzymes (ADK and APRT) were clearly upregulated at the DFB stage (Fig. 10B and Table S13), strongly suggesting that mycelia and fruiting bodies may have higher levels of adenosine. In addition, with regard to those proteins involved in the biosynthesis of other bioactive ingredients, nonribosomal peptide synthetase (NRPS) was upregulated in primordia and fruiting bodies compared to the mycelia, whereas non-ribosomal peptide synthases (NRPs) was clearly downregulated. Among the eight polyketide synthases (PKSs), half of them exhibited a significant upregulation expression in primordia and fruiting bodies. Further studies need to be performed to investigate the production of polysaccharides, adenosine, and other bioactive metabolites at the differential developmental stages of T. guangdongense.

3.9. Verification of the transcriptome reliability using qRT-PCR

The relative expression levels of 16 genes were used to verify the reliability of the transcriptome data. The expression trends of 14 genes (87.5%) across the differential developmental stages of T. guangdongense were consistent with the transcriptome data, with only the absolute fold changes in gene expression showing slight differences (Table S14). Compared to the My stage, DnaJ17 (CTABGD1-15_A4807) at the YFB stage was not differentially expressed by qRT-PCR, but was downregulated in the transcriptome analysis. Compared to the mycelia, Hsp98 (CTABGD1-15_A4807) was differentially expressed in YFB and DFB by qRT-PCR, but was not different in the transcriptome analysis.

4. Discussion

Because of its medicinal and nutritional values, O. sinensis has historically been regarded as the most valued traditional Chinese medicinal fungus. Recently, numerous transcriptomic, proteomic and metabolomic data have revealed fruiting body and sexual development-related genes [24], the protein uniformity between wild and artificial fruiting bodies [44] and metabolite differences (Table S13). Based on a correlation analysis between the transcriptomic and proteomic data (Fig. 4C and Table S6), A total of 18 heat shock proteins were found (Table S12). Among them, one C2H2 protein was clearly upregulated at the P stage compared to the My and DFB stages, and one GATA proteins were downregulated at the P stage compared to the My stage (Fig. 7C), which is consistent with previous studies that C2H2 showed a higher expression levels in primordia and fruiting [23,58]. Therefore, we speculate that the P stage is a critical stage for formation of the fruiting body, in which a large number of genes are significantly differentially expressed.

4.1. Changes in transcription factors during the developmental process

Transcription factors (TFs) play vital roles in the development and the response of organisms to abiotic and biotic stresses. C2H2 TFs, one type of the common zinc finger proteins, play an important role in regulating ascospore formation, sexual reproduction, and conidial and mycelial growth in Neurospora crassa [30–51]. Knockout of C2H2 TF Tha00974 decreased biomass and spore production, and reduced its pathogenicity to tomato [52]. In contrast, the fruiting body yield per day of C2H2 overexpressing strains peaked one day earlier than the wild type Agaricus bisporus strain [53]. GATA TFs in fungi have been shown to have various functions and are involved in light response, siderophore biosynthesis, mating-type switching, asexual-sexual development, and appressoria formation [54–57]. In the present study, a WGCNA analysis showed that 25 TFs including 16 C2H2, 5 bHLH, 2 bZIP, 1 GATA, and 1 HSF were all significantly differentially expressed during the developmental process of T. guangdongense (Fig. 4C and Table S6). Based on a correlation analysis between the transcriptomic and proteomic data, 8 TFs including 1 bHLH, 1 GATA, 2 MYB, and 4 C2H2 proteins were found (Table S12). Among them, one C2H2 protein was clearly upregulated at the P stage compared to the My and DBF stages, and one GATA showed a significant downregulation in expression at both the mRNA transcript and protein levels at the P stage compared to the My stage (Fig. 7C), which is consistent with previous studies that C2H2 showed a higher expression levels in primordia and fruiting [23,58]. Therefore, we speculate that the P stage is a critical stage for formation of the fruiting body, in which a large number of genes are significantly differentially expressed.

4.2. Changes in heat shock proteins during the developmental process

The low correlation between the transcriptomic and proteomic suggested that multiple protein modifications, including protein folding and degradation, occur in the developmental process of T. guangdongense. Heat shock proteins, a stress-responsive family of proteins, are involved in posttranslational modifications, protein folding, aggregation and disaggregation [59]. Based on the WGCNA analysis of the transcriptome data, we found that three DnaJ genes and one Hsp30 gene were significantly downregulated at the P stage compared to the My stage, and that Hsp30 was differentially expressed in another three comparisons (YFB vs. P, DBF vs. YFB and MFB vs. DFB) (Fig. 4C and Table S7). A total of 18 heat shock proteins were found in the correlation analysis between transcriptome and proteome, and Hsp30 showed an obvious downregulation expression at both the transcript and translation levels at the P stage (Fig. 7C and Table S11). In addition, three Hsp70 proteins and one Hsp90 protein were downregulated at the P stage, and one DnaJ protein showed an obvious downregulation at the DBF stage. It has been reported that the mRNA expression level of heat shock protein 9 in Grifola frondosa increased during fruiting body differentiation [60]. In addition, a variety of heat shock proteins
such as Hsp25 and DnaJ were involved in mycelial growth, conidial formation and germination, fungal morphogenesis, and the response to environmental factors [61–64]. The decreased expression of heat shock proteins may be unfavorable for protein and membrane stability during the developmental process of T. guangdongense, especially in the presence of detrimental environmental factors.

4.3. Changes in the MAPK pathway during the developmental process

Mitogen-activated protein kinase (MAPK) pathways, being composed of three serine/threonine protein kinases (MAPKKK, MAPKK and MAPK), can transduce extracellular signals to regulate transcriptional events and other cellular processes [65]. They have been reported to be involved in the formation of the ascocarp in ascomycetes, pathogenicity in insects, sexual reproduction, osmolarity resistance, and cell wall integrity in fungi [66–69]. A total of 12 genes encoding serine/threonine protein kinases were identified using a correlation between the transcriptomic and proteomic data (Table S12). Among them, four MAPK proteins showed obvious differential expression at the P stage compared to the My and DFB stages, indicating there is a relationship between primordium formation and environmental factors such as light and temperature. In addition, the differential expression of MAPKs can result in altered regulation of TFs and other related proteins, and thus may be involved in the developmental process of T. guangdongense.

4.4. Changes in proteins related to IAA and bioactive metabolites during development process

As the first discovered plant hormone, IAA was also found in basidiomycete and ascomycete fungi [70–71]. In plants, a low concentration of IAA effectively promotes the growth of roots and leaves and resistance to salt, low temperatures, and heavy metal stresses [72–74]. The current study showed that many genes related to IAA biosynthesis were obviously downregulated at the P and DFB stages, and that the IAA content showed a continuous decrease with the development of fruiting bodies (Fig. 8). In the basidiomycete fungus L. edodes, exogenous auxin and its analogues can alleviate the effects of oxidative damage induced by heat stress [75]. Therefore, we speculate that IAA may regulate the development and the response of T. guangdongense toward abiotic stresses, but this needs to be further confirmed.

The main bioactive ingredients identified in T. guangdongense are adenosine, polysaccharides, terpenoids, etc. [10–11]. In the adenosine biosynthesis pathway of O. sinensis, most of the key genes involved in this pathway were significantly downregulated at the late stage of fermentation [76]. Similarly, our results suggest that many proteins involved in adenosine biosynthesis are significantly downregulated at the P and DFB stages compared to the My stage (Fig. 10B and Table S13). Of all the medicinal fungus polysaccharides, the water-soluble 1,3-β-glucans and 1,6-β-glucans show the best immunomodulatory and antioxidant effects [77–78]. Three genes involved in 1,3-β-glucan and 1,6-β-glucan biosynthesis showed clearly differential expression at the P stage compared to the My or DFB stages (Table S13). Seven of 15 key proteins involved in the biosynthesis of terpenoids and ergosterol were also differentially expressed across the developmental process of T. guangdongense. These results indicate that T. guangdongense may be able to produce more ergosterol, terpenoids, and adenosine at the My stage than at the P and DFB stages.

Cordycepin is an important metabolite in C. militaris, and possesses the potential as a medicine, food and herbicide [79–81]. In T. guangdongense, extremely low levels of cordycepin were detected, and the genes involved in cordycepin biosynthesis were not identified in the genome of T. guangdongense. Nevertheless, it contained high levels of cordycepic acid, adenosine, ergosterol, polysaccharides, and vitamins, suggesting the potential of T. guangdongense as a food and a medicine, although this need further research. In future studies, we will comprehensively analyze cordycepin using multiple approaches, and explore the biosynthesis mechanisms of metabolites using reverse genetics methods. In addition, we will conduct in vivo studies using extracts of T. guangdongense to investigate the functions of T. guangdongense bioactive metabolites.

5. Conclusions

In this study, a combination of transcriptomic and proteomic analyses were used to provide a view of hub genes, key pathways, development-related and metabolite-related gene/protein profiles during the differential developmental stages of T. guangdongense. It was found that the transcriptomic profile at the My stage was significantly different from the P and FB stages. Many functional genes and proteins, especially associated with transcriptional regulation, protein modification, signal transduction and the biosynthesis of bioactive metabolites, were identified. Most of the differentially expressed TFs belonged to the C2H2 family, and numerous heat shock proteins such as Hsp30 and DnaJ showed significant differential expression during the developmental process. Moreover, IAA levels and key enzymes involved in its biosynthesis were downregulated, suggesting a potential effect on the growth and development of T. guangdongense. Furthermore, numerous proteins related to the biosynthesis of adenosine, terpenoid, ergosterol, and polysaccharide were differentially downregulated at the P and FB stages. In summary, the present study sheds lights on the developmental process of T. guangdongense and provides information on the biosynthesis of several metabolites. It is hoped that these data will lead to a further dissection of the functions of the DEGs identified here, and this will benefit the development of both industrial and medical applications of this medicinal fungus.

CRediT authorship contribution statement

Gangzheng Wang: Conceptualization, Methodology, Data curation, Validation, Writing - original draft. Min Li: Resources, Validation, Data curation. Chenghua Zhang: Conceptualization, Methodology. Huijiao Cheng: Resources, Investigation. Yu Gao: Resources, Investigation. Wangqiu Deng: Conceptualization, Writing - review & editing. Taihui Li: Conceptualization, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Editage (www.editage.cn) for English language editing. This research was funded by the National Natural Science Foundation of China (grant No. 31900015; 31800012), the GDAS’ Special Project of Science and Technology Development (2020GDASYL-2020010304), the Science and Technology Planning Project of Guangzhou (No. 201804020018), and the Science and Technology Planning Project of Guangdong Province, China (Nos.2018A0303130164, 2019B121202005).
Ohm RA, de Jong JF, de Bekker C, Lugones Wösten HA, LG. Transcription factor genes of *Schizophyllum commune* involved in regulation of mushroom formation. Mol Microbiol 2011;81(6):1433–45.

Hartl FU, Bracher A, Hayer-Hartl M. Molecular chaperones in protein folding and proteostasis. Nature 2011;475:324–32.

Kurahashi A, Sato M, Nishibori K, Fujimori F. Heat shock protein 9 mRNA expression increases during fruiting body differentiation in *Grifola frondosa* and other edible mushrooms. Mycoscience 2014;55(2):98–102.

Liao XG, Lu HL, Fang WG, St Leger RJ. Overexpression of a *Metarhizium robertsii* HSP25 gene increases thermotolerance and survival in soil. Appl Microbiol Biotechnol 2013;98(2):777–83.

Wang J, Ying SH, Hu Y, Feng MG. Mas5, a homologue of bacterial DnaJ, is indispensable for the host infection and environmental adaptation of a filamentous fungal insect pathogen. Environ Microbiol 2016;18(3):1037–47.

Xie JL, Bohovych I, Wong EG, Lambert JP, Gingras AC et al. Ydj1 governs fungal morphogenesis and stress response, and facilitates mitochondrial protein import via Mas1 and Mas2. Microbial Cell 4(10): 342-61.

Wang GZ, Zhou SS, Luo Y, Ma CJ, Gong YH et al. The heat shock protein 40 LeDnaJ regulates stress resistance and indole-3-acetic acid biosynthesis in *Lentinula edodes*. Fungal Genet Biol 2018;118:37–44.

Paul A, Wilson S, Belham CM, Robinson CJM, Scott PH, et al. Stress-activated protein kinases: activation, regulation and function. Cellular Signaling 1997;9(6):403.

Bayram Ö, Bayram ÖS, Ahmed YL, Maruyama J, Valerius O et al. The *Aspergillus nidulans* MAPK module AnSte11-Ste50-Ste7-Fus3 controls development and secondary metabolism. PLoS Genetics 8(7): e1002816.

Wang M, Dong Y, Zhao Q, Wang F, Liu K, et al. Identification of the role of a map kinase tmk2 in *Hypocre a jecorina* (Trichoderma reesei). Sci Rep-UK 2015;4(1):6732.

Martine Soto D, Ruiz-Herrera J. Functional analysis of the MAPK pathways in fungi. Micolgia 2017;34(4):192–202.

Jin K, Han L, Xia Y, Mamk1, a fus3/kss1-type mitogen-activated protein kinase gene, is required for appressorium formation, and insect cuticle penetration of the entomopathogenic fungus *Metarhizium acridum*. J Invertebr Pathol 2014;115:68–75.

Boland W. Biosynthesis and secretion of indole-3-acetic acid and its morphological effects on *Tricholoma vaccinum*-spore ectomycorrhiza. Appl Environ Microbiol 2015;81(20):7003–11.

Jaturong K, Nakarin S, Kenji M, Saisamorn L. Biosynthetic pathway of indole-3-acetic acid in ectomycorrhizal fungi collected from northern Thailand. PLoS ONE 2020;15(1):e0227478.

Shibasaki K, Uemura M, Tsurumi S, Rahman A. Auxin response in Arabidopsis under cold stress: underlying molecular mechanisms. Plant Cell 2009;21(12):3823–38.

Macej O, Anna C, Anna J (2015) The auxin conjugate indole-3-acetyl-aspartate affects responses to cadmium and salt stress in *Psium sativum* L. J Plant Physiol 2015;191:63–72.

Ran Jakang, Zheng Wen, Wang Hongbin, Wang Hajuan, Li Qinchun. Indole-3-acetic acid promotes cadmium (Cd) accumulation in a Cd hyperaccumulator and a non-hyperaccumulator by different physiological responses. Ecotoxicol Environ Saf 2020;191:110213. https://doi.org/10.1016/j.ecoenv.2020.110213.

Zhou SS, Wang GZ, Luo Y, Ma CJ, Gong YH et al. Auxin and auxin analogues enhanced the thermotolerance of *Lentinula edodes*. Mycosystema 37(12); 1723-30.

Lin S, Zou ZC, Zhou CB, Zhang HC, Cai ZM. Transcriptome analysis reveals the molecular mechanisms underlying adenosine biosynthesis in anamorph strain of caterpillar fungus. Biomed Res Int 2019;2016168.

Chen S, Xu J, Liu C, Zhu Y, Nelson DR, et al. Genome sequence of the model medicinal mushroom *Ganoderma lucidum*. Nat Commun 2012;3:913.

Mizuno M, Nishitani Y. Immunomodulating compounds in Basidiomycetes. J Clin Biochem Nutr 2013;52:202–7.

Chen DD, Xu R, Zhou JY, Chen JQ, Wang L, et al. Cordyceps isolated from *Cordyceps militaris*: Its Newly Discovered Herbicidal Property and Potential Plant-Based Novel Alternative to Glyphosate. Molecules 2019;24:2901.

Chen DD, Xu R, Zhou JY, Chen JQ, Wang L, et al. Cordyceps militaris polysaccharides exerted protective effects on diabetic nephropathy in mice via regulation of autophagy. Food Funct 2019;10:5102–14.

Jo EB, Jang HJ, Yang KE, Jang MS, Huh YH, et al. Cordyceps militaris induces apoptosis in ovarian cancer cells through TNF-α/TNFRI-mediated inhibition of NF-κB phosphorylation. BMC Compl Med Therapies 2020;20(1).