Microbiological Research 218 (2019) 12–21

Contents lists available at ScienceDirect

Microbiological Research

journal homepage: www.elsevier.com/locate/micres

Cordycepin and pentostatin biosynthesis gene identified through transcriptome and proteomics analysis of *Cordyceps kyushuensis* Kob

Xuan Zhao[a], Guoying Zhang[b], Caiyi Li[a], Jianya Ling[a,*]

[a] *State Key Laboratory of Microbial Technology, Shandong University, Jinan, 250100, China*
[b] *Shandong University of Traditional Chinese Medicine, Jinan, 250014, China*

ARTICLE INFO

Keywords:
Transcriptome
Proteomics
Cordycepin
Pentostatin
HPLC analysis

ABSTRACT

Cordyceps kyushuensis is the only species of cordyceps growing on the larvae of *Clanis bilineata* Walker, and has been demonstrated that there are lots of pharmacological components including cordycepin. Cordycepin shows lots of pharmacological action but it could be converted to 3′-deoxyinosine by adenosine deaminase *in vivo*, which weakens the efficiency of cordycepin. That pentostatin, which has been reported to inhibit adenosine deaminase, combining cordycepin could enhance the efficiency of cordycepin *in vivo*. During transcriptome and proteomics analysis of *Cordyceps kyushuensis*, a single gene cluster including four genes we named ck1-ck4 which can synthesis both cordycepin and pentostatin has been identified using BLAST. Meanwhile, KEGG, KOG, GO analysis and differentially expressed genes were analyzed in transcriptome and proteomics. This study first sequenced transcriptome and proteomics of *C. kyushuensis*, and demonstrated that there is a single gene cluster related to biosynthesis of cordycepin and pentostatin, which can be employed to improve the yield of cordycepin and find more functional proteins.

1. Introduction

Cordyceps genus is a kind of important fungi consisting of fungi and caterpillar. *Cordyceps sinensis* and *Cordyceps militaris*, which both belong to Cordyceps genus, have been used as Chinese traditional medicine for several thousand years (Zhang et al., 2015; Zhou et al., 2008). Pharmacological studies have proved that there are lots of biological activity ingredients in Cordyceps genus, for example, cordycepin, cordycepic acids, polysaccharide, cyclic peptides, etc. (Yin et al., 2012), which exhibit various of clinical health effects, inlcuding anticancer, antioxidant, anti-inflammatory, immunomodulatory (Ahn et al., 2000), anti-microbial activities(Ahn et al., 2000), etc.. *Cordyceps kyushuensis*, the only species of cordyceps growing on the larvae of *bilineata* Walker was classified by Japanese scientist Y. kobayasi in 1983. Subsequently, *C. kyushuensis* was discovered in Mount Meng area, China and was verified as a new species, which were also found that there are lots of pharmacological activity ingredients (Li et al., 2002; Fan et al., 2006; Wang et al., 2017). *C. kyushuensis* in this manuscript is referred as *C. kyusyuensis*. In 2008, Dai et.al. revised the checklist of medicinal fungi in China, and *Cordyceps kyushuensis* was renamed to *Cordyceps kyusyuensis*. But in recent years, the two nomenclatures of *Cordyceps kyushuensis* and *Cordyceps kyusyuensis* are both used, so in this manuscript, the authors still use the Latin name "*Cordyceps kyushuensis*".

Cordycepin is 3′-deoxyadenosine, which was originally isolated from *Cordyceps militaris* in 1950 (Cunningham et al., 1950), compared with adenosine it lacks a 3′ hydroxyl group. Researches showed that cordycepin interacts in biochemical processes, for example, nucleic acid synthesis, platelet aggregation, apoptosis, etc. (Tuli et al., 2013). In vitro studies showed that cordycepin exhibited anticancer activity through incorporating into RNA synthesis to break off it (Holbein et al., 2009; Tuli et al., 2013; Wen et al., 2016), *in vivo* studies showed that cordycepin could treat mice infected with *Trypanosoma evansi* (Dalla Rosa et al., 2013). Because of its pharmacological value, cordycepin becomes more and more popular and expensive. In 2016, the market price of cordycepin was approximately \$12,000 kg^{-1}, which has increased almost 20-fold in the last 10 years (Ni et al., 2009). But until now, the clear and exact cordycepin biosynthesis pathway is not elucidated. At the same time, cordycepin could be deaminated to 3′-deoxyinosine by adenosine deaminase *in vivo* (Rosa et al., 2013; Vodnala et al., 2013). The studies showed that the half-life of cordycepin elimination is significantly short ($t_{1/2}$ = 1.6 min) *in vivo* in rats (Tsai et al., 2010), which weakens its activity and limits its use.

Pentostatin (2′-deoxycoformycin) is an irreversible inhibitor of adenosine deaminase to prevent the deaminaton of cordycepin to 3′-deoxyinosine (Brogden and Sorkin, 1993; Dalla Rosa et al., 2015), which was first isolated from a fermentation culture of the bacterium

---

* Corresponding author at: State Key Laboratory of Microbial Technology, School of Life Sciences, Shandong University, Jinan, China.
E-mail address: lingjian-ya@sdu.edu.cn (J. Ling).

https://doi.org/10.1016/j.micres.2018.09.005
Received 15 May 2018; Received in revised form 17 August 2018; Accepted 14 September 2018

Streptomyces antibioticus in the 1970s (Woo et al., 1974). It has been used to treat leukemia as a chemotherapeutic drug (Sarvaria et al., 2016) and there was a report which showed that a combination of cordycepin and pentostatin could increase the antitumor effect over 3-fold (Yoshikawa et al., 2007). Recent researches showed that cordycepin and pentostatin were biosynthesized by a single gene cluster in C. militaris (Xia et al., 2017), which meant that it is promising to use the combination of cordycepin and pentostatin to treat illnesses *in vivo* and produce cordycepin and pentostatin in a higher level.

Recent study showed that there are no cordycepin-related genes in *C. sinensis* (Xia et al., 2017), which implies that there was no cordycepin in it. While *C. militaris* has been demonstrated to have a single gene cluster which can synthesis both cordycepin and pentostatin. *Cordyceps kyushensis*, which is close relative to *C. militaris*, was reported to produce both cordycepin and pentostatin (Yoshikawa et al., 2007). As researches moves along, *C. kyushuensis* shows more and more pharmacological activities, which means intensive study is necessary to discover its function mechanism.

High-throughput RNA-Seq is a more efficient, quick and sensitive method to deeply study and analyse transcriptome of organism (Metzker, 2010), which can provide a whole view to widely study the gene expression at RNA levels, no matter lowly expressed or differentially expressed genes. The information provided by the analysis of transcriptome reveals gene expression profiles and speculates unknown gene functions (Liu et al., 2015). In different development stages, the organism has different RNA levels. It is necessary to research samples from different development stages to dig genes related to active metabolites we are interested to. Although gene decides protein level, but genes don’t join vital activities directly, proteins are the main undertakers. Proteins participate in regulation of genes, most enzymes are proteins and some important hormones including insulin, thymin are proteins too. It is inadequate to learn about organism’s function only from genes’ point of view. Researchers showed that some genes which have the same abundance of mRNAs, but the abundance of proteins differ ten times, so we can’t illuminate organism’s function by the numbers. As an important part of post-genomics, proteomics is a forceful tool to study proteins expressed in specific biological state (Liu et al., 2017), which could characterize target protein’s functions, protein-protein interactions and protein modification to provide all-around view and comprehension (Lao et al., 2014).

To the best of our knowledge, the transcriptome and proteomics of *Cordyceps kyushuensis* have not been reported. Through gene level, transcriptional level and protein level, some genes regarded as cordycepin-related genes before have been corrected now. For example, 5′-nucleotidase (CCM_00622, cmNT5E), adenine phosphoribosyltransferase (CCM_00088, cmAPRT) and nucleoside triphosphate pyrophosphatase (CCM_03944, cmNTP) were verified that they are not responsible for cordycepin production in C. militaris. It has to say that combination transcriptome and proteomics analyses is a good way to lucubrate metabolite and biosynthesis of active compounds. In this study, transcriptome and proteome analysis of four different developmental stages of *C. kyushuensis* Kob (mycelium, colouring, initial stage of fruiting body and mature period of fruiting body cultured on rice medium) were conducted to excavate differentially expressed genes between the four samples. Meanwhile, with the deeply analysis of transcriptome and proteomics, some noteworthy and valuable genes and proteins were detected, especially cordycepin-related genes have been found in *C. kyushuensis* Kob. All these could help to lay the foundation of further research to study traditional Chinese medicine.

2. Materials and methods

2.1. Strain, reagent and cultural conditions

C. kyushuensis strain was conserved in potato dextrose agar (PDA) media in our lab. and potato dextrose liquid medium was used as fermentation medium. The mycelium from the active potato dextrose liquid medium was inoculated to rice medium in glass jars and cultured at 22°C for 7 days in dark and then for 10 days at 22°C with a 10:14 h light/dark cycle to conversion of the fungi and forming stromata, maintained at 22°C during the day and 10°C at night for 28 days to form the mature fruiting body. During the alternation of generations in *C. kyushuensis*, four samples were selected for the further research, including CKK1 (mycelium stage at seventh day), CKK2 (coloring stage at seventeenth day), CKK3 (stromata-forming intial stage at thirtieth day), CKK4 (fruiting body stage at forty-fifth day).

Cordycepin was purchased from Meilun Biotechnology Co., LTD. (Dalian, China). Pentostatin was purchased from An Apoptosis and Epigenetics Company (USA).

2.2. RNA isolation and transcriptome sequencing

Total RNAs of *C. kyushuensis* Kob from different growth stages (CKK1, CKK2, CKK3 and CKK4) were extracted using mirVanaTM miRNA ISOlation Kit (Ambion-1561, USA) according to the instructions of the manufacturer. The cDNA libraries were constructed and after quality testing by Agilent 2100 Bioanalyzer, Illumina HiseqTM2500 was used to sequence. We have 3 technical repetitions for each sample, and 3 biological repetitions for each stage.

2.3. Functional annotation and DEGs analysis

Unigenes were obtained from overlapping reads with the help of Trinity (Cattoi, 2013) and TGICL (Pertea et al., 2003). Functional annotation of these unique sequences was performed against different nucleotide and protein databases, including NR, KOG, SWISSPROT, GO and KEGG. FoldChange and FDR were two criterions to analyze if the same gene has differential expression between two samples. Differential expression genes were selected by $p < 0.05$ and FoldChange > 2. Afterwards, GO and KEGG analyses were carried out to describe relevant functions. The RNA-Seq results were then confirmed by semi-quantitative RT-PCR.

2.4. The extraction of cordycepin and pentostatin, HPLC analysis

Both medium and fungi from four different growth stages are collected, and then powdered to sift by 60 mesh. 0.5 g powders of each growth stages (CKK1-CKK4) were extracted using 10-fold volume purified water to obtain cordycepin and pentostatin. HPLC analysis was performed using Agilent 1260 Infinity II LC-system equipped with a TC-C18 column (particle size: 5 μm; length: 250 × 4.6 mm; Agilent, USA). The standards for cordycepin and pentostatin were performed. Based on the preliminary experiments, the elution conditions for cordycepin and pentostatin were performed with the solvents of deionized water (A) and methanol (B, 10–30%) at a flow rate of 0.8 ml/min for 60 min, which was monitored at a wavelength of 259 nm for cordycepin and 280 nm for pentostatin. The column oven was maintained at a constant temperature of 40°C.

2.5. Quantitative RT-PCR verification

2 μg total RNA extracted from each sample was used for cDNA synthesis by Abm. The specific primers of selected genes for quantitative RT-PCR were listed in Table 1. cDNA amplication was performed in a 20 μl reaction system containing 10 μl qPCR mix (Transgen Biotech, China). The cycling conditions were 94°C for 30 s followed by 40 cycles of 94°C for 5 s, 55°C for 15 s and 72°C for 10 s. Actin gene of *C. kyushuensis* Kob was used as the reference gene. Three replicates were performed independently for each gene in quantitative RT-PCR reactions. Differentially expressed genes were analyzed by the $2^{-\triangle\triangle CT}$ method.

Table 1
Primers used for quantitative RT-PCR.

Gene	Primers (5’→3’)	
actin	for	GCCCTTGACTTTGAGCAG
	rev	GGACATCGACATCGCACT
ck1	for	CGGAGTTGAAGTCGTTGGA
	rev	AGAGGAGCCTTGTGATTGATT
ck2	for	AGAGGTTCAGCGACTTCAGG
	rev	GGCTAATGCCGTGTTTGG
ck3	for	CAGCACCGAAATCACCCAG
	rev	AAGGCGTTCTTGCCCACA
ck4	for	TACCGAGCGACAGGTTAGA
	rev	CGGCGAGTTTGGATAGTTG

2.6. Protein Extraction, alkylation, digestion and iTRAQ labeling

Two samples from different stages (CKK1 and CKK4, 0.5 g per sample) were ground thoroughly tos fine powders in liquid nitrogen and protein were extracted with Tris saturated phenol method. Add appropriate BPP (butyl benzyl phthalate) and vortex for 10 min at 4℃, then add Tris saturated phenol in the same volume and vortex for 10 min at 4℃. Centrifuge at 4 ℃, add BPP in the same volume into phenolic phase and vortex for 10 min at 4℃. Continue to centrifuge at 4℃, add precool 0.1 M ammonium acetate methanol into phenolic phase, and then precipitate protein at -20℃ overnight. Centrifuge at 4 ℃, mix sediment with precool 90% acetone to get supernatant twice, and then dissolve protein with lysate (8 M urea with 1% SDS), finally centrifuge to get supernatant. The supernatant was used for protein quantification by Pierce BCA method.

One hundred microgram of proteins for each sample were incorporated into 100 μl lysate (8 M urea with 1% SDS). Add 10 mM TCEP and react for 60 min at 37℃, and then react for 40 min in dark at room temperature with 40 mM Iodoacetamide. Precool acetone was added into per tubes (acetone: sample = 6: 1), subside for 4 h at -20℃. Centrifuge for 20 min with 10,000 g to get sediment, and then dissolve samples with 240 μl 100 mM TEAB. Finally add trypsin (trypsin: protein = 1:25) into proteins and digest at 37℃ overnight. iTRAQ labeling performed according to the manufacturer’s instructions.

2.7. High pH RPLC first dimension separation, LC–MS/MS and HPLC analysis

Labeled peptides were dried in a vacuum concentrator, and then dissolved in UPLC buffer solution, loaded onto a reversed C18 column (1.7 μm, 3 mm × 150 mm, Waters, USA) at a flow rate of 400 μl/min with gradient elution buffer (2% ACN with 0.1% formic acid, A and 80% ACN with 0.1% formic acid, B). Finally, according to peak and time, peptide fractions were pooled together to obtain 10 fin. l fractions. All were concentrated by a rotation vacuum concentrator, dissolved in MS buffer and waited for LC–MS/MS analysis.

Each fraction was injected for LC–MS/MS analysis using a Q-Exactive MS (Thermo, USA) equipped with Easy-nLC 1200. The peptide mixture was loaded onto the C18 column (75 μm × 25 cm, Thermo, USA) with gradient buffer (2% ACN with 0.1% formic acid, A and 80% ACN with 0.1% formic acid, B) at a flow rate of 300 nL/min. MS data were acquired using DDA, top 20 method, and survey scan was 350–1300 Th. The resolution of MS was 70,000 for HCD fragmentation and the resolution of secondary MS was 17,500 whose dynamic exclusion duration was 18 s.

2.8. Proteomic data analysis and bioinformatics

Proteins identification was performed by searching against the protein database from the most relevant sequenced species, C. militaris (http://www.uniprot.org/uniprot/?query=reviewed:no%20taxonomy:983644) using ProteomeDiscovererTM Software 2.1. After obtaining the data, Kyoto Encyclopedia of Genes and Genome (KEGG) pathway and gene ontology (GO) analyses of full spectrum and differentially expressed proteins (DEPs) were conducted according to the information from the KEGG pathway and GO databases, respectively. And all the information got above was used for the further study.

Table 2
Output statistics of sequencing results of transcriptome.

Feature	CKK1	CKK2	CKK3	CKK4
Reads				
Raw Reads	46712874	46889362	47157344	46838934
Clean Reads	45286970	45451456	45701448	45540482
Q30 (%)	92.76%	92.82%	92.79%	92.98%
GC content (%)	56.00%	56.00%	56.00%	56.00%
Unigenes				
Number of unigenes				
≥300 bp	21351	17858	17869	17088
≥500 bp	14683	13259	13360	12773
≥1000 bp	9824	9223	9287	8876
Average length of	1463.87	1456.52	1424.58	1429.75
N50	2478	2226	2124	2146

3. Results

3.1. Summary of RNA-Seq data sets

RNA-Seq was performed for RNA samples extracted from four different developmental stages of *C. kyushuensis* Kob to obtain an overview to do further analysis. After filtering out low-quality reads with quality threshold 30 and reads shorter than 35 bp, 45286970, 45451456, 45701448, 45540482 clean reads were obtained for the samples of CKK1, CKK2, CKK3 and CKK4, respectively. Trinity software was used to assemble these high-quality reads, and 21351, 17858, 17,869 and 17,088 unigenes (≥300bp) with mean sizes of 1463, 1456, 1424 and 1429bp, respectively were obtained as the reference sequences to do the further analysis (Table 2).

Annotation of unigenes in different developmental stages were given based on BLAST algorithm against protein databases including non-redundant (nr), SwissProt, KOG (Clusters of orthologous groups for eukaryotic complete genomes), GO (Gene Ontology) and KEGG (Kyto Encyclopedia of Genes and Genomes) using a cut-off E-value of 1.0e-5.

In the nr database, 16,279 (CKK1), 14,124 (CKK2), 14,510 (CKK3) and 13,858 (CKK4) unigenes were matched, and the E-value distribution of the top hits were shown in Fig. 1. For species distribution, 13481 (82.81%), 11,744 (83.15%), 11,639 (80.21%) and 11,685 (84.32%) unigenes were matched to C. militaris CM01 which was the most matched species (Fig. 2).

3.2. GO, KEGG and KOG classification

GO (Gene Ontology) was established to provide an international standard to classify the unigene functions of *C. kyushuensis* Kob. There are three ontologies which are used to describe molecular function, cellular component and biological process. On basis of the blast NR annotation, a total of 78,563 (CKK1), 69,402 (CKK2), 73,632 (CKK3) and 67,386 (CKK4) unigenes were assigned to 64 GO terms. To four samples (CKK1, CKK2, CKK3, CKK4), for biological process, unigenes involved in “cellular process”, “metabolic process” and “single-organism process” accounted for the majority, while unigenes involved in “rhythmic process” and “cell killing” accounted for the minority. For cellular component, unigenes involved in “cell part”, “cell”and “organelle” accounted for the majority, while unigenes involved in “virion” and “virion part” accounted for the minority. For molecular function, unigenes involved in “catalytic activity” and “binding” accounted for

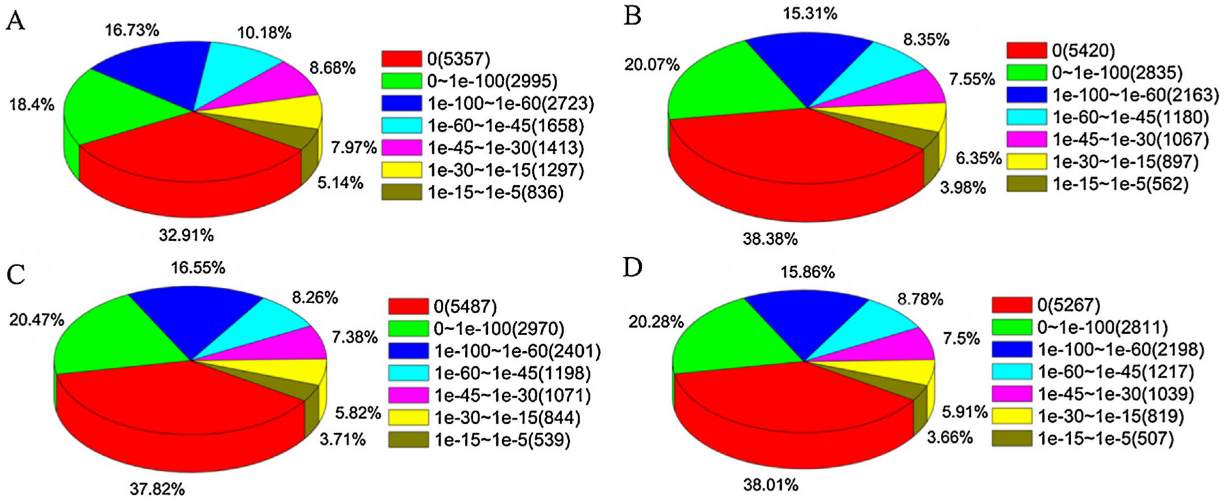

Fig. 1. E-value distribution of transcriptome A) CKK1, B) CKK2, C) CKK3, D) CKK4.

the majority, while unigenes involved in “channel regulator activity” accounted for the minority. It’s noteworthy that no unigenes involved in “mitochondrion-associated adherens complex” in cellular component ontology, and no unigenes involved in “morphogen activity” and “D-alanyl carrier activity” in molecular function ontology (Fig. 3).

16,279, 14,124, 14,510 and 13,858 nr hits were searched against the KOG database to further analysis the transcriptome of *C. kyushuensis* Kob. KOG of genes provided detailed information on the functional groups of predicted protein. 8141, 7266, 7608 and 7032 unigenes from the four samples (CKK1, CKK2, CKK3 and CKK4) have KOG classifications. Among 25 KOG categories, the cluster of “general function prediction only” represents the largest group followed by “posttranslational modification, protein turnover, chaperones” and “translation, ribosomal structure and biogenesis” in four samples. Meanwhile, “extracellular structures” and “cell motility” represented the small groups in four samples (Fig. 4).

To learn more about the biological pathways that are active in *C. kyushuensis* Kob, 11181, 6818, 7552 and 6616 unigenes from the four samples (CKK1-CKK4) were mapped to 350, 207, 207 and 207 KEGG pathways, respectively. “Ribosome”, “biosynthesis of amino acids” and “carbon metabolism” were the most representative pathways which included the most unigenes.

All three of these annotations provided a valuable information to study specific processes, functions and pathways in *C. kyushuensis* Kob.

3.3. Functional annotation and analysis of differentially expressed genes (DEGs)

RNA-Seq data was used to identify genes showing significant changes in expression during the four stages (CKK1-CKK4). CKK1-VS-CKK2, CKK1-VS-CKK3, CKK1-VS-CKK4, CKK2-VS-CKK3, CKK2-VS-CKK4 and CKK3-VS-CKK4 were the six groups to analysis the differential expression genes between them. There were 878 DEGs and 552 up-regulated genes (CKK1-VS-CKK2), 811 DEGs and 536 up-regulated genes (CKK1-VS-CKK3), 718 DEGs and 479 up-regulated genes (CKK1-VS-CKK4), 1057 DEGs and 566 up-regulated genes (CKK2-VS-CKK3), 912 DEGs and 493 up-regulated genes (CKK2-VS-CKK4), 907 DEGs and 453 up-regulated genes (CKK3-VS-CKK4).

GO and KEGG databases were carried out to further analysis the DEGs obtained above among the six groups. All DEGs were mapped to the GO terms in three main categories of GO database. Top 30 GO terms were selected in six groups by the standard of terms which included more than two genes and $-\log_{10}P$ value from high to low. During different growth stages of *C. kyushuensis* Kob, after stimulated by illumination, the mycelium starts to secrete xanthein, turns yellow and enters to coloring stage. Coloring stage is an important stage which concerns if *Cordyceps kyushuensis* Kob could smoothly enter the next developmental stages. There are 552 up-regulated genes between CKK1 and CKK2 (CKK1-VS-CKK2), 491 down-regulated genes between CKK2 and CKK3 (CKK2-VS-CKK3), 419 down-regulated genes between CKK2-VS-CKK4 (CKK2-VS-CKK4). It’s worth nothing that “ribosome”, “translation”, “ligase”, “transcription” and “helicase” were significantly enrichment terms in coloring stage, which implies that more cells were generated

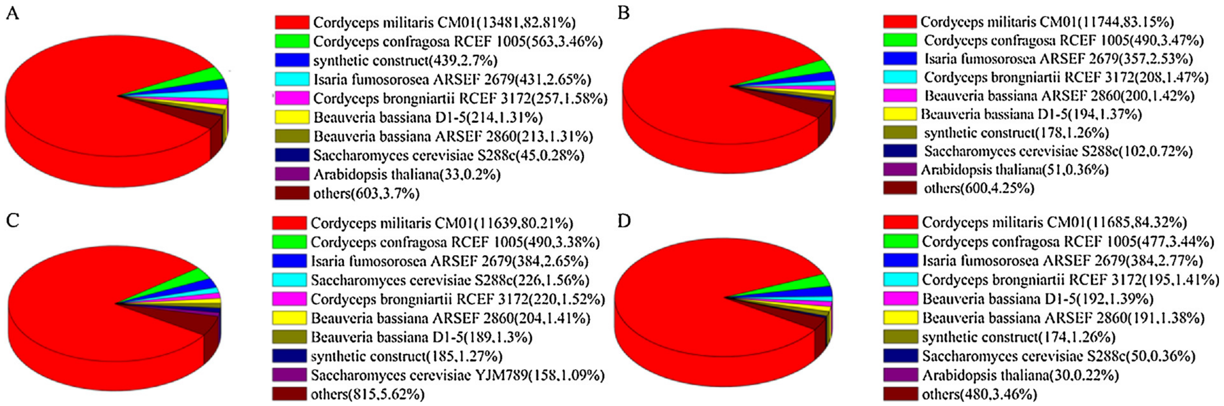

Fig. 2. Species distribution of transcriptome A) CKK1, B) CKK2, C) CKK3, D) CKK4.

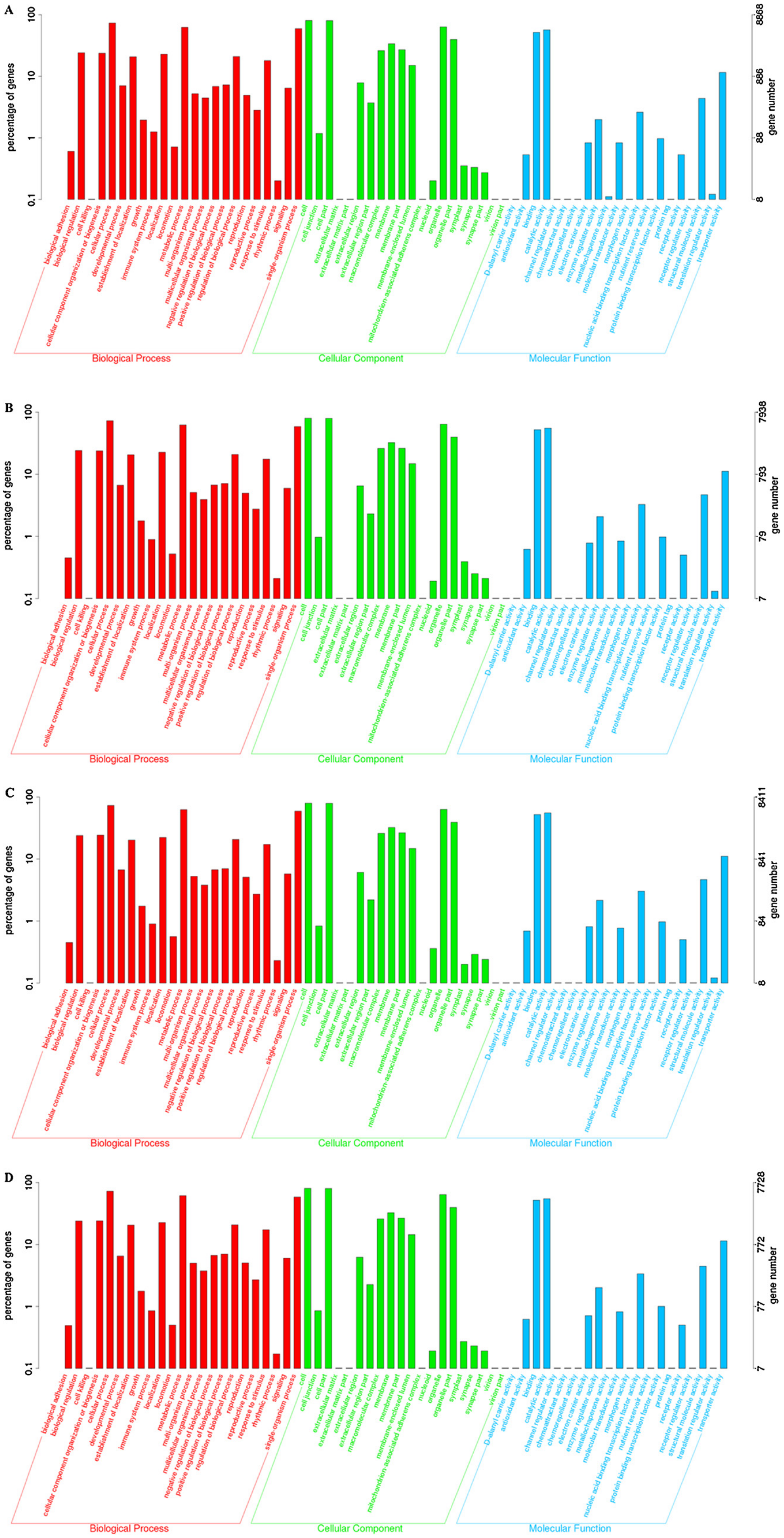

Fig. 3. Gene Ontology (GO) analysis of transcriptome A) CKK1, B) CKK2, C) CKK3, D) CKK4.

during this stage. Stromata-forming initial stage decides if fruiting body could grow successfully. There are 536 up-regulated genes between CKK1 and CKK3 (CKK1-VS-CKK3), 566 up-regulated genes between CKK2 and CKK3 (CKK2-VS-CKK3), 454 down-regulated genes between CKK3-VS-CKK4 (CKK3-VS-CKK4). It’s worth nothing that “ATP”, “carbohydrate”, “helicase”, “ligase” and “ATP-dependent” were

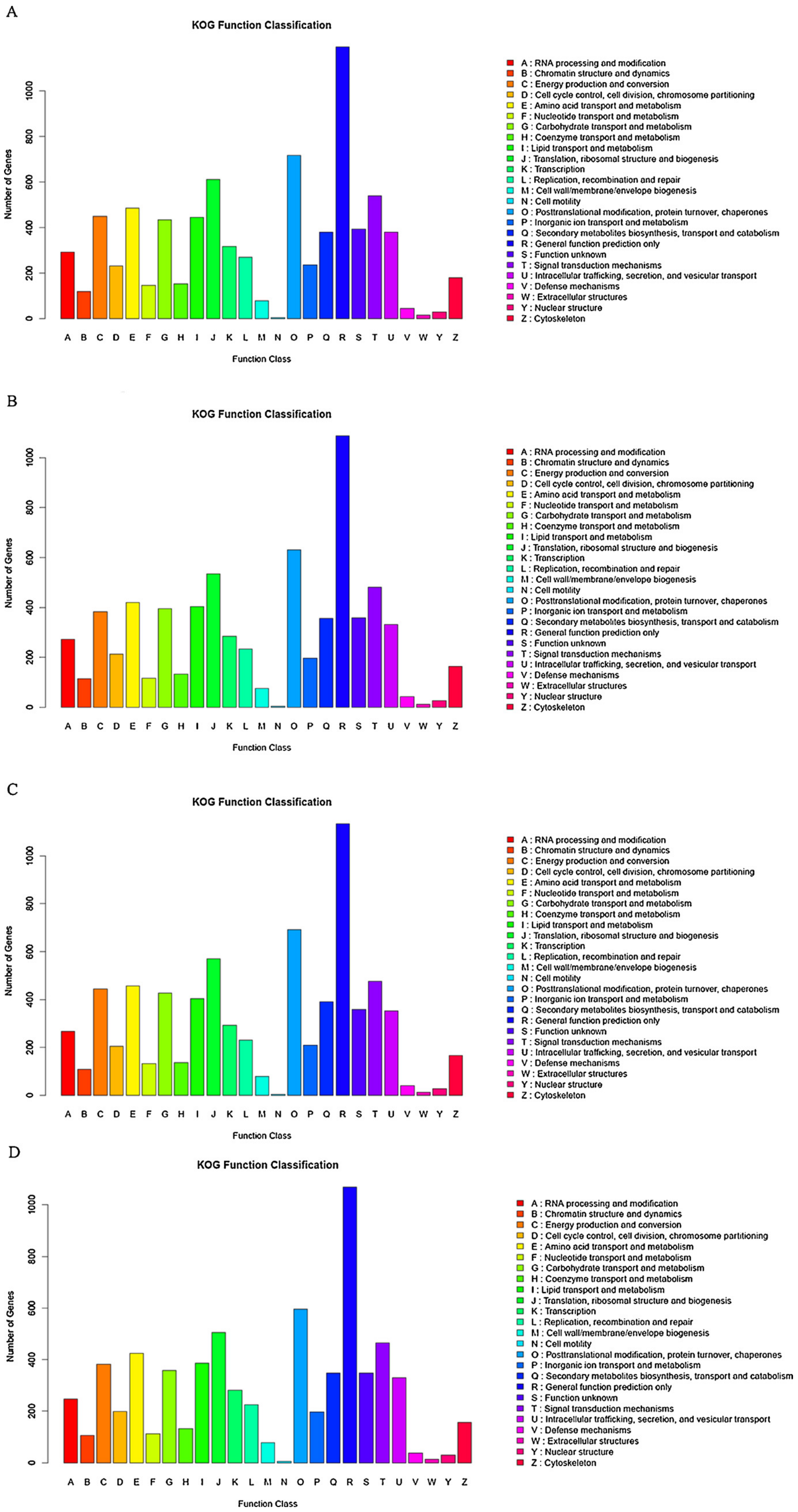

Fig. 4. KOG analysis of transcriptome A) CKK1, B) CKK2, C) CKK3, D) CKK4.

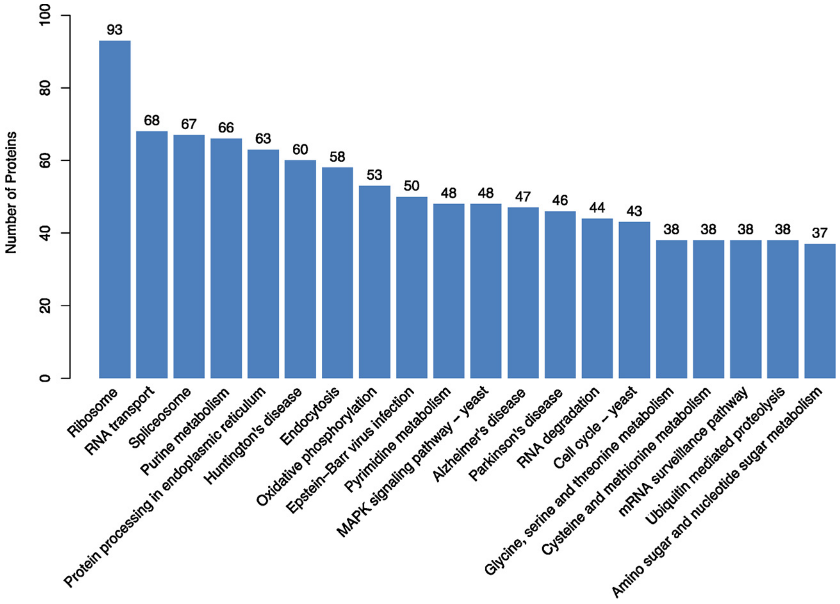

Fig. 5. KEGG analysis of differential expressed proteins in proteomics.

significantly enrichment terms in stromate-forming initial stage, which implies that this stage needs more energy to form stromate. Fruiting body forming stage decides if strong fruiting body could be produced. There are 479 up-regulated genes between CKK1 and CKK4 (CKK1-VS-CKK4), 493 up-regulated genes between CKK2 and CKK4 (CKK2-VS-CKK4), 453 up-regulated genes between CKK3-VS-CKK4 (CKK3-VS-CKK4). It's worth nothing that "ATP", "integral", "structural" were significantly enrichment terms in fruiting body forming stage, which implied that fast growth was going during this stage.

At the same time, different genes usually cooperate to play roles in different biological processes. KEGG pathway enrichment analysis helps us to understand more about the biological functions of genes. The results showed that genes up-regulated in CKK2 (coloring stage) were specifically located in the pathways of "mTOR signaling pathway", "protein processing in endoplasmic reticulum" and "ribosome biogenesis in eukaryotes", genes up-regulated in CKK3 (stromata-forming initial stage) were specifically located in the pathways of "carbon metabolism" and "biosynthesis of amino acids", and genes up-regulated in CKK4 (fruiting body stage) were specifically located in the pathways of "purine metabolism" and "RNA transport", which implied that these up-regulated genes mainly promoted cell growth and proliferation.

3.4. *Functional annotation analysis of proteome difference*

To deeply understand the difference between different developmental stages of *C. kyushuensis* Kob, protein sequencing was done by iTRAQ. Proteins from mycelium stage and fruiting body stage were extracted to be processed following the instruction, then 35,778 unique peptides corresponding to 4185 proteins were identified by searching against C. militaris genome database using ProteomeDiscovererTM Software 2.1. GO, KEGG and KOG were following used to analysis functions of all proteins. According to the standard: FC < 0.67 and FC > 1.5, there were 445 differential expression proteins were selected, including 224 up-regulated proteins and 221 down-regulated proteins between mycelium stage and fruiting body stage.

Among the 445 DEPs, Goatools (http://github.com/tanghaibao/GOatools) was used as the tool to do GO enrichment analysis. Between CKK1 and CKK4, DEPs were mainly enriched in "cellular process", "metabolic process" and "single-organism process" of biological process. At the same time, DEPs were mainly distributed to "cell", "cell part", "membrane", "membrane part" and "organelle" of cellular component, and under the category of molecular function, "binding" and "catalytic activity" were the mostly related terms. All above showed that during different developmental stages, the results all above showed that the expressions in protein level were in keeping with the results in mRNA level.

Except the enrichment analysis of GO, KEGG is also an important database to analysis biological functions of proteins. KOBAS (http://kobas.cbi.pku.edu.cn/home.do) were used to analysis KEGG pathway enrichment. According to Fig. 5, "metabolism" was the mostly significantly class in KEGG pathway enrichment, following were "organismal systems", "cellular processes", "environmental information processing", "Genetic information processing" and "cellular processes". The DEPs were mostly involved in "amino sugar and nucleotide sugar metabolism", "glycolysis/gluconeogenesis", "starch and sucrose metabolism", "carbon fixation pathways in prokaryotes" and "glycerophospholipid metabolism" which accorded with the enrichment of KEGG in transcriptome. Clustering analysis of differential expression protein was processing using hcluster, the result was shown in Fig. 6.

3.5. *Interested proteins and putative genes involved in cordycepin biosynthesis*

There are some interested genes which were deserved to pay more attention on them. In this study, 54 genes related to polyketide synthase which was an important kind of secondary metabolite to catalyzed synthesis polyketide, 1 gene related to calcineurin which was the only serine/threonine phosphatase to be regulated by Ca^{2+} and calmodulin as immunosuppressor, 5 genes related to chitinase which could degrade chitin.

According to previous study, C. militaris can produce both cordycepin and pentostatin by a single gene cluster, which includes four genes designated as cns1-cns4 (CCM_04436 – CCM_04439) [12]. *In vivo*, pentostatin could protect cordycepin from deaminized by adenosine deaminase, which enhances the efficiency of cordycepin. In transcriptome database of *C. kyushuensis* Kob, four related genes (called ck1-ck4) which can synthesis cordycepin and pentostatin were discovered using nucleotide blast and protein blast. In four different developmental stages, culture medium and strain were obtained to extract cordycepin in sterile water with ultrasonic. High-performance liquid chromatography (HPLC) analyses of water extract were performed to detect the existence of cordycepin and pentostatin (Fig. 7). Quantitative

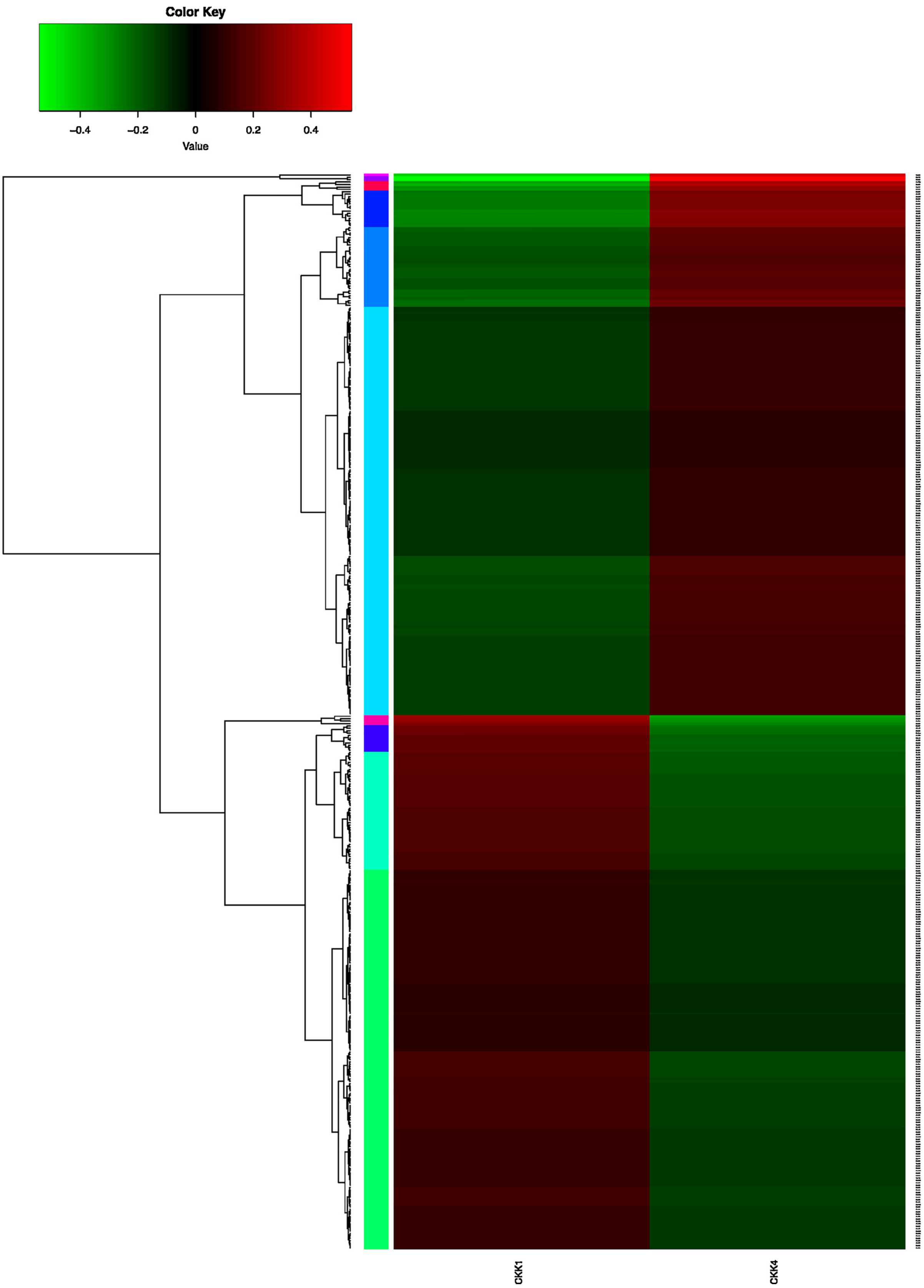

Fig. 6. Clustering analysis of differential expression proteins in CKK1 and CKK4.

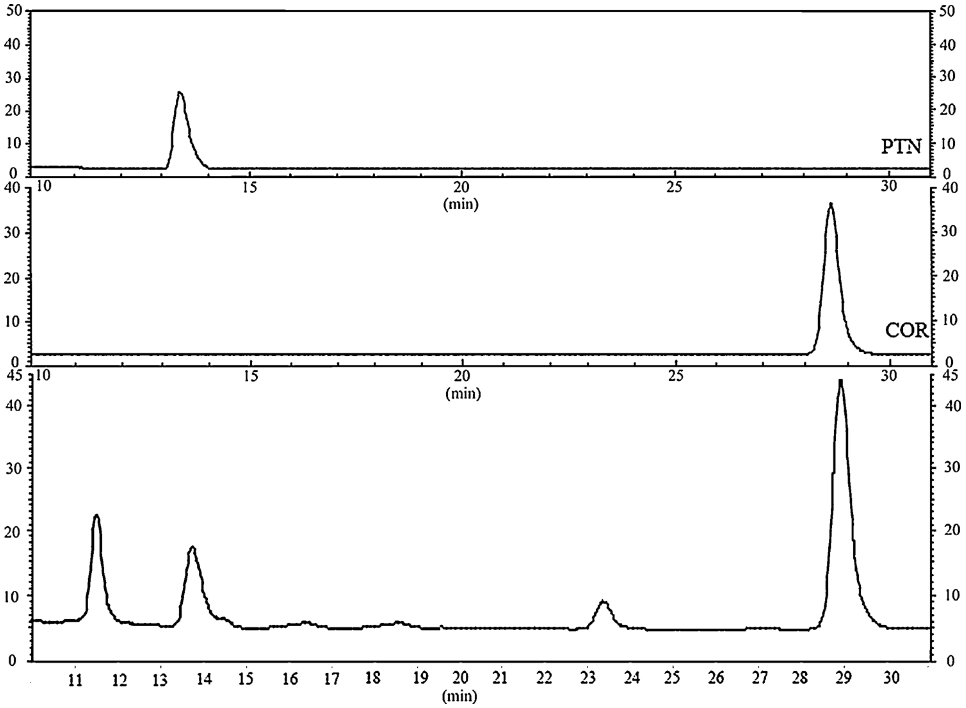

Fig. 7. HPLC analysis of A) pentostatin standard, B) cordycepin standard and C) water extracts of *Cordyceps kyushuensis* which indicated the existence of cordycepin and pentostatin.

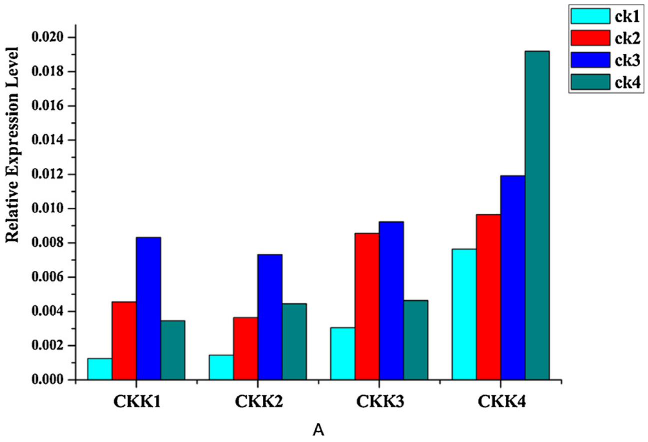

Fig. 8. qPCR analysis of four genes in four developmental stages.

RT-PCR were performed to detect the expression rate of ck1-ck4 (Fig. 8). It showed that ck4 was significantly up-regulated during CKK1 and CKK4 stages, while ck1 was significantly up-regulated during CKK2 stage, ck2 was significantly up-regulated during CKK3 stage. 5′-nucleotidase (CCM_00622, cmNT5E), adenine phosphoribosyltransferase (CCM_00088, cmAPRT) and nucleoside triphosphate pyrophosphatase (CCM_03944, cmNTP) which were reported to be related to cordycepin synthesis are verified that they are not responsible for cordycepin production in C. militaris. Adenosine kinase, which are reported to be related to cordycepin, is absence *C. kyushuensis*, which means that it's independent of the synthesis of cordycepin. The results showed that ck1-ck4 are related to the synthesis of cordycepin and pentostatin.

4. Discussion

In this study, we first reported the transcriptome and proteomics of *C. kyushuensis* Kob, and analyzed the difference among four different developmental stages of it. 16,279 (CKK1), 14,124 (CKK2), 14,510 (CKK3) and 13,858 (CKK4) unigenes were matched in the nr database and C. militaris CM01 was the most matched species in this analysis. Go analysis was performed on basis of the blast NR annotation. For biological process, cellular component and molecular function, most unigenes were involved in "cellular process", "metabolic process", "cell part", "cell" and "catalytic activity", which means *C. kyushuensis* Kob was growing from mycelium stage to fruting body stage. What is noteworthy is that no unigenes were involved in "morphogen activity", morphogen is a kind of protein which is coded by maternal effect gene, which transcribes in oogenesis. That means there was no spores during the four developmental stages. As for KOG and KEGG analysis, the largest group "general function prediction only", "posttranslational modification", "translation, ribosomal structure and biogenesis" in KOG categories and the largest pathways "Ribosome", "biosynthesis of amino acids" and "carbon metabolism" in KEGG pathways were accorded with the GO analysis, which implied that cell growth is lasting during the four developmental stages.

Besides integral analysis, differential expression genes were analyzed among the six groups (CKK1-VS-CKK2, CKK1-VS-CKK3, CKK1-VS-CKK4, CKK2-VS-CKK3, CKK2-VS-CKK4 and CKK3-VS-CKK4). In GO analysis, compared to mycelium stage, coloring stage showed more unigenes involved in cell growth. Compared to mycelium stage and coloring stage, stromata-forming initial stage and fruiting-body stage showed more unigenes involved in producing more energy and fast growth. The analysis results of KEGG pathway enrichment were accorded with the GO analysis, showed that up-regulated genes mainly promoted cell growth and proliferation.

To make the results of transcriptome convincing, proteomics analysis was proceeded. The result of proteomics analysis was in keeping with the results in mRNA level.

Finally, the greatest concern was that biosynthesis genes of cordycepin and pentostatin were identified through transcriptome and proteomics analysis. As a promising compound, cordycepin has lots of pharmacological activities. It is a pity that cordycepin could be deaminized by adenosine deaminase (ADA) *in vivo*, which loses its activity. According to previous studies, pentostatin, as an adenosine deaminase inhibitor, could protects cordycepin from deamination. During this study, four related genes we named ck1-ck4 which synthesis cordycepin and pentostatin were identified using blast in transcriptome database of *C. kyushuensis* Kob. HPLC and quantitative RT-PCR were used to prove the existence of cordycepin and pentostatin in *C. kyushuensis* Kob. According to previous studies, Cns1 and Cns2 interact tightly with each other and that one enzyme cannot function without the other (Xia et al., 2017). We didn't illuminate how they interact with each other here, it requires future study.

In addition to all the above, some genes which were reported to be related to cordycepin biosynthesis, for example 5′-nucleotidase (CCM_00622, cmNT5E), adenine phosphoribosyltransferase (CCM_00088, cmAPRT) and nucleoside triphosphate pyrophosphatase (CCM_03944, cmNTP), are verified that they are not responsible for cordycepin production in C. militaris. It's particular interesting that there is no adenosine kinase biosynthesis gene in *C. kyushuensis*, although the enzyme was reported to be related to cordycepin biosynthesis before.

5. Conclusion

This research reported the transcriptome and proteomics of *C. kyushuensis* Kob for the first time and provided the related analysis to support the following studies. Biosynthesis genes of cordycepin and pentostatin in *C. kyushuensis* Kob were identified, which verified the studies reported before. Improve the yield of cordycepin and enhace the activity of cordycepin *in vivo*. We believe that this study will be helpful to improve the yield of cordycepin and enhance the activity of cordycepin *in vivo*, and provide more useful information to do contribution in the field of medicine.

Conflict of interest

The authors have declared no conflict of interest.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 30770041, 30970012 and 21877075), the China Postdoctoral Science Foundation funded Project, Natural Science Foundation of Shandong Province (No. ZR2015CM028), and the Shandong Provincial Science and Technology Development Program (No. 2009GG10002063 and 2013GSF11715). The authors appreciate their help in earnest.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.micres.2018.09.005.

References

Ahn, Y.J., Park, S.J., Lee, S.G., Shin, S.C., Choi, D.H., 2000. Cordycepin: selective growth inhibitor derived from liquid culture of Cordyceps militaris against Clostridium spp. J. Agric. Food Chem. 48 (7), 2744–2748. https://doi.org/10.1021/Jf990862n.

Brogden, R.N., Sorkin, E.M., 1993. Pentostatin. Drugs 46 (4), 652–677. https://doi.org/10.2165/00003495-199346040-00006.

Cattoi, T., 2013. The trinity and theodicy: the trinitarian theology of Von Balthasar and the problem of evil. Theol. Stud. 74 (1), 212–214. https://doi.org/10.1177/004056391307400121.

Cunningham, K.G., Manson, W., Spring, F.S., Hutchinson, S.A., 1950. Cordycepin, a metabolic product isolated from cultures of cordyceps-militaris (Linn) link. Nature 166 (4231), 949. https://doi.org/10.1038/166949a0.

Dalla Rosa, L., Da Silva, A.S., Gressler, L.T., Oliveira, C.B., Dambros, M.G., Miletti, L.C., et al., 2013. Cordycepin (3' -deoxyadenosine) pentostatin (deoxycoformycin) combination treatment of mice experimentally infected with Trypanosoma evansi. Parasitology 140 (5), 663–671. https://doi.org/10.1017/S0031182012001990.

Dalla Rosa, L., Da Silva, A.S., Oliveira, C.B., Gressler, L.T., Arnold, C.B., Baldissera, M.D., et al., 2015. Dose finding of 3' deoxyadenosine and deoxycoformycin for the treatment of Trypanosoma evansi infection: an effective and nontoxic dose. Microb. Pathogene. 85, 21–28. https://doi.org/10.1016/j.micpath.2015.05.005.

Fan, H., Li, S.P., Xiang, J.J., Lai, C.M., Yang, F.Q., Gao, J.L., Wang, Y.T., 2006. Qualitative and quantitative determination of nucleosides, bases and their analogues in extraction and high performance liquid chromatography-electrospray ionization tandem mass spectrometry. Anal. Chim. Acta 567 (2), 218–228. https://doi.org/10.1016/j.aca.2006.03.032.

Holbein, S., Wengi, A., Decourty, L., Freimoser, F.M., Jacquier, A., Dichtl, B., 2009. Cordycepin interferes with 3′ end formation in yeast independently of its potential to terminate RNA chain elongation. RNA 15, 837–849. https://doi.org/10.1261/rna./1458909.

Lao, Y.Z., Wang, X.Y., Xu, N.H., Zhang, H.M., Xu, H.X., 2014. Application of proteomics to determine the mechanism of action of traditional Chinese medicine remedies. J. Ethnopharmacol. 155 (1), 1–8. https://doi.org/10.1016/j.jep.2014.05.022.

Li, D.S., Liu, X.Z., Zhang, Y., Zhang, C.K., 2002. Contrast analsis of mainly chemical ingredients of Cordyceps mengshanensis and Cordyceps sinensis. Edible Fungi China 21 (5), 35–37. https://doi.org/10.13629/j.cnki.53-1054.2002.05.025.

Liu, Z.Q., Lin, S., Baker, P.J., Wu, L.F., Wang, X.R., Wu, H., et al., 2015. Transcriptome sequencing and analysis of the entomopathogenic fungus Hirsutella sinensis isolated from Ophiocordyceps sinensis. BMC Genomics 16 doi: Artn 106 10.1186/S12864-015-1269-Y.

Liu, J.Y., Chang, M.C., Meng, J.L., Feng, C.P., Liu, Y.N., 2017. iTRAQ-based comparative proteomics analysis of the fruiting dikaryon and the non-fruiting monokaryon of Flammulina velutipes. Curr. Microbiol. 74 (1), 114–124. https://doi.org/10.1007/s00284-016-1164-z.

Metzker, M.L., 2010. Applications of next-generation sequencing sequencing technologies- the next generation. Nat. Rev. Genet. 11 (1), 31–46. https://doi.org/10.1038/nrg2626.

Ni, H., Zhou, X.H., Li, H.H., Huang, W.F., 2009. Column chromatographic extraction and preparation of cordycepin from Cordyceps militaris waster medium. J. Chromatogr. B 877 (22), 2135–2141. https://doi.org/10.1016/j.jchromb.2009.06.009.

Pertea, G., Huang, X.Q., Liang, F., Antonescu, V., Sultana, R., Karamycheva, S., et al., 2003. TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. Bioinformatics 19 (5), 651–652. https://doi.org/10.1093/bioinformatics/btg034.

Rosa, L.D., Da Silva, A.S., Ruchel, J.B., Gressler, L.T., Oliveira, C.B., Franca, R.T., et al., 2013. Influence of treatment with 3' -deoxyadenosine associated deoxycoformycin on hematological parameters and activity of adenosine deaminase in infected mice with Trypanosoma evansi. Exp. Parasitol. 135 (2), 357–362. https://doi.org/10.1016/j.exppara.2013.07.019.

Sarvaria, A., Topp, Z., Saven, A., 2016. Current therapy and new directions in the treatment of hairy cell leukemia a review. JAMA Oncol. 2 (1), 123–129. https://doi.org/10.1001/jamaoncol.2015.4134.

Tsai, Y.J., Lin, L.C., Tsai, T.H., 2010. Pharmacokinetics of adenosine and cordycepin, a bioactive constituent of cordyceps sinensis in rat. J. Agric. Food Chem. 58 (8), 4638–4643. https://doi.org/10.1021/jf100269g.

Tuli, H.S., Sharma, A.K., Sandhu, S.S., Kashyap, D., 2013. Cordycepin: a bioactive metabolite with therapeutic potential. Life Sci. 93 (23), 863–869. https://doi.org/10.1016/j.lfs.2013.09.030.

Vodnala, S.K., Lundback, T., Yeheskieli, E., Sjoberg, B., Gustavsson, A.L., Svensson, R., Olivera, G.C., Eze, A.A., de Koning, H.P., Hammarstrom, L.G., Rottenberg, M.E., 2013. Structure-activity relationships of synthetic cordycepin analogues as experimental therapeutics for African trypanosomiasis. J. Med. Chem. 56 (24), 9861–9873. https://doi.org/10.1021/jm401530a.

Wang, Y.M., Zhang, G.Y., Zhao, X., Ling, J.Y., 2017. Genome shuffling improved the nucleosides production in Cordyceps kyushuensis. J. Biotechnol. 260 (20), 42–47. https://doi.org/10.1016/j.jbiotec.2017.08.021.

Wen, T.C., Jiang, C.Y., Shen, D.T., Li, C.Y., Lu, Q.N., Liu, X.H., et al., 2016. Evaluation of the antitumor activity by cordycepin with gold nanoparticle. J. Nanosci. Nanotechnol. 16 (7), 7134–7139. https://doi.org/10.1166/jnn.2016.11364.

Woo, P.W.K., Dion, H.W., Lange, S.M., Dahl, L.F., Durham, L.J., 1974. Novel adenosine and ara-a deaminase inhibitor, R)-3-(2-deoxy-beta-D-erythro-pentofuranosyl)-3,6,7,8-tetrahydroimidazo[4,5-D] [1,3] diazepin-8-Ol. J. Heterocyclic Chem. 11 (4), 641–643. https://doi.org/10.1002/jhet.5570110438.

Xia, Y.L., Luo, F.F., Shang, Y.F., Chen, P.L., Lu, Y.Z., Wang, C.S., 2017. Fungal cordycepin biosynthesis is coupled with the production of the safeguard molecule pentostatin. Cell Chem. Biol. 24 (12). https://doi.org/10.1016/j.chembiol.2017.09.001. 1479-+.

Yin, Y.L., Yu, G.J., Chen, Y.J., Jiang, S., Wang, M., Jin, Y.X., et al., 2012. Genome-wide transcriptome and proteome analysis on different developmental stages of Cordyceps militaris. PLoS One 7 (12). https://doi.org/10.1371/journal.pone.0051853. ARTN e51853.

Yoshikawa, N., Nakamura, K., Yamaguchi, Y., Kagota, S., Shinozuka, K., Kunitomo, M., 2007. Reinforcement of antitumor effect of Cordyceps sinensis by 2' -deoxycoformycin, an adenosine deaminase inhibitor. In Vivo (Brooklyn) 21 (2), 291–295.

Zhang, G.Y., Yin, Q.S., Han, T., Zhao, Y.X., Su, J.J., Li, M.Z., Ling, J.Y., 2015. Purification and antioxidant effect of novel fungal polysaccharides from the stroma of Cordyceps kyushuensis. Ind. Crop Prod. 69, 485–491. https://doi.org/10.1016/j.indcrop.2015.03.006.

Zhou, X.X., Luo, L.P., Dressel, W., Shadier, G., Krumbiegel, D., Schmidtke, P., Zepp, F., Meyer, C.U., 2008. Cordycepin is an immunoregulatory active ingredient of Cordyceps sinensis. Am. J. Med. 36 (05), 967. https://doi.org/10.1142/S0192415X08006387.