12-2014

A specific type of cyclin-like F-box domain gene is involved in the cryogenic autolysis of *Volvariella volvacea*

Ming Gong

Mingjie Chen

Hong Wang

Qiuming Zhu
University of Nebraska at Omaha, qzhu@unomaha.edu

Qi Tan

Mycologia, 107(2), 2015, pp. 313–318. DOI: 10.3852/14-159

A specific type of cyclin-like F-box domain gene is involved in the cryogenic autolysis of *Volvariella volvacea*

Ming Gong

National Engineering Research Center of Edible Fungi, Ministry of Science and Technology; Key Laboratory of Edible Fungi Resources and Utilization (South), Ministry of Agriculture; Institute of Edible Fungi, Shanghai Academy of Agricultural Sciences, Shanghai 201403, P.R. China; Department of Computer Science, University of Nebraska at Omaha, Omaha, Nebraska 68182

Mingjie Chen
Hong Wang

National Engineering Research Center of Edible Fungi, Ministry of Science and Technology; Key Laboratory of Edible Fungi Resources and Utilization (South), Ministry of Agriculture; Institute of Edible Fungi, Shanghai Academy of Agricultural Sciences, Shanghai 201403, P.R. China

Qiuming Zhu[1]

Department of Computer Science, University of Nebraska at Omaha, Omaha, Nebraska 68182

Qi Tan[1]

National Engineering Research Center of Edible Fungi, Ministry of Science and Technology; Key Laboratory of Edible Fungi Resources and Utilization (South), Ministry of Agriculture; Institute of Edible Fungi, Shanghai Academy of Agricultural Sciences, Shanghai 201403, P.R.China

Abstract: Cryogenic autolysis is a typical phenomenon of abnormal metabolism in *Volvariella volvacea*. Recent studies have identified 20 significantly up-regulated genes via high-throughput sequencing of the mRNAs expressed in the mycelia of *V. volvacea* after cold exposure. Among these significantly up-regulated genes, 15 annotated genes were used for functional annotation cluster analysis. Our results showed that the cyclin-like F-box domain (FBDC) formed the functional cluster with the lowest P-value. We also observed a significant expansion of FBDC families in *V. volvacea*. Among these, the FBDC3 family displayed the maximal gene expansion in *V. volvacea*. Gene expression profiling analysis revealed only one FBDC gene in *V. volvacea* (*FBDV1*) that was significantly up-regulated, which is located in the FBDC3 family. Comparative genomics analysis revealed the homologous sequences of *FBDV1* with high similarity were clustered on the same scaffold. However, *FBDV1* was located far from these clusters, indicating the divergence of duplicated genes. Relative time estimation and rate test provided evidence for the divergence of *FBDV1* after recent duplications. Real-time RT-PCR analysis confirmed that the expression of the *FBDV1* was significantly up-regulated ($P < 0.001$) after cold-treatment of *V. volvacea* for 4 h. These observations suggest that the *FBDV1* is involved in the cryogenic autolysis of *V. volvacea*.

Key words: abnormal metabolism, autolysis, gene expansion, F-box domain, *Volvariella volvacea*

Submitted 16 Jun 2014; accepted for publication 3 Nov 2014.

[1] Corresponding authors. E-mail addresses: qzhu@mail.unomaha.edu, tanqsaas@gmail.com

Introduction

Volvariella volvacea is an important edible mushroom that has been cultivated for more than 300 y. *V. volvacea* is a tropical fungus that requires high temperatures (28–35 C) for vegetative growth and fruiting (Sun et al. 2006). Temperatures below 15 C cause damage to the fruiting body and adversely affect the viability of the *V. volvacea* mycelia (Bao et al. 2013). Routine storage at low temperatures (4 C) causes autolysis of the *V. volvacea* mycelium, and the fruiting body becomes soft, liquified and even rotten (Chang 1978). Thus cryogenic autolysis is a typical phenomenon of abnormal metabolism.

In previous studies the cloning and isolation of cold-induced genes (Chen et al. 2001, Guo et al. 2004) were employed to study the cryogenic autolysis of *V. volvacea*. Knowledge of cold-induced genes such as the cold-induced plant *COR* gene (Artus et al. 1996) was further used to analyze the expression of these genes (Qiao et al. 2009). However, cold-induced genes from other species do not represent the molecular features found in *V. volvacea*. Genome sequencing and gene annotation of multiple edible fungi, especially *V. volvacea* (Bao et al. 2013), have enabled the investigation of the molecular mechanism of the cryogenic autolysis of *V. volvacea* at the genome level.

In the present study the mycelia of *V. volvacea* (V23 strain) were subjected to 4 C for 0, 2 and 4 h. High-throughput sequencing was then carried out to detect the mRNAs expressed in each phase to study the cryogenic autolysis of *V. volvacea* (Bao et al. 2013). The mRNA expression observed in cold exposed *V. volvacea* is representative of the molecular features of cryogenic autolysis, which constitutes the basis for our study of cryogenic autolysis.

MATERIALS AND METHODS

Functional annotation clustering analysis.—The foundation of enrichment analysis is that, if a biological process is abnormal in a given study, the co-functioning genes should exhibit a higher potential (i.e. enrichment) to be selected as a relevant group via high-throughput screening technologies (Huang da et al. 2009a, b). Based on this type of analysis, cold-induced gene expression profiles were determined via high-throughput sequencing analysis of mRNAs expressed in the mycelia of *V. volvacea* (V23 strain) after exposure to 4 C for 0, 2 and 4 h (Bao et al. 2013). DEGseq software (Wang et al. 2010) was used to normalize the data and assess significant differences in gene expression, and $|\log_2 \text{(fold-change)}| \geq 1$, FDR < 0.001 was used as a threshold to determine significant changes at 2 and 4 h. InterproScan was employed for functional annotation. PFAM (http://pfam.sanger.ac.uk/) was used for PFAM ID information analysis. DAVID (Huang da et al. 2009b) was employed to test the enrichment of biological processes in the significantly regulated genes ($|\log2(\text{fold-change})| \geq 1$, FDR < 0.001). Because backgrounds containing large numbers of genes tend to give rise to more significant *P* values in enrichment analysis (Huang da et al. 2009b), human genes (30 000 genes total) were used as a background. A modified Fisher's exact test was used to examine the significance of gene-term enrichment (Huang da et al. 2009b). Only functional annotation clusters with *P* value $< 1e \times 10^{-50}$ were considered to be enriched.

Protein family analysis.—Because basidiomycete and ascomycete fungi show independent evolutionary trajectories (Gong et al. 2013), eight representative basidiomycete species in various niches were selected for comparative genomic analysis. The details regarding the generation of genomic data and the construction of the phylogenomic tree can be found in (Bao et al. 2013). Multigene families were generated from all the predicted proteins in eight representative basidiomycete genomes using SCPS tools (Paccanaro et al. 2006) with the default settings (BLASTp, cut-off e-value $\leq 1e^{-30}$).

The cyclin-like F-box domain (FBDC) sequences present in the *V. volvacea* genome were obtained through InterproScan (http://www.ebi.ac.uk/interpro/) analysis. The homologous sequences of these FBDC sequences in *V. volvacea* were obtained via BLAST queries (BLASTp, cut-off e-value $\leq 1e^{-35}$) in seven other representative basidiomycete genomes. The obtained homologous sequences and *V. volvacea* FBDC sequences then were subjected to BLASTp queries against the eight representative basidiomycete genomes (BLASTp, cut-off e-value $\leq 1e^{-50}$). Next, the distribution of the multigene FBDC families was generated from BLAST query results with SCPS tools (Paccanaro et al. 2006) with default settings.

Phylogenetic analysis.—Protein sequences were aligned with Clustal W (Thompson et al. 1994) with default parameters. WAG model is best with respect to maximum-likelihood (ML) values in the phylogenetic inference from protein sequence data using ML estimation (Whelan and Goldman 2001). Thus the topology of the phylogenetic tree was constructed via the ML estimation (bootstrap = 1000, WAG + Gamma2) using MEGA 5 software (Tamura et al. 2011). The topology acts as the user-supplied topology for generating the following linearized tree.

Relative time estimation and rate test.—DNA sequences were aligned with Clustal W (Thompson et al. 1994) with default parameters for the relative time estimation and rate test. The linearized tree obtained under the assumption of a molecular clock can be useful to estimate the relative timing of sequence divergence events (Takezaki et al. 1995). Thus ML branch lengths for all branching points in the user-supplied topology are estimated by assuming equal evolutionary rate among lineages (Tamura-Nei model + Gamma2) using MEGA 5 (Tamura et al. 2011) for generating the linearized tree. The analysis involved 17 nucleotide sequences. A total of 2035 positions were in the final dataset.

Sequences A, B and C (an outgroup) were selected for the Tajima's relative rate test (Tamura and Nei 1993) using MEGA 5. *P* value less than 0.05 is often used to reject the null hypothesis of equal rates between lineages. The analysis involved three nucleotide sequences. All positions containing gaps and missing data were eliminated. A total of 1529 positions were in the final dataset.

Analysis of qPCR assays.—To assess cold-induced gene expression in *V. volvacea* (Bao et al. 2013), fungal mycelia from cultures of the *V. volvacea* strain V23 were grown in 250 mL flasks containing 100 mL potato dextrose broth (PDB) at 32 C for 4 d. The flasks were exposed to 4 C for 0, 2, 4, 6 and 8 h. The treated mycelia subsequently were collected and lyophilized and total RNA was extracted with the TRIzol reagent (Invitrogen) following the manufacturer's instructions. DNA was removed by treatment with RNase-free DNase, and cDNA was synthesized in a reaction mixture containing random primers (Invitrogen), a dNTP mixture, an RNase inhibitor, DTT, SuperScript III Reverse Transcriptase (Invitrogen) and DEPC-treated water. Real-time quantitative PCR (qPCR) was performed on a 7500 fast sequence detection system (Applied Biosystems, Foster City, California) following the manufacturer's instructions. The primers employed for qPCR are provided (SUPPLEMENTARY TABLE I). The qPCR assay was performed in triplicate. Glyceraldehyde-3-phosphate dehydrogenase (*GPD*) gene expression is not induced by cold-stress treatment in *V. volvacea* (Sun et al. 2008); therefore *GPD* was used as a reference (GenBank accession number DQ140384) to normalize the expression of target genes. The relative quantification of transcripts was performed via the comparative CT method ($2^{-\Delta\Delta CT}$), calibrated to *GPD*. One-way ANOVA, followed by the Holm-Sidak test for multiple comparisons, was conducted using SigmaStat Windows 3.5 to evaluate significant differences in the mean values between groups ($P < 0.001$).

RESULTS

Enrichment of the FBDC annotation cluster.—Based on the high-throughput sequencing of the mRNAs

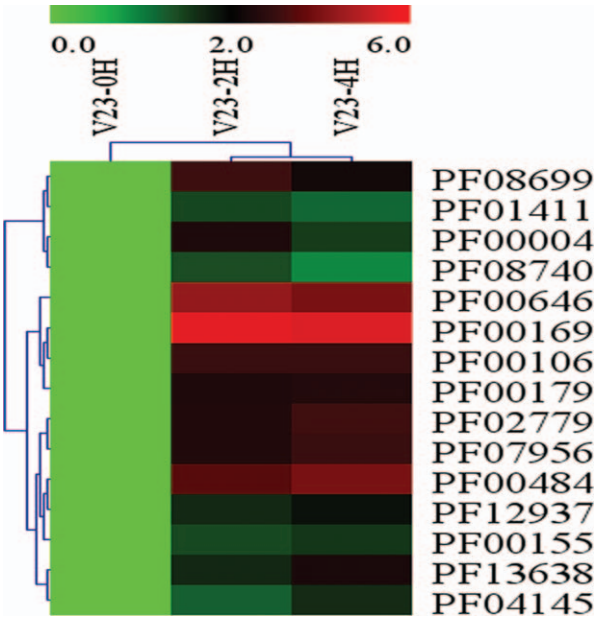

FIG. 1. Analysis of the cold-induced expression profile of *V. volvacea*. Pfam ID information for the annotated genes is listed in the right column. The gene's relative intensity is represented using a logarithmic scale ($\log_2$X) of the fold increase, as shown in the scale bar. Pfam ID information for significantly up-regulated genes was used for DAVID (Huang da et al. 2009b) enrichment analysis.

expressed in the mycelia of *V. volvacea* (V23 strain) (Bao et al. 2013), we identified 20 genes that were significantly up-regulated [$\log_2$(fold-change) $\geq$ 1, FDR $<$ 0.001] at 2 and 4 h with DEGseq software, and 15 of these were annotated with PFAM ID information (FIG. 1). Because DAVID is suited for analyzing smaller gene lists (Huang da et al. 2009b), we employed this resource to analyze the functional annotation enrichment of the 15 PFAM ID information. We found two annotation clusters (*P* value $<$ 1e $\times$ 10^{-50}) associated with FBDC and ubiquitin-conjugating enzyme E2 (UBE2). The FBDC functional annotation cluster had the lowest *P* value (4.8e-69). These results indicate that FBDC is likely associated with the cryogenic autolysis of *V. volvacea*.

Obvious expansion of FBDC families.—On the basis of the contraction tendency of *V. volvacea* gene families (Bao et al. 2013), the number of gene families was compared at different scales between eight representative basidiomycete species. The results revealed that an abnormal expansion of large families (>200 members) has occurred in *V. volvacea* (SUPPLEMENTARY TABLE II), which was observed for fam1, fam4 and fam6. Further analysis with InterproScan showed that the fam1 of *V. volvacea* consisted predominantly of FBDC genes. Together with the enriched FBDC functional annotation cluster identified with DAVID, the significant expansion of the FBDC genes in *V. volvacea* fam1 further indicates that FBDC is closely associated with the cryogenic autolysis of *V. volvacea*.

A total of 105 FBDC genes were identified in the *V. volvacea* genome through InterproScan analysis, and their homologous sequences were obtained via BLAST queries of representative basidiomycete genomes. The multigene families of the FBDC homologous genes identified with SCPS tools (Paccanaro et al. 2006) showed that the number of genes (211) belonging to FBDC families in *V. volvacea* was obviously higher than the average gene number (27) found in all other species (FIG. 2A). Further hierarchical clustering revealed that the gene numbers of FBDC families were mainly enriched in *V. volvacea* (FIG. 2B), such as those of the FBDC1, FBDC3, FBDC6 and FBDC7 families. Among these families, FBDC3 exhibited the maximum gene expansion (24) in *V. volvacea* (FIG. 2B).

A specific type of FBDC showing significantly up-regulated expression.—Gene expression profiling analysis revealed only one FBDC gene in *V. volvacea* (*FBDV1*) that was significantly up-regulated [$\log_2$ (fold-change) $\geq$ 1, FDR $<$ 0.001] among the FBDC families (SUPPLEMENTARY TABLE III). The *FBDV1* is located in the FBDC3 family, which indicates that *FBDV1* likely is associated with the cryogenic autolysis of *V. volvacea*.

BLASTp results (cut-off e-value $\leq$ $1e^{-35}$) indicated that existence of 25 homologous sequences of FBDV1 in *V. volvacea*. However, no homologous sequences of FBDV1 were found in other species. Further BLASTp queries of FBDV1 conducted in the NCBI nr database (cut-off e-value $\leq$ 10^{-5}) also revealed no sequences homologous to FBDV1. Thus, 17 homologous sequences of FBDV1 (VVO_06128) in *V. volvacea* were used for phylogenetic analysis. Obviously long and short sequences as well as sequences with a cut-off e-value $\geq$ $1e^{-50}$ were not included in the analysis to improve the accuracy of the phylogenetic tree. The phylogenetic tree clearly showed the existence of three groups (cluster 1, 2 and 3) (FIG. 3A), located at the corresponding three scaffolds (VVO_S00011, VVO_S00006 and VVO_S00022, respectively). Analysis of the nucleotide percent identity showed that the homologous sequences of *FBDV1* with high nucleotide percent identities were distributed in the two clusters on the VVO_S00011 scaffold (FIG. 3B). However, *FBDV1* (VVO_06218) was far from two clusters on the same scaffold (FIG. 3B), indicating the divergence of duplicated genes.

The relative time estimation indicated that recent duplications resulted in the divergence of *FBDV1*

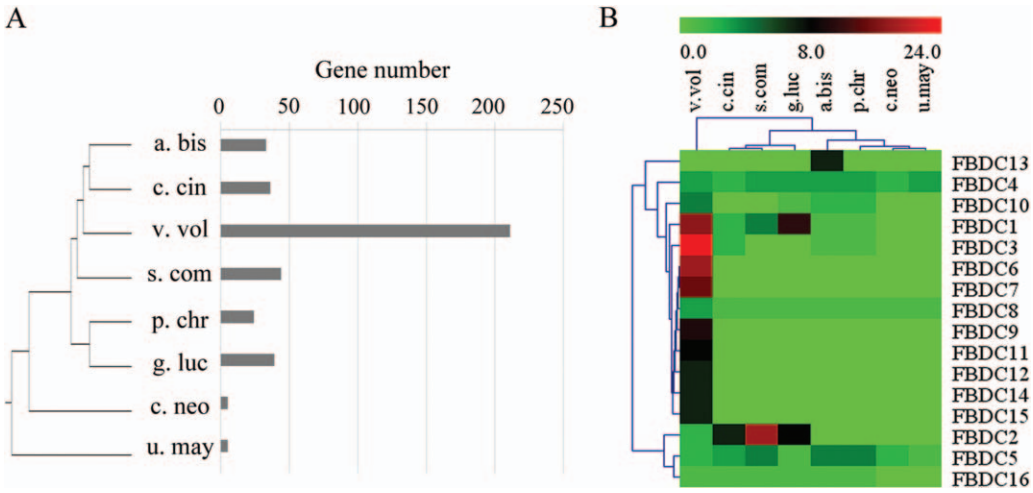

FIG. 2. Comparison of gene numbers and hierarchical clustering of FBDC multigene families. A. Gene numbers of FBDC multigene families in representative basidiomycetes. The phylogenomic tree was constructed using representative basidiomycetes. All bootstrap values are 100%. B. Hierarchical clustering of FBDC multigene families in representative basidiomycetes. Acronyms of basidiomycete species are provided above, and the names of FBDC families are in the right column. The bar at the top of the panel represents the numbers of genes belonging to FBDC families, 0–24. The following abbreviations are used: A. bis = *Agaricus Bisporus*, c. cin = *Coprinopsis cinerea*, v. vol = *Volvariella volvacea*, s.com = *Schizophyllum commune*, p. chr = *Phanerochaete chrysosporium*, g. luc = *Ganoderma lucidum*, c. neo = *Cryptococcus neoformans*, and u. may = *Ustilago maydis*.

(VVO_06218), VVO_06196 and VVO_06181 (SUPPLEMENTARY FIG. 1). Together with the same scaffold distribution (FIG. 3B), sequences A (VVO_06218), B (VVO_06196) and C (VVO_06181, an outgroup) were selected for the Tajima's relative rate test. The results revealed that the null hypothesis of equal evolutionary rate was rejected at an extremely significant level (P = 0.00161). These observations provided evidence for the divergence of *FBDV1* after recent duplications.

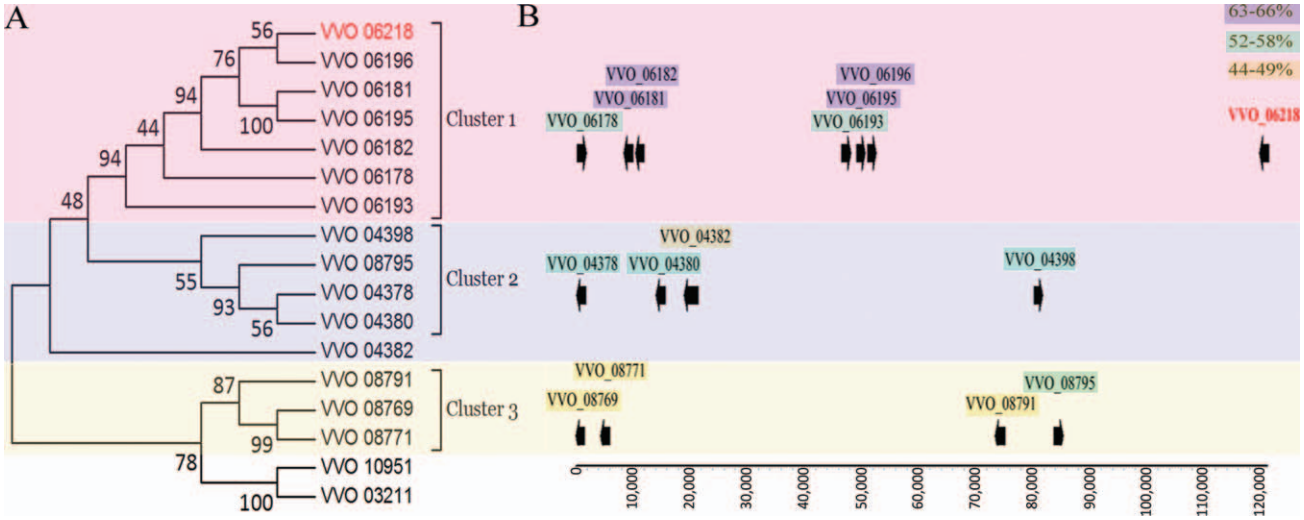

FIG. 3. Homologous sequences of *FBDV1* and their genomic distributions. A. Phylogenetic tree of the homologous sequences of FBDV1 (VVO_06218). The topology of the phylogenetic tree was constructed via the maximum likelihood estimation (bootstrap = 1000, WAG + Gamma2) with MEGA5 software (Tamura et al. 2011). VVO_06218 is indicated at the top in lightface gray (print image) or in red (online image). B. Genomic distributions of the homologous genes of *FBDV1* in *V. volvacea*. ChromoMapper 1.0 (Niculita-Hirzel and Hirzel 2008) was used for plotting the genomic distribution. The three scaffolds were indicated by different background colors. The percent nucleotide identity between VVO_06218 and its homologous sequences was calculated with clustal omega (Sievers et al. 2011).

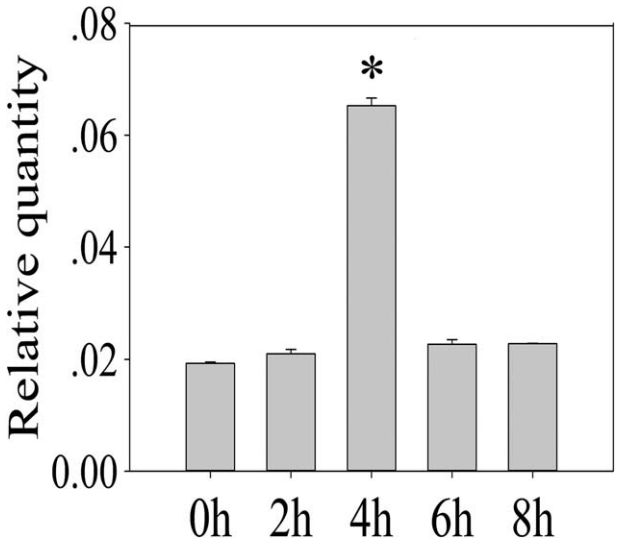

FIG. 4. qPCR analysis of *V. volvacea FBDV1*. Bars represent the mean ± SEM, while an asterisk indicates a significant difference ($P < 0.001$) in comparison with 0 h.

Real-time RT-PCR analysis confirmed that the expression of *FBDV1* was significantly up-regulated by 4 h of cold treatment (FIG. 4; $P < 0.001$). The expression of *FBDV1* showed the highest concentration at 4 h, when it was 3.56-fold higher than the concentration at 0 h. This provides additional support for the possibility that the *FBDV1* is involved in the cryogenic autolysis of *V. volvacea*.

DISCUSSION

The FBDC proteins control protein stability and regulate a wide-variety of cellular processes by targeting diverse substrates for ubiquitination (Ho et al. 2006). The significantly up-regulated gene expression of *FBDV1* observed after cold exposure (FIG. 4) indicates that *FBDV1* may function in cryogenic autolysis through ubiquitination. In addition to the FBDC enrichment annotation cluster, UBE2 constituted another enrichment annotation cluster. Of note, UBE2 functions at the heart of the ubiquitin transfer pathway and is responsible for much of the diversity of ubiquitin cellular signaling (Wenzel et al. 2011). Therefore the significant up-regulation of *UBE2* (Bao et al. 2013) and *FBDV1* (FIG. 4) observed after cold-shock treatment of *V. volvacea* further suggests that low temperatures may induce protein ubiquitination and subsequently initiate myriad cellular processes.

Significant expansion of the three large (>200 members) gene families also has been found in fam4 and fam6 in addition fam1. Further analysis with InterproScan revealed that fam4 consisted mainly of NACHT nucleoside triphosphatases and fam6 consisted mainly of reverse transcriptase enzymes. The cryogenic autolysis of *V. volvacea* also may be associated with NACHT nucleoside triphosphatases or reverse transcriptase enzymes besides FBDC.

We also conducted the enrichment analysis of the functional annotation of the 33 significantly down-regulated genes annotated with PFAM ID information with DAVID (SUPPLEMENTARY FIG. 2). We found four annotation clusters (P value $< 1e\times10^{-50}$), which were associated with adenyl nucleotide binding, ATP binding, AGC-kinase and the heat-shock protein DnaJ. The first three clusters are all directly or indirectly correlated with ATP binding according to gene ontology annotations, suggesting that cryogenic autolysis of *V. volvacea* also may be correlated with the ATP binding.

Our observations indicate that the cryogenic autolysis of *V. volvacea* is not a simple process but instead involves complex and abnormal metabolic activities. The detailed molecular mechanisms of the cryogenic autolysis of *V. volvacea* require further elucidation.

Based on the gene expression profiling and comparative genomics analysis, we identified the specific *FBDV1* that is closely correlated with the cryogenic autolysis of *V. volvacea*. Real-time RT-PCR analysis provided further experimental evidence. Thus we conclude that the *FBDV1* is involved in the cryogenic autolysis of *V. volvacea*. This finding will aid in understanding the molecular mechanism of the cryogenic autolysis of *V. volvacea*.

ACKNOWLEDGMENTS

This work was supported by the grants from the National Science and Technology Pillar Program (2013BAD16B02), Shanghai Municipal Science and Technology Commission [10DZ2212400], Shanghai Agriculture Committee (HuNongKeGongZi2011D-1-2), and Shanghai Agricultural Science Development 2013(10).

LITERATURE CITED

Artus NN, Uemura M, Steponkus PL, Gilmour SJ, Lin C, Thomashow MF. 1996. Constitutive expression of the cold-regulated Arabidopsis thaliana COR15a gene affects both chloroplast and protoplast freezing tolerance. Proc Natl Acad Sci U S A 93:13404–13409, doi:10.1073/pnas.93.23.13404

Bao D, Gong M, Zheng H, Chen M, Zhang L, Wang H, Jiang J, Wu L, Zhu Y, Zhu G, Zhou Y, Li C, Wang S, Zhao Y, Zhao G, Tan Q. 2013. Sequencing and comparative analysis of the straw mushroom (*Volvariella volvacea*) genome. PLoS One 8:e58294, doi:10.1371/journal.pone.0058294

Chang ST. 1978. Volvariella volvacea. In: Chang ST, Hayes WA, eds. The biology and cultivation of edible mushrooms. London, UK: Academic Press. p 573—603.

Chen MJ, Tan Q, Cao H, Wang N, Huang WY, Pan YJ. 2001. Isolation of cold-induced gene from *Volvariella volvacea* by mRNA differential display. Mycosystema 20:342–346.

Gong M, Bao D, Wang H, Tan Q, Chen M. 2013. Taxonomic status and divergence time of Volvariella volvacea. Acta Edulis Fungi 20:8–12.

Guo LQ, Lin JF, Yang LQ, Liu RR. 2004. Isolation of cold-induced genes by modified cDNA-AFLP with single enzyme digestion from Volvariella volvacea. Mycosystema 23:241–247.

Ho MS, Tsai PI, Chien CT. 2006. F-box proteins: the key to protein degradation. J Biomed Sci 13:181–91, doi:10.1007/s11373-005-9058-2

Huang da W, Sherman BT, Lempicki RA. 2009a. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 37:1–13, doi:10.1093/nar/gkn923

———. 2009b. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4:44–57, doi:10.1038/nprot.2008.211

Niculita-Hirzel H, Hirzel AH. 2008. Visualizing the gene order conservation among genomes with ChromoMapper. Int J Comput Intell Bioinform Syst Biol 1.

Paccanaro A, Casbon JA, Saqi MA. 2006. Spectral clustering of protein sequences. Nucleic Acids Res 34:1571–1580, doi:10.1093/nar/gkj515

Qiao N, Wang H, Chen MJ. 2009. Expression of cold-induced gene Cor1 in Volvariella volvacea during low temperature treatment. Mycosystema 28:213–219.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega. Mol Syst Biol 7:539, doi:10.1038/msb.2011.75

Sun X, Chen M, Pan Y. 2008. Cloning, sequence analysis and putative promoter region of a glyceraldehyde-3-phosphate dehydrogenase gene from *Volvariella volvacea*, and expression pattern under cold stress. Acta Edulis Fungi 15:17–31.

Sun XH, Feng AP, Chen MJ, Pan YJ. 2006. Cloning and sequence analysisof cold induced genes in Chinese straw mushroom *Volvariella volvacea*. Mycosystema 25: 88–93.

Takezaki N, Rzhetsky A, Nei M. 1995. Phylogenetic test of the molecular clock and linearized trees. Mol Biol Evol 12:823–833.

Tamura K, Nei M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 10:512–526.

———, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA 5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731–2739, doi:10.1093/molbev/msr121

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680, doi:10.1093/nar/22.22.4673

Wang L, Feng Z, Wang X, Wang X, Zhang X. 2010. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. Bioinformatics 26:136–138, doi:10.1093/bioinformatics/btp612

Wenzel DM, Stoll KE, Klevit RE. 2011. E2s: structurally economical and functionally replete. Biochem J 433: 31–42, doi:10.1042/BJ20100985

Whelan S, Goldman N. 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol 18:691–699, doi:10.1093/oxfordjournals.molbev.a003851