Genome-wide characterization of the Zn(II)$_2$Cys$_6$ zinc cluster-encoding gene family in *Pleurotus ostreatus* and expression analyses of this family during developmental stages and under heat stress

Zhihao Hou$^{1,2}$, Qiang Chen$^{1,2}$, Mengran Zhao$^{1,2}$, Chenyang Huang* Corresp.$^{1,2}$, Xiangli Wu* Corresp.$^{1,2}$

$^1$ Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing, China
$^2$ Key Laboratory of Microbial Resources, Ministry of Agriculture and Rural Affairs, Beijing, China

Corresponding Authors: Chenyang Huang, Xiangli Wu
Email address: huangchenyang@caas.cn, wuxiangli@caas.cn

*Pleurotus ostreatus* is one of the most widely cultivated mushrooms in China. The regulatory mechanisms of fruiting body formation and the response to heat stress in *P. ostreatus* are main research focuses. The Zn(II)$_2$Cys$_6$ family is one of the largest families of transcriptional factors and plays important roles in multiple biological processes in fungi. In this study, we identified 66 zinc cluster proteins in *P. ostreatus* (PoZCPs) through a genome-wide search. The PoZCPs were classified into 15 types according to their zinc cluster domain. Physical and chemical property analyses showed a huge diversity among the PoZCPs. Phylogenetic analysis of PoZCPs classified these proteins into six groups, and conserved motif combinations and similar gene structures were observed in each group. The expression profiles of these PoZCP genes during different developmental stages and under heat stress were further investigated by RNA-sequencing (RNA-seq), revealing diverse expression patterns. Thirteen PoZCPs that may participate in development or the heat stress response were selected for validation of their expression levels through real-time quantitative PCR (RT-qPCR) analysis, and some developmental stage-specific and heat stress-responsive candidates were identified. The findings contribute to our understanding of the roles and regulatory mechanisms of ZCPs in *P. ostreatus*. 
Genome-wide characterization of the Zn(II)$_2$Cys$_6$ zinc cluster-encoding gene family in *Pleurotus ostreatus* and expression analyses of this family during developmental stages and under heat stress

Zhihao Hou $^{1,2}$, Qiang Chen $^{1,2}$, Mengran Zhao $^{1,2}$, Chenyang Huang $^{1,2}$, Xiangli Wu $^{1,2}$

1 Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing, China
2 Key Laboratory of Microbial Resources, Ministry of Agriculture and Rural Affairs, Beijing, China

Corresponding Author:
Chenyang Huang $^{1,2}$
Xiangli Wu $^{1,2}$

12 Zhongguancun South Street, Beijing 100081, China.

Email address:
huangchenyang@caas.cn (C.H.)
wuxiangli@caas.cn (X.W.)
Abstract

*Pleurotus ostreatus* is one of the most widely cultivated mushrooms in China. The regulatory mechanisms of fruiting body formation and the response to heat stress in *P. ostreatus* are main research focuses. The Zn(II)$_2$Cys$_6$ family is one of the largest families of transcriptional factors and plays important roles in multiple biological processes in fungi. In this study, we identified 66 zinc cluster proteins in *P. ostreatus* (PoZCPs) through a genome-wide search. The PoZCPs were classified into 15 types according to their zinc cluster domain. Physical and chemical property analyses showed a huge diversity among the PoZCPs. Phylogenetic analysis of PoZCPs classified these proteins into six groups, and conserved motif combinations and similar gene structures were observed in each group. The expression profiles of these PoZCP genes during different developmental stages and under heat stress were further investigated by RNA-sequencing (RNA-seq), revealing diverse expression patterns. Thirteen PoZCPs that may participate in development or the heat stress response were selected for validation of their expression levels through real-time quantitative PCR (RT-qPCR) analysis, and some developmental stage-specific and heat stress-responsive candidates were identified. The findings contribute to our understanding of the roles and regulatory mechanisms of ZCPs in *P. ostreatus*.

Introduction

*Pleurotus ostreatus* is a widely cultivated mushroom in China because of its high nutritional value, uncomplicated cultivation techniques, and broad adaptability (Miles and Chang, 2004; Khan and Tania, 2012). In 2017, the annual yield of *P. ostreatus* in China was 5.46 tons, which was the third largest yield among all edible mushrooms, according to the statistics of the Chinese Edible Fungi Association. Because cultivation of *P. ostreatus* is mainly carried out in traditional greenhouses without environmental control systems, extreme and continuous high temperature always result in “spawn-burning” and easily causes infection by green mold, which ultimately affects quality and yield (Zhang et al., 2016a; Qiu et al., 2017). The mechanisms of the heat stress response in mushrooms have become a hot research topic, with studies focusing on signaling molecules (Zhang et al., 2016b; Liu et al., 2018), thermoprotective components (Wang et al., 2018a; Lei et al., 2019; Liu et al., 2019), antioxidant enzymes (Wang et al., 2017), and morphological characteristics (Song et al., 2014). Furthermore, transcription factors play important roles in heat stress responses. For example, heat shock transcription factors (HSFs) regulate heat tolerance in wheat and *Arabidopsis thaliana* (Jiang et al., 2018; Tian et al., 2019). Dehydration responsive element binding protein (DREB) improves heat resistance in *A. thaliana* and *Nicotiana tabacum* (Sakuma et al., 2006; Arroyo-Herrera et al., 2016). Transcription factors such as those of the NAC (Shahnejat-Bushehri et al., 2012; Fang et al., 2015; Guo et al., 2015), MYB (El-Kereamy et al., 2012; Casaretto et al., 2016), WRKY (Li et al., 2009; Cai et al., 2015; He et al., 2016) and zinc finger (Kim et al., 2015; Koguchi et al., 2017; Agarwal and Khurana, 2018) families are involved in tolerance to high temperature in many eukaryotes.
Fruiting bodies are arguably the most complex multicellular structures that are produced by fungi, and it is important to gain further insight into fruiting body morphogenesis (Nowrousian, 2018). Transcription factors, such as SQUAMOSA promoter-binding protein (SBP), homeodomain transcription factors and NAC, are involved in a variety of developmental processes (Cardon et al., 1999; Ernst et al., 2004; Vonk and Ohm, 2018).

The Zn(II)$_2$Cys$_6$ protein family, which is also called the zinc cluster protein (ZCP) family, is a large transcription factor family found exclusively in fungi. Members of these families bind two zinc atoms with a DNA-binding domain (DBD) consisting of six cysteine residues (Vallee et al., 1991; MacPherson et al., 2006). Previous studies have shown that ZCPs act as multifunctional regulators in many biological processes. Indeed, ZCPs have been shown to participate in primary metabolism, including sugar, amino acid, vitamin, and uracil metabolism (Losson and Lacroute, 1981; Nishimura et al., 1992; Iraqui et al., 1999), as well as in secondary metabolism, such as ergosterol biosynthesis and melanin biosynthesis (Crowley et al., 1998; Tsuji et al., 2000). Moreover, ZCPs have been implicated in fungal development and the response to stresses, such as heat shock, oxidative stress, and high osmotic stress (Akache et al., 2001; Ohm et al., 2011; Schumacher et al., 2018).

The ZCP gene family has been identified in many fungi, and it is well characterized in Saccharomyces cerevisiae (Akache et al., 2001), Candida albicans (Schillig and Morschhäuser, 2013), Aspergillus flavus (Chang and Ehrlich, 2013) and Tolypocladium guangdongense (Zhang et al., 2019). However, similar studies in Agaricales and even basidiomycetes are, to our knowledge, very limited. Genome-wide identification of ZCP genes in different species not only provides useful information regarding the biological functions of this important gene family but also reveals targets for molecular breeding. In this study, ZCP genes in P. ostreatus were identified by a genome-wide method, and classification and domain analysis were performed. Furthermore, protein properties, phylogenetic analyses, motif distributions and gene architectures among members of the PoZCP family were assessed. Finally, the expression patterns of PoZCP genes during different developmental stages and under heat stress at different time points were investigated based on previously obtained RNA-seq data. Additionally, ten heat-responsive and six development-related PoZCPs were selected for validation of their expression levels through real-time quantitative PCR (RT-qPCR) analysis. In conclusion, this study offers useful information about the PoZCP gene family, and some potential family members that might be involved in development and heat stress were identified, providing targets for molecular breeding of P. ostreatus.

Materials & Methods

Identification of zinc cluster genes

The zinc cluster domain (PF00172) was downloaded from the Pfam database (http://pfam.xfam.org/) (El-Gebali et al., 2018). Then, an HMM search was performed using the hmmssearch tools embedded in HMMER3.21 (Wheeler and Eddy, 2013) with default parameters
based on the protein sequences of *P. ostreatus* strain CCMSSC0389 (China Center for Mushroom Spawn Standards and Control, Beijing, China) (Qu et al., 2016). The zinc cluster domain region of each potential positive sequence was extracted and multialigned using the Clustal Omega online tool (Madeira et al., 2019). Sequences without complete six cysteine residues in the zinc cluster domain were regarded as redundant sequences and removed. The gene structures of the remaining sequences were corrected by checking the mapped reads positions in the RNA-seq data using the Integrative Genomics Viewer tool (Robinson et al., 2011). After manual removal of redundant sequences and correction, the remaining sequences were submitted to the NCBI Conserved Domain Database (https://www.ncbi.nlm.nih.gov/cdd) and the PFAM database to determine the presence and integrity of the zinc cluster domain. Transcription factor annotation was performed by the hmmsearch tool (E-value cut-off = 0.01) using 52 transcription factor Pfam models (Table S1).

**Gene structure and protein analyses of the zinc cluster family**

Gene structures were obtained by exon and intron information from the *P. ostreatus* genome. PoZCPs were classified according to the number of amino acid residues between each two cysteine residues in the typical zinc cluster domain. The isoelectric point (pI) and molecular weight (MW) of PoZCPs were estimated using IPC tools (Kozlowski, 2016). The subcellular localization of each PoZCP was predicted by BaCelLo online tools (Pierleoni et al., 2006).

**Sequence alignment, motif analysis and phylogenetic tree construction of the zinc cluster family**

The domain region of the zinc cluster sequences was characterized by the NCBI-CDD and Pfam results and classified by the amino acid counts of variable regions. Multiple sequence alignments of the protein sequences of PoZCPs were carried out using the Clustal Omega online tool (Madeira et al., 2019), and the poorly aligned regions were removed by the alignment utility trimAl (Capella-Gutiérrez et al., 2009). The aligned sequences were used to construct a phylogenetic tree based on the maximum likelihood (ML) method by IQ-TREE using the best fitting model and branch support computed with 1000 UltraFast Bootstrap (Trifinopoulos et al., 2016). Conserved motifs in the PoZCP family were assessed by the MEME online program (Bailey et al., 2009). The gene structure and motif distribution of the PoZCP family were displayed using Tbtools (Chen et al., 2018).

**RNA-sequencing and analysis**

RNA-sequencing (RNA-seq) was conducted to gain insight into the expression patterns of the zinc cluster genes at several developmental stages and heat stress. Fruiting bodies of the dikaryotic strain CCMSSC0389 were grown, and samples, including mycelia, primordia, and fruiting bodies, were collected at different developmental stages as described previously (Wang et al., 2017). The mycelia were inoculated on PDA medium and cultured at 28 °C for 6 days and...
then transferred to 35 °C for 0.5, 1, 2, and 3 h, respectively, for heat stress treatment. Three replicates for each sample were frozen in liquid nitrogen immediately after collection.

Sample collections, RNA extraction and RNA-seq were carried out as previously described (Wang et al., 2018b), except that TruSeq Stranded mRNA LTSample Prep Kit (Illumina, San Diego, CA, USA) was used for sequencing library construction and the Illumina HiSeq 2500 platform was used for sequencing. Transcript expression of the zinc cluster genes was analyzed according to the RNA-seq profiles using fragments per kilobase per million mapped reads (FPKM) values. A heatmap was constructed using Tbtools with log and row scales (Chen et al., 2018).

Validation of gene expression by RT-qPCR

To validate the RNA-seq results, we performed an RT-qPCR analysis of 13 genes that showed different expression patterns and large fold changes during various developmental stages or under heat stress. The experimental procedures were described in previous studies (Wang et al., 2017; Wang et al., 2018b). Beta-actin and beta-tubulin were used as reference genes. The primer sequences are listed in Table S2. The relative expression level of each gene was calculated using the 2$^{-}\Delta\Delta C_{t}$ method (Livak and Schmittgen, 2001).

Results

Identification of ZCP genes in P. ostreatus

After manual correction, a total of 66 unique genes with conserved zinc cluster domains were identified as ZCP genes in P. ostreatus, accounting for 23.82% of all transcription factors. The 66 genes were named PoZCP01 to PoZCP66, and the coding sequences were submitted to GenBank under accession numbers MN652925- MN652990. Fifty-nine (89.39%) family members contain zinc cluster domains in the N-terminus (front 1/3 part of the protein), three (4.55%) family members have zinc cluster domains in the middle (1/3- 2/3 of the protein) of the protein, and four (6.06%) family members possess zinc cluster domains in the C-terminus (last 1/3 part of the protein) (Fig. 1). Fungal-specific transcription factor domains (Fungal_trans, PF04082) are present in 43 (65.15%) family members, and C2H2-type zinc transcription factor domains (PF00096) are also present in three family members. One gene, PoZCP55, encodes three septin domains (PF00735), which are reported to be involved in cytokinesis and septum formation (Longtine et al., 1996), with two zinc cluster domains in the C-terminus.

Classification and characterization of ZCP genes in P. ostreatus

Previous studies in yeast have revealed that in the zinc cluster domain, the numbers of amino acids between C1-C2, C2-C3 and C4-C5 are always 2, 6, and 2, respectively, but that numbers between C3-C4 and C5-C6 vary (Todd and Andrianopoulos, 1997). The PoZCP family was classified into 15 subgroups named types A-O based on the different patterns between C3-C4 and C4-C5 (Table 1), 10 of which belong to the typical zinc cluster motif C-X$_2$-C-X$_6$-C-X$_{5,12}$-C-
However, five other patterns, including C-X$_2$-C-X$_6$-C-X$_5$-C-X$_2$-C-X$_4$-C (type A), C-X$_2$-C-X$_6$-C-X$_6$-C-X$_2$-C-X$_9$ (type G), C-X$_2$-C-X$_6$-C-X$_{13}$-C-X$_2$-C-X$_6$ (type M), C-X$_2$-C-X$_6$-C-X$_{14}$-C-X$_2$-C-X$_6$ (type N) and C-X$_2$-C-X$_6$-C-X$_{15}$-C-X$_2$-C-X$_6$ (type O), were also found in the PoZCP family (Table 1). The motifs of 15 types of zinc cluster domains were analyzed, and the motif logos of the seven major types (more than three members) are shown in Fig. 2. For types A, D, F, I and J, amino acids are conserved at position 5 (arginine) and position 8 (lysine); for types A, B, C and F, amino acids at position 12 (aspartate) and position 16 (proline) are conserved (Fig. 2).

The characteristics, including exon and intron length, protein length, isoelectric point (pI), molecular weight (MW), and predicted subcellular localization are shown in detail in Table S3. The protein length of putative PoZCP genes ranges from 195 to 1386 amino acids, with MWs ranging from 21 to 152 kDa and theoretical pIs ranging from 4.86 to 9.45, revealing significant differences among PoZCPs in terms of their physical and chemical properties. According to the predicted subcellular localization, 62.12% of PoZCPs are located in the nucleus, 31.34% in the cytoplasm and 6.06% in secretory pathways (Table S3).

**Phylogenetic, motif and structure analysis of PoZCPs**

A phylogenetic tree was constructed using amino acid sequences of the 66 PoZCP family members. The family members were divided into six well-conserved groups (I–VI) with significant bootstrap support according to phylogenetic relationships (Fig. 3A). PoZCP55 forms a separate clade, which imply its self-dependent origin. The PoZCP family members belonging to types A, B, C and F (Table 1) clustered respectively, suggesting genes in each class may have expanded from a single gene. Conversely, PoZCP family members of types D, I, and J (Table 1) were found to be distributed across different clades, indicating that these categories may have different ancestors.

To further investigate the diversity of the ZCP in *P. ostreatus*, we analyzed PoZCP motifs using the MEME online server. Ten conserved motifs named motif 1 to motif 10 were identified. Motif 1, which forms the zinc cluster domain, was the most common motif found in all family members. The combinations of different motifs differed among the groups (Fig. 3A). For example, the combination of motif 3, motif 7 and motif 9 occurs in most proteins of group II, the combination of motifs 7-9 appeared in most proteins of group IV, the combination of motif 3, motif 7 and motif 8 existed in all members of group V and motifs 2-6 were found in all members of group VI (Fig. 3B). These results suggest that the function of the PoZCPs of different groups may differ.

We deduced the gene structures of individual PoZCP genes to examine the relationship between their evolutionary origins and structural diversity. The results showed that the exon number of PoZCP genes changed over a wide range (from 2–23); however, each group in Fig. 3A is relatively conserved in gene structure (Fig. 3C). For example, group II consists of genes of approximately 2500 bp with similar architecture and 8-10 exons that encode approximately 600 amino acids. Group III is composed of short genes of approximately 1600 bp that encode
approximately 450 amino acids. Group IV and group VI are rich in exons, numbering approximately 14, which is significantly greater than the average exon number of 9 for all family members. Most of the genes in group VI contain many short exons of fewer than 50 bp.

**Expression patterns of PoZCPs during different developmental stages and under heat stress**

To gain insight into the expression profiles of PoZCP genes during different developmental stages and under heat stress at different time points, RNA-seq data obtained from different developmental stages as well as different heat treatment times were used to analyze expression patterns (Table S4, Table S5). PoZCP14 was not expressed in any expression profile, indicating that it may be a pseudogene.

The expression patterns during different developmental stages were categorized into four clusters according to RNA-seq profiles (Fig. 4A). Genes in cluster I and cluster IV were downregulated and upregulated, respectively, from the mycelial to fruiting body development stages (Fig. 4A), which suggests that these genes may act as negative or positive regulators during development. Genes in cluster II and cluster III were only upregulated in the fruiting body stage and primordia stage, respectively (Fig. 4A), therefore, these genes might be involved in the stage-specific process.

The expression patterns under heat stress at different time points were also categorized into four groups (Fig. 4B). Clusters I and IV were downregulated after heat stress treatment at different time points, which indicates their negative regulatory ability under long-term (2-3 h) and short-term (0.5-1 h) heat stress, respectively (Fig. 4B). Most genes in clusters II and III were upregulated after heat stress treatment at different time points, revealing their positive regulatory ability under short-term (0.5-1 h) and long-term (2-3 h) heat stress, respectively (Fig. 4B).

To verify the gene expression results for the identified PoZCPs, thirteen genes showing distinct variation at various developmental stages or heat treatment in RNA-seq analysis were selected for detection of their expression levels via RT-qPCR. All tested genes showed the same tendency as the transcript profiles (Fig. 5). PoZCP10, PoZCP26, PoZCP31, and PoZCP40 were dramatically upregulated (20-90-fold) in the fruiting body stage, suggesting that they are stage-specific genes. PoZCP15, PoZCP26, and PoZCP31 were markedly upregulated (5-140-fold) under 0.5-1 h heat stress (Fig. 5), which indicates their important roles in the response to short-term heat stress. PoZCP26 and PoZCP31 were upregulated during both developmental stages and under heat stress treatment and thus may have multiple functions in development and the heat stress response.

**Discussion**

Transcription factors play important roles in growth and development as well as in responses to various abiotic stresses by regulating downstream genes (Wu and Gallagher, 2012). As the genomes of an increasing number of organisms have been sequenced, these sequences have supplied valuable support for the discovery of potential transcription factors that may contribute
to important agronomic traits (Li et al., 2014). The ZCP family, which is the largest transcription factor family in many ascomycetes and basidiomycetes (Todd et al., 2014), has been previously found specifically in fungi. In this study, we identified 66 members of the ZCP family in *P. ostreatus*, which accounted for 23.82% of all transcription factors. In a previous study, 55 members of this family were identified in *S. cerevisiae* (Joshua and Höfken, 2017), and genome-wide analysis showed that 82 putative ZCPs exist in *C. albicans* (Schillig and Morschhäuser, 2013). For the pathogenic fungus *A. flavus* and medicinal fungus *T. guangdongense*, up to 304 and 139 family members, respectively, have been identified (Chang and Ehrlich, 2013; Zhang et al., 2019). The amount and proportion of PoZCPs is similar to the average for basidiomycetes, at 58 (20.49%), based on 31 basidiomycete genomes, but far below the average from ascomycetes, at 121 (33.43%), based on 77 ascomycete genomes. This finding suggests that this family developed after the split from ascomycetes, and that the underlying molecular mechanisms may be different (Todd et al., 2014). Most zinc cluster motifs are located in the N-terminus of the protein. However, four PoZCPs, including PoZCP14, PoZCP16, PoZCP22 and PoZCP55, are C-terminal zinc cluster proteins. At least two characterized C-terminal zinc cluster proteins also exist in *S. cerevisiae* and *C. albicans* (Whiteway et al., 1992).

Zinc cluster motifs display a wide spacing pattern of C-X$_2$-C-X$_6$-C-X$_{12}$-C-X$_2$-C-X$_6$-C, similar to that observed in yeasts. As shown in this study, 15 types of zinc cluster motifs, including five from different types against yeast, are found in PoZCPs. These findings are similar to those in *Aspergillus* species and *T. guangdongense*, which also exhibits several specific patterns of zinc cluster motifs (Chang and Ehrlich, 2013; Zhang et al., 2019). In addition to the six cysteine residues that coordinate the zinc ions, some amino acids, including arginine and lysine between C2 and C3 and aspartic acid and proline between C3 and C4, are also conserved in different types of zinc cluster patterns (Fig. 2). These residues play an important role in the formation of functional structures in *S. cerevisiae*. For example, the proline located in the loop between the two substructures confers the flexibility to the loop, and the lysine and arginine between C3 and C4 are involved in homodimerization (Marmorstein et al., 1992).

Diversity in physical, chemical and structural characteristics is very common among transcription factors. For example, the CCCH zinc finger family presents a wide variation of gene structures in *Brassica rapa* (Pi et al., 2018), and the physicochemical parameters of the ZCP family in *T. guangdongense* are also highly divergent (Zhang et al., 2019). The characteristics of PoZCPs, including protein length, pI, MW, exon and intron length, and number, show variation over a wide range, suggesting high diversity and multiple functions among ZCP genes. Regarding subcellular localization, the proportions of PoZCPs located in the nucleus, cytoplasm, and secretory pathways were predicted to be 62.12%, 31.34%, and 6.06%, respectively. This proportion is very similar to the ZCPs in *T. guangdongense* with 62.6% in the nucleus, 23% in the cytoplasm and 5% in the secretory pathway (Zhang et al., 2019). The ZCPs localized in the cytoplasm may be transported to the nucleus when they participate in transcriptional regulation (MacPherson et al., 2006). The phylogenetic relationship of the PoZCP family can be divided into six groups based on the phylogenetic tree. Some types of PoZCPs
were distributed in the same clade, indicating that they have originated from a single gene. In contrast, some types were distributed among all branches of the tree, which suggests that they may have distinct origins. The motifs in PoZCPs are very abundant, but only Motif 1 correlated with the Pfam domain. The potential functions of the other motifs are worthy of further investigation.

In previous studies, homologs of the zinc cluster transcription factor PRO1 were found to be required for fruiting body development in a number of ascomycete species, such as *A. nidulans* (Vienken and Fischer, 2006), *Cryphonectria parasitica* (Sun et al., 2009) and *Sordaria macrospora* (Steffens et al., 2016). In basidiomycetes, the transcription factors Fst3 and Fst4 of the ZCP family were shown to inhibit and induce mushroom development, respectively, in *Schizophyllum commune* (Ohm et al., 2011). We found that the expression patterns during different developmental stages differed among PoZCP family members. Many developmental stage-specific genes were up- or downregulated at specific developmental stages (Fig. 3A). For example, *PoZCP10*, *PoZCP26*, and *PoZCP31* were upregulated (20-90-fold) at the fruiting body stage, and *PoZCP59* was upregulated (30-fold) in the primordia stage compared to the mycelia stage (Fig. 4A). The gene function and downstream genes that are regulated during fruiting body formation should be further researched.

Many ZCPs have been presumed to be involved in the response to many environmental factors in different species. For example, Stb5 is a key player in the resistance to oxidative stress, and HAL9 is an important factor involved in salt tolerance in *S. cerevisiae* (Larochelle et al., 2006). Moreover, many members of the ZCP family in *Cordyceps militaris* (Yang et al., 2016) and *T. guangdongense* (Zhang et al., 2019) have an active role in the light response. The zinc cluster protein YFL052W regulates the thermal sensitivity of yeast on rich medium (Akache et al., 2001). The expression profiles of PoZCP genes under heat stress at different time points showed that some family members actively participate in the heat stress response. For example, *PoZCP15*, *PoZCP26*, and *PoZCP31* were dramatically upregulated (5-140-fold) under heat stress for 0.5-1 h, though their roles in the response to heat stress remain to be further investigated.

**Conclusions**

This study constitutes an exhaustive identification, characterization and expression analysis of the zinc cluster type transcription factor in the widely cultivated mushroom *P. ostreatus*. The findings provide useful information about the PoZCP gene family that can be utilized in other basidiomycetes. Some genes were identified as candidates for further study because of their potential functions in fruiting body formation and the heat stress response, which may provide targets for molecular breeding of *P. ostreatus*.

**References**
Agarwal, P. and Khurana, P., 2018. Characterization of a novel zinc finger transcription factor (TaZnF) from wheat conferring heat stress tolerance in Arabidopsis. Cell Stress and Chaperones 23, 253-267.

Akache, B., Wu, K. and Turcotte, B., 2001. Phenotypic analysis of genes encoding yeast zinc cluster proteins. Nucleic Acids Research 29, 2181-2190.

Arroyo-Herrera, A., Figueroa-Yáñez, L., Castano, E., Santamaría, J., Pereira-Santana, A., Espadas-Alcocer, J., Sánchez-Teyer, F., Espadas-Gil, F., Alcaraz, L.D. and López-Gómez, R., 2016. A novel Dreb2-type gene from Carica papaya confers tolerance under abiotic stress. Plant Cell Tiss. Organ. Cult. 125, 119-133.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W. and Noble, W.S., 2009. MEME Suite: tools for motif discovery and searching. Nucleic Acids Research 37, W202-W208.

Cai, H.Y., Yang, S., Yan, Y., Xiao, Z.L., Cheng, J.B., Wu, J., Qiu, A.L., Lai, Y., Mou, S.L. and Guan, D.Y., 2015. CaWRKY6 transcriptionally activates CaWRKY40, regulates Ralstonia solanacearum resistance, and confers high-temperature and high-humidity tolerance in pepper. Journal of Experimental Botany 66, 3163-3174.

Capella-Gutiérrez, S., Silla-Martínez, J.M. and Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972-1973.

Cardon, G., Höhmann, S., Klein, J., Nettesheim, K., Saedler, H. and Huijser, P., 1999. Molecular characterisation of the Arabidopsis SBP-box genes. Gene 237, 91-104.

Casaretto, J.A., El-kereamy, A., Zeng, B., Stiegelmeyer, S.M., Chen, X., Bi, Y.M. and Rothstein, S.J., 2016. Expression of OsMYB55 in maize activates stress-responsive genes and enhances heat and drought tolerance. BMC Genomics 17, 312.

Chang, P. and Ehrlich, K.C., 2013. Genome-wide analysis of the Zn(II)$_2$Cys$_6$ zinc cluster-encoding gene family in Aspergillus flavus. Applied Microbiology and Biotechnology 97, 4289-4300.

Chen, C.J., Xia, R., Chen, H. and He, Y., 2018. TBtools, a Toolkit for Biologists integrating various biological data handling tools with a user-friendly interface. bioRxiv, 289660.

Crowley, J.H., Leak, F.W., Shianna, K.V., Tove, S. and Parks, L.W., 1998. A mutation in a purported regulatory gene affects control of sterol uptake in Saccharomyces cerevisiae. Journal of Bacteriology 180, 4177-4183.

El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., Qureshi, M., Richardson, L.J., Salazar, G.A. and Smart, A., 2018. The Pfam protein families database in 2019. Nucleic Acids Research 47, D427-D432.

El-Kereamy, A., Bi, Y.-M., Ranathunge, K., Beatty, P.H., Good, A.G. and Rothstein, S.J., 2012. The rice R2R3-MYB transcription factor OsMYB55 is involved in the tolerance to high temperature and modulates amino acid metabolism. PLoS One 7, e52030.

Ernst, H.A., Nina Olsen, A., Skriver, K., Larsen, S. and Lo Leggio, L., 2004. Structure of the conserved domain of ANAC, a member of the NAC family of transcription factors. EMBO Rep. 5, 297-303.
Fang, Y.J., Liao, K.F., Du, H., Xu, Y., Song, H.Z., Li, X.H. and Xiong, L.Z., 2015. A stress-responsive NAC transcription factor SNAC3 confers heat and drought tolerance through modulation of reactive oxygen species in rice. Journal of Experimental Botany 66, 6803-6817.

Guo, W.W., Zhang, J.X., Zhang, N., Xin, M.M., Peng, H.R., Hu, Z.R., Ni, Z.F. and Du, J.K., 2015. The wheat NAC transcription factor TaNAC2L is regulated at the transcriptional and post-translational levels and promotes heat stress tolerance in transgenic Arabidopsis. PLoS One 10, e0135667.

He, G.H., Xu, J.Y., Wang, Y.X., Liu, J.M., Li, P.S., Chen, M., Ma, Y.Z. and Xu, Z.S., 2016. Drought-responsive WRKY transcription factor genes TaWRKY1 and TaWRKY33 from wheat confer drought and/or heat resistance in Arabidopsis. BMC Plant Biology 16, 116.

Iraqui, I., Vissers, S., André, B. and Urrestarazu, A., 1999. Transcriptional induction by aromatic amino acids in Saccharomyces cerevisiae. Molecular and Cellular Biology 19, 3360-3371.

Jiang, Y.L., Zheng, Q.Q., Chen, L., Liang, Y.N. and Wu, J.D., 2018. Ectopic overexpression of maize heat shock transcription factor gene ZmHsf04 confers increased thermo and salt-stress tolerance in transgenic Arabidopsis. Acta Physiol. Plant. 40, 9.

Joshua, I.M. and Höfken, T., 2017. From lipid homeostasis to differentiation: old and new functions of the zinc cluster proteins Ecm22, Upc2, Sut1 and Sut2. International Journal of Molecular Sciences 18, 772.

Khan, M.A. and Tania, M., 2012. Nutritional and medicinal importance of Pleurotus mushrooms: an overview. Food Reviews International 28, 313-329.

Kim, G.D., Cho, Y.H. and Yoo, S.D., 2015. Regulatory functions of evolutionarily conserved AN1/A20-like Zinc finger family proteins in Arabidopsis stress responses under high temperature. Biochemical and Biophysical Research Communications 457, 213-220.

Koguchi, M., Yamasaki, K., Hirano, T. and Sato, M.H., 2017. Vascular plant one-zinc-finger protein 2 is localized both to the nucleus and stress granules under heat stress in Arabidopsis. Plant Signal. and Behav. 12, e1295907.

Kozlowski, L.P., 2016. IPC–isoelectric point calculator. Biology Direct 11, 55.

Larochelle, M., Drouin, S., Robert, F. and Turcotte, B., 2006. Oxidative stress-activated zinc cluster protein Stb5 has dual activator/repressor functions required for pentose phosphate pathway regulation and NADPH production. Molecular and Cellular Biology 26, 6690-6701.

Lei, M., Wu, X.L., Huang, C.Y., Qiu, Z.H., Wang, L.N., Zhang, R.Y. and Zhang, J.X., 2019. Trehalose induced by reactive oxygen species relieved the radial growth defects of Pleurotus ostreatus under heat stress. Applied Microbiology and Biotechnology 103, 5379-5390.

Li, S.J., Fu, Q.T., Huang, W.D. and Yu, D.Q., 2009. Functional analysis of an Arabidopsis transcription factor WRKY25 in heat stress. Plant Cell Reports 28, 683-693.

Li, Y.H., Zhou, G.Y., Ma, J.X., Jiang, W.K., Jin, L.G., Zhang, Z.H., Guo, Y., Zhang, J.B., Sui, Y. and Zheng, L.T., 2014. De novo assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits. Nature Biotechnology 32, 1045-1052.
Liu, R., Shi, L., Zhu, T., Yang, T., Ren, A., Zhu, J. and Zhao, M.W., 2018. Cross talk between nitric oxide and calcium-calmodulin regulates ganoderic acid biosynthesis in *Ganoderma lucidum* under heat stress. Appl. Environ. Microbiol. 84, e00043-18.

Liu, X.M., Wu, X.L., Gao, W., Qu, J.B., Qiang, C., Huang, C.Y. and Zhang, J.X., 2019. Protective roles of trehalose in *Pleurotus pulmonarius* during heat stress response. J. Integr. Agr. 18, 428-437.

Livak, K.J. and Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔCT method. Methods 25, 402-408.

Longtine, M.S., DeMarini, D.J., Valencik, M.L., Al-Awar, O.S., Fares, H., De Virgilio, C. and Pringle, J.R., 1996. The septins: roles in cytokinesis and other processes. Current Opinion in Cell Biology 8, 106-119.

Losson, R. and Lacroute, F., 1981. Cloning of a eukaryotic regulatory gene. Molecular and General Genetics 184, 394-399.

MacPherson, S., Larochelle, M. and Turcotte, B., 2006. A fungal family of transcriptional regulators: the zinc cluster proteins. Microbiol. Mol. Biol. Rev. 70, 583-604.

Madeira, F., Park, Y.m., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A.R.N., Potter, S.C., Finn, R.D. and Lopez, R., 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Research 47, W636-W641.

Marmorstein, R., Carey, M., Ptashne, M. and Harrison, S.C., 1992. DNA recognition by GAL4: structure of a protein-DNA complex. Nature 356, 408.

Miles, P.G. and Chang, S.T., 2004. Mushrooms: cultivation, nutritional value, medicinal effect, and environmental impact, CRC press, Boca Raton, Florida, U.S.A.

Nishimura, H., Kawasaki, Y., Kaneko, Y., Nosaka, K. and Iwashima, A., 1992. Cloning and characteristics of a positive regulatory gene, THI2 (PHO6), of thiamin biosynthesis in *Saccharomyces cerevisiae*. FEBS Letters 297, 155-158.

Nowrousian, M., 2018. Genomics and transcriptomics to study fruiting body development: An update. Fungal Biology Reviews 32, 231-235.

Ohm, R.A., de Jong, J.F., de Bekker, C., Wösten, H.A. and Lugones, L.G., 2011. Transcription factor genes of *Schizophyllum commune* involved in regulation of mushroom formation. Molecular Microbiology 81, 1433-1445.

Pi, B.Y., He, X.H., Ruan, Y., Jang, J. and Huang, Y., 2018. Genome-wide analysis and stress-responsive expression of CCCH zinc finger family genes in *Brassica rapa*. BMC Plant Biology 18, 373.

Pierleoni, A., Martelli, P.L., Fariselli, P. and Casadio, R., 2006. BaCelLo: a balanced subcellular localization predictor. Bioinformatics 22, e408-e416.

Qiu, Z.H., Wu, X.L., Zhang, J.X. and Huang, C.Y., 2017. High temperature enhances the ability of *Trichoderma asperellum* to infect *Pleurotus ostreatus* mycelia. PLoS One 12, e0187055.

Qu, J.B., Zhao, M.R., Hsiang, T., Feng, X.X., Zhang, J.X. and Huang, C.Y., 2016. Identification and characterization of small noncoding RNAs in genome sequences of the edible fungus *Pleurotus ostreatus*. BioMed Res. Int. 2016, 9.
Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G. and Mesirov, J.P., 2011. Integrative genomics viewer. Nature Biotechnology 29, 24-26.

Sakuma, Y., Maruyama, K., Qin, F., Osakabe, Y., Shinozaki, K. and Yamaguchi-Shinozaki, K., 2006. Dual function of an Arabidopsis transcription factor DREB2A in water-stress-responsive and heat-stress-responsive gene expression. Proceedings of the National Academy of Sciences, USA 103, 18822-18827.

Schillig, R. and Morschhäuser, J., 2013. Analysis of a fungus - specific transcription factor family, the Candida albicans zinc cluster proteins, by artificial activation. Molecular Microbiology 89, 1003-1017.

Shahnejat-Bushehri, S., Mueller-Roeber, B. and Balazadeh, S., 2012. Arabidopsis NAC transcription factor JUNGBRUNNEN1 affects thermomemory-associated genes and enhances heat stress tolerance in primed and unprimed conditions. Plant Signal. and Behav. 7, 1518-1521.

Song, C., Chen, Q., Wu, X.L., Zhang, J.X. and Huang, C.Y., 2014. Heat stress induces apoptotic-like cell death in two Pleurotus species. Current Microbiology 69, 611-616.

Steffens, E.K., Becker, K., Krevet, S., Teichert, I. and Kück, U., 2016. Transcription factor PRO1 targets genes encoding conserved components of fungal developmental signaling pathways. Molecular Microbiology 102, 792-809.

Sun, Q.H., Choi, G.H. and Nuss, D.L., 2009. Hypovirus-responsive transcription factor gene pro1 of the chestnut blight fungus Cryphonectria parasitica is required for female fertility, asexual spore development, and stable maintenance of hypovirus infection. Eukaryotic Cell 8, 262-270.

Tian, X.J., Wang, F., Zhao, Y., Lan, T.Y., Yu, K., Zhang, L.Y., Qin, Z., Hu, Z.R., Yao, Y.Y., Ni, Z.F., Sun, Q.X., Rossi, V., Peng, H.R. and Xin, M.M., 2019. Heat shock transcription factor A1b regulates heat tolerance in wheat and Arabidopsis through OPR3 and jasmonate signalling pathway. Plant Biotechnol. J. 2019, 1-3.

Todd, R.B. and Andrianopoulos, A., 1997. Evolution of a fungal regulatory gene family: the Zn (II)2Cys6 binuclear cluster DNA binding motif. Fungal Genetics and Biology 21, 388-405.

Trifinopoulos, J., Nguyen, L.T., von Haeseler, A. and Minh, B.Q., 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Research 44, W232-W235.
*lagenarium* and Pig1p of *Magnaporthe grisea*, contain Cys2His2 zinc finger and Zn(II)2Cys6 binuclear cluster DNA-binding motifs and regulate transcription of melanin biosynthesis genes in a developmentally specific manner. Molecular Microbiology 38, 940-954.

Vallee, B.L., Coleman, J.E. and Auld, D.S., 1991. Zinc fingers, zinc clusters, and zinc twists in DNA-binding protein domains. Proceedings of the National Academy of Sciences, USA 88, 999-1003.

Vienken, K. and Fischer, R., 2006. The Zn(II)$_2$Cys$_6$ putative transcription factor NosA controls fruiting body formation in *Aspergillus nidulans*. Molecular Microbiology 61, 544-554.

Vonk, P.J. and Ohm, R.A., 2018. The role of homeodomain transcription factors in fungal development. Fungal Biology Reviews 32, 219-230.

Wang, G.Z., Zhou, S.S., Luo, Y., Ma, C.J., Gong, Y.H., Zhou, Y., Gao, S.S., Huang, Z.C., Yan, L.L. and Hu, Y., 2018a. The heat shock protein 40 LeDnaJ regulates stress resistance and indole-3-acetic acid biosynthesis in *Lentinula edodes*. Fungal Genetics and Biology 118, 37-44.

Wang, L.N., Gao, W., Wu, X.L., Zhao, M.R., Qu, J.B., Huang, C.Y. and Zhang, J.X., 2018b. Genome-wide characterization and expression analyses of *Pleurotus ostreatus* MYB transcription factors during developmental stages and under heat stress based on *de novo* sequenced genome. International Journal of Molecular Sciences 19, 2052.

Wang, L.N., Wu, X.L., Gao, W., Zhao, M.R., Zhang, J.X. and Huang, C.Y., 2017. Differential expression patterns of *Pleurotus ostreatus* catalase genes during developmental stages and under heat stress. Genes 8, 335.

Wheeler, T.J. and Eddy, S.R., 2013. nhmmer: DNA homology search with profile HMMs. Bioinformatics 29, 2487-2489.

Whiteway, M., Dignard, D. and Thomas, D.Y., 1992. Dominant negative selection of heterologous genes: isolation of *Candida albicans* genes that interfere with *Saccharomyces cerevisiae* mating factor-induced cell cycle arrest. Proceedings of the National Academy of Sciences, USA 89, 9410-9414.

Wu, S. and Gallagher, K.L., 2012. Transcription factors on the move. Current Opinion in Plant Biology 15, 645-651.

Yang, T., Guo, M.M., Yang, H.J., Guo, S.P. and Dong, C.H., 2016. The blue-light receptor CmWC-1 mediates fruit body development and secondary metabolism in *Cordyceps militaris*. Applied Microbiology and Biotechnology 100, 743-755.

Zhang, C.H., Huang, H., Deng, W.Q. and Li, T.H., 2019. Genome-wide analysis of the Zn(II)$_2$Cys$_6$ zinc cluster-encoding gene family in *Tolypocladium guangdongense* and its light-Induced expression. Genes 10, 179.

Zhang, R.Y., Hu, D.D., Zhang, Y.Y., Goodwin, P.H., Huang, C.Y., Chen, Q., Gao, W., Wu, X.L., Zou, Y.J. and Qu, J.B., 2016a. Anoxia and anaerobic respiration are involved in “spawn-burning” syndrome for edible mushroom *Pleurotus eryngii* grown at high temperatures. Sci. Hortic (Amsterdam). 199, 75-80.
Zhang, X., Ren, A., Li, M.J., Cao, P.F., Chen, T.X., Zhang, G., Shi, L., Jiang, A.L. and Zhao, M.W., 2016b. Heat stress modulates mycelium growth, heat shock protein expression, ganoderic acid biosynthesis, and hyphal branching of *Ganoderma lucidum* via cytosolic Ca$^{2+}$. Appl. Environ. Microbiol. 82, 4112–4125.
Figure 1

The distribution of conserved domains in 66 members of the PoZCP family.

The domain organization of the zinc cluster proteins was determined from the Pfam and NCBI conserved domain databases. The X-axis represents the amino acid position from the N-terminus to the C-terminus. The gray rectangles represent amino acid regions with no identified domain, and the rectangles in different colors represent different domains, as shown in the upper right corner.
Figure 2

The motif logos of major types of zinc cluster domains in the PoZCP family.

The capital letters on the left side of the logos represent the PoZCP family types shown in Table 1. The X-axis represents the positions beginning with the first cysteine of the zinc cluster domain, and the Y-axis represents the number of bits for each amino acid. The amino acids are shown with one-letter abbreviations.
Figure 3

Phylogenetic tree, motif composition and gene structures of the PoZCP family.

(A) Phylogenetic tree obtained using the maximum likelihood (ML) method by IQ-TREE with best fitting model branch support computed with 1000 UltraFast Bootstrap. Different colors of nodes represent different ranges of bootstrap support values. The classification of different types of PoZCPs is shown with polygons of different shapes and colors. (B) Motif composition of the PoZCP proteins, obtained by the MEME program. The gray line represents the length of the proteins, and the rectangles in different colors represent different motifs from motif 1 to motif 10. (C) Gene structures of PoZCP genes. The exons are represented by red rectangles, and the black lines connecting two exons represent introns.
Figure 4

Expression levels of PoZCPs under different conditions using RNA-seq analysis

(A) Expression levels of PoZCPs during different developmental stages. M = mycelia (control); P = primordia; F = fruiting bodies. The heatmap was constructed based on RNA-seq analysis of PoZCP genes. The color scale at the top right corner represents high to low expression levels of each gene from red to blue. Different colors of clades represent different expression patterns. I: Downregulated during developmental stages. II: upregulated at the fruiting body stage. III: upregulated at the primordia stage. IV: upregulated during developmental stages. (B) Expression levels of PoZCPs under heat stress at different time points. The heatmap was constructed based on RNA-seq analysis of PoZCP genes. Mycelia without heat treatment (0 h) were analyzed as a control. The color scale at the top right corner represents high to low expression levels of each gene from red to green. Different colors of clades represent different expression patterns. I: Downregulated during heat stress. II: upregulated during 0.5-1 h of heat treatment. III: upregulated during 2-3 h of heat treatment. IV: upregulated during 0.5-1 h and downregulated during 2-3 h of heat treatment.
Manuscript to be reviewed
Figure 5

Expression analysis of 13 PoZCP genes using RT-qPCR.

(A-I) The expression levels of six PoZCP genes during different developmental stages. M = mycelia (control); P = primordia; F = fruiting bodies. (J-P) The expression levels of ten PoZCP genes under heat stress. Mycelia without heat treatment (0 h) were analyzed as a control. The data are representative of three independent experiments. The error bars represent the standard deviations of three replicates. Statistical analysis was performed using the t test. * represents $p < 0.05$ and ** represents $p < 0.01$. 
Table 1 (on next page)

Classification of PoZCP family members.

The classification of different types of PoZCP family members was based on the variable regions of different zinc cluster domains. The variable regions are shown as underlined characters.
**Table 1.** Classification of PoZCP family members.

| Type | Pattern | Counts |
|------|---------|--------|
| A    | C-X2-C-X6-C-X5-C-X2-C-X4-C | 8      |
| B    | C-X2-C-X6-C-X5-C-X2-C-X6-C | 7      |
| C    | C-X2-C-X6-C-X5-C-X2-C-X6-C | 10     |
| D    | C-X2-C-X6-C-X6-C-X2-C-X6-C | 8      |
| E    | C-X2-C-X6-C-X6-C-X2-C-X7-C | 1      |
| F    | C-X2-C-X6-C-X6-C-X2-C-X6-C | 6      |
| G    | C-X2-C-X6-C-X6-C-X2-C-X6-C | 2      |
| H    | C-X2-C-X6-C-X7-C-X2-C-X6-C | 1      |
| I    | C-X2-C-X6-C-X8-C-X2-C-X6-C | 8      |
| J    | C-X2-C-X6-C-X9-C-X2-C-X6-C | 7      |
| K    | C-X2-C-X6-C-X10-C-X2-C-X6-C | 1      |
| L    | C-X2-C-X6-C-X11-C-X2-C-X6-C | 3      |
| M    | C-X2-C-X6-C-X12-C-X2-C-X6-C | 1      |
| N    | C-X2-C-X6-C-X14-C-X2-C-X6-C | 2      |
| O    | C-X2-C-X6-C-X15-C-X2-C-X6-C | 1      |

The classification of different types of PoZCP family members was based on the variable regions of different zinc cluster domains. The variable regions are shown as underlined characters.