Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Genome-wide association studies disclosed LePC4 gene as a negative regulator of protein content in shiitake mushroom, Lentinula edodes

Chenjie Xu
Huazhong Agricultural University
Wei Zheng
Huazhong Agricultural University
Tianji Huang
Suizhou Vocational & Technical College
Jian Zhang
Suizhou Vocational & Technical College
Yang Xiao
xiaoyang@mail.hzau.edu.cn
Huazhong Agricultural University

Research Article

Keywords: Lentinula edodes, crude protein content, protein fractions, genome-wide association studies, comparative transcriptome

Posted Date: June 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10053066/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at World Journal of Microbiology and Biotechnology on August 4th, 2026. See the published version at https://doi.org/10.1007/s11274-026-05169-5.

Abstract

Lentinula edodes, the most widely cultivated edible mushroom worldwide, is of considerable significance for both nutritional health and commercial applications owing to its high crude protein content (CPC). However, the genetic mechanisms underlying the regulation of CPC in edible mushrooms remain poorly understood. In this study, genome-wide association studies (GWAS) identified twenty single-nucleotide polymorphism (SNP) loci significantly associated with CPC in *L. edodes*, implicating 31 candidate genes. Integrating GWAS results with comparative transcriptomic analysis of two *L. edodes* strains exhibiting markedly divergent CPC levels, we prioritized *LePC4* as a key candidate gene. Structural and functional domain analyses indicated that *LePC4* encodes a protein containing apoptosis-related domains. Subsequent functional characterization via both overexpression and RNA interference (RNAi) revealed that interference of *LePC4* significantly increased CPC in transformed strains by 13.06% − 17.11%, accompanied by elevated relative abundances of albumin (7.41% − 38.73%) and globulin (13.78% − 70.77%). Notably, *LePC4* modulation did not adversely affect mycelial growth or other major nutritional growth parameters. Our findings would help elucidate the genetic mechanisms underlying protein content variation in *L. edodes* and provide scientific guidance for the breeding of high-protein strains in it.

1. Introduction

Proteins, as essential macronutrients required for sustaining vital physiological functions, are increasingly subject to dual pressures arising from rapid global population growth and mounting environmental sustainability constraints (Li et al., 2025). Against this backdrop, apart from relying on the protein resources provided by traditional husbandry, the exploration and development of diversified alternative protein sources has become a prominent topic (Sexton et al., 2019; Sobczak et al., 2023). Edible mushrooms enriched with comprehensive amino acids and proteins, which attaining the standards of high-quality protein in terms of nutritional indicators (Bach et al., 2017). Moreover, researches showed that edible mushroom protein equipped physiological functions such as immune regulation and chronic disease prevention (Sousa et al., 2023). These nutritional characteristics and functional advantages suggest that the development and utilization of edible mushroom protein resources hold promising prospects (González et al., 2020).

Lentinula edodes, known as shiitake or xianggu, was widely cultivated in various countries (Roszczyk et al., 2022). The protein content of *L. edodes* may account for up to 40% of its dry weight, and the species is rich in 18 amino acids, including all nine essential amino acids, with a compositional profile that closely matches human nutritional requirements (Turło et al., 2008; Vetter, 1995). Moreover, *L. edodes* harbors a diverse array of bioactive proteins including lectins and antimicrobial or immunomodulatory peptides, which exhibit functional properties with functional protein resource. (Liu et al., 2025b; Zhou et al., 2019). Previous studies have demonstrated that proteins extracted from *L. edodes* exhibit favorable physicochemical and functional properties (Bing et al., 2023). Incorporating an appropriate proportion of this protein can significantly enhance the taste, flavor, and quality of food products (Uthayakumaran et

al., 1999). Despite these promising prospects for applications in the food processing and pharmaceutical fields, foundational research on *L. edodes* proteins remains limited.

Elucidating the molecular mechanisms governing amino acid metabolism is essential for rationally enhancing protein-related agronomic and nutritional traits in edible mushrooms. However, genes specifically associated with crude protein content (CPC) have rarely been characterized. For instance, the transcription factor GCN4 contributes to intracellular amino acid homeostasis in *Ganoderma lucidum* by activating genes involved in amino acid biosynthesis and transport (Lian et al., 2023). In *Flammulina filiformis*, amino acid transporter genes (*AAT3* and *AAT4*), the *Saccharopine* dehydrogenase gene (*Fvsdh*), and the AMP-activated protein kinase gene (*AMPK*) have been implicated in regulating lysine metabolism (Fan et al., 2023; Liu et al., 2019; Tian et al., 2022). Nevertheless, direct genetic dissection of CPC variation in *L. edodes* remains scarce, which hampers targeted breeding and the development of high-protein strains.

Genome-wide association studies (GWAS) and comparative transcriptomics are important methods for analyzing the genetic basis of important traits in edible mushrooms. GWAS has been successfully applied to analyze the genetic architecture of important agronomic traits of edible mushroom, such as fruiting body morphological traits and stress resistance traits (Liu et al., 2024; Yu et al., 2022). Integration of GWAS with population genomic analyses has proven instrumental in deciphering domestication histories and pinpointing candidate genes underlying key agronomic traits across major cultivated mushroom species. Illustrative examples include cap color and stipe length in *F. filiformis* (Liu et al., 2025a) and fruiting body-related traits in *L. edodes* (Wang et al., 2024; Zhang et al., 2022), which provided a basic theoretical framework for the development of molecular markers for breeding. Development of trait-associated molecular markers is essential for the high-throughput, genotype guided selection and precise identification of elite germplasm in edible mushroom breeding programs.

In this study, CPC was quantified across a diverse panel of 133 *L. edodes* strains. Integrating GWAS with comparative transcriptomic profiling of high- and low-CPC strains, we identified and prioritized candidate genes functionally implicated in CPC regulation. Subsequent functional validation through overexpression and RNA interference (RNAi) confirmed the regulatory role of a key gene (*LePC4*) in modulating protein accumulation. Collectively, our findings provide mechanistic insights into the genetic basis of natural CPC variation in *L. edodes* and establish a foundation for marker-assisted selection of high-protein strains.

2. Materials and Methods

2.1 Determination of crude protein contents in *L. edodes*

A total of 133 *L. edodes* strains were cultivated at Hubei Changjiu Edible Fungi Co., Ltd. All the strains are preserved in the Institute of Applied Mycology, Huazhong Agricultural University. Cultivation experiments employed a randomized complete block design, with the CPC serving as the primary quantitative trait.

The cultivation experiments were carried out followed by Gong et al. (2014). For each strain, four independent biological replicates were established; from each replicate, 3–5 mature fruiting bodies were harvested for CPC analysis. CPC quantification was conducted by Qingdao STD Standard Testing Co., Ltd. using the AOAC-approved Kjeldahl method (Kelrich, 1990). Protein content was calculated from total nitrogen using a nitrogen-to-protein conversion factor of 6.25, following the USDA National Nutrient Database.

2.2 Genome-wide association studies

The re-sequencing data of the 133 *L. edodes* strains was conducted in our previous study (Zhang et al., 2022), with accession numbers of PRJNA320211 and PRJNA535345 in the NCBI Sequenced Read Archive. The re-sequencing data of the 133 *L. edodes* strains were used for SNP calling based on the reference genome of *L. edodes* strain Wpm-1 (Shen et al., 2024). 1,099,843 SNPs were filtered by bcftools (v1.22) with minor allele frequency (MAF) ≥ 0.05 for subsequently analysis (Danecek et al., 2021). GWAS was performed using a linear mixed model implemented in FaST-LMM, integrating the CPC trait with the filtered SNPs (Lippert et al., 2011). Target SNP were filtered with the threshold 1×10^{-5}. R package 'CMplot' in R version 4.3.2 were used to drawn the Manhattan and Quantile-Quantile (QQ) plots (https://CRAN.R-project.org/package=CMplot). Finally, genes within ± 2 kb region of the significantly associated SNP were selected as the candidate genes (Zhang et al., 2022).

2.3 Analysis of differential expression gene in *L. edodes*

According to the CPC of *L. edodes* fruiting bodies, strain YS19 with high CPC (29.083 g/100g) and strain YS66 with low CPC (17.237 g/100g) were selected for RNA sequencing. To investigate gene expression differences across developmental stages, both mycelium and fruiting body samples of these two strains were subjected to RNA sequencing. Mycelial samples were inoculated and collected in sawdust medium (77% hardwood powder, 20% wheat bran, 1% gypsum, 2% agar) at 25℃ in darkness and for 10 days. The fruiting body cultivation were employed with the same methodology as the cultivation experiment. All biological samples were immediately frozen in liquid nitrogen and stored at -80℃. RNA sequencing performed by Wuhan Frasergen Co., Ltd, and raw sequencing data have been deposited in Genome Sequence Archive in National Genomics Data Center under accession number of CRA043948.

Low-quality reads and adapter sequences were trimmed using Trimmomatic v0.33 (Bolger et al., 2014), yielding high-quality clean reads. These reads were aligned to the *L. edodes* Wpm-1 reference genome using HISAT2 (Kim et al., 2015), with alignment quality assessed via Qualimap (García-Alcalde et al., 2012). Mapped clean reads were converted to TPM (transcripts per million) values using TBtools (Chen et al., 2018a). Differentially expressed genes (DEGs) were identified using edgeR (Chen et al., 2014), with the selection criteria set as $|\log_2(\text{fold change})| > 1$ and a false discovery rate (FDR) < 0.01. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed by ClusterProfiler with p-value threshold of 0.01 (Yu et al., 2012).

2.4 Candidate gene cloning and sequence analysis

Integrating GWAS results with differential expression profiling, we identified *LePC4* as a candidate gene significantly associated with CPC in *L. edodes*. High CPC strains YS19 (29.083 g/100g) and standard strain WX-1 (22.267 g/100g) were selected for molecular characterization. Genomic DNA and total RNA were extracted from fresh mycelia by CTAB method (Murray and Thompson, 1980) and RNAiso Plus reagent, respectively. First-strand cDNA synthesis was performed using Hiscript Q RT SuperMix (+ gDNA wiper) (Vazyme, Nanjing, China). Based on the genome database of the monokaryon Wpm-1 from *L. edodes* strain WX-1 (Shen et al., 2024), the DNA and coding sequence (CDS) of *LePC4* were amplified using the primers listed in Table S1 with DNA and cDNA as templates, respectively. The recovered and purified gene fragments were integrated into the pEASY-Blunt vector (TransGen Biotech, Beijing, China) for sequencing. The domains of *LePC4* were predicted using NCBI Conserved Domains database (CDD) (https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml) (Wang et al., 2023), and its homologous protein sequences in fungi were obtained through NCBI BLASTP (Altschul et al., 1997). A phylogenetic tree was constructed based on the neighbor-joining method using MEGA 11.0 (Kumar et al., 2016).

2.5 Vector construction and genetic transformation

Overexpression (OE) and RNA interference (RNAi) vectors were constructed as previous methods (Gong et al., 2014; Wang et al., 2018b), based on the binary vector pCAMBIA1300-g (Fig. S1A). For overexpression, the genomic fragment of the candidate gene *LePC4* was fused with the *Legpd* promoter via homologous recombination and inserted into the pCAMBIA1300-g backbone to generate the *LePC4*-OE construct (Fig. S1B). For gene silencing, an *LePC4*-RNAi fragment was cloned into the pCAMBIA1300-g vector containing a bidirectional promoter (*Leactin* and *Legpd*) to construct the *LePC4*-RNAi (Fig. S1C).

Agrobacterium tumefaciens-mediated method was used to genetic transformation (Wang et al., 2018b). Specifically, the *LePC4*-OE and *LePC4*-RNAi vector was transformed into the *L.edodes* strain YS19 and WX-1, respectively. The empty vector pCAMBIA1300-g was transformed into both strains as controls. Transgenic strains were selected using hygromycin B (HygB) for three rounds, which the treatment concentrations of 5 μg/mL for YS19, and 4 μg/mL for WX-1.

PCR amplification was performed to identify positive transformants: the *LePC4*-OE strains were validated using primers *LePC4*-cDNA-F and hyg-R; *LePC4*-RNAi lines were screened with primers *LePC4*-yan-F/R; and control strains were confirmed with primers HYG and gpdR11 (Table S1). Relative expression levels of *LePC4* were analyzed by quantitative real-time PCR (qRT-PCR), using *Leactin* as the internal reference gene. The $2^{-\Delta\Delta CT}$ method was applied for expression quantification (Luo et al., 2019).

2.6 Phenotypic analysis of transformant strains

The filtered transformants including *LePC4*-OE, *LePC4*-RNAi, and CK, were cultured on MYG medium (with Hygromycin B at the same concentration above) and in sawdust medium. Mycelial growth rate and morphology were evaluated to assess the impact of *LePC4* on strain growth characteristics. Three biological replicates were set for each treatment. To measure total protein content, *LePC4*-OE, *LePC4*-RNAi, and the CK transformants were inoculated into MYG liquid medium and shaking cultured at 25°C with 120 rpm for 16 days. The mycelia were harvested, washed with ddH_2O, and freeze-dried. The CPC of

the dried mycelia were detected with the Kjeldahl method (Kelrich, 1990). Three biological replicates were prepared for each strain.

Protein fractionation was performed following the Osborne method to extract four types of protein fractions, i.e., albumin, globulin, prolamin and glutelin (Fan et al., 2021; Zheng et al., 2026). 0.1 g of mycelial powder was transferred into a 1.5 mL centrifuge tube and mixed thoroughly with ddH_2O (v/g ratio: 9:1). The mixture was incubated in a water bath at 45°C for 1.5 h, then centrifuged at 8,000 rpm for 10 min. The supernatant was collected as the albumin fraction. Globulin, prolamin, and glutelin were subsequently extracted using 2% NaCl solution, 70% ethanol, and 0.03 mol/L NaOH solution, respectively, at volume-to-mass ratios of 9:1, 7:1, and 10:1. Protein concentrations of the four fractions were measured using the Bradford Protein Assay Kit (Beyotime, China). Absorbance was read at 595 nm, and protein content was calculated based on a standard curve. Three biological replicates were included for each strain.

3. Results

3.1 Candidate genes associated with CPC of *L. edodes* detected by GWAS

124 out of the 133 *L. edodes* strains successfully produced fruiting bodies in the cultivation experiment, and were thus included in CPC phenotyping (Table S2). CPC values across these strains ranged from 13.25 to 29.08 g/100 g dry weight (Fig. 1A), with a mean of 21.47 g/100 g dry weight and a coefficient of variation (CV) of 15.19%. Shapiro Wilk normality testing confirmed that CPC distribution did not significantly deviate from normality ($p = 0.5013$) (Fig. 1A), satisfying a key assumption for GWAS analysis.

To dissect the genetic basis of CPC in *L. edodes*, GWAS analysis was conducted using high-quality SNP markers and CPC genotype across the 124 strains. A total of 1,099,843 biallelic SNPs were identified for association analysis. Among these, 20 SNPs significantly associated with the CPC, mapping to 31 genes (Fig. 1B, C; Table S4). Functional annotation revealed that 12 of these genes are implicated in stress response and carbohydrate metabolism pathways (Table S3). For instance, *Lewpm1004763* encodes a phospholipase C (PLC) - like phosphodiesterase predicted to modulate inositol triphosphate which mediated signaling during environment stress adaptation (Ali et al., 2022). *Lewpm1006155* encodes DnaJ-domain-containing protein essential for maintaining proteostasis under thermal stress (Wang et al., 2018a). Candidate genes *Lewpm1001358, Lewpm1005711, Lewpm1011380* and *Lewpm1011382* were involved in hydrolyzing the carbohydrates and lignocellulose, which would provide a basic energy foundation for protein synthesis.

3.2 Analysis of differential expressed genes

A total of 91,640,425 reads from YS19 and 93,595,592 reads from YS66 were obtained (Table S4). The Pearson correlation coefficients results indicated that the reproducibility between biological replicates

was highly enough for further analysis (Fig. S2A).

In the mycelial stage, there were 1,173 DEGs between YS66 and YS19, including 860 up-regulated genes and 313 down-regulated ones. In the fruiting body stage, there were 1,973 DEGs between YS66 and YS19, including 878 up-regulated and 1,095 down-regulated genes. There were 435 DEGs that were differentially expressed in both the mycelial and the fruiting body stages, including 293 up-regulated genes and 142 down-regulated ones (Fig. 2A, B).

GO enrichment results showed that in mycelial stage, DEGs mainly enriched in oxidoreductase activity, monooxygenase activity and carbohydrate metabolic process (Fig. S3A). In fruiting body stage, DEGs mainly enriched in transmembrane transporter, organonitrogen compound metabolic process and sulfur compound metabolic process (Fig. S3B). KEGG enrichment results showed that DEGs in mycelial stage enriched in tyrosine metabolism, lysine degradation and amino acid metabolism (Fig. S3C). DEGs in fruiting body stage enriched in ribosome, cyanoamino acid metabolism and tryptophan metabolism (Fig. S3D). Above results indicated that DEGs enriched in amino acids metabolism pathway such as tyrosine, lysine and tryptophan may contribute to the variation in CPC among *L. edodes* strains.

Twelve DEGs were randomly selected for validation by quantitative real-time PCR (qRT-PCR), and the results showed strong concordance with the RNA-seq data, thereby corroborating the reliability and reproducibility of the transcriptomic profiling (Fig. S2B).

3.3 Identification and sequence analysis of the *LePC4*

Among the DEGs between YS66 and YS19, gene *Lewpm1006156* (named as *LePC4*) was identified as significantly associated with CPC in GWAS. This convergence of transcriptomic and genetic evidence strongly supports *LePC4* as a high-priority candidate gene for functional validation.

The *LePC4* locus spans 1,292 bp in the Wpm-1 reference genome, comprising a 684 bp CDS interrupted by four introns and organized into five exons (Fig. 3A). Conserved domain prediction using NCBI (CDD) revealed a single DUF5910 superfamily domain in the predicted protein (Fig. 3B), which linked in CDD to an entry annotated as being associated with apoptosis in animal cells. Phylogenetic tree analysis based on homologous protein sequences showed that *LePC4* clustered with a homolog from *Rhodocollybia butyracea* (Fig. 3C).

3.4 *LePC4* negatively regulates the crude protein content in *L. edodes*

To investigate the role of *LePC4* in regulating the CPCs of *L. edodes*, a high CPC strain YS19 and moderate CPC strain WX-1 with telomere-to-telomere genome used here were selected for genetic transformation. The YS19-OE and WX-1-RNAi vector were designed and verified by PCR (Fig. S4), then the gene expression of *LePC4* in both transformants were determined by qRT-PCR (Fig. 4A).

Subsequently, mycelial growth rates were quantified on both MYG agar medium and sterilized sawdust substrate to assess the functional impact of *LePC4*. Statistical analysis revealed no significant difference in growth rate on MYG or sawdust medium between each *LePC4* transformant and its wild-type control ($p > 0.05$) (Fig. 4B). In addition, colony morphology of all the transformants was comparable to those of the corresponding wild-type strains, with no obvious changes observed (Fig. 4C). Overall, these results indicated that the overexpression and RNAi of *LePC4* rarely affect mycelial growth rate and colony morphology in *L. edodes*.

The mycelia from each transformants were collected for CPCs determination. For *LePC4* RNAi transformants, the CPCs of WX-1-RNAi transformants showed a significantly increase, which ranged from 13.06% to 17.11% ($p < 0.05$) (Fig. 5A). In contrast, *LePC4* overexpression transformants of YS19 exhibited a significantly reduction of CPC, which decreased by 6.91% to 18.67% ($p < 0.05$) (Fig. 5B). Collectively, the above results suggested that *LePC4* showed no influence on the mycelia growth but negatively regulated the CPC in *L. edodes*.

3.5 Effect of *LePC4* expression on protein composition in *L. edodes*

The impact of *LePC4* on protein composition in *L. edodes* were further investigated, including the protein fractions -- albumin, globulin, prolamin and glutelin.

For the RNAi strains, all the WX-1 RNAi transformants exhibited significantly higher levels of albumin (increased by 7.41%–38.73%) and globulin (increased by 13.78%–70.77%) compared to the CK ($p < 0.05$). On the contrast, the content of glutelin in all the transformants significantly decreased by 16.27%–36.07% (Fig. 5C). For the overexpression strains, *LePC4* overexpression led to a significant reduction in albumin (9.64%–25.80%), globulin (15.97%–21.88%), and prolamin (12.85%–38.22%) compared to the CK in YS19. On the contrast, glutelin content in these overexpression transformants significantly increased by 36.23%–155.69% (Fig. 5D).

Taken together, our results indicated that interfering *LePC4* promotes the accumulation of albumin and globulin while reducing glutelin content. Conversely, *LePC4* overexpression exerts the opposite effect, shifting the protein profile toward increased glutelin and reduced albumin and globulin fractions.

4. Discussion

Edible mushrooms proteins possess remarkable thermal and pH stability, making them valuable for various protein processing applications (Komárek et al., 2018). Moreover, mushroom proteins contain a comprehensive profile of amino acids, including all the nine essential amino acids required by the human body (González et al., 2020; Pashaei et al., 2024). Therefore, they serve as a valuable dietary protein source and function synergistically with plant and animal proteins to enhance overall protein quality and nutritional completeness (Assemie and Abaya, 2022). The CPC and yield trait are critical breeding target in genetic improvement of *L. edodes*. In contrast to staple crops such as soybean and wheat, the

molecular mechanisms regulating the CPC in mushrooms remains poorly. Studies on elucidating the molecular mechanisms underlying CPC are required to further protein utilization in *L. edodes*.

Four candidate genes involved in hydrolyzing carbohydrates and lignocellulose were associated with CPC by GWAS. In particular, manganese peroxidase (*Lewpm1005711*) and cellobiohydrolase (*Lewpm1001358*) may contribute to CPC indirectly by improving the degradation and utilization of lignocellulosic substrates. Manganese peroxidase is one of the key lignin-degrading enzymes produced by white-rot fungi (Gold and Alic, 1993). Selective lignin degradation by fungi such as *Inonotus obliquus* and *Pleurotus* spp. can be accompanied by improved nutritional value, including increased crude protein and digestibility of crop residues (Zang et al., 2021). Similarly, cellobiohydrolase is a key cellulase that releases cellobiose from cellulose, and co-expression of endoglucanase and cellobiohydrolase has been reported to enhance cellulose degradation during whole-plant corn silage fermentation and increase CPC (Wan et al., 2024). These findings suggest that variation in CPC may be partly related to differences in lignocellulose degrading capacity, which could affect carbon source availability, fungal growth, and accumulation of fungal protein.

The transcriptome results further supported a potential metabolic basis for CPC variation. DEGs were enriched in oxidoreductase activity and carbohydrate/carbon metabolism, which is consistent with previous studies in *L. edodes* showing that increased CPC was accompanied by enhanced carbohydrate metabolism, oxidoreductase activity and amino acid metabolism (Zhao et al., 2025; Zhao et al., 2024). The DEGs enrichment in amino acid metabolism is directly related to protein biosynthesis, and similar associations between amino acid metabolic pathways and protein-content variation have been reported in wheat (Zhang et al., 2023). In addition, DEGs in sulfur compound metabolic process may influence CPC by regulating sulfur-containing amino acids such as cysteine and methionine, which are important for protein quality (Liu et al., 2022). Together, these findings suggest that CPC differences among *L. edodes* strains may be associated with coordinated changes in carbon utilization, amino acid metabolism and sulfur-related protein biosynthesis.

Functional validation further confirmed that *LePC4* acts as a negative regulator of total mycelial protein content. Conserved domain analysis revealed that *LePC4* encodes a protein homologous to the *Cordyceps militaris* protein (CMP), which has been reported to induce apoptosis and increase the proportion of cells in the sub-G1 phase (Bai and Sheu, 2018; Kim et al., 2020). The G1 phase is critical for RNA and protein biosynthesis, providing the basis for subsequent DNA replication and metabolic activity (Yuan et al., 2005). Based on this functional similarity, it is possible that *LePC4* may be involved in cellular processes related to cell cycle regulation in *L. edodes*. Reduced expression of *LePC4* may therefore mitigate cell cycle perturbation and favor a cellular state conducive to macromolecular accumulation, thereby potentially enhancing precursor availability for protein biosynthesis and contributing to increased CPC.

Beyond CPC, silencing of *LePC4* promoted the accumulation of albumin and globulin while reducing glutelin. Similar associations between total protein variation and coordinated remodeling of protein

fractions have been widely reported in plants. Cereal seed storage proteins exhibit dynamic compositional changes, in which alterations in one fraction are often accompanied by compensatory adjustments in others (Shewry and Halford, 2002). Consistently, in wheat cultivars, total protein content ranges from 9.32% to 12.60%, while albumin, globulin, gliadin and glutenin fractions also show substantial genotype-dependent variation (Siddiqi et al., 2021). Likewise, rice displays extensive natural variation in albumin, globulin, prolamin, glutelin and total storage protein contents, indicating that both total protein and individual fractions are genetically variable traits (Chen et al., 2018b). Furthermore, reduction of specific seed storage proteins has been shown to trigger compensatory increases in other protein fractions at both the transcript and protein levels (Kawakatsu et al., 2010). Together, these findings support the possibility that *LePC4* silencing may influenced CPC partly through coordinated remodeling of mycelial protein fractions, although the underlying regulatory mechanisms in *L. edodes* remain to be elucidated.

Previous study improved that grain yield and CPC correlated negatively in cereal crops (He et al., 2021). Therefore, future breeding programs for *L. edodes* should adopt a more precise, trait-informed approach to preserve alleles conferring optimal performance across multiple agronomic and quality-related traits. In this study, *LePC4* was found to negatively regulate CPC in *L. edodes* without influencing mycelial growth (Fig. 4B, C), indicating its potential utility for molecular breeding aimed at improving mycelial protein content and developing *L. edodes* mycelia as fermented protein resource.

5. Conclusion

In this study, *LePC4* was identified as a negatively regulator of CPC in *L. edodes* by integrating GWAS and comparative transcriptomic analysis. GWAS analysis identified 31 candidate genes associated with CPC in *L. edodes*, suggesting a polygenic basis in which variation in carbohydrate and lignocellulose metabolism may contribute to protein accumulation. Transcriptomic analysis suggested that CPC variation may be associated with oxidoreductase activity, carbohydrate/carbon metabolism, amino acid metabolism, and sulfur compound metabolism. Moreover, *LePC4* appears to regulate protein-fraction composition, as its silencing increased albumin and globulin while decreasing glutelin, whereas overexpression showed the opposite trend, suggesting a role in modulating protein fraction distribution during CPC formation. Our findings provide a basis for understanding CPC regulation and for breeding high-protein strains in *L. edodes*.

Declarations

Declaration of Competing Interest

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

Author Contribution

Chenjie Xu: Investigation, Writing–review & editing. Wei Zheng: Investigation, Writing–original draft, Writing–review & editing. Tianji Huang: Resources. Jian Zhang: Resources. Yang Xiao: Conceptualization, Funding acquisition, Supervision, Writing–review & editing.

Acknowledgement

This work was financially supported by the National Key Research and Development Program of China (Grant No. 2023YFF1000800), and the Supporting Projects for High-quality Development of Seed Industry in Hubei Province (Grant No. HBZY2023B004-2).

Data Availability

The raw sequencing data generated in this study are publicly available in the Genome Sequence Archive (GSA) of the National Genomics Data Center under accession number CRA043948.

References

1. Ali U, Lu S, Fadlalla T, Iqbal S, Yue H, Yang B, Hong Y, Wang X, Guo L (2022) The functions of phospholipases and their hydrolysis products in plant growth, development and stress responses. Prog Lipid Res 86(1):101158. https://doi.org/10.1016/j.plipres.2022.101158
2. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25(17):3389–3402. https://doi.org/10.1093/nar/25.17.3389
3. Assemie A, Abaya G (2022) The effect of edible mushroom on health and their biochemistry. Int. J. Microbiol. 2022 (1), 8744788. https://doi.org/10.1155/2022/8744788
4. Bach F, Helm CV, Bellettini MB, Maciel GM, Haminiuk CWI (2017) Edible mushrooms: a potential source of essential amino acids, glucans and minerals. Int J Food Sci Technol 52(1):2382–2392. https://doi.org/10.1111/ijfs.13522
5. Bai KC, Sheu F (2018) A novel protein from edible fungi Cordyceps militaris that induces apoptosis. J Food Drug Anal 26(1):21–30. https://doi.org/10.1016/j.jfda.2016.10.013
6. Bing S-J, Li YQ, Sun GJ, Wang CY, Liang Y, Hua DL, Chen L, Mo HZ (2023) Effect of different acidic or alkaline environments on structural characteristics, functional and physicochemical properties of *Lentinus edodes* protein. Process Biochem 130:15–25. https://doi.org/10.1016/j.procbio.2023.04.001
7. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30(15):2114–2120. https://doi.org/10.1093/bioinformatics/btu170

8. Chen C, Chen H, He Y, Xia R (2018a) TBtools, a toolkit for biologists integrating various biological data handling tools with a user-friendly interface. bioRxiv 289660. https://doi.org/10.1101/289660
9. Chen P, Shen Z, Ming L, Li Y, Dan W, Lou G, Peng B, Wu B, Li Y, Zhao D, Gao G, Zhang Q, Xiao J, Li X, Wang G, He Y (2018b) Genetic basis of variation in rice seed storage protein (albumin, globulin, prolamin, and glutelin) content revealed by genome-wide association analysis. Front Plant Sci 9:612. https://doi.org/10.3389/fpls.2018.00612
10. Chen Y, Lun ATL, Smyth GK (2014) Differential expression analysis of complex RNA-seq experiments using edgeR. In: Datta S, Nettleton D (eds) Statistical Analysis of Next Generation Sequencing Data. Springer, Cham, pp 51–74. https://doi.org/10.1007/978-3-319-07212-8_3.
11. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H (2021) Twelve years of SAMtools and BCFtools. GigaScience 10, giab008. https://doi.org/10.1093/gigascience/giab008
12. Fan HB, Zou Y, Huang SY, Liu YL, Zheng QW, Guo LQ, Lin JF (2021) Study on the physicochemical and emulsifying property of proteins extracted from *Pleurotus tuoliensis*. LWT 151, 112185. https://doi.org/10.1016/j.lwt.2021.112185
13. Fan H, Ge F, Wu T, Liu Y, Tian L, Liu Y, Xiang T, Yu H, Shi L, He Q (2023) The AMP-activated protein kinase (AMPK) positively regulates lysine biosynthesis induced by citric acid in *Flammulina filiformis*. J Fungi 9:340. https://doi.org/10.3390/jof9030340
14. García-Alcalde F, Okonechnikov K, Carbonell J, Cruz LM, Götz S, Tarazona S, Dopazo J, Meyer TF, Conesa A (2012) Qualimap: evaluating next-generation sequencing alignment data. Bioinformatics 28(20):2678–2679. https://doi.org/10.1093/bioinformatics/bts503
15. Gold MH, Alic M (1993) Molecular biology of the lignin-degrading basidiomycete *Phanerochaete chrysosporium*. Microbiol Rev 57(3):605–622. https://doi.org/10.1128/mr.57.3.605-622.1993
16. Gong W, Xu R, Xiao Y, Zhou Y, Bian Y (2014) Phenotypic evaluation and analysis of important agronomic traits in the hybrid and natural populations of *Lentinula edodes*. Sci Hortic 179:271–276. https://doi.org/10.1016/j.scienta.2014.09.044
17. González A, Cruz M, Losoya C, Nobre C, Loredo A, Rodríguez R, Contreras J, Belmares R (2020) Edible mushrooms as a novel protein source for functional foods. Food Funct 11(8):7400–7414. https://doi.org/10.1039/D0FO01746A
18. He T, Angessa TT, Hill CB, Zhang XQ, Chen K, Luo H, Wang Y, Karunarathne SD, Zhou G, Tan C, Wang P, Westcott S, Li C (2021) Genomic structural equation modelling provides a whole-system approach for the future crop breeding. Theor Appl Genet 134(9):2875–2889. https://doi.org/10.1007/s00122-021-03865-4
19. Kawakatsu T, Hirose S, Yasuda H, Takaiwa F (2010) Reducing rice seed storage protein accumulation leads to changes in nutrient quality and storage organelle formation. Plant Physiol 154(4):1842–1854. https://doi.org/10.1104/pp.110.164343
20. Helrich K (1990) Official Methods of Analysis, 15th edn. Association of Official Analytical Chemists, Arlington, VA

21. Kim D, Langmead B, Salzberg SL (2015) HISAT: a fast spliced aligner with low memory requirements. Nat Methods 12(3):357–360. https://doi.org/10.1038/nmeth.3317
22. Kim MJ, Kim C-H, An M-J, Lee J-H, Shin G-S, Song M, Kim JW (2020) Ethylparaben induces apoptotic cell death in human placenta BeWo cells via the Caspase-3 pathway. Anim Cells Syst 24(1):34–43. https://doi.org/10.1080/19768354.2020.1711804
23. Komárek J, Ivanov Kavková E, Houser J, Horáčková A, Ždánská J, Demo G, Wimmerová M (2018) Structure and properties of AB21, a novel *Agaricus bisporus* protein with structural relation to bacterial pore-forming toxins. Proteins 86(9):897–911. https://doi.org/10.1002/prot.25522
24. Kumar S, Stecher G, Tamura K (2016) Mol Biol Evol 33(7):1870–1874. https://doi.org/10.1093/molbev/msw054. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets
25. Li X, Cao Q, Liu G (2025) Advances, applications, challenges and prospects of alternative proteins. J Food Compos Anal 137:106900. https://doi.org/10.1016/j.jfca.2024.106900
26. Lian L, Qiao J, Guo X, Xing Z, Ren A, Zhao M, Zhu J (2023) The transcription factor GCN4 contributes to maintaining intracellular amino acid contents under nitrogen-limiting conditions in the mushroom *Ganoderma lucidum*. Microb Cell Fact 22:205. https://doi.org/10.1186/s12934-023-02213-z
27. Lippert C, Listgarten J, Liu Y, Kadie CM, Davidson RI, Heckerman D (2011) FaST linear mixed models for genome-wide association studies. Nat Methods 8(10):833–835. https://doi.org/10.1038/nmeth.1681
28. Liu F, Ma XB, Han B, Wang B, Xu JP, Cao B, Ling ZL, He MQ, Zhu XY, Zhao RL (2025a) Pan-genome analysis reveals genomic variations during enoki mushroom domestication, with emphasis on genetic signatures of cap color and stipe length. J Adv Res 75:199–212. https://doi.org/10.1016/j.jare.2024.11.005
29. Liu J, Li Q, Jiang P, Xu Z, Zhang D, Zhang L, Zhang M, Yu H, Song C, Tan Q (2019) Overexpression of the saccharopine dehydrogenase gene improves lysine biosynthesis in *Flammulina velutipes*. J Basic Microbiol 59:890–900. https://doi.org/10.1002/jobm.201900189
30. Liu P, Zhang Z, Wu D, Li W, Chen W, Yang Y (2025b) The prospect of mushroom as an alternative protein: From acquisition routes to nutritional quality, biological activity, application and beyond. Food Chem 469:142600. https://doi.org/10.1016/j.foodchem.2024.142600
31. Liu Z, Liu D, Fu X, Du X, Zhang Y, Zhen W, Li S, Yang H, He S, Li R (2022) Integrated transcriptomic and metabolomic analyses revealed the regulatory mechanism of sulfur application in grain yield and protein content in wheat (*Triticum aestivum* L.). Front. Plant Sci 13:935516. https://doi.org/10.3389/fpls.2022.935516
32. Luo Y, Wang G, Wang C, Gong Y, Bian Y, Zhou Y (2019) Selection and validation of reference genes for qRT-PCR in *Lentinula edodes* under different experimental conditions. Genes 10:647. https://doi.org/10.3390/genes10090647
33. Murray M, Thompson W (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8(19):4321–4326. https://doi.org/10.1093/nar/8.19.4321

34. Pashaei KHA, Irankhah K, Namkhah Z, Sobhani SR (2024) Edible mushrooms as an alternative to animal proteins for having a more sustainable diet: a review. J Health Popul Nutr 43:205. https://doi.org/10.1186/s41043-024-00701-5
35. Roszczyk A, Turło J, Zagożdżon R, Kaleta B (2022) Immunomodulatory properties of polysaccharides from *Lentinula edodes*. Int J Mol Sci 23:8980. https://doi.org/10.3390/ijms23168980
36. Sexton AE, Garnett T, Lorimer J (2019) Framing the future of food: The contested promises of alternative proteins. Environ Plan E Nat Space 2:47–72. https://doi.org/10.1177/2514848619827009
37. Shen N, Xie H, Liu K, Li X, Wang L, Deng Y, Chen L, Bian Y, Xiao Y (2024) Near-gapless genome and transcriptome analyses provide insights into fruiting body development in *Lentinula edodes*. Int J Biol Macromol 263:130610. https://doi.org/10.1016/j.ijbiomac.2024.130610
38. Shewry PR, Halford NG (2002) Cereal seed storage proteins: structures, properties and role in grain utilization. J Exp Bot 53(370):947–958. https://doi.org/10.1093/jexbot/53.370.947
39. Siddiqi RA, Singh TP, Rani M, Sogi DS (2021) Electrophoretic characterization and proportion of different protein fractions in wheat cultivars of North-India. J Agric Food Res 4:100137. https://doi.org/10.1016/j.jafr.2021.100137
40. Sobczak P, Grochowicz J, Łusiak P, Żukiewicz-Sobczak W (2023) Development of alternative protein sources in terms of a sustainable system. Sustainability 15:12111. https://doi.org/10.3390/su151612111
41. Sousa AS, Araújo-Rodrigues H, Pintado ME (2023) The health-promoting potential of edible mushroom proteins. Curr Pharm Des 29:804–823. https://doi.org/10.2174/1381612829666221223103756
42. Tian L, Ma Z, Qiu H, Liu X, Wu T, Ge F, Liu R, Zhu J, Shi L, Jiang A (2022) Chitosan increases lysine content through amino acid transporters in *Flammulina filiformis*. Foods 11:2163. https://doi.org/10.3390/foods11142163
43. Turło J, Gutkowska B, Herold F, Krzyczkowski W, Błażewicz A, Kocjan R (2008) Optimizing vitamin B12 biosynthesis by mycelial cultures of *Lentinula edodes* (Berk.) Pegl. Enzyme Microb Technol 43:369–374. https://doi.org/10.1016/j.enzmictec.2008.05.005
44. Uthayakumaran S, Gras P, Stoddard F, Bekes F (1999) Effect of varying protein content and glutenin-to-gliadin ratio on the functional properties of wheat dough. Cereal Chem 76(3):389–394. https://doi.org/10.1094/CCHEM.1999.76.3.389
45. Vetter J (1995) Mineralstoff-und Aminosäuregehalte des eßbaren, kultivierten Pilzes Shii-take (*Lentinus edodes*). Z Lebensm Unters Forsch 201(1):17–19. https://doi.org/10.1007/BF01193193
46. Wan X, SunKang Y, Chen Y, Zhang Z, Gou H, Xue Y, Wang C, Wei Y, Yang Y (2024) Co-expression of endoglucanase and cellobiohydrolase from yak rumen in lactic acid bacteria and its preliminary application in whole-plant corn silage fermentation. Front Microbiol 15:1442797. https://doi.org/10.3389/fmicb.2024.1442797

47. Wang GZ, Ma CJ, Luo Y, Zhou SS, Zhou Y, Ma XL, Cai YL, Yu JJ, Bian YB, Gong YH (2018a) Proteome and transcriptome reveal involvement of heat shock proteins and indoleacetic acid metabolism process in *Lentinula edodes* thermotolerance. Cell Physiol Biochem 50:1617–1637. https://doi.org/10.1159/000494784
48. Wang G, Zhou S, Luo Y, Ma C, Gong Y, Zhou Y, Gao S, Huang Z, Yan L, Hu Y (2018b) The heat shock protein 40 LeDnaJ regulates stress resistance and indole-3-acetic acid biosynthesis in *Lentinula edodes*. Fungal Genet Biol 118:37–44. https://doi.org/10.1016/j.fgb.2018.07.002
49. Wang J, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu S, Marchler GH, Song JS, Thanki N, Yamashita RA, Yang M, Zhang D, Zheng C, Lanczycki CJ, Marchler-Bauer A (2023) The conserved domain database in 2023. Nucleic Acids Res 51:D384–D388. https://doi.org/10.1093/nar/gkac1096
50. Wang L, Gong W, Xie H, Liu G, Xiao Y (2024) NUA4 histone acetylase gene LeEaf6 regulates precocity of shiitake mushroom, *Lentinula edodes*. Sci Hortic 328:112912. https://doi.org/10.1016/j.scienta.2024.112912
51. Yu G, Wang LG, Han Y, He Q-Y (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 16(5):284–287. https://doi.org/10.1089/omi.2011.0118
52. Yu H, Zhang L, Shang X, Peng B, Li Y, Xiao S, Tan Q, Fu Y (2022) Chromosomal genome and population genetic analyses to reveal genetic architecture, breeding history and genes related to cadmium accumulation in *Lentinula edodes*. BMC Genomics 23:120. https://doi.org/10.1186/s12864-022-08325-x
53. Yuan X, Shan Y, Zhao Z, Chen J, Cong Y (2005) G0/G1 arrest and apoptosis induced by SARS-CoV 3b protein in transfected cells. Virol J 2(1):66. https://doi.org/10.1186/1743-422X-2-66
54. Zang Q, Chen X, Zhang C, Lin M, Xu X (2021) Improving crude protein and methionine production, selective lignin degradation and digestibility of wheat straw by *Inonotus obliquus* using response surface methodology. J Sci Food Agric 102:1146–1154. https://doi.org/10.1002/jsfa.11451
55. Zhang C, Zhang P, Zhang X, Wang Q, Liu J, Li L, Cheng S, Qin P (2023) Integrated metabolome and transcriptome analyses reveal amino acid biosynthesis mechanisms during the physiological maturity of grains in Yunnan hulled wheat (*Triticum aestivum* ssp. yunnanense King). Int J Mol Sci 24:13475. https://doi.org/10.3390/ijms241713475
56. Zhang J, Shen N, Li C, Xiang X, Liu G, Gui Y, Patev S, Hibbett DS, Barry K, Andreopoulos W et al (2022) Population genomics provides insights into the genetic basis of adaptive evolution in the mushroom-forming fungus *Lentinula edodes*. J Adv Res 38:91–106. https://doi.org/10.1016/j.jare.2021.09.008
57. Zhao Y, Li H, Yao Y, Wei Q-J, Hu T, Li X, Zhu B, Ma H (2025) Combined analysis of proteomics and metabolism reveals critical roles of oxidoreductase activity in mushrooms stimulated by wolfberry and sea buckthorn substrates. Front Nutr 12:1543240. https://doi.org/10.3389/fnut.2025.1543240
58. Zhao Y, Yao Y, Li H, Han Z, Ma X (2024) Integrated transcriptome and metabolism unravel critical roles of carbon metabolism and oxidoreductase in mushroom with Korshinsk peashrub substrates. BMC Genomics 25:763. https://doi.org/10.1186/s12864-024-10666-8

59. Zheng W, Xu C, Liu G, Zhou Y, Xiao Y (2026) Physicochemical and functional properties of isolated proteins from shiitake mushroom, *Lentinula edodes*. Int J Biol Macromol 338:149648. https://doi.org/10.1016/j.ijbiomac.2025.149648
60. Zhou R, Liu ZK, Zhang YN, Wong JH, Ng TB, Liu F (2019) Research progress of bioactive proteins from the edible and medicinal mushrooms. Curr Protein Pept Sci 20(3):196–219. https://doi.org/10.2174/1389203719666180625113036
61. Supplementary Figures

Figures

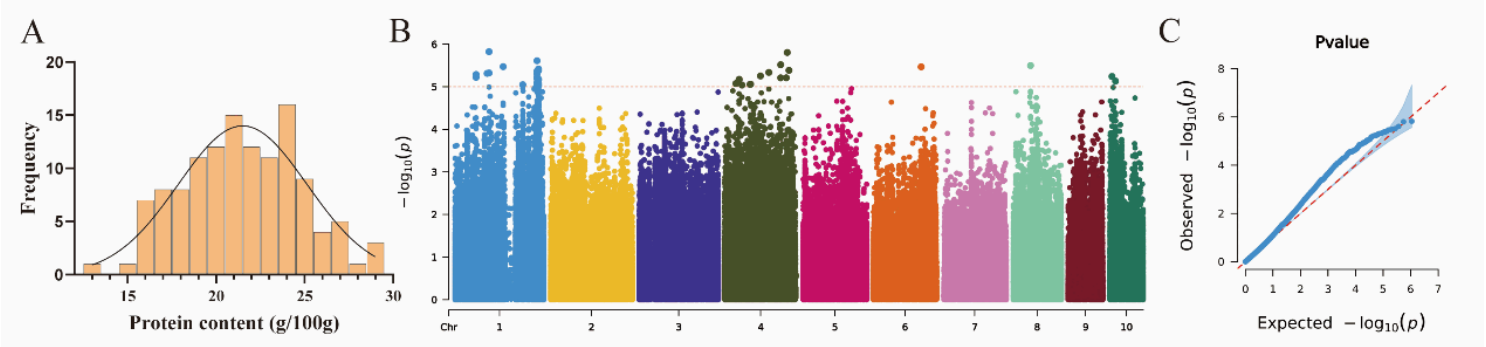


Figure 1

GWAS of the CPC in *L. edodes*.

The frequency distribution of CPC (g/100g) among the 124 strains (A). Manhattan plot (B) and Quantile-Quantile plot (C).

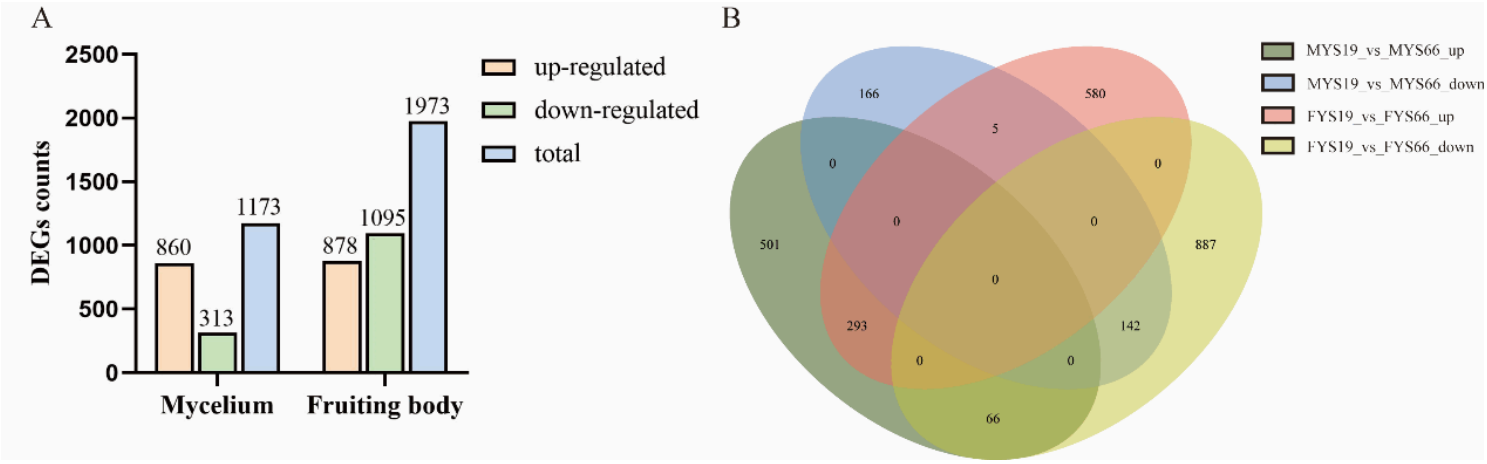


Figure 2

DEGs between YS66 and YS19 at mycelium (M) and fruiting body stages (F).

Number of up-regulated, down-regulated and total DEGs identified in mycelium and fruiting body stages (A). Venn diagram of DEGs between YS66 and YS19 (B).

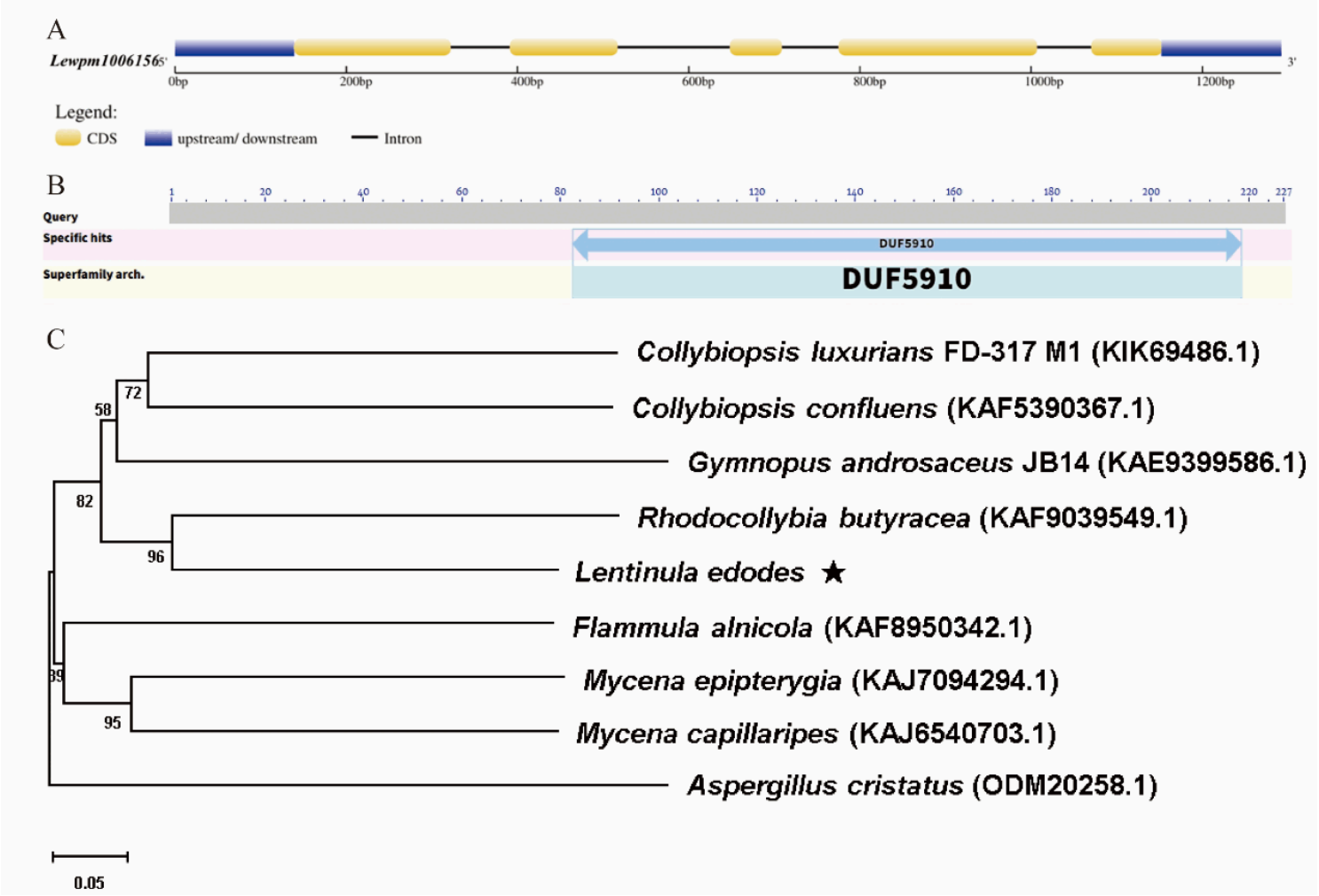


Figure 3

The gene structure and phylogenetic tree of *LePC4*.

Gene structure of *LePC4*in WX-1 strain (A). Protein domain of LePC4 (B). Phylogenetic tree of LePC4 protein and other fungal homologs (C).

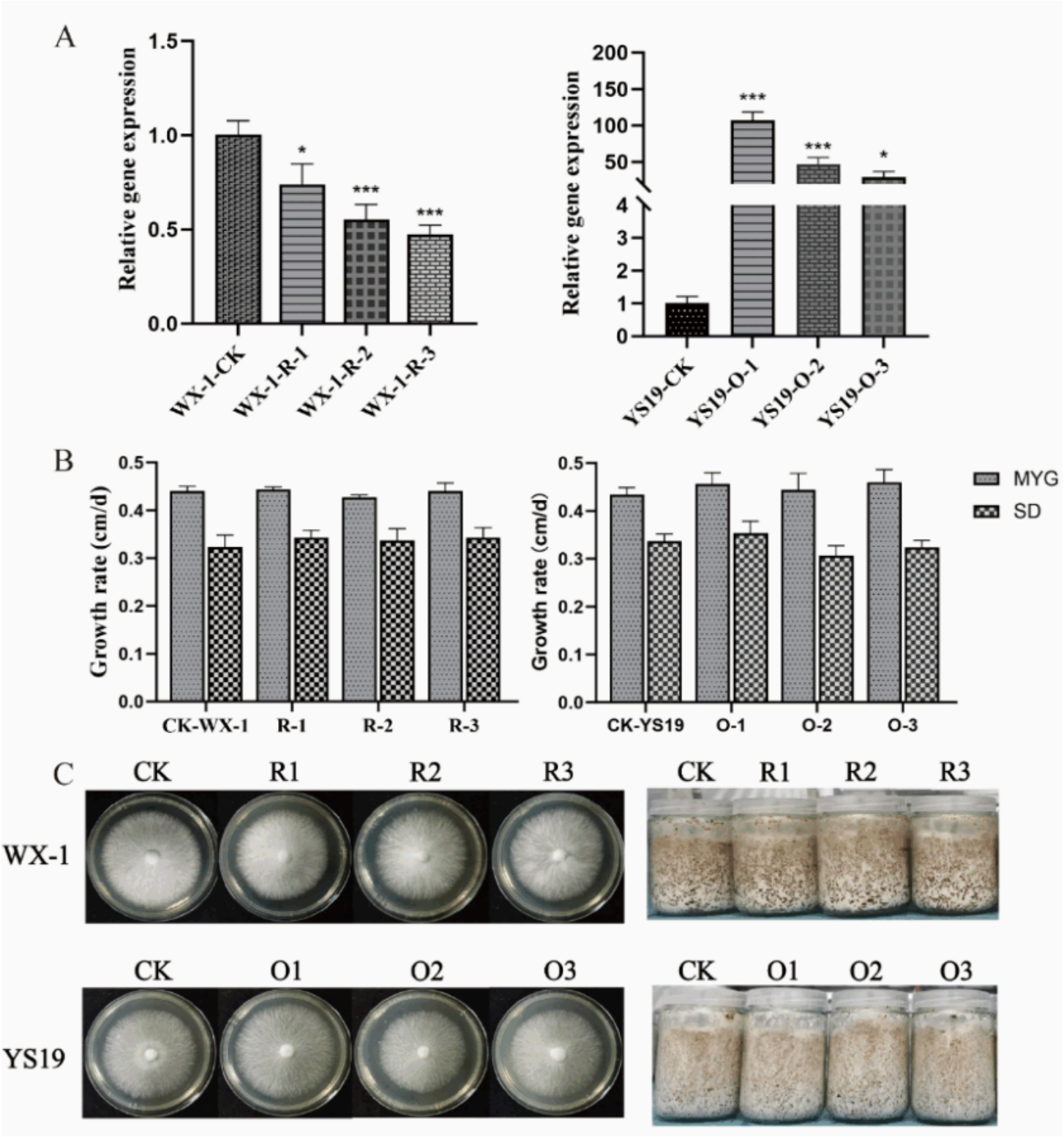


Figure 4

Transformants verification of *LePC4* in WX-1 and YS19

Relative gene expression of transformants (A). Mycelial growth rate of transformants and the control (B). The colony morphology of the transformants on MYG medium (left) and sawdust medium (right) (C).

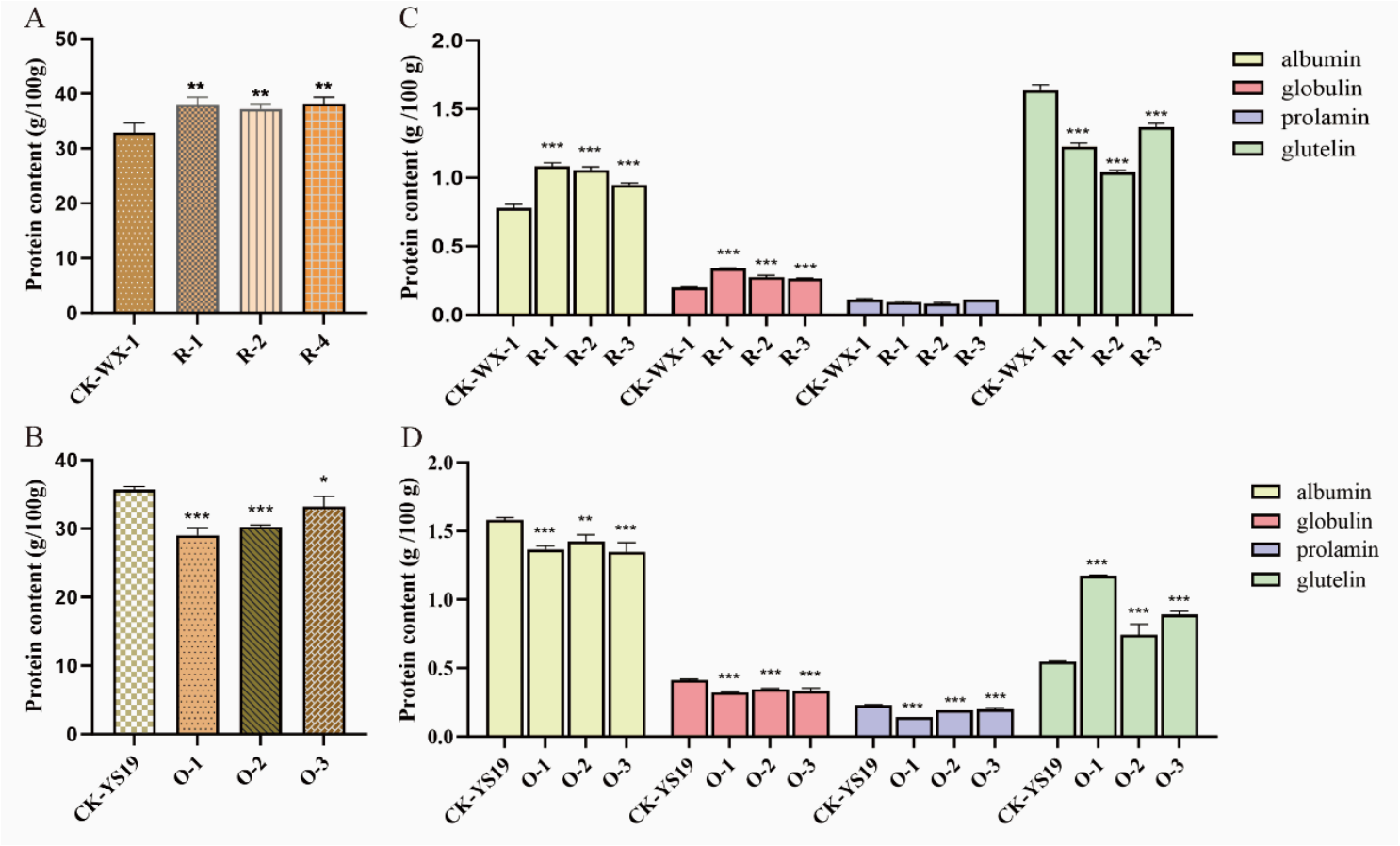


Figure 5

CPC and protein fraction composition determination of *LePC4* transformants

The CPC of WX-1(A) and YS19 (B) transformants. The protein fraction composition of WX-1 (C) and YS19 (D) transformants.

Supplementary Files

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

- SupplementaryTables.xls.xlsx
- SupplementaryFigures.docx