Received: 26 April 2021 | Revised: 2 October 2021 | Accepted: 18 October 2021

DOI: 10.1002/pmic.202100115

Proteomics
Proteomics and Systems Biology

DATASET BRIEF

Quantitative proteomic analysis of hepatic tissue in allotetraploid hybridized from red crucian carp and common carp identified crucial proteins and pathways associated with metabolism and growth rate

Yujie Yan[1,2] | Junting Wang[2] | Xiaoping Dong[2] | Yisheng Cai[2] | Yude Wang[1] | Li Ren[1] | Chun Zhang[1] | Min Tao[1] | Kaikun Luo[1] | Yong Zeng[1,2] | Shaojun Liu[1]

[1] The State Key Laboratory of Developmental Biology of Freshwater Fish, College of Life Science, Hunan Normal University, Changsha, Hunan, China

[2] National & Local Joint Engineering Laboratory of Animal Peptide Drug Development, College of Life Sciences, Hunan Normal University, Changsha, Hunan, China

Correspondence
Yong Zeng and Shaojun Liu, National & Local Joint Engineering Laboratory of Animal Peptide Drug Development, College of Life Sciences, Hunan Normal University, Changsha, 410081 Hunan, China.
Email: yongz@hunnu.edu.cn; lsj@hunnu.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Number: 31873038; Outstanding Youth Foundation of Hunan Education Department, Grant/Award Number: 19B356; Natural Science Foundation of Hunan Province, Grant/Award Number: 2019JJ50396

Abstract

Allotetraploid is a new species produced by distant hybridization between red crucian carp (*Carassius auratus* red var., abbreviated as RCC) and common carp (*Cyprinus carpio* L., abbreviated as CC). There is a significant difference in growth rate between allotetraploid and its parents. However, the underlying molecular mechanism is largely unknown. In this study, to find direct evidence associated with metabolism and growth rate in protein level, we performed quantitative proteomics analysis on liver tissues between allotetraploid and its parents. A total of 2502 unique proteins were identified and quantified by SWATH-MS in our proteomics profiling. Subsequently, comprehensive bioinformatics analyses including gene ontology enrichment analysis, pathway and network analysis, and protein–protein interaction analysis (PPI) were conducted based on differentially expressed proteins (DEPs) between allotetraploid and its parents. The results revealed several significant DEPs involved in metabolism pathways in liver. More specifically, the integrative analysis highlighted that the DEPs *ACSBG1*, *OAT*, and *LDHBA* play vital roles in metabolism pathways including "pentose phosphate pathway," "TCA cycle," and "glycolysis and gluconeogenesis." These could directly affect the growth rate in fresh water fishes by regulating the metabolism, utilization, and exchange of substance and energy. Since the liver is the central place for metabolism activity in animals, we firstly established the comprehensive and quantitative proteomics knowledge base for liver tissue from freshwater fishes, our study may serve as an irreplaceable reference for further studies regarding fishes' culture and growth.

KEYWORDS
allotetraploid, liver, metabolism, network analysis, SWATH-MS

Abbreviations: 4nAT, allotetraploid; BCA, bicinchoninic acid; CC, common carp; DDA, data dependent acquisition; DIA, data independent acquisition; FDR, false discovery rate; RCC, red crucian carp

Allotetraploid (4nAT) is a bisexual and fertile freshwater tetraploid fish with stable inheritable trait. It was firstly produced in vertebrates by distant hybridization between red crucian carp (RCC, *Carassius auratus* red var.) and common carp (CC, *Cyprinus carpio* L.). Liu has bred 4nAT F_{18} by distant hybridization between RCC and CC [1]. Allotetraploid plays a significant role in genetics and breeding. Triploid fish is an economical fish which can be bred by 4nAT and diploid fish. In morphological characteristics, there are many similarities between allotetraploid and its parents, as well as differences. Many traits in allotetraploid such as body length, body height, and the numbers of vertebrae are balanced between its parents RCC and CC [2]. For example, there are two pairs of whiskers in allotetraploid which are shorter than its male parent CC, while there are no whiskers in its female parent RCC. The growth rate of 4nAT is slower than its male parents CC, while it is faster than its female parents RCC. The reproduction of allotetraploid not only offers new insight to understand the evolution of vertebrae but also provides researchers with parents' fish for triploid reproduction [2].

The liver is the largest digestive organ in vertebrate, and the liver cells are organized around lobules as functional structural units. Generally, the liver is comprised of the left lobe and the right lobe in vertebrate. But there are differences in fishes including CC, RCC, and 4nAT. Their livers are diffuse glands scattering on the mesentery with pancreatic to form hepatopancreas. The liver is a hub organ for many biological processes including nutrient metabolism, blood volume regulation, and immune response [3, 4]. The liver could regulate systematic glucose and lipid flux even during fasting to maintain energy homeostasis and glycemia level [5, 6]. Liver serves as a vital role in regulating growth, immunity, and metabolism response to bacterial infection [7].

SWATH-MS is a newly developed data acquisition mode of mass spectrometry. It has been one of the most eye-catching mass spectrometry technologies in recent years. Comparing to traditional data-dependent acquisition (DDA), SWATH-MS could identify more proteins with very low abundance in complex samples. All the precursor ions are fragmented in each acquisition window at high speed cyclically during data collection in SWATH-MS [8]. Then, collect the information of every fragmented precursor ions [9, 10]. Therefore, with high-throughput, high resolution, high reproducibility, and accurate quantitation, SWATH-MS is suitable for proteomics profiling on complex samples.

As a homemade characteristic freshwater fish, 4nAT has not been studied as much as other fishes. The object of our study is to find the core genes that play vital roles in metabolism and immunity from liver by SWATH-MS. The study may provide useful information to those researchers who work on developmental biology and fish immunology.

4nAT, RCC, and CC were obtained from the State Key Laboratory of Developmental Biology of Freshwater Fish, College of Life Science. The license number for animal manipulation was 2020318. Animal experimenters were certified under a professional training course for laboratory animal practitioners held by the Institute of Experimental Animals, Hunan Province, China. All fish were euthanized using 2-phenoxyethanol (Sigma, USA) before dissection. The fish were humanely treated following the regulations of the Administration of Affairs Concerning Experimental Animals for the Science and Technology Bureau of China. The liver samples were taken and collected from above three species by professional technicians. Fresh tissue samples were immediately frozen by liquid nitrogen and then stored at −80°C for further analysis. All the fishes used in this study were cultured in similar environment with the same temperature, illumination, and feeding patterns. The tested fish were reared in three 100 m^2 ponds (each pond farmed one kind of fish). All the tested fish were fed with commercial diets, and CC, RCC, and 4nAT aged 10–11 months were collected for the present experiments. Three fishes per species and three liver samples from each fish were used for proteomics analysis in this study.

Bicinchoninic acid (BCA) quantitative kit (ThermoScientific) was used to extract the total proteins. A total protein extract of 50 μg were mixed with 30 μL reducing buffer (10 mM DTT, 8 M Urea, 50 mM NH_4HCO_3, pH 8.5) and filtered in 10 KD ultrafiltration tube. The solution was incubated at 55°C for 1 h and then added IAA to the solution with the final concentration of 50 mM. Then, all samples were incubated for 20 min at room temperature in the dark. Then centrifuge the solution on the filters at 12,000 rpm for 20 min at 4°C and then discard the flow-through from the collection tube. Added 200 μL 8 M urea containing 50 mM NH_4HCO_3 buffer and centrifuged at 12,000 rpm for 20 min. After washing step, the filter units were transferred into new collection tubes and 200 μL 50 mM NH_4HCO_3 was added, then repeat this step with 1 μL sequencing-grade trypsin (1 μg/μL) in each tube [11]. And the solutions were incubated for digestion at 37°C for 14 h. Finally, centrifuged the collections of digested peptides at 12,000 rpm for 10 min. Added 100 μL 50 mM NH_4HCO_3 and then centrifuged again. The peptide solution was mixed and then the mixture solution was completely dried in a SpeedVac.

For high pH reversed phase peptide prefractionation based on HPLC, 20 μg tryptic peptides of each sample was added and then each sample was loaded onto an Agilent ZORBAX Extend-C18 column (5 μm C18, 2.1 × 150 mm) with buffer A (10 mM NH_4OH in 95% H_2O, pH 10.0) at 0.3 mL/min, by using an Agilent 1100 HPLC system (Agilent, Santa Clara, USA) connected to a sample collector. The peptides were eluted from the column on a 0–35% ramp gradient with a 70 min gradient with buffer B (10 mM NH_4OH in 90% ACN, pH 10.0) into 32 fractions, mixed and added into 12 pools.

The pool of the tryptic digestion samples were fractionated by HPLC and isolated into 12 fractions. According to manufacturer's instructions, the iRT peptides (Biognosys, Schlieren, Switzerland) were spiked into each peptides fraction. Peptide sample of 1.5 μg of each prefractionation was loaded onto the trap column (ChromXP C18, 3 μm, 120 Å, Sciex, USA). The online chromatography separation was performed on the Ekspert nanoLC 415 system (SCIEX, Concord, ON). The trapping, desalting procedures were carried out at a flow rate of 4 μL/min for 5 min with 100% solvent A (water/acetonitrile/formic acid [98/2/0.1%; B, 2/98/0.1%]). Then, an elution gradient of 5–38% solvent B in 100 min was used on an analytical column (75 μm × 15 cm, C18, 3 μm 120 Å, ChromXP, Eksigent) with a total run time of 120 min. The Triple TOF 6600 mass spectrometer (Sciex, Concord, Ontario, Canada) fitted with a Nanospray III ion source was operated in positive DDA mode with the following settings: an ion spray voltage of 2.3 kV, curtain gas of 35 PSI,

nebulizer gas of 12 PSI, and an interface heater temperature of 150°C. The MS was operated with TOF-MS scans. If the charge state was 2–4 and the 260 cps threshold was exceeded, survey scans were performed within 250 ms and up to 40 ion products scans were collected (50 ms). A rolling collision energy setting was applied to all precursor ions for collision-induced dissociation. Dynamic exclusion was set for 18 s.

iRT calibration peptides (Biognosys, Schlieren, Switzerland) were spiked into the DIA samples according to manufacturer's instructions. SWATH quantification analysis parameters were the same as for spectral identification library analyses, with the following exceptions: the mass-range *m/z* 350–1350 was divided into 60 variable windows based on *m/z* density, with 1 Da overlap between the two adjacent mass windows, and the spread collision energy (CES) was set at 8 V. The ProteinPilot Software v.5.0 (Sciex Inc., USA) was used for protein identification. The global false discovery rate (FDR) at peptides and proteins level was estimated with the PSPEP tool integrated in the ProteinPilot Software to be 1.0% after searching against a combined database about 62,000 entries including homemade transcriptome data of 4nAT and the protein database of Danio Rerio from Uniprot. The search parameters were as follows: Sample type: Identification; Cys Alkylation: Iodoacetamide; Digestion: Trypsin; Instrument: TripleTOF 6600; Species: None; Search effort: Thorough ID. All SWATH-MS data were analyzed with Spectronaut 11 (Biognosys) using default settings which included dynamic peak detection, automatic precision non-linear iRT calibration, interference correction, and cross run normalization (total peak area) enabled. The spiked-in HRM peptides for *m/z* and retention time calibration was utilized by Spectronaut. All results were filtered by a *q*-value of 0.01 (equal to an FDR of 1% on peptide level).

Due to the limitation of sensitivity of proteomics technology, missing value was an inevitable problem in quantitative proteomics study. To minimize the error and reduce the FDR, three biological and technological runs were performed respectively for each subject. Only those proteins identified in triplicate runs were considered to perform further analysis. The average value of triplicate analyses was used to represent the protein expression level in each sample.

Proteins identified in triplicate runs were selected for further analysis. The mean value from triplicate analyses was used to represent the protein expression level in each sample. To identify the differentially expressed proteins (DEPs), student's t-test and fold change were performed to estimate DEPs among 4nAT, CC, and RCC. *p* value <0.05 and ratio <0.5 or >2 was set as the statistical cutoff line. Detailed information of all DEPs was illustrated in Table S3.

ClusterProfiler was an R package that can perform enrichment analysis of gene clusters with an integrative strategy [12]. In order to ensure maximal covering of the significant biological information, all the DEPs were used for functional annotation and enrichment analysis.

DAVID (Database for Annotation, Visualization, and Integrated Discovery) was a comprehensive database that can provide many online tools for integrative and systematic functional annotation and gene enrichment analysis [13, 14]. All the DEPs were imported into DAVID for keywords enrichment analysis.

All DEPs were imported into STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) with default parameters for

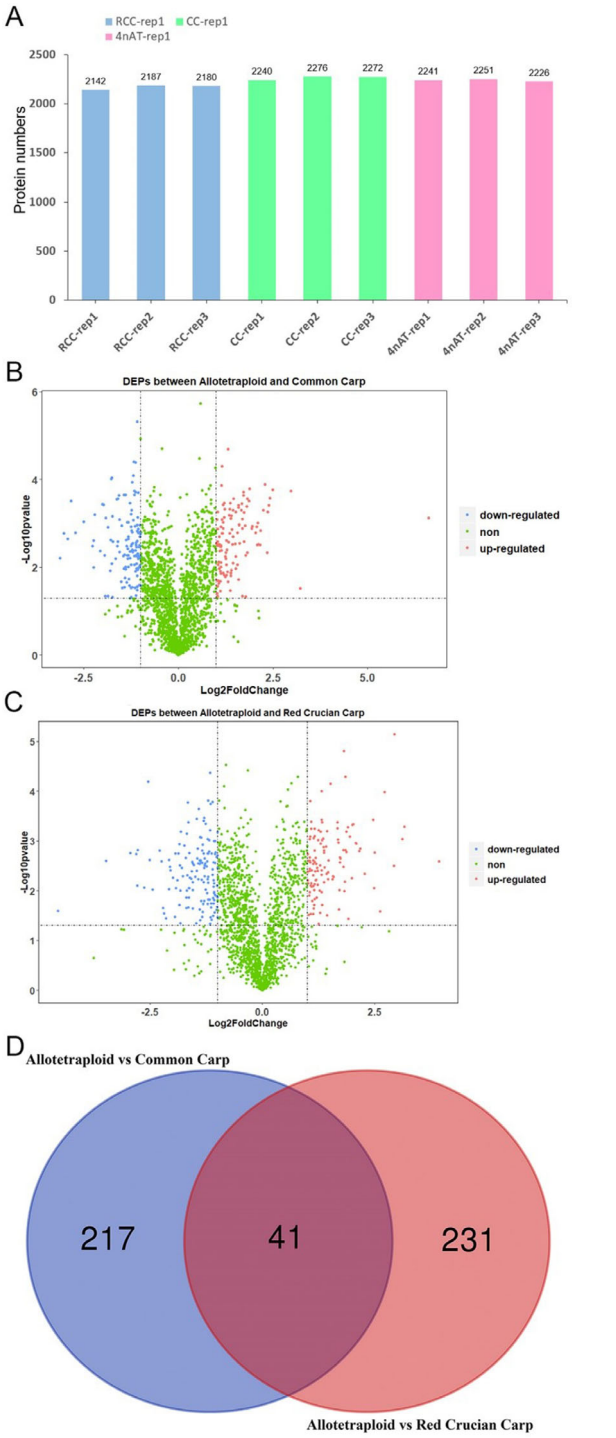

FIGURE 1 Statistical information of proteomics profiling. (A) The total proteins identified by SWATH-MS in CC, RCC, and 4nAT. There are three replicates with each sample of fish. (B) The DEPs identified between 4nAT and CC (ratio <0.5 or >2, *p* value < 0.05). Blue dots represent downregulated proteins and red for upregulated proteins. (C) The DEPs identified between 4nAT and RCC (ratio <0.5 or >2, *p* value < 0.05). Blue dots represent downregulated proteins and red for upregulated proteins. (D) The overlapped DEPs between two groups (Group 1: DEPs between 4nAT and CC; Group 2: DEPs between 4nAT and RCC). We highlighted the DEPs and overlapped DEPs in these three fishes. CC, common carp; DEP, differentially expressed protein; 4nAT, allotetraploid; RCC, red crucian carp.

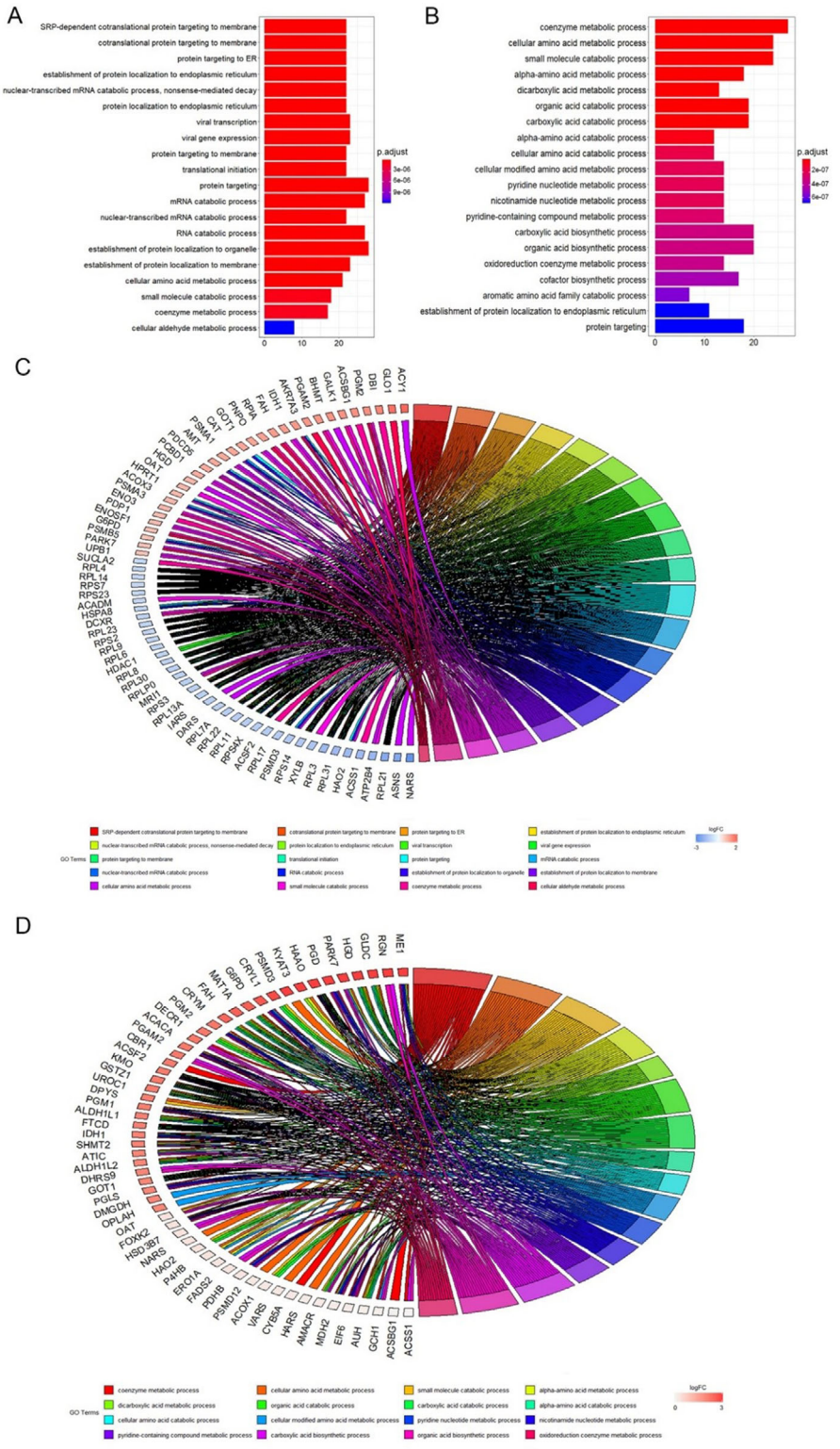

FIGURE 2 The results of gene functional annotation and enrichment analysis. (A) The results of significantly enriched GO terms based on adjust *p* values between 4nAT and CC. (B) The results of significantly enriched GO terms based on adjusted *p* values between 4nAT and RCC. (C) The DEPs participate in these processes including upregulated and downregulated proteins. *ACSBG1* and *OAT* were upregulated in 4nAT compared to CC. (D) The DEPs participated in these processes were upregulated in 4nAT. *ACSBG1* and *OAT* were downregulated in 4nAT compared to RCC. CC, common carp; DEP, differentially expressed protein; 4nAT, allotetraploid; RCC, red crucian carp.

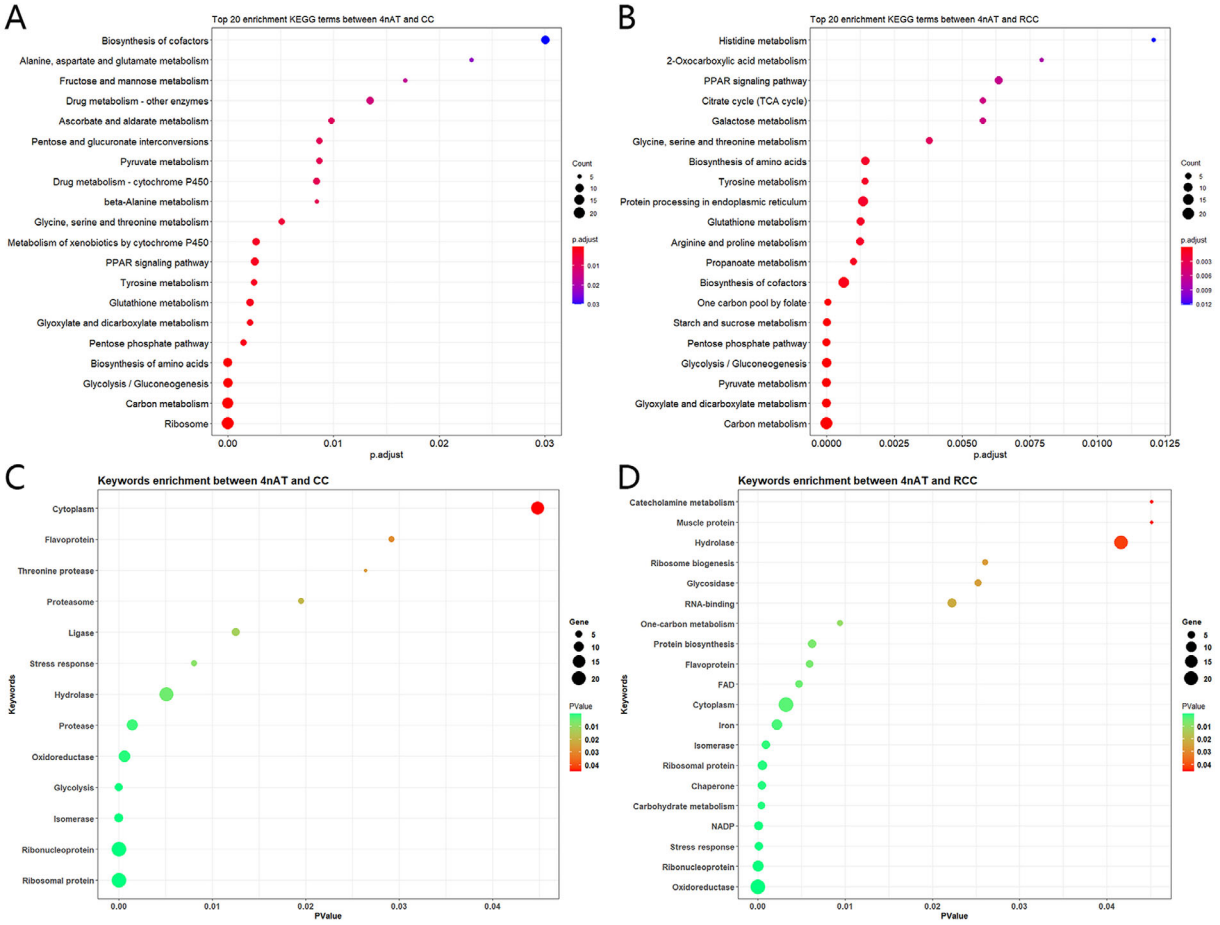

FIGURE 3 The results of pathway analysis. (A) The enriched pathways based on DEPs between CC and 4nAT. (B) The KEGG terms enriched based on DEPs between RCC and 4nAT. (C) The enriched keywords terms based on DEPs between CC and 4nAT. (D) The enriched keywords terms based on DEPs between RCC and 4nAT. The results of pathway analysis showed that the DEPs mainly enriched in metabolism related terms. CC, common carp; DEP, differentially expressed protein; 4nAT, allotetraploid; RCC, red crucian carp.

protein–protein interaction (PPI) and pathway analysis. PPI and pathway analysis could cover more information in a topological and systematical level. The FDR < 0.05 were set as statistical cutoff line in pathway enrichment analysis. The STRING database was used to evaluate and disseminate information of protein–protein association [15, 16]. Subsequently, some plugins such as ClueGO [17] and STRING [18] based on Cytoscape were used for network and subnetwork reconstruction, visualization, and analysis [19, 20]. The organism which we chose was Danio rerio. The pV correction method was Bonferroni step down, and the kappa score was 0.4 in ClueGO. We set the default parameter from Cytoscape with *p* value <0.05. Proteins without any interactions with other proteins were removed from the network. Network reconstruction and visualization based on DEPs allowed us to better understand the biological function of DEPs in a direct way.

In this study, a total of 2502 proteins were identified and quantified in all the samples from 4nAT, CC, and RCC based on SWATH-MS. For data quality control, only those proteins identified and quantified in triplicated runs were used for further analysis. Among them, 2000, 2124, and 2090 proteins were identified and qualified confidently in liver samples from RCC, CC, and 4nAT, respectively (Figure 1A).

Comparative proteomic analysis among 4nAT and its parents was performed to identify significant DEPs. A total of 258 proteins were differentially expressed between 4nAT and its male parents according to the significant statistical cutoff line: *p*-value < 0.05, ratio > 2, or <0.5. Among them, 123 proteins were upregulated and 135 proteins were downregulated between 4nAT and CC (Figure 1B). Meanwhile, 276 proteins were differentially expressed between 4nAT and RCC, and among them, 127 are upregulated proteins and 149 are downregulated proteins (Figure 1C). A total of 41 proteins were identified as significant DEPs between 4nAT and RCC, 4nAT and CC (Figure 1D).

All the DEPs were imported into R for functional annotation and enrichment analysis. The *p*-value < 0.01 and *q*-value < 0.05 were set as the significant statistical cutoff line. All the DEPs between 4nAT and CC were annotated into 124 GO terms (Table S1) based on GO database. Meanwhile, the DEPs identified between 4nAT and RCC were annotated into 198 GO terms (Table S1). Most of the DEPs were enriched in metabolism and catabolic-related pathways and terms as expected. With the integrative consideration of biological function and statistical results, we highlighted several GO terms involved in liver catabolism

TABLE 1 The KEGG terms enriched between 4nAT and CC

Term	Gene count	*p* value	Significant genes
Ribosome	24	3.30E-13	*RPL17*, *RPS2*
Biosynthesis of antibiotics	26	1.20E-09	*LDHBA*, *OAT*
Carbon metabolism	19	2.98E-09	*ACSS1*, *G6PD*
Glycolysis/gluconeogenesis	13	3.41E-07	*ACSS1*, *LDHBA*
Metabolic pathways	56	2.56E-05	*LDHBA*, *ACSBG1*, *OAT*
Biosynthesis of amino acids	11	5.72E-05	*IDH1*, *PGAM2*
Pentose phosphate pathway	6	1.18E-03	*PGM2*, *G6PD*
Tyrosine metabolism	6	1.58E-03	*GOT1*
Glutathione metabolism	7	3.08E-03	*G6PD*, *IDH1*
Pyruvate metabolism	6	5.18E-03	*ACSS1*, *LDHBA*
Glycine, serine and threonine metabolism	6	6.30E-03	*PGAM2*
PPAR signaling pathway	7	8.79E-03	*ACSBG1*
Beta-alanine metabolism	5	9.50E-03	*ACADM*
Glyoxylate and dicarboxylate metabolism	5	0.01	*GLULB*, *HAO2*
Metabolism of xenobiotics by cytochrome P450	5	0.02	*GSTR*
Alanine, aspartate and glutamate metabolism	5	0.02	*GLULB*, *GOT1*
Fructose and mannose metabolism	5	0.02	*ALDOCA*
Fatty acid degradation	5	0.03	*ACSBG1*
Cysteine and methionine metabolism	5	0.03	*LDHBA*
Propanoate metabolism	4	0.03	*ACSS1*, *LDHBA*

Abbreviations: CC, common carp; 4nAT, allotetraploid.

and metabolism such as "coenzyme metabolic process," "cellular amino acid metabolic process," "small molecule catabolic process," and "carboxylic acid catabolic process" (Figure 2A, B). The results indicated that both down- and upregulated DEPs between 4nAT and CC (Figure 2C), and most of upregulated DEPs between 4nAT and RCC (Figure 2D) are closely associated with multiple metabolism processes.

The DEPs between 4nAT and CC were enriched in 45 pathways (Table S2), and most of the pathways were related to metabolism (Figure 3A, Table 1). Although the DEPs between RCC and 4nAT were enriched in 49 pathways (Table S2), and most of them are associated with metabolism process as well (Figure 3B, Table 2). PPI network showed that DEPs mainly enriched in metabolism-related pathways including amino acid metabolism, pentose phosphate pathway, glycolysis/gluconeogenesis, citrate cycle, drug metabolism, and fatty acid metabolism (Tables 1 and 2).

Functional annotation and enrichment analysis were performed based on 41 overlapped DEPs between 4nAT and CC, 4nAT and RCC. There were 10 GO terms and seven KEGG pathways related to metabolism. There were three genes *LDHBA*, *OAT*, and *ACSBG1* with different expression level in 4nAT compare to its parents RCC and CC. *ACSBG1* and *OAT* were upregulated in 4nAT compared with CC, and they were downregulated in 4nAT compared with RCC. And *LDHBA* was downregulated in 4nAT compared with CC, and it was upregulated in 4nAT compared with RCC. Thus, *OAT* and *ACSBG1* may inhibit the metabolism and growth rate, and *LDHBA* may promote metabolism and growth rate in freshwater fish according to our results.

Significant DEPs were imported into STRING for PPI and network analysis. The network constructed based on DEPs between 4nAT and CC contains 254 nodes and 1198 edges which the average clustering coefficient is 0.442, while another network contains 269 nodes and 1049 edges based on DEPs between 4nAT and RCC (p-value $< 1 \times 10^{-16}$). The result suggested that DEPs between 4nAT and CC, 4nAT and RCC were significantly enriched in 18 and 31 functional keywords including "glycolysis," "oxidoreductase," and "hydrolase" (Tables 3 and 4). Most of the terms are closely associated with well-known biological functions of liver (Figure 3C, D). Network reconstruction and visualization showed that DEPs between 4nAT and CC were enriched in metabolism-related pathways such as glycolysis/gluconeogenesis, pentose phosphate pathway, etc. (Figure 4A). PPAR pathway which *ACSBG1* participated in also played vital role in glycose and lipid metabolism (Figure 5A). Similarly, the DEPs between 4nAT and RCC played crucial roles in pathways such as TCA cycle, pentose phosphate pathway, pyruvate metabolism, etc. (Figure 4B). The metabolism of pyruvate, metabolism and biosynthesis of amino acids, and others metabolism-related pathways were directly regulated by these DEPs including *ACSBG1*, *LDHBA*, and *OAT* (Figure 5).

Quantitative proteomics based on SWATH-MS with favorable quantitative characteristics has enabled the study of systematic PPIs in living body [21–23]. Pathway and network analysis based on quantitative proteomics has been proved to be an efficient technology in deciphering the underlying molecular mechanisms of differential expressed proteins and genes [24]. Since amino acid sequences matching for protein

TABLE 2 The KEGG terms enriched between 4nAT and RCC

Term	Gene count	*p* value	Significant genes
Biosynthesis of antibiotics	31	3.06E-14	*LDHBA*, *IDH1*
Metabolic pathways	68	1.15E-11	*LDHBA*, *ACSBG1*, *OAT*
Carbon metabolism	19	1.15E-09	*ACSS1*, *GP6D*
Glyoxylate and dicarboxylate metabolism	9	1.01E-06	*HAO2*, *MDH2*
Pyruvate metabolism	9	8.54E-06	*ACSS1*, *LDHBA*
Pentose phosphate pathway	7	9.56E-05	*G6PD*, *PGM2*
Glycolysis/gluconeogenesis	9	4.22E-04	*ACSS1*, *LDHBA*
Protein processing in endoplasmic reticulum	14	5.30E-04	*CALR3A*, *GANAB*
One carbon pool by folate	5	8.68E-04	*SHMT2*, *ATIC*
Starch and sucrose metabolism	6	1.06E-03	*PGM2*, *PGM1*
Glutathione metabolism	7	2.31E-03	*G6PD*, *IDH1*
Biosynthesis of amino acids	8	4.44E-02	*SHMT2*, *IDH1*
Propanoate metabolism	5	4.78E-03	*ACSS1*, *LDHBA*
Glycine, serine and threonine metabolism	6	3.94E-03	*SHMT2*, *GRHPRB*
Tyrosine metabolism	5	8.69E-03	*GOT1*, *FAH*
Citrate cycle (TCA cycle)	5	8.69E-03	*IDH1*, *MDH2*
2-Oxocarboxylic acid metabolism	4	0.01	*GOT1*, *IDH1*
Arginine and proline metabolism	6	0.01	*GOT1*, *OAT*
Cysteine and methionine metabolism	5	0.02	*GOT1*, *LDHBA*
Galactose metabolism	4	0.04	*PGM2*, *GLA*
Ribosome	8	0.04	*RPL35*, *RPS12*

Abbreviations: 4nAT, allotetraploid; RCC, red crucian carp.

TABLE 3 The keywords enrichment between 4nAT and CC

Terms	Gene count	*p* value
Ribosomal protein	22	8.60E-20
Ribonucleoprotein	22	8.54E-16
Isomerase	8	5.63E-06
Glycolysis	6	1.46E-05
Oxidoreductase	13	6.04E-04
Protease	12	1.48E-03
Hydrolase	20	5.10E-03
Stress response	4	8.08E-03
Ligase	6	0.01
Proteasome	4	0.02
Threonine protease	3	0.03
Flavoprotein	4	0.03
Cytoplasm	17	0.04

Abbreviations: CC, common carp; 4nAT, allotetraploid.

identification depends on an established database, high-throughput proteomics profiling in non-model animal including freshwater fish is relatively lacking. The purpose of our study is to elaborate the molecular mechanism of proteins associated with metabolism and growth rate in 4nAT and its parents. Quantitative proteomics integrated with multiple bioinformatics analysis such as gene functional annotation and enrichment analysis, pathway and network analysis revealed that a series of significant DEPs including *LDHBA*, *OAT*, *ACSBG1*, *IDH1*, *PGAM2*, and *G6PD* played multiple roles in metabolism-related processes and further regulating the growth rate of freshwater fishes. By using an integrative database consisting of homemade and public transcriptomic and proteins sequence data, the proteome draft map established in this study may serve as a significant reference for further studies in fresh fishes.

A total of 41 proteins were identified as overlapped DEPs in both comparative groups (4nAT and CC, 4nAT and RCC) (Figure 1D). The growth rate of 4nAT was slower than its male parents CC, and it was faster than its female parents RCC. The proteins with different expression trend may contribute to the differences of growth rate between 4nAT and its parents. Among these 41 proteins, there are 14 proteins with different expression trend which contributed to differences of growth rate between 4nAT and its parents (Figure 6). These proteins played vital roles in regulating metabolism processes in a positive or negative way. *LDHBA*, *OAT*, and *ACSBG1* were three of the 14 proteins which could regulate metabolism in positive and negative ways, respectively. Herein, we will mainly discuss the molecular mechanisms of DEPs *LDHBA*, *OAT*, and *ACSBG1* in regulating metabolism and growth rate.

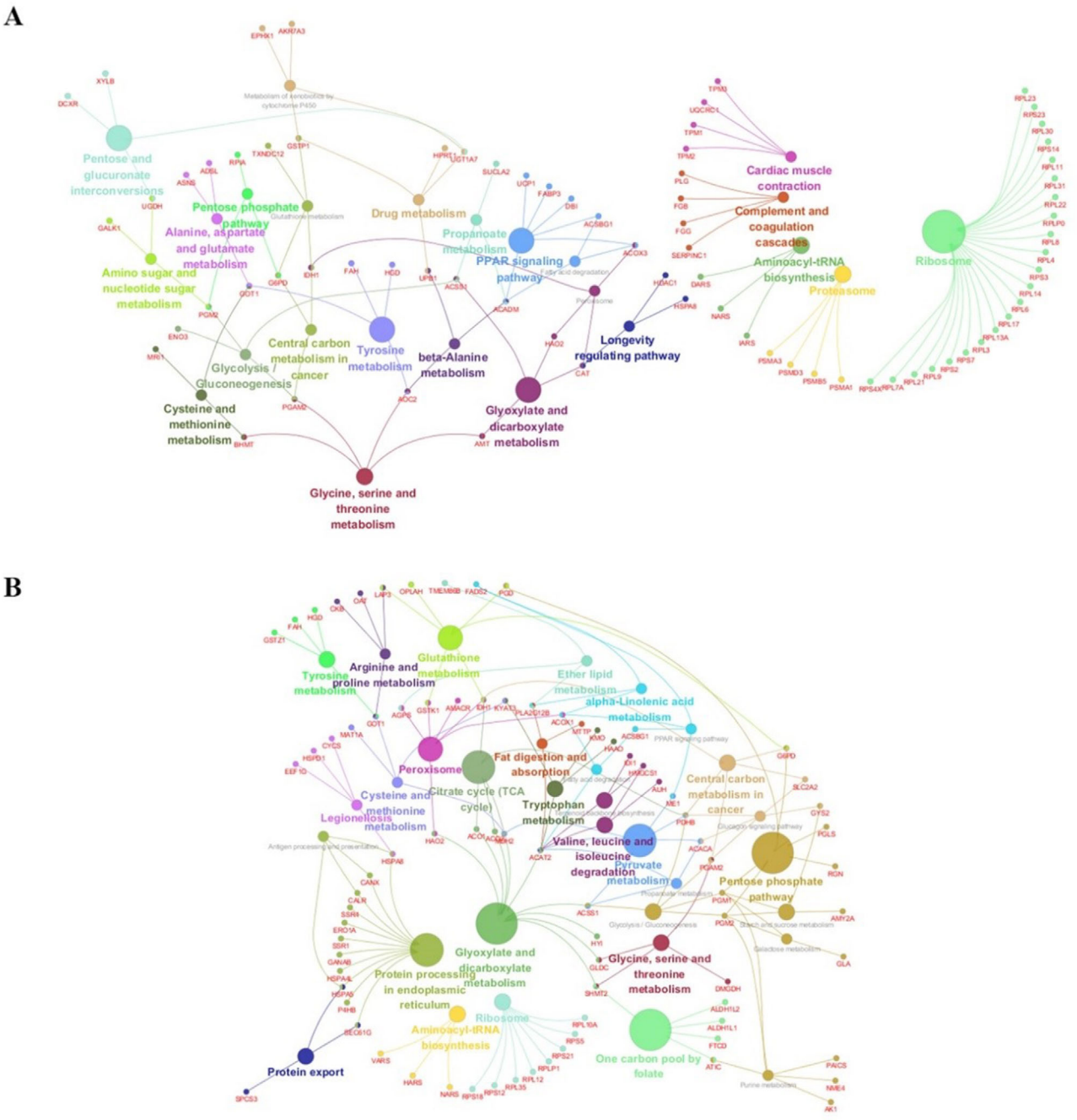

FIGURE 4 The results of network analysis and reconstruction. (A) The PPI network based on DEPs between CC and 4nAT. (B) The PPI network based on DEPs between RCC and 4nAT. The network showed that *ACSBG1*, *LDHBA*, and *OAT* could regulate vital metabolism processes in these three fishes. These networks were constructed by ClueGO. CC, common carp; DEP, differentially expressed protein; 4nAT, allotetraploid; PPI, protein–protein interaction; RCC, red crucian carp.

LDHBA (L-lactate dehydrogenase B-A chain) is a protein involved in the sub-pathway which synthesizes lactate from pyruvate [25, 26]. This protein catalyzes the pyruvate generated from glycolysis/gluconeogenesis to lactate. Compared to CC, *LDHBA* was downregulated in 4nAT. And it was upregulated in 4nAT compared with RCC. The protein expression level in 4nAT and its parents were consistent with the phenotype growth rate in these three fishes, respectively. It suggested that *LDHBA* may promote metabolism and growth rate through pyruvate metabolism pathways.

OAT (*ENSDARG00000078425*) encodes an enzyme ornithine aminotransferase that is a key enzyme in the pathway to convert arginine and ornithine into glutamate [27]. It controls the L-ornithine level by catalyzing the transfer of the delta-amino group of L-ornithine to 2-oxoglutarate in tissue. And the products of the reaction are L-glutamate-gamma-semialdehyde and L-glutamate [28]. Glutamine not

TABLE 4 The keywords enrichment between 4nAT and RCC

Terms	Gene count	*p* value
Oxidoreductase	23	3.44E-10
Ribonucleoprotein	11	6.08E-05
Stress response	6	8.01E-05
NADP	7	1.10E-04
Carbohydrate metabolism	5	4.17E-04
Chaperone	6	4.53E-04
Ribosomal protein	8	5.02E-04
Isomerase	6	9.01E-04
Iron	10	2.18E-03
Cytoplasm	22	3.18E-03
FAD	5	4.70E-03
Flavoprotein	5	5.93E-03
Protein biosynthesis	6	6.19E-03
One-carbon metabolism	3	9.41E-03
RNA-binding	7	0.02
Glycosidase	4	0.02
Ribosome biogenesis	3	0.02
Hydrolase	18	0.04
Muscle protein	2	0.04
Catecholamine metabolism	2	0.04

Abbreviations: 4nAT, allotetraploid; RCC, red crucian carp.

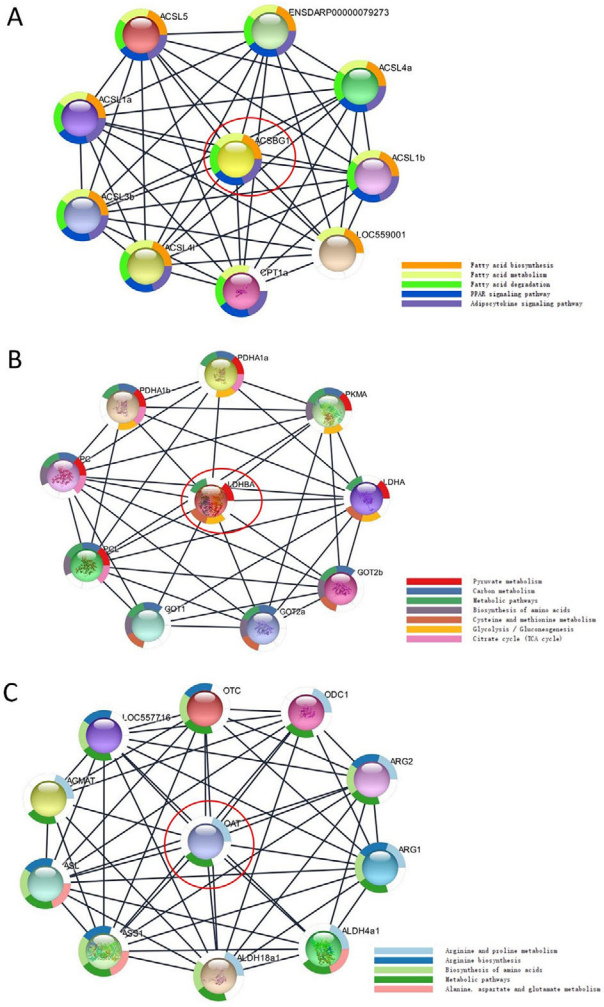

FIGURE 5 Sub-networks of three candidate DEPs. (A, B, C) The subnetworks of three candidate genes including *ACSBG1*, *LDHBA*, and *OAT*, respectively. They played significant roles in metabolism-related pathways including "fatty acid metabolism," "PPAR pathway," "pyruvate metabolism," "carbon metabolism," and "amino acid metabolism and biosynthesis." *ACSBG1*, *LDHBA*, and *OAT* were marked with red circle.

only plays essential role in inter-organ nitrogen transport but also is the preferential substrate for rapidly dividing cells like lymphocytes, endothelial cells, and fibroblasts [29]. Arginine is involved in the immune response as a substrate in nitric oxide synthesis, and regulates protein synthesis by stimulating secretion of growth hormone and insulin [30]. Glutamine and arginine share a metabolic pathway in which the enzyme *OAT* could play a pivotal role [31]. The different expression level of *OAT* between 4nAT and its parents contributes to the different growth rate and immune response. Different expression level of *OAT* regulates metabolism of arginine and glutamine which is consistent with its phenotypic growth rate. And the results suggest that *OAT* may inhibit the metabolism and growth rate through arginine and proline metabolism.

ACSBG1 (acyl-CoA synthetase bubblegum family member 1) is a protein which possesses long-chain acyl-CoA synthetase activity. This protein could catalyze the conversion of fatty acids such as long-chain and very long-chain fatty acids to their active form acyl-CoAs for both synthesis of cellular lipids, and degradation via beta-oxidation [32–34]. Acyl-CoA synthetases play a vital role in fatty acid metabolism, because it provides activated substrates for fatty acid catabolic and anabolic pathways. *ACSBG1* can also activate diverse saturated, monosaturated, and polyunsaturated fatty acids [34]. As our results displayed that *ACSBG1* was upregulated in 4nAT between 4nAT and CC. And it was downregulated in 4nAT between 4nAT and RCC. Metabolism in 4nAT which is consistent with the phenotype growth rate is balanced between its parents CC and RCC. *ACSBG1* may inhibit the metabolism and growth rate in these freshwater fishes through PPAR signaling pathway.

Furthermore, other metabolism-related proteins such as *G6PD*, *ACSS1*, *IDH1*, and *PGAM2* were differentially expressed. All these proteins are involved in central metabolism pathways such as pentose phosphate, glycolysis, and gluconeogenesis. And these proteins directly regulate metabolism among 4nAT, RCC, and CC which contribute to the difference in rate of growth and metabolism.

Based on SWATH-MS, quantitative proteomics profiling was performed between 4nAT and its parents RCC and CC, respectively. Compare to CC, *OAT* and *ACSBG1* were upregulated and *LDHBA* was downregulated in 4nAT. Although *OAT* and *ACSBG1* were downregulated and *LDHBA* was upregulated in 4nAT compared with RCC. Enrichment analysis and network analysis suggested *ACSBG1*, *OAT* and *LDHBA*

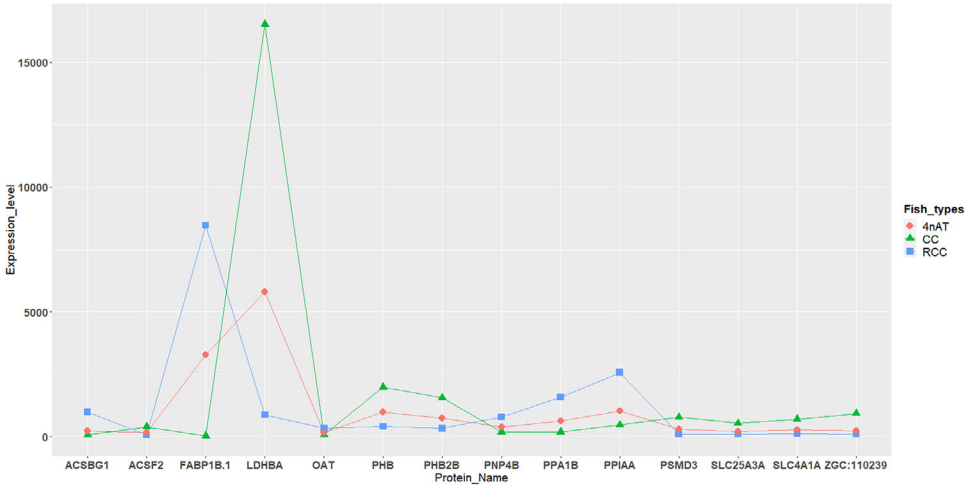

FIGURE 6 The expression level of these 14 proteins from 41 overlapped proteins in 4nAT, RCC, and CC, respectively. CC, common carp; 4nAT, allotetraploid; RCC, red crucian carp.

played vital roles in metabolism in a synergistic and direct way. Our study demonstrated that *ACSBG1*, *OAT*, and *LDHBA* contribute to the metabolism and growth in 4nAT in a synergetic way with metabolic genes and pathways. And we firstly established the comprehensive and quantitative proteomics knowledge base. It may serve as an indispensable reference for further studies on the molecular mechanisms underlying polyploidy breeding and individual differences.

ACKNOWLEDGEMENTS

This research was financially supported by the National Natural Science Foundation of China (grant no. 31873038), Outstanding Youth Foundation of Hunan Education Department (19B356) and the Natural Science Foundation of Hunan Province (grant no.2019JJ50396).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Yong Zeng and Shaojun Liu contributed to the conceptual framework of the study and experimental design. Yude Wang, Li Ren, Chun Zhang, Min Tao, and Kaikun Luo contributed to sample acquisition and preparation. Yujie Yan, Junting Wang, Xiaoping Dong, and Yong Zeng participated in data collection and data analysis. Yujie Yan, Junting Wang, Xiaoping Dong, Yisheng Cai, and Yong Zeng drafted and revised the manuscript.

DATA AVAILABILITY STATEMENT

The proteomics data have been deposited to ProteomeXchange Consortium via the iProX partner repository with the dataset identifiers IPX0002621000.

ORCID

Yong Zeng https://orcid.org/0000-0003-0119-2852

REFERENCES

1. Liu, S. (2010). Distant hybridization leads to different ploidy fishes. *Science China Life Sciences*. *53*(4), 416–425.
2. Duan, W., Qin, Q., Chen, S., Liu, S., Wang, J., Zhang, C., Sun, Y., & Liu, Y. (2007). The formation of improved tetraploid population of red crucian carp A common carp hybrids by androgenesis. *Science China Series C: Life Sciences*, *50*(6), 753–761.
3. Trefts, E., Gannon, M., & Wasserman, D H. (2017). The liver. *Current Biology*, *27*(21), R1147–R1151.
4. Kubes, P., & Jenne, C. (2018). Immune responses in the liver. *Annual Review of Immunology*, *36*, 247–277.
5. Jones, J. G. (2016). Hepatic glucose and lipid metabolism. *Diabetologia*, *59*(6), 1098–1103.
6. Oosterveer, M H., & Schoonjans, K. (2014). Hepatic glucose sensing and integrative pathways in the liver. *Cellular and Molecular Life Sciences*, *71*(8), 1453–1467.
7. Causey, D. R., Pohl, M. A. N., Stead, D. A., Martin, S. A. M., Secombes, C. J., & Macqueen, D. J. (2018). High-throughput proteomic profiling of the fish liver following bacterial infection. *BMC Genomics*, *19*(1), 719.
8. Ludwig, C., Gillet, L., Rosenberger, G., Amon, S., Collins, B. C., & Aebersold, R. (2018). Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. *Molecular Systems Biology*, *14*(8), e8126.
9. Egertson, J. D., Maclean, B., Johnson, R., Xuan, Y., & Maccoss, M. J. (2015). Multiplexed peptide analysis using data-independent acquisition and Skyline. *Nature Protocols*, *10*(6), 887–903.
10. Hu, A., Noble, W S., & Wolf-Yadlin, A. (2016). Technical advances in proteomics: New developments in data-independent acquisition. *F1000Res*, *5*, 419. https://doi.org/10.12688/f1000research.7042.1.
11. Wiśniewski, J. R., Zougman, A., Nagaraj, N., & Mann, M. (2009). Universal sample preparation method for proteome analysis. *Nature Methods*, *6*(5), 359–362.
12. Yu, G., Wang, L.-G., Han, Y., & He, Q.-Y., (2012). clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS*, *16*(5), 284–287.
13. Huang da, W., Sherman, B. T., & Lempicki, R A. (2009). Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research*, *37*(1), 1–13.
14. Huang da, W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, *4*(1), 44–57.

15. Szklarczyk, D., Morris, J. H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., Santos, A., Doncheva, N. T., Roth, A., Bork, P., Jensen, L J., & Von Mering, C. (2017). The STRING database in 2017: Quality-controlled proteina protein association networks, made broadly accessible. *Nucleic Acids Research*, *45*(D1), D362–D368.
16. Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., Jensen, L. J., & Mering, C. V. (2019). STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*, *47*(D1), D607–D613.
17. Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W.-H., Pagès, F., Trajanoski, Z., & Galon, J. (2009). ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics*, *25*(8), 1091–1093.
18. Doncheva, N. T., Morris, J. H., Gorodkin, J., & Jensen, L. J. (2019). Cytoscape StringApp: Network analysis and visualization of proteomics data. *Journal of Proteome Research*, *18*(2), 623–632.
19. Cline, M. S., Smoot, M., Cerami, E., Kuchinsky, A., Landys, N., Workman, C., Christmas, R., Avila-Campilo, I., Creech, M., Gross, B., Hanspers, K., Isserlin, R., Kelley, R., Killcoyne, S., Lotia, S., Maere, S., Morris, J., Ono, K., Pavlovic, V.,… Bader, G. D. (2007). Integration of biological networks and gene expression data using Cytoscape. *Nature Protocols*, *2*(10), 2366–2382.
20. Shannon, P. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research*, *13*(11), 2498–2504.
21. Collins, B. C., Hunter, C. L., Liu, Y., Schilling, B., Rosenberger, G., Bader, S. L., Chan, D. W., Gibson, B. W., Gingras, A.-C., Held, J. M., Hirayama-Kurogi, M., Hou, G., Krisp, C., Larsen, B., Lin, L., Liu, S., Molloy, M. P., Moritz, R. L., Ohtsuki, S., … Aebersold, R. (2017). Multi-laboratory assessment of reproducibility, qualitative and quantitative performance of SWATH-mass spectrometry. *Nature Communications*, *8*(1), 291.
22. Collins, B. C., Gillet, L. C., Rosenberger, G., Röst, H. L., Vichalkovski, A., Gstaiger, M., & Aebersold, R. (2013). Quantifying protein interaction dynamics by SWATH mass spectrometry: Application to the 14-3-3 system. *Nature Methods*, *10*(12), 1246–1253.
23. Lambert, J.-P., Ivosev, G., Couzens, A. L., Larsen, B., Taipale, M., Lin, Z.-Y., Zhong, Q., Lindquist, S., Vidal, M., Aebersold, R., Pawson, T., Bonner, R., Tate, S., & Gingras, A.-C. (2013). Mapping differential interactomes by affinity purification coupled with data-independent mass spectrometry acquisition. *Nature Methods*, *10*(12), 1239–1245.
24. Khatri, P., Sirota, M., & Butte, A. J. (2012). Ten years of pathway analysis: Current approaches and outstanding challenges. *PLoS Computational Biology*, *8*(2), e1002375.
25. Li, Y.-J., Tsoi, S. C.-M., Mannen, H., & Shoei-lung Li, S. (2002). Phylogenetic analysis of vertebrate lactate dehydrogenase (LDH) multigene families. *Journal of Molecular Evolution*, *54*(5), 614–624.
26. Lucitt, M. B., Price, T. S., Pizarro, A., Wu, W., Yocum, A. K., Seiler, C., Pack, M. A., Blair, I. A., Fitzgerald, G. A., & Grosser, T. (2008). Analysis of the zebrafish proteome during embryonic development. *Molecular & Cellular Proteomics*, *7*(5), 981–994.
27. Zielonka, M., Breuer, M., Okun, J. G., Carl, M., Hoffmann, G. F., & Kölker, S. (2018). Pharmacologic rescue of hyperammonemia-induced toxicity in zebrafish by inhibition of ornithine aminotransferase. *PLoS One*, *13*(9), e0203707.
28. Seiler, N. (2000). Ornithine aminotransferase, a potential target for the treatment of hyperammonemias. *Current Drug Targets*, *1*(2), 119–154.
29. Curi, R., Lagranha, C. J., Doi, S. Q., Sellitti, D. F., Procopio, J., Pithon-Curi, T. C., Corless, M., & Newsholme, P. (2005). Molecular mechanisms of glutamine action. *Journal of Cellular Physiolgy*, *204*(2), 392–401.
30. Barbul, A., Wasserkrug, H. L., Yoshimura, N., Tao, R., & Efron, G. (1984). High arginine levels in intravenous hyperalimentation abrogate post-traumatic immune suppression. *Journal of Surgical Research*, *36*(6), 620–624.
31. Ventura, G., Moinard, C., Segaud, F., Le Plenier, S., Cynober, L., & De Bandt, J. P. (2010). Adaptative response of nitrogen metabolism in early endotoxemia: Role of ornithine aminotransferase. *Amino Acids*, *39*(5), 1417–1426.
32. Pei, Z., Oey, N. A., Zuidervaart, M. M., Jia, Z., Li, Y., Steinberg, S. J., Smith, K. D., & Watkins, P. A. (2003). The acyl-CoA synthetase a bubbleguma (Lipidosin). *Journal of Biological Chemistry*, *278*(47), 47070–47078.
33. Ohkuni, A., Ohno, Y., & Kihara, A. (2013). Identification of acyl-CoA synthetases involved in the mammalian sphingosine 1-phosphate metabolic pathway. *Biochemical and Biophysical Research Communications*, *442*(3-4), 195–201.
34. Steinberg, S. J., Morgenthaler, J., Heinzer, A. K., Smith, K. D., & Watkins, P. A. (2000). Very long-chain Acyl-CoA synthetases. *Journal of Biological Chemistry*, *275*(45), 35162–35169.

SUPPORTING INFORMATION

Additional supporting information may be found online https://doi.org/10.1002/pmic.202100115 in the Supporting Information section at the end of the article.

How to cite this article: Yan, Y., Wang, J., Dong, X., Cai, Y., Wang, Y., Ren, Li, Zhang, C., Tao, M., Luo, K., Zeng, Y., & Liu, S. (2022). Quantitative proteomic analysis of hepatic tissue in allotetraploid hybridized from red crucian carp and common carp identified crucial proteins and pathways associated with metabolism and growth rate. *Proteomics*, *22*, e2100115. https://doi.org/10.1002/pmic.202100115