Open Access

SCHOLARENA

Scholarena Journal of Biotechnology

RESEARCH ARTICLE

ISSN: 2375-6713

Protein Composition of Foot Muscles of the Razor Clam *Sinonovacula constricta* based on Proteomics Analysis

Zhi-Hao Zhou, Guo-Fang Jiang*, Bo Fan and Fang Hong

College of Ocennology and Food Science, Quanzhou Normal School University, Quanzhou 362000, China

***Corresponding Author:** Guo-Fang Jiang, College of Ocennology and Food Science, Quanzhou Normal School University, Quanzhou 362000, China, Tel.: 15060936569, E-mail: 370718493@qq.com

Citation: Zhi-Hao Zhou, Guo-Fang Jiang, Bo Fan, Fang Hong (2024) Protein Composition of the Foot Muscles of the Razor Clam Sinonovacula Constricta Based on Proteomics Analysis, SAJ Biotechnol 9: 103

Abstract

Proteins contained in the foot of the razor clam Sinonovacula constricta were analyzed and identified by non-labeled quantitative technology, and nutritional and functional components were explored in this study. The identified proteins were compareed to relevant databases for annotation, protein properties were described in the foot muscles of S. constricta, and the coordinated function of different proteins in their biological behavior were explored. The results showed that a total of 1,674 proteins were identified in the muscle tissue of the feet of S. constricta. The biological processes of protein annotation were mainly concentrated in cellular and single-cell processes, and the cell composition was mainly concentrated in the cellular and cellular parts. The molecular function was manifested as binding and catalytic activity. There are 1,190 proteins annotated to the GO database, 1,350 proteins annotated to the KEGG database, and 1,500 proteins annotated to the KOG database. Among the identified proteins, actin and myosin have the highest expression levels. In addition, among the identified amino acids and their derivatives of S. constricta, the composition of standard amino acids is complete. Meanwhile, Tryptophan, which is prone to decomposition and destruction under gas phase detection, was measured in this study.

Keywords: *Sinonovacula constricta*; Foot Muscles; Non Labeled Proteomics; Shellfish Food Products; Nutritional Ingredient

Introduction

In recent years, with the improvement of proteomics technology and lower testing costs, proteomic-based techniques have been used in a variety of research fields, including the detection of diagnostic markers, vaccine candidates, pathogenic mechanisms, changes in response expression patterns of various signals, and functional protein pathways that explain the immune response of organisms under viral infection. Integrate proteome data to obtain more comprehensive and accurate data compared to traditional methods. Proteomics is often used to analyze the protein composition in aquatic products, screen relevant marker proteins at the protein level, and apply it to the mechanism of protein action in food quality changes. It analyzes meat quality related issues from different perspectives [1], providing a new perspective for studying meat quality and nutrient composition.

In terms of meat quality of marine animals, proteomics technology is mainly used to study the mechanism of meat product formation, improve meat quality, and detect meat quality [2]. Shellfish farming is an important economic activity that provides society with a valuable source of food. Selected tandem mass spectrometry (MS/MS) ion monitoring (SMIM) is the most suitable scanning mode to detect known peptides in complex samples when an ion-trap mass spectrometer is the instrument used for the analysis [3]. Campos et al. [4] compared different protein extraction protocols commonly used in two-dimensional gel electrophoresis (2DE) research and select the most suitable for the analysis of gill and digestive gland proteomes from the marine mussel Mytilus galloprovincialis, and concluded that important protein markers for contaminant and quality assessment of shellfish food products can be analysed using 2DE. The crude extracts from the edible parts of two species of clams were analyzed proteomically by two-dimensional gel electrophoresis combined with high performance liquid chromatography-tandem mass spectrometer, and the biomarkers useful for the identification and nutritional characteristics of these two isolates were selected [5].

The razor clam *Sinonovacula constricta*, belonging to Pharidae of Adapedonta in the class Bivalvia [6], is one of the important economic mollusks in China. It has a wide distribution range and abundant biological resources. Chen's team has applied two-dimensional polyacrylamide gel electrophoresis and image analysis techniques to search for proteomic differences in the hemolymph of S. constricta under low salt environmental pressure, verifying the key role of hemolymph in the defense system of S. constricta [7]; has analyzed the changes in the proteome of hemolymph under high temperature stress, the differential expression of hemolymph proteins in *S. constricta* under high temperature environment, and speculated on the resistance mechanism of *S. constricta* under high temperature stress [8], and has also explored the changes in the proteome of the hemolymph of *S. constricta* under parasitic pressure, searched for the resistance mechanism of the hemolymph of *S. constricta*, and clarified the role of hemolymph in the body's defense against parasitic pressure [9].

Clam meat has the characteristics of low fat and high protein. Protein unlabeled quantification technology and liquid chromatography-mass spectrometry (LC-MS) were used for directly analyzing the enzymatically hydrolyzed peptide segments of the foot muscle proteins of *S. constricta* in this study. The mass spectrometry data generated during large-scale protein identification were analyzed, and the signal intensity of corresponding peptide segments in different samples was compared to relatively quantify the corresponding proteins of the peptide segments

Materials and Methods

Materials and Reagents

The razor clam used in this article was purchased from the Quanzhou Farmers' Market in Fujian Province, China. The main reagents used in this experiment include sample lysate, DTT, iodoacetamide (IAA), acetone, trypsin (Promega), Buffer A, Buffer B, and Solvent A. The instruments and equipment used in this study included a grinder, ultra-low temperature refrigerator, rapid dye removal instrument, vacuum concentration instrument, liquid chromatography, and Orbitrap Lumis mass spectrometer (Thermo Fisher Scientific, MA, USA).

Sample Preparation

Lysate (7M urea, 2% SDS, 0.1% PMSF, 65 mM DTT) was added to the samples of *S. constricta* and ultrasonic lysis was performed. The homogenate was centrifuged at 15,000 rpm for 15 mins at 4 ℃, and the supernatant was collected. The protein concentration was determined by BCA protein quantitative kit. DTT and iodoacetamide were used to break the disulfide bond and reduce the alkylation of the protein, so as to fully enzymolize the protein. 50 μg protein was diluted to 50 μL, then 1 μL of 1M DTT was added and incubated at 55°C for 1h. 5 μ L of 1M iodoacetamide (IAA) was added, and avoid light at room temperature for 1 h. 300 μL pre-cooled acetone was added to precipitate for 2 h. Finally the proteins were digested with sequence-grade modified trypsin (Promega, Madison, WI) at a substrate/enzyme ratio of 50:1 (w/w) at 37 °C for 16 h.

High pH Reverse Phase Separation

The peptide mixture was re-dissovled in the buffer A (buffer A: 20mM ammonium formate in water, pH10.0, adjusted with ammonium hydroxide), and then fractionated by high pH separation using Ultimate 3000 system (ThermoFisher scientific, MA, USA) connected to a reverse phase column (XBridge C18 column, 4.6mm x 250 mm, 5μm (Waters Corporation, MA, USA). High pH separation was performed using a linear gradient, starting from 5% B to 45% B in 40 min (B: 20mM ammonium formate in 80% ACN, pH 10.0, adjusted with ammonium hydroxide). The column was re-equilibrated at the initial condition for 15 min. The column flow rate was maintained at 1mL/min and the column temperature was maintained at 30℃. Ten fractions were collected; each fraction was dried in a vacuum concentrator for the next step.

DDA: nano-HPLC-MS/MS Analysis

The peptides were re-dissolved in 30 μL solvent A (A: 0.1% formic acid in water) and analyzed by on-line nanospray LC-MS/MS on an Orbitrap Fusion™ Lumos™ coupled to EASY-nLC 1200 system (Thermo Fisher Scientific, MA, USA). 3 μL peptide sample was loaded onto the analytical column (Acclaim PepMap C18, 75 μm x 25 cm) and separated with a 120-min gradient, from 5% to 35% B (B: 0.1% formic acid in ACN). The column flow rate was maintained at 200 nL/min with the column temperature of 40°C. The ele ctrospray voltage of 2 kV versus the inlet of the mass spectrometer was used.

The mass spectrometer was run under data dependent acquisition mode, and automatically switched between MS and MS/MS mode. The parameters was: (1) MS: scan range (m/z) = 350–1,200; resolution = 120,000; AGC target = 400,000; maximum injection time = 50 ms; Filter Dynamic Exclusion: exclusion duration = 30s; (2) HCD-MS/MS: resolution = 15,000; AGC target = 50,000; maximum injection time = 35 ms; collision energy = 32. For this Mass spectrometry detection, Guangzhou Kedio Biotechnology Co., Ltd. was entrust for testing.

Database Search, Data Analysis and Protein Function Annotation

PEAKS Studio version X (Bioinformatics Solutions Inc, Waterloo, Canada) mass spectrometry was used for analysis. PEAKS DB was used to search the protein database. PEAKS DB were searched with a fragment ion mass tolerance of 0.05 Da and a parent ion tolerance of 10 ppm. Carbamidomethylation (C) was specified as the fixed modifications. Oxidation (M), and Acetylation (Protein N-term) were specified as the variable modifications. Qualitative and quantitative analysis of protein based on search results.

Qualitative analysis of proteins is to determine whether a protein is present in a sample and to identify the type of protein. To ensure the reliability of the results, peptide FDR is required to be less than 1% to evaluate the error discovery rate (FDR). In addition, a unique peptide refers to a peptide that has been identified and only comes from a single protein sequence or a sequence from the same group, requiring a protein's uniquepetide≥1. Peptides and proteins that meet the requirements are used for subsequent analysis.

Raw data were pooled and analyzed via Spectraut X (Biognosys AG) tusing Uniprot or the provided databases. In addition, the con-

taminated sequence library was also searched to determine whether the sample was contaminated, and Trypsin enzyme digestion was setup. Search database parameters: fixed modification: Carbamidomethyl (C), variable modification: methionine oxidation. The false positive rate (FDR) at both the mother ion and peptide levels was set to 1%. Function and classification of proteins were analyzed by searching the following databases: Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KOG), Kyoto Encyclopedia of Genes and Genomes (KEGG) database, and the NR library.

Subcellular Localization Analysis

The protein's function is usually related to its subcellular localization, the ability to predict subcellular localization directly from protein sequences will be useful for inferring protein functions. Subcellular localization refers to the specific location of a protein or expression product within a cell. The software WoLFPSort was used to predict the subcellular location of the protein and study the function of the protein.

Qualitative and Quantitative Analysis of Proteins

Qualitative analysis of proteins is the process of determining the presence of proteins in a sample and identifying the types of proteins involved. To ensure the reliability of the results and evaluate the error detection rate (FDR), the peptide FDR is required to be ≤ 1%. In addition, a unique peptide refers to the identified peptide segment that has and only comes from a single protein sequence or from the same group. The protein's unique peptide is required to be ≥ 1, and the peptide segment and protein that meet the conditions are used for subsequent analysis. The peak area of the chromatogram increases with the concentration of injected peptide segments. The identified proteins are sorted in descending order according to the peak area of the chromatogram, and the top 40 proteins with larger peak area are selected for subsequent analysis.

Protein quantitative analysis is the determination of the content of each protein in a sample. Measuring the changes in protein types and abundance in organisms under different conditions can be used to study protein function.

Screening and Identification of Actin and Myosin

The corresponding amino acid sequences of actin and myosin were screened out from related species to *S. constricta* of the class Bivalve on the NCBI website. The closely related species were *Mytilus coruscus*, *M. galloprovincialis*, *Pecten maximus* and *Mizuhopecten yessoensis*. The four amino acid sequences with actin domains in *S. constricta* were compared, and PoyloSuite software was used for constructing an evolutionary tree by combining the amino acid sequences obtained from NCBI and the four amino acid sequences of *S. constricta*. The sequences screened by BLAST were compared with the amino acid sequences of the constricted razor clam protein using MEGA v11 [10], while conserved domain was used to screen amino acid sequences with actin and myosin domains. Then, the software ExPASy was used to analyze the physicochemical properties of actin based on its amino acid sequence

Results

Protein Composition and Quantification

A total of 1,674 proteins were identified in the proteomic data of *S. constricta*, among which 17,594 were identified as spectral number with the only peptide segment, 21,494 were identified as peptide segments, and 1,669 proteins identified as the same peptide segment were classified into the same group (Figure 1).

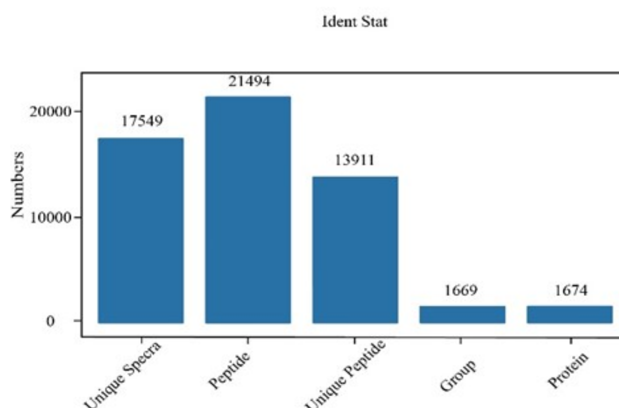

Figure 1: Statistics of protein and peptide identification results

Protein quantitative analysis is the determination of the content of each protein in a sample. Measuring the changes in protein types and abundance in organisms under different conditions can be used for studing protein function. Select the first 40 proteins with higher expression levels (Table 1) for analysis, among which actin is the highest protein content, followed by myosin, indicating a higher protein content related to exercise in the muscle tissue of the feet of *S. constricta*. In addition, calmodulin was identified with a peak area of 3,676,100,000 on the chromatogram, which may be related to its regulation of smooth muscle contraction, because tropomyosin is also expressed in the foot, and calmodulin can bind to actin and tropomyosin [17-18].

Table 1: Sequencing of Chromatographic peak area of protein

Protein ID	Chromatographic peak area	Symbol
Unigene0001481	46,031,000,000.00	act-4
Unigene0008779	33,598,000,000.00	MYH16
Unigene0005600	20,606,000,000.00	PMY
Unigene0001276	10,800,000,000.00	FLNC
Unigene0007077	8,630,500,000.00	TRO
Unigene0045759	7,644,600,000.00	cher
Unigene0024274	6,465,000,000.00	gxcB
Unigene0028708	6,238,200,000.00	unc-22
Unigene0020903	5,338,900,000.00	Smp_194770
Unigene0037878	4,917,400,000.00	Rlc-a
Unigene0022698	4,086,200,000.00	LKD
Unigene0010019	3,676,100,000.00	CAL
Unigene0033722	2,751,600,000.00	COL1A1
Unigene0001883	2,410,300,000.00	PEPCK
Unigene0056796	2,276,200,000.00	Actn
Unigene0054527	2,154,700,000.00	FIB
Unigene0040337	2,090,500,000.00	mlc-3
Unigene0060935	2,075,200,000.00	Hspg2
Unigene0008126	1,980,400,000.00	Fhl2

Unigene0058881	1,896,500,000.00	GAPDH
Unigene0036069	1,667,300,000.00	ATP2A2
Unigene0056797	1,627,700,000.00	X14
Unigene0058571	1,546,500,000.00	FBPA
Unigene0012144	1,327,100,000.00	RET
Unigene0019685	1,265,700,000.00	act-2b
Unigene0008276	1,247,600,000.00	HSPB6
Unigene0000493	949,470,000.00	PDLIM3
Unigene0028694	931,920,000.00	ENO
Unigene0005309	913,930,000.00	act
Unigene0057026	764,650,000.00	unc-22
Unigene0061939	717,600,000.00	CAL
Unigene0055388	717,320,000.00	MDH1
Unigene0020099	714,300,000.00	Hil
Unigene0000375	665,830,000.00	CYIA
Unigene0002127	661,960,000.00	let-2
Unigene0024905	614,650,000.00	PDLIM3
Unigene0003575	598,420,000.00	gelsolin
Unigene0061938	529,880,000.00	CAL
Unigene0005378	515,280,000.00	GlyP
Unigene0001582	514,450,000.00	TTN

Protein Subcellular Localization

Proteomic data in this study indicated that the proteins that play a role in the muscle tissue of the feet of *S. constricta* mainly exist in the cytoplasm and nucleus (Table 2). Classifying the identified proteins according to their functions and comparing them with related species can provide a better explanation of the functions of the proteins contained in the foot muscles of *S. constricta*. This helps to further understand the coordinated exercise of different proteins in their biological behavior, and determine the main biochemical metabolic and signal transduction pathways involved in proteins. In *S. constricta*, different proteins coordinate with each other to exercise their biological behavior [9-13]. The results of this article indicate that the actin of *S. constricta* is located in the cytoskeleton and participates in the structural composition of cells as a cytoskeleton.

GO Annotation of Proteins

GO is a method and tool that comprehensively describes the properties of various genes and gene products in organisms, explaining proteins from different aspects of biological effects. The GO annotation of proteins in the foot muscles of *S. constricta* is shown in Figure 2. In GO classification, there are 995 proteins involved in biological processes, accounting for approximately 35.60% of total proteins, with proteins mainly concentrated in biological, cellular, and organic metabolic processes. There are 974 proteins assigned to their cellular components, accounting for approximately 34.90% of the total proteins, with proteins concentrated in relation to the formation of intracellular organelles. There are 824 proteins enriched in the metabolic process, accounting for approximately 29.50% of the total protein, which are mostly related to catalytic activity, protein, organic cyclic compounds, heterocyclic compounds, and nucleic acid binding (Figure 2).

Table 2: Protein composition, subcellular localization and annotation

Protein ID	Subcellular localization	Description	Species compared
Unigene0008779	nuclear	Myosin	*Platevindex* sp. XL-2018
Unigene0045759	nuclear	Filamin	*Mizuhopecten yessoensis*
Unigene0033722	nuclear	Collagenous fiber	*Haliotis discus hannai*
Unigene0008126	nuclear	Acanthin	*Crassostrea virginica*
Unigene0056797	nuclear	LOC105341623 subtype X14	*Crassostrea gigas*
Unigene0000493	nuclear	PDZ and LIMdomain protein 3	*M. yessoensis*
Unigene0057026	cytoplasm,nuclear	Twichin	*C. virginica*
Unigene0024905	nuclear	PDZ and LIM domain protein 3 like subtype X1	*C. virginica*
Unigene0003575	nuclear	Gelsolin	*C. gigas*
Unigene0061938	nuclear	Calmodulin	*M. yessoensis*
Unigene0001582	nuclear	Connectin	*M. yessoensis*
Unigene0010019	mitochondria	Calmodulin	*Hymenolepis microstoma*
Unigene0012144	mitochondria	Retrograde protein	*C. virginica*
Unigene0061939	mitochondria	Calmodulin	*L. anatina*
Unigene0005600	cytoplasm	Paramyosin	*C. gigas*
Unigene0001276	cytoplasm	Filamin	*M. yessoensis*
Unigene0007077	cytoplasm	Tropomyosin	*S. constricta*
Unigene0028708	cytoplasm	Twichin	*C. virginica*
Unigene0024274	cytoplasm	Transgelin	*C. virginica*
Unigene0020903	cytoplasm	Arginine kinase	*Solen strictus*
Unigene0037878	cytoplasm	Myosin	*Pseudocardium sachalinense*
Unigene0022698	cytoplasm	LKD-rich protein-1	*Mytilus galloprovincialis*
Unigene0001883	cytoplasm	Phosphoenolpyruvate carboxykinase	*M. yessoensis*
Unigene0056796	cytoplasm	Actin	*C. gigas*
Unigene0040337	cytoplasm	Myosin	*Hyriopsis cumingii*
Unigene0060935	cytoplasm	Glycan core proteins	*Pomacea canaliculata*
Unigene0058881	cytoplasm	Glyceraldehyde-3-phosphate dehydrogenase	*Petromyzon marinus*
Unigene0036069	cytoplasm	Muscular/endoplasmic reticulum Ca^{2+} ATPase	*Tridacna squamosa*
Unigene0058571	cytoplasm	Fructose-1, 6-diphosphatase	*Meretrix meretrix*
Unigene0008276	cytoplasm	Heat shock protein	*S. constricta*
Unigene0028694	cytoplasm	Enolase	*C. gigas*
Unigene0055388	cytoplasm	Cytoplasmic malate dehydrogenase	*Mytilus californianus*
Unigene0020099	cytoplasm	Fusion protein	*M. yessoensis*
Unigene0000375	cytoplasm	Actin	*Monopterus albus*

Unigene0005378	cytoplasm	Phosphorylase	*Diplodon chilensis*
Unigene0001481	cytoskeleton	Actin	*C. virginica*
Unigene0005309	cytoskeleton	Actin	*M. yessoensis*
Unigene0019685	cytoskeleton	Actin	*M. meretrix*
Unigene0054527	extracellular	Collagen	*C. virginica*
Unigene0002127	extracellular	Collagen-α2	*C. virginica*

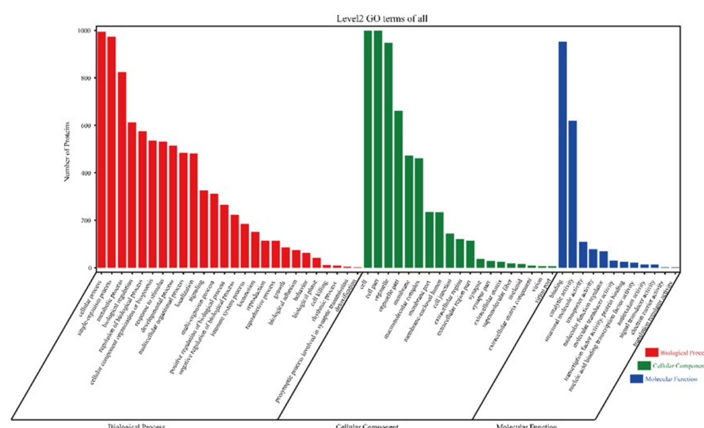

Figure 2: GO enrichment analysis

Protein KEGG Pathway Enrichment Analysis

In the KEGG pathway analysis, a total of 909 proteins were enriched in the KEGG pathway, of which 318 proteins were enriched in the metabolic pathway, accounting for approximately 18.99% of the total proteins involved in enzyme generation. 109 proteins are enriched in the Alzheimer's disease pathway, accounting for approximately 6.51% of the total protein, and 90 proteins are enriched in Huntington's disease, accounting for approximately 5.37% of the total protein (Figure 3).

Protein is an important raw material for constructing and repairing the body, and an important component of human tissue. It plays an important role in the repair of damaged cells and the formation of antibodies during the growth and development stage of the human body. In addition, when the human body lacks carbohydrates and fats, protein can participate in heat production to meet the body's heat needs [14-17]. The protein content of *S. constricta* is rich and easy for the human body to absorb and digest. It exists as a nutrient before it is ingested into the human body. After ingested into the human body, it is digested, absorbed, decomposed into amino acids, and then synthesized into new proteins, which exist as living substances. Proteins related to exercise, such as myoglobin and its derivatives, can participate in transporting oxygen, generating energy for the body's tissues and organs, and providing energy for the body's tissues and organs. In this study, the mutual transformation between myoglobin and its derivatives is also closely related to the exercise ability of the solen's foot muscles.

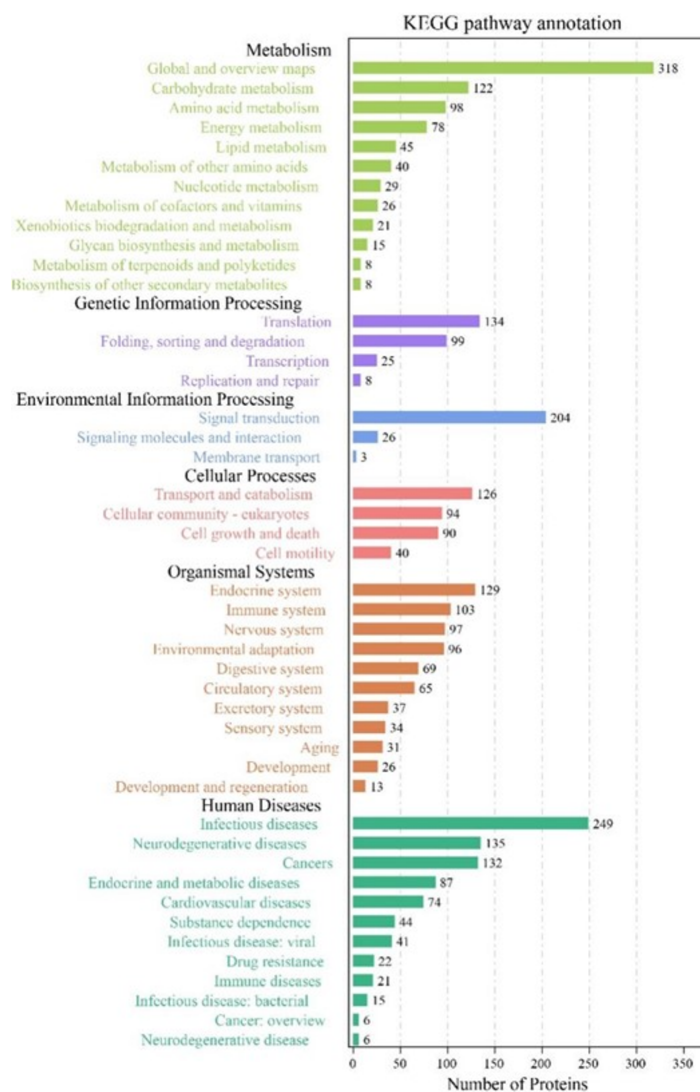

Figure 3: KEGG enrichment analysis

Table 3: Protein composition, subcellular localization and annotation

Protein name	Number of amino acids	Molecular weight	Theoretical PI	Instability index	Aliphatic index	GRAVY
ScACTB	376	41,921.94	5.30	35.09	80.93	-0.23
ScACT01	277	31,025.59	5.25	33.74	83.14	-0.20
ScACT02	376	41,791.75	5.30	36.01	82.50	-0.20
ScACT03	380	42,446.44	5.52	42.21	79.26	-0.21

Actin

The data analysis in this paper shows that there are four kinds of actin in *S. constricta*, namely ScACTB, ScACT1 ScACT2 and ScACT3. By searching, actin sequences of the bivalves, *Mytilus coruscus, M. galloprovincialis*, *Pecten maximus,* and *Mizuhopecten yessoensis*, are currently published on NCBI. The species include. As shown in Figure 4, the actin ScACT03 of *S.constricta* converges with the actin of shrimp scallops and oysters (*Ostrea gigas*), indicating a close genetic relationship. ScACT02 is closely related to its related species Cg in the ACT family. ScACTB is also clustered with the ACTB of its closely related species, the Mediterranean mussel, and is adjacent to ScACT01. ScACT01 and ScACTB belong to the same large branch system.

The results are shown in Table 3. Monomer actin is a globulin that can self polymerizeinto active double helix actin filaments in solution. It is generally composed of 375-377 amino acid residues, while ScACT01 is relatively unique with 277 amino acid numbers.

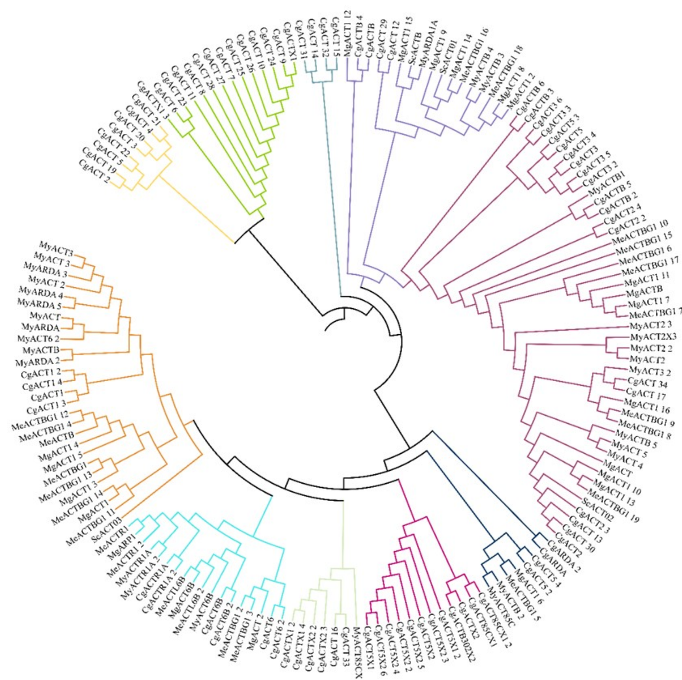

Figure 4: Phylogenetic tree of Actin in Sinonovacula constricta

Myosin

The amino acid sequences of myosin domains in *S. constricta* were compared, and the software PoyloSuite was used to construct an evolutionary tree by combining the amino acid sequences obtained from NCBI and 14 amino acid sequences from *S. constricta*. As shown in Figure 5, the corresponding phylogenetic relationships between different species or within the same species were clearly classified. From the evolutionary tree, it can be seen that the myosin of *S. constricta* is mostly divided into the same branch or in adjacent lineages as the long oyster and the shrimp scallop, and the genetic relationship is relatively close. The physicochemical properties of the selected proteins were analyzed using bioinformatics analysis software, and the results are shown in Table 4.

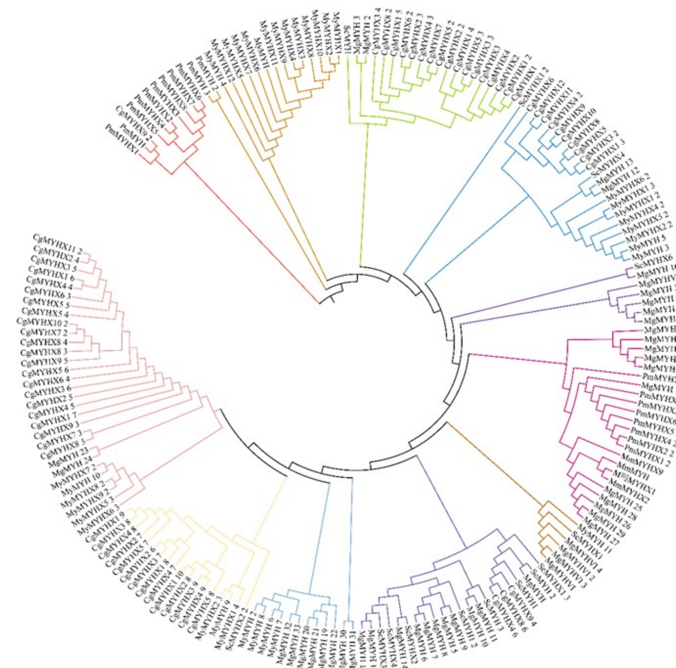

Figure 5: Phylogenetic tree of Myosin in *Sinonovacula constricta*

Table 4: Protein composition, subcellular localization and annotation

Protein name	Number of amino acids	Molecular weight	Theoretical PI	Instability index	Aliphatic index	GRAVY
ScMYHX1	1270	145979.94	9.12	40.61	76.62	-0.652
ScMYHX2	1024	118051.85	8.97	36.76	88.04	-0.408
ScMYHX6	1542	177222.49	5.82	48.71	80.73	-0.738
ScMYHX4	1955	226676.28	5.39	48.88	80.00	-0.843
ScMYH1	1776	202688.16	6.55	53.82	83.01	-0.494
ScMYH1	1009	116191.98	9.40	32.22	83.25	-0.424
ScMYHX1	900	103468.56	5.44	48.39	85.43	-0.927
ScMYHX1	577	66527.44	5.66	35.27	83.31	-0.428
ScMYH	1946	222840.76	5.59	48.06	80.39	-0.820
ScMYHX2	1081	122828.53	8.60	44.59	78.59	-0.473
ScMYH	2182	251623.95	8.79	44.88	82.62	-0.457
ScMYH	1094	124463.23	9.14	47.59	78.23	-0.527
ScMYHX4	561	65657.44	10.33	42.27	88.77	-0.454

ScMYHX3	380	42446.44	5.52	42.21	79.26	-0.213

Identification of Amino Acids and their Derivatives

The amino acids and their derivatives of the muscle tissue of the foot of *S. constricta* were detected, and 36 amino acids involved in amino acid metabolism were screened out (Table S1). According to their peak area values, they are sorted from high to low: Tryptophan, Lysine (148,404,388), Threonine, Valine (80,850,514), Leucine (27,070,522), Isoleucine (23,648,264), Methionine (20,723,750), and Phenylalanine (10,648,844), which are eight essential amino acids for the human body. The top six of the 12 non-essential amino acids for the human body are Tlutamic acid (3,425,265,032), Arginine (259,759,426), Histidine (402,582,366), Tyrosine (149,712,954), Serine (67,236,734), and Aspartic acid (681,262).

Discussion

To compare with the amino acid content of marine fish [20-23], the content of Arginine, Lysine, Leucine, and Aspartic acid in S. constricta are higher (Table 5). It can be seen from the data in the table that the contents of lysine and leucine are higher in other 7 species of Marine fish, and the top 2 amino acids with higher contents in S. constricta are Methionine and Phenylalanine, respectively. From the perspective of human essential amino acid content, the content of human essential amino acids in 7 Marine fish species is similar, but they all lack Tryptophan, while S. constricta contains 8 kinds of human essential amino acids, and its protein composition is more rich and comprehensive, which is a kind of complete protein.

Tryptophan has antioxidant effects and can alleviate stress reactions in animals. Tryptophan and its metabolites can enhance the body's immune system by regulating the production of immunoglobulins and lymphocytes, can regulate growth performance by regulating animal feed intake and intestinal peristalsis function, and can also regulate protein synthesis. When it binds to nucleic acid Tryptophan receptors, it has a certain promoting effect on the synthesis of liver nuclear proteins, promoting protein synthesis [24-27]. Based on this, it is preliminarily speculated that the antioxidant effect of S. constricta may be stronger than that of fish.

Table 5: Comparison of amino acid content between *Sinonovacula constricta* and 7 species of marine fishes

Note: * indicates essential amino acids. The amino acid content of 7 marine fish species is ranked according to the mg of amino acid contained in each gram of sample, and the amino acid content of S. constricta is ranked according to the area of amino acid peak map measured based on proteomic data.

Amino acid type	*Larimichthys crocea* [24]	*Collichthys lucidus* [23]	*Acanthopagrus schlegelii* [22]	*Pagrus major* [22]	*Thunnus albacares* [21]	*Thunnus obesus* [21]	*Thunnus maccoyii* [21]	*S. constricta*
essential amino acids	Lys*	Lys*	Lys*	Lys*	Lys*	Lys*	Lys*	Trp*
	Leu*	Leu*	Leu*	Leu*	Leu*	Leu*	Leu*	Met*
	Val*	Val*	Thr*	Val*	Val*	Val*	Thr*	Thr*
	Ile*	Ile*	Val*	Thr*	Thr*	Thr*	Val*	Lys*
	Thr*	Thr*	Phe*	Ile*	Ile*	Ile*	Ile*	Val*
	Phe*	Phe*	Ile*	Phe*	Phe*	Phe*	Phe*	Ile*
	Met*	Met*	Met*	Met*	Met*	Met*	Met*	Phe*
	-	-	-	-	-	-	-	Leu*
non-essential amino acid	Glu	Glu	Arg	Arg	Glu	Glu	Glu	Ser
	Ala	Asp	His	His	Asp	Asp	Asp	Glu
	Gly	Pro	Glu	Glu	Arg	Arg	Arg	Tyr

	Ser	Arg	Asp	Ala	Ala	Ala	Ala	Arg
	Tyr	Gly	Ala	Ser	Gly	Gly	Gly	Asp

Clam meat can present a unique fresh taste, and among the identified amino acids, Aspartic acid, Arginine, Phenylalanine, and Tyrosine are high in content, making them special flavoring substances [28-30]. The essential amino acids in the human body contain high levels of Lysine and Threonine, which can regulate metabolism. Adding an appropriate amount of Lysine to food stimulates the secretion of pepsin and enhances appetite [31-33]. Threonine has a water holding effect on human skin [34]. Among non-essential amino acids, Citrulline can soften blood vessels and is important for maintaining muscle and liver health. Among the identified amino acid derivatives, ketoglutaric acid can participate in the tricarboxylic acid cycle in the body, convert into amino acids, and serve as a dietary supplement. Taurocholic acid can reduce the permeability of inflammatory tissue capillaries, inhibit inflammatory swelling, and inhibit the production of inflammatory mediators. Taurine is an antioxidant and singlet oxygen scavenger that protects cells from oxidative damage [35]. Pantothenic acid can participate in fatty acid metabolism reactions and has a good protective effect on cells suffering from lipid peroxidation damage. Glutathione can maintain normal immune system function and has antioxidant, integrated detoxification effects. Dehydroascorbic acid plays an important role in the recycling of ascorbic acid and the protection of cellular components against oxidative damage. This study identified a high content of Taurine. Taurine is a structurally simple sulfur-containing amino acid in animals that plays an important role in the development of the infant nervous system, improving nerve conduction and visual function [36-38]. In the circulatory system, Taurine can inhibit platelet aggregation, reduce blood lipids, maintain normal blood pressure, protect myocardial cells, affect lipid absorption, absorb lipids in the digestive tract, and inhibit the formation of cholesterol stones; improving endocrine status and enhance human immunity. Taurine can affect glucose metabolism, bind to insulin receptors, accelerate cell utilization of glucose, and accelerate glycolysis [39-40]. Taurine also has a certain preventive and therapeutic effect on iron deficiency anemia, promoting the absorption of iron in the intestine, and optimizing the structure of gut bacterial flora [41]. Therefore, Taurine is widely used in food additives.

S. constricta uses its feet to dig tubular holes on the mudflat and drill into them. The depth of its hiding is affected by factors such as the size of the razor clam. The razor clam body moves up and down in the hole with the rise and fall of the tide. When *S. constricta* is submerged by the sea water, it stretches out the inlet and outlet pipes to breathe, feed, excrete and other life activities. But when the clam is exposed on the beach, it descends to the middle or bottom of the cave. The larger and more athletic clam often has soil uplift around the outlet where it resides. From the above proteomic data, it can be seen that the content of actin and myosin in the razor clam is the highest. We speculates that it is closely related to the movement of the razor clam and the contraction of the foot muscles. When *S. constricta* uses its feet to contact seawater and mudflat, it is inevitably affected by microorganisms. In this study, in addition to proteins closely related to exercise, there are also proteins involved in regulation under short--term or long-term environmental pressure, such as GAPDH. The difference in protein composition is also the reason why the stress resistance strategies of different tissues in the constricted clam have different focuses. The synthesis and decomposition of proteins are also the body's response to environmental stress, which is consistent with the speculation of Chen's team [7].

Conclusion

The protein composition and content of *S. constricta* were detected using protein non labeled quantitative technology. The research results showed that the expression levels of actin and myosin were the highest in the muscle tissue of *S. constricta* feet, making it a high protein, low fat economic shellfish. The foot muscles of *S. constricta* contain 8 essential amino acids and 12 non-essential amino acids, as well as 2 nonprotein amino acids, indicating that the protein of *S. constricta* is a complete protein with high nutritional value. This article further confirms that proteomics technology can conduct physiological and ecological research on trace samples of *S. constricta*, and efficiently and sensitively detect the proteins contained in the samples in a short period of time. Meanwhile, this article found that the actin of *S. constricta* is closely related to the actin of scallops and oysters, suggesting that they may have similar structures and functions.

References

1. Liao Z, Li N, Wang X, et al. (2012) Identification of novel foot proteins from Mytilus coruscus Byssus through proteomic analysis. Chinese Journal of Biochemistry and Molecular Biology 28: 870-878.

2. Wang X, Sun J, Fan M, et al. (2012) Proteomic analysis of granule of hemocytes from Mytilus coruscus. Southern Fisheries Science 8: 8.

3. Ortea I, Cañas B, Gallardo JM (2011) Selected tandem mass spectrometry ion monitoring for the fast identification of seafood species. JOURNAL OF CHROMATOGRAPHY A, 1218: 4445-4451.

4. Campos A, Puerto M, Prieto A, Cameán A, Almeida AM, et al. (2013) Protein extraction and two-dimensional gel electrophoresis of proteins in the marine mussel Mytilus galloprovincialis: an important tool for protein expression studies, food quality and safety assessment. Journal of the Science Of Food And Agriculture, 93: 1779-87.

5. Vincenzetti S, Felici A, Ciarrocchi G, Pucciarelli S, Ricciutelli M, et al. (2017) Comparative proteomic analysis of two clam species: Chamelea gallina and Tapes philippinarum. Food Chemistry 219: 223-9.

6. GBIF Secretariat: GBIF Backbone Taxonomy.
https://doi.org/10.15468/39omei Accessed via https://www.gbif.org/species/5284517 [07 October 2023].

7. Zhang L, Chen Y, Peng X (2004) The comparative proteomics in hemolymph of Sinonovacula constricta（Lamarck）under temperature stress. Animal Husbandry and Veterinary Medicine 36: 1-3.

8. Zhang L, Chen Y, Peng X (2004) Study of resistance in hemolymph of Sinonovacula constricta under parasitizing with proteomic approaches. Transactions of Oceanology and Limnology 2: 80-6.

9. Zhang L, Chen Y, Peng X (2004) Influence mechanisms of high temperature and low salinity on hemolymph basic proteins of Sinonovacula constricta. Chinese Journal of Marine Drugs 23: 23-7.

10. Tamura K, Stecher G, Kumar S (2021) MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Mol Biol Evol 38: 3022–7.

11. Meltretter J, Wust J, Pischetsrieder M (2013) Comprehensive analysis of nonenzymatic post-translational β-lactoglobulin modifications in processed milk by ultrahigh-performance liquid chromatography-tandem mass spectrometry. Journal of Agricultural and Food Chemistry 61: 6971 - 81.

12. Divilov K, Schoolfield B, Morga B, et al. (2019) First evaluation of resistance to both a California OsHV-1 variant and a French OsHV-1 microvariant in Pacific oysters. BMC Genetics 20: 96.

13. Martenot C, Segarra A, Baillon L, et al. (2016) In situ localization and tissue distribution of ostreid herpesvirus 1 proteins in infected Pacific oyster, Crassostrea gigas. Journal of Invertebrate Pathology 136: 124–35.

14. Arce C, Lucena C, Moreno A, et al. (2014) Proteomic analysis of intestinal mucosa responses to Salmonella enterica serovar typhimurium in naturally infected pig. Comparative Immunology Microbiology and Infectious Diseases 37: 59-67.

15. Bondzio A, Pieper R, Gabler C, et al. (2013) Feeding low or pharmacological concentrations of zinc oxide changes the hepatic proteome profiles in weaned piglets. PLoS One 8: e81202.

16. Collado-Romero M, Martins R P, Arce C, et al. (2012) An in vivo proteomic study of the interaction between Salmonella typhimurium and porcine ileum mucosa. Journal of Proteomics 75: 2015-26.

17. Etske B, De Vriesa R, Valenberg H, et al. (2014) Factors influencing casein micelle size in milk of individual cows: Genetic variants and glycosylation of κ-casein. International Dairy Journal 34: 135-41.

18. Landi C, Bargagli E, Bianchi L, et al. (2013) Towards a functional proteomics approach to the comprehension of idiopathic pulmonary fibrosis, sarcoidosis, systemic sclerosis and pulmonary Langerhans cell histiocytosis. Journal of Proteomics 83: 60-75.

19. Martenot C, Faury N, Morga B, et al. (2019) Exploring first interactions between ostreid herpesvirus 1 (OsHV-1) and its host, Crassostrea gigas: Effects of specific antiviral antibodies and dextran sulfate. Frontiers in Microbiology 10: 1128.

20. Matulova M, Varmuzova K, Sisak F, et al. (2013) Chicken innate immune response to oral infection with Salmonella enterica serovar Enteritidis. Veterinary Research 44: 37.

21. Yang C, Cao X, Zhang Y, et al. (2019) Comparative composition and nutritional evaluation of different head parts of three tuna species. Meat Research 33: 8-14.

22. Cao G, Sun R, Gao Bo, et al. (2020) Comparative study of the growth and nutrient composition of muscle tissue of Acanthopagrus schlegelii, Pagrus major and the reciprocal F1 hybrids. Fisheries Science and Technology Information 47: 61-7.

23. Cao P (2020) Analysis of the nutritional components of Collichthys lucidus and nutritional status of its broodstocks. Shanghai Ocean University.

24. Gao S (2020) Analysis of nutrient composition and gene expression related to growth of hybridization between Larimichthys polyactis ♀ and L.corcea ♂. Zhejiang Ocean University.

25. Wang X, Tan B, Ding H, et al. (2020) Research progresses on physiological function of tryptophan and its metabolites. Feed Research 43: 130-3.

26. Walter K, Grosskopf H, Karkossa I, et al. (2021) Proteomic characterization of the cellular effects of AhR activation by microbial tryptophan catabolites in endotoxin-activated human macrophages. International Journal of Environmental Research and Public Health 18: 10336.

27. Ramprasath T, Han YM, Zhang D, et al. (2021) Tryptophan catabolism and inflammation: A novel therapeutic target for Aortic Diseases. Front Immunology 12: 731701.

28. Zhang W, Ma J, Wang Z, et al. (2020) Metabolism and function of tryptophan and its application in animal production. Journal of Domestic Animal Ecology 41: 73-8.

29. Huang R, Ai C, Lin X, et al. (2014) Nutritional component analysis and quality evaluation of Sinonovacula rivularis from Changle sea area of Fujian. Journal of Applied Oceanography 33: 96-104.

30. Qin N, Zhang Y, Luo Y (2015) Comparison and analysis of major nutritional components in different body parts of Changfeng silver carp. Journal of China Agricultural University 20: 129-34.

31. Xiong M, Wu Z, Lin X (2016) Meat quality charateristics of spotted knifejaw (Oplegnathus punctatus) cultured under different aquaculture modes. Food Science 37: 17-21.

32. Biswas S, Bellare J (2021) Adaptive mechanisms induced by sparingly soluble mercury sulfide (HgS) in zebrafish: Behavioural and proteomics analysis. Chemosphere 270: 129438.

33. Tian M, Shen X, Meng X, et al. (2010) Study on the proteomics of maricultured Litopenaeus vannamei muscle tissue. Journal of Anhui Agricultural Sciences 38: 760-2.

34. Lepretre M, Almunia C, Armengaud J, et al. (2020) Identification of immune-related proteins of Dreissena polymorpha hemocytes and plasma involved in host-microbe interactions by differential proteomics. Scientific Reports 10: 6226.

35. Ren W, Zhang D, Wang G, et al. (2014) Mechanical and thermal properties of bamboo pulp fiber reinforced polyethylene composites. Biology Resources 9: 4117-27.

36. Rose-Martel M, Du J, Hincke MT (2012) Proteomic analysis provides new insight into the chicken eggshell cuticle. Journal of Proteomics 75: 2697-706.

37. Borchers C H, Christoph H, Petrotchenko E V, et al. (2014) Modern mass spectrometry-based structural proteomics. Advances in Protein Chemistry and Structural Biology 95: 193-213.

38. Luo W, Qian X, Yang S, et al. (2006) Study on the nutritive content of Sinonovacula Constricta. Jiangxi Science 24: 360-2.

39. Yang Q, Jin Q, Li W, et al. (2021) Inhibitory mechanism of lectin (MCL-T) against Shewanella putrefaciens in Manila clam Ruditapes philippinarum by proteomics approach. Journal of Dalian Fisheries University 36: 653-60.

40. Zhao Y, Li N, Li L, et al. (2016) Application of proteomics in regulation of aquatic products quality and safety: a review. Journal of Dalian Fisheries University 31: 344-50.

41. Luo M, Guan R, Jin H (2015) Analysis on the ratio of flesh content and the nutritional composition in the muscle of five species of eel. Acta Hydrobiologica Sinica 39: 714-22.