



OPEN ACCESS

Edited by:

Yuehuan Zhang,
South China Sea Institute
of Oceanology, Chinese Academy
of Sciences (CAS), China

Reviewed by:

Ryan Spengler,
University of Wisconsin–Madison,
United States
Xiaoting Huang,
Ocean University of China, China
Yinghui Dong,
Zhejiang Wanli University, China

***Correspondence:**

Weijun Wang
wwj2530616@163.com
Jianmin Yang
ladderup@126.com

Specialty section:

This article was submitted to
Marine Fisheries, Aquaculture
and Living Resources,
a section of the journal
Frontiers in Marine Science

Received: 15 June 2021

Accepted: 06 October 2021

Published: 28 October 2021

Citation:

Wang X, Wang W, Li Z, Sun G,
Xu T, Xu X, Feng Y, Luo Q, Li B and
Yang J (2021) Comprehensive
Analysis of Differentially Expressed
mRNA, Non-coding RNA, and Their
Competitive Endogenous RNA
Network of Pacific Oyster *Crassostrea*
gigas With Different Glycogen Content
Between Different Environments.
Front. Mar. Sci. 8:725628.
doi: 10.3389/fmars.2021.725628

Comprehensive Analysis of Differentially Expressed mRNA, Non-coding RNA, and Their Competitive Endogenous RNA Network of Pacific Oyster *Crassostrea gigas* With Different Glycogen Content Between Different Environments

Xue Wang¹, Weijun Wang^{1*}, Zan Li¹, Guohua Sun¹, Tao Xu², Xiaohui Xu¹, Yanwei Feng¹, Qihao Luo¹, Bin Li^{1,3} and Jianmin Yang^{1,3*}

¹ School of Agriculture, Ludong University, Yantai, China, ² General Station of Fishery Development and Resource Conservation in Shandong Province, Jinan, China, ³ Yantai Haiyu Marine Science and Technology Co., Ltd., Yantai, China

Glycogen content is a quantitative trait, its phenotype differences are found between individual oysters due to genetic effects and environmental factors which were including food, water temperature, salinity, and so on. In this study, a full sibling family of Pacific oyster *Crassostrea gigas* showed different phenotypes with high and low glycogen content between South Huanghai Sea (Rizhao offshore area, RZ) and North Huanghai Sea (Kongtong Dao area, KTD), respectively. At the same time, the content of 11 glucogenic amino acids and 13 fatty acids were also significant differences between RZ and KTD. RNA-seq and small RNA-seq technologies were used for transcriptome sequencing and functional enrichment analysis of differentially expressed RNA were used by Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway. A total of 2,084 mRNAs, 1,080 long non-coding RNAs (lncRNAs), 34 circular RNAs (circRNAs), and 7 microRNAs (miRNAs) were differentially expressed. Based on these differentially expressed genes (DEGs), miRNA target interactions (lncRNA/circRNA–miRNA pairs and miRNA–mRNA pairs) were predicted using the miRanda software. The differentially expressed mRNAs in this network were mainly shown to be involved in calcium signaling pathway and insulin signaling pathway. These findings could help to speculate that environmental factors may be epigenetically regulated by non-coding RNA in *C. gigas*, thereby further affecting glycogen content.

Keywords: *Crassostrea gigas*, glycogen, ncRNA, ceRNA, epigenetic regulation

INTRODUCTION

Pacific oyster (*Crassostrea gigas*) has the characteristics of strong adaptability to the external environment, fast growth, high yield, rich nutrition, and delicious flavor. Therefore, it is widely cultivated worldwide as a high-economic value aquatic product. At present, the improvement of oyster meat quality is one of the important research topics for its industrial development, because the quality of meat is often measured by people's basic sensory evaluation of taste. Glycogen content accounts for 20–40% of the dry weight of oysters, which directly affects the taste of oysters and is an important quality trait. Previous studies have found that on the one hand, the content of glycogen metabolism-related enzymes directly affects the changes in glycogen content (Hata et al., 1993), on the other hand, the metabolic level of amino acids and fatty acids also affects glycogen metabolism (Randle et al., 1963; Fromentin et al., 2011). Therefore, genetic and environmental factors (temperature, nutrition, salinity, etc.) may affect the glycogen content (Li et al., 2017b). At present, several key genes that have been cloned in the oyster glycogen metabolism pathway, such as: glycogen phosphorylase (Hata et al., 1993), phosphoglucomutase (Tanguy et al., 2006), glycogen synthase kinase (Zeng et al., 2013), glycogenin (Li et al., 2017a), protein phosphatase 1 regulatory subunit 3B (Liu et al., 2019), etc. Glycogen content fluctuates seasonally with gonad development, and glycogen-related genes express differences in different seasons and tissues (Bacca et al., 2005). And the increase of free fatty acid content could lead to reduce the consumption of glycogen. Similarly, high amino acid content also might increase glycogen content, because some amino acids could guide glycogen synthesis (Li et al., 2017b).

With the development of sequencing technology and the continuous improvement of the genome project, the human genome sequencing revealed that there are only about 20,000 protein-coding genes, and the proportion of the total genome sequence is <2%, and the remaining about 98% belong to non-coding RNA (ncRNA) (Chan and Tay, 2018). At present, ncRNA is divided into two types: housekeeping ncRNA and regulatory ncRNA. The common regulatory ncRNAs include microRNA (miRNA), long non-coding RNA (lncRNA), and circular RNA (circRNA). Studies have found that regulatory ncRNA is involved in the larval metamorphosis, shell pigmentation, and immune-related mechanisms (Yu et al., 2016; Feng et al., 2018). However, whether ncRNA mediates the genetic mechanism that causes changes in glycogen content due to environmental factors is still unknown.

Therefore, in order to increase our understanding of the regulation of glycogen content, we performed ncRNA transcriptome analysis. During the research process, we placed a full sibling family of *C. gigas* in two offshore areas and produced two different phenotypes: high glycogen content and low glycogen content. The gonads of individuals with high and low glycogen content were selected for analysis. These findings may help us to further understand the regulatory effects of ncRNA on glycogen content in different environments, and provide new information for future genetic breeding of *C. gigas* in aquaculture.

MATERIALS AND METHODS

Ethics Statement

In this study, all experiments were performed according to local and central government regulations. No endangered species were used in the study. No special permission was required to collect oysters or to perform the described experiments.

Experimental Animals and Sample Collection

In this study, we used a near-infrared (NIR) model (Wang et al., 2015) which constructed in our laboratory to determine the phenotypic value of the glycogen content of *C. gigas*. The families with high breeding value were selected through EBV calculation, and comprehensively consider the genetic relationship between families (the genetic relationship is less than 0.2) which could prevent the occurrence of inbreeding decline. During the oyster breeding season, the phenotypic value of glycogen content of the parents of each candidate family was determined by a NIR spectrometer (MicroNIR 1700 JDSU, United States) on site, and the male and female individuals with high glycogen content between each family were selected. Through mating between families, male and female parents from different families were fertilized to construct a full-sib family. Later, a full sibling family of *C. gigas* was farmed in the South Huanghai Sea (Rizhao offshore area, RZ) and North Huanghai Sea (Kongtong Dao area, KTD). In March 2018, we collected nine two-year-old oysters from RZ and KTD areas, and measured the glycogen content. In the RZ group and KTD group, three samples with similar growth characteristics were selected for follow-up experiments. We calculated the average glycogen content of RZ and KTD oysters, and each group selected three oyster individuals whose glycogen content was close to the average glycogen content of each group, and collected the gonadal tissues of these samples for RNA-seq and small RNA-seq, and determine its fatty acid content and amino acid composition. All samples were quick-frozen in liquid nitrogen and stored at -80°C .

Determination of Amino Acid Concentration

According to the Chinese National Standard GB 5009.124-2016, the amino acid analysis of the gonadal tissue was performed at the SGS-CSTC Standards Technical Services Co., Ltd. (Qingdao, China). The main steps were as follows. First, 15 mL of 6 mol/L HCl solution was hydrolyzed for 100 mg of sample at $110 \pm 1^{\circ}\text{C}$ for 22 h. Second, filter the acid hydrolysate with filter paper, and then used a tubular concentrator to evaporate under reduced pressure at 50°C . Third, 1.0 mL of sodium citrate buffer (pH = 2.2) was added to the dried hydrolysate, and the resulting solution was passed through a 0.22 μm filter. The external standard method is used in the amino acid analyzer to calculate the amino acid concentration in the sample solution from the peak area.

Determination of Fatty Acid Composition

According to the Chinese National Standard GB 5009.168-2016, the fatty acid composition of the gonadal tissue was measured

by gas chromatography (GC). Simply, 2 ml methanol solution of hydrochloric acid, 1 mL n-hexane were added into 100 mg powder sample and then hydrolyzed at 80°C for 2 h. After cooling down, 3 mL 6% K₂CO₃ was added to the samples. After filtered using a 0.22- μ m membrane, the samples were analyzed on a high performance gas chromatography analyzer (GC-2010, Shimadzu, Japan), each injection volume was 1 μ L. The types and relative contents of fatty acids were calculated according to the standard peak time screening and the corrected area normalization method.

Total RNA Isolation and Illumina Sequencing

RNA-seq and small RNA-seq were completed by Novogene Bioinformatics Technology Co., Ltd. (Beijing, China). The total RNA from gonadal tissue was isolated using TRIzol reagent (Invitrogen, CA, United States) according to the manufacturer's instructions. The RNA degradation and contamination was monitored on 1% agarose gels, purity was checked using the NanoPhotometer Spectrophotometer® (IMPLEN, CA, United States), concentration was measured using Qubit® RNA Assay Kit in Qubit® 2.0 Fluorometer (Life Technologies, CA, United States), integrity was assessed using the Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, United States).

According to the manufacturer's instructions, a total amount of 3 μ g RNA per sample was used as input material for the RNA sample preparations. Firstly, ribosomal RNA was removed by Epicentre Ribo-zero™ rRNA Removal Kit (Epicentre, United States), and rRNA free residue was cleaned up by ethanol precipitation. Subsequently, sequencing libraries were generated using the rRNA depleted RNA by NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina® (NEB, United States) following manufacturer's recommendations, products were purified (AMPure XP system) and library quality was assessed on the Agilent Bioanalyzer 2100 system. The clustering of the index-coded samples was performed on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumina). After cluster generation, the libraries were sequenced on an Illumina HiSeq 4000 platform and 150 bp paired-end reads were generated.

A total amount of 3 μ g total RNA per sample was used as input material for the small RNA library. Sequencing libraries were generated using NEBNext® Multiplex Small RNA Library Prep Set for Illumina® (NEB, United States), library quality was assessed on the Agilent Bioanalyzer 2100 system using DNA High Sensitivity Chips. The clustering using TruSeq SR Cluster Kit v3-cBot-HS (Illumina) After cluster generation, the library preparations were sequenced on an Illumina HiSeq 2500/2000 platform and 50 bp single-end reads were generated.

Sequencing Results

In RNA-seq, raw data were obtained by removing reads containing adapter, reads on containing ploy-N and low quality reads. At the same time, Q20, Q30, and GC contents of the clean data were calculated. Raw sequencing reads were submitted to

Sequence Read Archive in NCBI; the SRA accession numbers were SRR15293269, SRR15293270, SRR15293271, SRR15293272, SRR15293273, and SRR15293274¹.

The reference genome and gene model annotation files are downloaded from the genome website.² Index of the reference genome was built using bowtie2 v2.2.8 and paired-end clean reads were aligned to the reference genome using HISAT2 (Langmead et al., 2009) v2.0.4. HISAT2 was run with “-rna-strandness RF” other parameters were set as default. The mapped reads of each sample were assembled by StringTie (v1.3.1) (Pertea et al., 2016) in a reference-based approach.

In small RNA-seq, raw data were obtained by removing reads containing ploy-N, with 5' adapter contaminants, without 3' adapter or the insert tag, containing ploy A or T or G or C and low quality reads. At the same time, Q20, Q30, and GC contents of the clean data were calculated. Raw sequencing reads were submitted to Sequence Read Archive in NCBI; the SRA accession numbers were SRR15293267 and SRR15293268 (see text footnote 1).

The reference genome and gene model annotation files are downloaded from the genome website (see text footnote 2). The small RNA tags were mapped to reference sequence by Langmead et al. (2009) without mismatch to analyze their expression and distribution on the reference.

Identification of Long Non-coding RNA and Circular RNA

The spliced transcripts were screened for lncRNA according to the following steps: first, selected transcripts with exon number ≥ 2 and length > 200 bp; second, Cuffcompare software was used to screen annotation lncRNA (Trapnell et al., 2010); third, selected the transcripts with FPKM ≥ 0.5 ; finally, CPC (Coding Potential Calculator) and PfamScan were used to screen for coding potential (Kong et al., 2007; Mistry et al., 2007). CircRNA was screened and identified in the constructed lncRNA library. Because of the high false positives in circRNA identification, circRNA was screened using find_circ and CIRI2 (Gao et al., 2018), the results were merged and the intersection of the two was taken.

Identification of MicroRNA

The mapped small RNA tags were compared with the *Danio rerio* sequence in miRBase 20.0 to find known miRNAs. The modified software mirdeep2 (Friedlander et al., 2012) and srna-tools-cli³ were used to obtain potential miRNAs and draw secondary structures. To remove tags originating from protein-coding genes, repeat sequences, rRNA, tRNA, snRNA, and snoRNA, small RNA tags were mapped to RepeatMasker, Rfam database. The software miREvo (Wen et al., 2012) and mirdeep2 were integrated to predict novel miRNA. To make every unique small RNA mapped to only one annotation, we follow the following priority rule: known

¹https://www.ncbi.nlm.nih.gov/Traces/study/?acc=SRP330386&o=acc_s%3Aa

²https://www.ncbi.nlm.nih.gov/datasets/genomes/?txid=29159&source_db=RefSeq

³<http://srna-workbench.cmp.uea.ac.uk/>

miRNA > rRNA > tRNA > snRNA > snoRNA > repeat > gene > NAT-siRNA > gene > novel miRNA > ta-siRNA.

Target Gene Prediction

The lncRNA *cis*-acting target genes were screened for protein coding genes within 100 kb upstream and downstream; the lncRNA *trans*-acting target genes were screened for the Pearson correlation coefficient between lncRNA and mRNA than 0.95. Used miRanda (Enright et al., 2003) to predict the target gene of miRNA. The prediction was based on the miRNA binding site in the 3'UTR region of the gene. If the gene was uncertain or there was no 3'UTR region, then the miRNA binding site was determined according to the gene's stop codon within 1,000 bp downstream point.

Quantification of Gene Expression Level

Cuffdiff (v2.1.1) was used to calculate FPKMs of both lncRNAs and coding genes in each sample (Trapnell et al., 2010). Gene FPKMs were computed by summing the FPKMs of transcripts in each gene group. MiRNA expression levels were estimated by TPM (transcript per million) through the following criteria (Zhou et al., 2010).

Differentially Expressed RNA Gene Ontology and Kyoto Encyclopedia of Genes and Genomes Pathway Analysis

During the analysis, Ballgown was used for differential expression of lncRNA and mRNA analysis (Pertea et al., 2016), DESeq R package (1.10.1) was used for differential expression circRNA analysis, and DESeq R package (1.8.3) was used for differential expression miRNA analysis. The $p < 0.05$ and $|\log_2\text{FoldChange}| > 1$ as the threshold for differential screening to identify differential genes.

The GOseq R software package (Young et al., 2010) and KOBAS software (Mao et al., 2005) were used to perform Gene Ontology (GO) enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis on differentially expressed genes (DEGs) or ncRNA target genes, respectively.

Construction of the Long Non-coding RNA/Circular RNA–MicroRNA–mRNA Network

Used miRanda to predict the interaction between DEMis and DEG, DELs and DECs (Friedlander et al., 2012) and Cytoscape V3.2 software to visualize it, and choosed to express the trend as “up-down-up” or “down-up-down” network for further research (Saito et al., 2012).

Quantitative Real-Time PCR Validation

To verify the results of RNA-seq and small RNA-seq, quantitative real-time PCR (qRT-PCR) was performed. The sample used for qRT-PCR was the same as the sequencing sample. The verification primers involved were designed using Primer Premier 5 software (Table 1), and were specifically compared by NCBI Primer Blast, synthesized by Sangon Biotech Co., Ltd. (Shanghai, China). First, total RNA was used as the initial

TABLE 1 | The primers for qRT-PCR analysis.

Type	Gene name	Sequences (5'-3')
mRNA	PPP1R3B	Forward: TCCCGGAGACAGCAGCAGAAG Reverse: ACTCCCGCCTGATGAACATCC
	INSR	Forward: CCACACGACCACTGCCATTG Reverse: GCGATGAACCACTCCGATGAC
	CAM	Forward: ACTTGACAGGGAAGACCACGAG Reverse: TTCATCATAACAGGGGAGTTCC
	IP3KB	Forward: TGCTCAGACCTCACTGCTGGAC Reverse: GCAAGACGGCGGTGGAATGG
	LncRNA	XLOC-003125
XLOC-097516		Forward: CGCTCGGCTGTCTCTGTAATCTC Reverse: CCCTCGTTCAACCAGACTCTTAGC
XLOC-020486		Forward: TCCACACAGTCCACAAAGATTGCG Reverse: GTAGCGATTGCGGCAGGTCAG
MiRNA	novel-miR-154	Forward: TTGAAGTCGTTGCTGGGACATG
	novel-miR-225	Forward: CCGAGGATGTGAAGAAAGACAAAGACGT
	novel-miR-78	Forward: CGGACCTGATTTTGTACTACAGTCTGT
CircRNA	novel-miR-34	Forward: ACGTCTGTCTGAGGGTCGG
	novel-circ-000916	Forward: GCTGGTGTGGATGAACAGA Reverse: TACGGGAGTGAATGTTGGC
Reference gene	EF-1	Forward: CAAGAACGGAGATGCTGGTATGG Reverse: TTTCACTCTTTCCACCGGCTTT
	U6	Forward: GGAACGATACAGAGAAGATTAGC Reverse: TGAACGCTTACGAATTTGCG

template, using PrimeScriptTM RT reagent Kit with gDNA Eraser (TaKaRa, Beijing, China) (for mRNA and lncRNA), and Evo M-MLV RT Kit for qPCR (Accurate Biotechnology, Hunan, China) (for circRNA), Mir-X miRNA First-Strand Synthesis Kit (TaKaRa, Beijing, China) (for miRNA), convert total RNA into cDNA, with Oligo dT (for mRNA and lncRNA), random hexamer (For circRNA), poly-A RT primer (for miRNA). Second, qPCR used TB Green[®] Premix Ex TaqTM (TaKaRa, Dalian, China) for mRNAs, lncRNAs, and circRNAs, TB Green[®] Premix Ex TaqTM II (Tli RNaseH Plus) (TaKaRa, Dalian, China) for miRNAs. Extension factor 1 Genes (EF-1, used for mRNA, lncRNA, and circRNA) and U6 (used for miRNA) were used as endogenous controls (Renault et al., 2011). Finally, the Light Cycler 480 real-time PCR instrument (Roche Diagnostics, Burgess Hill, United Kingdom) was used for amplification, and $2^{-\Delta\Delta Ct}$ method was used to calculate relative gene expression levels (Ballester et al., 2013).

Statistical Analyses

The Student's *t*-test analysis in SPSS statistical software v.22 (SPSS, Inc., Chicago, IL, United States) was used to analyze the data of glycogen content, fatty acid content and amino acid content. Quantitative data were expressed as means \pm standard deviation (SD), and comparisons between different groups were completed using *t*-tests. The significance level for the statistical analyses was $p < 0.05$.

RESULTS

Phenotypic Statistics

The average glycogen content of the selected samples (fresh samples) in the RZ group and the KTD group were 7.173 ± 0.165

and 2.973 ± 0.055 mg/g, respectively. The differences in glycogen content between the two comparison groups were significantly different ($p < 0.001$). Therefore, in the follow-up analysis, we divided the two groups of oysters into high glycogen content of RZ (RZGH) group and low glycogen content of KTD (KTDGL) group.

As shown in **Table 2**, a total of 16 amino acid compositions were measured. 11 glucogenic amino acids were identified that were more abundant in KTDGL oysters ($p < 0.05$), such as aspartic acid, threonine, serine, glycine, alanine, valine, methionine, isoleucine, tyrosine, phenylalanine, and arginine. As shown in **Table 3**, there are differences in the content of 13 fatty acids in the RZGH and KTDGL groups ($p < 0.05$), and the content of 9 fatty acids in the RZGH group was significantly higher than that in the KTDGL group.

Expressional Differences in mRNAs, Long Non-coding RNAs, and MicroRNAs Between High Glycogen Content of Rizhao and Low Glycogen Content of Kongtong Dao Oysters

Compared with KTDGL oysters, there were 2,084 mRNAs (including 844 up- and 1,240 down-regulated), 1,080 lncRNAs (including 923 up- and 157 down-regulated), 34 circRNAs (including 26 up- and 8 down-regulated), and 7 miRNAs (including 4 up- and 3 down-regulated) that were differentially expressed in RZGH oysters (**Figures 1A–D**).

Functional Analysis of Differentially Expressed mRNAs, Long Non-coding RNAs, and Circular RNAs Gene Ontology Enrichment Analysis on Differentially Expressed Genes

It mainly includes three categories: biological process (BP) category, cellular component (CC) category, and molecular

TABLE 2 | Amino acid composition of RZGH and KTDGL oysters.

Amino acids	RZGH (g/100 g)	KTDGL (g/100 g)	p-value
Aspartic acid	0.583 ± 0.015	0.983 ± 0.225	0.037
Threonine	0.360 ± 0.017	0.460 ± 0.082	0.004
Serine	0.287 ± 0.021	0.433 ± 0.090	0.003
Glutamic acid	1.153 ± 0.115	1.570 ± 0.453	0.198
Glycine	0.433 ± 0.023	0.623 ± 0.097	0.005
Alanine	0.517 ± 0.012	0.613 ± 0.165	0.011
Valine	0.340 ± 0.010	0.457 ± 0.095	0.008
Methionine	0.157 ± 0.006	0.267 ± 0.055	0.001
Isoleucine	0.290 ± 0.000	0.433 ± 0.115	0.005
Leucine	0.427 ± 0.006	0.717 ± 0.216	0.003
Tyrosine	0.133 ± 0.107	0.340 ± 0.066	0.021
Phenylalanine	0.237 ± 0.006	0.377 ± 0.055	0.012
Lysine	0.520 ± 0.020	0.787 ± 0.236	0.004
Histidine	0.190 ± 0.017	0.230 ± 0.020	0.059
Arginine	0.437 ± 0.015	0.697 ± 0.208	0.001
Proline	0.593 ± 0.023	0.563 ± 0.095	0.384
Sum of 16 amino acids	6.657 ± 0.099	9.550 ± 2.184	0.005

Bold fonts were differentially expressed of glucogenic amino acids.

TABLE 3 | Fatty acid content of RZGH and KTDGL oysters.

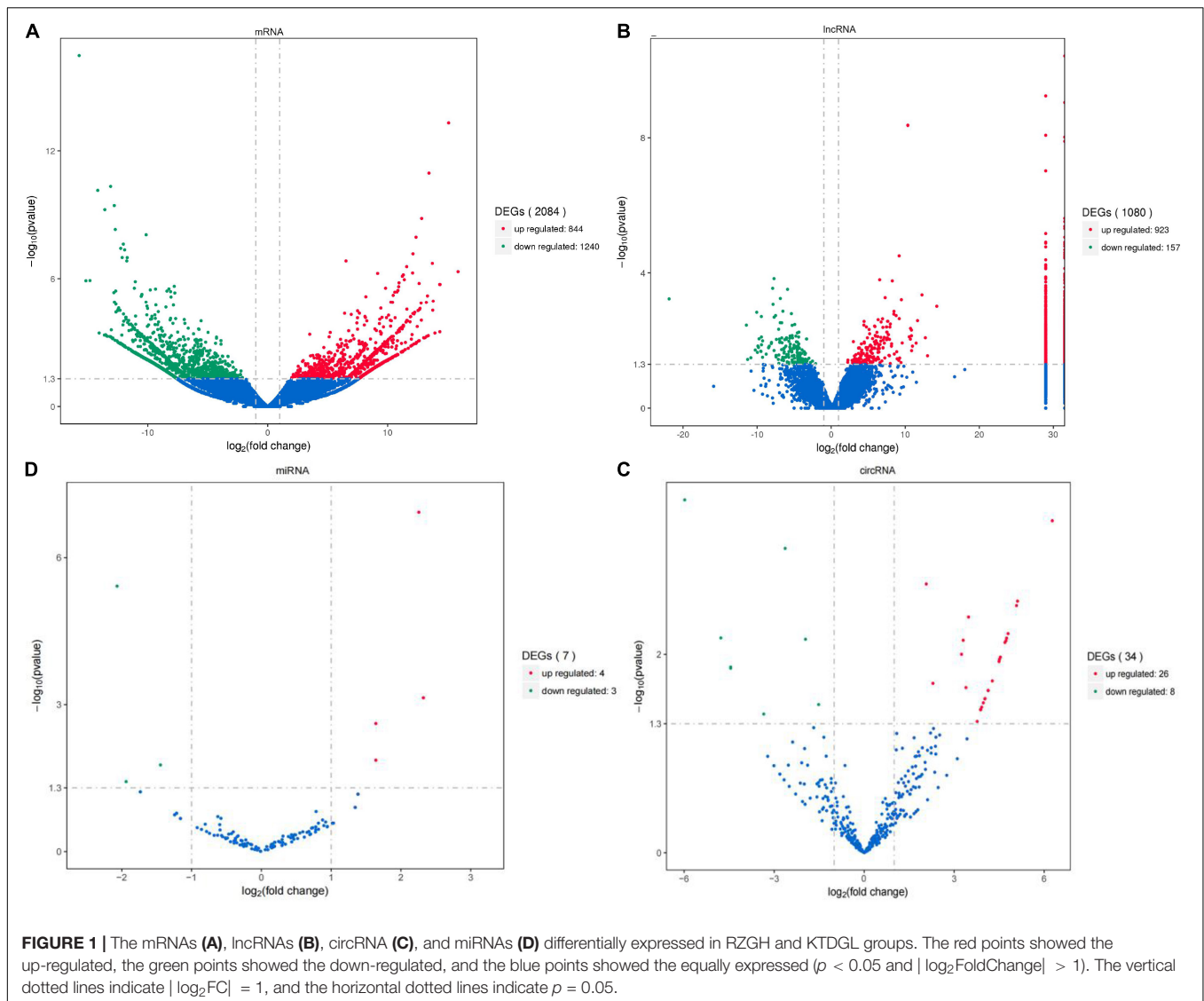
Fatty acid	RZGH (g/100 g)	KTDGL (g/100 g)	p-value
C14:0	3.302	4.206	0.042
C15:0	0.506	0.447	0.042
C16:0	0.429	0.230	0.1347
C17:0	3.409	3.003	0.617
C18:0	3.409	2.709	0.027
C20:0	0.102	0.143	0.198
C21:0	0.405	0.276	0.666
C22:0	0.045	0.061	0.424
C14:1	0.044	0.051	0.600
C15:1	18.412	18.105	0.655
C16:1	3.073	6.145	0.003
C17:1	0.434	0.149	0.187
C18:1n-9c	5.669	6.864	0.024
C18:1n-7	0.062	0.067	0.800
C20:1	4.575	3.699	0.014
C22:1n-9	0.172	0.078	0.009
C24:1	0.063	0.097	0.110
C18:2n-6c	1.258	1.758	0.060
C18:3n-6	0.127	0.100	0.031
C20:2	2.575	2.318	0.509
C20:3n-3	2.738	0.074	0.003
C20:3n-6	0.152	2.793	0.021
C22:2	0.601	0.471	0.160
C22:5n-3	0.323	0.257	0.023
ARA	2.030	1.758	0.414
EPA	20.166	23.097	0.048
DHA	10.991	8.447	0.027

Bold fonts were differentially expressed of fatty acids.

function (MF) category. GO enrichment analysis found that 120 GO terms were significantly enriched ($p < 0.05$). Among the three categories, microtubule-based movement, proteinaceous extracellular matrix and intramolecular oxidoreductase activity, and interconverting aldoses and ketoses were the most significant ones, which were selected separately the top 10 terms of significance are displayed (**Figure 2A**). KEGG enrichment analysis was performed on these DEGs to further determine the metabolic process pathway. The results showed that DEGs were significantly enriched in 13 pathways, such as calcium signaling pathway, ECM–receptor interaction, and insulin signaling pathway, among which 29 genes were enriched into the calcium signaling pathway, and 17 genes were enriched into the insulin signaling pathway (**Figure 2B**). Including protein phosphatase 1 regulatory subunit 3B (LOC105340237, PPP1R3B), calmodulin (LOC105327344, CAM), protein tyrosine phosphatase (LOC105329458, PTP), insulin-like peptide receptor (LOC105348544, INSR), acetyl coenzyme A carboxylase (LOC105343342, ACCA), inositol Triphosphate 3-kinase B (LOC105336889, IP3KB), and so on.

The *Cis* and *Trans* Role of Long Non-coding RNAs in Target Genes

To investigate the possible functions of lncRNAs, the *cis* and *trans* roles of DELs in target genes were identified, and the filtered mRNAs were analyzed in combination with DEGs. A total of 471 lncRNA–mRNA interactions were found to have *cis* roles and 1,327 lncRNA–mRNA interactions were found to



have *trans* roles. The target genes were analyzed by GO and KEGG pathway enrichment. In the results of GO enrichment analysis, it was found that fucose metabolic process, mRNA cap binding complex, and L-fucose isomerase activity were the most significant among the three categories, respectively. The top 10 terms of significance were selected for display (Figure 3A). KEGG path analysis of lncRNA target genes to explore its function. Similar to the KEGG pathway of DEGs, it was found that 22 genes were enriched in the calcium signaling pathway, and 14 genes were enriched in the insulin signaling pathway, including PPP1R3B, CAM, PTP, INSR, DDO, etc. (Figure 3B).

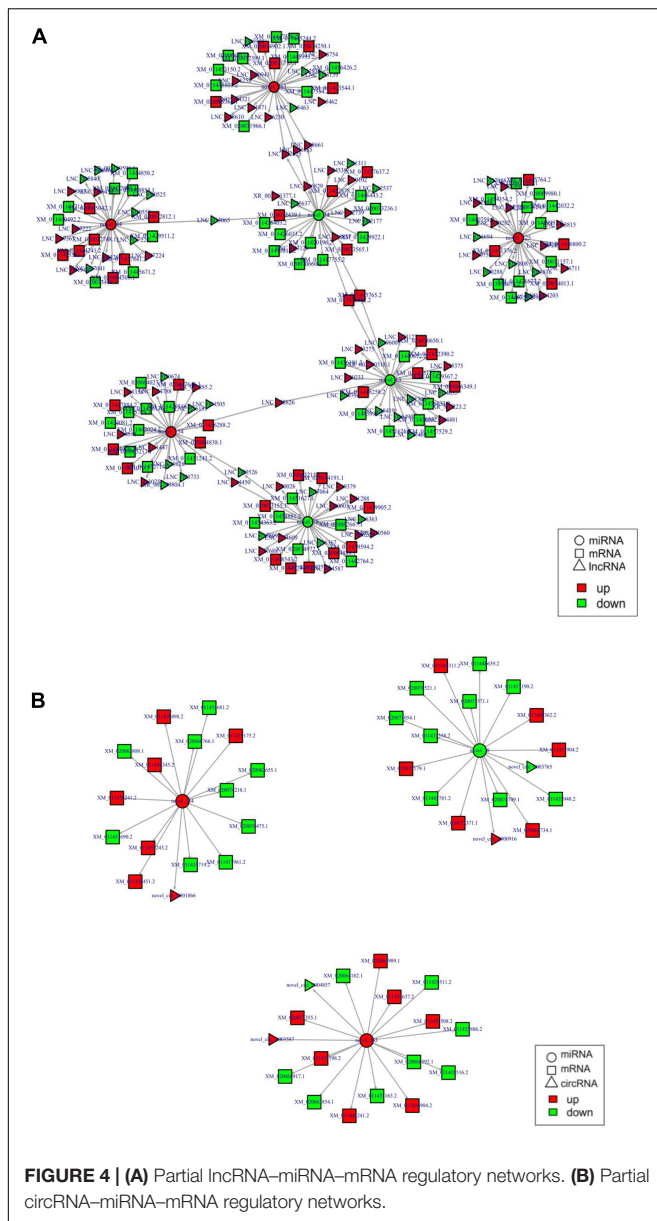
Function Analysis of Differentially Expressed Circular RNAs

The functional enrichment of circRNA host genes was explored, 33 GO terms were significantly enriched ($p < 0.05$). KEGG pathway analysis was performed on the host genes

of DEGs to explore its function. The two pathways were significantly enriched ($p < 0.05$), other glycan degradation and propanoate metabolism. And it was found that novel-circ-0004211 were transcribed from acetyl-CoA synthase 2 (LOC105319828, ACAS2).

Construction of a Potential Long Non-coding RNA/Circular RNA–MicroRNA–mRNA Regulatory Network

A total of seven DEMis (four up- and three down regulated) were screened out with small RNA-seq. By analyzing the predicted target genes and DEGs, a total of 1,077 target genes were found. Analysis found that three down-regulated DEMis, novel-miR-34, novel-miR-78, and novel-miR-13 target a total of 283 up-regulated DELs; four up-regulated DEMis, novel-miR-154, novel-miR-44, novel-miR-225, and novel-miR-263 target a total



In addition, two up-regulated DEMs were found, novel-miR-154 targeted novel-circ-0001866; novel-miR-263 targeted novel-circ-0004857 and novel-circ-0003587. One down-regulated DEMi, novel-miR-78 targets novel-circ-0003765 and novel-circ-0000916. In the interacting network, the other novel-circ-0000916 and INSR gene were all negatively regulated by novel-miR-78. Some circRNA-miRNA-mRNA network interactions were shown in **Figure 4B**.

Gene Ontology and KEGG pathway analyses of the functions of DEGs were included in the network. A total of 106 GO terms were significantly enriched ($p < 0.05$), and the top 10 terms of significance in the three categories were selected to display, as shown in **Figure 5A**. The BP mainly involves the response of microtubule-based motion, fucose metabolism, etc. The MFs mainly involved in deoxyribonuclease activity,

mercury ion transmembrane transport protein activity, and intramolecular involvement of aldose, etc. A total of 14 pathways were significantly enriched ($p < 0.05$), of which 25 genes were enriched into calcium signaling pathway, and 15 genes were enriched into insulin signaling pathway (**Figure 5B**).

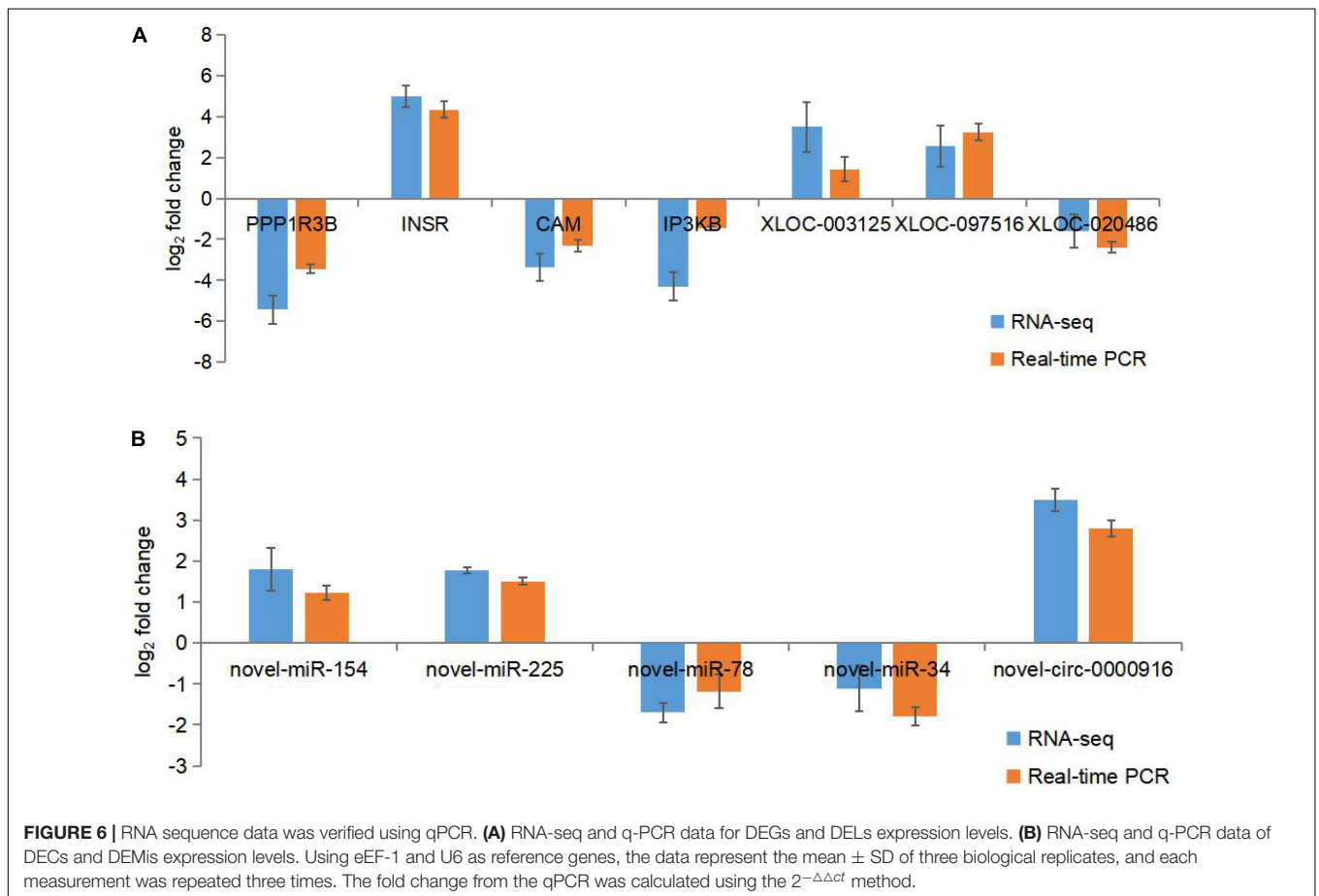
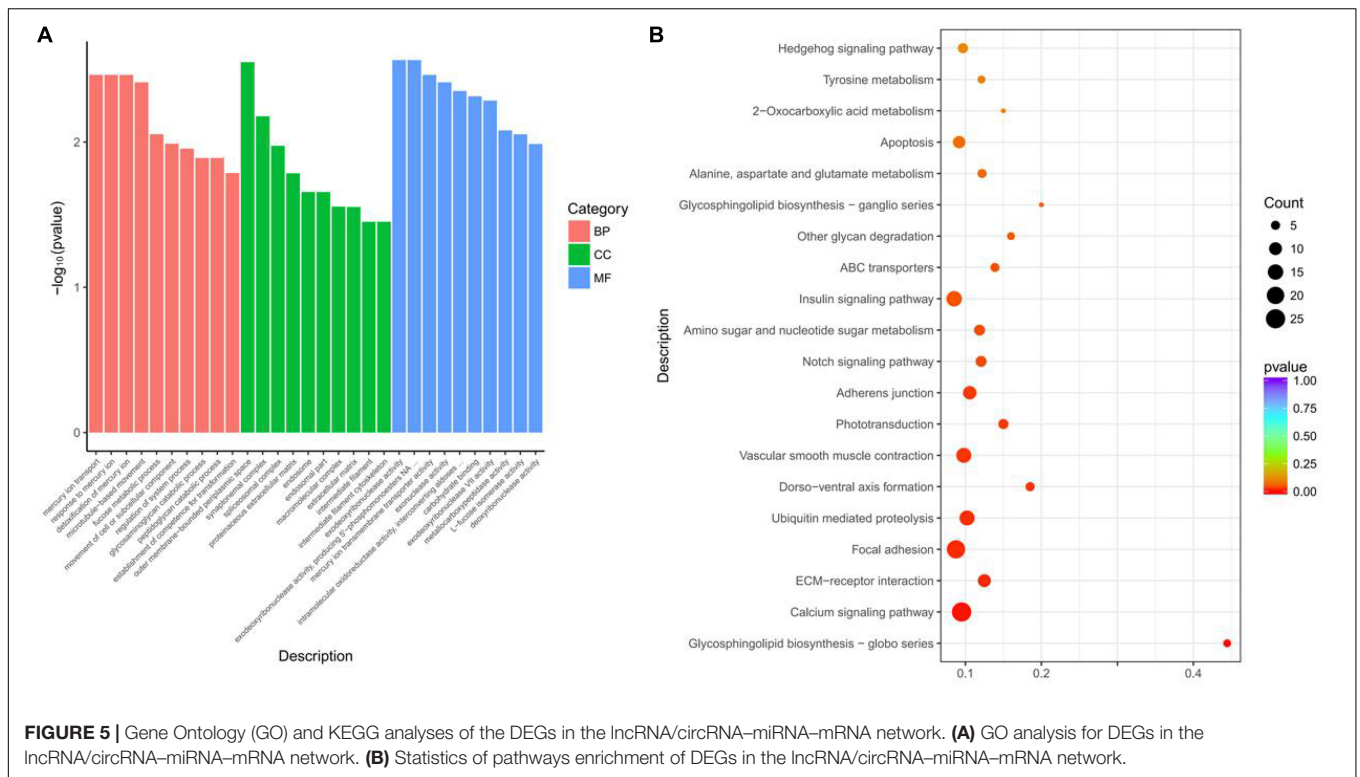
Quantitative Real-Time PCR Validation

The expression of twelve different genes was selected to confirm the utility of RNA-seq in qPCR analysis. The results showed that these transcripts were differentially expressed in oysters with different glycogen contents and were generally consistent with RNA sequencing data (**Figure 6**).

DISCUSSION

Oyster glycogen content is a quantitative trait, which varies greatly among individuals (Bacca et al., 2005; Aagesen and Haese, 2014). It is a complex and important trait that is affected by genetic and environmental factors, including seasons, nutrients, salinity, and temperature (Bacca et al., 2005; Varis et al., 2016). Liu's (2019) previous research showed that March is the proliferation period of oyster gonads, and the glycogen content of oysters is higher than the mature period of oyster gonads. Therefore, in this study, oysters of a full sibling family in two different sea areas RZ and KTD were farmed, to eliminate any potential genetic influence. In March 2018, we collected samples separately and found that the oyster glycogen in the RZ area was significantly higher than that of oysters in the KTD area. Since carbohydrates such as glycogen, amino acids, and fatty acids are the three major nutrients in organisms, their metabolic pathways are interrelated. An increase in free fatty acid levels may lead to an increase in glycogen content by reducing glycogen consumption (Randle et al., 1963). Similarly, a high-protein diet can increase glycogen content, because a large part of dietary amino acids are directed to glycogen synthesis (Fromentin et al., 2011). Therefore, we also measured the amino acid composition and fatty acid composition of the oysters in the RZGH and KTDGL groups, and the results found that there were significant differences in the amino acid composition and fatty acid composition of the oyster gonads between the two groups.

In this study, in order to get insight into the molecular regulation mechanism of glycogen content changes in oysters under different environments, we compared the differences in the expression profiles of mRNA, miRNA, lncRNA, and circRNA in the oyster gonads in the RZGH and KTDGL groups of a full sibling family. To determine the key factors involved in the regulation of glycogen content. Functional analysis of 2,084 differentially expressed genes showed that the insulin-like family genes were significantly enriched. Previous studies had shown that insulin could promote glycogen synthesis in mammals (Li et al., 2019). *Lymnaea stagnalis* was the first shellfish to discover insulin-like family genes, and for invertebrates, the insulin-like family was widespread and highly conserved (Li et al., 1992). It was mainly involved in growth, development, and metabolism (Brogiolo et al., 2001). However, whether insulin-like in *C. gigas* was related to glycogen content had not been



reported. INSR is an insulin-like receptor protein tyrosine kinase substrate (Payankulam et al., 2019), PTP dephosphorylates the tyrosine residues on the insulin receptor or its substrate, which prevents the insulin receptor from binding to insulin and negatively regulates insulin signal transduction (Haeusler et al., 2018). During the research, it was found that PTP was lowly expressed in RZGH oyster gonads, and INSR was highly expressed in RZGH oyster gonads. Therefore, it is speculated that the low expression of PTP in RZGH oyster gonads reduces the negative regulation of insulin-like signal transduction and further promotes glycogen synthesis. There are several genes related to glycogen synthesis in the insulin signaling pathway. For example, PPP1R3B was highly expressed in RZGH oyster gonad, it could cause glycogen synthase dephosphorylation, promote glycogen synthase activity, and further promote glycogen synthesis (Liu et al., 2019). lncRNA performs its function mainly by regulating mRNA, and analysis of the potential function of lncRNA is mainly predicted by analyzing their protein-coding target genes (Zhang et al., 2016). In the RZGH vs KTDGL group, it was found that lncRNA *cis* targets 471 mRNAs and *trans* targets 1,327 mRNAs. Functional enrichment analysis found that it was similar to differentially expressed mRNA and was related to insulin-like synthesis. For example, some of differentially expressed lncRNAs, XLOC-019254, XLOC-003125, XLOC-097516, and XLOC-020486 regulate CAM, PTP, INSR, IP3KB, etc., respectively. Therefore, lncRNA may play a certain regulatory role in insulin-like metabolism.

In eukaryotes, miRNAs have been reported to affect a variety of BPs, including glycogen metabolism (Dou et al., 2015), and their role in regulation is very important. With the development of high-throughput sequencing technology, studies on the miRNA transcriptome profiles of various mollusks have been reported (Xu et al., 2014; Zhou et al., 2014), and a new theory competitive endogenous RNA (ceRNA) has been proposed. However, in oysters, there is no report about miRNAs targeting glycogen metabolism in different environments. In this study, functional analysis of the seven differentially expressed miRNA target genes revealed that the differentially expressed miRNA might also be related to the insulin signaling pathway. The analysis found that in the network interaction analysis, INSR genes and lncRNAs XLOC-116536, XLOC-104062, etc., were all negatively regulated by novel-miR-13, novel-miR-34, and novel-circ-0000916 and INSR gene were all negatively regulated by novel-miR-78. These reports and our research results indicate that the ceRNA network formed by the interaction of lncRNA-miRNA-mRNA and circRNA-miRNA-mRNA related to the

insulin-like signaling pathway may play a role in the regulation of glycogen content in different environments, but the specific function needs further study.

CONCLUSION

In summary, we compared the expression characteristics of mRNA and ncRNA in the same high-glycogen content sibling family produced two different phenotypes of *C. gigas* with high and low glycogen content. The results show that there are a large number of ncRNAs in gonadal tissues, and bioinformatics analysis shows that ncRNAs are involved in insulin-like metabolism. These include novel-miR-78, novel-miR-154, novel-miR-225, and novel-circ-0000916 and XLOC-116536, XLOC-105344, and XLOC-104062, etc. These findings may help us to further understand the regulatory effects of ncRNA on glycogen content in different environments, and provide new information for selective breeding of *C. gigas* in future.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: https://www.ncbi.nlm.nih.gov/Traces/study/?acc=SRP330386&o=acc_s%3Aa.

AUTHOR CONTRIBUTIONS

WW and JY contributed to the conception and design of the study. WW, XW, ZL, QL, and BL completed the material preparation, and data collection and analysis. GS, TX, XX, and YF supervised the data collection and analysis. WW and XW wrote the first draft. All authors read and approved the final manuscript.

FUNDING

This work was supported by the National Nature Science Foundation of China, No. 31402298, the Earmarked Fund for Agriculture Seed Improvement Project of Shandong Province, China, No. 2020LZGC016, and Innovation and Entrepreneurship Program of Jiangsu Province, China.

REFERENCES

- Aagesen, A. M., and Haese, C. C. (2014). Seasonal effects of heat shock on bacterial populations, including artificial vibrio parahaemolyticus exposure, in the pacific oyster, *crassostrea gigas*. *Food Microbiol.* 38, 93–103. doi: 10.1016/j.fm.2013.08.008
- Bacca, H., Huvet, A., Fabioux, C., Daniel, J.-Y., Delaporte, M., Pouvreau, S., et al. (2005). Molecular cloning and seasonal expression of oyster glycogen phosphorylase and glycogen synthase genes. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 140, 635–646. doi: 10.1016/j.cbpc.2005.01.005
- Ballester, M., Cordon, R., and Folch, J. M. (2013). DAG expression: high-throughput gene expression analysis of real-time PCR data using standard curves for relative quantification. *PLoS One* 8:e80385. doi: 10.1371/journal.pone.0080385
- Broggiolo, W., Stocker, H., Ikeya, T., Rintelen, F., Fernandez, R., and Hafen, E. (2001). An evolutionarily conserved function of the *Drosophila* insulin receptor and insulin-like peptides in growth

- control. *Curr. Biol.* 11, 213–221. doi: 10.1016/S0960-9822(01)0068-9
- Chan, J. J., and Tay, Y. (2018). Noncoding RNA:RNA Regulatory Networks in Cancer. *Int J Mol Sci.* 19, 1310. doi: 10.3390/ijms19051310
- Dou, L., Meng, X., Sui, X., Wang, S., Shen, T., and Huang, X. (2015). MiR-19a regulates PTEN expression to mediate glycogen synthesis in hepatocytes. *Sci. Rep.* 5:11602. doi: 10.1038/srep11602
- Enright, A., John, B., Gaul, U., Tuschl, T., Sander, C., Marks, D., et al. (2003). MicroRNA targets in *Drosophila*. *Genome Biol.* 4, 1–27. doi: 10.1186/gb-2003-4-11-p8
- Feng, D., Li, Q., Yu, H., Kong, L., and Du, S. (2018). Transcriptional profiling of long non-coding RNAs in mantle of *Crassostrea gigas* and their association with shell pigmentation. *Sci. Rep.* 8:1436. doi: 10.1038/s41598-018-19950-6
- Friedlander, M. R., Mackowiak, S. D., Li, N., Chen, W., and Rajewsky, N. (2012). miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res.* 40, 37–52. doi: 10.1093/nar/gkr688
- Fromentin, C., Azzout-Marniche, D., Tomé, D., Even, P., Luengo, C., Piedcoq, J., et al. (2011). The postprandial use of dietary amino acids as an energy substrate is delayed after the deamination process in rats adapted for 2 weeks to a high protein diet. *Amino Acids* 40, 1461–1472. doi: 10.1007/s00726-010-0756-3
- Gao, Y., Zhang, J., and Zhao, F. (2018). Circular RNA identification based on multiple seed matching. *Brief Bioinform.* 19, 803–810. doi: 10.1093/bib/bbxx014
- Haeusler, R. A., McGraw, T. E., and Acili, D. (2018). Biochemical and cellular properties of insulin receptor signalling. *Nat. Rev. Mol. Cell Biol.* 19, 31–44. doi: 10.1038/nrm.2017.89
- Hata, K., Hata, M., and Matsuda, K. (1993). Purification and properties of glycogen phosphorylase from the adductor muscle of the oyster, *Crassostrea gigas*. *Comp. Biochem. Physiol. B* 105, 481–486. doi: 10.1016/0305-0491(93)90077-1
- Kong, L., Zhang, Y., Ye, Z. Q., Liu, X. Q., Zhao, S. Q., Wei, L., et al. (2007). CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res.* 35, W345–W349. doi: 10.1093/nar/gkm391
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10:R25. doi: 10.1186/gb-2009-10-3-r25
- Li, B., Meng, J., Li, L., Liu, S., Wang, T., and Zhang, G. (2017a). Identification and functional characterization of the glycogen synthesis related gene glycogenin in pacific oysters (*Crassostrea gigas*). *J. Agric. Food Chem.* 65, 7764–7773. doi: 10.1021/acs.jafc.7b02720
- Li, B., Song, K., Meng, J., Li, L., and Zhang, G. (2017b). Integrated application of transcriptomics and metabolomics provides insights into glycogen content regulation in the Pacific oyster *Crassostrea gigas*. *BMC Genomics* 18:713. doi: 10.1186/s12864-017-4069-8
- Li, K. W., Geraerts, W. P., and Joosse, J. (1992). Purification and sequencing of molluscan insulin-related peptide II from the neuroendocrine light green cells in *Lymnaea stagnalis*. *Endocrinol* 130, 3427–3432. doi: 10.1210/en.130.6.3427
- Li, Q., Zhao, Q., Zhang, J., Zhou, L., Zhang, W., Chua, B., et al. (2019). The protein phosphatase 1 complex is a direct target of AKT that links insulin signaling to hepatic glycogen deposition. *Cell Rep.* 28, 3406–3422. doi: 10.1016/j.celrep.2019.08.066
- Liu, P. (2019). *Study on Nutrition Quality, Flavor and Plankton Correlation of Crassostrea gigas*. Ph.D. thesis. Shanghai: Shanghai Ocean University.
- Liu, S., Li, L., Meng, J., Song, K., Huang, B., Wang, W., et al. (2019). Association and functional analyses revealed that PPP1R3B plays an important role in the regulation of glycogen content in the pacific oyster *Crassostrea gigas*. *Front. Genet.* 10:106. doi: 10.3389/fgene.2019.00106
- Mao, X., Cai, T., Olyarchuk, J. G., and Wei, L. (2005). Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics* 21, 3787–3793. doi: 10.1093/bioinformatics/bti430
- Mistry, J., Bateman, A., and Finn, R. D. (2007). Predicting active site residue annotations in the Pfam database. *BMC Bioinformatics* 8:298. doi: 10.1186/1471-2105-8-298
- Payankaulam, S., Raicu, A. M., and Arnosti, D. N. (2019). Transcriptional regulation of INSR, the insulin receptor gene. *Genes* 10:984. doi: 10.3390/genes10120984
- Pertea, M., Kim, D., Pertea, G. M., Leek, J. T., and Salzberg, S. L. (2016). Transcript-level expression analysis of rna-seq experiments with hisat, stringtie and ballgown. *Nat. Protoc.* 11, 1650–1667. doi: 10.1038/nprot.2016.095
- Randle, P. J., Garland, P. B., Hales, C. N., and Newsholme, E. A. (1963). The glucose fatty-acid cycle. its role in insulin sensitivity and the metabolic disturbances of diabetes mellitus. *Lancet* 281, 785–789. doi: 10.1016/S0140-6736(63)91500-9
- Renault, T., Faury, N., Barbosa-Solomieu, V., and Moreau, K. (2011). Suppression subtractive hybridisation (SSH) and real time PCR reveal differential gene expression in the Pacific cupped oyster, *Crassostrea gigas*, challenged with ostreid herpesvirus 1. *Dev. Comp. Immunol.* 35, 725–735. doi: 10.1016/j.dci.2011.02.004
- Saito, R., Smoot, M. E., Ono, K., Ruschinski, J., Wang, P. L., Lotia, S., et al. (2012). A travel guide to cytoscape plugins. *Nat. Methods* 9, 1069–1076. doi: 10.1038/nmeth.2212
- Tanguy, A., Boutet, I., Boudry, P., Degremont, L., Laroche, J., and Moraga, D. (2006). Molecular identification and expression of the phosphoglucomutase (PGM) gene from the Pacific oyster *Crassostrea gigas*. *Gene* 382, 20–27. doi: 10.1016/j.gene.2006.06.005
- Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., et al. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515. doi: 10.1038/nbt.1621
- Varis, J., Haverinen, J., and Vornanen, M. (2016). Lowering temperature is the trigger for glycogen build-up and winter fasting in crucian carp (*Carassius Carassius*). *Zool Sci* 33, 83–91. doi: 10.2108/zs150072
- Wang, W. J., Yang, J. M., Li, Q., Ji, R., Gong, X. H., and Li, L. (2015). Development of calibration models for rapid determination of chemical composition of pacific oyster (*Crassostrea gigas*) by near infrared reflectance spectroscopy. *J. Shellfish Res.* 34, 303–309. doi: 10.2983/035.034.0212
- Wen, M., Shen, Y., Shi, S., and Tang, T. (2012). miREvo: an integrative microRNA evolutionary analysis platform for next-generation sequencing experiments. *BMC Bioinformatics* 13:140. doi: 10.1186/1471-2105-13-140
- Xu, F., Wang, X., Feng, Y., Huang, W., Wang, W., Li, L., et al. (2014). Identification of conserved and novel microRNAs in the Pacific oyster *Crassostrea gigas* by deep sequencing. *PLoS One* 9:e104371. doi: 10.1371/journal.pone.0104371
- Young, M. D., Wakefield, M. J., Smyth, G. K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 11:R14. doi: 10.1186/gb-2010-11-2-r14
- Yu, H., Zhao, X., and Li, Q. (2016). Genome-wide identification and characterization of long intergenic noncoding RNAs and their potential association with larval development in the Pacific oyster. *Sci. Rep.* 6:20796. doi: 10.1038/srep20796
- Zeng, Z., Ni, J., and Ke, C. (2013). Expression of glycogen synthase (GYS) and glycogen synthase kinase 3beta (GSK3beta) of the Fujian oyster, *Crassostrea angulata*, in relation to glycogen content in gonad development. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 166, 203–214. doi: 10.1016/j.cbpb.2013.09.003
- Zhang, S., Qin, C., Cao, G., Xin, W., Feng, C., and Zhang, W. (2016). Systematic analysis of long noncoding RNAs in the senescence-accelerated mouse prone 8 brain using RNA sequencing. *Mol. Ther. Nucleic Acids* 5:e343. doi: 10.1038/mtna.2016.57
- Zhou, L., Chen, J., Li, Z., Li, X., Hu, X., Yi, H., et al. (2010). Integrated profiling of micrornas and mrnas: micrornas located on xq27.3 associate with

clear cell renal cell carcinoma. *PLoS One* 5:e15224. doi: 10.1371/journal.pone.0015224

Zhou, Z., Wang, L., Song, L., Liu, R., Zhang, H., Huang, M., et al. (2014). The identification and characteristics of immune-related microRNAs in haemocytes of oyster *Crassostrea gigas*. *PLoS One* 9:e88397. doi: 10.1371/journal.pone.0088397

Conflict of Interest: BL and JY were employed by the company Yantai Haiyu Marine Science and Technology Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Wang, Wang, Li, Sun, Xu, Xu, Feng, Luo, Li and Yang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.