The Israeli Journal of Aquaculture – Bamidgeh • ISSN 0792-156X • IJA.74.2022.1722195, 15 pages
CCBY-NC-ND-4.0 • https://doi.org/10.46989/001c.37738

The *IJA* is a peer-reviewed open-access, electronic journal, freely available without charge to users
-
Produced by the AquacultureHub non-profit Foundation
Sale of *IJA* papers is strictly forbidden

Transcriptome Analysis of *Metapenaeus affinis* Reveals Genes Involved in Gonadal Development

Yan Zhang [1,2#], Yun-Dong Li [2], Qi-Bin Yang [2], Song Jiang [2], Shi-Gui Jiang [2], Fa-Lin Zhou [2*]

[1] *College of Fisheries and Life Sciences, Shanghai Ocean University, Shanghai 201306, P.R. China*

[2] *Key Laboratory of South China Sea Fishery Resources Exploitation and Utilization, Ministry of Agriculture and Rural Affairs, South China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, 231 Xingang West Road, Guangzhou 510300, P.R. China*

Keywords: DEGs, *Metapenaeus affinis*, Molecular mechanism, Ovary, RNA-Seq

Abstract

Metapenaeus affinis is a crustacean with important commercial value in the fishery of the South China Sea. Overfishing has resulted in the decline of the wild population and germplasm degradation. However, there is little background knowledge about its gonadal development, and there is a lack of research on the development of this species. To better understand the molecular regulatory mechanisms during gonadal development, here, we performed RNA-Seq on immature and mature ovaries and compared their transcriptomic signatures. 126,930,488 and 122,677,356 clean sequencing reads were obtained from the Illumina sequencing platform, respectively. 394 differentially expressed genes (DEGs) were identified, of which 136 were up-regulated, and 258 were down-regulated. Further analysis revealed rich transcriptional sequences, which have homology with genes related to reproduction and development. Expression patterns of COX, GPX, E3s, PCNA, STPK, and other genes were changed during ovarian development. Validation by qRT-PCR demonstrated the reliability of RNA-Seq. This study has made a significant contribution to the currently available sequence data of *M. affinis* and provided reference data for the development of genetic and breeding work.

* Corresponding author. Fa-Lin Zhou. e-mail: zhoufalin0925@163.com

Introduction

As one of the important species of Penaeus, *Metapenaeus affinis* (H. Milne Edwards, 1837) (Decapoda, Penaeidae) is origins in Indo-West Pacific, distributed in Pakistan, Bangladesh, India, Malaysia, and China (Holthuis, 1980). In the South China Sea, *M. affinis* is an important Decapod resource with important commercial value. Its meat is plump and nutritious, containing more protein, minerals, and saturated fatty acids (Dancer et al., 2014). However, due to the excessive dependence of *M. affinis* on wild fisheries, the number of wild populations has decreased, which harms their population health. As a management measure, aquaculture can not only reduce the fishing pressure on wild populations, but also improve the yield, which is a method to maintain the sustainable development of resources (González-Castellano et al., 2019). The premise of developing shrimp aquaculture technology is to understand the molecular regulation mechanism that occurs during the reproduction process (Lo et al., 2007). Understanding the genetic basis for the regulation of traits such as gonad development and maturation in species can provide information for aquaculture programs and promote the development of high-quality broodstock (Santos et al., 2014).

The ovary is a multifunctional organ, which mainly plays a role in the reproduction and hormone secretion of female shrimp and regulates its growth and development (Chu et al., 2005). Some scholars have described the morphology and structure of ovary and germ cells of *M. affinis*, as well as their histological changes during gonadal development, but little is known about the molecular regulatory mechanism of ovarian maturation and the expression of related genes (Safaie, 2009). The first step to understanding the molecular mechanism of gonadal development in species is to identify and describe gonad-related genes and corresponding regulatory pathways (Peng et al., 2015). In recent years, great progress has been made in this area. Some genes have been confirmed to be key factors in the growth and development of crustaceans, including Cyclin B (Feng et al., 2020), Sox9 (Wan et al., 2021), Crustacean hyperglycemia hormone (CHH) (Guo et al., 2020), gem associated protein 2-like isoform X1 (GEM) (Jin et al., 2019), Crustacean female sex hormone (CFSH), (Zhuang et al., 2020), mitogen- activated protein kinase 1 (Mapk1) (Ponza et al., 2011), sex determinant transformer-2 (tra-2) (Wangl et al., 2019).

The development of next-generation sequencing technology has radically accelerated biological research by providing large amounts of data in a short period of time and at low cost. Transcriptome sequencing enables the generation of high-throughput double-stranded cDNA fragments that are rapidly assembled into sequences for annotation. Through transcriptome sequencing, functional gene information involved in specific biological processes can be obtained from specific tissues, which broadens the understanding of gene networks, especially in non model organisms with unknown genomes (Cahais et al., 2012; Bar, I et al., 2016; Mardis, 2008). The genes related to reproduction and development were successfully identified in some commercially important crustacean species by transcriptome sequencing, such as *Fenneropenaeus merguiensis* (Powell et al., 2015), *Macrobrachium nipponense* (Qiao et al., 2017), *Portunus trituberculatus* (Meng et al., 2015), *Eriocheir sinensis* (Li and Qian, 2017) and *Litopenaeus vannamei* (Peng et al., 2015). Up to now, the genomics of *M. affinis* is still an area of research that has not yet been explored, and there are few reports on whole-genome sequencing and next-generation sequencing, resulting in a lack of research information related to reproductive processes and gonad development. Studying the mechanism of ovarian development and transcriptome information is one of the effective methods to solve the depletion of species germplasm resources, while the transcriptome sequencing of ovarian-related tissues can reflect the dynamic level of gene transcription, thus providing a basis for molecular biology research (Bissonnette et al., 2009; Chalmel et al., 2007; Wang, 2009).

In this study, RNA-Seq technology was used to prepare a transcriptome library, and then gene expression changes in immature ovaries and mature ovary tissues were analyzed to identify gonadal maturation-related genes. The generation of gonad specific library of *M. affinis*

is helpful to understand the molecular mechanism of gonadal development. These data provide a theoretical basis for exploring the regulation mechanism of the ovary development of *M. affinis* at the molecular level, which is helpful for genetic breeding research.

Materials and methods

Sample collection

The females of *M. affinis* used in this study were purchased from the seafood wholesale market in Sanya, Hainan Province, China. Then through the observation of the size and color of the ovaries on the back (Safaie, 2009; Ayub and Ahmed, 2002), they were divided into immature ovaries (average weight 10-15g) and mature ovaries (average weight 20-30g). Immature ovaries are translucent and colorless with an opaque line visible, while mature ovaries are green and can be clearly seen through the exoskeleton. Three biological replicates in each group. Before dissection, *M. affinis* were anesthetized in an ice tray for about 1-2 minutes. Ovarian tissue were took out and immediately placed in liquid nitrogen, then stored at -80°C until RNA extraction.

RNA extraction and sequencing

Extraction of total RNA from ovarian tissue by using Trizol reagent (Sangon, China). After that, the samples of RNA were detected. Detection of RNA degradation or contamination on a 1% agarose gel. RNA purity and concentration (OD260/280: 1.8-2.2, OD260/230: 2.0-2.2) were detected by Nanodrop (Thermo Scientific NC2000, USA). RNA integrity was assessed by the RNA Nano 6000 Assay Kit from the Agilent 2100 Bioanalyzer (Agilent Technologies, CA, USA).

After the RNA samples were qualified, the mRNA was enriched using magnetic beads with Oligo(dT), and then cDNA was synthesized by reverse transcription. Double-stranded cDNA was purified using the AMPure XP system (Beckman Coulter, Beverly, USA), followed by end-repair and the addition of "A" bases. Afterward, 250-300bp cDNA fragments were selected by AMPure XP beads. Then, Agilent 2100 was used to check the quality of the cDNA library constructed by PCR amplification. After the library passed the quality inspection, Illumina's high-throughput sequencing platform (Illumina novaseq 6000) was used for sequencing.

De novo assembly and functional annotation

Clean reads were obtained by filtering out a small amount of linker contamination and low-quality reads in the raw data. Reads containing adapters were removed by using Trimmomatic software (v0.33); Filtered out reads with a ratio of more than 10% of N (N means that the base information cannot be determined); Reads with more than 50% low-quality bases (Q≤20) were deleted. The clean reads were assembled by Trinity software (Grabherr et al., 2011). The reads were first spliced into contigs, which were grouped and joined into transcripts (unigenes) whose ends could not be extended.

To further obtain more comprehensive gene function information, the reference-free genome analysis method was used to annotate the gene function of seven databases on the sequencing data. These databases include Nr (NCBI non-redundant protein sequences) , Nt (NCBI non-redundant nucleotide sequences) (http://www.ncbi.nlm.nih.gov/), Pfam (Protein family) (http://pfam.xfam.org/), KOG/COG (Clusters of Orthologous Groups of proteins) (http://www.ncbi.nlm.nih.gov/COG/KOG), Swiss-Prot (A manually annotated and reviewed protein sequence database) (http://www.ebi.ac.uk/uniprot/), KEGG (Kyoto Encyclopedia of Genes and Genome) (http://www.genome.jp/kegg/), KO (KEGG Ortholog) (http://www.genome.jp/kegg/), GO (Gene Ontology) (http://www.geneontolo-gy.org/).

Differentially expressed genes (DEGs) and enrichment analysis

Differential expression analysis of the two groups was performed using the DESeq R package (1.10.1). DESeq provides statistical routines for determining differential expression

in digital gene expression data using a model based on the negative binomial distribution (Langmead and Salzberg, 2012). The resulting P-values were adjusted to control the false discovery rate using the method of Benjamini-Hochberg (Roberts and Pachter, 2013). Genes with adjusted P-values < 0.05 found by DESeq were designated as differentially expressed. GO functional enrichment analysis and KEGG pathway enrichment analysis were performed on the differential gene sets using the software GOseq (Young et al., 2010) and KOBAS (Mao et al., 2005), and significance was set at qvalue<0.05. The enrichment analysis is based on the principle of hypergeometric distribution. All genes in the differential gene set are analyzed for significant differences and annotated into the GO and KEGG database. The enrichment analysis obtains all differential genomes, up-regulated differential genes and down-regulated genes for each differential comparison combination.

Real-time quantitative PCR (qRT-PCR) validation

Ten DEGs were randomly selected for quantitative real-time polymerase chain reaction (qRT-PCR) validation of sequencing quantitative data. Elongation factor 1 α (EF1-α) was used as an internal reference gene, three parallel samples per group, and each parallel sample was repeated three times. Gene-specific primers were designed based on the identified transcript sequences by the primer software Premier 6.0 (**Table 1**). Roche Light Cycler 96 (Roche Applied Science, Germany) was used for the assay. The reaction mixture (12.5 μL) contained 6.25 μL of 2 × SYBR Premix Ex Taq, 0.5 μL of upstream and downstream primers, 1.0 μL of cDNA, and 4.25 μL of sterile water. The qPCR cycling parameters were as follows: 300 s at 95°C; 40 cycles of 10 s at 95°C, and 30 s at 60°C; followed by the melting curve. Finally, the expression of related genes was analyzed by the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001).

Table 1 Primers used for quantitative real-time polymerase chain reaction (qRT-PCR)

Primer name	Forward Primer(5′-3′)	Reverse primer (5′-3′)	Purpose
AST	ATCCTGGGAGTGACTGAGGCTTTC	GCTTGCCACTGTCATCACGGTAA	qRT-PCR
PAOX	GGCGTCTCCAACAAGATCCAACTT	CACTTGCCACGTTCACGATGTCT	qRT-PCR
1**α-HYD**	GCTTAGCAGTGATGTGGTGATGTCT	ATCGCTCAGGAAGGAACTCGTCT	qRT-PCR
NOS	TGCCAGAGTGAAGATGTCCTCCTT	GCCACTGATTGTCCATCGGTAGAAT	qRT-PCR
DHP	AGTGCTGCTGACATAGTCTCTGCTA	GTGCCAGTAGTGGTTACCTGTTGT	qRT-PCR
GAT2	TCGTGACTGCCATCGTTGACTTG	TACATTCCACCCTGCGTGACCAT	qRT-PCR
SMPD	CGCCTGTCAATAGTTTCCCACCTC	CTGCCATTGTCTGTCCAACTCATCA	qRT-PCR
KYNU	AGACTGTGCAATCCTCCTCCTATCC	CTGCTTCTCCACAATTCGGTCCATT	qRT-PCR
KSR2	AGCACCAGTCAGTCAGCGTAGT	AAGGGTAGAGGAAGTGTGGAGAGG	qRT-PCR
PCK	GTGCCTGAGACGAACGAGATCATC	AGCATGTGTTCCGCCAACCATC	qRT-PCR
EF1-α	AAGCCAGGTATGGTTGTCAACTTT	CGTGGTGCATCTCCACAGACT	Reference gene

Results

Transcriptome assembly

After Illumina sequencing, low-quality and contaminating reads were filtered in the raw data. The immature ovary and mature ovary obtained 122,677,356 and 126,930,488 clean reads, the clean base numbers of the two groups were 19.04 G and 19.50 G, respectively. The Q30 of both groups was >92%, and the GC content of sequence bases was between 45.31% and 48.14% (**Table 2**). In the analysis using Trinity software, 28,065 unigenes were identified, with the longest length of 20,335 bp, the shortest length of 201 bp, and the N50 length of 2591 bp (**Table 3**). Most unigenes sequences and transcripts ranged from <300 bp, 10,338

transcripts were >1000 bp, and 10,265 unigenes were >1000 bp (**Figure 1**). These results indicated that the data quality was high and that unigenes were suitable for further analysis.

Table 2 Sequencing data statistics.

Sample	Raw reads	Clean reads	Clean bases	Q30%	GC%
I_1	44,037,684	43,260,110	6.94G	93.12%	48.14%
I_2	46,260,800	45,549,508	6.83G	92.82%	47.80%
I_3	38,801,802	38,120,870	5.72G	92.38%	47.11%
M_1	42,890,866	42,241,208	6.34G	92.87%	47.12%
M_2	42,617,366	41,838,862	6.28G	92.70%	46.68%
M_3	39,303,116	38,597,286	5.79G	92.62%	45.31%

I and M represent female ovaries of immature stage ovary and mature stage ovary, respectively. _1, _2 and _3 represent three independent biological replicates for each group. Q30 indicates that the base calling accuracy rate is greater than 99.9%.

Table 3 Distribution of splicing lengths.

	Transcripts	Unigenes
Total number	32,833	28,065
Total length	46,249,254	35,882,966
Max length	20,335	20,335
Median length	661	589
Min length	201	201
N50	2819	2591
N90	535	470

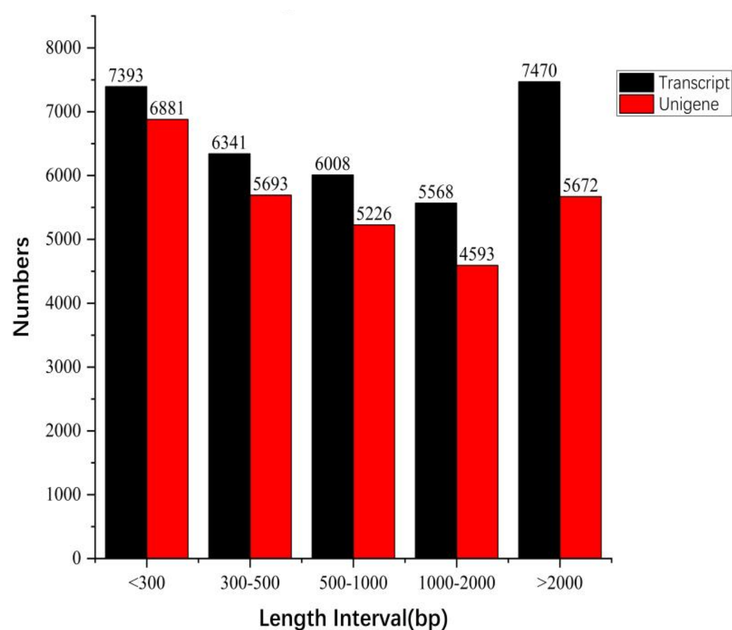

Figure 1 Length distribution of transcripts and unigenes

Functional Annotation and Classification

To evaluate the function of single gene sequences, gene function annotation was performed in seven databases, namely Nr, Nt, Pfam, KOG, Swiss-Prot, KEGG, and GO. 9026 (32.16%), 1962 (6.99%), 8244 (29.37%), 10,647 (37.94%), 10,551 (37.59%), 10,647 (37.94%), 10,551 (37.59%), 12,886 (45.91%) and 4660 (16.60%) genes with sequence homology. Of the 28,065 single genes, 15,222 (54.24%) were annotated in at least one database, and 681 genes (2.43%) were annotated in all databases (**Table 4**).

Table 4 Statistical of seven databases

Database	Number of Unigenes	Percentage (%)
Annotated in NR	10551	37.59
Annotated in NT	8244	29.37
Annotated in KO	1962	6.99
Annotated in SwissProt	10647	37.94
Annotated in Pfam	12886	45.91
Annotated in GO	4660	16.60
Annotated in COG/KOG	9026	32.16
Annotated in all Databases	681	2.43
Annotated in at least one Database	15222	54.24
Total Unigenes	28065	100

In this analysis by Gene Ontology, 4660 (16.60%) genes were successfully annotated into three main functional categories of GO: biological processes (26 functional groups, 12,407), cellular components (18 functional groups, 10,517), and molecules function (12 function groups, 4197) (**Figure 2**). The number of single genes of biological processes and cellular components was significantly larger than the number of single genes determined by GO analysis. One interpretation of these results is that a single gene is assigned to multiple functional groups. Among biological processes, cellular processes (2553) and metabolic processes (2162) are the most representative terms. Cells (2373) and cell fractions (2286) were significantly enriched in cellular components, as were binding (1763) and catalytic activity (1777) in molecular functions.

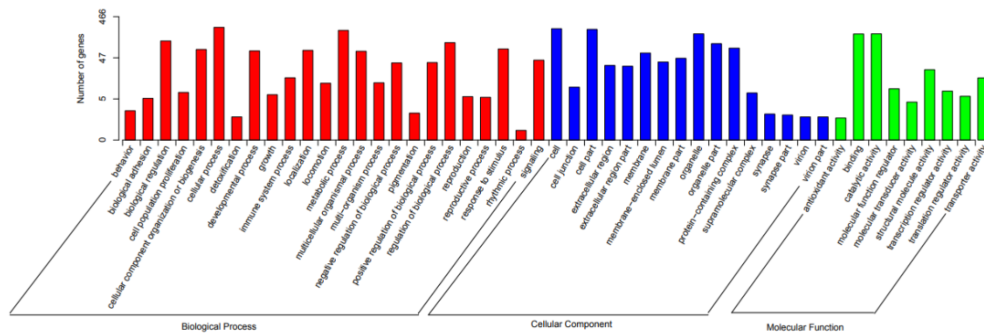

Figure 2 Gene Function Classification (GO)

In the KOG database, 9026 single genes were grouped into 26 clusters. The most abundant clusters were only (R) general function predictions (2078, 23.02%), followed by (T) signal

transduction mechanisms (1243, 13.77%) and (O) post-translational modifications, protein turnover and chaperones (883, 9.78%) (**Figure 3**).

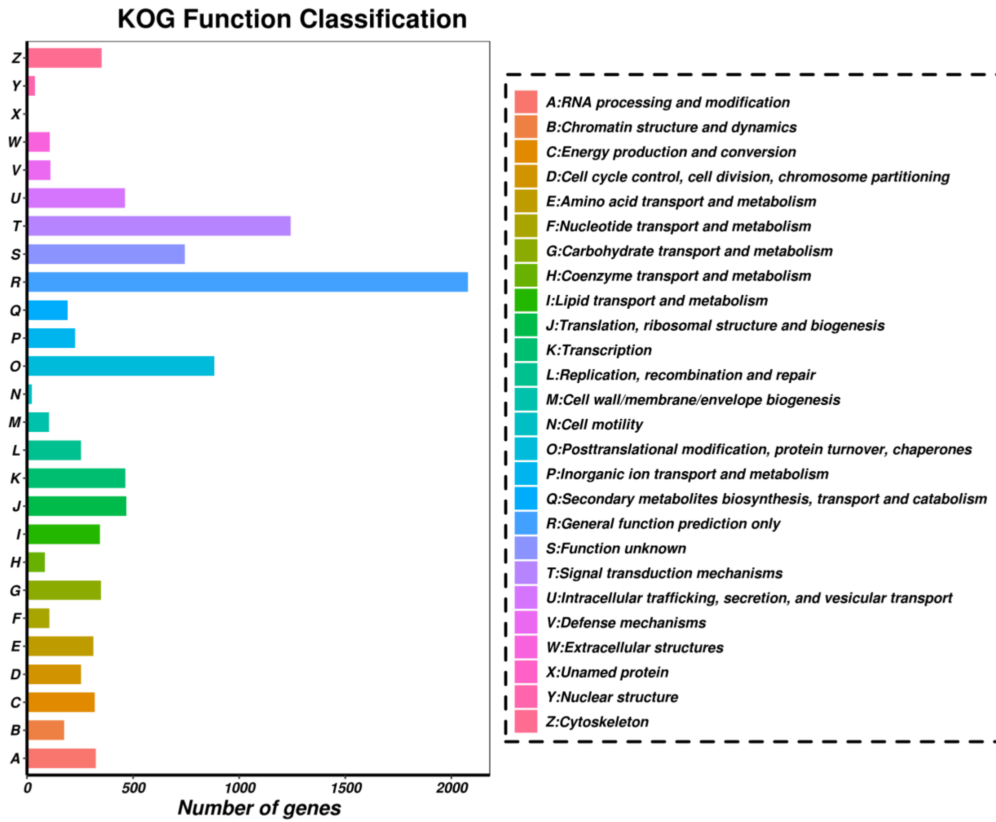

Figure 3 KOG Function Classification

The KEGG database divides 901 single genes involved in metabolic pathways into 5 main clusters (32 subclusters): A. Cellular processes (4 subclusters, 98); B. Environmental information processing (3 subclusters, 59); C, genetic information processing (4 subclusters, 299); D, metabolism (12 subclusters, 309); and E, organic systems (9 subclusters, 136) (**Figure 4**).

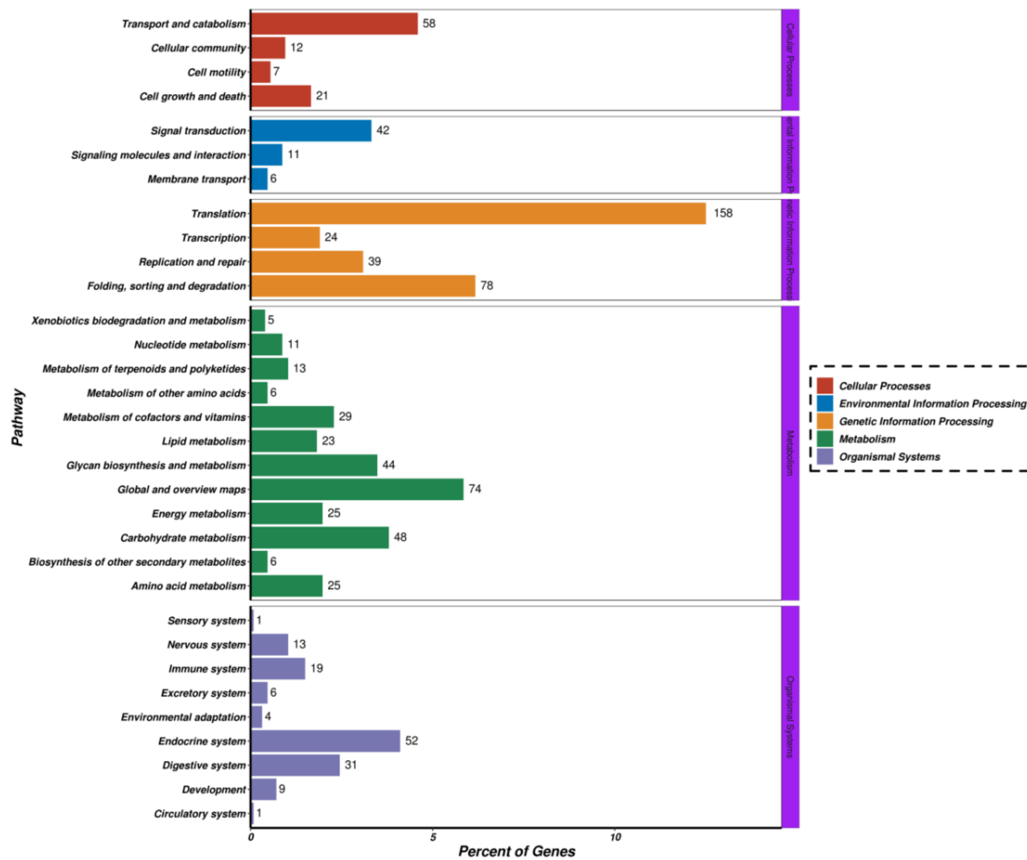

Figure 4 KEEG Function Classification

DEGs expression analysis

Gene expression was changed in mature ovaries compared to immature ovaries. 394 DEGs were identified from immature and mature stage ovaries, of which 136 were up-regulated, and 258 were down-regulated (**Figure 5**).

Enrichment analysis of GO and KEGG metabolic pathways

Mapping DEGs to GO terms, the most enriched terms were ion transport, nucleoside biosynthesis process, nucleoside phosphate biosynthesis process, nucleoside phosphate metabolic process, and transporter activity. In organisms, different genes coordinate with each other to perform their biological functions. The most important biochemical metabolic pathways and signal transduction pathways involved in DEGs can be determined through the significant enrichment of Pathway. KEGG pathway analysis was performed on assembled single genes to reveal biochemical pathways. 394 DEGs were annotated as 105 distinct pathways. Among them, the number of sequences related to metabolic pathways was the largest (53, 13.5%), followed by oxidative phosphorylation (14, 3.6%). In addition, some DEGs were also localized to multiple pathways related to reproduction and development, such as MAPK signaling pathway, insulin signaling pathway, cell cycle pathway, ribosomal pathway, retinol metabolism, steroid hormone biosynthesis, and GnRH signaling pathway. These pathways may play important roles in studying specific biochemical and developmental processes (**Figure 6**).

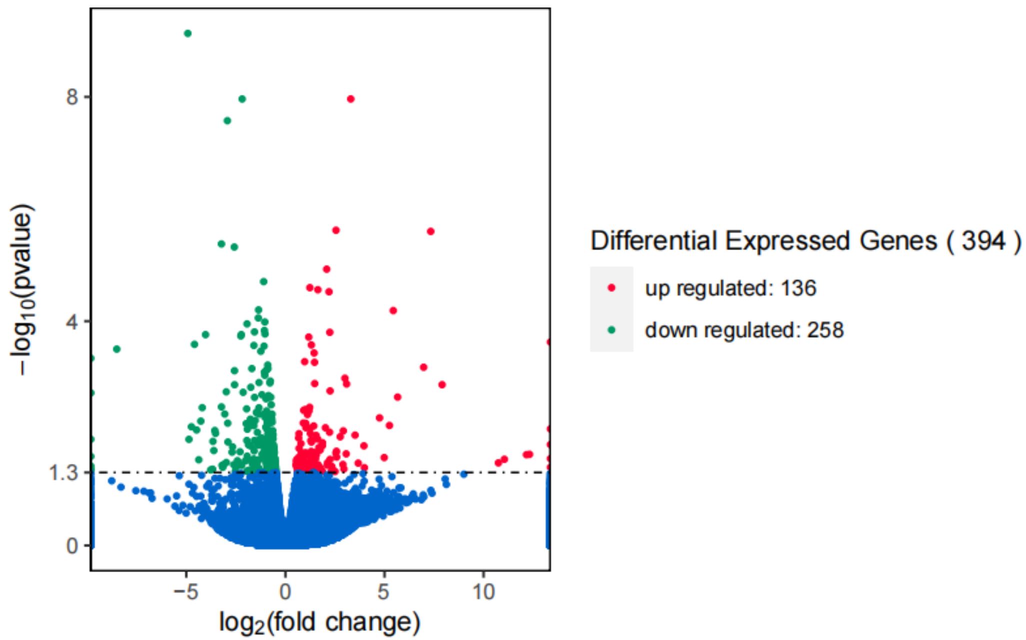

Figure 5 Differential gene volcano map

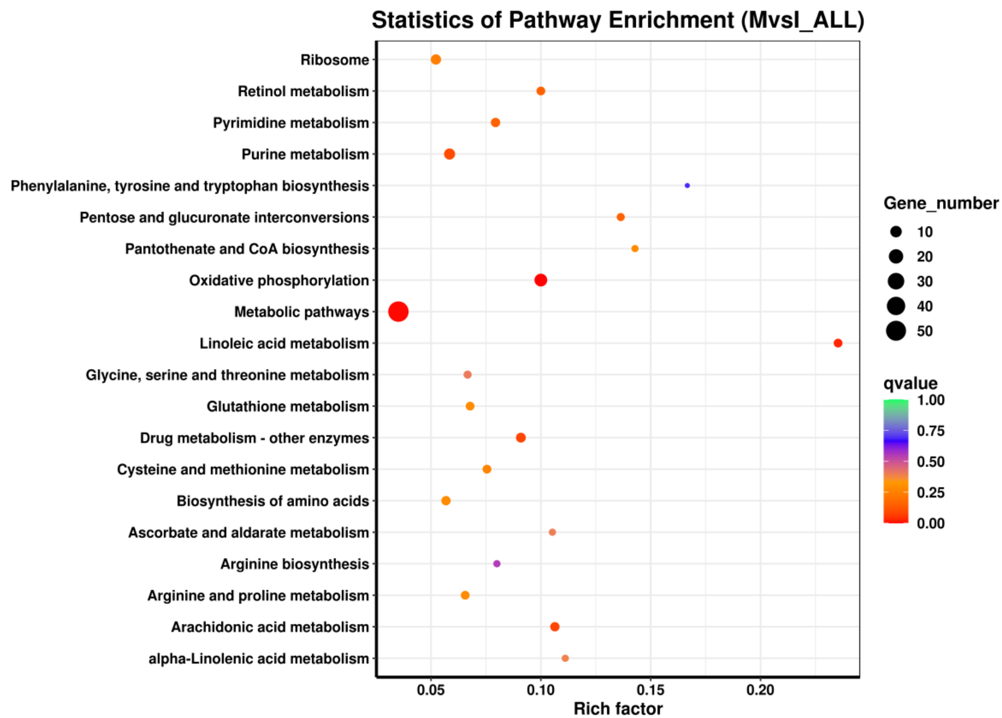

Figure 6 Statistics of Pathway Enrichment

Validation of transcriptome data by qRT-PCR

Five up-regulated and five down-regulated genes were selected for qRT-PCR validation of RNA-seq. Overall, the expression patterns were consistent with RNA-seq trends (**Figure 7**). The result showed that RNA-seq analysis could accurately and reliably identify the differential gene expression in the ovarian tissue development of *M. affinis*.

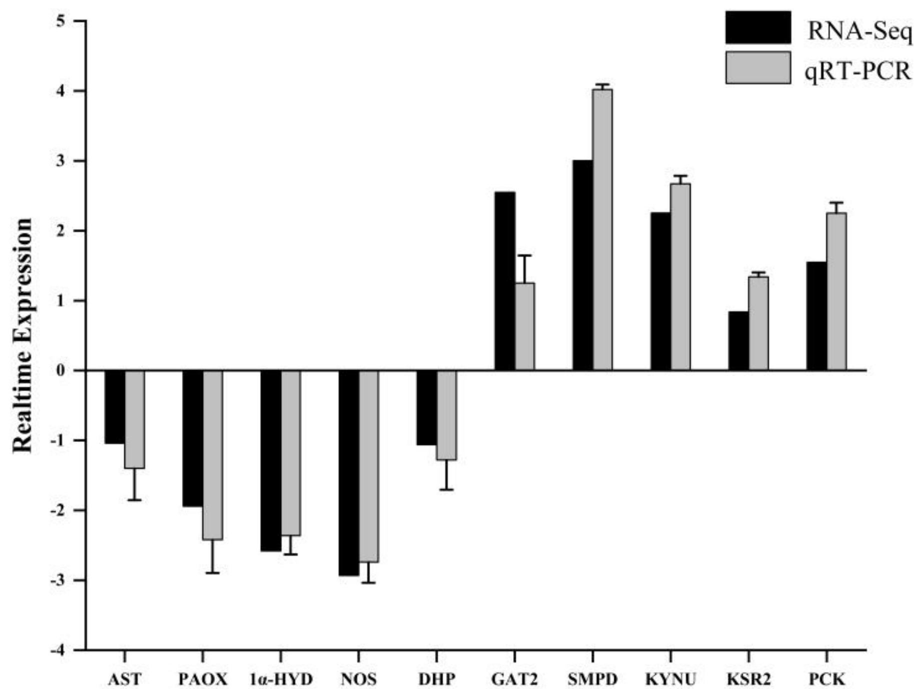

Figure 7 qRT-PCR validation of DEGs

Discussion

To investigate the gene expression in the ovary of *M. affinis* during growth and spawning, we screened key regulatory genes related to these two developmental stages. Clean data of 38.54 G was obtained from ovarian tissue by RNA-Seq, and 28,065 unigenes were identified with Q30>92%, so the sequencing quality was considered reliable. The N50 length was 2591 bp, which indicated good assembly. Further biological function annotation of unigenes showed that 15,222 unigenes were homologous to known genes, accounting for 54.24% of the entire assembled transcripts, and the annotation rate was more than half. The remaining 45.76% of the sequences have no annotation information, but these genes may be involved in important biological processes, which need to be further studied.

Analysis of transcriptome results showed that many genes in the energy metabolism pathway were down regulated. Oxidative phosphorylation occurs in the mitochondrial respiratory chain and is an important biochemical process of cell energy metabolism. In the oxidative phosphorylation pathway, most genes are down regulated, including cytochrome c oxidase gene (COX), acyl carrier protein gene, ATP synthase and NADH dehydrogenase gene (ubiquinone). These genes are important genes in the electron transfer chain and play an important role in energy transfer (Fontanesi et al., 2008). Cox plays a key role in the cellular energy production of mitochondrial respiratory chain. COX transfers electrons to molecular oxygen, which is then coupled with proton transfer from the mitochondrial matrix to the intermembrane space to form H_2O and release ATP to meet the body's normal energy needs; COX also promotes energy storage in the form of an electrochemical gradient, This electrochemical gradient will be used by the oxidative phosphorylation system for ATP synthesis (Verner et al., 2014). The down-regulation of these genes showed that the energy metabolism of *M. affinis* changed during ovarian development. Metabolism of energy is generally considered to be related to growth rate (Dayton et al., 2007; Vahl, 1984), and the expression of COX genes was higher in immature ovaries than in mature ovaries, suggesting a large storage of energy in immature ovaries, which ultimately led to rapid ovarian

maturation. However, the true expression levels of COX genes in ovaries at different stages need further study.

Cells in organisms produce reactive oxygen species (ROS) during aerobic metabolism. Studies have shown that the production of reactive oxygen species can induce oocyte maturation in the follicle and promote the development of the ovary, and is a key signaling molecule involved in the biological reproduction process (Behrman et al., 2001; Arenas-Rios et al., 2007). At the same time, its excessive accumulation will cause damage to cellular components such as proteins and DNA (Lu et al., 2018). Cells have a well-established system of enzyme scavengers, including antioxidant enzymes such as superoxide dismutase (SOD), catalase (CAT), and glutathione peroxidase (GPX), which can reduce excess ROS oxidative damage caused. SOD acts as the first line of defense against the generation of reactive oxygen species in the antioxidant system, catalyzing the disproportionation of superoxide radicals to H_2O_2, CAT catalyzing the decomposition of H_2O_2 into H_2O and O_2, and GPX catalyzing the reduction of H_2O_2 or organic hydroperoxides to water or the corresponding alcohols (Al-Shehri, 2021; Rubio-Riquelme et al., 2020; Gao et al., 2014). GPX was only expressed in the ovaries of *Metapenaeus ensis* and at high levels was early ovaries but not detected in late yolk-forming oocytes (Wu et al., 2010). This was similar to the expression pattern of GPX in this study, where GPX gene expression was down-regulated in mature ovaries compared with immature ovaries. At the same time, SOD expression was also down-regulated, while CAT gene expression did not change between the two groups. Under the regulation of antioxidant enzymes, ROS generation and disproportionation reached a relative balance to avoid oxidative damage to ovarian cells. The low expression of SOD gene and GPX gene disrupted this balance, resulting in excessive accumulation of ROS, which further promoted ovarian maturation as a signaling molecule.

The development and maturation of the ovary is influenced by factors such as gene expression and the external environment, and involves the regulation of molecular networks of many signaling pathways. Rapid growth is always accompanied by frequent cell division and gene expression (Krieger, 1978), which is reflected in the enrichment of DEGs in genetic information processing pathways. In this study, many genes related to reproductive activities such as oocyte proliferation, development and maturation were identified, including ubiquitin protein ligase (E3s), proliferating cell nuclear antigen (PCNA), serine-threonine protein kinase (STPK), eukaryotic translation initiation factor, translationally-controlled tumor protein, heat shock protein and cAMP-dependent protein kinase catalytic subunit.

Genes related to the ubiquitin system have been studied, which play important roles in the reproductive process. In the ovarian cDNA library of Crayfish, the ubiquitin-related homologous genes have different expression levels, which may play a role in the reproductive process of *Procambarus clarkii* (Jiang et al., 2014). The expression of ubiquitin-ribosomal proteins S27 and L40 is significantly altered in the developing ovary of *Saccharomyces cerevisiae*, which played a key role in gametogenesis and ovarian development (Wang et al., 2012). The ubiquitin-proteasome pathway (UPP) is involved in the life activities of animals and plants, including cell proliferation, differentiation, regulation of the cell cycle, etc., and is important for every aspect of cell life (Nabavi et al., 2018; Gao and Karin., 2005). This system typically includes three ubiquitinases: ubiquitin-activating enzymes (E1s), ubiquitin-conjugating enzymes (E2s) and E3s. E3 ubiquitin protein ligase is mainly located in the chromatin region of meiosis, and participates in protein degradation by binding ubiquitin-conjugating enzyme E2 and substrates. In our ovarian transcriptome database, E3 ubiquitin protein ligase was found, and it was highly expressed in mature ovaries, and it was speculated that it was closely related to the development of the ovary of *M. affinis*.

PCNA is widely present in proliferating cells of various biological tissues and is considered to be an important part of eukaryotic replicators. It participates in DNA replication through the coenzyme form of DNA polymerase, thereby regulating cell proliferation (Moldovan et al., 2007; Bravo et al., 1987). As a cell proliferation marker protein, PCNA regulates the growth

and development of follicles in vertebrates, indicating the growth of oocytes. For example, the number of cells in PCNA is a standard indicator to measure the growth of follicular epithelium at different stages of zebrafish oogenesis (Korfsmeier, 2002). Among crustaceans, studies have shown that PCNA of *Penaeus monodon* is highly expressed in the ovary and its expression peaks in the third stage of ovarian development. It is speculated that PCNA may play an important role in regulating the ovarian development of *P. monodon* (Wang et al., 2016). The expression of PCNA in *Marsupenaeus japonicus* was lowest in stage Ⅰ and highest in stage Ⅱ the in developing testis and ovary. The data suggest that PCNA plays an important role in gonad development in *M. japonicus*, especially during mitosis and meiosis (Zhang et al., 2010). In this study, PCNA expression was lower in mature ovaries than in immature ovaries, suggesting a reduction in cellular proliferative activity. Elongation factor (EF-1ɑ) and eukaryotic translation initiation factors (EIFS) as genes that play a very important role in promoting cell protein translation (Jakobsson et al., 2018; Hao et al., 2020). The down-regulation of their expression further indicated that they did not promote cell proliferation. This may be that ovarian development had reached the late stage and was fully mature, but anyway, the change of its expression pattern indicated that it was related to ovarian development.

Conclusion

This is the first report on the transcriptomic analysis of immature ovary and mature ovary tissue of *M. affinis*. A transcriptome library was created using RNA-Seq technology, a total of 28,065 unigenes were detected, and 394 DEGs were obtained from differential expression analysis. Some candidate genes related to gonadal development, such as Cox, GPX, E3s, PCNA, STPK, were preliminarily identified, and their specific roles in the development process were attempted to elucidate. The next step will be the full-length amplification and functional verification of candidate genes. In conclusion, the analysis of transcriptome data in this study not only provides information for the study of gene expression in the ovarian tissue of *M. affinis*, but also contributes important research data to the molecular regulation mechanism in the process of development and maturation, which will help to better understand the regulation of reproductive cycle of *M. affinis*.

Acknowledgments

This work was supported by the China Agriculture Research System of MOF and MARA(CARS-48); the Financial Fund of the Ministry of Agriculture and Rural Affairs of China (NHYYSWZZZYKZX2020); Guangdong Provincial Special Fund for Modern Agriculture Industry Technology Innovation Teams (NO. 2019KJ149); Central Public-interest Scientific Institution Basal Research Fund, CAFS (NO. 2020TD30).

References

Al-Shehri, SS., 2021. Reactive oxygen and nitrogen species and innate immune response. *Biochimie.* 181, 52-64. https://doi.org/10.1016/j.biochi.2020.11.022.

Arenas-Rios, E., Leon-Galvan, MA., Mercado, PE., Lopez-Wilchis, R., Cervantes, DLMI., Rosado, A., 2007. Super-oxide dismutase, catalase, and glutathione peroxidase in the testis of the Mexican big-eared bat (*Corynorhinus mexicanus*) during its annual reproductive cycle. *Comp. Biochem. Phys. A.* 148, 150-158. https://doi.org/10.1016/j.cbpa.2007.04.003.

Ayub, Z., Ahmed, M., 2002. A description of the ovarian development stages of penaeid shrimps from the coast of Pakistan. *Aquac. Res.* 33, 767-776. https://doi.org/10.1046/j.1365-2109.2002.00715.x.

Bar, I., Cummins, S., Elizur, A., 2016. Transcriptome analysis reveals differentially expressed genes associated with germ cell and gonad development in the Southern bluefin tuna (*Thunnus maccoyii*). *BMC Genomics.* 17, https://doi.org/10.1186/s12864-016-2397-8.

Behrman, HR., Kodaman, PH., Preston, SL., Gao, SP., 2001. Oxidative stress and the ovary. *J. Soc. Gynecol. Invest.* 8, S40-S42. https://doi.org/10.1016/S1071-5576(00)00106-4.

Bissonnette, N., Lesque-Sergerie, JP., Thibault, C., Boissonneault, G., 2009. Spermatozoal transcriptome profiling for bull sperm motility: a potential tool to evaluate semen quality. *Reproduction.* 138, 65–80. https://doi.org/10.1530/REP-08-0503.
Bravo, R., Frank, R., Blundell, P., Macdonald-Bravo, H., 1987. Cyclin/PCNA is the auxiliary protein of DNA polymerase-δ. *Nature*. 326, 515–517. https://doi.org/10.1038/326515a0.
Cahais, V., Gayral, P., Tsagkogeorga, G., Melo-Ferreira, J., Ballenghien, M., Weinert, L., Chiari, Y., Belkhir, K., Ranwez, V., Galtier, N., 2012. Reference-free transcriptome assembly in non-model animals from next-generation sequencing data. *Mol. Ecol. Resour.* 12, 834–845. https://doi.org/10.1111/j.1755-0998.2012.03148.x.
Chalmel, F., Rolland, A.D., Niederhauser-Wiederkehr, C., Chung, S.S.W., Demougin, P., et al., 2007. The conserved transcriptome in human and rodent male gametogenesis. *Proc. Natl. Acad. Sci.* 104, 8346–8351. https://doi.org/10.1073/pnas.0701883104.
Chu, SL., Weng, CF., Hsiao, CD., Hwang, PP., Chen, YC., Ho, JM., Lee, SJ., 2005. Profile analysis of expressed sequence tags derived from the ovary of tilapia, *Oreochromis mossambicus*. *Aquaculture.* 251, 537-548. https://doi.org/10.1016/j.aquaculture.2005.05.040.
Dayton, W.R., White, M.E., 2007. Cellular and molecular regulation of muscle growth and development in meat animals. *J. Anim. Sci.* 86, E217–E225. https://doi.org/10.2527/jas.2007-0456.
Feng, HY., Dong, YT., Liu, X., Qiu, GF., 2020. Cyclin B protein undergoes increased expression and nuclear relocation during oocyte meiotic maturation of the freshwater prawn Macrobrachium rosenbergii and the Chinese mitten crab *Eriocheir sinensis*. Gene. 758, 144955-144955. https://doi.org/10.1016/j.gene.2020.144955.
Fontanesi, F., Soto, IC., Barrientos, A., 2008. Cytochrome c oxidase biogenesis: new levels of regulation. *IUBMB life.* 60, 557-68. https://doi.org/10.1002/iub.86.
Gao, M., Karin, M., 2005. Regulating the regulators: control of protein ubiquitination and ubiquitin-like modifications by extracellular stimuli. *Mol. Cell.* 19, 581–593. https://doi.org/10.1016/j.molcel.2005.08.017.
Gao, F., Chen, J., Ma, TT., Li, HY., Wang, N., Li, ZL., Zhang, ZC., Zhou, YJ., 2014. The Glutathione Peroxidase Gene Family in Thellungiella salsuginea: Genome-Wide Identification, Classification, and Gene and Protein Expression Analysis under Stress Conditions. *Int. J. Mol. Sci.* 15, 3319-3335. https://doi.org/10.3390/ijms15023319.
González-Castellano, I., Manfrin, C., Pallavicini, A. Andrés Martínez-Lage., 2019. De novo gonad transcriptome analysis of the common littoral shrimp *Palaemon serratus*: novel insights into sex-related genes. *BMC Genomics.* 20, 757. https://doi.org/10.1186/s12864-019-6157-4.
Guo, Q., Li, SH., Lv, XJ., Xiang, JH., Manor, R., Sagi, A., Li, FH., 2020. Sex-Biased CHHs and Their Putative Receptor Regulate the Expression of IAG Gene in the Shrimp *Litopenaeus vannamei*. *Front. Physiol.* 10, 1525. https://doi.org/10.3389/fphys.2019.01525.
Grabherr, MG., Haas, BJ., Yassour, M., Levin, JZ., Thompson, DA., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, QD., Chen, ZH., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, BW., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644-652. https://doi.org/10.1038/nbt.1883.
Hao, PQ., Yu, JJ., Ward, R., Lin, L., Qiao, H., Su, A., Tianrui, X., 2020. Eukaryotic translation initiation factors as promising targets in cancer therapy. *Cell Commun. Signaling.* 18, 175. https://doi.org/10.1186/s12964-020-00607-9.
Holthuis, L. B. 1980. FAO species catalogue: Shrimps and prawns of the world. Anannotated catalogue of species of interest to fisheries. FAO Species Catalogue, 1. *FAO Fish. Synop.,* l (125), 1-271. http://refhub.elsevier.com/S1687-4285(21)00101-1/h0065.
Jakobsson, ME., Malecki, J., Falnes, PO., 2018. Regulation of eukaryotic elongation factor 1 alpha (eEF1A) by dynamic lysine methylation. *RNA Biology.* 15, 314-319, https://doi.org/10.1080/15476286.2018.1440875.
Jiang, HC., Xing, ZJ., Wei, L., Qian, ZJ., Yu, HW., Li, JL., 2014. Transcriptome analysis of red swamp crawfish *Procambarus Clarkii* reveals genes involved in gonadal development. *PLoS ONE.* 9, e105122. https://doi.org/10.1371/journal.pone.0105122.
Jin, SB., Hu, YN., Fu, HT., Jiang,SF., Xiong, YW., Qiao,H., Zhang, WY., Gong, YS., Wu, Y., 2019. Potential Functions of Gem-Associated Protein 2-Like Isoform X1 in the Oriental River Prawn *Macrobrachium nipponense*: Cloning, qPCR, In Situ Hybridization, and RNAi Analysis. *Int. J. Mol. Sci.* 20, 3995. https://doi.org/10.3390/ijms20163995.

Korfsmeier, KH., 2002. PCNA in the ovary of zebrafish (Brachydanio rerio, Ham.Buch.). Acta Histochem. 104, 73-76. https://doi.org/10.1078/0065-1281-00632.
Krieger, I., 1978. Relation of specific dynamic action of food (SDA) to growth in rats. Am. J. Clin. Nutr. 31, 764–768. https://doi.org/10.1093/ajcn/31.5.764.
Langmead, B., Salzberg, SL., 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods.* 9, 357–359. https://doi.org/10.1038/NMETH.1923.
Li, GL., Qian, H., 2017. Transcriptome using Illumina sequencing reveals the traits of spermatogenesis and developing testes in *Eriocheir sinensis*. *PLoS ONE.* 12, e0172478. https://doi.org/10.1371/journal.pone.0172478.
Livak, K.J., Schmittgen, T.D., 2001. Analysis of Relative Gene Expression Data using Real-Time Quantitative PCR. *Methods.* 25, 402–408. https://doi. org/10.1006/meth.2001.1262.
Lo, T.S., Cui, Z., Mong, J.L.Y., Wong, Q.W.L., Chan, S.M., Kwan, H.S., Chu, K.H., 2007. Molecular coordinated regulation of gene expression during ovarian development in the penaeid shrimp. *Mar. Biotechnol*. 9, 459-468. https://doi.org/10.1007/s10126-007-9006-4.
Lu, JY., Wang, ZX., Cao, J., Chen, YX., Dong, YL., 2018. A novel and compact review on the role of oxidative stress in female reproduction. *Reprod. Biol. Endocrinol.* 16, 80. https://doi.org/10.1186/s12958-018-0391-5.
Mao, X., Cai, T., Olyarchuk, J.G., Wei, L., 2005. Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics.* 21, 3787–3793. https://doi.org/10.1093/bioinformatics/bti430.
Mardis, ER., 2008. The impact of next-generation sequencing technology on genetics. *Trends. Genet.* 24, 133-141. https://doi.org/10.1016/j.tig.2007.12.007.
Meng, Xl., Liu, P., Jia, Fl., Li, J., Gao, BQ., 2015. De novo Transcriptome Analysis of *Portunus trituberculatus* Ovary and Testis by RNA-Seq: Identification of Genes Involved in Gonadal Development. *PLoS ONE.* 10, e0128659. https://doi.org/10.1371/journal.pone.0128659.
Moldovan, GL., Pfander, B., Jentsch, S., 2007. PCNA, the maestro of the replication fork. *Cell.* 129, 665–679. https://doi.org/10.1016/j.cell.2007.05.003.
Nabavi, S., Atanasov, AG., Khan, H., Barreca, D., Trombetta, D., Testai, L., Sureda, A., Tejada, S., Vacca, RA., Pittala, V., Gulei, D., Berindan-Neagoe, I., Shirooie, S., Nabavi, SM., 2018. Targeting ubiquitin-proteasome pathway by natural, in particular polyphenols, anticancer agents: Lessons learned from clinical trials. *Cancer Lett.* 434, 101-113. https://doi.org/10.1016/j.canlet.2018.07.018.
Palomares M. L. D. and D. Pauly. Editors. 2021. Sea Life Base. *World Wide Web* electronic publication. <http://www.sealifebase.org>, version (08/2021).
Peng, JX., Wei, PY., Zhang, B., Zhao, YZ., Zeng, DG., Chen, XL., Li, M., Chen, XH., 2015. Gonadal transcriptomic analysis and differentially expressed genes in the testis and ovary of the Pacific white shrimp (*Litopenaeus vannamei*). *BMC Genomics.* 16, 1006. https://doi.org/10.1186/s12864-015-2219-4.
Ponza, P., Yocawibun, P., Sittikankaew, K., Hiransuchalert, R., Yamano, K., Klinbunga, S., 2011. Molecular cloning and expression analysis of the Mitogen-activating protein kinase 1 (MAPK1) gene and protein during ovarian development of the giant tiger shrimp *Penaeus monodon*. *Mol. Reprod. Dev.* 78, 47–60. https://doi.org/10.1002/mrd.21310.
Powell, D., Knibb, W., Remilton, C., Elizur, A., 2015. De-novo transcriptome analysis of the banana shrimp (*Fenneropenaeus Merguiensis*) and identification of genes associated with reproduction and development. *Mar. Genomics.* 22, 71-78. https://doi.org/10.1016/j.margen.2015.04.006.

Qiao, H., Fu, HT., Xiong, YW., Jiang, SF., Zhang, WY., Sun, SM., Jin, SB., Gong, YS., Wang, YB., Shan, DY., Li, F., Wu, Y., 2017. Molecular insights into reproduction regulation of female Oriental River prawns *Macrobrachium nipponense* through comparative transcriptomic analysis. *Sci. Rep.* 7, 1-11. https://doi.org/10.1038/s41598-017-10439-2.
Roberts, A., Pachter, L., 2013. Streaming fragment assignment for real-time analysis of sequencing experiments. *Nat. Methods.* 10, 71-73. https://doi.org/10.1038/NMETH.2251.
Rubio-Riquelme, N., Huerta-Retamal, N., Gomez-Torres, MJ., Martinez-Espinosa, RM., 2020. Catalase as a Molecular Target for Male Infertility Diagnosis and Monitoring: An Overview. *Antioxidants*. 9, 78-78. https://doi.org/10.3390/antiox9010078.
Safaie, M., 2009. Histological study of singa shrimp ovaries (*Metapenaeus affinis*) in coastal water of the hormozgan province. *Pajouhesh and Sazandegi.* 81, 168-171.

Santos, CA., Blanck, DV., de Freitas, PD., 2014. RNA-seq as a powerful tool for penaeid shrimp genetic progress. Front. *Genet.* 5, 298. https://doi.org/10.3389/fgene.2014.00298.

Tolga Dincer, İlker Aydin, 2014. Proximate composition and mineral and fatty acid profiles of male and female jinga shrimps (*Metapenaeus affinis*, H. Milne Edwards, 1837). *Turk. J. Vet. Anim. Sci.* 38, 445-451. https://doi.org/10.3906/vet-1301-15.

Vahl, O. 1984. The relationship between specific dynamic action (SDA) and growth in the common starfish, *Asterias rubens* L. *Oecologia.* 61, 122–125. https://doi.org/10.1007/BF00379097.

Verner, Z., Cermakova, P., Skodova, I., Kovacova, B., Lukes, J., Horvath, A., 2014. Comparative analysis of respiratory chain and oxidative phosphorylation in *Leishmania tarentolae*, *Crithidia fasciculata*, *Phytomonas serpens* and procyclic stage of *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 193, 55-65. https://doi.org/10.1016/j.molbiopara.2014.02.003.

Wan, HF., Liao, JQ., Zhang, ZP., Zeng, XY., Liang, KY., Wang, YL., 2021. Molecular cloning, characterization, and expression analysis of a sex-biased transcriptional factor sox9 gene of mud crab *Scylla paramamosain*. *Gene.* 774, 145423. https://doi.org/10.1016/j.gene.2021.145423.

Wang, CY., Zhao, C., Fu, MJ., Qiu, LH., 2016. Molecular cloning and expression analysis of PCNA from black tiger shrimp, *Penaeus monodon*. *Journal of Fishery Sciences of China.* 23, 1052-1062. https://doi.org/10.3724/SP.J.1118.2016.15449.

Wang, Q., Chen, L., Wang, Y., Li, W., He, L., Jiang, H., 2012. Expression characteristics of two ubiquitin/ribosomal fusion protein genes in the developing testis, accessory gonad and ovary of Chinese mitten crab. *Eriocheir sinensis*. *Mol. Biol. Rep.* 39, 6683-6692. https://doi.org/10.1007/s11033-012-1474-6.

Wang, YB., Jin, SB., Fu, HT., Qiao, H., Sun, SM., Zhang, WY., Jiang, SF., Gong, YS., Xiong, YW., Wu, Y., 2019. Molecular cloning, expression pattern analysis, and in situ hybridization of a Transformer-2 gene in the oriental freshwater prawn, *Macrobrachium nipponense* (de Haan, 1849). 3 *Biotech.* 9, 205. https://doi.org/10.1007/s13205-019-1737-1.

Wang, Z., Gerstein, M., Snyde, M., 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10, 57-63. https://doi.org/10.1038/nrg2484.

Wu, LT., Chu, KH., 2010. Characterization of an ovary-specific glutathione peroxidase from the shrimp *Metapenaeus ensis* and its role in crustacean reproduction. *Comp. Biochem. Phys. B.* 155, 26-33. https://doi.org/10.1016/j.cbpb.2009.09.005.

Young, MD., Wakefield, MJ., Smyth, GK., Oshlack, A., 2010. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 11. https://doi.org/10.1186/gb-2010-11-2-r14.

Zhang, ZP., Shen, BL., Wang, YL., Chen, Y., Wang, GD., Lin, P., Zou, ZH., 2010. Molecular Cloning of Proliferating Cell Nuclear Antigen and Its Differential Expression Analysis in the Developing Ovary and Testis of Penaeid Shrimp *Marsupenaeus japonicus*. *DNA Cell Biol.* 29, 163-170. https://doi.org/10.1089/dna.2009.0958.

Zhuang, MG., Jiang, SG., Zhou, FL., Huang, JiH., Yang, QB., Jiang, S., Yang, LS., 2020. Molecular cloning and multifunction exploration of CFSH gene in *Penaeus monodon*. *South China Fish. Sci.* 16, 90-99. https://doi.org/10.12131/20200011.