

Transcriptome Analysis of The Whole Body Provides Novel Insights Into The Molecular Mechanisms of Growth, Immunity, and Metabolism During The Molting Cycle of The Juvenile Mud Crab *Scylla Paramamosain*

Lei Liu

Ningbo University

Xiao Liu

Ningbo University

Yuanyuan Fu

Ningbo Institute of Oceanography

Wei Fang

Ningbo University

Chunlin Wang (✉ wangchunlin@nbu.edu.cn)

Ningbo University

Research Article

Keywords: *Scylla paramamosain*, Transcriptome, Molting cycle, Molecular mechanism, Crustacean, Hormone regulation, Cuticle reconstruction

Posted Date: January 18th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-146494/v1>

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

[Read Full License](#)

Abstract

Background: Molting is an inherent cyclic process that occurs in all crustaceans and is essential for growth, metamorphosis, and reproduction. Recent studies have shown that many organs and systems are involved in molting regulation, including the hepatopancreas, the immune system, and production of ecdysteroids by Y-organs. However, the molecular mechanisms underlying the dynamic process of crab molting remain poorly understood at the whole-individual level. Therefore, we investigated global expression changes in the transcriptomes of the mud crab *Scylla paramamosain*, the most cultured crab species worldwide. We revealed a cascade of sequential expression events of molting-related genes involved in various aspects of the molting process by using whole-body sequencing of juvenile crabs.

Results: RNA-sequencing (RNA-seq) produced 139.49 Gb of clean reads and 20,436 differentially expressed genes (DEGs) among different molting stages. The expression patterns for genes involved in several molecular events critical for molting, such as cuticle reconstruction, cytoskeletal structure remodeling, hormone regulation, immune responses, and metabolism, were characterized and considered as mechanisms underlying molting in *S. paramamosain*. Among these genes, we identified 10,695 DEGs in adjacent molting stages. In further GO and KEGG analyses, the significantly enriched pathways included structural constituents of cuticle, binding and chitin metabolic processes, steroid hormone biosynthesis, insulin resistance, and amino sugar metabolic processes. The expression profiles of 12 functional genes detected via RNA-Seq were corroborated via real-time RT-PCR assays. The results indicated gene expression profiles across the molting cycle and identified possible activation pathways for future investigation of the underlying molecular mechanisms.

Conclusions: This study establishes a comprehensive transcriptomic repertoire for the molting process in *S. paramamosain*. The data generated in this study will serve as an important transcriptomic resource for the crab research community to facilitate gene and genome annotation and provide fundamental support for future research on molecular mechanisms in this species. Furthermore, the comprehensive analysis results will improve our understanding of morphological variation in the molting cycle and serve as a potential blueprint for future research on molting in crustaceans and other animals.

1. Introduction

Molting is an important and ongoing process of physiological change in the life history of all crustaceans, with periodic shedding of the old cuticle and subsequent reconstruction of a new rigid exoskeleton [1, 2]. Several factors are common to the molt cycle in all crustaceans, including transcription of cuticular genes, chitin synthesis, exoskeleton reconstruction, and hormone regulation [3, 4]. Previous studies have made significant advances in our understanding of the molting cycle concerning growth characteristics, material changes, energy metabolism, and hormone regulation at the ecological and physiological levels [5–9]. Moreover, recent studies have explored molting-related genes and their effects on protein synthesis [10], chitin metabolism [11], and muscle hydration control [12] during the molting

cycle. However, the molecular events and mechanisms associated with the dynamic process of molting remain poorly understood in crustaceans, especially in crabs.

Scylla paramamosain (Crustacea: Decapoda: Portunidae), commonly known as the mud crab, is widely distributed along the temperate, subtropical, and tropical areas of the Pacific Ocean and Indian Ocean, and this species is cultivated in many countries. *S. paramamosain* undergoes molting up to 18 times during its life cycle, including during larval development, during growth-related molting, and in adult reproductive molting [13]. The economic significance of this species makes the mud crab a valuable system for exploring the genes involved in the molting cycle. Several researchers have described the effects of environmental factors and hormone-related genes on the molting cycle of *S. paramamosain*. Previous study evaluated the effects of temperature, salinity, starvation, and autotomy on the molting of early juvenile mud crabs and found that changes in ecdysone receptor gene expression levels appeared to play an important role in regulating the molting process [14]. Molt-inhibiting hormone (MIH) genes and the crustacean hyperglycemic hormone (CHH) genes that serve as key regulators of hormones controlling the molting process in crustaceans have been cloned and their distribution explored in different tissues during the molting cycle of the mud crab [15, 16]. Moreover, the expression profiles of the chitinase gene family and the $\text{Na}^+/\text{K}^+/\text{2Cl}^-$ cotransporter of mud crabs have been explored during molting [13, 17]. Most recently, a transcriptome study on the hepatopancreas immune response during the molting cycle in *S. paramamosain* has provided a better understanding of the hepatopancreas immune system [17], but our knowledge of the molecular mechanisms remains limited. Therefore, it is important to understand the mechanisms underlying crab molting and to characterize the cascade of sequential expression events of molt-related genes involved in various aspects of the molting process. This will facilitate further investigation of functional genomics for *S. paramamosain* and other closely related species.

The present study attempted to construct a comprehensive transcriptional landscape for further studies of the molecular mechanisms governing the molting cycle process in *S. paramamosain* using RNA-sequencing. The assembled unigenes were annotated against public protein databases followed by NR-NCBI, NT, PFAM, Swiss-Prot, GO, KOG, and KEGG classification to identify DEGs among all molting stages. To investigate expression profiles, we identified genes that were differentially expressed during separate molting stages by sequencing whole juvenile crabs. This resource and the associated findings will provide insights into biological processes of the molting cycle in this species and will help to improve its performance in aquaculture.

2. Materials And Methods

2.1. Experimental design and animal sampling

Healthy juvenile mud crabs from a full-sib family with an average carapace width of 1 cm at growth stage II (Fig. 1) were sampled from our research facility at Ningbo University. These individuals were cultured in tanks with adequate aeration, temperature (25 °C), and food (given twice daily). All procedures

involving animals throughout the experiments were conducted in strict accordance with the the National Institutes of Health guide for the care and use of Laboratory animals.

All crabs were molting staged by examination of pleopod paddles for epidermal retraction, pigmentation, setae development, and the presence of matrix or internal cones in the setal lumen. Molting stages were classified as follows: pre-molt (PrM, Fig. 1A), post-molt (PoM, Fig. 1B), and intermolt (InM-I, InM-II, InM-III, InM-IV, Fig. 1C–F) [3, 4, 18]. According to the descriptions and definitions of molting stages in previous studies [19, 20], PoM should correspond to stage B; InM-I, InM-II, InM-III, and InM-IV correspond to stage C, and PrM corresponds to stage D. We collected three whole crabs from each of these six molting stages and combined them to generate a single sample for sequencing. Three samples were taken as biological replicates for each molting stage. Nine individual crabs were sampled at each molting stage. All samples were immediately frozen and stored in liquid nitrogen until RNA isolation.

2.2. Transcriptome library constructing and Illumina sequencing

The total RNA was isolated using the TRIzol method (Invitrogen, USA) according to the manufacturer's protocol. Total RNA was treated with DNase to remove DNA contamination. The RNA integrity was assessed using the RNA Nano 6000 Assay Kit of the Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA). Only high-quality RNA samples ($OD_{260}/_{280} = 1.8 \sim 2.2$, $OD_{260}/_{230} \geq 2.0$, $RIN \geq 6.5$, $28S:18S \geq 1.0$, $> 10 \mu\text{g}$) were used to construct the sequencing library [21, 22].

A total of $1.5 \mu\text{g}$ RNA per sample was used as input material for the RNA sample preparation. Sequencing libraries were generated using the NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA) following the manufacturer's recommendations, and index codes were added to attribute sequences to each sample. First-strand cDNA was synthesized using random hexamer primers, and M-MuLV Reverse Transcriptase (RNase H-) Second-strand cDNA synthesis was subsequently performed using DNA Polymerase I and RNase H. After adenylation of the 3' ends of DNA fragments, NEBNext Adaptor with a hairpin loop structure was ligated to prepare for hybridization. In order to preferentially select cDNA fragments of 250–300 bp in length, the library fragments were purified with an AMPure XP system (Beckman Coulter, Beverly, USA). Then, $3 \mu\text{l}$ USER Enzyme (NEB, USA) was used with size-selected, adaptor-ligated cDNA at 37°C for 15 min followed by 5 min at 95°C before PCR. The PCR was performed with Phusion High-Fidelity DNA polymerase, Universal PCR primers, and Index (X) Primer. Finally, PCR products were purified (AMPure XP system), and library quality was assessed on the Agilent Bioanalyzer 2100 system [21].

The clustering of the index-coded samples was performed on a cBot Cluster Generation System using a TruSeq PE Cluster Kit v3-cBot-HS (Illumina) according to the manufacturer's instructions. After cluster generation, the library preparations were sequenced on an Illumina HiSeq platform, and paired-end reads were generated [22].

2.3. Assembly of sequencing data, gene annotation, and classification

The raw paired end reads were trimmed and subjected to quality control using SeqPrep (<https://github.com/jstjohn/SeqPrep>) and Sickle (<https://github.com/najoshi/sickle>) with default parameters. Then, clean data from the three samples were used to perform RNA de novo assembly using Trinity (<http://trinityrnaseq.sourceforge.net/>) [23]. All the assembled transcripts were searched against the NCBI (<http://www.ncbi.nlm.nih.gov/>) protein NR and non-redundant nucleotide (Nt), Swissprot (<http://www.ebi.ac.uk/uniprot/>), the KOG database (<http://www.ncbi.nlm.nih.gov/COG/>), and the KEGG (<http://www.genome.jp/kegg/>) database and KO (KEGG Ortholog database, <http://www.geneontology.org>). The remaining unmatched unigenes were further analyzed using ESTscan, and the predicted CDSs were translated into peptide sequences. GATK3 software was used to perform SNP calling. SSRs of the transcriptome were identified using MISA (<http://pgrc.ipk-gatersleben.de/misa/misa.html>), and the primers for each SSR were designed using Primer3 (<http://primer3.sourceforge.net/releases.php>).

2.4. Differential gene identification, enrichment, and pathway analysis

The gene expression levels were calculated by using FPKM (expected number of Fragments Per Kilobase of transcript sequence per Million base pairs sequences) [24]. Differential expression analysis of two samples was performed using the DESeq (2010) R package. DESeq provides statistical routines for determining differential expression in digital gene expression data. Genes with an adjusted P-value < 0.05 found by DESeq were assigned as differentially expressed using a model based on the negative binomial distribution. The P-value was adjusted using the q value [25]. The q value < 0.005 & |log₂ (foldchange)| > 1 were set as the threshold criteria for significantly differential expression between adjacent samples (InM-I vs PoM, InM-II vs InM-I, InM-III vs InM-II, InM-IV vs InM-III, PrM vs InM-IV, PoM vs PrM). Gene Ontology (GO) enrichment analysis of the differentially expressed genes (DEGs) was implemented by the GO seq R packages-based Wallenius non-central hyper-geometric distribution that can adjust for gene length bias in DEGs [26]. We used KOBAS software [27] to test the statistical significance of enrichment of differentially expressed genes in KEGG pathways.

2.5. Quantitative reverse-transcription PCR (qRT-PCR) analysis

To validate the results from RNA-Seq differential gene expression analysis, qRT-PCR was carried out among all six molting stages. Twelve genes from different gene expression clusters were randomly selected for qRT-PCR assays. PCR primers were designed based on the assembled transcriptome sequences (Table S2). The β -actin gene was amplified in parallel as an internal control.

Total RNA was extracted, and the first-strand cDNA was synthesized by using a PrimeScript RT Reagent Kit with gDNA Eraser (Takara). All samples were run in triplicate in separate tubes; each cDNA sample

was run in duplicate. The qRT-PCR was conducted by using SYBR Green Premix Ex Taq (Takara, Japan) in an ABI 7500 Sequence Detection System (Applied Biosystems). A standard curve was initially generated to assess accuracy, and primers with efficiency of amplification between 95% and 105% were chosen for the following qRT-PCR. The comparative Ct method (with the formula $2^{-\Delta\Delta Ct}$) was used to analyze the expression levels of different genes [28]. Student's t-test was conducted using SPSS 20.0 (<http://www-01.ibm.com/software/analytics/spss/>), and P-values less than 0.05 were considered statistically significant.

3. Results

3.1. Illumina sequencing and transcriptome assembly

The assembly generated 953,784,150 raw reads from 18 samples taken at the six molting stages. A draft assembly of the *S. paramamosain* transcriptome was constructed with 139.49 Gb of quality filtered short-read Illumina sequence data (Table S1). All transcriptome data are available in the NCBI Short Read Archive (SRA) database under accession PRJNA687923. Finally, we obtained an average of 51,664,527 clean reads for our 18 samples, with Q20 of all samples greater than 97% (Table S1). The numbers of unigenes with FPKM > 0.3 ranged from 70,319 to 78,247 among the eight molting stages (Fig. 2A). The number of reads mapped to the stage-specific transcriptome was more than 100 million clean reads (Fig. 2B). More than 70% of the reads were mapped to the filtered reference transcriptome and the stage-specific corresponding transcriptome (Fig. 2B, Table S1).

In our assembly, 350,531 transcripts (each with more than 200 bp) with a total length of 339,582,267 bp were obtained, and these were assembled into 248,744 unigenes totaling 312,762,368 bp. A large portion of the transcripts were between 200 bp and 500 bp (Table 1). The transcripts had an N50 size of 2,224 bp, mean transcript length of 969 bp, maximum transcript length of 33,949 bp, and mean unigene length of 1,257 bp (Table 1).

Table 1
Summary statistics of *S. paramamosain*
transcriptome.

Length	Transcripts	Unigenes
200–500 bp	208,438	107,506
500-1kbp	58,411	57,615
1 k-2kbp	38,919	38,866
> 2kbp	44,763	44,757
Total number	350,531	248,744
Mean Length	969	1,257
Max Length	33,949	33,949
N50	2,224	320
N90	2,505	455
Total Nucleotides	339,582,267	312,762,368

3.2. Transcriptome annotation

We used the databases NR, NT, Swiss-Prot, PFAM, KOG, GO, and KEGG for unigene annotation. A total of 248,744 genes were successfully annotated by aligning to one or more of the seven databases (Table S2). Of these 248,744 unigenes, 78,617 (31.6%) matched to known proteins in the NR database (Table S3), and 32,401 (13.02%) matched to putative homologues in the NT database. Among all annotated databases, the largest match of annotation hits (39.01%) was obtained with the GO database, followed by the PFAM database (38.74%, Table S2, Fig. 2C). In addition, the KOG database provided annotation for 38,663 (15.54%) unigenes; the Swissprot database confirmed matches for 63,177 (25.39%) unigenes, and 38,478 (15.46%) unigenes found putative homologues in the KEGG database (Table S2, Fig. 2C). Of the unigenes, 120,837 (48.57%) exhibited a positive match against at least one of the seven databases, whereas 13,062 (5.25%) had the best BLAST matches to proteins in all four databases (Table 2).

Table 2
Expression levels of carapace, muscle, and hemolymph-related molting genes.

Factor name	Gene Number	Average FPKM value					
		PoM	InM- \bar{x}	InM- \bar{x}	InM- \bar{x}	InM- \bar{x}	PrM
Carapace							
Arthrodial cuticle protein	23	180.01	295.92	318.88	305.36	105.33	210.27
BCS-1	30	2.69	1.89	0.77	0.30	0.11	152.75
Calcified cuticle protein	38	964.13	1062.33	1539.60	1647.34	337.25	41.86
Chitin binding peritrophin-A	182.72	174.85	138.95	54.01	25.03	57.85	
Chitin binding protein-1	9	16.24	7.83	1.58	1.99	0.09	96.17
Chitin deacetylase 1 precursor	7	219.78	190.60	77.53	18.70	10.43	49.65
Chitin deacetylase 4	4	94.38	31.87	43.42	1.78	2.61	10.96
Chitin deacetylase 9 precursor	0.33	0.16	0.22	0.28	0.51	1.40	
Chitin synthase	65.01	44.06	36.94	5.80	6.30	8.62	
Chitinase	5	189.24	163.82	195.94	216.37	34.51	21.20
Chitinase 1	7	49.65	83.76	227.83	297.75	103.36	27.40
Chitinase 2	2	0.28	0.41	0.43	0.38	36.71	0.53
Chitinase 3	2	38.69	40.69	86.10	122.44	45.79	271.32
Chitinase 4	7	50.12	44.74	39.82	16.77	4.69	10.41
Chitinase 5	1	0.50	0.06	0.47	0.05	0.17	0.01
Cuticle proprotein proCP	49	4941.49	6435.35	7016.03	5764.07	1163.86	13.67
Cuticle protein	15	3.63	3.13	3.73	3.05	3.30	36.34
Cuticle protein 19	5	11.11	5.32	0.82	0.48	0.00	94.38

Factor name	Gene Number	Average FPKM value					
		PoM	InM- \bar{x}	InM- \bar{x}	InM- \bar{x}	InM- \bar{x}	PrM
Cuticle protein 19.8	5	1.47	1.96	0.91	0.15	0.07	27.11
Cuticle protein 6	2	4.72	5.81	8.66	0.84	2.28	17.97
Cuticle protein 7	11	7.53	4.75	2.03	0.70	1.46	117.80
Cuticle protein 8	2	16.01	18.83	12.11	2.90	0.46	2354.32
Cuticle protein AM1239	3	62.68	91.21	76.12	89.91	24.30	3.01
Cuticle protein AMP	6	366.01	610.65	991.85	1369.20	702.14	392.23
Cuticle protein BD1	2	3683.55	3366.82	1882.50	785.89	202.75	5.80
Cuticle protein CB5	3	40.18	35.96	33.41	21.54	0.11	55.53
Cuticle protein CB6	46	6.31	4.98	0.43	0.22	0.05	443.36
Cuticle protein CUT2	3	11523.74	12540.51	22623.31	12243.07	1145.52	0.00
Cuticle protein CUT5	1	1.23	1.86	0.55	0.64	145.83	0.46
Cuticlin 1	2	6.39	7.06	11.38	0.67	2.13	2.08
Cuticular protein	6	11.54	15.58	33.97	28.00	4.63	62.88
Cuticular protein 15	2	539.13	2145.49	4054.25	3704.62	370.98	0.88
Early cuticle protein 2	11	2.69	0.77	0.60	0.45	0.00	292.64
Early cuticle protein 3	16	5.71	3.50	1.53	1.62	0.19	829.86
Early cuticle protein 5	5	0.09	0.02	0.02	0.01	0.00	238.07
Early cuticle protein 6	13	1.87	1.57	1.27	0.86	0.14	162.58

Factor name	Gene Number	Average FPKM value					
		PoM	InM- \bar{x}	InM- \bar{x}	InM- \bar{x}	InM- \bar{x}	PrM
Gastrolith protein	2	3481.04	3394.12	1242.40	373.16	141.77	1159.36
Gastrolith protein 10	3	2770.05	1522.79	606.51	274.01	180.67	6.80
Gastrolith protein 30	1	4773.74	4287.93	3622.48	503.81	11.52	1294.13
Muscle and skelemin							
Alpha tubulin	21	11.54	2.96	6.82	6.24	4.91	6.49
Beta tubulin	17	13.16	6.72	6.76	2.80	6.54	4.22
Titin	10	4.87	3.16	4.77	1.35	5.05	1.73
aquaporin-3	6	19.64	12.81	9.46	3.71	7.74	7.69
aquaporin-12	11	1.35	0.64	2.12	0.72	1.26	0.50
aquaporin	9	64.03	85.23	41.95	18.49	33.73	15.49
Actin	26	5.52	6.13	5.81	5.11	8.63	19.87
Actin, cytoplasmic	1	33.05	37.55	38.11	17.66	41.05	62.70
Actin, muscle-like isoform	2	487.66	580.10	467.80	715.96	364.22	124.11
Actin-binding protein anillin	2	0.00	0.15	0.04	0.07	0.72	0.26
Actin-related protein 5	2	0.17	0.44	0.12	0.24	0.22	0.20
Beta-actin	17	114.94	139.16	115.52	100.68	109.08	141.14
Cardiac muscle actin	1	4.80	4.61	7.28	2.84	1.29	0.07
Muscle LIM protein	1	24.75	19.45	17.12	8.05	23.75	13.68
Muscle and skeletal receptor tyrosine protein kinase	1	3.82	3.02	3.43	0.69	1.11	0.65
troponin C	3	17.18	18.38	7.64	10.12	33.65	40.37

Factor name	Gene Number	Average FPKM value					
		PoM	InM- \bar{x}	InM- \bar{x}	InM- \bar{x}	InM- \bar{x}	PrM
tropomyosin	14	0.29	0.14	0.20	0.12	0.13	0.24
Muscle-specific calpain	3	1.51	2.96	2.79	7.02	6.47	4.66
Myostatin-like protein	4	11.79	4.06	4.33	0.77	0.58	0.39
Myosin	13	2.76	2.31	2.00	0.85	0.96	1.50
Myosin H	2	8.89	8.47	10.94	8.49	22.94	36.62
Myosin heavy chain	26	167.19	106.64	83.48	41.63	60.38	52.32
Myosin light chain	1	682.33	512.57	391.75	195.35	168.86	190.14
Myosin light chain kinase	2	18.51	13.86	4.81	3.29	4.84	3.13
Hemolymph							
Cryptocyanin 1	1	0.45	2.63	0.53	0.27	1070.55	29.25
Hemocyte kazal-type proteinase inhibitor	1	94.07	38.16	27.73	4.01	29.41	5.37
Hemocyte homeostasis associated protein	1	4.62	2.33	4.59	1.12	1.40	1.00
Hemocyanin subunit	10	63.16	67.50	201.55	132.20	1105.81	264.94
Hemocyanin	3	63.81	53.01	514.09	314.22	2377.74	579.84
Spz3	2	29.84	31.30	34.99	10.30	0.46	32.27

The similarity analysis between *S. paramamosain* unigenes and NR protein databases was conducted using BLAST matches with a cut-off E-value of 1.0E-5 (Fig. S1A). For the main species distribution matched against the NR database, only 10.2% of the matched unigenes showed similarities with *Zootermopsis nevadensis*, followed by *Daphnia pulex* (5.4%), *Tribolium castaneum* (3.2%), *Stegodyphus mimosarum* (2.5%), *Crassostrea gigas* (2.2%), and others (76.4%) (Fig. S1B). There were 1.5%, 34.3%, 44.2%, 16.3%, and 3.6% putative proteins showing 18–40%, 40–60%, 60–68%, and 80–100% similarity, respectively, with the known proteins in the NR protein database (Fig. S1C).

In our study, a total of 97,051 unigenes were assigned to 56 sub-categories of GO terms belonging to the following three main categories: cellular component, molecular function, and biological process. These main categories included 20, 10, and 26 subcategories, respectively (Fig. S2A). For cellular components, cell (38,969 genes), cell part (38,968 genes), organelle (28,306 genes), membrane (25,539 genes), and macromolecular complex (24,009 genes) represented most of the categories. Binding (48,876 genes), catalytic activity (35,003 genes), transporter activity (11,160 genes), structural molecule activity (5,699 genes), and molecular transducer activity (4,038 genes) represented the activity categories of molecular function. Cellular process (59,345 genes), metabolic process (48,887 genes), single-organism process (43,713 genes), biological regulation (25,401 genes), and localization (23,965 genes) represented a high percentage of the biological process category. The results of KEGG pathway analysis showed that 38,478 unigenes were annotated into five main categories and 32 sub-categories (Fig. S2B). Among the main categories, organismal systems had the largest number of unigenes (13,843 unigenes), followed by metabolism (7,698 unigenes), genetic information processing (8,824 unigenes), cellular processes (7,956 unigenes), and environmental information processing (6,318 unigenes). Among the 32 sub-categories, translation was the maximum group with 5,668 unigenes, followed by signal transduction (5,365 unigenes). The smallest group containing only 53 unigenes was biosynthesis of other secondary metabolites. As a consequence, 38,663 annotated genes in KOG were grouped into 26 KOG categories, among which the cluster for “translation, ribosomal structure and biogenesis” (5,704 unigenes) represented the largest group, followed by “signal transduction mechanisms” (5,570 unigenes), “general function prediction only” (5,297 unigenes), “post translational modification, protein turnover, chaperones” (4,161 unigenes) and “cytoskeleton” (2,513 unigenes); “nuclear structure” (144 unigenes) and “cell motility” (50 unigenes) represented the smallest groups (Fig. S2C).

3.3. Expression profile of DEGs across the molting cycle of *S. paramamosain*

A total of 20,436 DEGs were identified across the six molting stages (Table S4). Expression levels of DEGs during the molting cycle were divided into 22 categories based on K-means clustering that revealed different expression patterns of DEGs among the six molting stages (Fig. S3). The largest group, cluster 3, contained 3,028 DEGs with gene expression levels decreasing during the InM-I vs PoM and InM-III vs InM-II transitions. Other clusters containing over 1,000 members were cluster 10 (1,458 DEGs), cluster 11 (1,474 DEGs), cluster 12 (1,038 DEGs), cluster 16 (1,091 DEGs), cluster 19 (2,739 DEGs), and cluster 20 (1,528 DEGs) (Fig. S3).

GO enrichment analysis (FDR < 0.005) of these DEGs demonstrated enrichment of multiple terms of biological processes, cellular components, and molecular functions associated with the molting cycle (e.g., structural constituents of cuticle, amino sugar metabolic processes, chitin metabolic processes, and hormone receptor binding) (Fig. 3). Profiling gene expression and GO enrichment analysis of DEGs found that genes in clusters 3, 7, 8, 10, 11, 18, and 20 were predominantly up-regulated in the molting stages InM-I, InM-II, and PoM and were down-regulated in InM-III, InM-IV, and PrM, stages that were mainly associated with “myosin complex,” “motor activity,” and “actin cytoskeleton”; and “structural molecule

activity,” “structural constituent of cuticle,” and “chitin binding”; and “amino sugar metabolic process.” Conversely, DEGs in clusters 12, 16, 17, and 19 were predominantly down-regulated in the molting stages InM-I, InM-II, and PoM and were up-regulated in InM-III, InM-IV, and PrM, stages that were mainly associated with “glucosamine-containing compound metabolic process,” “amino sugar metabolic process,” “GTP binding,” “oxidoreductase activity,” “guanyl ribonucleotide binding,” “3-beta-hydroxy-delta5-steroid dehydrogenase activity,” “C21-steroid hormone metabolic process,” and “regulation of hormone levels.” Interestingly, DEGs in clusters 2, 4, and 13 were only up-regulated in stage PrM, and these were strongly enriched in “structural constituent of cuticle,” “structural molecule activity,” “chitin metabolic process,” “glucosamine-containing compound metabolic process,” and “amino sugar metabolic process.” Moreover, DEGs in clusters 1 and 16 showed that the expression of genes increased with the number of days following molting. Details of the top 30 GO categories of the 22 clusters are shown in Table S5.

3.4. Expression profiles of DEGs associated with growth, immunity, and metabolism among molting stages

The molting of crustaceans is accompanied by structural changes and material metabolism in various tissues of the body, among which muscles and the carapace undergo the most dramatic changes during the molting cycle. In this study, dozens of DEGs associated with carapace, muscle, and hemolymph were identified (Table 2). Of these DEGs, we identified many transcripts related to the composition and organization of the shell in crustaceans, including arthroal cuticle protein, BCS, calcified cuticle protein, chitin-binding protein, chitin deacetylase, chitinase, cuticle protein, cuticular protein, early cuticle protein, and gastrolith protein. Members of different gene families had their own characteristics during the molting cycle. For example, chitinase 1 and 3 showed up-regulated expression trends from PoM to InM-III, but both decreased at stage InM-IV. Moreover, all the four early cuticle protein family members had down-regulated expression patterns from PoM to InM-IV and then increased suddenly at stage PrM. The muscle-and skeleton-related genes included many aspects of muscle formation and function from structural protein to muscle contraction such as actin, muscle LIM protein, muscle and skeletal receptor tyrosine protein kinase, cardiac muscle actin, and myosin. In addition, six other genes were found to be associated with hemocyte homeostasis and composition, including cryptocyanin, hemocyte kazal-type proteinase inhibitor, hemocyte homeostasis associated protein, hemocyanin, and spz3 (Table 2). Interestingly, some DEGs showed large changes in expression. For example, the FPKM values of cuticle protein CUT2 increased more than 10,000-fold from InM-I to InM-II, but decreased from InM-III, reaching zero at stage PrM. Additionally, the FPKM values of cuticle proprotein proCP and cuticular protein 15 decreased more than 3,000-fold from InM-III to InM-IV, and the same pattern was found for gastrolith protein 30, with a more than 3,000-fold decrease from InM-II to InM-III.

In our study, 21 important hormone-related genes that may be directly or indirectly involved in controlling the crustacean molting cycle were identified (Table 2). Of these genes, the expression levels of ecdysis triggering hormone receptor subtype-B and ecdysone induced protein showed characteristic patterns of down-regulation during the entire molting cycle from PoM to PrM. The expression of ecdysteroid-regulated protein was up-regulated from PoM to InM-II and then declined from InM-III to PrM. Estrogen

sulfotransferase was up-regulated from PoM to InM-IV but increased suddenly at stage PrM, and the same pattern was found for estradiol 17-beta-dehydrogenase 8 and juvenile hormone epoxide hydrolase (Table 3).

Table 3
Expression levels of molting regulation-related hormone genes.

Factor name	Gene Number	Average FPKM value					
		PoM	InM- \square	InM- \square	InM- \square	InM- \square	PrM
Androgen-induced protein 1	1	2.41	1.57	3.21	7.69	8.67	4.9
Crustacean hyperglycemic hormone	1	0.94	0.19	0.55	0.28	4.22	5.18
Ecdysis triggering hormone receptor subtype-B	2	13.42	8.05	7.63	4.43	4.14	3.14
Ecdysone receptor	4	1.91	2.37	1.75	2.22	1.38	1.34
Ecdysone induced protein	4	4.01	0.67	0.56	0.14	0.1	0.09
Ecdysteroid regulated protein	4	208.13	368.08	471.96	448.12	56.74	28.86
Estradiol 17-beta-dehydrogenase 8	4	1.75	4.43	14.39	18.52	109.8	14.37
Estradiol receptor-like protein 1	1	1.01	0.04	0.13	0	0	0.04
Estrogen sulfotransferase	7	4.14	5.74	13.03	20.13	35.54	6.71
Farnesoic acid O-methyltransferase long isoform	1	9.29	12.56	12.53	75.18	51.38	62.54
Female sterile homeotic protein	2	8.17	4.4	7.6	2.91	5.51	2.41
Hormone receptor 4	1	0.42	0.49	0.34	0.06	0.03	0.21
Hormone receptor 4 isoform X	4	14.79	9.85	7.6	2.2	0.65	9.28
Hormone receptor 78	2	6.87	6.52	7.5	3.53	1.24	0.63
Hormone receptor hr3	3	0.27	0.12	0.17	0.18	0.21	21.81
Juvenile hormone acid methyltransferase	2	2.88	0.63	2	0.14	1.13	0.81
Juvenile hormone epoxide hydrolase	1	13.26	25.85	36.46	52.27	153.99	33.96
Lutropin-choriogonadotropic hormone receptor	2	0.77	0.12	0.66	0	0.54	0.08

Factor name	Gene Number	Average FPKM value					
		PoM	InM-I	InM-II	InM-III	InM-IV	PrM
Molting fluid carboxypeptidase A precursor	2	0.24	0.87	0.06	0.05	1.05	8.1
Vitelline membrane outer layer 1-like protein	2	80.75	99.5	93.34	153.24	139.87	156.19
Vitellogenin	1	0.15	0.08	3.38	0.03	0.02	7.11

To identify the possible factors involved in the molting process, we also investigated the expression profiles of members of the immune system and metabolism during molting. There were 21 and 28 genes found to be related to the immune system and metabolism, respectively. Of the immune-related genes, the C-type lectin receptor was down-regulated in PoM but up-regulated in PrM, and the same pattern was found for crustin 2, crustin 3, beta-thymosin 4, and beta-thymosin 5. Interestingly, most immune system genes reached their maximum expression levels in InM-II; these included C-type lectin domain family 4-member, C-type lectin-2, C-type lectin receptor, C-type lectin, crustin-2, crustin 4, and ALF1. Twenty-eight genes related to substance metabolism, storage, and transportation were detected, including four sugar transporters, one sterol regulatory element-binding protein, six serine proteinases, three serine proteinase inhibitor accases, six cytochrome P450s, three lipoprotein receptors, one lipid storage droplet protein, and one lipase (Table S6).

3.5. GO and KEGG enrichment analysis of DEGs in adjacent molting stages

Our analysis specifically investigated the interactions among DEGs between adjacent molting stages and their functioning in biological processes by GO and KEGG enrichment. A total of 10,695 DEGs between any two adjacent molting stages (i.e., InM-I vs PoM, InM-II vs InM-I, InM-III vs InM-II, InM-IV vs InM-III, PrM vs InM-IV, and PoM vs PrM) were identified (Table S7, Fig. 2D). The 10,695 DEGs were further subjected to GO enrichment to predict their potential functions and to KEGG enrichment to obtain relevant metabolic pathways. To associate different expression patterns of these DEGs with morphological and physiological changes during the molting cycle, we performed GO enrichment analysis and sorted DEGs into different sub-categories that belonged to three main GO categories: biological processes, cellular components, and molecular functions. GO enrichment analysis revealed that structural constituents of the cuticle were the most significantly enriched pathway for 10,695 DEGs accounting for 451 unigenes, followed by chitin-binding and chitin metabolic process (Table S8, Fig. S4A). Notably, these top three enrichment pathways were all related to the molecular function and biological process of shell formation regulation during crab molting.

The 20 most enriched KEGG pathways are listed in Table S9 and Figure S4B. Among these, the ten most important and representative pathways that are potentially relevant to gene interactions underlying the molting cycle included 616 unigenes that were matched to “steroid hormone biosynthesis” (63 unigenes),

“insulin resistance” (102 unigenes), “linoleic acid metabolism” (33 unigenes), “drug metabolism-other enzymes” (60 unigenes), “retinol metabolism” (55 unigenes), “hedgehog signaling pathway” (60 unigenes), “adipocytokine signaling pathway” (67 unigenes), “protein digestion and absorption” (77 unigenes), “bile secretion” (69) and “hematopoietic cell lineage” (30) (Table S9). In addition, the pathway of “steroid hormone biosynthesis” was the most representative pathway, with high hormone-related gene numbers in accordance with the characteristics of molting control; this is likely to be useful in future investigations focusing on their functions in the molting cycle of *S. paramamosain* (Fig. S4C).

In the comparison of InM-I vs PoM, 597 DEGs were up-regulated, and 711 DEGs were down-regulated. These were sorted into 2,585 GO sub-categories with six significantly enriched sub-categories (Corrected P-value < 0.05) of amino sugar metabolic processes, structural constituents of cuticle, chitin metabolic processes, glucosamine-containing compound metabolic processes, chitin binding, and viral release from host cell (Fig. 2D, Fig. S5). In the comparison of InM-II vs InM-I, 338 DEGs were up-regulated, and 410 DEGs were down-regulated; these were sorted into 2,152 GO sub-categories with four significantly enriched sub-categories (Corrected P-value < 0.05) of structural constituents of cuticle, histidine biosynthetic processes, histidine metabolic processes, and imidazole-containing compound metabolic processes (Fig. 2D, Fig. S5). A total of 322 DEGs were up-regulated and 627 DEGs were down-regulated in the comparison of InM-III vs InM-II, and no significantly enriched sub-categories (Corrected P-value < 0.05) were observed. In the comparison of InM-IV vs InM-III, 789 DEGs were up-regulated, and 780 DEGs were down-regulated; these were sorted into 2,506 GO sub-categories with 13 significantly enriched sub-categories (Corrected P-value < 0.05) (Fig. 2D, Fig. S5). There were 1947, 2958 up-regulated DEGs and 681, 3197 down-regulated DEGs for the comparisons of PrM vs InM-IV and PoM vs PrM, respectively. The major significantly enriched sub-categories (Corrected P-value < 0.05) for the comparisons of PrM vs InM-IV and PoM vs PrM are shown in Figure S5.

To more precisely analyze the expression patterns of DEGs in the molting cycle, we examined the expression profiles (Table S10) and listed the top 10 GO enrichment categories of DEGs in pairwise comparisons among molting stages of PrM, PoM, and InM (Table S11). Interestingly, Venn diagram analysis revealed a total of 73 and 1418 DEGs for pairwise comparisons between the four intermolt stages (InM-I, InM-II, InM-III, InM-IV) and PrM, PoM, respectively (Fig. 4). In InM vs PoM, 73 DEGs were sorted into the most enriched GO sub-categories associated with energy metabolism (e.g., adenosine kinase activity, AMP biosynthetic process, AMP metabolic process, and riboflavin synthase complex). In InM vs PrM, 1418 DEGs were sorted into the most enriched GO sub-categories associated with shell formation in the molting cycle (e.g., structural constituent of cuticle, structural molecule activity, chitin binding, and chitin metabolic process).

3.6. qRT-PCR validation

For validation purposes, we selected 12 representative genes from the 22 clusters identified above and examined their expression levels during the molting cycle by using qRT-PCR. The qRT-PCR results confirmed the findings of RNA-Seq, with correlation coefficient (*r*) values from 0.844 to 0.979 depending on the specific genes and reference genes (Fig. 5).

4. Discussion

Establishment of a comprehensive transcriptomic repertoire for the molting process in *S. paramamosain*

In order to produce a comprehensive reference transcriptome, next-generation sequencing was used to construct a genome-wide transcriptional landscape for further studies of the developmental and molecular aspects of the molting process in *S. paramamosain*. An *S. paramamosain* molt cycle-specific transcriptome derived from whole juvenile crabs, including individual organs such as the brain, gills, nerves, eyestalks, hepatopancreas, muscle, MO, and Y organs from all molt cycle stages, was developed for this study. There were some similarities to previous studies [3,4] involving molting cycle transcriptome analyses in crustaceans compared to the results of our work; for example, some common factors related to the molting cycle were identified, such as the cuticular genes, chitin synthase, exoskeleton reconstruction, and hormone regulation. However, several differences exist between our results and those of previous studies. These differences may be due to the materials and techniques used. The transcriptome database obtained using Illumina RNA-seq in our study was larger than those obtained by microarray-based analysis in other studies, and more genes were detected at each stage of the molting process in our study. Second, previous studies mainly focused on the functional analysis of single organs during the molting cycle or on a specific stage, for example, focusing on muscle [29], the hepatopancreas [4,17], the eyestalks and heart [30], or the mandibular organ and Y-organ [31]. Moreover, our analysis revealed a cascade of sequential expression events of genes involved in various aspects of the molting process, including chitin degradation and synthesis, tissue development, hormone regulation, immune response, and energy metabolism. The present transcriptomic data are of high quality compared with previous studies and provide a comprehensive survey from each molting cycle stage of *S. paramamosain*.

Gene Ontology functional annotation analysis classified the predicted functions of the assembled unigenes. Cell, binding, and cellular processes were the most highly represented categories, consistent with transcriptome results of the whole claw muscles in *Eriocheir sinensis* [29] and *Penaeus monodon* [30]. Our results further demonstrated that proliferation of epithelial tissues [32,33], muscle [34], and other development-related processes accompanying muscle and exoskeleton reconstruction were prominent during molting. Moreover, KEGG analysis demonstrated that the top enriched pathways were organismal systems instead of being related to metabolism as in previous studies [29,30]. These pathways included the circulatory system, environmental adaptation, the digestive system, endocrine system, the immune system, and the nervous system. Also, translation, ribosomal structure, and biogenesis were identified as the most present categories. Therefore, the present results suggest the biological functions and interactions of annotated unigenes involved in organ development, tissue reconstruction, and adjustment of environmental pressure during the molting cycle. These annotations are the starting point for investigating the molecular mechanisms of molting-related genes in the molting cycle of *S. paramamosain*, and they also provide a valuable resource for further research into specific functions and pathways of crustacean molting.

The molting cycle is a complex process that is normally divided into four stages: intermolt, premolt, ecdysis, and postmolt [35,36]. Overall, our clustering analysis and GO enrichment of DEGs across the molting cycle of *S. paramamosain* revealed the variational characteristics of molecular changes responding to morphological and physiological changes. For example, previous studies have indicated that postmolt (PoM) is a critical phase in the molting cycle of crustaceans, a period in which the animal is recovering from molting and during which its exoskeleton is quickly hardened to avoid predation [4,29]. The most enriched GO terms of DEGs in PoM, InM-I, and InM-II of *S. paramamosain* included “myosin complex,” “motor activity,” “actin cytoskeleton,” and “structural molecule activity,” revealing a gradual up-regulation of muscle-related genes and further indicating the important functions of these genes in muscle regeneration during the postmolt and intermolt periods [6,31,37]. Moreover, up-regulation of genes in the premolt stage being enriched in the GO terms “3-beta-hydroxy-delta5-steroid dehydrogenase activity,” “C21-steroid hormone metabolic process,” and “regulation of hormone levels” demonstrated that the premolt stage, a critical phase in preparation for ecdysis, is controlled by hormonal regulation in the molting cycle of *S. paramamosain*.

Expression patterns of DEGs associated with cuticle and cytoskeletal proteins among molting stages

Similar to other crustaceans, growth of *S. paramamosain* is a stepwise process comprising periodic shedding, subsequent reconstruction of a rigid external exoskeleton and cuticle, and muscle growth to fill the new body during a molting cycle [3,38]. Some functional genes involved in the molting cycle of crabs were identified as associated with a series of biological processes in previous studies, for example, in cuticle reconstruction, cytoskeletal structure remodeling, protein synthesis, hormone regulation, immune response, and metabolism [4,13,15,17]. Among these processes, cuticle reconstruction and cytoskeletal structure remodeling have been considered to be essential for all four phases of the molting cycle [33]. More than 50 different cuticle-and cytoskeletal-related genes were identified as being differentially expressed across the molting cycle in this study. For example, chitinase digests the old exoskeleton to regenerate a new shell via chitin metabolism that involves chitin synthase genes, chitinase genes, chitin deacetylase genes, and a number of genes whose gene products contain chitin-binding or other chitin metabolism-related domains [11,39,40]. Research studies on *Penaeus aztecus* [41] and *P. monodon* [42] have identified chitinase 2 as a direct factor in the molting process that is up-regulated in stage PrM and down-regulated in PoM. Moreover, previous study in juvenile *Penaeus chinensis* [43] has shown that the expression changes of chitinase 1 and 3 can be observed at all stages of the molting cycle. Finally, we identified six chitinases and found that they showed differential expression patterns in stages PoM, InM, and PrM, suggesting different physiological roles and modes of action of during the molting cycle. The chitin deacetylase gene is another chitin metabolism-related gene involved in the catalysis of the acetamido group in the N-acetyl-d-glucosamine units of chitin [44,45]. To date, five classes of chitin deacetylase have been characterized, three of which were identified in this study (chitin deacetylase 1, chitin deacetylase 4, and chitin deacetylase 9). In insects, it has been demonstrated that the chitin deacetylase gene is involved not only in growth but also in the immune system [45,46]. However, few studies have characterized the chitin deacetylase genes or studied their function in the molting cycle of crustaceans. Most recently, the first study concerning the cDNA cloning of chitin deacetylase in a crab

species has been carried out in *Chionoecetes japonicus* [47]. The authors of the latter study hypothesized that chitin deacetylase was involved in ecdysis, as it was expressed only in the epidermis. Moreover, transcriptome analysis of molting-related tissues in *Cherax quadricarinatus* [11] and *E. sinensis* [40] revealed the up-regulated expression of chitin deacetylase during PoM compared to its levels during InM. In our study, chitin deacetylase 1 and chitin deacetylase 4 had similar expression patterns to those found in previous studies, reaching the highest expression peaks at stage PoM, decreasing in the four periods of InM, and increasing again in PrM. In addition, we also observed similar transcriptional up-regulation of chitin synthase and chitin-binding protein genes, i.e., reaching the highest expression peaks at stage PoM, decreasing in the four periods of InM, and increasing again in PrM. The functions of these chitin metabolism-related genes in the molting cycle of *S. paramamosain* need further exploration in the future.

Dozens of cuticle protein-related genes have also been isolated, including calcified cuticle protein, cuticle protein AMP, cuticle protein CUT, and cuticular protein, some of which have extremely high expression levels and similar expression patterns as in chitin genes during the molting cycle. These include cuticle protein CUT2, cuticle protein BD1, cuticle proprotein proCP, and cuticular protein 15. The expression levels of these genes have been shown to be associated with cuticle calcified by the role of inhibition or promotion in the molting process [48-50]. Moreover, gastrolith protein and gastrolith protein 30 showed expression profiles of up-regulation in the PoM stage and then decline during the four InM stages followed by a sharp increase in the PrM stage. Gastrolith protein 10 clearly declined from PoM to PrM. Our results suggest that the molting cycle process is complex, involving up-or down-regulation of these cuticle transcripts, indicating that formation and/or repair of the exoskeleton may need them to operate separately. The new description of DEGs and the determination of their timing in different molting stages provide temporal markers for future studies of molting progress and regulation.

Cytoskeletal reconstruction is an important process for body recovery after molting [33]. However, factors participating in enlargement of the cytoskeleton during the molting process are not yet completely clear. At the level of gene expression, we found a number of genes involved in osmotic regulation, muscle growth, and cytoskeletal structure remodeling that were up-or down-regulated at different molting stages. For example, three aquaporin genes were identified with lower expression levels during the PrM stage but higher levels during PoM and early InM, possibly due to rapid water uptake that occurs after ecdysis. Previous gene expression studies in *Palaemon argentinus* [12] and *S. paramamosain* [2] have explored the role of aquaporin with Na⁺/K⁺-ATPase and Na⁺/K⁺/2Cl⁻ cotransporter in the uptake of water during the molting processes; these are osmoregulation-related genes with the same trends of expression levels in each of the molting stages in our study. We inferred that aquaporin and other osmoregulation-related genes act synergistically to regulate osmotic pressure and water absorption after molting. Cytoskeletal reconstruction involves not only fast water absorption but also expansion of the cytoskeleton to provide a scaffold and muscle to fill the new body [38]. We observed up-regulation of many cytoskeletal genes during early PrM and PoM that have been shown to be related to constituents of the microtubules, growth, and movement of muscle, including titin, tubulin, tropomyosin, and myosin. Among these cytoskeletal-related genes, titin binds to filamentous actin and provides elasticity to muscles in insects [51], and tubulin is another important cytoskeletal group of proteins, the major constituents of the

microtubules [52]. Actin is a globular multi-functional protein that forms microfilaments and affects muscle plasticity in crustaceans [53,54]. As the major contractile protein in vertebrates and invertebrates, myosin was identified to be associated with muscle atrophy during molting and is characterized by decreases in fiber width and myofibril cross-sectional areas, an increase in interfibrillar spaces, and degradation [34]. Moreover, hemocyanin is thought to play a role in cuticle formation and ecdysone transport during molting regulation and antigen non-specific immune defense by reversibly binding oxygen [55] and displaying PO, ecdysone binding, and transportation activity [50,56], which is significantly increased together with other hemolymph-associated genes during the late InM and PrM stages. Additionally, energy reserves, including glycogen and lipids, are also accumulated in the hemolymph for the next molt [18].

Expression patterns of DEGs associated with hormones among molting stages

The regulatory mechanism of molting is a complex process that includes a network of signals involving many hormone genes such as molt-inhibiting hormone (MIH) and crustacean hyperglycemic hormone (CHH) peptide families that control the molting process by inhibiting YO ecdysteroidogenesis and secretion [57,58]. Although it is known that CHH is released from gut endocrine cells immediately before ecdysis in crustaceans, we know little regarding the stage-to-stage expression variation in molting, although this is clearly necessary to understand the significance of these hormones. In this study, CHH was up-regulated more than fifteenfold during the InM-IV and PrM stages. During the same period, the expression levels of ecdysteroid-regulated proteins were sharply decreased by more than tenfold, which may be attributed to the suppression effect of CHH. The expression of other ecdysteroid-regulated genes also obviously decreased during the InM-IV and PrM stages, for example, ecdysis triggering hormone receptor subtype-B, ecdysone receptors, and ecdysone-induced proteins. The similar expression results for CHH and ecdysteroid-related genes have been observed in other crustaceans just prior to ecdysis [33,59]. The same trend of up-regulated transcript levels was found in the vitellogenin gene (VG), which has been suggested to be an ecdysteroid-responsive gene in the molting process [60]. Farnesoic acid o-methyltransferase is another crustacean molting hormone-related category that may play key roles in growth regulation of crustaceans [3,61]. These enzymes were up-regulated from InM-III to PrM in the present study. In *P. Chinensis*, a stage-specific expression profile revealed that the highest expression level of farnesoic acid o-methyltransferase occurred at the intermolt stage, implying that the conversion of farnesoic acid to methyl farnesoate may be involved in the onset of molting processes [61]. Moreover, many other growth and development-related hormone genes were also identified as differentially expressed during the molting cycle, including estradiol 17-beta-dehydrogenase 8, estrogen sulfotransferase, hormone receptor, juvenile hormone epoxide hydrolase, and lutropin-choriogonadotropic hormone receptor. Although the functions of these genes that are potentially involved in hormone regulation are still not well characterized, it appears that *S. paramamosain* produces a rather complex regulation profile of hormone genes as in other crustacean species. In summary, we infer that hormone-regulated molting signals are likely play distinctly different roles in the *S. paramamosain* molting cycle, and these will serve as promising candidates for future analyses.

Expression patterns of DEGs associated with immune response and metabolism among molting stages

During early postmolt, *S. paramamosain* still has a soft cuticle and thus may be more vulnerable to bacteria, viruses, or predators [3,35,50]. As with all crustaceans, *S. paramamosain* does not have an adaptive immune system and instead relies upon an innate immune system to avoid exogenous stresses. Antimicrobial peptides (AMPs) are important effectors in innate immunity, and these were up-regulated from PoM to InM in *S. paramamosain*. The AMPs included crustin 1, crustin 4, and ALF 1. Moreover, other transcripts homologous to a number of immune-related genes, including c-type lectin gene families, were identified as differentially expressed during the molting cycle. The functions of these DEGs need to be further confirmed as immune and protective tactics that allow *S. paramamosain* to avoid stress during the molting cycle.

It is well known that crustaceans accumulate nutrients prior to molting to provide enough energy for ecdysis. Various changes in protein, lipid, and carbohydrate content have been detected during the course of the molting cycle in crustaceans [62,63]. In our study, changes in the lipid composition were found to correspond well with the functions of these lipids during the molting cycle. For example, lipid metabolism-related genes, including lipase 3, alcohol dehydrogenase, and fatty acid-binding protein, were accumulated by the crab before molting, reached a peak during PrM, and were depleted during PoM. Carbohydrates such as glucose and glycogen being used primarily as a direct source of metabolic energy have been observed in several crustacean species prior to the molt [4]. The expression profile of carbohydrate metabolism-related transcripts appears to reflect an increase in the energy requirements of *S. paramamosain* as the molt cycle progresses. Enzymes involved include UDP-glucuronosyltransferase and 6-phosphogluconolactonase. Most of the protein metabolism-related DEGs displayed relatively high expression during InM followed by a gradual decrease across the rest of the molting cycle. These genes included arginine kinase, glutamine synthetase 2, and glutamate dehydrogenase. In summary, the expression profiles of these transcripts indicated that molting induction creates stress that may impact on metabolic function.

5. Conclusion

To our knowledge, this is the first systematic transcriptome derived from whole crab bodies, including all the individual organs, during the molting cycle in crustaceans. This transcriptome dataset significantly expands the available genomic information for *S. paramamosain* and provides fundamental support for future research on the molecular mechanisms of the molting cycle in this species. Moreover, the new description of major transcriptional events during different molting stages and the determination of their timing provide temporal markers for future studies of molting progress and regulation in crustaceans.

Declarations

Acknowledgments

We thank LetPub (www.letpub.com) for its linguistic assistance during the preparation of this manuscript.

Authors' contributions

CLW and LL designed the study. LL, XL, YYF, WF performed the research and analyzed the data. LL wrote the paper. All authors have read and approved the manuscript.

Funding

This research was supported by the National Key R & D Program of China [grant number 2019YFD0900405], the Sci & Tech Project for Common Wealth of Ningbo [grant number 2017C10022], the Zhejiang Provincial Natural Science Foundation of China [grant number LY20C190005], and the K.C. Wong Magna Fund of Ningbo University.

Availability of data and materials

The datasets generated during the current study were deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (<https://www.ncbi.nlm.nih.gov/sra/PRJNA687923>) with the accession number PRJNA687923, SRR13347083, SRR13347093, SRR13347084, SRR13347091, SRR13347092, SRR13347088, SRR13347089, SRR13347085, SRR13347082, SRR13347090, SRR13347087, SRR13347080, SRR13347086, SRR13347081, SRR13347077, SRR13347079, SRR13347078 and SRR13347076.

Ethics approval and consent to participate

Not applicable (does not include any human data, research, or human tissue).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Ghanawi J, Saoud IP. Molting, reproductive biology, and hatchery management of redclaw crayfish *Cherax quadricarinatus* (von Martens 1868). *Aquaculture*. 2012;358:183-195.
2. Xu BP, Tu DD, Yan MC, Shu MA, Shao QJ. Molecular characterization of a cDNA encoding Na⁺/K⁺/2Cl⁻ cotransporter in the gill of mud crab (*Scylla paramamosain*) during the molt cycle: Implication of its function in osmoregulation. *Comparative Biochemistry and Physiology, Part A*. 2017;203:115-125.

3. Gao Y, Zhang X, Wei J, Sun X, Yuan J, Li F, Xiang J. Whole transcriptome analysis provides insights into molecular mechanisms for molting in *Litopenaeus vannamei*. PLoS ONE. 2015;10(12):e0144350. <https://doi.org/10.1371/journal.pone.0144350>.
4. Huang S, Wang J, Yue WC, Chen J, Gaughan S, Lu WQ, Lu Guoqing Q, Wang CH. Transcriptomic variation of hepatopancreas reveals the energy metabolism and biological processes associated with molting in Chinese mitten crab, *Eriocheir sinensis*. Scientific Reports. 2015;5:14015. <https://doi.org/10.1038/srep14015>.
5. Aiken DE. Photoperiod endocrinology and crustacean molt cycle. Science. 1969;164:149-155.
6. Skinner DM. Moulting and regeneration, in: Bliss, D.E. (Eds), The biology of crustacea integument, pigments, and hormonal processes. New York, pp. 1985a;44-128.
7. Wilder MN, Huong DTT, Jasmani S, Jayasankar V, Kaneko T, Aida K, Hatta T, Nemoto S, Wigginton A. Hemolymph osmolality, ion concentrations and calcium in the structural organization of the cuticle of the giant freshwater prawn *Macrobrachium rosenbergii*. Changes with the molt cycle. Aquaculture. 2009;292:104-110.
8. de Oliveira Cesar JR, Zhao B, Malecha S, Ako H, Yang J. Morphological and biochemical changes in the muscle of the marine shrimp *Litopenaeus vannamei* during the molt cycle. Aquaculture. 2006;261:688-694.
9. Yamasaki-Granados S, Ruíz-Fregozo M, Vega-Villasante F, Espinosa-Chaurand LD. Contributions to the biology of molting and growth of the Longarm river prawn *Macrobrachium tenellum* (Decapoda: Palaemonidae) in Mexico. Arch. Biol. Sci., Belgrade, 2012;64(2):651-658.
10. Covi JA, Bader BD, Chang ES, Mykles DL. Molt cycle regulation of protein synthesis in skeletal muscle of the blackback land crab, *Gecarcinus lateralis*, and the differential expression of a myostatin-like factor during atrophy induced by molting or unweighting. The Journal of Experimental Biology. 2010;213:172-183.
11. Shai A, Lilah G, Jenny T, Inbar P, Vered CC, Isam K, Aflalo ED, Amir S. Binary gene expression patterning of the molt cycle: the case of chitin metabolism. PLoS ONE. 2015;10(4):e0122602. <https://doi.org/10.1371/journal.pone.0122602>.
12. Foguesatto K, Boyle RT, Rovani MT, Freire CA, Souza MM. Aquaporin in different moult stages of a freshwater decapod crustacean: Expression and participation in muscle hydration control. Comparative Biochemistry and Physiology. Part A. 2017;208:61-69.
13. Zhou ZK, Gu WB, Wang C, Zhou YL, Tu DD, Liu ZP, Zhu QH, Shu MA. Seven transcripts from the chitinase gene family of the mud crab *Scylla paramamosain*: Their expression profiles during development and moulting and under environmental stresses. Aquaculture Research. 2018;49(8):3296-3308.
14. Gong J, Yu K, Shu L, Ye HH, Li SJ, Zeng CS. Evaluating the effects of temperature, salinity, starvation and autotomy on molting success, molting interval and expression of ecdysone receptor in early juvenile mud crabs, *Scylla paramamosain*. Journal of Experimental Marine Biology and Ecology. 2015;464:11-17.

15. Fu CR, Huang XS, Gong J, Chen XL, Huang HY, Ye HH. Crustacean hyperglycaemic hormone gene from the mud crab, *Scylla paramamosain*: cloning, distribution and expression profiles during the moulting cycle and ovarian development. *Aquaculture Research*. 2014;47(7):2183-2194.
16. Huang HY, Fu CR, Chen XL, Gong J, Huang XS, Ye HH. Molt-inhibiting hormone (MIH) gene from the green mud crab *Scylla paramamosain* and its expression during the molting and ovarian cycle. *Aquaculture Research*. 2015;46:2665-2675.
17. Xu ZN, Liu A, Li SK, Wang GZ, Ye HH. Hepatopancreas immune response during molt cycle in the mud crab, *Scylla paramamosain*. *Scientific Reports*. 2020;10(1):13102. <https://doi.org/10.1038/s41598-020-70139-2>.
18. Kuballa A, Elizur A. Novel molecular approach to study moulting in crustaceans. *Bull. Fish. Res. Agency*. 2007;20:53-57.
19. Cameron JN. Molting in the blue crab. *Scientific American*. 1985;252(5):102-109.
20. Phlippen MK, Webster SG, Chung JS, Dircksen AH. Ecdysis of decapod crustaceans is associated with a dramatic release of crustacean cardioactive peptide into the haemolymph. *J Exp Biol*. 2000;203:521-536.
21. Zhang Y, Wu Q, Fang S, Li S, Zheng H, Zhang Y, Ikhwanuddin M, Ma H. mRNA profile provides novel insights into stress adaptation in mud crab megalopa, *Scylla paramamosain* after salinity stress. *BMC Genomics*. 2020;21:559. <https://doi.org/10.1186/s12864-020-06965-5>.
22. Liu L, Fu Y, Zhu F, Mu C, Li R, Song W, Shi C, Ye Y, Wang C. Transcriptomic analysis of *Portunus trituberculatus* reveals a critical role for WNT4 and WNT signalling in limb regeneration. *Gene*. 2018;658:113-122.
23. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*. 2011;29:644-652.
24. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotech*. 2010;28:511-515.
25. Storey JD, Tibshirani R. Statistical significance for genome wide studies. *Proc. Nat. Acad. Sci*. 2003;100(16):9440-9445.
26. Young MD, Wakefield MJ, Smyth GK, Oshlack A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biology*. 2010;11:R14.
27. Mao X, Cai T, Olyarchuk JG, Wei LP. Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics*. 2005;21:3787-3793.
28. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta C_t$ method. *Methods*. 2001;25:402-408.
29. Tian ZH, Jiao CZ. Molt-dependent transcriptome analysis of claw muscles in Chinese mitten crab *Eriocheir sinensis*. *Genes & Genomics*. 2019;41:515-528.

30. Nguyen C, Nguyen GT, Nguyen VL, Pham QH, Nguyen HT, Pham TH, Nguyen TH, Ha TT, Dau HT, Vu TH, Nguyen DD, Nguyen TTN, Nguyen HN, Quyen VD, Chu HH, Dinh DK. De novo assembly and transcriptome characterization of major growth-related genes in various tissues of *Penaeus monodon*. *Aquaculture*. 2016;464:545-553.
31. Kuballa AV, Merritt DJ, Elizur A. Gene expression profiling of cuticular proteins across the moult cycle of the crab *Portunus pelagicus*. *BMC Biology*. 2007;5(1):45.
32. Johnson PT. Histology of the blue crab, *Callinectes sapidus*. A Model for the Decapoda. 1980. New York, NY: Praeger Publications.
33. Tom M, Manfrin C, Chung SJ, Sagi A, Gerdol M, De Moro G, Pallavicini A, Giulianini GP. Expression of cytoskeletal and molt-related genes is temporally scheduled in the hypodermis of the crayfish *Procambarus clarkii* during premolt. *The Journal of Experimental Biology*. 2014;217:4193-4202.
34. Lv JJ, Zhang DN, Gao BQ, Liu P, Li J. Transcriptome and MassARRAY analysis for identification of transcripts and SNPs for growth traits of the swimming crab *Portunus trituberculatus*. *Gene*. 2015;566:229-235.
35. Skinner DM. Interacting factors in the control of the crustacean molt cycle. *Am. Zool.* 1985b;25:275-284.
36. Chang ES. Physiological and biochemical changes during the molt cycle in decapod crustaceans, an overview. *J. Exp. Mar. Biol. Ecol.* 1995;193:1-14.
37. El Haj AJ. Regulation of muscle growth and sarcomeric protein gene expression over the intermolt cycle. *Am Zool.* 1999;39(3):570-579.
38. Relaix F, Zammit PS. Satellite cells are essential for skeletal muscle regeneration: the cell on the edge returns centre stage. *Development*. 2012;139:2845-2856.
39. Hornung DE, Stevenson JR. Changes in the rate of chitin synthesis during the crayfish molting cycle. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*. 1971;40(2):341-346.
40. Li JJ, Sun JS, Dong XW, Geng XY, Qiu GF. Transcriptomic analysis of gills provides insights into the molecular basis of molting in Chinese mitten crab (*Eriocheir sinensis*). 2019;PeerJ 7:e7182. <https://doi.org/10.7717/peerj.7182>.
41. Zou E. Effects of hypoxia and sedimentary naphthalene on the activity of N-acetyl- β -glucosaminidase in the epidermis of the brown shrimp, *Penaeus aztecus*. *Bull Environ Contam Toxicol*. 2009;82:579-582
42. Porrañee P, Anchalee T, Vichien R. Chitinases from the black tiger shrimp *Penaeus monodon*: phylogenetics, expression and activities. *Comp Biochem Physiol B*. 2010;156:86-96
43. Priya TA, Li F, Zhang J, Bing W, Xiang J. Molecular characterization and effect of RNA interference of retinoid X receptor (RXR) on E75 and chitinase gene expression in Chinese shrimp *Fenneropenaeus chinensis*. *Comp Biochem Physiol B*. 2009;153:121-129.
44. Cohen E. Chitin biochemistry: synthesis, hydrolysis and inhibition. In: Robert, H.S., (Eds.), *Advances in insect physiology*. Israel: Faculty of Agriculture, Food and Environment. Rehovot, pp. 2010;5-74.

45. Fan B, Li YM, Huang YH, Zhang M, Liu ZQ, Fan WJ, Zhao YL. Cloning and expression of chitin deacetylase 1 from *Macrobrachium nipponense*, and the effects of dietary protein on growth, body composition and digestive enzymes. *Aquaculture Nutrition*. 2018;24(6):1-15.
46. Sarmiento KP, Panes VA, Santos MD. Molecular Cloning and Expression of Chitin Deacetylase 1 Gene from the Gills of *Penaeus monodon* (Black Tiger Shrimp). *Fish & Shellfish Immunology*. 2016;55:484-489.
47. Fujimori K, Fukushima H, Matsumiya M. Molecular cloning and phylogenetic analysis of a chitin deacetylase isolated from the epidermis of the red snow crab *Chionoecetes japonicus*. *Advances in Bioscience and Biotechnology*. 2018;9:52-62.
48. Inoue H, Ohira T, Ozaki N, Nagasawa H. Cloning and expression of a cDNA encoding a matrix peptide associated with calcification in the exoskeleton of the crayfish. *Comp Biochem Physiol B*. 2003;136:755-765.
49. Faircloth LM, Shafer TH. Differential expression of eight transcripts and their roles in the cuticle of the blue crab, *Callinectes sapidus*. *Comp Biochem Physiol B*. 2007;146:370-383.
50. Seear PJ, Tarling GA, Burns G, Goodall-Copestake WP, Gaten E, Özkaya Ö, Rosato E. Differential gene expression during the moult cycle of *Antarctic krill* (*Euphausia superba*). *BMC Genomics*. 2010;11:582. <https://doi.org/10.1186/1471-2164-11-582>.
51. Ayme-Southgate AJ, Southgate RJ, Philipp RA, Sotka EE, Kramp C. The myofibrillar protein, projectin, is highly conserved across insect evolution except for its PEVK domain. *J. Mol. Evol.* 2008;67:653-669.
52. Verdier-Pinard P, Pasquier E, Xiao H, Burd B, Villard C, Lafitte D, Miller LM, Angeletti RH, Horwitz SB, Braguer D. Tubulin proteomics: towards breaking the code. *Anal. Biochem.* 2009;384:197-206.
53. Schafer DA, Gill SR, Cooper JA, Heuser JE, Schroer TA. Ultrastructural analysis of the dynactin complex: an actin-related protein is a component of a filament that resembles F-actin. *Journal of Cell Biology*. 1994;126:403-412.
54. Uddowla MH, Salma U, Kim HW. Molecular characterization of four actin cDNAs and effects of 20-hydroxyecdysone on their expression in swimming crab, *Portunus trituberculatus* (Miers, 1876). *Animal Cells & Systems*. 2013;17:203-212.
55. Rocha J, Garcia-Carreno FL, Muhlia-Almazan A, Peregrino-Uriarte AB, YepizPlascencia G, Cordova-Murueta JH. Cuticular chitin synthase and chitinase mRNA of whiteleg shrimp *Litopenaeus vannamei* during the molting cycle. *Aquaculture*. 2012;330:111-115.
56. Su J, Heng J, Huang T, Peng L, Yang C, Li Q. Identification, mRNA expression and genomic structure of TLR22 and its association with GCRV susceptibility/resistance in grass carp (*Ctenopharyngodon idella*). *Dev. Comp. Immunol.* 2012;36:450-462.
57. Chung JS, Webster SG. Dynamics of in vivo release of molt-inhibiting hormone and crustacean hyperglycemic hormone in the shore crab, *Carcinus maenas*. *Endocrinology*. 2005;146(12):5545-5551.

58. Chang ES, Mykles DL. Regulation of crustacean molting: a review and our perspectives. *Gen. Comp. Endocrinol.* 2011;172:323-330.
59. Chung JS. Hemolymph ecdysteroids during the last three molt cycles of the blue crab, *Callinectes sapidus*: quantitative and qualitative analyses and regulation. *Arch. Insect Biochem. Physiol.* 2010;73:1-13.
60. Shechter A, Aflalo ED, Davis C, Sagi A. Expression of the reproductive female-specific vitellogenin gene in endocrinologically induced male and intersex *Cherax quadricarinatus* crayfish. *Biology of Reproduction.* 2005;73:72-79.
61. Li ZX, Xu XF, Wang JY, Wang CD. Possible roles of farnesoic acid o-methyltransferase in regulation of molting in the shrimp, *Penaeus Chinensis*. *Journal of the World Aquaculture Society.* 2013;44(6):826-834.
62. Stoner AW, Ottmar ML, Copeman LA. Temperature effects on the molting, growth, and lipid composition of newly-settled red king crab. *Journal of Experimental Marine Biology and Ecology.* 2010;393:138-147.
63. Carter CG, Mente E. Protein synthesis in crustaceans: a review focused on feeding and nutrition. *Cent. Eur. J. Biol.* 2014;9(1):1-10.

Figures

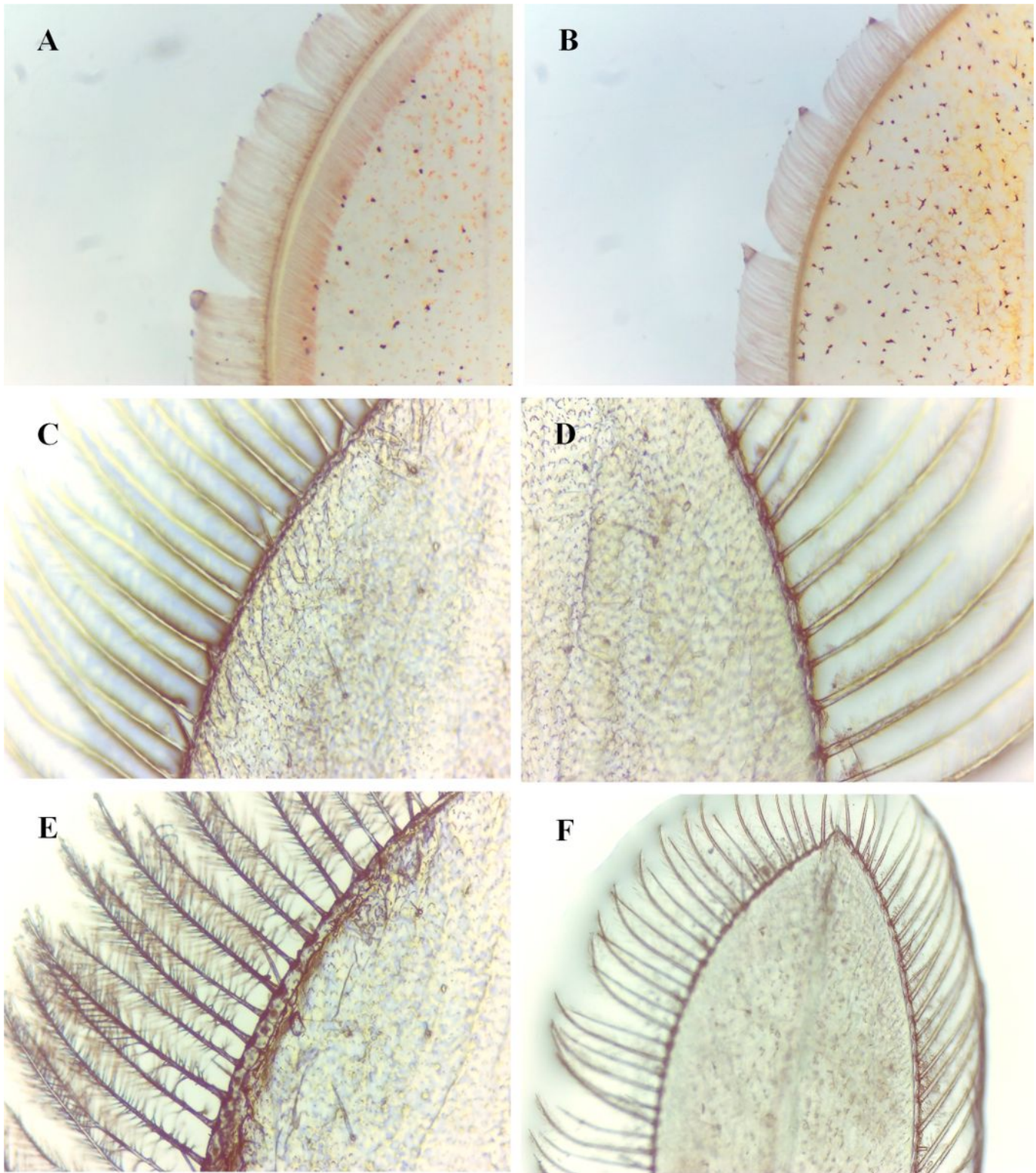


Figure 1

Morphological changes of crab uropods during molting stages under a light microscope. (A and B in 50 × magnification; C, D, E, and F in 400 × magnification) Observed physical characteristics of each stage were as follows: Pre-molt stage (A) a noticeably wider clear zone between the setal cones and the epidermis; new setae are fully generated. Post-molt stage (B) soft and delicate setae, dense connective tissue. Intermolt stage (C-F) C, setae harden, absence of setal cones; D, fully-spread epidermis; E, a clear margin

of epidermal tissue at the base of the setal cones; F, a narrow zone between the setal cones and the epidermis; setal cones arise.

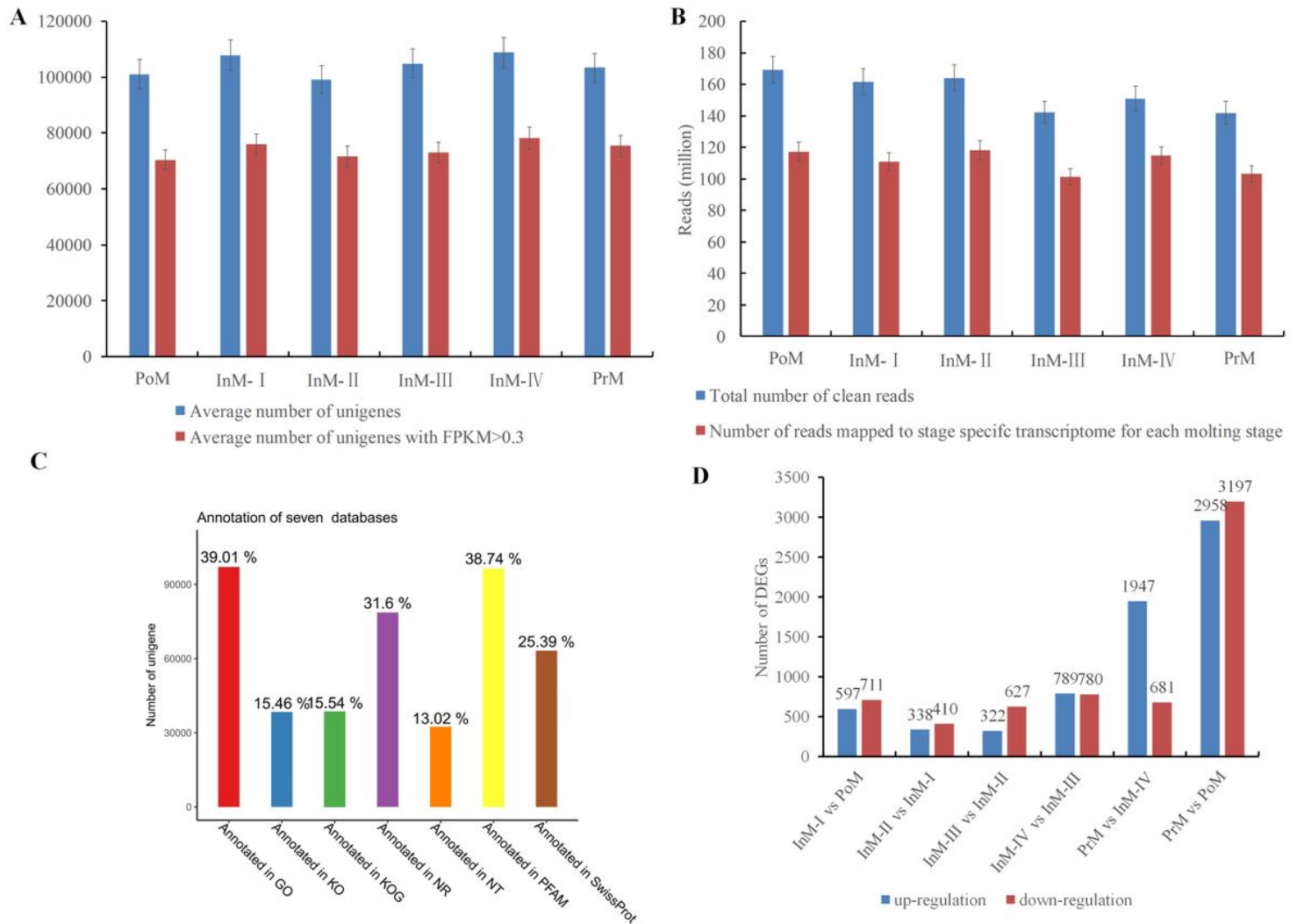


Figure 2

Assembly, mapping, annotation of transcriptomes, and differentially expressed genes (DEGs) between six comparisons of adjacent molting stages in *S. paramamosain*. (A) Average number of unigenes and number of genes with $FPKM \geq 1$; (B) Total number of reads, number of reads mapped to the reference transcriptome, and number of reads mapped to the stage-specific transcriptome for each molting stage; (C) Summary of unigenes annotated in different databases. NR NCBI non-redundant protein sequences, NT NCBI non-redundant nucleotide sequences, PFAM Protein family; Swiss-Prot A manually annotated and reviewed protein sequence database, KOG Clusters of Orthologous Groups of proteins, GO Gene Ontology, KEGG Kyoto Encyclopedia of Genes and Genomes database; All, total number of unigenes that were successfully annotated in at least one database; (D) DEGs between six comparisons of adjacent molting stages (InM-I vs PoM, InM-II vs InM-I, InM-III vs InM-II, InM-IV vs InM-III, PrM vs InM-IV, PrM vs

PoM). The number of upregulated and down-regulated genes between comparisons is given. The x-axis indicates adjacent stages in comparisons. The y-axis indicates the number of DEGs.

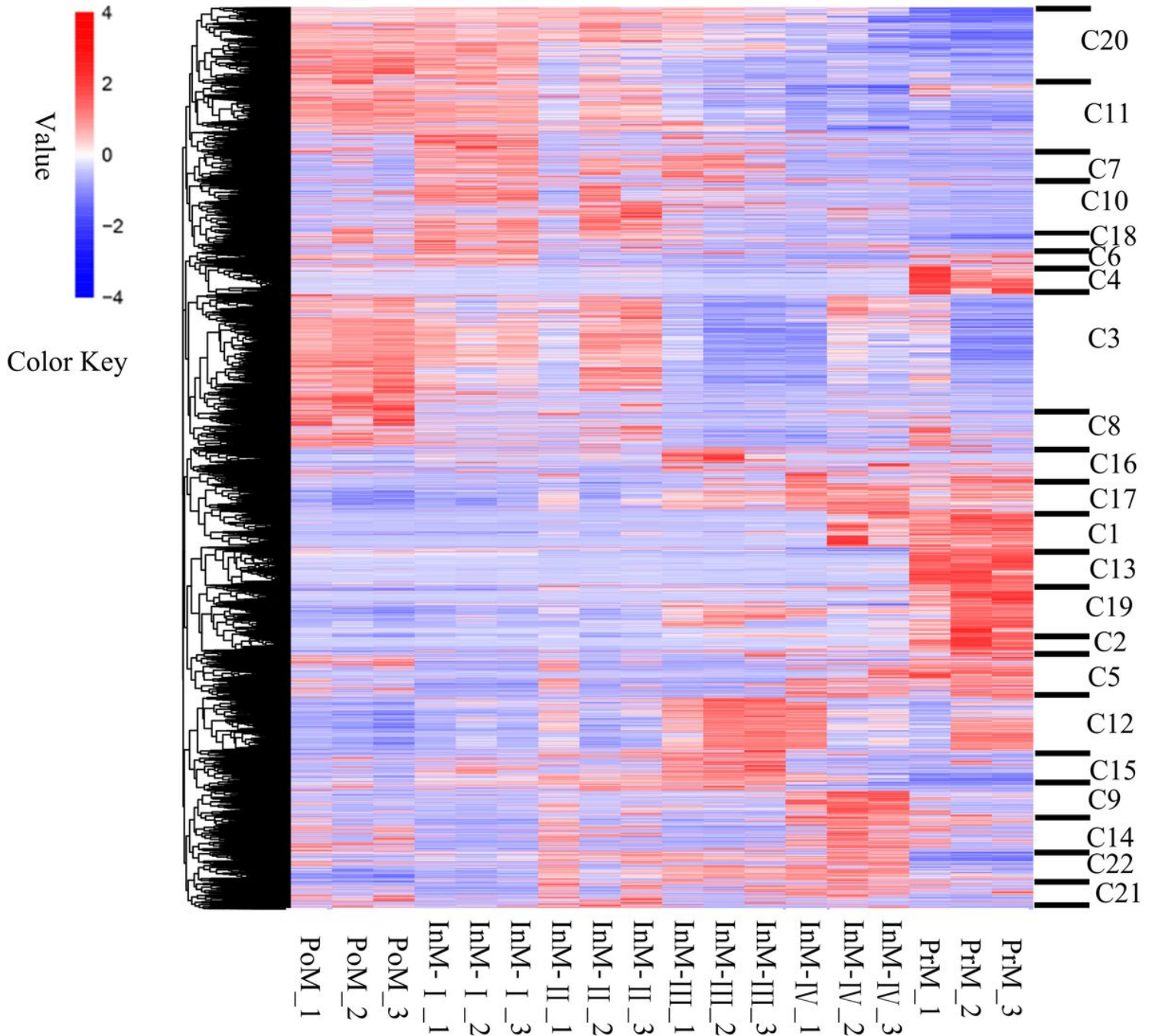
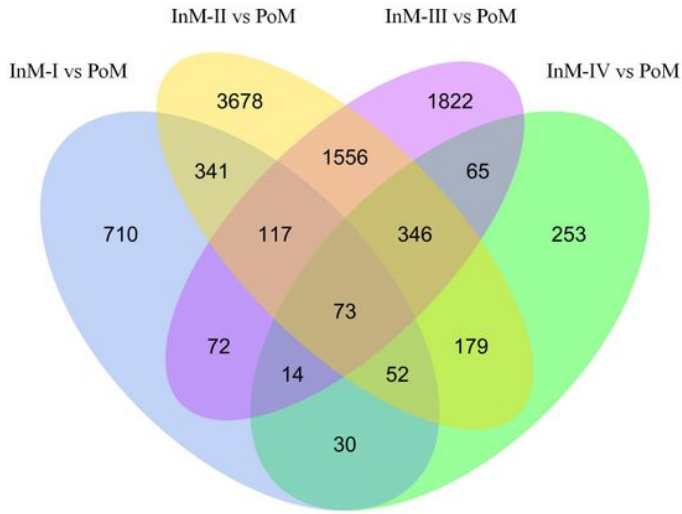


Figure 3

Heat map of differentially expressed genes clustered in two ways -molting stages and GO categories. Each stage has three replicate samples. C1–C22 -functional clusters of differentially expressed genes. Color key value -FPKM fold change.

A



B

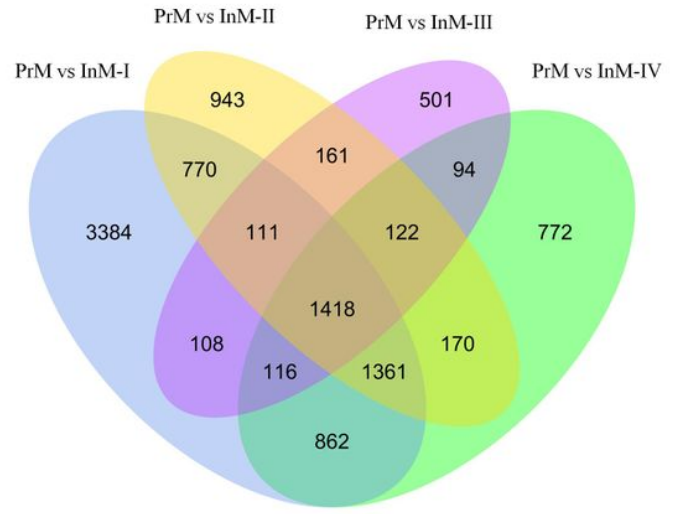


Figure 4

Venn diagram for the number of DEGs between comparisons of InM vs PoM and PrM vs InM. (A) InM-I vs PoM, InM-II vs PoM, InM-III vs PoM, InM-IV vs PoM; (B) PrM vs InM-I, PrM vs InM-II, PrM vs InM-III, PrM vs InM-IV.

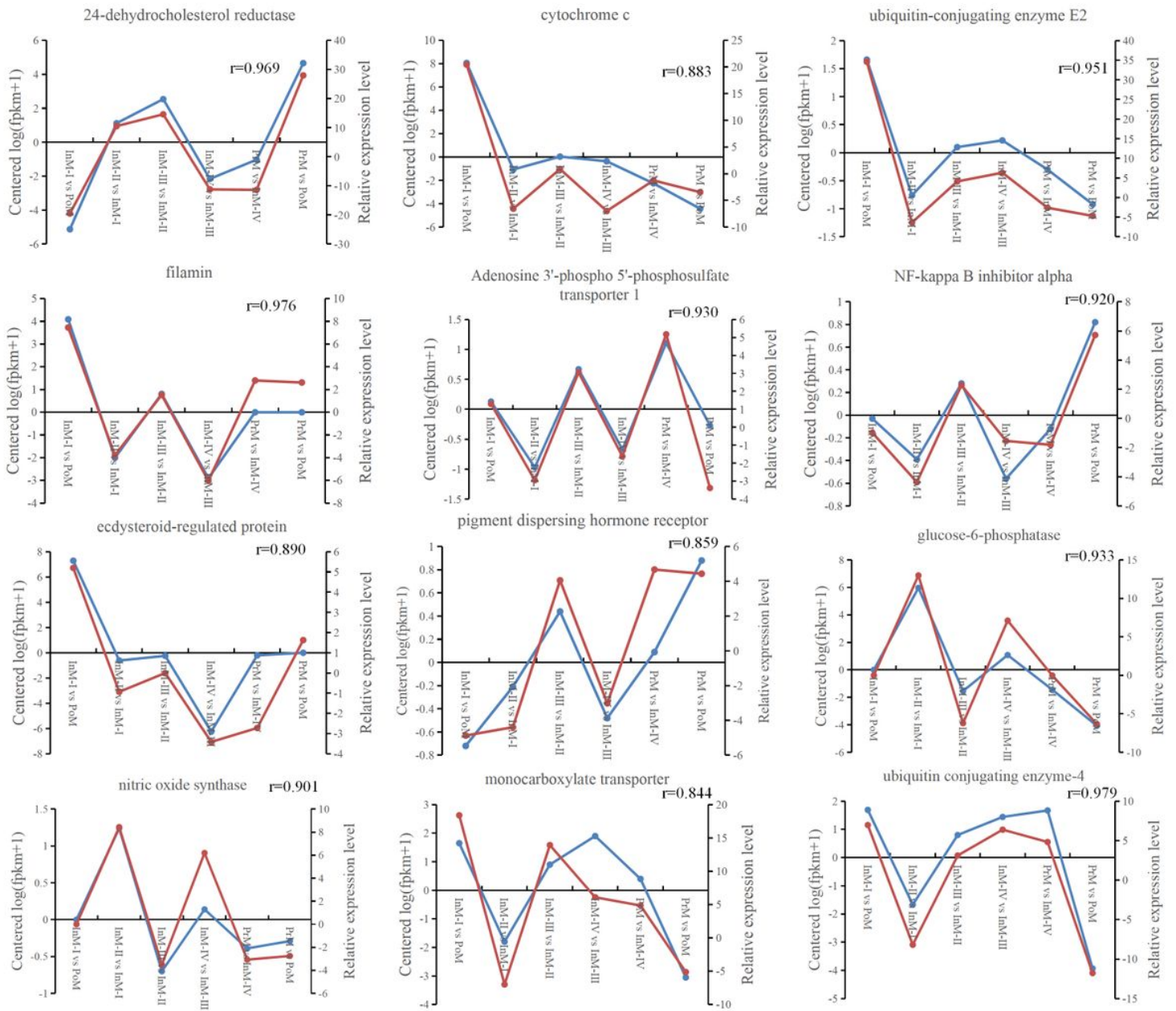


Figure 5

Expression profiles of 12 differentially expressed genes from RNA-Seq (blue) and qRT-PCR (red) with alpha-tubulin as a reference gene in different molting stages.

Supplementary Files

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

- [FigureS1.jpg](#)
- [FigureS2.jpg](#)
- [FigureS3.jpg](#)

- [FigureS4.jpg](#)
- [FigureS5.jpg](#)
- [TableS1.xlsx](#)
- [TableS2.xlsx](#)
- [TableS3.xlsx](#)
- [TableS4.xlsx](#)
- [TableS5.xlsx](#)
- [TableS6.xlsx](#)
- [TableS7.xlsx](#)
- [TableS8.xlsx](#)
- [TableS9.xlsx](#)
- [TableS10.xlsx](#)
- [TableS11.xlsx](#)
- [TableS12.xlsx](#)