Article

Not peer-reviewed version

Decoding the Impacts of Mating Behavior on Ovarian Development in Mud Crab (*Scylla paramamosain*): Insights from SMRT RNA-Seq

Chenyang Wu , Sadek Md Abu , Xiyi Zhou , Yang Yu , Mhd Ikhwanuddin , Waqas Waqas * , Hongyu Ma *

Posted Date: 18 August 2025

doi: 10.20944/preprints202508.1328.v1

Keywords: *Scylla paramamosain*; mating behavior; ovarian development; SMRT RNA-seq; alternative splicing

 doi:10.20944/preprints202508.1328.v1

Article

Decoding the Impacts of Mating Behavior on Ovarian Development in Mud Crab (*Scylla paramamosain*): Insights from SMRT RNA-Seq

Chenyang Wu [1,2,3], Sadek Md Abu [1,2,3], Xiyi Zhou [1,2,3], Yang Yu [1,2,3], Mhd Ikhwanuddin [4], Waqas Waqas [1,2,3,*] and Hongyu Ma [1,2,3,*]

[1] Guangdong Provincial Key Laboratory of Marine Biotechnology, Shantou University, Shantou 515063, China
[2] International Joint Research Center for the Development and Utilization of Important Mariculture Varieties Surrounding the South China Sea Region, Shantou University, Shantou 515063, China
[3] STU-UMT Joint Shellfish Research Laboratory, Shantou University, Shantou 515063, China
[4] Higher Institute Centre of Excellence (HICoE), Institute of Tropical Aquaculture and Fisheries, Universiti Malaysia Terengganu, Kuala Nerus 21030, Malaysia
* Correspondence: 21waqas@stu.edu.cn (W.W.); mahy@stu.edu.cn (H.M.); Tel.: +86-0754-86503471

Simple Summary

This study aims to explore how mating behavior influences ovarian development in the mud crab *Scylla paramamosain*, an important species in aquaculture. The single molecule real-time (SMRT) RNA-seq was employed to analyze the full-length transcriptome of female crabs at four stages: before molting, unmated, one day post-mating, and three days post-mating. A total of 59,156 high-quality isoforms were identified, with the most frequent alternative splicing events occurring in the first exon. The analysis found 1,157 differentially expressed transcripts and 21 differential splicing events between the unmated and one-day post-mating stages, mainly linked to amino acid metabolism. By the third day post-mating, 1,963 differentially expressed transcripts and 34 splicing events were identified, primarily associated with energy metabolism and detoxification. These findings provide valuable insights into the alternative splicing events and the gene expression changes triggered by mating behavior, which could have important applications in reproductive regulation and broodstock management for crustacean aquaculture.

Abstract

The mud crab *Scylla paramamosain* is a commercially valuable aquaculture species whose reproductive success depends heavily on ovarian maturation. However, the molecular effects of mating behavior on ovarian development remain poorly understood. In this study, we used Single Molecule Real-Time (SMRT) RNA-seq to construct a high-quality full-length transcriptome of female crabs at four stages: before molting (BM), unmated (UM), 1-day post-mating (M1), and 3-days post-mating (M3). A total of 59.6 million subreads (34.0 GB) yielded 880,044 circular consensus sequences (CCSs), from which 59,156 high-quality isoforms (average length: 2,074.8 bp) were obtained. After collapsing, 51,637 isoforms remained. Alternative splicing analysis identified alternative first exon (AF) as the most frequent event, followed by A5, A3, AL, RI, SE, and MX. Gene Ontology enrichment indicated protein-related processes as dominant. Comparative analysis revealed 21 differential alternative splicing (DAS) events and 1,157 differentially expressed transcripts (DETs) between UM and M1, mainly enriched in amino acid metabolism pathways. Between UM and M3, 34 DAS events and 1,963 DETs were associated with energy metabolism and detoxification. These findings highlight mating behavior-induced shifts in gene expression and splicing, particularly in metabolic and detoxification pathways, offering new insights into reproductive regulation and broodstock management in crustacean aquaculture.

 doi:10.20944/preprints202508.1328.v1

Keywords: *Scylla paramamosain*; mating behavior; ovarian development; SMRT RNA-seq; alternative splicing

1. Introduction

The genus *Scylla* includes commercially important aquaculture crab species cultured across Indo-Pacific regions [1,2]. Among the *Scylla* crabs, the mud crab (*Scylla paramamosain*), a key species, has emerged as one of the most economically important crustaceans and is widely cultured across China and Vietnam [3,4]. *S. paramamosain* is valued for its rapid growth, high fecundity, and strong market demand. In the year 2023, its production was recorded as 157,012 tons in China, showcasing its significance in crustacean aquaculture [5]. Female crabs are important from both ecological and aquaculture perspectives due to their economic value and reproduction capabilities [6]. In hatcheries, female crabs are the primary source of larval production, while their developed ovaries make them highly sought in seafood markets [7,8]. Additionally, their reproductive traits serve as sensitive indicators of environmental stress, making female crabs key models in environmental and molecular research [9]. Ovarian development is a complex process regulated by the neuroendocrine network involving various hormones and enzymes such as ecdysone and vitellogenin [10]. Ovarian development in female crabs after mating demands a higher energy and material supply than unmated individuals [11]. This process involves a multifaceted interplay of molecular entities, including genes, hormones, enzymes, lipids, and proteins, forming a complicated interactive network [12]. However, the molecular profiling of post-mated female mud crabs is largely unexplored, leaving a critical gap in the literature on how mating behavior influences the molecular dynamics of ovarian development.

In mud crabs, ovarian development follows a gradual process consisting of five stages: starting from the undeveloped phase (Stage I), progressing through pre-vitellogenic (Stage II), early and late vitellogenic stages (Stages III and IV), and culminating in the fully mature stage (Stage V) [13]. Before undergoing pubertal molt and mating, mud crabs usually have ovaries at Stage II, which are still small and appear translucent to milky white [14]. In the early stages of ovarian development, oogonia actively proliferate and then begin to differentiate into primary oocytes [15]. As Stage II progresses, follicular cells begin to proliferate around the cytoplasm of the developing oocyte [16]. Following mating, the mud crab undergoes rapid ovarian development and yolk accumulation [17]. Thus, mating behavior serves as a potential catalyst of ovary development and vitellogenesis in female crabs following pubertal molt [18]. Eubrachyuran crabs exhibit six mating strategies, differentiated by the timing of mating relative to molting (soft- or hard-shell), growth pattern (determinate or indeterminate), and seminal receptacle position (dorsal or ventral) [19]. The mud crabs mate during their soft-shell stage and has dorsal seminal receptacles. In the genus *Scylla*, mating involves courtship, extended guarding, sperm plugs, and the absence of hinged vulval coverings [20].

Next-generation sequencing (NGS) has made it possible to quickly and accurately analyze genetic material, opening new doors for understanding how genes function and interact [21,22]. NGS has enabled fast, affordable DNA sequencing, driving the development of tools like RNA-seq, exome sequencing, ChIP-seq, miRNA-seq, RAD-seq, and small RNA sequencing to tackle diverse biological [23–25]. However, short-read NGS poses limitations for accurate transcript reconstruction in eukaryotes due to complex splicing and transcript isoforms [26]. This complexity stems from the ability of a single gene to produce multiple isoforms through alternative transcription start sites and RNA processing. Long-read sequencing overcomes this limitation by capturing full-length isoforms, enabling more accurate transcriptome analysis [27]. SMRT RNA-seq produces highly accurate long reads, enabling analysis of alternative splicing, polyadenylation, genome annotation, fusion transcripts, isoform phasing, and long noncoding RNAs [28]. Its real-time, PCR-free workflow minimizes bias, reduces costs, and lowers computational requirements [29]. Full-length transcripts refine genome annotation by precisely mapping gene structures, regulatory elements, and coding regions, facilitating detailed isoform analysis across the transcriptome [30].

 doi:10.20944/preprints202508.1328.v1

Previously, Yu et al. [31] reported that ovarian development in unmated mud crabs is prolonged, taking up to 60 days with smaller oocytes, while mated females reach stage III (proliferative stage) within 23 days and develop larger oocytes, indicating that mating behavior triggers ovarian maturation. However, the molecular mechanism linking mating behavior to ovarian development remains unclear. In this study, we employed full-length transcriptome analysis using SMRT RNA-seq to explore differentially expressed genes, alternative splicing events, and regulatory pathways involved in ovarian development in *Scylla paramamosain* before and post-mating stages.

2. Materials and Methods

2.1. Ethics Statement

All animal handling and experimental procedures were carried out with care and in strict accordance with ethical guidelines, as approved by the Animal Care and Use Committee of Shantou University and in compliance with the Guide for the Care and Use of Laboratory Animals.

2.2. Crabs Collection and Mating

A total of 80 female and 30 male mud crabs were collected from a local seafood market in Chaozhou city (23.26◦ N, 116.54◦ E) of Guangdong Province in China, with detail information in our previous research [31]. Briefly, crabs were reared individually in water tanks containing filtered seawater maintained at 29 °C and fed twice daily with clams (Sinonovacula constricta). Pubertal molting and courtship were monitored using a real-time video surveillance system [32]. Artificial mating between male and female crabs was carried out, following the method of Fazhan et al. [17]. Immediately after reproductive molt, female crabs were paired with potential males in dark and oxygenated mating boxes for 1–2 days to facilitate copulation. Clean water and food were provided to maintain water quality and reduce the risk of male cannibalism during mating. Each female crab was used only once for mating, while the male crabs were used repeatedly. Additionally, four experimental groups were established to represent key stages of reproductive development: (1) BM – before pubertal molting, (2) UM – unmated females exhibiting courtship behavior without successful copulation, (3) M1 – females at 1-day post-mating, and (4) M3 – females at 3 days post-mating. Mating was confirmed by detecting spermatophores in the seminal receptacles of female crabs anesthetized on ice [33,34]. The crabs were then dissected, and approximately 50 mg of ovarian tissue was collected and preserved in RNA Keeper Tissue Stabilizer (Vazyme Biotech Co., Ltd., China). Following the incubation at 4 ℃ for 24 h, the tissue samples were stored -80 ℃ until RNA sequencing. A total of 12 ovarian tissue samples—three biological replicates each from the BM, UM, M1, and M3 groups—were subjected to RNA sequencing.

2.3. RNA Extraction and Sequencing

Total RNA was extracted using the TRIzol Reagent kit (ComWin Biotech, China). RNA quality was assessed by measuring concentration and integrity using an Agilent 2100 Bioanalyzer (Agilent Technologies, USA), and a Nanodrop 2000 (Agilent Technologies, CA, USA). Only high-quality RNA samples with a RNA integrity number (RIN) ≥ 8 and a 28S/18S ratio ≥ 1.5 were selected for subsequent library construction. For SMRT sequencing, libraries were constructed and sequenced using the PacBio Sequel platform. Briefly, PacBio single-molecule long reads were processed using RS_IsoSeq (v2.3) in the SMRT Analysis package (v2.3.0.140936.p4.150482) via command line to obtain insert reads. Full-length transcript characterization was performed using the pbtranscript.py script from the same package. The SMRT library was constructed from pooled RNA extracted from all 12 crab ovarian tissue samples. PacBio single-molecule long reads were processed using RS_IsoSeq (v2.3) in the SMRT Analysis package (v2.3.0.140936.p4.150482) via command line to obtain insert reads. Full-length transcript characterization was performed using the pbtranscript.py script from the same package. The Clontech kit was used to identify the 5′ and 3′ primers, with the poly(A) tail upstream of the 3′ primer serving as a key signal for identifying strand-specific full-length reads. Sequencing

 doi:10.20944/preprints202508.1328.v1

errors in consensus reads were corrected using LSC 2.0 (https://www.healthcare.uiowa.edu/labs/au/LSC/) with Illumina short reads derived from ovarian tissues of *S. paramamosain* (parameters: runLSC.py –long_reads SQ_SMRT.fa –short_reads SQ_Illumina.fa –output output). The completeness of the SMRT-based transcriptome assembly was then evaluated using DETONATE and Ex90N50 metrics. For Illumina sequencing, three libraries were prepared per group (BM, UM, M1, and M3) and sequenced on the Illumina HiSeq 2500 platform to generate 150 bp paired-end reads. Library construction and sequencing were carried out by Beijing Biomarker Technologies Co., Ltd. The RNA-seq date were deposited in the Sequence Read Archive of the National Center for Biotechnology Information (NCBI) with the accession number of PRJNA1306745.

2.4. SMRT Sequencing Data Processing

To identify transcript isoforms, we processed the subread sequence data using a structured bioinformatics pipeline. Firstly, circular consensus sequences (CCS) were generated from the raw subreads using ccs (v4.2.0) with parameters set to —minLength 50, —maxLength 15,000, and —minPasses 1 to ensure high-quality reads. Next, full-length (FL) reads were obtained by removing primers and demultiplexing with lima (v1.11.0), applying the —dump-clips and —peek-guess parameters for accurate read processing. To eliminate noise, FL reads were refined using the isoseq3 (v3.3.0) refine module, requiring polyadenylation (—require-polya) and a minimum poly(A) tail length of 20 nucleotides (—min-polya-length 20). These filtered reads were then clustered into unpolished transcript consensus sequences using the same IsoSeq3 refine module with default settings. Finally, the unpolished transcripts were polished using the isoseq3 polish module (default parameters) to generate high- and low-confidence transcript isoforms, ensuring robust isoform identification for downstream analysis.

2.5. Collapsing Redundant Transcripts Isoforms

To reduce redundancy among transcript isoforms, we implemented a two-step collapsing approach. Firstly, high-quality isoforms were aligned to the *Scylla paramamosain* reference genome [35] using minimap2 (v2.18) (default parameters), generating SAM files [36]. Redundant mapped isoforms were then collapsed using cDNA_Cupcake (https://github.com/Magdoll/cDNA_Cupcake) with parameters (—min_aln_coverage 0.95, —min_aln_identity 0.85, —dun-merge-5-shorter). For unmapped transcripts, we employed Cogent (v8.0) (https://github.com/Magdoll/Cogent) and cDNA_Cupcake (default parameters) to identify distinct gene families. A "fake genome" was constructed by concatenating Cogent-derived contigs, allowing unmapped transcripts to be realigned and collapsed following the same criteria. Finally, CD-HIT (v4.8.1) [37] was applied to both mapped and unmapped isoforms to remove highly similar sequences (identity threshold: 95%), ensuring a non-redundant transcriptome for downstream analysis.

2.6. Completeness and Characteristics Analysis of Reconstructed Transcriptomes

To evaluate the quality and completeness of the full-length transcriptomes, we conducted benchmarking universal single-copy orthologs (BUSCO) (v5.1.3) [38] analysis in transcriptome mode using the Arthropoda lineage dataset (arthropoda_odb10) to assess the representation of evolutionarily conserved single-copy orthologs [39]. The full-length transcripts were then systematically classified by comparing them to the reference genome annotation using gffcompare (v0.12.2) (http://ccb.jhu.edu/software/stringtie/gffcompare.shtml) [40] (Table 1). To identify shared and unique transcripts across the datasets, we performed pairwise comparisons using BLASTn (v2.11.0+) with stringent parameters (*–e* value ≤ 1×10^{-10} and percent identity ≥ 95%), where each transcriptome was alternately used as both the reference database and query sequence. This comprehensive analytical approach enabled us to thoroughly characterize the transcriptomes and identify both conserved and sample-specific transcripts.

 doi:10.20944/preprints202508.1328.v1

Table 1. The 15 types of class codes and their descriptions.

Class code	Description
=	complete, exact match of intron chain
c	contained in reference (intron compatible)
k	containment of reference (reverse containment)
m	retained intron(s), all introns matched or retained
n	retained intron(s), not all introns matched/covered
j	multi-exon with at least one junction match
e	single exon transfrag partially covering an intron.possible pre-mRNA fragment
o	other same strand overlap with reference exons
s	intron match on the opposite strand (likely amapping error)
x	exonic overlap on the opposite strand (like o or ebut on the opposite strand)
i	fully contained within a reference intron
y	contains a reference within its intron(s)
p	possible polymerase run-on (no actual overlap)
r	repeat (at least 50% bases soft-masked)
u	none of the above (unknown, intergenic)

2.7. Gene Functional Annotation

To elucidate the biological functions of the full-length transcripts, we performed comprehensive gene functional annotation through a multi-step bioinformatics. Firstly, we employed TransDecoder to identify potential open reading frames (ORFs) from the transcripts, selecting the first ORF when multiple candidates were present in a single transcript. These ORFs were then functionally annotated using eggNOG-mapper (v2.1.4) [41] to obtain classifications from multiple databases, including Clusters of Orthologous Groups (COG), Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and Protein families database (Pfam). To further validate and expand the functional annotations, we conducted homology searches against four major protein databases (Swiss-Prot, TrEMBL, UniRef90, and NR) using DIAMOND (v2.0.4.142) with stringent parameters (blastp, -outfmt 6, -max_target_seqs 1, $-e$ value 1×10^{-5}). Finally, all annotation results were consolidated into a comprehensive, tab-delimited summary file to facilitate further functional analyses.

2.8. Alternative Splicing (AS) Events Analysis

Alternative splicing (AS) events were analyzed using SUPPA2 software [42], and seven major types of AS events were identified using the generateEvents function of SUPPA2 with default parameters. These included skipped exon (SE), mutually exclusive exon (MX), alternative 5′ splice site (A5), alternative 3′ splice site (A3), retained intron (RI), alternative first exon (AF), and alternative last exon (AL). The distribution and frequency of these AS events were compared across different mating groups to identify potential splicing differences linked to genetic background. Furthermore, Gene Ontology (GO) enrichment analysis of genes associated with AS events was conducted using the clusterProfiler package, based on functional annotations obtained through eggNOG-mapper.

2.9. Quantification of Identified Transcripts

To investigate differential alternative splicing (AS) events among the BM, UM, M1, and M3 groups, subread sequences from each group were processed to reconstruct group-specific transcriptomes, following the pipeline previously described. Transcript and gene expression quantification were then performed using Salmon (mapping-based mode) [43]. Firstly, a reference

 doi:10.20944/preprints202508.1328.v1

transcriptome index was generated by constructing a k-mer hash (k = 31) from the reconstructed transcript sequences. Subsequently, clean RNA-seq reads were aligned to the corresponding group-specific transcriptomes using Salmon's quant module with the parameters —l IU and —validateMappings for fast and accurate alignment. Finally, transcript abundance was estimated using the quantmerge module, which produced transcript per million (TPM) values for each sample. These TPM values were subsequently used to calculate percent spliced-in (PSI) values, providing a quantitative measure of alternative splicing across the different groups.

2.10. Differential Alternative Splicing (DAE) Events, Differential Expressed Transcripts (DETs) and Their Enrichment Analysis

To identify differential alternative splicing (DAS) events between four groups (BM, UM, M1, and M3), PSI values for each AS event were calculated using the psiPerEvent function of SUPPA2, based on the transcript per million (TPM) values and the corresponding AS event annotations. Differential splicing analysis was then performed using the diffSplice function in SUPPA2, with statistical significance defined by a threshold of $|\Delta PSI| \geq 0.15$ and $p < 0.05$. In parallel, differential expression analysis of transcripts between groups (BM, UM, M1, and M3) was conducted using DESeq2 [44], based on TPM. Differentially expressed transcripts (DETs) were identified using the cutoff $|\log_2(\text{fold change})| \geq 1$ and $p < 0.05$. To explore the functional implications of both significant DAS events and DETs, Gene Ontology (GO) and KEGG pathway enrichment analyses were performed. Gene annotations were first retrieved from eggNOG-mapper results, and enrichment analysis was carried out using clusterProfiler [45]. Only GO terms or KEGG pathways with $p < 0.05$ were considered significantly enriched.

2.11. Validation of Differentially Expressed Genes

To validate the transcriptome sequencing results, five differentially expressed genes were selected for quantitative real-time PCR (qRT-PCR). Specific primers for the selected five genes were designed using Primer Premier 5.0. To validate differential gene expression between (BM, UM, M1, and M3), two qPCR kits, including the miRcute Plus miRNA qPCR Kit (SYBR Green) and the Talent qPCR Premix (SYBR Green) (both from TIANGEN Biotech, Beijing, China), were used to perform qRT-PCR for the five selected genes based on a LightCycler® 480 system (Roche Applied Science, Indianapolis, IN, USA). 18S rRNA served as the internal control (reference gene). Each gene was amplified in three biological replicates and three technical replicates. Relative expression levels were calculated using the 2^−ΔΔCt method [46]. Statistical significance was determined using a Student's t-test ($p < 0.05$) implemented in R software.

3. Results

3.1. Summary of PacBio Iso-Seq Data and Collapsing Redundant Isoforms

In the present study, PacBio SMRT RNA-seq was performed on RNA samples obtained from four experimental groups (BM, UM, M1, and M3), generating a total of 59,566,665 subreads (approximately 34.0 GB) and 880,044 circular consensus sequences (CCSs). Following IsoSeq3 refinement, clustering, and polishing, a total of 59,156 high-quality isoforms (average length: 2,074.8 bp) and 36 low-quality isoforms (average length: 2,293.1 bp) were obtained. Given their minimal proportion, the low-quality isoforms were excluded from subsequent analyses (Table 2). Reference-guided collapsing of redundant transcripts yielded 50,170 unique isoforms. Additionally, pseudo-reference genome construction identified 1,483 novel isoforms. The final integration of isoforms using CD-HIT resulted in a non-redundant transcriptome comprising 51,637 isoforms, with an average length of 2,056 bp (Table 3).

 doi:10.20944/preprints202508.1328.v1

Table 2. The summary of PacBio sequencing data in ovary.

Types	Numbers of Sequences	Length of Isoforms			N50[6]
		Min	Mean	Max	
Subreads	59,566,665	51	1,705.6	240,063	2,300
CCS[1]	880,044	97	2,222.4	14,561	2,721
FL[2]	717,465	51	2,057.4	10,605	2,681
FLCN[3]	713,486	50	2,012.4	10,573	2,654
HQ[4]	59,156	51	2,074.8	8,730	2,769
LQ[5]	36	140	2,293.1	8,310	3,319

[1] CCS = circular consensus sequence; [2] FL = full-length; [3] FLNC = full-length-non-chimeric; [4] HQ = high-quality isoforms; [5] LQ = Low-quality isoforms; [6] N50 = 50% of reads are longer than this value.

Table 3. The summary of features of transcript isoforms after collapsing redundant isoforms with cDNA cupcake, cogent, and CD-HIT.

Numbers of Transcript Isoforms after Collapsing Redundants				Length of Collapsing Redundant Isoforms			N50[1]
Reference Genome	Fake Genome	Unmap-Ped	Merge	Min	Max	Mean	
50,170	1,483	25	51,637	82	8,730	2,056	2,761

[1] N50 = 50% of reads are longer than this value.

3.2. Evaluation of Reconstructed Transcriptomes

In this study, the completeness and structural characteristics of the reconstructed transcriptome were comprehensively evaluated (Figure 1). Using BUSCO analysis, we identified 25.8% complete and single-copy orthologs (n = 261), 57.7% complete and duplicated (n = 584), 1.9% fragmented (n = 20), and 14.6% missing orthologs (n = 148) (Figure 1A), indicating a reasonably complete representation of conserved gene content. Comparative analysis with the reference genome annotation revealed that the reconstructed transcriptome comprised 11,101 transcripts matching known isoforms (category "c"), 14,824 potentially novel isoforms sharing splice junctions with known genes (category "j"), and 6,289 entirely novel transcripts from previously unannotated loci (category "u") (Figure 1B). These findings suggest that the reconstructed transcriptome not only recovers a substantial proportion of the reference annotation but also uncovers a considerable number of novel and potentially functionally relevant isoforms.

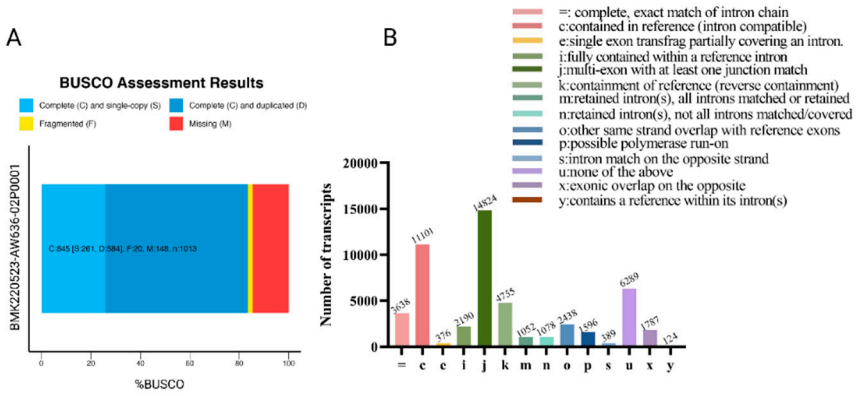

Figure 1. The completeness and characteristics analysis of reconstructed transcriptomes. (A) BUSCO assessment results of collapsed redundant transcripts. The Y-axis represents the reconstructed transcriptome. Both the X-

 doi:10.20944/preprints202508.1328.v1

axis and the different colors of the box represent the proportion of different categories, including complete and single-copy, complete and duplicated, fragmented, or missing. (B) The comparison of reconstructed transcriptomes with reference genome annotation using gffcompare v0.12.2 software. The Y-axis represents the number of genes. The X-axis and the different colored bars represent different categories.

3.3. Functional Annotation

To gain biological insight into the reconstructed transcriptome, functional annotation was performed and aligned with multiple well-established databases. The number of transcript isoforms annotated varied across different databases, ranging from 26,827 (Swiss-Prot) to 37,883 (TrEMBL), demonstrating broad but variable annotation coverage (Figure 2A). The overlap in annotation results across four databases — Swiss-Prot, TrEMBL, UniRef90, and NR. A total of 38,124 transcript isoforms were annotated in at least one of these databases, among which 26,826 isoforms were commonly identified by all four, reflecting a high degree of annotation consistency and supporting the reliability of the reconstructed transcriptome (Figure 2B). To further characterize transcript functions, isoforms were classified into Clusters of Orthologous Groups (COG) categories, as presented in Figure 2C. The largest group of transcripts was assigned to Category S (Function unknown), followed by Category T (Signal transduction mechanisms) and Category O (Posttranslational modification, protein turnover, chaperones). This functional distribution suggests the transcriptome encompasses a wide range of cellular roles, with a substantial portion representing uncharacterized or potentially novel functions.

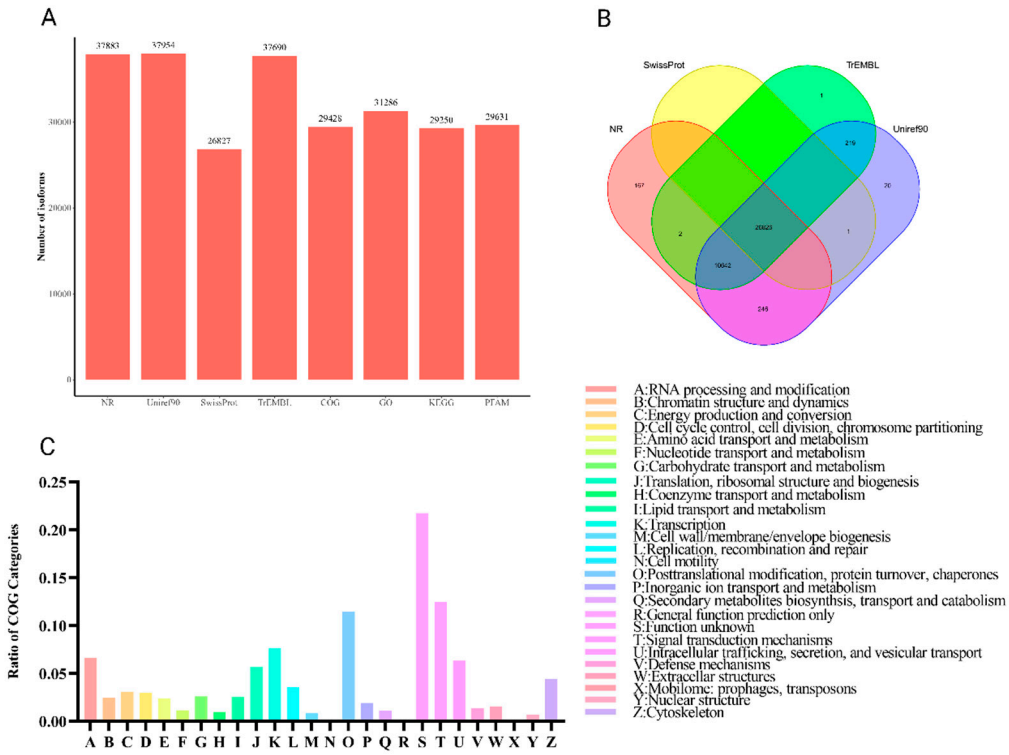

Figure 2. The summary of gene functional annotation using different databases. (A) Statistics of isoforms annotation results using different databases, including NR, Uniref90, Swiss-Prot, TrEMBL, COG, GO, KEGG, and PFAMs. The Y-axis represents the number of annotated isoforms. The X-axis represents different databases. (B) Venn diagrams showing the overlapping isoforms annotation results obtained using a different database. (C) COG profiles of transcript isoforms.

 doi:10.20944/preprints202508.1328.v1

3.4. Alternative Splicing Events

To further characterize transcriptomic complexity, we investigated alternative splicing (AS) events in the reconstructed transcriptome. A total of seven distinct AS event types were identified: alternative 3′ splice site (A3), alternative 5′ splice site (A5), alternative first exon (AF), alternative last exon (AL), mutually exclusive exon (MX), retained intron (RI), and skipped exon (SE). A detailed analysis of AS patterns in ovarian tissue revealed that AF (alternative first exon) was the most prevalent event, followed by A5, A3, AL, RI, SE, and MX (Figure 3A; Table 4). To elucidate the biological significance of these splicing events, Gene Ontology (GO) enrichment analysis was performed on all genes undergoing alternative splicing. As shown in Figure 3B, the top ten significantly enriched biological process (BP) terms were predominantly associated with protein regulation and chromatin remodeling. These included peptidyl-serine dephosphorylation, positive regulation of histone deacetylation, and histone monoubiquitination, suggesting that alternative splicing in ovarian tissue may contribute substantially to post-translational modification and epigenetic control mechanisms.

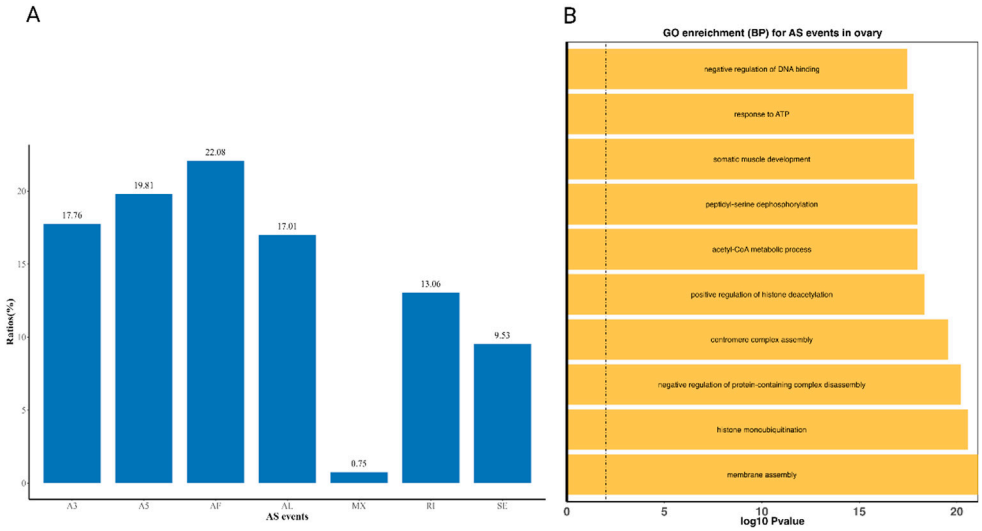

Figure 3. The summary overview of variable splicing events in the reconstructed transcript. (A) The proportion of each AS type. The Y-axis represents the proportion of different AS events. The X-axis represents different AS event types, including SE (skipped exon), MX (mutually exclusive exon), A5 (alternative 5′ splice site), A3 (alternative 3′ splice site), RI (retained intron), AF (alternative first exon), and AL (alternative last exon). (B) The top ten significantly biological processes (BP) obtained from Gene Ontology (GO) enrichment analysis using genes with AS events in the reconstructed transcript. The Y-axis represents different BP categories. The X-axis represents the corresponding −log10 transformed p-value.

Table 4. The number of each AS events.

AS events	A3	A5	AF	AL	MX	RI	SE
Numbers	2400	2676	2983	2298	101	1764	1288

3.5. Differential Alternative Splicing (DAS) Events, Differential Expressed Transcripts (DETs), and Their Enrichment Analysis

To better understand transcriptomic differences between experimental groups, we examined both differential alternative splicing (DAS) events and differentially expressed transcripts (DETs) across the BM, UM, M1, and M3 samples. The DAS analysis revealed that the number of significant splicing events varied among the six pairwise comparisons. Most of these events were linked to protein-coding genes and were dominated by four splicing types: retained intron (RI), alternative 5′

 doi:10.20944/preprints202508.1328.v1

splice site (A5), alternative 3′ splice site (A3), and alternative first exon (AF) (Table 5; Figure 4A). In terms of differential expression, the UM vs M3 comparison showed the largest number of significantly different transcript isoforms (1,963), while M1 vs M3 had the fewest (921) (Table 5; Figure 4B). Interestingly, the most highly differentially expressed transcripts (based on p-value) varied depending on the comparison group. For instance, in the UM vs M1 group, notable transcripts included PB.3431.2, PB.10089.7 (with domains found in myosin and kinesin tails), PB.10152.5, PB.552.5 (a member of the Rho family of small GTPases), and PB.9323.2. In the UM vs M3 group, top transcripts included PB.8738.1 (belonging to the uncharacterized UPF0029 protein family), along with PB.5936.9, PB.229.2, PB.9078.1, and PB.9829.3. For the M1 vs M3 comparison, prominent transcripts included PB.5147.8 (containing a BRIX domain), PB.6586.23 (with a VWC_def domain), PB.2909.2 (from the glycosyl hydrolase family 31), PB.6759.12 (from the class-I aminoacyl-tRNA synthetase family), and PB.705.6 (a DnaJ homolog subfamily B member).

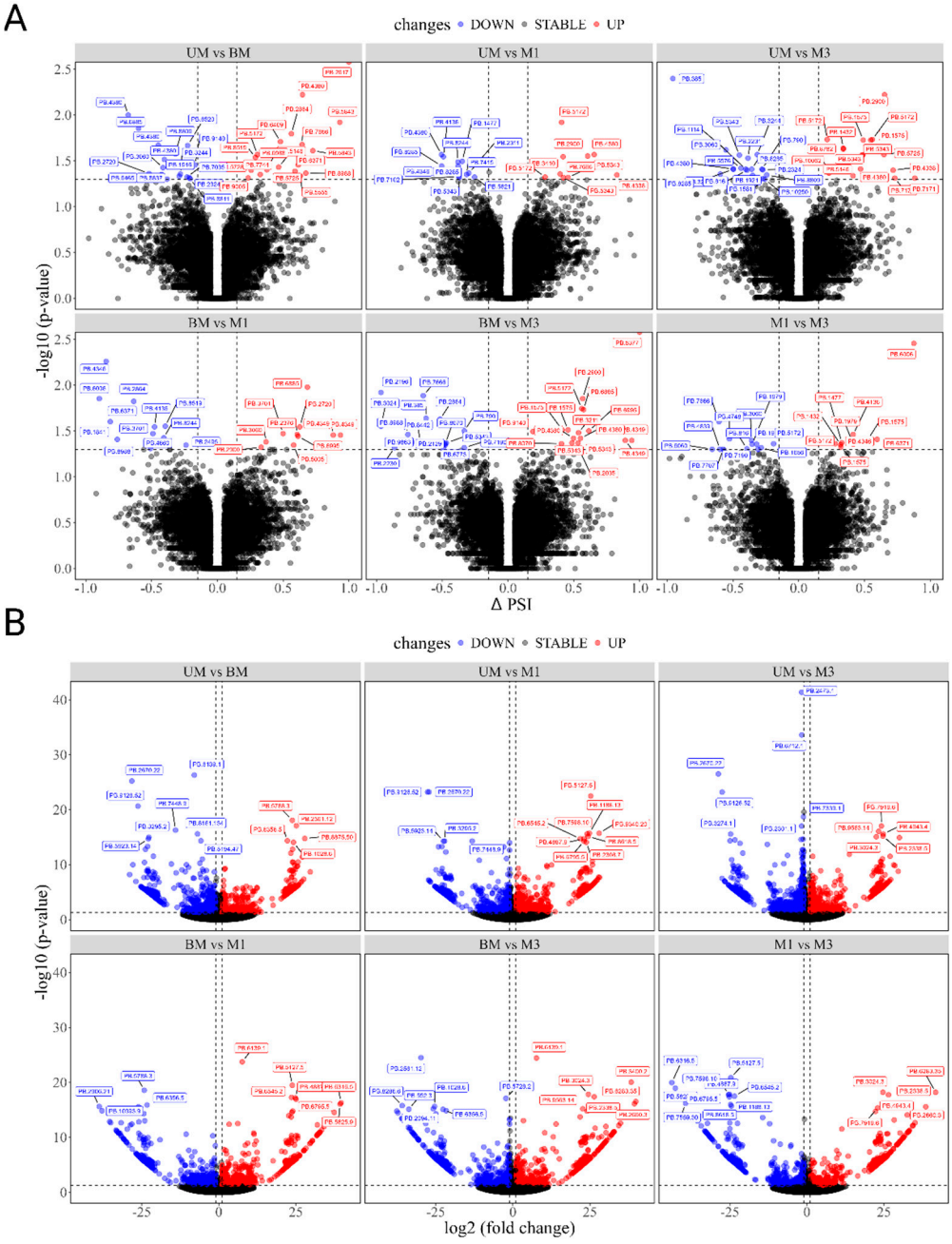

Figure 4. Differential expression analysis among different groups. (A) The volcano plot indicates p-values with minus log10-transformed for AS events (Y-axis) against their corresponding difference in inclusion levels (ΔPSI)

 doi:10.20944/preprints202508.1328.v1

of each AS event (X-axis). The horizontal gray dotted line represents the significant threshold (0.05). The red, blue, and gray points represent up-regulated, down-regulated, and non-regulated AS events in groups, respectively. (B) The volcano plot indicates with minus log10-transformed for genes (Y-axis) against their corresponding log2(|fold change|) of echo gene (X-axis).

Table 5. The summary of DAS and DETs characteristics of the six comparison groups.

Group	DAS				DETs		
	numbers	upregulated	downregulated	Protein-coding Genes	numbers	upregulated	downregulated
UM vs BM	34	18	16	30	1032	490	542
UM vs M1	21	9	12	16	1157	575	582
UM vs M3	34	16	18	27	1963	1004	959
BM vs M1	22	10	12	19	975	515	460
BM vs M3	32	17	15	26	1245	664	581
M1 vs M3	22	10	12	18	921	457	464

3.6. DAS and DETs Enrichment Analysis

To gain deeper insights into the molecular mechanisms underlying transcriptomic changes, both Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed using genes associated with differentially expressed transcripts (DETs) and differentially alternative splicing (DAS) events. In the UM vs M1 comparison, the top ten significantly enriched GO biological process terms were primarily linked to signaling pathway activation and neurochemical processes. These included primary amino compound metabolic processes, serotonin metabolism, positive regulation of catecholamine secretion, and negative regulation of cardiac contractility via chemical signaling (Figure 5A).

KEGG pathway analysis further revealed enrichment in biosynthesis and metabolic pathways, such as betalain biosynthesis, isoquinoline alkaloid biosynthesis, indole alkaloid biosynthesis, arachidonic acid metabolism, phenylalanine metabolism, and pyruvate metabolism (Figure 5B). In the UM vs M3 comparison, GO enrichment pointed predominantly to pathways involved in energy metabolism, including oxidative phosphorylation, ribonucleoside triphosphate biosynthetic process, ATP biosynthetic process, respiratory electron transport chain, and mitochondrial ATP synthesis coupled with electron transport (Figure 5C). The KEGG results highlighted oxidative phosphorylation, non-alcoholic fatty liver disease (NAFLD), and carbon fixation in photosynthetic organisms as the most enriched pathways (Figure 5D). A broader integrative analysis of DAS events and DETs across all six group comparisons revealed a recurring enrichment in functional categories related to cellular energy production, protein complex formation, detoxification, and developmental regulation. The top GO terms included protein tetramerization, oxidative phosphorylation, cellular detoxification, biogenic amine metabolism, glucose catabolic process, and regulation of endodermal cell differentiation (Figure 5E). Correspondingly, KEGG pathway enrichment across comparisons indicated consistent involvement in pathways such as ECM–receptor interaction, oxidative phosphorylation, glycosaminoglycan binding, complement and coagulation cascades, protein digestion and absorption, phenylalanine metabolism, tyrosine metabolism, maturity-onset diabetes of the young, arginine biosynthesis, and isoquinoline alkaloid biosynthesis (Figure 5F). These findings collectively underscore the interplay between transcriptomic regulation, metabolic activity, and tissue-specific functional remodeling.

 doi:10.20944/preprints202508.1328.v1

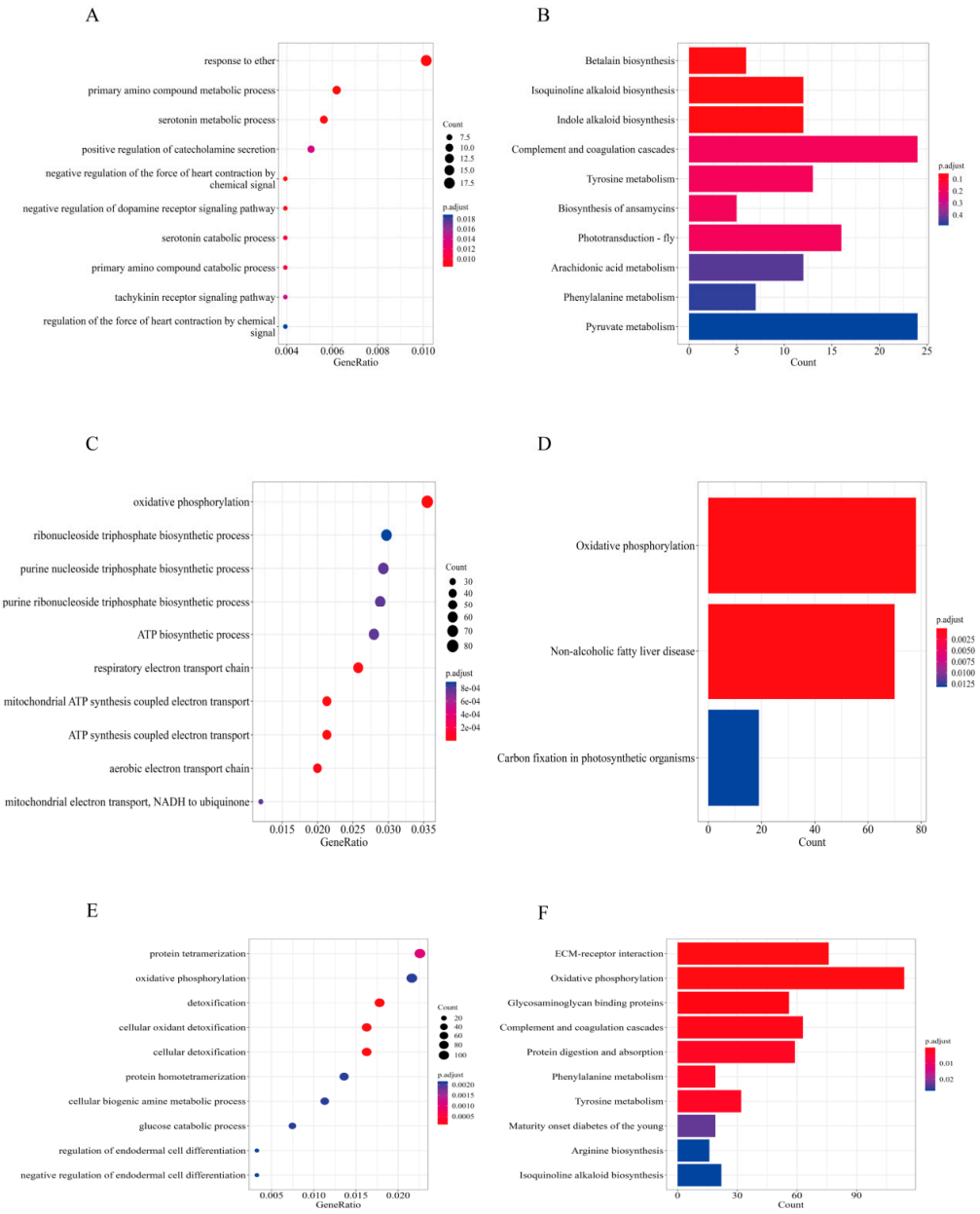

Figure 5. Results of GO and KEGG analyses. (A) The top ten significant gene ontology (GO) terms obtained from GO enrichment analysis using genes with DAS events or DETs in UM vs M1. The Y-axis represents different GO term categories. The X-axis represents the proportion of significantly expressed genes in the list of corresponding GO terms (GeneRatio). Different sizes and colors of circles represent the number of significantly expressed genes and the corresponding adjusted p-value of GO terms. (B) The top ten significant pathways obtained from KEGG enrichment analysis using genes with DAS events or DEGs in UM vs M1. The Y axis represents different pathway categories. The X-axis represents the number of significantly expressed genes in the corresponding pathway. Different colors represent the different adjusted p-values of the pathway. (C) The top ten significant gene ontology (GO) terms obtained from GO enrichment analysis using genes with DAS events or DETs in UM vs M3. (D) The top ten significant pathways obtained from KEGG enrichment analysis using genes with DAS events or DEGs in UM vs M3. (E) The top ten significant gene ontology (GO) terms obtained from GO enrichment analysis using genes with DAS events or DETs in six groups. (F) The top ten significant pathways obtained from KEGG enrichment analysis using genes with DAS events or DEGs in six groups.

 doi:10.20944/preprints202508.1328.v1

3.7. Validation of Significant Differential Expression Transcripts

To ensure the accuracy of the transcriptome profiling results, we conducted quantitative real-time PCR (qRT-PCR) validation on a subset of differentially expressed genes. Five genes—PB.10563, PB.6079, PB.10649, PB.2978, and PB.3758—were randomly selected for this purpose. The expression patterns obtained from qRT-PCR closely paralleled those from the RNA-seq analysis across the different group comparisons (Figure 6). Specifically, both approaches confirmed that PB.10563, PB.6079, PB.10649, and PB.2978 were significantly upregulated, while PB.3758 was consistently downregulated. These findings confirm the reliability and consistency of the transcriptomic data used in this study.

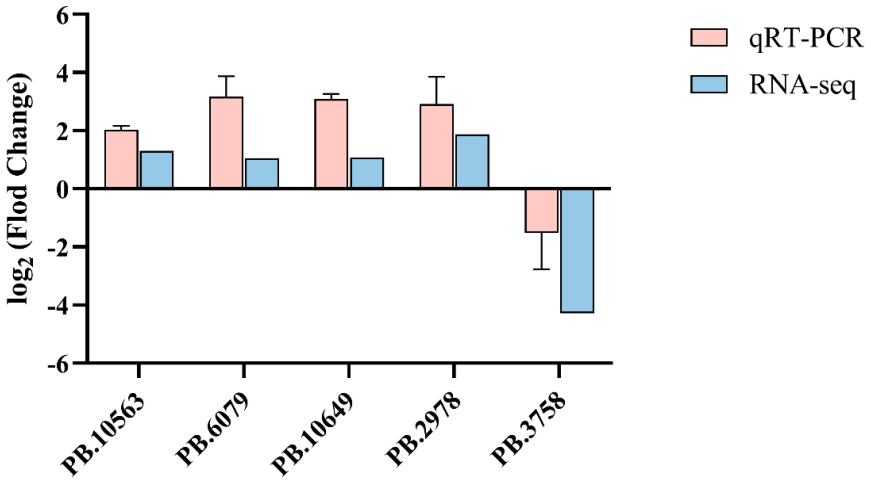

Figure 6. Validation of significantly differentially expressed genes. The Y-axis is $\log_2$ Flod Change, X-axis is different genes.

4. Discussion

Pubertal molting marks a critical transition in the reproductive cycle of crustaceans, signaling the onset of ovarian development. In the mud crabs, mating is known to accelerate this process dramatically [31,32,47]. Yet, despite its biological importance, the molecular landscape underlying this transition has remained largely unexplored. In this study, we applied full-length transcriptome sequencing to comprehensively map the changes in gene expression and alternative splicing (AS) following mating. Our results offer new insights into the transcriptional profile and signaling pathways that drive post-mating ovarian maturation.

A high-quality transcriptome is essential for investigating dynamic developmental processes like oogenesis. The BUSCO completeness of the reconstructed transcriptome was 25.8%, which was significantly lower than that of the reference genome [35]. This difference likely reflects the tissue-specific and temporally restricted nature of ovarian gene expression [48,49]. Encouragingly, the proportion of fragmented transcripts in our dataset was only 1.9%, which is lower than a previous SMRT-sequencing-based transcriptome (3.2%) [50], highlighting the reliability and accuracy of the sequencing and data processing. The reconstructed transcriptome also revealed substantial isoform diversity, including over 26,000 transcripts not previously annotated—many of which may be novel regulators of ovarian development.

Alternative splicing analysis revealed extensive transcript diversity, with the most frequent AS events being alternative first exon (AF) and alternative 5′ splice site (A5). These splicing patterns were enriched in genes associated with chromatin remodeling, protein regulation, and metabolic processes, suggesting a shift in cellular priorities to support oocyte maturation. The presence of such complex isoform architecture implies that AS plays a fundamental role in fine-tuning gene function during ovarian development [51].

 doi:10.20944/preprints202508.1328.v1

Through differential expression and AS analyses, we observed significant transcriptional shifts across developmental stages. The highest number of differentially expressed transcripts (1,963) was found in the UM vs M3 comparison, highlighting how mating triggers a profound molecular reprogramming in the ovary. Several upregulated genes stood out due to their functional importance. For instance, glutathione S-transferase (GST) may reflect increased antioxidant defense during rapid tissue growth [52], ATP synthase supports high energy demands during oocyte proliferation [53,54], and LPIN2-like is likely involved in lipid redistribution and membrane biosynthesis essential for yolk accumulation [53,55,56].

Pathway enrichment analyses offered even deeper insights. In the UM vs M1 comparison, GO enrichment pointed to the activation of neurochemical signaling pathways, especially those involving serotonin and catecholamines. These amino acid-derived compounds are known to regulate reproductive hormones in crustaceans and fish alike [57]. For example, serotonin stimulates gonadotropin release in the thoracic ganglia [58], while tyrosine-derived metabolites like dopamine and N-acetylserotonin fluctuate with ovarian stage in Chinese sturgeon, peaking during active vitellogenesis [59]. The observed enrichment of these pathways in our data suggests that serotonin and catecholamine signaling may be among the earliest molecular responses to mating, serving as upstream activators of reproductive gene networks [60].

In contrast, the UM vs M3 comparison highlighted enrichment of energy metabolism-related pathways, including oxidative phosphorylation, ATP biosynthesis, and electron transport chain activity. These pathways reflect the high energy demands associated with vitellogenesis and oocyte growth [61]. Supporting this, KEGG analysis revealed significant enrichment in pathways like pyruvate metabolism, phenylalanine metabolism, and even non-alcoholic fatty liver disease—systems that are increasingly recognized for their roles in nutrient sensing and energy regulation during reproduction [62]. This post-mating shift toward bioenergetic activation is consistent with observations in *Portunus trituberculatus*, where energy metabolism genes are downregulated in later stages, likely indicating a transition from energy storage to energy utilization [63].

Structural remodeling of the ovary also emerged as a central theme. The extracellular matrix (ECM)-receptor interaction pathway, enriched in our analysis, is critical for follicular development and oocyte differentiation [64]. In crustaceans and other animals, ECM components such as laminin and collagen interact with integrin receptors to mediate key processes like steroidogenesis, cell migration, and structural integrity [65,66]. These interactions activate intracellular pathways—including TGF-β and PI3K-Akt—that orchestrate follicular growth and survival [67,68]. Given the enrichment of these pathways in our dataset, it's likely that ECM remodeling contributes to the physical restructuring of the ovary required for oocyte maturation. We also observed enrichment in the HIF-1 and IL-17 signaling pathways. HIF-1 is known to regulate oxygen availability in growing tissues and has been implicated in modulating oocyte developmental timing and blastocyst viability [69,70]. The IL-17 pathway, though traditionally associated with immune function, has recently been linked to follicular maturation in eels [71,72], and may serve a similar role in *S. paramamosain* by regulating ovarian inflammation and tissue remodeling during maturation.

Lipids are vital during ovarian development, not just as structural components but as key energy reserves for embryos and larvae. In *Scylla olivacea*, fatty acids such as palmitic acid increase significantly during vitellogenesis [73], supporting the hypothesis that lipid mobilization supports later-stage oocyte development. In our study, enrichment of pathways like arachidonic acid metabolism and nitrogen metabolism suggests that similar mechanisms are at play in *S. paramamosain*. The upregulation of LPIN2-like, known to regulate phospholipid and triglyceride synthesis, further supports the role of lipid remodeling during this critical phase.

Finally, oxidative stress management appears essential for successful oogenesis. While reactive oxygen species (ROS) can serve as signaling molecules that promote oocyte maturation [74,75], their overaccumulation can cause cellular damage. Our data revealed strong enrichment in GO terms related to cellular oxidant detoxification, with genes such as SOD2, CAT, GPX4, and GSTO1 being significantly upregulated. These enzymes form the core of the antioxidant defense system and likely

 doi:10.20944/preprints202508.1328.v1

help maintain redox balance in the rapidly developing ovarian tissue [76,77]. Reduced antioxidant expression in other species has been linked to disrupted ROS balance and altered oocyte development, suggesting that proper ROS regulation is a conserved feature of reproductive biology.

5. Conclusions

This study shows that SMRT RNA-seq provides a powerful and accurate method to explore the complex process of ovarian development in *S. paramamosain*. By generating a high-quality, full-length transcriptome, we identified 1,963 differentially expressed transcripts and 34 alternative splicing events between unmated and post-mating stages, revealing major changes in gene activity during ovarian maturation. Consistent patterns across all six group comparisons pointed to key roles for energy metabolism, lipid processing, extracellular matrix signaling, and oxidative stress regulation. In particular, genes linked to oxidative phosphorylation, serotonin and catecholamine signaling, and detoxification pathways were significantly enriched, suggesting a coordinated effort to support the energy needs, structural changes, and protection required for oocyte growth and yolk formation. Altogether, these findings provide valuable insights into the molecular mechanism guiding crustacean reproduction. This work also offers a rich genomic resource for future studies aimed at improving broodstock management and reproductive control in aquaculture.

Author Contributions: Conceptualization, M.I. and H.M.; methodology, C.W., S.M.A., X.Z., Y.Y., M.I., and H.M.; validation, C.W., and H.M.; formal analysis, C.W., S.M.A., and W.W.,; investigation, C.W., S.M.A., X.Z., Y.Y., W.W., and H.M.; resources, H.M.; data curation, X.Z.; writing—original draft preparation, C.W., and W.W.; writing—review and editing, M.I., and H.M.; visualization, W.W.; supervision, H.M.; project administration, H.M.; funding acquisition, H.M.. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (32473150), and the Program of Agricultural and Rural Department of Guangdong Province (2024-SPY-00-013).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: We express our gratitude to Dr. Noah Esmaeili for his valuable input in enhancing the linguistic quality of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Sanda, T.; Shimizu, T.; Iwasaki, T.; Dan, S.; Hamasaki, K. Effect of Temperature on Survival, Intermolt Period, and Growth of Juveniles of Two Mud Crab Species, *Scylla Paramamosain* and *Scylla Serrata* (Decapoda: Brachyura: Portunidae), under Laboratory Conditions. *Nauplius* **2022**, *30*, e2022012, doi:10.1590/2358-2936e2022012.
2. Zhao, Y.; Waqas, W.; Cui, W.; Ye, S.; Gao, W.; Zhang, Q.; Lin, Z.; Zhu, D.; Lin, F.; Ikhwanuddin, M. Comparative Analysis of Embryonic Development and Growth Performance among Two Mud Crab Species and Their Hybrids. *Aquaculture* **2025**, *596*, 741795, doi:10.1016/j.aquaculture.2024.741795.
3. Ye, H.; Tao, Y.; Wang, G.; Lin, Q.; Chen, X.; Li, S. Experimental Nursery Culture of the Mud Crab *Scylla Paramamosain* (Estampador) in China. *Aquac. Int.* **2011**, *19*, 313–321, doi:10.1007/s10499-010-9399-3.
4. Nghia, T.T.; Wille, M.; Binh, T.C.; Thanh, H.P.; Van Danh, N.; Sorgeloos, P. Improved Techniques for Rearing Mud Crab *Scylla Paramamosain* (Estampador 1949) Larvae. *Aquac. Res.* **2007**, *38*, 1539–1553, doi:10.1111/j.1365-2109.2007.01814.x.
5. Fishery Bureau of Ministry of Agriculture of China. China Fishery Statistical Yearbook 2024. Chinese Agricultural Press, Beijing, June, **2024**.

 doi:10.20944/preprints202508.1328.v1

6. Quinitio, E.T.; De Pedro, J.; Parado-Estepa, F.D. Ovarian Maturation Stages of the Mud Crab *Scylla Serrata*. *Aquac. Res*. **2007**; *38*, 1434–1441, doi:10.1111/j.1365-2109.2007.01650.x.
7. Christy, J.H. Timing of Hatching and Release of Larvae by Brachyuran Crabs: Patterns, Adaptive Significance and Control. *Integr. and Comp. Biol.* **2011**, *51*, 62–72, doi:10.1093/icb/icr013.
8. Waiho, K.; Fazhan, H.; Shahreza, M.S.; Moh, J.H.Z.; Noorbaiduri, S.; Wong, L.L.; Sinnasamy, S.; Ikhwanuddin, M. Transcriptome Analysis and Differential Gene Expression on the Testis of Orange Mud Crab, *Scylla Olivacea*, during Sexual Maturation. *PLoS One* **2017**, *12*, e0171095, doi:10.1371/journal.pone.0171095.
9. Green, B.S.; Gardner, C.; Hochmuth, J.D.; Linnane, A. Environmental Effects on Fished Lobsters and Crabs. *Rev. Fish Biol. Fish.* **2014**, *24*, 613–638, doi:10.1007/s11160-014-9350-1.
10. Esmaeili, N.; Ma, H.; Kadri, S.; Tocher, D.R. Protein and Lipid Nutrition in Crabs. *Rev. Aquac.* **2024**, *16*, 1499-1519, doi:10.1111/raq.12908.
11. Colpo, K.D.; López-Greco, L.S. Dynamics of Energy Reserves and the Cost of Reproduction in Female and Male Fiddler Crabs. *Zoology* **2018**, *126*, 11–19, doi:10.1016/j.zool.2018.01.004.
12. Usman, M.; Li, A.; Wu, D.; Qinyan, Y.; Yi, L.X.; He, G.; Lu, H. The Functional Role of lncRNAs as ceRNAs in Both Ovarian Processes and Associated Diseases. *Non-coding RNA Res.* **2024**, *9*, 165–177, doi:10.1016/j.ncrna.2023.11.008.
13. Bao, C.; Yang, Y.; Huang, H.; Ye, H. Inhibitory Role of the Mud Crab Short Neuropeptide F in Vitellogenesis and Oocyte Maturation via Autocrine/paracrine Signaling. *Front. Endocrinol* **2018**, *9*, 390, doi:10.3389/fendo.2018.00390.
14. Islam, M.S.; Kodama, K.; Kurokora, H. Ovarian Development of the Mud Crab *Scylla Paramamosain* in a Tropical Mangrove Swamps, Thailand. *J. Sci. Res.* **2010**, *2*, 380–389, doi:10.3329/jsr.v2i2.3543.
15. Ye, H.; Song, P.; Ma, J.; Huang, H.; Wang, G. Changes in Progesterone Levels and Distribution of Progesterone Receptor during Vitellogenesis in the Female Mud Crab (*Scylla Paramamosain*). *MAR FRESHW BEHAV PHY* **2010**, *43*, 25–35, doi:10.1080/10236241003654113.
16. Sharifian, S.; Kamrani, E.; Safaie, M.; Sharifian, S. Oogenesis and Ovarian Development in the Freshwater Crab *Sodhiana Iranica* (Decapoda: Gecarcinuaidae) from the South of Iran. *TISSUE CELL* **2015**, *47*, 213–220, doi:10.1016/j.tice.2014.11.006.
17. Fazhan, H.; Waiho, K.; Wan Norfaizza, W.I.; Megat, F.H.; Ikhwanuddin, M. Inter-Species Mating among Mud Crab Genus *Scylla* in Captivity. *Aquaculture* **2017**, *471*, 49–54, doi:10.1016/j.aquaculture.2017.01.005.
18. Koolkalya, S.; Thapanand, T.; Tunkijjanujij, S.; Havanont, V.; Jutagate, T. Aspects in Spawning Biology and Migration of the Mud Crab *Scylla Olivacea* in the Andaman Sea, Thailand. *FISHERIES MANAG ECOL* **2006**, *13*, 391–397, doi:10.1111/j.1365-2400.2006.00518.x.
19. McLay, C.L.; López Greco, L.S. A Hypothesis about the Origin of Sperm Storage in the Eubrachyura, the Effects of Seminal Receptacle Structure on Mating Strategies and the Evolution of Crab Diversity: How Did a Race to Be First Become a Race to Be Last? *Zool. Anz.* **2011**, *250*, 378–406, doi:10.1016/j.jcz.2011.06.002.
20. McLay, C.L.; Becker, C. Reproduction in Brachyura. *Treatise on Zoology - Anatomy, Taxonomy, Biology. The Crustacea,* **2015**, *9*, 185–243.
21. Kulski, J.K. Next-Generation Sequencing — An Overview of the History, Tools, and "Omic" Applications. *Next Generation Sequencing - Advances, Applications and Challenges* **2016, 3-60**.
22. Satam, H.; Joshi, K.; Mangrolia, U.; Waghoo, S.; Zaidi, G.; Rawool, S.; Thakare, R.P.; Banday, S.; Mishra, A.K.; Das, G.; et al. Next-Generation Sequencing Technology: Current Trends and Advancements. *Biology* **2023**, *12*, 997, doi:10.3390/biology12070997.

 doi:10.20944/preprints202508.1328.v1

23. Nguyen, T.V.; Jung, H.; Rotllant, G.; Hurwood, D.; Mather, P.; Ventura, T. Guidelines for RNA-Seq Projects: Applications and Opportunities in Non-Model Decapod Crustacean Species. *Hydrobiologia* **2018**, *825*, 5–27, doi:10.1007/s10750-018-3682-0.
24. Wang, Z.; Gerstein, M.; Snyder, M. RNA-Seq: A Revolutionary Tool for Transcriptomics. *Nat. Rev. Genet.* **2009**, *10*, 57–63, doi:10.1038/nrg2484.
25. Mykles, D.L.; Burnett, K.G.; Durica, D.S.; Joyce, B.L.; McCarthy, F.M.; Schmidt, C.J.; Stillman, J.H. Resources and Recommendations for Using Transcriptomics to Address Grand Challenges in Comparative Biology. *Integr. Comp. Biol.* **2016**; *56*, 1183–1191, doi:10.1093/icb/icw083.
26. Pop, M.; Salzberg, S.L. Bioinformatics Challenges of New Sequencing Technology. *TRENDS GENET* **2008**, *24*, 142–149, doi:10.1016/j.tig.2007.12.006.
27. Wenger, A.M.; Peluso, P.; Rowell, W.J.; Chang, P.C.; Hall, R.J.; Concepcion, G.T.; Ebler, J.; Fungtammasan, A.; Kolesnikov, A.; Olson, N.D.; et al. Accurate Circular Consensus Long-Read Sequencing Improves Variant Detection and Assembly of a Human Genome. *Nat. Biotechnol.* **2019**, *37*, 1155–1162, doi:10.1038/s41587-019-0217-9.
28. Zhang, D.; Li, W.; Chen, Z. jian; Wei, F. gang; Liu, Y. long; Gao, L.Z. SMRT- and Illumina-Based RNA-Seq Analyses Unveil the Ginsinoside Biosynthesis and Transcriptomic Complexity in Panax Notoginseng. *Sci. Rep.* **2020**, *10*, 15310, doi:10.1038/s41598-020-72291-1.
29. Ye, S.; Yu, X.; Chen, H.; Zhang, Y.; Wu, Q.; Tan, H.; Song, J.; Saqib, H.S.A.; Farhadi, A.; Ikhwanuddin, M.; et al. Full-Length Transcriptome Reconstruction Reveals the Genetic Mechanisms of Eyestalk Displacement and Its Potential Implications on the Interspecific Hybrid Crab (Scylla Serrata ♀ × S. Paramamosain ♂). *Biology.* **2022**, *11*, 1026, doi:10.3390/biology11071026.
30. Byrne, A.; Cole, C.; Volden, R.; Vollmers, C. Realizing the Potential of Full-Length Transcriptome Sequencing. *PHILOS T R. SOC B* **2019**, *374*, *20190097*, doi:10.1098/rstb.2019.0097.
31. Yu, Y.; Zhang, M.; Wang, D.; Xiang, Z.; Zhao, Z.; Cui, W.; Ye, S.; Fazhan, H.; Waiho, K.; Ikhwanuddin, M. Whole Transcriptome RNA Sequencing Provides Novel Insights into the Molecular Dynamics of Ovarian Development in Mud Crab, *Scylla Paramamosain* after Mating. *Comp. Biochem. Physiol. Part D Genomics Proteomics* **2024**, *51*, 101247, doi:10.1016/j.cbd.2024.101247.
32. Li, W.; Li, S.; Wang, X.; Chen, H. yun; Hao, H.; Wang, K.J. Internal Carbohydrates and Lipids as Reserved Energy Supply in the Pubertal Molt of *Scylla Paramamosain*. *Aquaculture* **2022**, *549*, 737736, doi:10.1016/j.aquaculture.2021.737736.
33. Farhadi, A.; Shi, X.; Zhang, Y.; Zhang, Y.; Li, S.; Zheng, H.; Ikhwanuddin, M.; Ma, H. A Novel Imprinted Gene (Sp-Pol) With Sex-Specific SNP Locus and Sex-Biased Expression Pattern Provides Insights Into the Gonad Development of Mud Crab (*Scylla Paramamosain*). *Front. Mar. Sci.* **2021**, *8*, 727607, doi:10.3389/fmars.2021.727607.
34. Wang, D.; Yu, Y.; Gao, W.; Xiang, Z.; Zhao, Z.; Fazhan, H.; Waiho, K.; Ikhwanuddin, M.; Ma, H. Dynamic Changes Characteristics of the Spermatozoon during the Reproductive Process of Mud Crab (*Scylla Paramamosain*): From Spermatophore Formation, Transportation to Dispersion. *Aquac. Reports* **2023**, *33*, 101866, doi:10.1016/j.aqrep.2023.101866.
35. Zhang, Y.; Yuan, Y.; Zhang, M.; Yu, X.; Qiu, B.; Wu, F.; Tocher, D.; Zhang, J.; Ye, S.; Cui, W. et al. High-resolution chromosome-level genome of *Scylla paramamosain* provides molecular insights into adaptive evolution in crabs. *BMC Biol.* **2024**, *22*, 255, doi:10.1186/s12915-024-02054-1.
36. Li, H. Minimap2: Pairwise Alignment for Nucleotide Sequences. *Bioinformatics* **2018**, *34*, 3094–3100, doi:10.1093/bioinformatics/bty191.

 doi:10.20944/preprints202508.1328.v1

37. Huang, Y.; Niu, B.; Gao, Y.; Fu, L.; Li, W. CD-HIT Suite: A Web Server for Clustering and Comparing Biological Sequences. *Bioinformatics* **2010**, *26*, 680–682, doi:10.1093/bioinformatics/btq003.
38. Update, B. Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. *Mol Biol Evol* **2021**, *38*, 10–1093, doi:10.1093/molbev/msab199.
39. Kriventseva, E. V.; Kuznetsov, D.; Tegenfeldt, F.; Manni, M.; Dias, R.; Simão, F.A.; Zdobnov, E.M. OrthoDB v10: Sampling the Diversity of Animal, Plant, Fungal, Protist, Bacterial and Viral Genomes for Evolutionary and Functional Annotations of Orthologs. *Nucleic Acids Res.* **2019**, *47*, D807–D811, doi:10.1093/nar/gky1053.
40. Pertea, G.; Pertea, M. GFF Utilities: GffRead and GffCompare. *F1000Res* **2020**, *9*, 304, doi:10.12688/f1000research.23297.1.
41. Huerta-Cepas, J.; Forslund, K.; Coelho, L.P.; Szklarczyk, D.; Jensen, L.J.; Von Mering, C.; Bork, P. Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. *Mol. Biol. Evol.* **2017**, *34*, 2115–2122, doi:10.1093/molbev/msx148.
42. Trincado, J.L.; Entizne, J.C.; Hysenaj, G.; Singh, B.; Skalic, M.; Elliott, D.J.; Eyras, E. SUPPA2: Fast, Accurate, and Uncertainty-Aware Differential Splicing Analysis across Multiple Conditions. *Genome Biol.* **2018**, *19*, 40, doi:10.1186/s13059-018-1417-1.
43. Patro, R.; Duggal, G.; Love, M.I.; Irizarry, R.A.; Kingsford, C. Salmon Provides Fast and Bias-Aware Quantification of Transcript Expression. *Nat. Methods* **2017**, *14*, 417–419, doi:10.1038/nmeth.4197.
44. Love, M.I.; Huber, W.; Anders, S. Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. *Genome Biol.* **2014**, *15*, 550, doi:10.1186/s13059-014-0550-8.
45. Yu, G.; Wang, L.G.; Han, Y.; He, Q.Y. ClusterProfiler: An R Package for Comparing Biological Themes among Gene Clusters. *Omi. A J. Integr. Biol.* **2012**, *16*, 284–287, doi:10.1089/omi.2011.0118.
46. Schmittgen, T.D.; Livak, K.J. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **2001**, *25*, 402–408, doi:10.1006/meth.2001.1262.
47. Long, X.; Guo, Q.; Wang, X.; Francis, D.S.; Cheng, Y.; Wu, X. Effects of Fattening Period on Ovarian Development and Nutritional Quality of Adult Female Chinese Mitten Crab *Eriocheir Sinensis*. *Aquaculture* **2020**, *519*, 734748, doi:10.1016/j.aquaculture.2019.734748.
48. Johnson, B.R.; Atallah, J.; Plachetzki, D.C. The Importance of Tissue Specificity for RNA-Seq: Highlighting the Errors of Composite Structure Extractions. *BMC Genomics* **2013**, *14*, 586, doi:10.1186/1471-2164-14-586.
49. Naumova, O.Y.; Lee, M.; Rychkov, S.Y.; Vlasova, N. V.; Grigorenko, E.L. Gene Expression in the Human Brain: The Current State of the Study of Specificity and Spatiotemporal Dynamics. *Child Dev.* **2013**, *84*, 76–88, doi:10.1111/cdev.12014.
50. Wan, H.; Jia, X.; Zou, P.; Zhang, Z.; Wang, Y. The Single-Molecule Long-Read Sequencing of *Scylla Paramamosain*. *Sci. Rep.* **2019**, *9*, 12401, doi:10.1038/s41598-019-48824-8.
51. Dioken, D.N.; Ozgul, I.; Erson-Bensan, A.E. The 3′ End of the Tale—neglected Isoforms in Cancer. *FEBS Lett.* **2025**, doi:10.1002/1873-3468.70122.
52. Ma, M.; Zhang, Y.X.; Chen, D.; Smagghe, G.; Wang, J.J.; Wei, D. Functional Characterization of a Glutathione S-Transferase Gene GSTe10 That Contributes to Ovarian Development in *Bactrocera Dorsalis* (Hendel). *Entomol. Gen.* **2022**, *42*, 539–547, doi:10.1127/entomologia/2022/1363.
53. Wang, T.; Wang, T.; Zhang, M.; Shi, X.; Zhang, M.; Wang, H.; Yang, X.; Yu, Z.; Liu, J. The Ovarian Development Genes of Bisexual and Parthenogenetic Haemaphysalis Longicornis Evaluated by Transcriptomics and Proteomics. *Front. Vet. Sci.* **2021**, *8*, 783404, doi:10.3389/fvets.2021.783404.

 doi:10.20944/preprints202508.1328.v1

54. Moreira, H.N.S.; Barcelos, R.M.; Vidigal, P.M.P.; Klein, R.C.; Montandon, C.E.; Maciel, T.E.F.; Carrizo, J.F.A.; Costa de Lima, P.H.; Soares, A.C.; Martins, M.M.; et al. A Deep Insight into the Whole Transcriptome of Midguts, Ovaries and Salivary Glands of the *Amblyomma Sculptum* Tick. *Parasitol. Int.* **2017**, *66*, 64–73, doi:10.1016/j.parint.2016.10.011.
55. Trapp, J.; Almunia, C.; Gaillard, J.C.; Pible, O.; Chaumot, A.; Geffard, O.; Armengaud, J. Proteogenomic Insights into the Core-Proteome of Female Reproductive Tissues from Crustacean Amphipods. *J. Proteomics* **2016**, *135*, 51–61, doi:10.1016/j.jprot.2015.06.017.
56. Yepiz-Plascencia, G.; Vargas-Albores, F.; Higuera-Ciapara, I. Penaeid Shrimp Hemolymph Lipoproteins. *Aquaculture* **2000**, *191*, 177–189, doi:10.1016/S0044-8486(00)00427-0.
57. Barcellos, L.J.G.; Volpato, G.L.; Barreto, R.E.; Coldebella, I.; Ferreira, D. Chemical Communication of Handling Stress in Fish. *Physiol. Behav.* **2011**, *103*, 372–375, doi:10.1016/j.physbeh.2011.03.009.
58. Kulkarni, G.K.; Nagabhushanam, R.; Amaldoss, G.; Jaiswal, R.G.; Fingerman, M. In Vivo Stimulation of Ovarian Development in the Red Swamp Crayfish, *Procambarus Clarkii* (Girard), by 5-Hydroxytryptamine. *Invertebr. Reprod. Dev.* **1992**, *21*, 231–240, doi:10.1080/07924259.1992.9672242.
59. Zhu, Y.; Wu, J.; Leng, X.; Du, H.; Wu, J.; He, S.; Luo, J.; Liang, X.; Liu, H.; Wei, Q.; et al. Metabolomics and Gene Expressions Revealed the Metabolic Changes of Lipid and Amino Acids and the Related Energetic Mechanism in Response to Ovary Development of Chinese Sturgeon (*Acipenser Sinensis*). *PLoS ONE* **2020**, *15*, e0235043, doi:10.1371/journal.pone.0235043.
60. Esmaeili, N.; Kadri, S.; Kumar, V.; Ma, H. Serotonin: A Multifunctional Molecule in Crustaceans. *Rev. Aquac.* **2025**, *17*, e70046, https://doi.org/10.1111/raq.70046.
61. Kleppe, L.; Edvardsen, R.B.; Furmanek, T.; Taranger, G.L.; Wargelius, A. Global Transcriptome Analysis Identifies Regulated Transcripts and Pathways Activated during Oogenesis and Early Embryogenesis in Atlantic Cod. *Mol. Reprod. Dev.* **2014**, *81*, 619–635, doi:10.1002/mrd.22328.
62. Inigo, M.; Deja, S.; Burgess, S.C. Ins and Outs of the TCA Cycle: The Central Role of Anaplerosis. *Annu. Rev. Nutr.* 2021, *41*, 19–47, doi:10.1146/annurev-nutr-120420-025558.
63. Zhang, Y.; Chen, Y.; Hou, C.; Wang, C.; Mu, C. Analysis of cDNA microarrays revealed the effects of mating on the ovary and hepatopancreas of female swimming crab (*Portunus trituberculatus*) during the late stage of ovarian develoment. *COMP BIOCHEM PHYS D* **2025**, *55*, 101520, doi: org/10.1016/j.cbd.2025.101520.
64. Nagyová, E.; Němcová, L.; Camaioni, A. Cumulus Extracellular Matrix Is an Important Part of Oocyte Microenvironment in Ovarian Follicles: Its Remodeling and Proteolytic Degradation. *Int. J. Mol. Sci.* 2022, *23*, 54, doi:10.3390/ijms23010054.
65. Curry, T.E.; Osteen, K.G. The Matrix Metalloproteinase System: Changes, Regulation, and Impact throughout the Ovarian and Uterine Reproductive Cycle. *Endocr. Rev.* 2003, *24*, 428–465, doi:10.1210/er.2002-0005.
66. Dong, Y.; Lyu, L.; Zhang, D.; Li, J.; Wen, H.; Shi, B. Integrated lncRNA and mRNA Transcriptome Analyses in the Ovary of Cynoglossus Semilaevis Reveal Genes and Pathways Potentially Involved in Reproduction. *Front. Genet.* **2021**, *12*, 671729, doi:10.3389/fgene.2021.671729.
67. Kreeger, P.K.; Deck, J.W.; Woodruff, T.K.; Shea, L.D. The in Vitro Regulation of Ovarian Follicle Development Using Alginate-Extracellular Matrix Gels. *Biomaterials* **2006**, *27*, 714–723, doi:10.1016/j.biomaterials.2005.06.016.
68. Li, Z.; Tan, S.; Qi, L.; Chen, Y.; Liu, H.; Liu, X.; Sha, Z. Genome-Wide Characterization of Integrin (ITG) Gene Family and Their Expression Profiling in Half-Smooth Tongue Sole (*Cynoglossus Semilaevis*) upon *Vibrio Anguillarum* Infection. *COMP BIOCHEM PHYS D* **2023**, *47*, 101099, doi:10.1016/j.cbd.2023.101099.

 doi:10.20944/preprints202508.1328.v1

69. Song, H.; Zhu, B.; Dong, T.; Wang, W.; Hu, M.; Yan, X.; Xu, S.; Hu, H. Whole-Genome Resequencing Reveals Selection Signatures for Caviar Yield in Russian Sturgeon (*Acipenser Gueldenstaedtii*). *Aquaculture* **2023**, *568*, 739312, doi:10.1016/j.aquaculture.2023.739312.
70. Turhan, A.; Pereira, M.T.; Schuler, G.; Bleul, U.; Kowalewski, M.P. Hypoxia-Inducible Factor (HIF1alpha) Inhibition Modulates Cumulus Cell Function and Affects Bovine Oocyte Maturation in Vitro. *Biol. Reprod.* **2021**, *104*, 479–491, doi:10.1093/biolre/ioaa196.
71. Lai, X.; Peng, S.; Feng, J.; Zou, P.; Wang, Y. Immune Function Modulation during Artificial Ovarian Maturation in Japanese Eel (*Anguilla Japonica*): A Transcriptome Profiling Approach. *Fish Shellfish Immunol.* **2022**, *131*, 662–671, doi:10.1016/j.fsi.2022.10.048.
72. Lai, X.J.; Peng, S.; Wang, Y.L. Dynamic Transcriptome Analysis of Ovarian Follicles in Artificial Maturing Japanese Eel (*Anguilla Japonica*). *Theriogenology* **2022**, *180*, 176–188, doi:10.1016/j.theriogenology.2021.12.031.
73. Ghazali, A.; Azra, M.N.; Noordin, N.M.; Abol-Munafi, A.B.; Ikhwanuddin, M. Ovarian Morphological Development and Fatty Acids Profile of Mud Crab (*Scylla Olivacea*) Fed with Various Diets. *Aquaculture* **2017**, *468*, 45–52, doi:10.1016/j.aquaculture.2016.09.038.
74. Arenas-Ríos, E.; León-Galván, M.A.; Mercado, P.E.; López-Wilchis, R.; Cervantes, D.L.M.I.; Rosado, A. Superoxide 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* **2007**, *148*, 150–158, doi:10.1016/j.cbpa.2007.04.003.
75. Behrman, H.R.; Kodaman, P.H.; Preston, S.L.; Gao, S. Oxidative Stress and the Ovary. *J. Soc. Gynecol. Investig.* **2001**, *8*, S40–S42, doi:10.1016/S1071-5576(00)00106-4.
76. Lu, J.; Wang, Z.; Cao, J.; Chen, Y.; Dong, Y. A Novel and Compact Review on the Role of Oxidative Stress in Female Reproduction. *Reprod. Biol. Endocrinol.* 2018, *16*, 80, doi:10.1186/s12958-018-0391-5.
77. Zhang, Y.; Li, Y.D.; Yang, Q. Bin; Jiang, S.; Jiang, S.G.; Zhou, F.L. Transcriptome Analysis of *Metapenaeus Affinis* Reveals Genes Involved in Gonadal Development. *ISR J AQUACULT - BAMID* **2022**, *74*, 15, doi:10.46989/001c.37738.