Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Transcriptomic profiling of long non-coding RNAs and circular RNAs during pupal development and early adulthood in tissues containing mycangial primordia of the ambrosia beetle Euwallacea interjectus

Mingzhen Zhao[1]
Huzhou Normal University
Mengjia Shi
Nanjing Forestry University
Weiyi Pan[2]
Nanjing Forestry University
Lulu Dai[2]
dailulu@njfu.edu.cn

Nanjing Forestry University

Research Article

Keywords: Ambrosia beetle, Euwallacea interjectus, Long non-coding RNA, Circular RNA, Mycangium, Pupal development

Posted Date: November 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11767475/v1

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

Additional Declarations: No competing interests reported.

Transcriptomic profiling of long non-coding RNAs and circular RNAs during pupal development and early adulthood in tissues containing mycangial primordia of the ambrosia beetle *Euwallacea interjectus*

Mingzhen Zhao[1], Mengjia Shi[1,2], Weiyi Pan[2], Lulu Dai[2]*

[1] Zhejiang Key Laboratory of Biology and Ecological Regulation of Crop Pathogens and Insects, College of Life Science, Huzhou University, Huzhou 313000, China

[2] Co-Innovation Center for Sustainable Forestry in Southern China, College of Forestry and Grassland, Nanjing Forestry University, Nanjing, Jiangsu 210037, China

*Correspondence: dailulu@njfu.edu.cn

Abstract

Background Ambrosia beetles maintain an obligate mutualism with fungi through female-specific mycangial organs. Although mycangial development begins at the pupal stage and completes by adult emergence, the non-coding RNA landscape associated with this process remains unexplored. Here we profiled long non-coding RNAs (lncRNAs) and circular RNAs (circRNAs) across four developmental stages of the invasive pest *Euwallacea interjectus*.

Results Using rRNA-depleted strand-specific libraries, we identified 2,692 putative lncRNAs (2,261 high-confidence) and 2,820 circRNA candidates from female head–thorax tissues containing mycangial primordia. lncRNAs exhibited shorter lengths and fewer exons than mRNAs, and showed pronounced stage-associated separation at the mid-pupal stage. circRNA candidates displayed dynamic stage-dependent patterns. Functional enrichment of lncRNA cis -targets and circRNA host genes highlighted cytoskeletal organization, muscle-related processes, and immune signaling; however, muscle-related pathways should not be interpreted as direct evidence of mycangial maturation because pupal samples necessarily included thoracic tissues. Predicted lncRNA–mRNA co-expression networks implicated cell cycle regulation and spliceosome function.

Conclusions This study provides the first non-coding RNA resource for a mycangium-bearing tissue in an ambrosia beetle. The results define stage-associated lncRNA and circRNA expression patterns during pupal–adult transition and offer a foundation for future functional tests of candidate RNAs in mycangial development.

Keywords Ambrosia beetle, *Euwallacea interjectus*, Long non-coding RNA, Circular RNA, Mycangium, Pupal development

Introduction

Ambrosia beetles (Coleoptera: Curculionidae: Scolytinae and Platypodinae) have evolved an obligate mutualism with fungi, relying on cultivated fungal hyphae and spores as their sole nutritional source rather than feeding directly on wood [1]. A key innovation enabling this symbiosis is the mycangium—a specialized symbiotic organ that is female-specific in many ambrosia beetle lineages (including *Euwallacea* species), although both sexes possess mycangia in some other lineages [2]. The mycangium is essential for vertical transmission of mutualistic fungi and for successful colonization of new hosts. In *Euwallacea* ambrosia beetles, two pairs of preoral mycangia have been documented, and their structural development begins at the pupal stage and completes prior to adult emergence, independent of fungal signals [3, 4]. Recent studies in *Euwallacea validus* have begun to uncover the genetic underpinnings of mycangial development, with gene knockdown experiments suggesting that *breathless* (*btl*) and *trachealess* (*trh*) are required for mycangium initiation and that tubulogenesis may have been co-opted for early mycangial development [5]. Despite these advances, the molecular mechanisms governing mycangial morphogenesis remain largely unexplored, particularly regarding the roles of non-coding RNAs.

Euwallacea interjectus (Blandford) (Coleoptera: Curculionidae: Scolytinae) is a destructive and invasive pest of live trees native to Asia, and has now been found in the United States and Argentina [6]. In recent years, it has emerged as a major pest causing extensive mortality in poplar plantations in China [6, 7]. Its damage is mediated by the beetle–fungus complex, where female beetles carry symbiotic *Fusarium* fungi in their mycangia and inoculate them into host trees during tunneling, leading to wilting and death [8]. Like other *Euwallacea* species, the mycangia of *E. interjectus* develop during the pupal stage and are fully formed by adult emergence.

High-throughput sequencing technologies have been increasingly applied to ambrosia beetle research. Transcriptome analyses have been conducted in *Euwallacea* species to investigate developmental regulation and host–symbiont interactions. For instance, Hu et al. [9] performed transcriptome sequencing of *E. interjectus* across different developmental stages and identified differentially expressed unigenes related to development and egg laying, revealing pathways involved in protein metabolism such as neuroactive ligand–receptor interaction, endoplasmic reticulum, and RNA transport. In *Euwallacea validus*, transcriptomic approaches combined with RNA interference have identified key genes required for mycangium initiation [5]. Genomic and transcriptomic analyses of ambrosia beetle complexes (*Euwallacea* sp. near *fornicatus–Fusarium euwallaceae* and *Xyleborus glabratus–Raffaelea lauricola*) have revealed evolutionary signatures associated with mutualism, including positive selection in genes related to nutrient homeostasis, gene expression regulation, and nervous system development [10]. Additionally, transcriptomic profiling of *Xylosandrus germanus* has established reference genes for normalizing gene expression in mycangium-related studies [11]. However, most transcriptomic studies in ambrosia beetles have focused on protein-coding genes, and the roles of non-coding RNAs in their development remain virtually unexplored.

Long non-coding RNAs (lncRNAs) are transcripts longer than 200 nucleotides that lack protein-coding potential but play critical regulatory roles in diverse biological processes, including chromatin remodeling, transcription, and post-transcriptional regulation [12]. In insects, lncRNAs have been implicated in metamorphosis, immunity, reproduction, and behavior. Genome-wide analyses in *Drosophila melanogaster* revealed that a substantial proportion of lncRNAs exhibit stage-specific expression during embryonic and larval development [13], and testis-specific

lncRNAs are essential for spermatogenesis [14]. In *Aedes aegypti*, lncRNA_1317 modulates antiviral immune responses against dengue virus [15]. More recently, lncRNAs have been shown to regulate cuticle formation and chitin metabolism in *Aedes albopictus* [16], and to participate in the behavioral transition from nursing to foraging in honey bees [17]. In *Apis cerana*, lncRNAs are differentially expressed between hygienic behaviors such as uncapping and removal of diseased brood, suggesting their involvement in sensory perception and stimulus response [18]. Furthermore, lncRNAs contribute to insecticide resistance, as demonstrated in *Plutella xylostella* where an antisense lncRNA stabilizes a detoxification gene transcript by blocking miRNA binding [19]. During insect metamorphosis, lncRNAs exhibit dynamic expression patterns, with many being upregulated at the onset of metamorphosis, suggesting their involvement in regulating developmental transitions [20, 21]. In the flesh fly *Sarcophaga peregrina*, a specific lncRNA (SP_lnc5000) has been identified as a potential regulator of metamorphosis by affecting pupal cuticle structure and composition [22]. However, whether lncRNAs regulate the development of specialized symbiotic organs such as the mycangium has not been investigated.

Circular RNAs (circRNAs), a class of covalently closed non-coding RNAs generated by back-splicing, have also emerged as important regulators in insect development and stress responses. They act as microRNA sponges, protein scaffolds, and transcriptional regulators. In *Drosophila*, circRNAs are dynamically expressed during metamorphosis and have been linked to muscle development and neuronal differentiation [23, 24]. In *Bombyx mori*, circRNAs participate in immune responses against viral infection by modulating miRNA availability [25]. In *Aedes albopictus*, a circRNA–miRNA–mRNA network regulates ovarian development [26]. Moreover, circRNAs derived from the juvenile hormone receptor gene *Met* in the migratory locust act

differentially on metamorphosis and reproduction, with one circRNA functioning as a miRNA sponge to sustain vitellogenesis [27]. In the aphid *Aphis citricidus*, the circ-379–miR-3050–DNAH3/RNF8 axis ensures proper wing formation [28]. Whole-transcriptome profiling across developmental stages of *Aedes albopictus* has provided insights into chitin-related non-coding RNA and competing endogenous RNA networks, demonstrating that circRNAs can serve as hub genes to manipulate chitin metabolism and promote growth and development [16]. These examples underscore the versatility of circRNAs in controlling key developmental transitions and organogenesis in insects. Their roles in symbiosis-related organ development, however, remain poorly understood.

In this study, we performed a comprehensive transcriptomic analysis across four key developmental stages—early pupa (EP), mid-pupa (MP), late pupa (LP), and early adult (EA)—to systematically identify and characterize lncRNAs and circRNAs expressed during mycangial development in *E. interjectus*. We compared the structural features of lncRNAs and mRNAs, profiled their global expression patterns, identified differentially expressed transcripts across developmental transitions, and performed functional enrichment analyses to infer their potential roles. Furthermore, we constructed lncRNA–mRNA regulatory networks (cis-, antisense-, and trans-acting) to uncover candidate interactions that may govern mycangial morphogenesis and symbiotic establishment. Our findings provide the first molecular landscape of non-coding RNAs associated with a specialized symbiotic organ in ambrosia beetles and lay a foundation for functional studies and potential pest management strategies targeting the symbiotic link.

Materials and methods

Insect rearing and developmental stage collection

All experimental insects (*Euwallacea interjectus*) were obtained from a laboratory colony maintained on semi-artificial diet at 28 °C. Third-instar larvae were individually transferred to Petri dishes (60 mm diameter) containing potato dextrose agar (PDA) pre-inoculated with the symbiotic fungus *Fusarium populicola* (strain isolated from the same beetle population in Dongtai, Jiangsu, China) and kept at 28 °C in darkness. The subsequent developmental stages were defined based on distinct morphological criteria: **Early pupa (EP)** – freshly pupated individuals with soft, pale-colored integument; **Mid-pupa (MP)** – pupae after 4 days of further development, with wing pads turning greyish; **Late pupa (LP)** – pupae after 1 additional day (total 5 days from EP), with deeply pigmented brownish body; **Newly emerged adult (EA)** – adults collected within 12 h after eclosion, with fully sclerotized cuticle (Figure S1). For each stage, individuals were snap-frozen in liquid nitrogen and stored at −80 °C until dissection.

Euwallacea interjectus exhibits pronounced sexual dimorphism. Males are haploid, occur at an extremely low frequency in natural populations, and are significantly smaller than females, allowing for reliable sex differentiation even at the pupal stage. Because the mycangium is a female-specific organ, only female pupae and adults were selected for this study; individuals with ambiguous sex characteristics were excluded.

The mycangia of *E. interjectus* are preoral cavities located on both sides of the mouthparts, making it impossible to dissect pure mycangial tissue without damaging surrounding structures. Therefore, we collected the head region to capture the mycangial primordia. Because the head and thorax are fused during the pupal stages (EP, MP, and LP), the dissected tissue for pupae comprised the head–thorax region. For the early adult stage (EA), the head was separated from the thorax, and only the head was collected. Because mycangial development initiates at the pupal stage and completes by

adult emergence, we focused on the EP, MP, LP, and EA stages, omitting the larval stage. Three biological replicates were collected for each stage, with each replicate pooled from 10 female individuals.

Tissue dissection and RNA extraction

For pupal stages (EP, MP, LP), the entire head-thorax blocks containing the mycangial primordia were excised under a stereomicroscope in liquid-nitrogen vapor. For adult samples (EA), only the complete heads were dissected. Each biological replicate comprised 15 individuals for pupal stages and 20 individuals for adults. Tissues were pooled into RNase-free tubes and immediately homogenized in liquid nitrogen. Total RNA was extracted using TRIzol Universal reagent (Tiangen, China) following the manufacturer's protocol. Ribosomal RNA (rRNA) was removed using a rRNA depletion kit, rather than poly(A) selection, to retain both polyadenylated and non-polyadenylated transcripts. The remaining RNA was fragmented, and strand-specific cDNA libraries were constructed according to the OmicsMaster standard protocol. The purified double-stranded cDNA underwent end repair, A-tailing, and adapter ligation before PCR amplification. Sequencing was performed on the MGI DNBSEQ-T7 or Illumina NovaSeq X Plus platform (Omicsmaster Biotechnology Co., Guangzhou, China). Because the libraries were rRNA-depleted, they were suitable for detecting circular RNAs (circRNAs) and long non-coding RNAs (lncRNAs) lacking poly(A) tails. RNA integrity and purity were assessed by 1% agarose gel electrophoresis and spectrophotometry (NanoDrop 2000, Thermo Fisher). Only samples with RNA integrity number (RIN) $\geq$ 8.0 were used for library construction.

cDNA library construction and sequencing

Poly(A)-enriched mRNA libraries were constructed using the NEBNext Ultra RNA Library Prep Kit (Illumina, USA) according to the standard protocol. Ribosomal RNA was removed, and mRNA was enriched using SPRI purification beads, followed by fragmentation via high-temperature treatment. The fragmented mRNA served as a template for first-strand cDNA synthesis, followed by second-strand cDNA synthesis. The purified double-stranded cDNA underwent end repair, A-tailing, and adapter ligation before PCR amplification. The libraries were sequenced on an Illumina NovaSeq 6000 platform (paired-end, 150 bp) by Guangzhou Aozhi Biotechnology Co., Ltd. Three biological replicates were performed for each stage, yielding a total of 12 libraries.

Quality control, read mapping, and transcriptome assembly

Raw reads were processed with fastp (v0.23.4) [29] to remove adapter sequences, reads with >10% ambiguous bases, reads consisting solely of poly-A, and low-quality reads (Q-score ≤20 over >50% of the read length). Clean reads were first aligned to the *Euwallacea* rRNA reference database using Bowtie2 (v2.5.1) [30] to eliminate ribosomal RNA contamination. The remaining reads were then mapped to the reference genome of *Euwallacea fornicatus* (GenBank assembly GCF_040115645.1) using HISAT2 (v2.2.1) [31] with the parameter "--ma-strandness RF" and other parameters set as default. Transcript assembly was performed with StringTie (v2.2.1) [32, 33], generating a comprehensive transcriptome including both known and novel transcripts.

Identification of long non-coding RNAs (lncRNAs) and circular RNAs (circRNAs)

Novel transcripts assembled by StringTie were identified using Cuffcompare and filtered to retain those with length ≥200 nt and at least two exons. Three computational tools—CPC2 (v0.1) [34], CNCI (v2.0) [35], and FEELnc (v0.2) [36]—were employed in parallel to assess coding potential. Only transcripts predicted as non-coding by all three tools were retained as high-confidence

lncRNAs. These lncRNAs were further classified into antisense, intronic, sense, bidirectional, and intergenic categories based on their genomic positions relative to annotated protein-coding genes [37].

For circRNA identification, the clean reads that passed rRNA depletion were aligned to the reference genome using HISAT2 [31], and the resulting SAM files were processed with CIRIquant (v1.0.0) [38] to detect back-spliced junction reads. CIRIquant was run with default parameters, and candidate circRNAs were initially identified based on the presence of at least two unique back-spliced reads. To obtain a high-confidence set, we applied additional filtering criteria: each circRNA had to be supported by ≥2 distinct junction reads and exhibit a minimum of two independent read pairs spanning the back-splice site, with a maximum circRNA length of 100 kb. The filtered circRNAs were then annotated against the reference genome annotation using the built-in annotation module of CIRIquant, which assigns genomic coordinates, flanking gene contexts, and host gene symbols. Only circRNAs with reliable junction signals and consistent splice sites across replicates were retained for downstream analyses.

Differential expression analysis

Expression levels for mRNAs and lncRNAs were quantified as transcripts per million (TPM) using StringTie [32, 33], and circRNA abundance was quantified as RPM (Reads Per Million mapped reads) based on back-spliced junction reads [38]. Differential expression analysis between developmental stages (EP vs. MP, MP vs. LP, LP vs. EA) was performed using DESeq2 (v1.38.3) [39] with the threshold of $|\log_2(\text{fold change})| > 1$ and false discovery rate (FDR) < 0.05. Volcano plots and Venn diagrams were generated to visualize shared and stage-specific differentially expressed transcripts.

Functional enrichment analysis

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were conducted on the sets of differentially expressed mRNAs and source genes of circRNAs [40]. Enrichment significance was assessed using a hypergeometric test with FDR correction (q-value ≤0.05) [41]. Bubble charts and circle plots were drawn to illustrate enriched terms and pathways.

Time-series trend analysis

To identify gene expression patterns that paralleled mycangial morphogenesis, the STEM (Short Time-series Expression Miner) algorithm [42] was applied to the TPM values across the four developmental stages. Clustering parameters were set to 10 profiles and a maximum unit change. Profiles with significant enrichment (q-value ≤0.05) were selected for downstream KEGG pathway analysis [40]. The expression trajectories were manually matched to biological events (organogenesis, immune maturation, symbiont accommodation) based on previous knowledge.

lncRNA–mRNA interaction networks

Three types of regulatory interactions were predicted:

Cis-regulation: lncRNAs located within 10 kb upstream or downstream of a protein-coding gene on the same genomic scaffold were considered potential cis targets [37].

Trans-regulation: Pearson correlation coefficients (PCC) were calculated between expression profiles of all lncRNA–mRNA pairs; pairs with $|PCC| > 0.95$ and p-value < 0.01 were retained as trans interactions.

Antisense regulation: lncRNAs with sequence complementarity to mRNA transcripts were identified using BLASTN with the Vienna RNA package to predict the best base pairing based on the calculation of minimum free energy through thermodynamic structure [43].

The resulting lncRNA–mRNA pairs were used to construct co-expression networks using Cytoscape (v3.7.1) [44]. Functional annotations of the network nodes were derived from GO and KEGG enrichment of the target mRNAs.

Quantitative real-time PCR (qPCR) validation

Total RNA from independently collected samples (three biological replicates per stage) was reverse-transcribed using the TransScript One-Step gDNA Removal and cDNA Synthesis SuperMix (TransGen, China). qPCR was performed on a QuantStudio™ 5 Real-Time PCR System using PerfectStart Green qPCR SuperMix (TransGen). Primers for three differentially expressed mRNAs and six lncRNAs were designed with Primer3 (Table S1) [45]. The *EiRPS3a* gene served as the internal reference, based on its stability in ambrosia beetle tissues [11]. Each reaction was run in triplicate. Relative expression levels were calculated using the 2^(−ΔCt) method. The qPCR results were compared with RNA-seq TPM values to validate the sequencing data.

Results

Sequencing data quality and transcriptome overview

After stringent quality filtering, each of the 12 libraries yielded >9.3 million clean reads, with clean read rates ranging from 97.69% to 98.69% (Table 1). Q20 and Q30 values exceeded 99.01% and 95.95%, respectively, indicating high sequencing quality. GC content ranged from 42.83% to 44.02%.

Identification and characterization of lncRNAs and circRNAs

A total of 2,692 transcripts were identified as putative lncRNAs and classified by genomic location: antisense (629), intergenic (501), intronic (523), sense (269), bidirectional (135), and unclassified (204) (Fig. 1A). Three computational tools (CNCI, CPC2, FEELnc) predicted 844 high-confidence lncRNAs common to all methods, while 447, 365, and 33 were uniquely identified by each tool, respectively (Fig. 1B). Coding potential scores and ORF cover values were consistently low, supporting non-coding identity.

Concurrently, 2,820 circRNAs were identified across samples. Expression counts varied by group: EP (745, 712, 717 for replicates), MP (1,061, 1,053, 1,081), LP (905, 1,011, 1,120), and EA (899, 826, 824) (Fig. 1C), indicating dynamic, condition-specific circRNA abundance with highest overall levels in MP.

Structural distinction between lncRNAs and mRNAs

LncRNAs showed a sharp concentration in shorter size ranges: 500–1,000 nt (38.17%), <500 nt (24.68%), and 1,000–1,500 nt (13.31%), with >76% shorter than 1,500 nt. In contrast, mRNAs displayed broader length distribution, with the largest fraction >4,500 nt (21.13%), and substantial proportions at 1,500–2,000 nt (15.3%) and 2,000–2,500 nt (12.03%) (Fig. 2A). Exon number also differed profoundly: most lncRNAs had two (45.73%) or three (34.72%) exons and none were mono-exonic, whereas mRNAs had wide exon counts (1 to >20), with only 2.77% mono-exonic and 9.47% with ≥20 exons (Fig. 2B). These features are consistent with known lncRNA architecture in insects.

Global expression profiling reveals group-specific clustering

PCA of mRNA expression showed clear separation of all four groups, with PC1 (80.9%) and PC2 (15.8%) explaining 96.7% variance (Fig. 3A; Table S2). EP was distinctly separated along PC1,

while MP separated from EA and LP along PC2; EA and LP clustered closely but were resolved by PC2. For lncRNA expression, PC1 (85.21%) and PC2 (12.35%) explained 97.56% variance (Fig. 3B; Table S2). MP showed the most striking separation along PC1, while EP, EA, and LP overlapped but were partially resolved along PC2. For circRNAs, PC1 (81.7%) and PC2 (5.9%) explained 87.6% variance (Fig. 3C; Table S2); EP clearly separated along PC1, while EA, LP, and MP overlapped, with EA showing high intra-group variability. These results indicate that each transcript type exhibits stage-specific expression profiles, with mRNAs and lncRNAs showing the most pronounced separations, particularly for EP and MP conditions.

Differential expression analysis of mRNAs, lncRNAs and circRNAs

Venn analysis of mRNAs revealed 266 commonly differentially expressed across all three comparisons (EP-vs-MP, MP-vs-LP, LP-vs-EA) (Fig. 4A). DEM numbers varied: EP-vs-MP had 1191 up- and 2966 down-regulated; MP-vs-LP had 427 up- and 563 down-regulated; LP-vs-EA had 437 up- and 477 down-regulated (Fig. 4B; Table S3). For lncRNAs, 12 were common to all comparisons (Fig. 4C); EP-vs-MP uniquely altered 145 lncRNAs, while MP-vs-LP and LP-vs-EA had 19 and 20 specific ones, respectively. Differential lncRNA counts mirrored mRNA patterns but with higher $\log_2$ fold changes (Fig. 4D; Table S3). For circRNAs, only 3 were common across all comparisons (Fig. 4F); EP-vs-MP uniquely regulated 114, MP-vs-LP 32, and LP-vs-EA 42. CircRNA differential expression was lower overall (Fig. 4E; Table S3), suggesting more stable profiles or distinct regulatory mechanisms.

GO and KEGG pathway enrichment analysis of differentially expressed mRNAs

EP vs MP Transition

GO enrichment for EP-vs-MP DEMs showed strong predominance of nuclear organization, chromosome biology, and gene regulation (Fig. 5A; Table S4). Top CC terms included *nuclear protein-containing complex*, *nucleoplasm*, and *nuclear chromosome*; BP terms were dominated by *chromosome organization*, *regulation of nucleobase-containing compound metabolic process*, and *regulation of RNA metabolic process*, along with cell-cycle terms (*regulation of cell cycle*; *mitotic cell cycle*). Most genes in these categories were down-regulated, indicating global reduction of nuclear and cell-cycle functions during early-to-middle pupal transition.

KEGG enrichment for EP-vs-MP revealed top pathways: *Cell cycle* (ko04110, 82 genes), *Spliceosome* (ko03040, 90 genes), *DNA replication* (ko03030, 31 genes), and *Nucleocytoplasmic transport* (ko03013, 61 genes) (Fig. 5D; Table S4). Other enriched pathways included *Proteasome*, *Nucleotide excision repair*, and *Mismatch repair*. The vast majority of genes were down-regulated, corroborating reduced DNA replication, transcription, and cell division.

MP vs LP Transition

MP-vs-LP GO terms were primarily associated with transmembrane transport and extracellular structures (Fig. 5B). Top MF terms: *xenobiotic transmembrane transporter activity*, *sugar transmembrane transporter activity*, and *carbohydrate transmembrane transporter activity*; BP included *xenobiotic transport* and *carbohydrate transport*. Cuticle-related terms (*structural constituent of cuticle*, *structural constituent of chitin-based cuticle*) were also enriched, indicating active cuticle remodeling.

KEGG enrichment shifted to metabolism and transport: *Lysosome* (ko04142, 23 genes), *Cholesterol metabolism* (ko04979, 11 genes), *Tyrosine metabolism* (ko00350, 8 genes), and *ABC transporters*

(ko02010) (Fig. 5E). *Metabolic pathways* (ko01100) contained 89 genes. These results suggest lysosomal degradation and metabolic remodeling for tissue breakdown and energy mobilization.

LP vs EA Transition

LP-vs-EA GO enrichment (Fig. 5C) featured xenobiotic transport, heme binding, and pigment biosynthesis. Top MF: *xenobiotic transmembrane transporter activity*, *heme binding*, and *tetrapyrrole binding*; BP included *pigment biosynthetic process* and *pigment metabolic process*, reflecting pigmentation and cuticle sclerotization during adult emergence. Many genes were up-regulated.

KEGG enrichment (Fig. 5F) was distinct: *Antifolate resistance* (ko01523, 15 genes), *ABC transporters* (ko02010, 17 genes), *Bile secretion* (ko04976, 19 genes), *Vitamin digestion and absorption* (ko04977, 10 genes), and *Cholesterol metabolism* (ko04979, 13 genes). Importantly, *Insect hormone biosynthesis* (ko00981, 6 genes) was enriched, implicating ecdysteroid/juvenile hormone pathways in the final pupal-to-adult transition.

GO and KEGG pathway enrichment analysis of differentially expressed circRNA

EP vs MP Transition

GO enrichment for circRNA-associated differentially expressed genes revealed nervous system development and cellular morphogenesis (Fig. 6A; Table S5). Top BP terms: *nervous system development*, *cell development*, *cell differentiation*, *anatomical structure morphogenesis*, and *neuron differentiation*. KEGG pathways (Fig. 6D; Table S5) were primarily muscle-related: *Vascular smooth muscle contraction* (ko04270), *Adrenergic signaling in cardiomyocytes* (ko04261), *Cardiac muscle contraction* (ko04260), and *Cytoskeleton in muscle cells* (ko04820) (all Q=0.190), suggesting circRNA involvement in muscle development during early pupal stage.

MP vs LP Transition

GO terms for MP-vs-LP (Fig. 6B) centered on regulation of neurogenesis: *regulation of neurogenesis*, *neurogenesis*, *anatomical structure morphogenesis*, *generation of neurons*, and *regulation of neuron projection development*. KEGG pathways showed no significant enrichment (all Q>0.2), indicating that circRNAs may affect individual genes rather than whole pathways during this transition (Fig. 6E).

LP vs EA Transition

LP-vs-EA GO enrichment (Fig. 6C) highlighted cytoskeletal dynamics and muscle function: *actin cytoskeleton organization* (Q=0.0095), *muscle contraction* (Q=0.0095), *actin filament-based process* (Q=0.0095), and *actomyosin structure organization* (Q=0.0095). KEGG (Fig. 6F) showed *Cytoskeleton in muscle cells* (ko04820, $Q=7.80\times10^{-4}$) as most significant, along with *Toll and Imd signaling pathway* (ko04624, Q=0.223), suggesting involvement in adult muscle maturation and immune preparation.

Time-series clustering of expression profiles

For mRNAs, the three largest significant profiles (profile0, profile1, profile8) accounted for ~70% of differentially expressed transcripts (Fig. 7A). **Profile0** (2,830 transcripts) showed stable expression; **profile1** (1,215) progressively decreased from EP to EA; **profile8** (431) peaked at MP. For lncRNAs (Fig. 7B), **profile0** (286) was stable, **profile1** (203) decreasing, and **profile9** (178) increasing, resembling mRNA patterns. For circRNAs (Fig. 7C), the dominant profile was **profile9** (675) showing progressive increase, followed by **profile0** (485) stable, **profile6** (430) peaking at MP, and **profile5** (387) gradual decrease. This indicates circRNAs are preferentially up-regulated toward later stages, potentially accumulating for adult-specific functions.

KEGG enrichment of key mRNA profiles

Profile0 (stable) was enriched in *Cell cycle* (ko04110, 78 genes), *ATP-dependent chromatin remodeling* (ko03082, 46 genes), and *Polycomb repressive complex* (ko03083, 32 genes) (Fig. 8A; Table S6), indicating constitutive expression of core cellular machinery. Profile1 (decreasing) enriched in *Spliceosome* (ko03040, 46 genes), *Ribosome biogenesis* (ko03008, 30 genes), and *Nucleocytoplasmic transport* (ko03013, 26 genes) (Fig. 8B), reflecting progressive shutdown of transcription/translation. Profile8 (MP peak) enriched in *Metabolic pathways* (ko01100, 54 genes), *ECM-receptor interaction* (ko04512, 7 genes), *Lysosome* (ko04142, 13 genes), and *Cytoskeleton in muscle cells* (ko04820, 9 genes) (Fig. 8C), indicating active metabolic and structural remodeling during mid-pupal development.

KEGG enrichment of key circRNA profiles

Profile5 (decreasing circRNAs) was enriched in muscle pathways: *Cytoskeleton in muscle cells* (ko04820, 13 genes), *Dilated cardiomyopathy* (ko05414, 10 genes), and *Adrenergic signaling* (ko04261, 11 genes) (Fig. 9A; Table S7). Profile6 (MP peak) enriched in signaling: *cGMP-PKG* (ko04022, 14 genes), *Insulin secretion* (ko04911, 9 genes), *Circadian entrainment* (ko04713), and *Hippo signaling* (ko04390) (Fig. 9B). Profile8 (valley pattern) enriched in *Cytoskeleton in muscle cells* (ko04820, 8 genes) and *cAMP signaling* (ko04024) (Fig. 9C). Profile9 (increasing circRNAs) showed the most diverse enrichment: *Cytoskeleton in muscle cells* (ko04820, 16 genes), *Adrenergic signaling* (ko04261, 12 genes), *Hypertrophic cardiomyopathy* (ko05410, 9 genes), *Dilated cardiomyopathy* (ko05414, 9 genes), along with *Insulin secretion*, *cAMP signaling*, and *Adherens junction* (Fig. 9D). These results underscore circRNA involvement in muscle development and signaling, with distinct temporal dynamics.

Functional annotation of lncRNA–mRNA regulatory networks

Cis-regulated lncRNA targets

Cis-targets (mRNAs within 100 kb of lncRNAs) showed GO enrichment in broad categories: BP – *cellular process* (221 genes), *metabolic process* (198), *biological regulation* (155); MF – *binding* (202), *catalytic activity* (141); CC – *cellular anatomical entity* (216), *protein-containing complex* (101) (Fig. 10A; Table S8). KEGG enrichment was modest (Q>0.1) with top pathways including *Autoimmune thyroid disease*, *Arginine biosynthesis*, and *Folate biosynthesis* (Fig. 10B), suggesting cis-lncRNAs regulate individual genes rather than entire pathways.

Antisense-regulated lncRNA targets

Antisense targets (overlapping opposite strand) had narrower GO scope (Fig. 10C): BP – *cellular process* (40), *metabolic process* (37), *biological regulation* (28); MF – *binding* (36), *catalytic activity* (23); CC – *cellular anatomical entity* (34). KEGG showed no significant enrichment (all Q>0.3) (Fig. 10D), indicating antisense lncRNAs may target scattered genes without converging on common pathways.

Trans-regulated lncRNA targets

Trans-targets (distally located) displayed the broadest GO categories (Fig. 10E): BP – *cellular process* (3,322), *metabolic process* (2,861), *biological regulation* (2,405), *multicellular organismal process* (1,849), *developmental process* (1,742); MF – *binding* (3,018), *catalytic activity* (1,879); CC – *cellular anatomical entity* (3,295). KEGG enrichment was highly significant (Fig. 10F; Table S9): *Cell cycle* (ko04110, 87 genes), *Nucleocytoplasmic transport* (ko03013, 71 genes), *Spliceosome* (ko03040, 88 genes), and other pathways including *ATP-dependent chromatin remodeling*, *Polycomb repressive complex*, *DNA replication*, *Proteasome*, and *mRNA surveillance*.

This indicates trans-lncRNAs exert widespread control over gene expression, genome stability, and developmental timing.

Validation of differentially expressed lncRNAs by qPCR

Six lncRNAs and three mRNAs were randomly selected for qPCR validation. Eight of them (MSTRG.8390.4, XR_010733211.1, MSTRG.10320.1, ncbi_136349018, ncbi_136346571, ncbi_136344757, MSTRG.2210.2, MSTRG.18735.1) showed expression patterns highly consistent with RNA-seq TPM profiles across the four stages (Fig. 11). One lncRNA (MSTRG.4263.2) showed discrepancy, possibly due to alternative splicing variants. Overall, qPCR confirmed the reliability of the transcriptomic data.

Discussion

In this study, we performed a comprehensive transcriptomic analysis to characterize the expression profiles of lncRNAs and circRNAs during the pupal development of *E. interjectus*, with a particular focus on the critical period of mycangial morphogenesis. Our results revealed that lncRNAs and circRNAs exhibit distinct structural features, stage-specific expression patterns, and potential regulatory functions that are closely associated with the developmental transitions from early pupa to adult. To our knowledge, this is the first study to systematically identify and characterize non-coding RNAs in the context of symbiotic organ development in ambrosia beetles, providing a foundational resource for understanding the molecular regulation of mycangial formation.

Structural features and identification of lncRNAs and circRNAs

The structural characteristics of the lncRNAs identified in *E. interjectus* are consistent with those reported in other insect species. Similar to findings in *Sogatella furcifera* [21] and *Aedes albopictus* [20], the lncRNAs in our study were generally shorter, possessed fewer exons, and exhibited lower

GC content compared to protein-coding mRNAs. The predominance of two-exon lncRNAs (45.73%) and the absence of mono-exonic transcripts reflect the stringent filtering criteria applied in our pipeline, ensuring a high-confidence lncRNA set for downstream analyses. In *Sarcophaga peregrina*, a genome-wide analysis identified 6,921 lncRNA transcripts across three pupal developmental stages, with the majority also exhibiting two exons and shorter transcript lengths compared to mRNAs [46]. Similarly, in the silkworm *Bombyx mori*, lncRNAs were found to have longer transcript lengths but fewer exons than mRNAs [47]. These consistent structural features across diverse insect taxa suggest that lncRNA architecture is evolutionarily constrained and may reflect fundamental aspects of their biogenesis and function. The identification of 2,820 circRNAs across all stages, with the highest abundance in the mid-pupa (MP) stage, suggests that circRNA accumulation may be particularly relevant during active morphogenesis, a phenomenon also observed during metamorphosis in *Drosophila* [23]. In *Bombyx mori*, circRNAs have been shown to be dynamically expressed during silk gland apoptosis, with specific circRNAs potentially functioning as miRNA sponges to regulate cell death processes [48].

Stage-specific expression and developmental trajectories

Principal component analysis (PCA) revealed clear stage-specific clustering for all three RNA types, with mRNAs and lncRNAs showing more pronounced separations than circRNAs. The distinct separation of EP along PC1 for mRNAs indicates a major transcriptomic shift as larvae transition to early pupae, which is consistent with the cessation of feeding and the initiation of metamorphosis. The striking separation of MP along PC1 for lncRNAs suggests that these non-coding transcripts may play a particularly active role during mid-pupal development, coinciding with the peak of mycangial morphogenesis and extensive tissue remodeling [3].

The time-series clustering analysis further supported the functional specialization of different transcript types. The progressive decrease of Profile1 transcripts (including mRNAs and lncRNAs) from EP to EA, enriched in pathways such as spliceosome and ribosome biogenesis, reflects the global shutdown of active transcription and translation as development proceeds toward the quiescent adult stage. This pattern is consistent with findings in *Sarcophaga peregrina*, where lncRNA expression profiles across pupal development revealed that many lncRNAs were downregulated during the mid-to-late pupal transition, while a specific lncRNA (SP_lnc5000) peaked at the mid-pupal stage and was implicated in regulating metamorphosis-related genes in the 20E signaling pathway [46]. In contrast, the increasing trend of circRNA Profile9, enriched in muscle-related pathways, suggests that circRNAs may accumulate to support the differentiation and maturation of adult-specific tissues, including flight muscles and the mycangial apparatus. This is consistent with the known stability and accumulation of circRNAs in post-mitotic tissues [23, 49]. Furthermore, in *Drosophila*, circRNAs have been shown to be abundantly localized in the brain and central nervous system, accumulating with age and potentially contributing to neural function and longevity [23].

Functional implications of differentially expressed mRNAs

The GO and KEGG enrichment analyses of differentially expressed mRNAs across the three developmental transitions provided a detailed molecular portrait of the pupal-to-adult metamorphosis. The EP-to-MP transition was characterized by a widespread downregulation of genes involved in cell cycle, DNA replication, and RNA processing, indicative of the termination of mitotic divisions and the entry into a differentiation-dominated phase. This is consistent with the

well-established role of ecdysone signaling in halting cell proliferation and initiating metamorphic gene cascades [50].

The MP-to-LP transition showed a shift toward lysosomal activity, cholesterol metabolism, and cuticle remodeling. The enrichment of lysosomal pathways likely reflects the massive histolysis of larval tissues, a hallmark of holometabolous metamorphosis. Concurrently, the upregulation of chitin-based cuticle synthesis genes suggests active deposition of the pupal cuticle, which is essential for protection during the quiescent pupal stage. The LP-to-EA transition was marked by the enrichment of pigment biosynthesis and insect hormone biosynthesis pathways, aligning with the final cuticle sclerotization and the emergence of the adult phenotype. These findings are consistent with the hormonal regulation of insect metamorphosis, where the decline of juvenile hormone and the surge of ecdysteroids orchestrate stage-specific gene expression programs [51, 52]. In *Bombyx mori*, 20-hydroxyecdysone (20E) has been shown to induce autophagy in the fat body during the larval-pupal transition by upregulating autophagy-related genes (Atg genes), and lncRNAs such as LNC_000560 were found to be involved in this process by potentially regulating the expression of Atg4B [47].

Potential roles of circRNAs in muscle and nervous system development

A notable finding of this study is the enrichment of circRNA-associated genes in muscle development and nervous system-related pathways. During the EP-to-MP transition, the source genes of differentially expressed circRNAs were enriched in terms such as nervous system development and neuron differentiation, as well as muscle contraction pathways. This suggests that circRNAs may be actively involved in the early differentiation of neural and muscular tissues, which are critical for adult locomotion and behavior. In *Drosophila*, circRNAs have been shown to

regulate muscle function and neuronal development through miRNA sponging or protein scaffolding [53, 54]. The circRNA Edis, derived from the Ect4 gene in *Drosophila*, has been demonstrated to regulate neurodevelopment and innate immunity, encoding a peptide that suppresses the immune deficiency (IMD) pathway [54]. Similarly, the circRNA Edis-Relish-castor axis has been shown to regulate neuronal development in *Drosophila* [55]. Our results extend these observations to a non-model insect and suggest that similar regulatory mechanisms may operate during the development of the mycangium-associated musculature.

The LP-to-EA transition exhibited a strong enrichment of circRNA-associated genes in actin cytoskeleton organization and actomyosin structure, processes that are fundamental for muscle cell maturation and contraction. The significant enrichment of the *Toll and Imd signaling pathway* at this stage also hints at a potential role for circRNAs in priming the immune system for adult life, as newly emerged adults may be exposed to novel microbial challenges [56]. In *Aedes albopictus*, circRNAs have been implicated in chitin metabolism and immune responses, with whole-transcriptome profiling revealing that circRNAs can serve as hub genes in competing endogenous RNA (ceRNA) networks to regulate developmental processes [16].

lncRNA–mRNA regulatory networks: cis, antisense, and trans interactions

The construction of lncRNA–mRNA interaction networks provided functional insights into the potential regulatory roles of lncRNAs during mycangial development. While cis-targets showed broad but modest functional enrichment, trans-targets displayed highly significant enrichment in pathways central to development, including the cell cycle, spliceosome, and nucleocytoplasmic transport. This suggests that trans-acting lncRNAs may exert a more widespread influence on gene expression networks than their cis-acting counterparts, possibly by modulating the activity of key

transcription factors or signaling molecules [57]. The absence of significant KEGG enrichment for antisense targets suggests that antisense lncRNAs may act on individual genes rather than entire pathways, consistent with their role in post-transcriptional regulation of specific transcripts [58]. In *Bombyx mori*, a study on 20E-regulated autophagy identified 6,493 lncRNA-mRNA cis pairs and 42,797 trans pairs, with the majority of trans interactions being positively correlated, suggesting that trans-acting lncRNAs may function primarily through positive regulation of their target genes [47]. The identification of lncRNAs with dynamic expression patterns correlating with key developmental transitions suggests that they may function as important regulators of metamorphosis. For instance, lncRNAs that are highly expressed during the MP stage and target genes involved in tissue remodeling and immune maturation could be critical for the successful establishment of the mycangium and its fungal symbionts. This is particularly intriguing given that mycangial development occurs autonomously in the absence of fungal cues [3], implying that endogenous regulatory programs, potentially involving lncRNAs, are sufficient to orchestrate this process. In *Sarcophaga peregrina*, RNA interference of the lncRNA SP_lnc5000 led to deformed pupae, stagnation of pupal-adult metamorphosis, and reduced expression of metamorphosis-related genes in the 20E signaling pathway, cuticle tanning pathway, and chitin-related pathway [46]. This provides direct functional evidence that specific lncRNAs can regulate pupal metamorphosis by affecting cuticle structure and composition. Furthermore, in *Bombyx mori*, the lncRNA lncR17454 was found to regulate metamorphosis through the let-7 miRNA cluster, with knockdown of lncR17454 leading to molting arrest and altered expression of ecdysone signaling pathway genes [59]. In honey bees, the lncRNA Nb-1 is expressed in a sex-specific and developmentally regulated

manner, with roles in embryogenesis, mushroom body neuroblast proliferation, and drone-specific retinal development, suggesting pleiotropic functions associated with the honey bee lifecycle [60].

Integrating non-coding RNAs into the model of mycangial development

Based on our findings, we propose a working model in which lncRNAs and circRNAs coordinate the molecular events leading to mycangial formation. During the EP-to-MP transition, the global downregulation of cell cycle and transcription-related genes, coupled with the upregulation of lncRNAs targeting developmental pathways, may facilitate the transition from cell proliferation to differentiation. The peak of circRNA abundance in MP, along with the enrichment of their source genes in morphogenesis-related pathways, suggests that circRNAs may play a role in stabilizing or fine-tuning the expression of genes essential for mycangial morphogenesis. In the LP-to-EA transition, the increasing circRNA profiles and the activation of immune and muscle-related pathways may prepare the mycangium and the adult beetle for symbiont inoculation and environmental challenges. This model is consistent with the observation that lncRNAs can function as precursors of miRNAs, as demonstrated for lncR17454 in *Bombyx mori*, where it serves as the precursor of the let-7 miRNA cluster and regulates metamorphosis through the ecdysone signaling pathway [59]. Similarly, in *Bombyx mori*, the lncRNA lncR26319 functions as a competing endogenous RNA (ceRNA) for miR-2834 to regulate Endophilin A expression, which is essential for vitellogenin endocytosis during oogenesis [61]. These examples highlight the diverse regulatory mechanisms through which lncRNAs can influence developmental processes in insects.

Despite the comprehensive nature of this study, several limitations should be acknowledged. First, our transcriptomic analysis was based on whole head-thorax blocks (for pupae) and whole heads (for adults), which, while containing the mycangial primordia, also include other tissues. This bulk

tissue approach may have obscured subtle tissue-specific changes in gene expression. Future studies employing laser-capture microdissection or single-cell RNA-seq could provide more precise spatial resolution of the transcriptional changes occurring within the mycangium itself. Second, our analyses are purely descriptive and computational; functional validation through RNA interference (RNAi) or CRISPR-based knockdown of candidate lncRNAs and circRNAs will be essential to confirm their roles in mycangial development.

Conclusions

This study provides the first comprehensive characterization of long non-coding RNAs and circular RNAs associated with pupal development and mycangial morphogenesis in the ambrosia beetle *Euwallacea interjectus*. Our transcriptomic analysis revealed that these non-coding RNAs exhibit distinct stage-specific expression patterns and are enriched in pathways critical for metamorphosis, tissue remodeling, and immune maturation, with circRNAs being particularly abundant during the mid-pupal stage and associated with muscle and nervous system development. Through the construction of lncRNA–mRNA regulatory networks, we identified that trans-acting lncRNAs are extensively involved in core developmental pathways including cell cycle regulation, spliceosome function, and nucleocytoplasmic transport. The temporal expression trajectories of multiple transcripts closely align with key developmental events such as mycangial morphogenesis, larval tissue histolysis, and adult cuticle formation, suggesting that non-coding RNAs may orchestrate pupal metamorphosis through complex competing endogenous RNA regulatory networks. These findings establish a foundational framework for understanding the molecular mechanisms underlying symbiotic organ development in ambrosia beetles and provide potential targets for developing novel pest management strategies that disrupt the beetle–fungus symbiosis.

Supplementary Information

Supplementary Figure S1: Schematic illustration of the key developmental stages of mycangia in *Euwallacea interjectus*

Supplementary Table S1: List of primers used for quantitative real-time PCR (qPCR) validation.

Supplementary Table S2: Principal component analysis (PCA) of transcriptome expression profiles. Principal component (PC) loadings for mRNA, lncRNA, and circRNA expression datasets across all samples. For each RNA type, the table provides the sample name and its coordinates along PC1, PC2, and PC3, with the percentage of variance explained by each principal component indicated in parentheses.

Supplementary Table S3: Differential expression analysis of mRNAs, lncRNAs, and circRNAs. Comprehensive list of differentially expressed transcripts across the three developmental comparisons (EP vs. MP, MP vs. LP, and LP vs. EA) for each RNA species (mRNA, lncRNA, circRNA).

Supplementary Table S4: GO and KEGG enrichment analysis of differentially expressed mRNAs. Functional enrichment analysis of differentially expressed mRNAs for each developmental transition (EP vs. MP, MP vs. LP, and LP vs. EA). The categories "Up" and "Down" indicate the number of genes up- or down-regulated within each term.

Supplementary Table S5: GO and KEGG enrichment analysis of differentially expressed circRNA-associated genes. GO and KEGG enrichment analysis of the source genes of differentially expressed circRNAs for each developmental transition (EP vs. MP, MP vs. LP, and LP vs. EA).

Supplementary Table S6: KEGG pathway enrichment analysis of major mRNA temporal expression profiles. KEGG pathway enrichment analysis for the three most prominent mRNA time-series expression profiles (Profile 0: stable, Profile 1: decreasing, and Profile 8: peaking at MP) identified by STEM clustering.

Supplementary Table S7: KEGG pathway enrichment analysis of major circRNA temporal expression profiles. KEGG pathway enrichment analysis for the four major circRNA time-series expression profiles (Profile 5: decreasing, Profile 6: peak at MP, Profile 8: valley at MP, and Profile

9: increasing) identified by STEM clustering.

Supplementary Table S8: Functional annotation of lncRNA–mRNA regulatory networks. Summary of GO and KEGG enrichment analyses for mRNAs targeted by lncRNAs through three distinct regulatory modes: *cis*-regulation, antisense-regulation, and *trans*-regulation.

Declarations

Ethics, Consent to Participate, and Consent to Publish declarations

Not applicable. This study did not involve human participants, their data, or biological specimens, nor did it involve vertebrate animals.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The RNA-seq datasets generated during the current study are available in the NCBI Sequence Read Archive (SRA) under accession number: SAMN62011601-SAMN62011612. All other data supporting the findings of this study are included in the manuscript and its supplementary information files.

Funding

This study received financial support from the National Natural Science Foundation of China (32572060).

Authors' contributions

M.Z. and L.D. conceived and designed the study. M.Z. and W.P. performed the experiments and data analysis. M.Z. and M.S. drafted the manuscript. L.D. revised the manuscript and supervised the study. All authors read and approved the final manuscript.

References

1. Hulcr J, Stelinski LL. The ambrosia symbiosis: From evolutionary ecology to practical management. Annu Rev Entomol. 2017; 62(1): 285–303.
2. Francke-Grosmann H. Hautdrüsen als Träger der Pilzsymbiose bei Ambrosiakäfern. Z Morphol Ökol Tiere. 1956; 45(3): 275–308.
3. Spahr E, Kasson MT, Kijimoto T. Micro-computed tomography permits enhanced visualization of mycangia across development and between sexes in *Euwallacea* ambrosia beetles. PLoS ONE. 2020; 15(9): e0236653.
4. Li Y, Ruan YY, Stanley EL, Skelton J, Hulcr L. Plasticity of mycangia in *Xylosandrus* ambrosia beetles. Insect Sci. 2019; 26(4) :732–42.
5. Spahr EJ, Wasef F, Kasson MT, Kijimoto T. Developmental genetic underpinnings of a symbiosis-associated organ in the fungus-farming ambrosia beetle *Euwallacea validus*. Sci Rep. 2023; 13: 14014.
6. Lai S, Jiang N, Wang J, Zhou Y, Yu X, Wu Y, et al. Phylogeography of the ambrosia beetle *Euwallacea interjectus* (Coleoptera: Curculionidae: Scolytinae): an emerging poplar pest and its *Fusarium* mutualists from poplar plantations in China. J Econ Entomol. 2024; 117(6): 2505-2517.
7. Wang L, Chen F, Qiu C, Tang J, Ding X, Ren J. Identification and risk analysis of *Euwallacea interjectus*. J Nanjing For Univ (Nat Sci Ed). 2021; 45: 201–8.
8. Lai S, Zhao C, Li Y, Zhou Y, Zhong L, Qiu C, et al. Three novel *Fusarium* mutualists of ambrosia beetle *Euwallacea interjectus* in China. Mycol Prog. 2022; 21(8): 70.

9. Hu J, Zhao C, Tan J, Lai S, Zhou Y, Dai L. Transcriptome analysis of *Euwallacea interjectus* reveals differentially expressed unigenes related to developmental stages and egg laying. Comp Biochem Physiol D-Genomics Proteomics. 2023; 47: 101100.

10. Blaz J, Barrera-Redondo J, Vazquez-Rosas-Landa M, Canedo-Téxon A, von Wobeser EA, Carrillo D, et al. Genomic signals of adaptation towards mutualism and sociality in two ambrosia beetle complexes. Life-Basel. 2018; 9(1): 2.

11. Patwa N, Ranger CM, Lehenberger M, Biedermann PH, Reding ME. Stability of nuclear and mitochondrial reference genes in selected tissues of the ambrosia beetle *Xylosandrus germanus*. Insects. 2021; 12(12): 1125.

12. Kopp F, Mendell JT. Functional classification and experimental dissection of long noncoding RNAs. Cell. 2018; 172(3): 393–407.

13. Chen B, Zhang Y, Zhang X, Jia S, Chen S, Kang L. Genome-wide identification and developmental expression profiling of long noncoding RNAs during *Drosophila* metamorphosis. Sci Rep. 2016; 6: 23330.

14. Wen K, Yang L, Xiong T, Di C, Ma DH, Wu MH, et al. Critical roles of long noncoding RNAs in *Drosophila* spermatogenesis. Genome Res. 2016; 26(9): 1233–44.

15. Etebari K, Asad S, Zhang G, Asgari S. Identification of *Aedes aegypti* long intergenic non-coding RNAs and their association with *Wolbachia* and dengue virus infection. PLoS Negl Trop Dis. 2016; 10(10): e0005069.

16. Liu W, An S, Cheng P, Zhang K, Gong M, Zhang Z, Zhang R. Whole-transcriptome profiling across different developmental stages of *Aedes albopictus* (Diptera: Culicidae) provides

671 insights into chitin-related non-coding RNA and competing endogenous RNA networks.
672 Parasit Vectors. 2023; 16: 33.
673 17. Liu F, Shi T, Qi L, Su X, Wang D, Dong J, Huang ZY. lncRNA profile of *Apis mellifera* and
674 its possible role in behavioural transition from nurses to foragers. BMC Genomics. 2019; 20:
675 393.
676 18. Li X, Yang X, Yu F, Miao C, Li M, Wang K, et al. Differences between uncapping and
677 removal behaviors in *Apis cerana* from the perspective of long non-coding RNAs. BMC
678 Genomics. 2024; 25: 912.
679 19. Zhu B, Li L, Wei R, Liang P, Gao X. Regulation of GSTu1-mediated insecticide resistance in
680 *Plutella xylostella* by miRNA and lncRNA. PLoS Genet. 2021; 17(10): e1009888.
681 20. Liu W, Cheng P, Zhang K, Gong M, Zhang Z, Zhang R. Systematic identification and
682 characterization of long noncoding RNAs (lncRNAs) during *Aedes albopictus* development.
683 PLoS Negl Trop Dis. 2022; 16(4): e0010245.
684 21. Chang ZX, Ajayi OE, Guo DY, Wu QF. Genome-wide characterization and developmental
685 expression profiling of long non-coding RNAs in *Sogatella furcifera*. Insect Sci. 2020; 27:
686 987–997.
687 22. Shang YJ, Feng YK, Ren LP, Zhang XY, Yang FQ, Zhang CQ, Guo YD. Genome-wide
688 analysis of long noncoding RNAs and their association in regulating the metamorphosis of the
689 *Sarcophaga peregrina* (Diptera: Sarcophagidae). PLoS Negl Trop Dis. 2023; 17(6): e0011411.
690 23. Westholm JO, Miura P, Olson S, Shenker S, Joseph B, Sanfilippo P, Celniker SE, Graveley
691 BR, Lai EC. Genome-wide analysis of *Drosophila* circular RNAs reveals their structural and
692 sequence properties and age-dependent neural accumulation. Cell Rep. 2014; 9(5): 1966–1980.

24. Li Y, Zhang Q, Chen J. Circular RNAs in insect development and stress responses. Insect Mol Biol. 2020;29(3):231–242.

25. Yin H, Zhang S, Shen M, Zhang Z, Huang H, Zhao Z, Guo X, Wu P. Integrative analysis of circRNA/miRNA/mRNA regulatory network reveals the potential immune function of circRNAs in the *Bombyx mori* fat body. J Invertebr Pathol. 2021; 179: 107537.

26. Gao Y, Yang L, Chen Y, Liu P, Zhou Y, Chen X, Gu J. Aal-circRNA-407 regulates ovarian development of *Aedes albopictus*, a major arbovirus vector, via the miR-9a-5p/Foxl axis. PLoS Pathog. 2023; 19(5): e1011374.

27. Gao L, Yan Q, Qu Y, Li W, Song J, Zhou S. circRNAs derived from a nuclear hormone receptor act differentially on insect metamorphosis and reproduction. J Genet Genomics. 2026; 53: 510–521.

28. Liu YJ, Lu JM, Xie XC, Xu QQ, Ding BY, Wang JJ, Shang F. Circ-379-miR-3050-DNAH3/RNF8 axis modulates wing development in *Aphis citricidus*. Insect Biochem Mol Biol. 2025; 184: 104393.

29. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018; 34(17): i884–890.

30. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012; 9(4): 357–359.

31. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015; 12(4): 357–360.

713 32. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables
714 improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015;33(3):
715 290–295.
716 33. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of
717 RNA-seq experiments with HISAT, StringTie and Ballgown. Nat Protoc. 2016; 11(9): 1650–
718 1667.
719 34. Kong L, Zhang Y, Ye ZQ, Liu XQ, Zhao SQ, Wei L, Gao G. CPC: assess the protein-coding
720 potential of transcripts using sequence features and support vector machine. Nucleic Acids
721 Res. 2007; 35: W345–349.
722 35. Sun L, Luo H, Bu D, Zhao G, Yu K, Zhang C, et al. Utilizing sequence intrinsic composition
723 to classify protein-coding and long non-coding transcripts. Nucleic Acids Res. 2013; 41(17):
724 e166.
725 36. Wucher V, Legeai F, Hédan B, Rizk G, Lagoutte L, Leeb T, et al. FEELnc: a tool for long
726 non-coding RNA annotation and its application to the dog transcriptome. Nucleic Acids Res.
727 2017; 45(8): e57.
728 37. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. The GENCODE v7
729 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and
730 expression. Genome Res. 2012; 22(9): 1775–1789.
731 38. Zhang J, Chen S, Yang J, Zhao F. Accurate quantification of circular RNAs identifies
732 extensive circular isoform switching events. Nat Commun. 2020; 11(1): 90.
733 39. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-
734 seq data with DESeq2. Genome Biol. 2014; 15(12): 550.

40. Kanehisa M, Araki M, Goto S, Hattori M, Hattori M, Itoh M, et al. KEGG for linking genomes to life and the environment. Nucleic Acids Res. 2008; 36(Database issue): D480–484.

41. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B. 1995; 57(1): 289–300.

42. Ernst J, Bar-Joseph Z. STEM: a tool for the analysis of short time series gene expression data. BMC Bioinformatics. 2006; 7: 191.

43. Tafer H, Hofacker IL. RNAplex: a fast tool for RNA-RNA interaction search. Bioinformatics. 2008; 24(22): 2657–63.

44. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003; 13(11): 2498–2504.

45. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. Primer3—new capabilities and interfaces. Nucleic Acids Res. 2012; 40(15): e115.

46. Xia Y, Wu H, Chen S, Wang Y, Sun J, Li Y, Guo Y, Shang Y. Temporal miRNA biomarkers for pupal age estimation in *Sarcophaga peregrina* (Diptera: Sarcophagidae). Insects. 2025; 16(8): 754.

47. Qiao H, Wang J, Wang Y, Yang J, Wei B, Li M, et al. Transcriptome analysis reveals potential function of long non-coding RNAs in 20-hydroxyecdysone regulated autophagy in *Bombyx mori*. BMC Genomics. 2021; 22: 374.

48. Chen RT, Xiao Y, Liu Z, Li LL, Lu Y, Jiao P, Miao YG. Three vital RNA functions and interactions in the process of silk gland apoptosis in silkworm *Bombyx mori*. Arch Insect Biochem Physiol. 2019; 100(3): e21511.

49. Weigelt CM, Sehgal R, Tain LS, Cheng J, Eßer J, Pahl A, Dieterich C, Grönke S, Partridge L. An insulin-sensitive circular RNA that regulates lifespan in *Drosophila*. Mol Cell. 2020; 79(2): 268–279.e5.

50. Truman JW, Riddiford LM. The evolution of insect metamorphosis: A developmental and endocrine view. Philos Trans R Soc B-Biol Sci. 2019; 374(1783): 20190070.

51. Jindra M, Palli SR, Riddiford LM. The juvenile hormone signaling pathway in insect development. Annu Rev Entomol. 2013; 58: 181–204.

52. Song J, Zhou S. Post-transcriptional regulation of insect metamorphosis and oogenesis. Cell Mol Life Sci. 2020; 77(10): 1893–909.

53. Pamudurti NR, Bartok O, Jens M, Ashwal-Fluss R, Stottmeister C, Ruhe L, et al. Translation of circRNAs. Mol Cell. 2017; 66(1): 9–21.e7.

54. Xiong XP, Liang W, Liu W, Xu S, Li JL, Tito A, et al. The circular RNA Edis regulates neurodevelopment and innate immunity. PLoS Genet. 2022;18(10): e1010429.

55. Liu W, Liang W, Xiong XP, Li JL, Zhou R. A circular RNA Edis-Relish-castor axis regulates neuronal development in *Drosophila*. PLoS Genet. 2022; 18(10): e1010433.

56. Lemaitre B, Hoffmann J. The host defense of *Drosophila melanogaster*. Annu Rev Immunol. 2007; 25: 697–743.

57. Li K, Tian Y, Yuan Y, Fan X, Yang M, He Z, Yang D. Insights into the functions of lncRNAs in *Drosophila*. Int J Mol Sci. 2019; 20(18): 4646.

58. Pease B, Borges AC, Bender W. Noncoding RNAs of the Ultrabithorax domain of the *Drosophila* bithorax complex. Genetics. 2013; 195(4): 1253–64.

59. Fu Y, Wang Y, Huang C, Zhao C, Li X, Kan Y, Li D. Long noncoding RNA lncR17454 regulates metamorphosis of silkworm through let-7 miRNA cluster. J Insect Sci. 2022; 22(3): 12.

60. Tadano H, Kohno H, Takeuchi H, Kubo T. Unique spatially and temporary-regulated/sex-specific expression of a long ncRNA, Nb-1, suggesting its pleiotropic functions associated with honey bee lifecycle. Sci Rep. 2024; 14: 8701.

61. Wang Y, Fu Y, Cheng H, Zhao C, Huang Q, Chang M, et al. lncR26319/miR-2834/EndophilinA axis regulates oogenesis of the silkworm, *Bombyx mori*. Insect Sci. 2022; 30: 65–80.

Figure legends

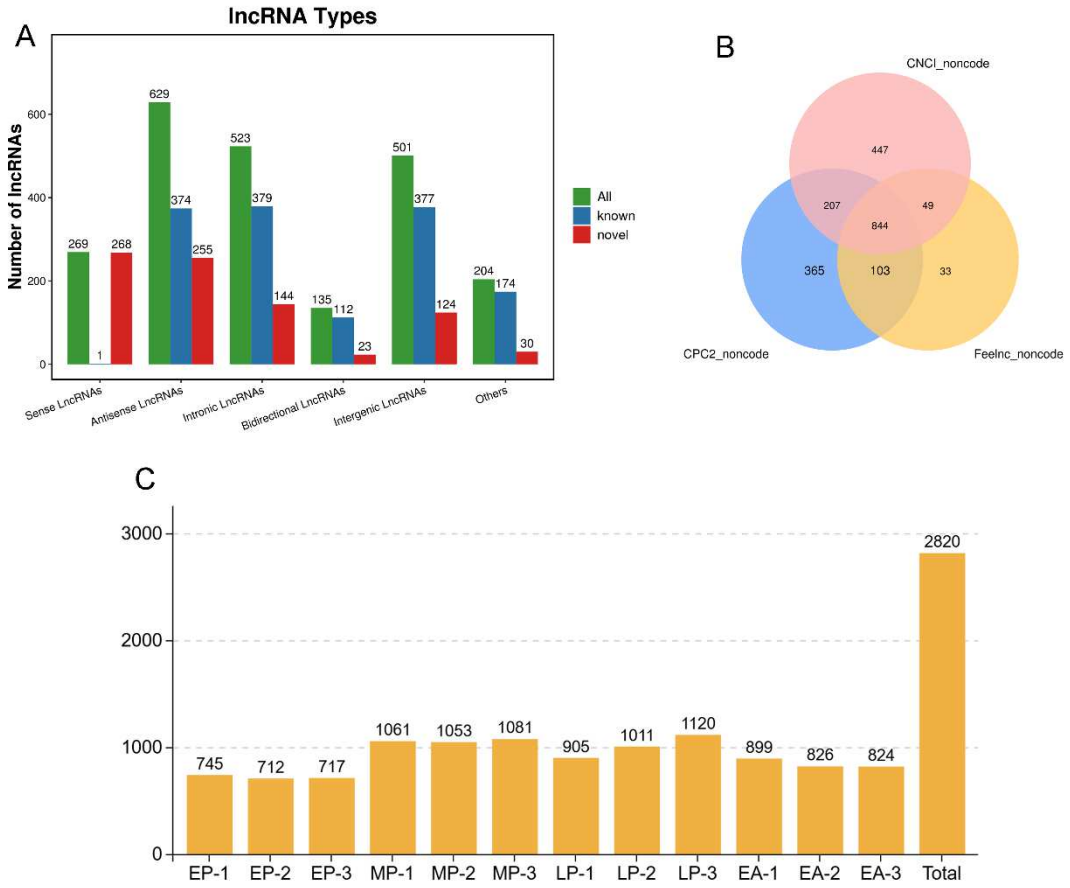


Figure 1. Identification and classification of lncRNAs and circRNAs.

(A) Classification of identified lncRNAs based on their genomic location (Sense, Antisense, Intronic, Bidirectional, Intergenic, and Others). (B) Venn diagram showing the overlap of lncRNA transcripts predicted by CNCI, CPC2, and FEELnc. (C) Bar plot showing the number of identified circRNAs in different sample groups (EP, MP, LP, EA) with three replicates each.

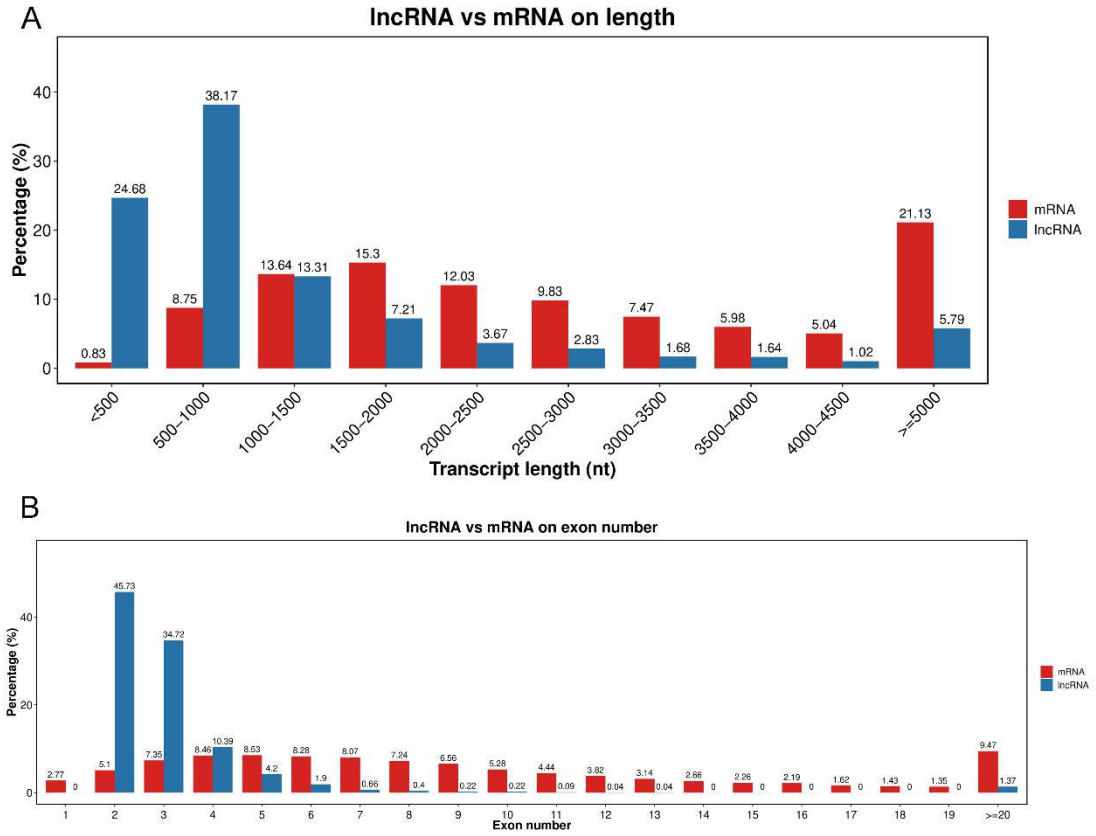
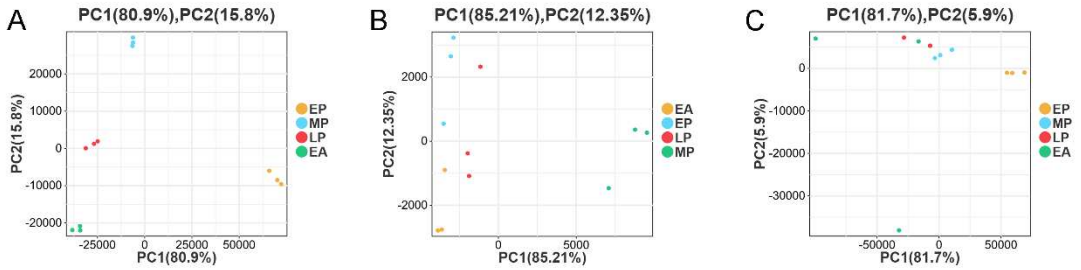


Figure 2. Comparative analysis of transcript structural features between mRNAs and lncRNAs. (A) Transcript length distribution. The percentage of transcripts is plotted across different length intervals (binned in 500-nt increments). (B) Exon number distribution. The percentage of transcripts is shown as a function of exon count.



803

804 Figure 3. Principal component analysis (PCA) of transcriptome expression profiles.

805 PCA plots for (A) mRNA, (B) lncRNA, and (C) circRNA expression data. Each point represents an

806 individual biological replicate, colored by group (EA, EP, LP, MP). The percentage of variance

807 explained by each principal component is indicated on the axes.

808

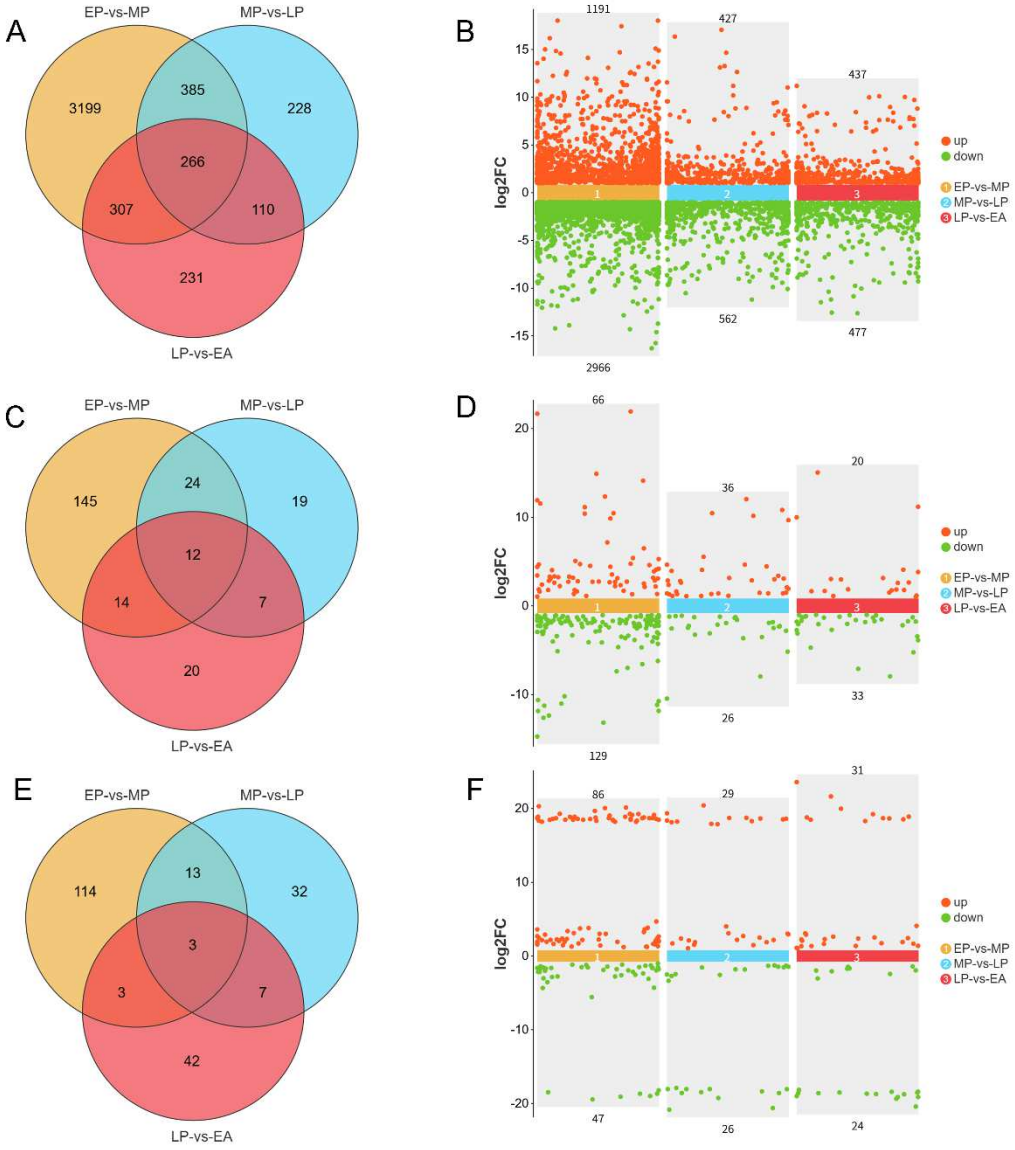


Figure 4. Differential expression profiles of mRNAs, lncRNAs, and circRNAs across developmental stages.

Volcano plots showing differentially expressed (A) mRNAs, (C) lncRNAs, and (E) circRNAs for each comparison (EP vs MP, MP vs LP, LP vs EA). Red and blue dots represent significantly up-regulated and down-regulated transcripts, respectively ($|\log_2 FC| \geq 1$ and FDR < 0.05). Venn diagrams illustrating the overlap of differentially expressed (B) mRNAs, (D) lncRNAs, and (F) circRNAs among the three pairwise comparisons.

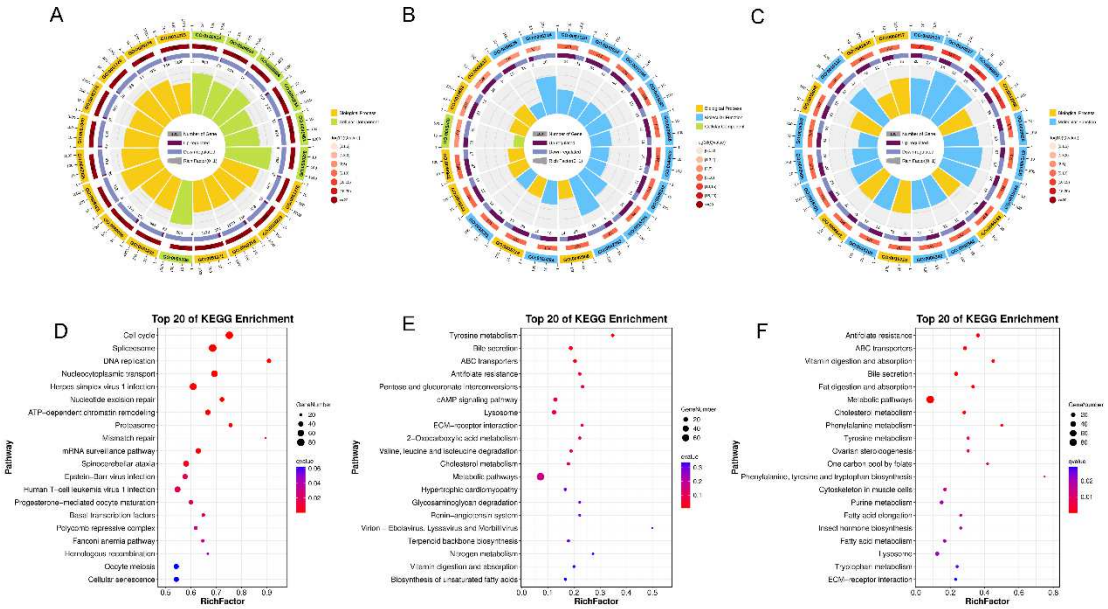


Figure 5. Functional enrichment analysis of differentially expressed genes across developmental stages.

GO enrichment circos plots for mRNA DEGs in (A) EP vs MP, (B) MP vs LP, and (C) LP vs EA comparisons. Outer ring: top 20 GO terms with coordinate scale representing gene counts; colors indicate ontology (BP, CC, MF). Second ring: bar length represents background gene count, with color gradient from red (lower Q-value) to gray (higher Q-value). Third ring: proportion of up-regulated (dark purple) and down-regulated (light purple) genes. Inner ring: Rich Factor (ratio of DEGs to total genes in each term), with grid lines at 0.1 intervals. KEGG enrichment bubble plots for mRNA DEGs in (D) EP vs MP, (E) MP vs LP, and (F) LP vs EA comparisons. Each bubble represents a pathway, with bubble size indicating gene count, color representing Q-value (red: lower Q-value), and x-axis showing Rich Factor.

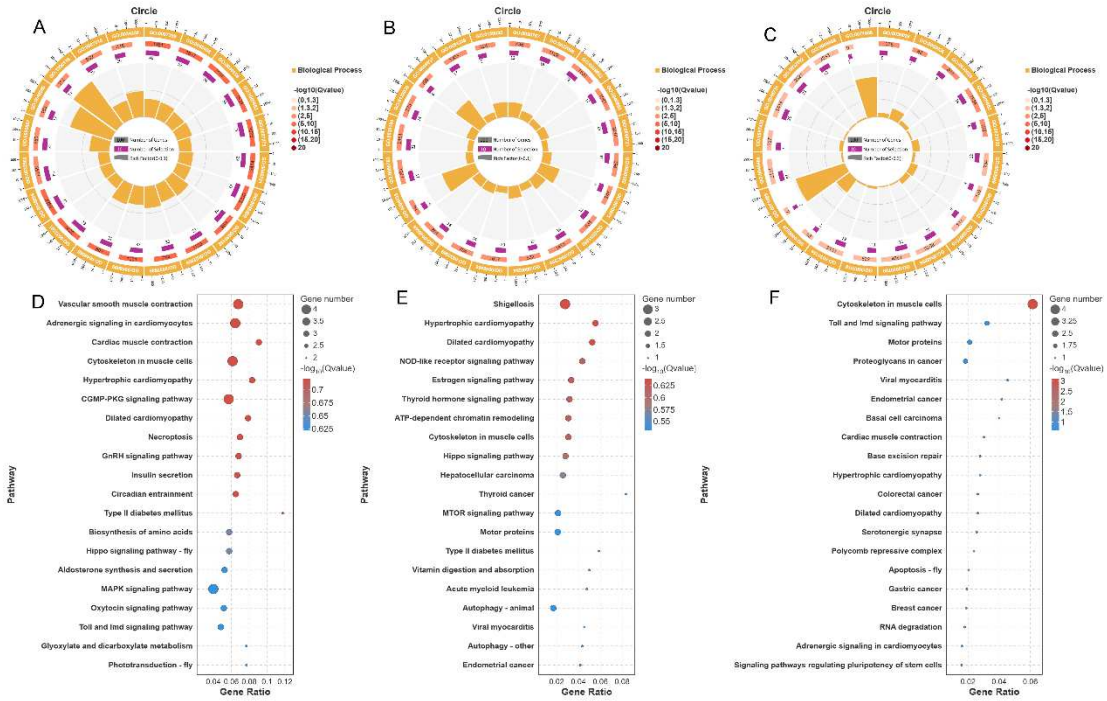


Figure 6. Functional enrichment of differentially expressed circRNA-associated genes across developmental transitions.

GO enrichment circle plots for (A) EP vs MP, (B) MP vs LP, and (C) LP vs EA comparisons, showing the top 20 enriched GO terms (first circle: scale of gene counts; second circle: term count and Q-value, with color gradient indicating significance; third circle: proportions of up- and down-regulated genes; fourth circle: Rich Factor with grid lines at 0.1 intervals). KEGG pathway enrichment bubble plots for (D) EP vs MP, (E) MP vs LP, and (F) LP vs EA, displaying the top 20 pathways (y-axis: pathway name; x-axis: Rich Factor; bubble size: gene count; color: Q-value).

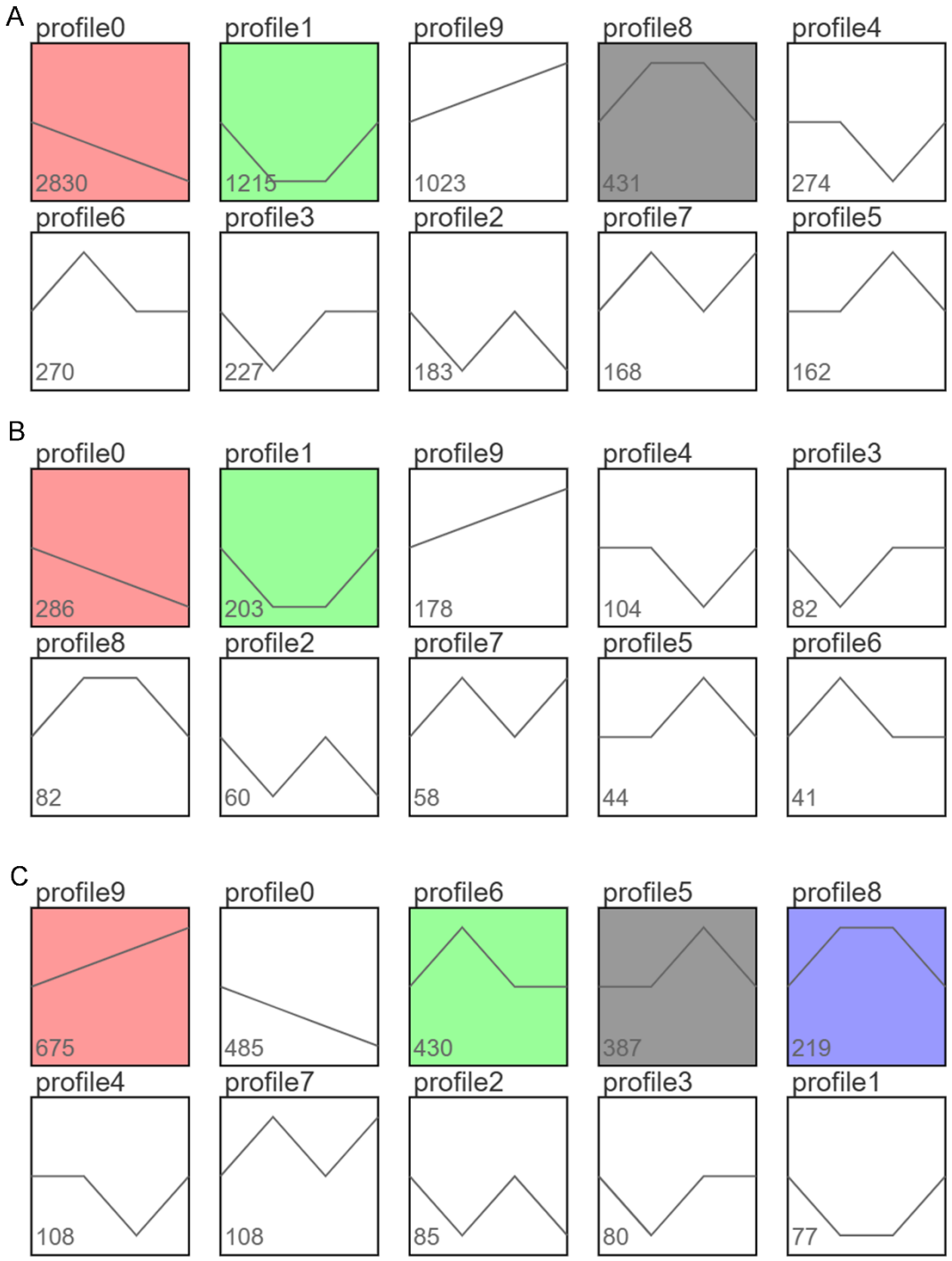


Figure 7. Temporal expression trends of (A) mRNAs, (B) lncRNAs, and (C) circRNAs across development.

Each profile shows standardized expression levels across stages (EP-MP-LP-EA); numbers indicate transcript counts assigned to each pattern. Only the major (colored) profiles with the highest counts are highlighted.

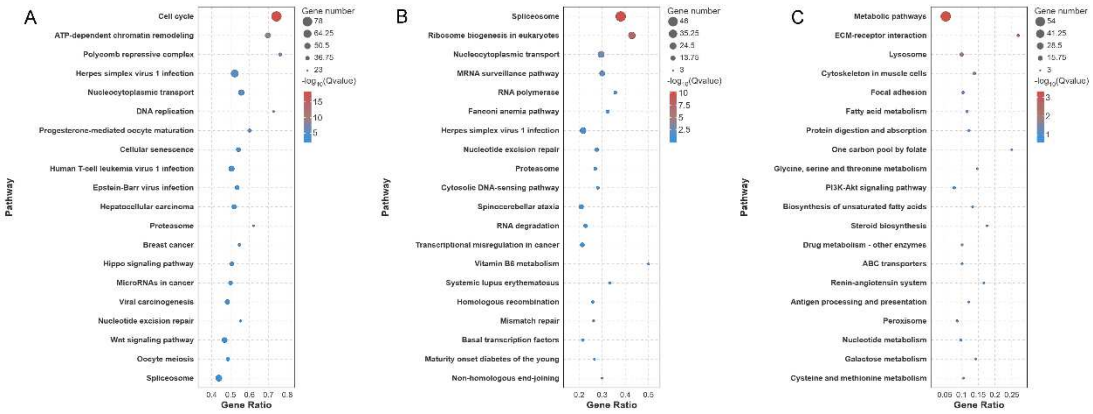


Figure 8. KEGG pathway enrichment of major mRNA expression profiles.

Bubble plots showing enriched pathways for (A) profile0 (stable expression), (B) profile1 (progressive decrease), and (C) profile8 (peak at MP). The x-axis represents the gene ratio (number of genes in the pathway divided by total genes in the profile), bubble size indicates the number of genes, and color intensity represents the Q-value (red: lower Q-value).

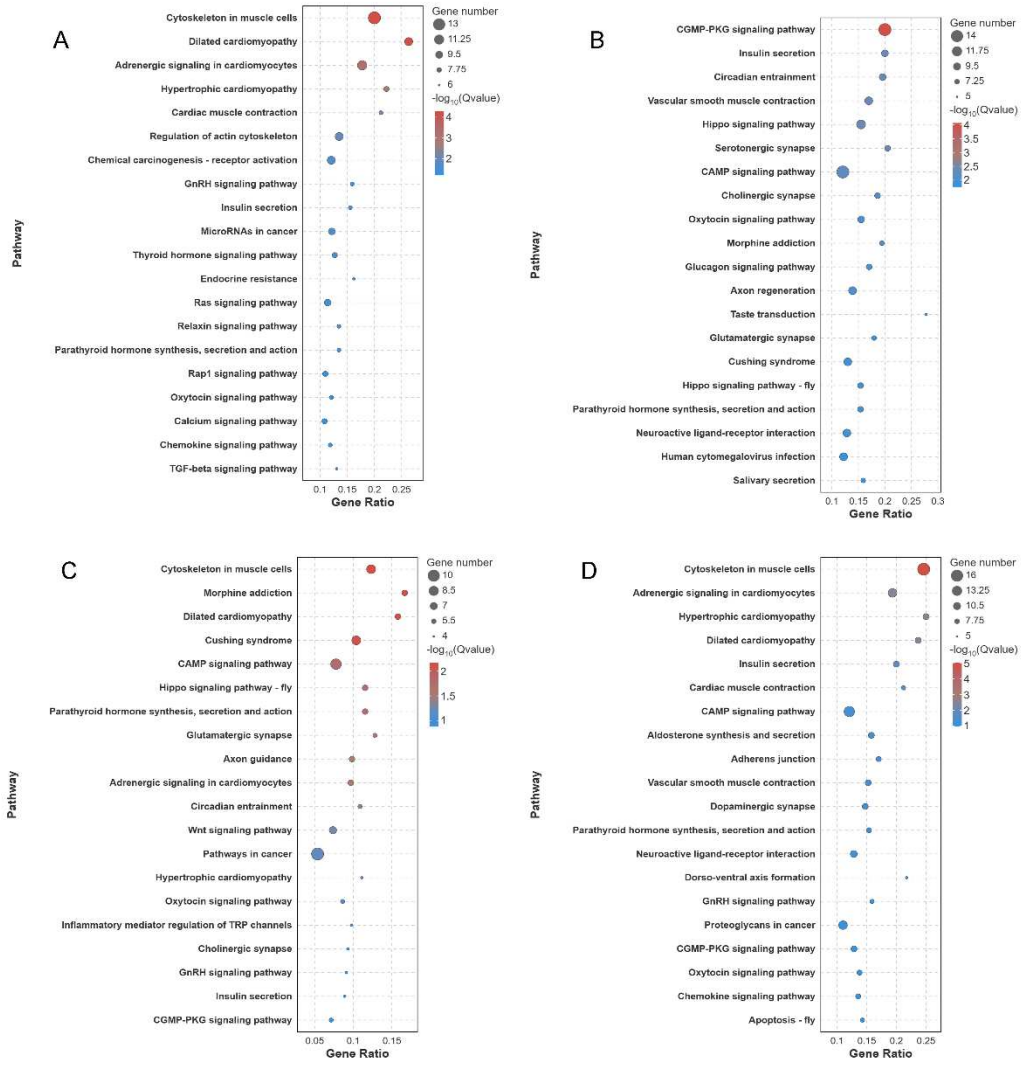


Figure 9. KEGG pathway enrichment of major circRNA expression profiles.

Bubble plots showing enriched pathways for (A) profile5 (gradual decrease), (B) profile6 (peak at MP), (C) profile8 (valley at MP), and (D) profile9 (progressive increase). The x-axis represents the gene ratio, bubble size indicates the number of genes, and color intensity represents the Q-value (red: lower Q-value).

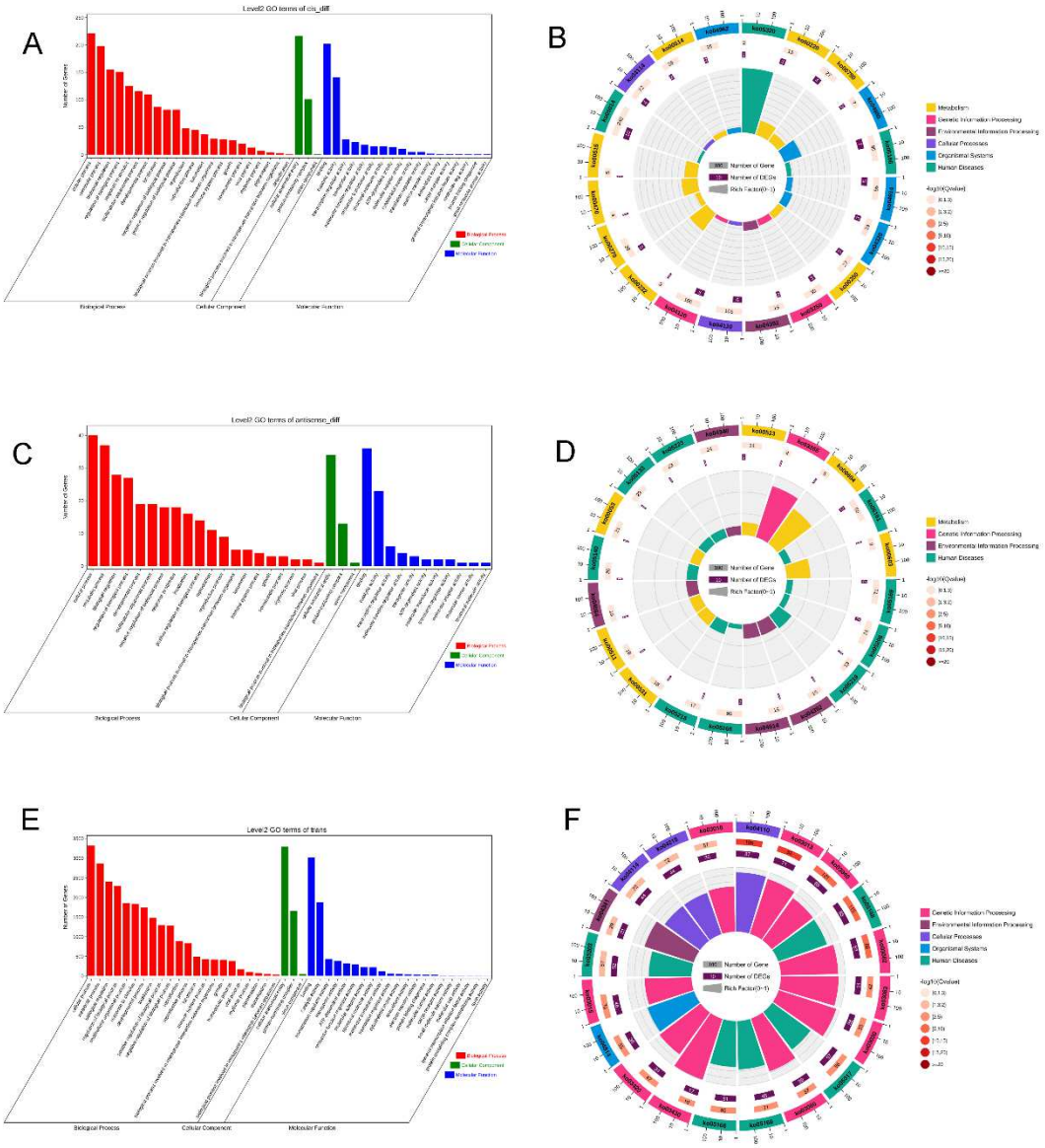


Figure 10. Functional enrichment of lncRNA target genes predicted by different regulatory modes. GO level-2 classification bar plots for (A) cis targets, (C) antisense targets, and (E) trans targets, showing the number of genes associated with each GO term (Biological Process, Cellular Component, Molecular Function). KEGG enrichment circle plots for (B) cis targets, (D) antisense targets, and (F) trans targets, displaying enriched pathways with their significance and gene counts.

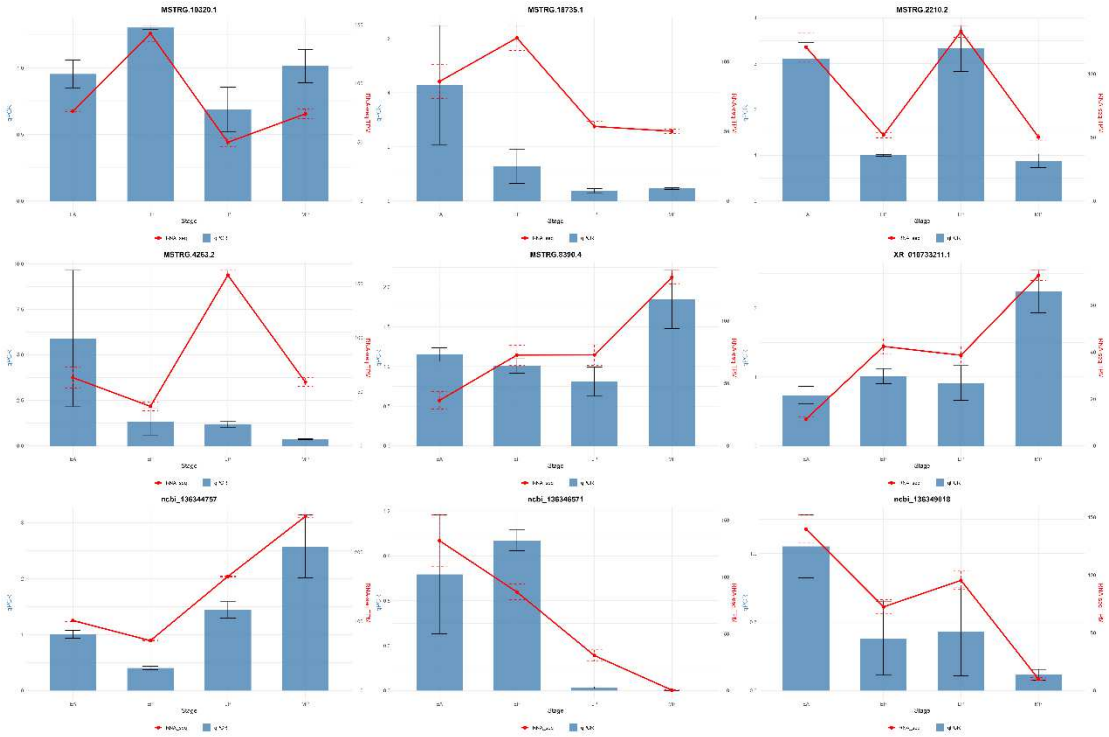


Figure 11 Comparative validation of target gene expression by qPCR and RNA-seq. The left Y-axis represents the relative expression level quantified by qPCR. The right Y-axis shows the transcripts per million (TPM) values derived from RNA-seq data. Error bars indicate the standard error of the mean.

880 Table 1 Transcriptome sequencing and assembly quality of *E. interjectus.*

Sample	Raw Reads	Clean Reads	Clean ReadsRate (%)	Clean Q20 BasesRate (%)	Clean Q30 BasesRate (%)	Clean GCpercent (%)
EA-1	11481849600	11244232804	97.93%	99.28%	96.87%	43.49%
EA-2	9571396500	9350192077	97.69%	99.23%	96.72%	44.02%
EA-3	9420516300	9297423252	98.69%	99.10%	96.17%	43.37%
EP-1	9008378100	8880036204	98.58%	99.16%	96.30%	43.22%
EP -2	10963824300	10798529517	98.49%	99.08%	96.10%	42.83%
EP -3	10138582800	9995055010	98.58%	99.01%	95.85%	42.97%
LP-1	10702058700	10501740827	98.13%	99.23%	96.65%	43.19%
LP-2	11254971900	11110326315	98.71%	99.02%	95.95%	43.15%
LP-3	10355799600	10228721729	98.77%	99.10%	96.13%	43.68%
MP-1	10227514200	10056030533	98.32%	99.14%	96.29%	42.96%
MP-2	9956481600	9778328721	98.21%	99.13%	96.36%	43.26%
MP-3	9908627700	9713101489	98.03%	99.20%	96.53%	42.96%

881

882

Supplementary Files

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

- SupplementaryMaterials.zip