


RESEARCH ARTICLE

Open Access



Genome-wide analysis reveals transcriptional and translational changes during diapause of the Asian corn borer (*Ostrinia furnacalis*)

Xingzhuo Yang^{1†}, Xianguo Zhao^{1†}, Zhangwu Zhao¹ and Juan Du^{1*} 

Abstract

Background Diapause, a pivotal phase in the insect life cycle, enables survival during harsh environmental conditions. Unraveling the gene expression profiles of the diapause process helps uncover the molecular mechanisms that underlying diapause, which is crucial for understanding physiological adaptations. In this study, we utilize RNA-seq and Ribo-seq data to examine differentially expressed genes (DEGs) and translational efficiency during diapause of Asian corn borer (*Ostrinia furnacalis*, ACB).

Results Our results unveil genes classified as “forwarded”, “exclusive”, “intensified”, or “buffered” during diapause, shedding light on their transcription and translation regulation patterns. Furthermore, we explore the landscape of lncRNAs (long non-coding RNAs) during diapause and identify differentially expressed lncRNAs, suggesting their roles in diapause regulation. Comparative analysis of different types of diapause in insects uncovers shared and unique KEGG pathways. While shared pathways highlight energy balance, exclusive pathways in the ACB larvae indicate insect-specific adaptations related to nutrient utilization and stress response. Interestingly, our study also reveals dynamic changes in the HSP70 gene family and proteasome pathway during diapause. Manipulating HSP protein levels and proteasome pathway by HSP activator or inhibitor and proteasome inhibitor affects diapause, indicating their vital role in the process.

Conclusions In summary, these findings enhance our knowledge of how insects navigate challenging conditions through intricate molecular mechanisms.

Keywords *Ostrinia furnacalis*, RNA-seq, Ribo-seq, Diapause, lncRNAs

Background

Diapause is an adaptation in the developmental process that enable survival under adverse conditions, offering insects selective advantages in various ways. It serves as a method of cold tolerance for insects undergoing winter diapause [1]. Additionally, diapause plays a crucial role in the storage of metabolites such as lipids, before the onset of winter. Moreover, the termination of diapause in the spring leads to synchrony within insect populations. Studying diapause has the potential to provide us with proper targets for further insect control development.

[†]Xingzhuo Yang and Xianguo Zhao contributed equally to this work.

*Correspondence:

Juan Du
dujuan9981@cau.edu.cn

¹ Department of Entomology, MOA Key Lab of Pest Monitoring and Green Management, College of Plant Protection, China Agricultural University, Beijing 100193, China



Insect diapause is a programmed dormancy predetermined by environmental cues and not immediately terminated in response to changes in the environment. It can be classified into different types based on the developmental stage in which it occurs: embryonic diapause (found in *Bombyx mori*), larval diapause (found in *Ostrinia furnacalis*), pupal diapause (found in *Helicoverpa armigera*), and adult diapause (found in *Culex pipiens*). By comparing the regulatory mechanisms of these different types of diapause, we can gain a better understanding of the molecular mechanisms involved in diapause control and developmental regulation.

The process of diapause is influenced by a combination of environmental cues and genetic factors. Studies have shown that the environmental factors influencing diapause include photoperiod and temperature. Genetic factors that affect diapause include circadian genes [2, 3], endocrine factors [4], epigenetic factors [5, 6], small non-coding RNAs [7, 8], and heat shock proteins [9, 10]. RNA-seq experiments were used to dissect the DEGs in stages of diapause [11]. While mechanisms governing the induction and regulation of diapause have been identified, several significant gaps in our knowledge persist. Notably, there is a dearth of research conducted at the translational level, despite the acknowledged importance of both transcriptional and post-transcriptional processes in modulating gene expression. This raises the intriguing question of the extent to which translational control plays a pivotal role in diapause. Moreover, the involvement of long non-coding RNAs (lncRNAs) in diapause regulation remains unexplored, as no such lncRNAs have been identified to date. Increasing evidence indicates the importance of lncRNAs in gene expression regulation. The discovery and characterization of differentially expressed lncRNAs across various diapause stages hold considerable promise for shedding light on the intricate regulatory network that governs diapause.

Diapause is under endogenous hormonal control. For example, juvenile hormone (JH) levels undergo dynamic changes during different stages of larval diapause. The upregulation of JH at the photosensitive pre-diapause stage plays a pivotal role in larval diapause in various insects, including the beet webworm (*Loxostege sticticalis* L.) [12–14], *Ostrinia nubilalis* [15], and *Laspeyresia pomonella* [16]. On the other hand, species undergoing stationary molts maintain a high JH titer during diapause. For example, a high JH titer has been reported for larval diapause in the Mediterranean corn borer, *Sesamia nonagrioides* [17–19], and the yellow-spotted longicorn beetle, *Psacothaea hilaris* [20]. Application of JH can trigger the termination of diapause, which was likely caused by its effect in stimulating the production of ecdysone [13, 21, 22]. It is an intriguing problem that to understand the

mechanism regulating JH production during diapause process.

Insects regulate the synthesis of heat shock proteins (HSPs) during normal growth, exposure to stress, and entry into diapause. HSPs are also known as stress proteins and molecular chaperones. During diapause and in response to temperature extremes [23], crowding [24], starvation [25], and hypoxia/anoxia [26, 27], insects produce HSPs in these different contexts to potentially counteract these stresses. HSPs display distinct patterns during different stages of diapause in insects, including in insect embryos [28], diapausing larvae [29], and prepupae and pupae [10], as well as during adult/reproductive diapause [30]. For example, in diapause of *Chilo suppressalis*, certain Hsps undergo dynamic changes, such as Hsp90, Hsp70, and Hsp60. Specifically, their expression is low in the early diapause stage, increases during the mid-diapause stage, and decreases in the late diapause stage [31]. This evidence suggested the importance of HSPs in diapause regulation. However, the detailed mechanism of their function in diapause needs further study.

In this study, we utilized RNA-seq and Ribo-seq data to examine differentially expressed genes (DEGs) and translational efficiency during diapause. We sampled the head of ACB due to the importance of the brain in the control of diapause [3]. The results revealed comprehensive differential gene expression at both transcriptional and translational levels. Furthermore, we explored the landscape of long non-coding RNAs (lncRNAs) during diapause and identified differentially expressed lncRNAs, suggesting their roles in diapause regulation. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses of lncRNA targets highlighted the role of lncRNAs in the regulation of the JH pathway. More importantly, a comparative analysis of diapause in different insects uncovered shared and unique KEGG pathways, emphasizing conserved metabolic processes during diapause. Interestingly, our study reveals dynamic changes in the HSP70 gene family and the Proteasome pathway during diapause, suggesting their regulatory roles. HSP70 genes exhibit a distinct pattern of downregulation followed by upregulation during diapause, impacting the diapause rate. Manipulating HSP protein levels and the Proteasome pathway affects diapause, indicating their vital roles in the process.

Results

RNA-seq of ACB head at different stages of diapause revealed a DEGs profile at transcriptional level

To better understand genome-wide changes of transcription and translation in ACB at various stages of diapause, we performed Ribo-seq and RNA-seq to

identify gene expressions in ACB at various stages of diapause (Additional file 1: Fig. S1A). After quality control, the RNA-seq had an average Q20 score above 97% and an average Q30 score above 92% (Supplementary Table 5). The average alignment rate was 85%, indicating that we obtained sufficient and high-quality data for subsequent analyses. First, we performed RNA-seq to detect differentially expressed genes at the transcriptional level. The gene density plots of the nine samples used in RNA-seq indicated that the main peak positions in each state are roughly the same, with the pre-diapause state having the most concentrated gene expression density distribution (Additional file 1: Fig. S1B). The principal component analysis (PCA) of the RNA-seq data showed a clear separation between groups (Additional file 1: Fig. S1C). The reliability of the data performance was evaluated through the repeatability study represented by the heat map, which showed the high correlation among the three biological replicates of each sample (Additional file 1: Fig. S1D). Transcriptome alterations were determined through differential expression analysis of RNA-seq data. Comparative analysis of the transcriptome of pre-diapause, diapause, and no-diapause samples of ACB showed DEGs between these samples. Comparing with the no-diapause sample, there are 974 significantly changed genes in the diapause sample, with comparable numbers of genes with increased abundance and genes with decreased abundance (Fig. 1A, D). When compared to pre-diapause period, there are 1455 DEGs in the diapause sample, with more than twice as many at genes with decreased abundance compared with genes with increased abundance (983 vs. 472) at diapause period (Fig. 1B, D). Comparing with the no-diapause sample, there are 201 significantly changed genes in the pre-diapause sample ($\log_2(\text{fold change})$, adjusted p -value < 0.05), including 168 genes with increased and 33 with decreased abundance (Fig. 1C, D). Comparative analysis showed that

35 genes were significantly changed between every two of the three samples (Fig. 1D).

To understand the biological implications underlying these differences, GO enrichment and KEGG enrichment analysis were carried out. GO enrichment analysis showed that the mitosis related term (such as mitotic spindle organization, microtubule cytoskeleton organization involved in mitosis, spindle midzone), negative chemotaxis, and synaptic target inhibition were specifically enriched in DEGs between diapause and no-diapause samples. The mitochondrial energy metabolism related terms (such as generation of precursor metabolites and energy, ATP metabolic process, mitochondrial protein-containing complex) were specifically enriched in DEGs between diapause and pre-diapause samples. It is interesting to note that the cuticle development, chitin-based cuticle development, and organic acid biosynthetic process were all enriched between each pair of the three samples (Fig. 1E). Specifically, 34, 52, and 13 genes were enriched in the cuticle development process in the comparisons between diapause vs. no-diapause, and diapause vs. pre-diapause and pre-diapause vs. no-diapause, respectively. In the comparisons between diapause and no-diapause, as well as between diapause and pre-diapause, the majority of these genes showed significantly decreased expression. Conversely, in the comparison between pre-diapause and no-diapause, most of these genes (495 of 947 in diapause vs no-diapause; 983 of 1455 in diapause vs pre-diapause; 33 of 210 in pre-diapause vs no-diapause) exhibited significantly increased expression (Additional files 3 and 4, Additional file 1: Fig. S2A–C). The KEGG enrichment analysis indicated that DNA replication, FoxO signaling pathway, and Wnt signaling pathway were specifically enriched between diapause and no-diapause samples. Purine metabolism and Insect hormone biosynthesis was enriched between each pair of the three samples (Fig. 1F). Specifically, 6, 11, and 5 differentially expressed genes were significantly enriched in the insect hormone biosynthesis process in the comparisons

(See figure on next page.)

Fig. 1 RNA-seq revealed gene profiles that are changed at transcriptional level. **A–C** Volcano plot representing the DEGs of RNA-seq candidates with the threshold of $|\log_2(\text{fold change})| \geq 1$ and adjusted p -value < 0.05 were regarded as significant differentially expressed genes. The genes with significant increased and decreased abundance were marked by orange and blue, respectively. **D** The upset plot represents the intersection of significantly differentially expressed genes identified in the RNA-seq analysis across three comparisons for **A–C**. The bars indicate the number of significantly differentially expressed genes in each intersection set, with the horizontal bars showing the total number of significantly differentially expressed genes for each comparison. **E, F** Compare GO and KEGG enrichment analysis of DEGs revealed by RNA-seq. The circle color indicated adjusted p -value of the enrichment and the dot size indicated the number of DEGs in the functional class or pathway. **G** Clustering and functional enrichment analysis of genome-wide gene expression at transcriptional levels. The heatmap show the eight profiles of the clustered genes, with the number of genes for each cluster were noted in the box just right to the heatmap. The genes in each cluster were presented in the heat map, which were normalized by Z-score. The representative trends of the expression of each cluster were shown in the mid box. The top five GO enrichment term of biological progress and KEGG enrichment pathway are shown to the right of each cluster. The text size was changed according to the p -value (the larger the font, the more significant the pathway)

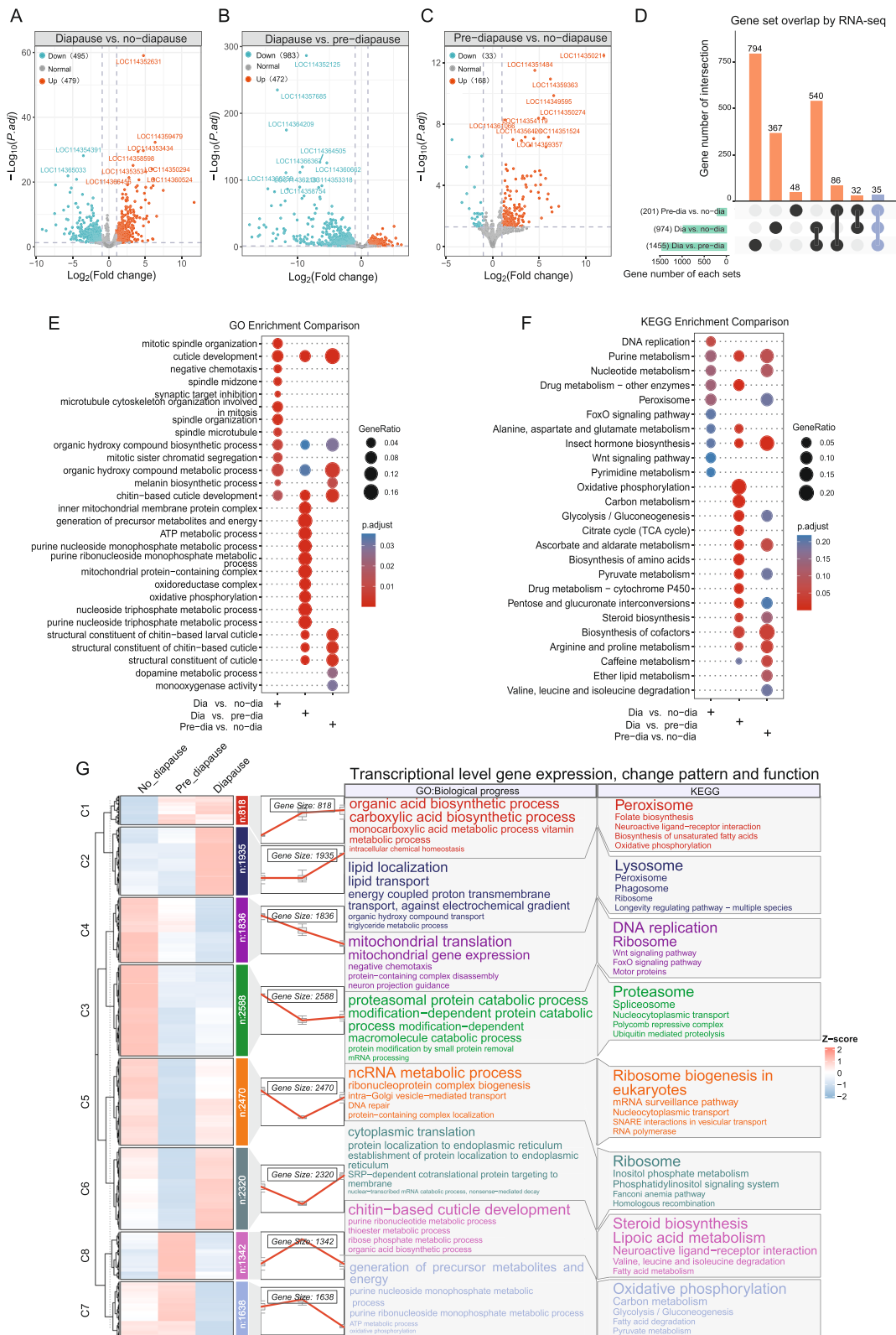


Fig. 1 (See legend on previous page.)

between diapause vs. no-diapause, diapause vs. pre-diapause, and pre-diapause vs. no-diapause, respectively. The majority of these genes exhibited decreased expression (Additional files 3 and 4, Additional file 1: Fig. S3A-C). All the gene lists used for analysis were included in Additional file 4. The specific gene lists of the enrichment analysis were included in Additional file 3.

To gain deeper insights into the temporal expression patterns of genes across the three time points represented by the samples in our study, we performed the gene temporal expression analysis by clustering and visualizing the time series of the gene expression data from the RNA-seq. The fuzzy K-means cluster method from Mfuzz package was used to obtain appropriate 8 clusters (Fig. 1G). For cluster 1, there were 818 genes tended to be up regulated from no diapause to pre-diapause to diapause and were primarily enriched in the organic acid biosynthetic process, peroxisome. A similar trend was found in cluster 2. Clusters 5 and 6 exhibit a common gene expression trend characterized by an initial decrease followed by a gradual increase. The genes significantly enriched in these clusters are mainly associated with biological processes including non-coding RNA metabolism, intracellular molecular biology of DNA, RNA, and protein synthesis, transport, repair, and degradation. In cluster 3, there is an initial and substantial decrease in gene expression levels, followed by a sustained, unchanging pattern. These genes are associated with cellular pathways related to proteasomes and ubiquitin-mediated degradation. GO enrichment highlights proteasomal protein catabolic process, protein modification by small protein removal, and mRNA processing. KEGG enrichment includes proteasome, spliceosome, nucleocytoplasmic transport, polycomb repressive complex, and ubiquitin-mediated proteolysis pathways. These suggest roles in proteolysis, protein modification, and RNA processing. We also found that transcript abundance of genes in clusters 7 and 8 initially increases and then decreases. GO enrichment analysis for genes in cluster 7 shows involvement in processes such as chitin-based cuticle development, purine metabolism, and various metabolic processes. KEGG enrichment analysis indicates these genes are involved in pathways related to steroid biosynthesis, lipoic acid metabolism, neuroactive ligand-receptor interaction, amino acid degradation, and fatty acid metabolism. The transcript abundance of genes in cluster 8 initially increases and then decreases. GO enrichment analysis shows that these genes are involved in energy production and metabolism processes, including ATP metabolism and oxidative phosphorylation. KEGG enrichment analysis indicates involvement in major metabolic pathways such as oxidative phosphorylation, carbon metabolism, and fatty acid degradation.

These results suggest that the genes in cluster 8 play a crucial role in energy production and overall metabolic regulation.

Similarly, we employed Gene Set Enrichment Analysis (GSEA) to investigate RNA-seq data from different diapause stages. Consistently, our analysis revealed the activation and suppression of pathways related to many cellular processes during the transition from the no-diapause state to diapause (Additional file 1: Fig. S4A-R). This indicates a concerted downregulation of these pathways as part of the diapause onset process. We found that numerous gene groups, including those related to cuticle development and proteasomes exhibit significant dynamic changes (Additional file 1: Fig. S4A-R). For instance, pathways related to cuticle development, oxidative phosphorylation, and glycolysis/gluconeogenesis were significantly activated during the pre-diapause stage, with a marked decrease of many core genes compared to the no-diapause state. However, when transitioning to the full diapause stage, these pathways were significantly inhibited (Additional file 1: Fig. S4A, E, F, G, H, K, L), with a substantial decrease of the core genes relative to the no-diapause state. This trend is also evident when compared to the pre-diapause stage (Additional file 1: Fig. S4M, N, Q, R). This indicates that the cuticle development, oxidative phosphorylation, and glycolysis/gluconeogenesis pathways exhibit a similar dynamic pattern: significant increase followed by a decrease, with the final levels being lower than the no-diapause state. Other pathways, such as the proteasome (Additional file 1: Fig. S4J, P) and mitochondrial energy metabolism-related pathways (Additional file 1: Fig. S4C), were also significantly enriched, corroborating the previous findings from the differentially expressed genes' GO/KEGG analysis and the clustering results.

These gene dynamics are consistent with the trends observed in the clustering profiles. Observing these trends from multiple perspectives, including significantly differentially expressed genes and dynamically changing genes, validates the accuracy of the observed changes.

Ribo-seq of ACB head at different stages around diapause revealed a DEGs profile at translational level

To understand translation changes during diapause, we used Ribo-seq to identify ribosome-protected footprints (RPFs). The Ribo-seq data had an average Q20 score above 99% and an average Q30 score above 97%, indicating that we obtained high-quality data for subsequent analyses (Additional file 4). Analysis showed the dominant RPF length was 28 nucleotides (nt) across all stages (Additional file 1: Fig. S5A-C). These 28-nt RPFs had higher counts in the dominant reading frame, ensuring significant 3-nt periodicity, a key quality control measure

(Additional file 1: Fig. S5A-C). Metagene profiles of RPF counts near translation initiation (TIS) and termination sites (TTS) confirmed proper P-site estimation. During diapause stages, read length distribution consistently showed 28-nt RPFs (Additional file 1: Fig. S5A-D). Pre-diapause and diapause stages had broader RPF length distributions, suggesting ribosomal differences. P-site analysis indicated that 28-nt RPFs predominantly occupied frame 0 of the CDS, over 95%, showing preferential translation within the CDS (Additional file 1: Fig. S5E). Lower P-site signals in 5' and 3' UTRs suggested weaker translation activity in these regions (Additional file 1: Fig. S5F).

We compared ORFs and found that over 90% annotated across all conditions. Actively translated ORFs were more common during no-diapause and diapause stages than pre-diapause (Additional file 1: Fig. S5G, H). Meta-gene heatmaps showed distinct three-nucleotide periodicity, emphasizing precise ribosomal positioning (Additional file 1: Fig. S5G). This highlighted dynamic changes in ribosomal activity during diapause, affecting protein synthesis regulation (Additional file 1: Fig. S5I, J).

Comparative analysis of the translation levels of pre-diapause, diapause and no-diapause samples of ACB also showed DEGs between these samples (Fig. 2A, B, C). Compared with the no-diapause sample, there are 383 significantly changed genes in the pre-diapause sample ($\log_2(\text{fold change}) \geq 1$, adjusted p -value < 0.05), including 297 genes with increased abundance and 86 down-regulated genes with decreased abundance (Fig. 2D). Comparing with the no-diapause sample, there are 1769 significantly changed genes in the diapause sample. When compared to pre-diapause period, there are 1538 DEGs in the diapause sample, the translation of most genes was decreased (647 genes with increased abundance vs. 891 genes with decreased abundance) at diapause period (Fig. 2D). Comparative analysis showed that 134 genes were significantly changed between every two of the three samples (Fig. 2D). To understand the

biological processes these genes were involved in, we performed GO and KEGG enrichment analyses. The gene lists used for analysis were included in Additional file 4. The specific gene lists of the enrichment analysis were included in Additional file 3. GO enrichment analysis showed a bunch of terms that were similarly enriched in the three samples. It mainly involves cuticle development and metabolism as well as metabolic processes such as ribonucleotide and amino acids metabolism etc. (Fig. 2E). The KEGG analysis showed that these genes were enriched in different pathways. For example, translation-related genes were mainly enriched in oxidative phosphorylation, Toll and Imd signaling pathway, glycolysis and gluconeogenesis, lysosome, and purine metabolism, when comparing the diapause sample to the no-diapause or pre-diapause samples (Fig. 2F). When compared, the no-diapause and pre-diapause samples, the pathways of drug metabolism, biosynthesis of nucleotide sugars, pyrimidine metabolism, etc., are enriched. The three pathways of glycine, serine and threonine metabolism, amino sugar and nucleotide sugar metabolism, and ascorbate and aldarate metabolism are enriched in differentially translated genes between every two of these three gene lists (Fig. 2F). These results indicate that the translational profiles changed during the different stages of diapause.

We also conducted a cluster analysis of the change patterns of translation levels genes in the three time points represented by the three samples (Fig. 2G). With the development of diapause, the expression of 1300 genes in cluster1 decreased at the pre-diapause stage and further decreased during diapause. Enrichment analysis showed that these pathways involve various metabolic processes, lipid synthesis, subcellular functions, molecular transport, and neural signal transduction during diapause. Similar patterns of change were observed in 1922 genes associated with encompassing various aspects of energy production, organic molecule synthesis, breakdown, and cellular

(See figure on next page.)

Fig. 2 Ribo-seq revealed gene profile that are changed at translation level. **A–C** Volcano plot representing the DEGs of Ribo-seq Candidates that satisfied the criteria of $|\log_2(\text{fold change})| \geq 0.265$ and adjusted p -value < 0.05 were regarded as significantly expressed genes, which are marked by orange and blue for up and downregulation, respectively. **D** The upset plot represents the intersection of significantly differentially expressed genes identified in the Ribo-seq analysis across three comparisons for **A–C**. The bars indicate the number of significantly differentially expressed genes in each intersection set, with the horizontal bars showing the total number of significantly differentially expressed genes for each comparison. **E, F** Compare GO and KEGG enrichment analysis of DEGs in revealed by Ribo-seq. The circle color indicates the enrichment adjusted p -value, and the dot size indicates the number of DEGs in the functional class or pathway. **G** Clustering and functional enrichment analysis of genome-wide gene expression at translational levels. The heatmap show the eight profiles of the clustered genes, and the number of genes for each cluster were noted in the box just right to the heat map. The genes in each cluster were presented in the heat map, which were normalized by Z-score. The representative trends of the expression of each cluster are shown in the mid box. The top five GO enrichment term of biological progress and KEGG enrichment pathway are shown to the right of each cluster. The text size corresponds to the p -value (the larger the font, the more significant the pathway)

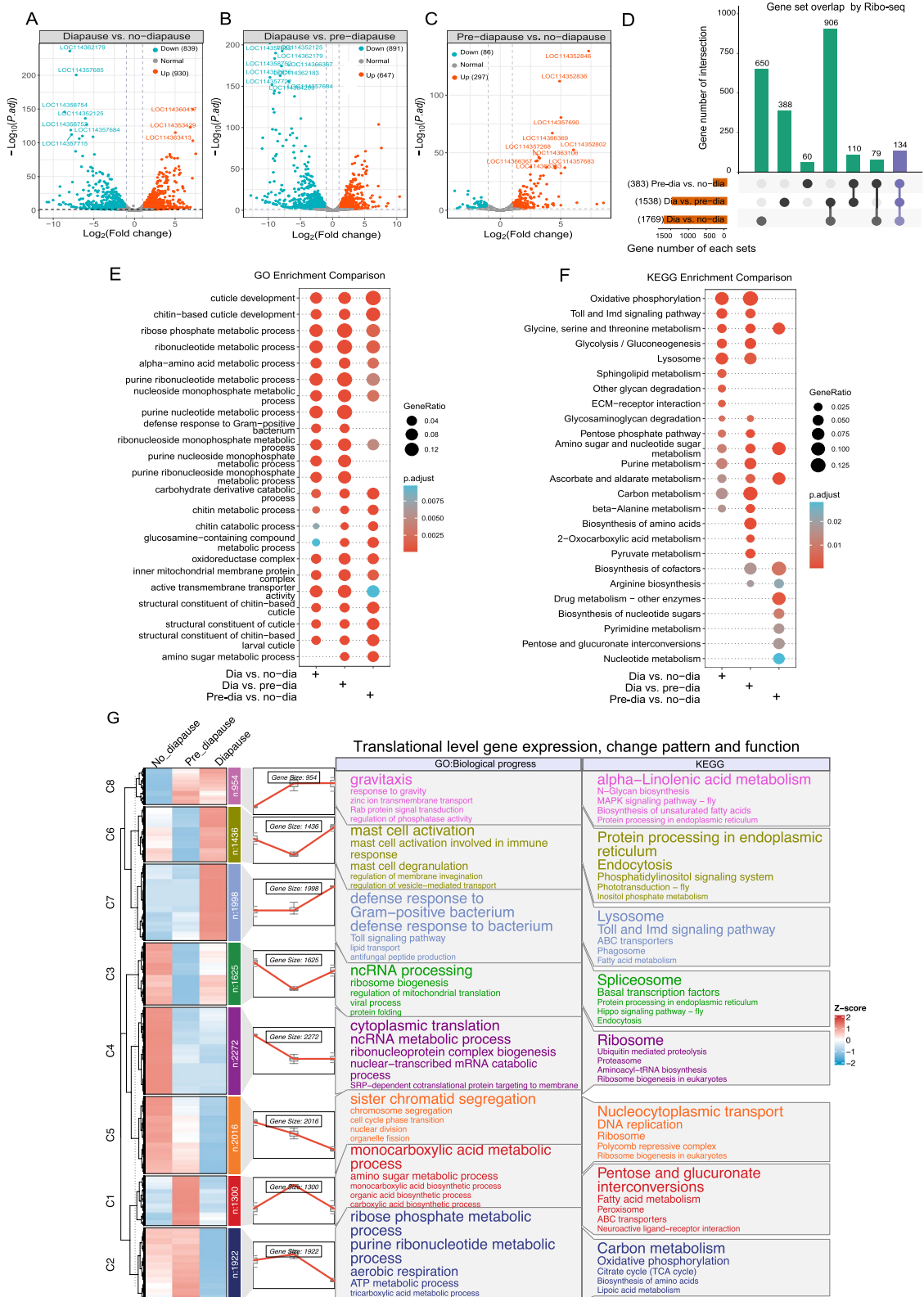


Fig. 2 (See legend on previous page.)

functions in cluster 2. There are also two gene sets with opposite change trends. For example, genes are primarily involved in various vital biological functions, including protein processing, cellular uptake, signal transduction, photoreception, metabolism, immune response, intracellular molecular transport, and membrane processes. These genes low expressed in pre-diapause process but are highly expressed in no-diapause and diapause process in cluster 6. Cluster 3 exhibits a similar trend. Additionally, many clusters with similar dynamic trends and enrichment analysis results show a high degree of similarity with the previous RNA-seq results, such as cluster 3 (Fig. 2G) and cluster 5 (Fig. 1G).

Combinational analysis of the RNA-seq and Ribo-seq data revealed differences in translational efficiency between different samples

Translation efficiency (TE) is an important aspect of translation process, which standardizes the change in translation level according to transcription level. The genes with significant RPF and mRNA but not significant TE belong to “Forwarded”. A total of 834 genes were “Forwarded” genes when comparing diapause vs. no-diapause samples (Fig. 3A). Similarly, diapause vs. pre-diapause and pre-diapause vs. no diapause had 956 and 102 such genes, respectively (Fig. 3B, C). The second group of genes was classified as “Exclusive”, whose change in RPF was not driven by changes in mRNA (yellow in Fig. 3A–C). Therefore, for these genes TE and RPF had significant

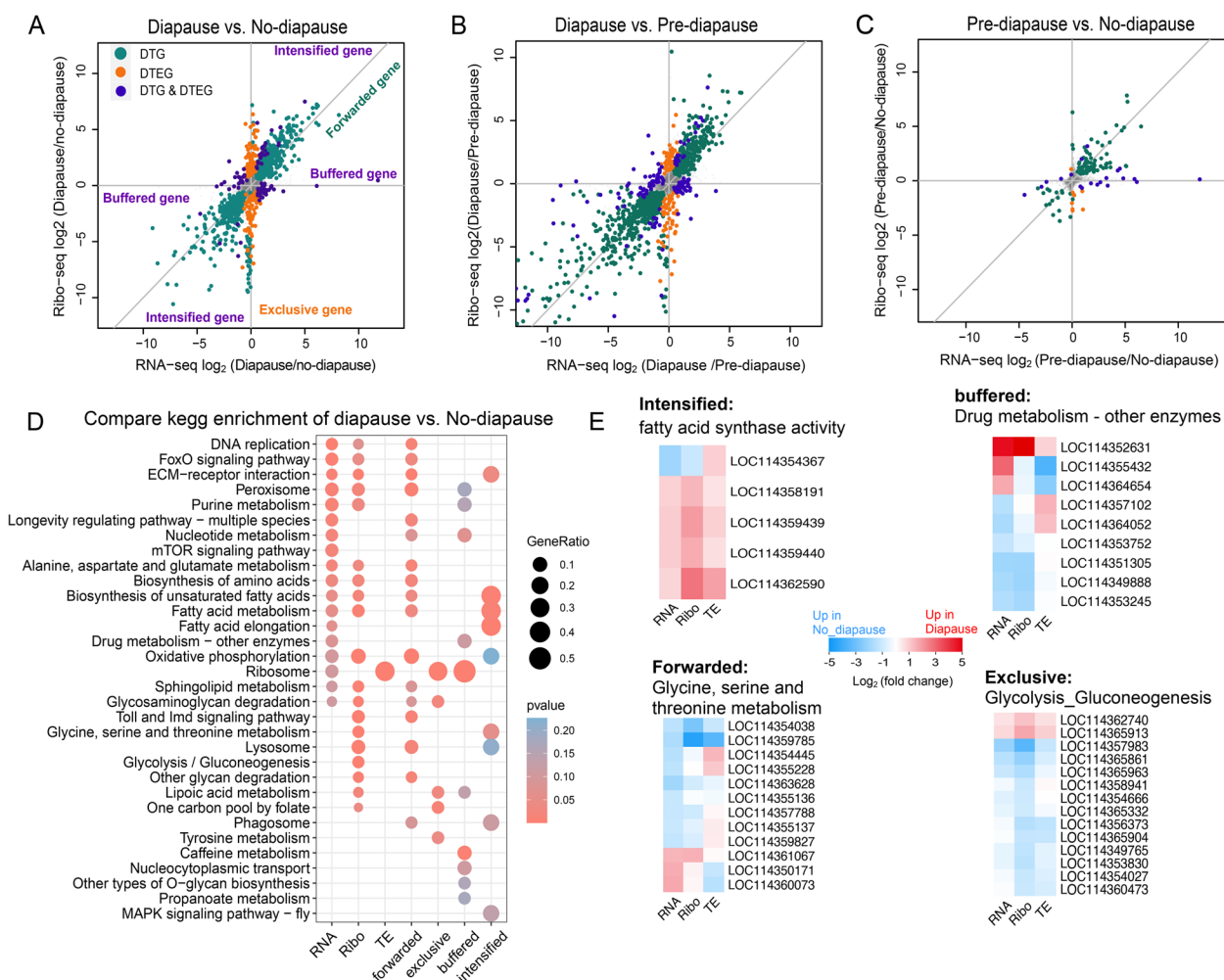


Fig. 3 Genome-wide transcriptional and translational regulation at different stages around diapause. **A–C** Scatter plot of fold changes between different stage of diapause for all ORFs in Ribo-seq data and the corresponding gene in RNA-seq data. The genes with DTGs and DTEGs are displayed. **D** Dot plot of the top enriched KEGG pathway in each regulatory category defined in **A**. The circle color indicates the enrichment adjusted *p*-value, and the dot size indicates the generation in the pathway. **E** Heat map of genes associated with the top KEGG enrichment pathway in each regulatory category identified in **A**

changes, but mRNA changes were not significant, which were called differentially translation efficiency genes (DTEGs). There were 155, 103, and 8 in each of the three comparisons (Fig. 3A–C). For example, we found glycolysis/gluconeogenesis related genes showed significantly decreased abundance only at translational level between the diapause and no-diapause by comparative enrichment analysis (Fig. 3D, E bottom right). The third group of genes was classified as “Intensified” that were regulated by both transcription and translation (significantly changed in mRNA, RPFs, and TE) (purple in Fig. 3A–C). In this case, mRNA changes acted with the change in TE. There were 28, 34, and 2 in each of the three comparisons. In this regard, we found ORFs encoding fatty acid synthase activity increases or decreases in both mRNA levels and translational efficiency between the diapause and no-diapause (Fig. 3D, E top left). Finally, the fourth group of genes was classified as “Buffered” whose transcriptional changes were offset by TE changes (blue in Fig. 3A–C). The transcriptional change and TE change of these genes were in the opposite direction. Thus, genes with significant TE and mRNA but not significant RPF are also considered translation buffered. There were 81, 141, and 18 genes in the three comparisons (Fig. 3A–C). In this regard, we found, for example, that developmental decreases in drug metabolism gene RNA levels were effectively buffered by corresponding increases in TE between the diapause and no-diapause (Fig. 3D, E top right).

We also performed the GO enrichment analysis for genes categorized by transcription, translation, translation efficiency, and deltaTE differences between diapause and no-diapause (Additional file 1: Fig. S6A).

Furthermore, we conducted functional enrichment analysis for differential expressed genes between diapause and pre-diapause (Additional file 1: Fig. S6B–C) and between pre-diapause and no-diapause (Additional file 1: Fig. S6D–E). Through comparative analysis, we identified several common and unique terms enriched in various gene groups. For instance, genes related to GO terms like vitamin binding, fatty acid synthase activity, and fatty acid elongase activity, as well as KEGG pathways like fatty acid metabolism and fatty acid elongation, were intensified in both diapause vs. no-diapause and diapause vs. pre-diapause comparisons, indicating the significant and widespread role of fatty acid synthesis metabolism in later stages of diapause. Conversely, genes related to the Toll and Imd signaling pathway, oxidative phosphorylation, and lysosome were categorized as forwarded in these comparisons, with lysosome specifically showing significant enrichment at the RNA level in pre-diapause vs. no-diapause, and in the RPF and forwarded groups in the other comparisons. The gene lists used for analysis were included in Additional file 4. The specific gene lists of the enrichment analysis were included in Additional file 3.

The lncRNAs profile were changed in the diapause process

To investigate the lncRNAs during diapause of ACB, we analyzed the transcriptome data of ACB in no-diapause, pre-diapause and diapause stages, and systematically identified and analyzed the lncRNAs of ACB (Fig. 4A). In total, 4561 high-confidence ACB lncRNAs were selected through four different lncRNA prediction software (Fig. 4B). Furthermore, by removing the already annotated lncRNAs in the insectbase2 database, 318

(See figure on next page.)

Fig. 4 Comprehensive analysis of lncRNAs identification and differential expression in ACB developmental stages. **A** Pipeline of lncRNAs identity. This part depicts the workflow for the identification of lncRNAs in ACB. **B** Venn diagram demonstrating the overlap and unique sets of lncRNAs identified through the utilization of filtering software, including CPC2, CAPT, CNCI, and Pfam, subsequent to the initial identification step. **C** Bar chart visualizes the distribution of exon counts within the identified lncRNAs. The blue bars represent known lncRNAs, the red bars denote novel lncRNAs, and the green bars signify mRNAs. **D** Density plots revealing the distribution of transcript sizes for known lncRNAs (in blue), novel lncRNAs (in red), and mRNAs (in green). **E** Box plots of the GC content for known lncRNAs, novel lncRNAs, and mRNAs. **F–H** The volcano plot shows the differential expression analysis results of lncRNAs among the three stages of ACB: no-diapause, pre-diapause, and diapause. The x-axis represents the \log_2 (fold change) value, and the y-axis represents the $-\log_{10}$ (adjusted *p*-value). Red dots indicate significantly upregulated lncRNAs, blue dots indicate significantly downregulated lncRNAs, and grey dots indicate lncRNAs with no significant differential expression. The number of upregulated and downregulated lncRNAs in each comparison is shown in the plot. **I** The relationship network among lncRNAs, target genes, and KEGG pathways of diapause vs. pre-diapause. The color represents the significance value of lncRNAs differential analysis or mRNA differential analysis. Different shapes represent different molecular types. The size of the shape proportionally maps the \log_2 (fold change) of lncRNAs differential analysis or mRNA differential analysis, with positive values indicating a significant increase during diapause and negative values indicating a significant increase during pre-diapause stage. **J–L** Validation of the relative expression levels of lncRNA using qPCR. Relative expression levels of lncRNA ofu_lnc_1066, ofu_lnc_0096 and ofu_lnc_004514 in pre-diapause vs. no-diapause (from left to right mean \pm SEM: 1.0716 \pm 0.29626, 3.7935 \pm 0.22433, 1.0351 \pm 0.17751, 1.5516 \pm 0.198, 1.0907 \pm 0.30674, 0.61699 \pm 0.17486, *n* = 3). **K** Relative expression levels of lncRNA MSTRG33, ofu_lnc_004834 and ofu_lnc_005214 in diapause vs. no-diapause (from left to right mean \pm SEM: 1.0672 \pm 0.271, 7.3254 \pm 1.5441, 1.0471 \pm 0.2332, 0.0007 \pm 0.0001, 1.1665 \pm 0.3857, 22.1190 \pm 5.5802, *n* = 3). **L** Relative expression levels of lncRNA MSTRG16991, ofu_lnc_001159 in diapause vs. pre-diapause (from left to right mean \pm SEM: 1.0369 \pm 0.20120, 0.0374 \pm 0.0045, 1.0164 \pm 0.1312, 0.0460 \pm 0.0142, *n* = 3)

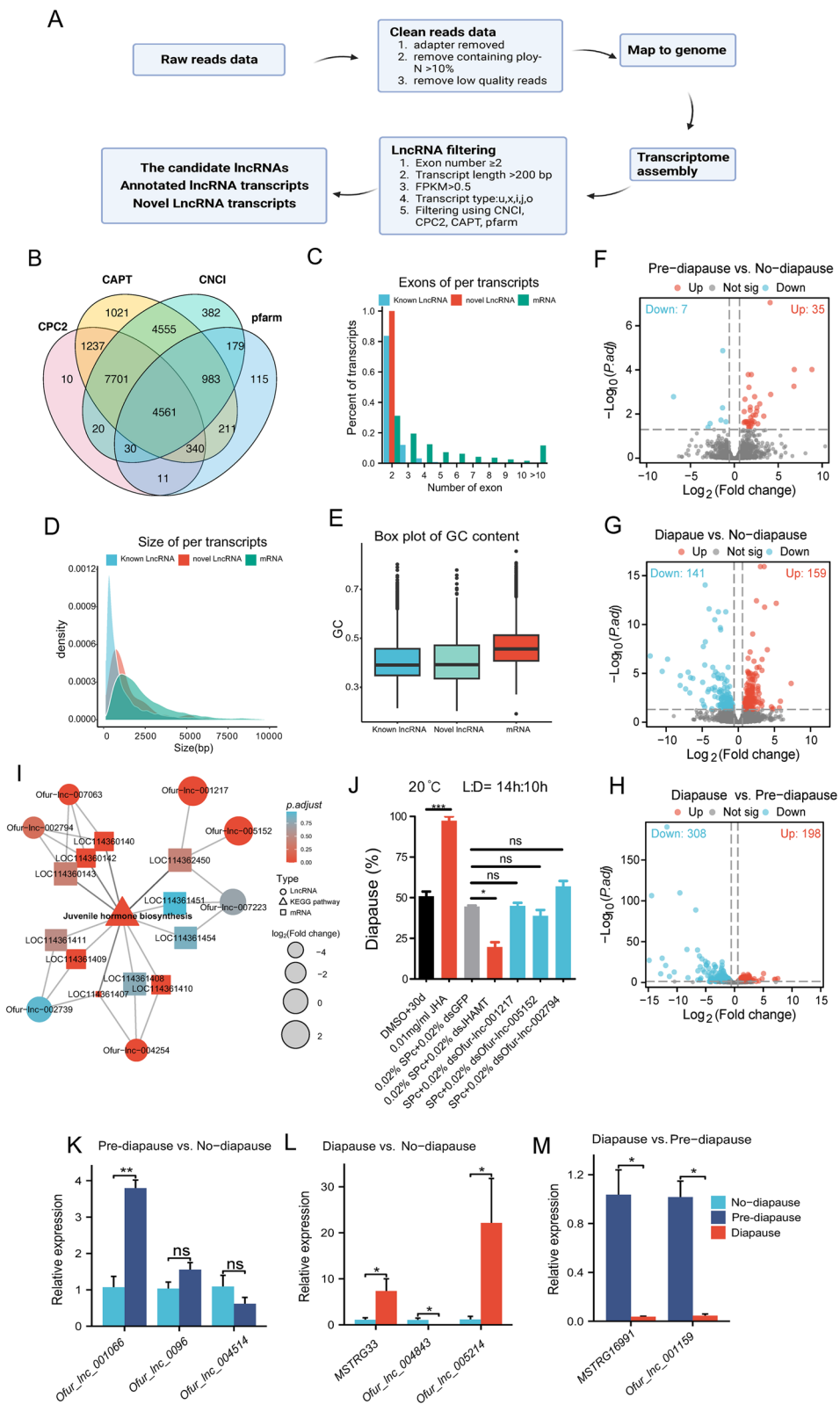


Fig. 4 (See legend on previous page.)

new high-confidence ACB lncRNAs were identified. All subsequent lncRNA analyses were conducted based on these 318 newly annotated and 7349 known lncRNAs. To characterize their genomic features, we compared the putative ACB lncRNAs with known protein-coding mRNAs. Generally, the number of exons in the ACB lncRNAs transcripts is mostly between 2 and 3 (Fig. 4C), and the transcript length is within 2500 bp (Fig. 4D), with a significantly lower GC content compared to mRNA (Fig. 4E).

After the lncRNAs identification and filtering steps, differential expression analysis was conducted to identify differentially expressed lncRNAs between no-diapause, pre-diapause, and diapause stages of ACB. PCA analysis was performed to visualize the variation between samples, and the results showed that samples from the same stage clustered together (Additional file 1: Fig. S7A-C), indicating high reproducibility and consistency between replicates.

Differentially expressed lncRNAs were identified based on the RNA-seq data (Fig. 4F-H). Comparing pre-diapause vs. no-diapause, there were 35 lncRNAs with increased abundance and 7 with decreased abundance. Comparing diapause vs. no-diapause, there were 159 differentially expressed lncRNAs with increased abundance and 141 with decreased abundance. Finally, comparing diapause vs. pre-diapause, there were 198 differentially expressed lncRNAs with increased abundance and 308 with decreased abundance. These results suggest that lncRNAs may play an important role in the regulation of diapause in ACB.

In order to gain a deeper understanding of the biological functions of lncRNAs, we conducted target predictions and GO and KEGG enrichment analyses of these lncRNA targets (Additional file 1: Fig. S7D, E). GO enrichment analysis showed that several processes were enriched in both the diapause vs. pre-diapause and diapause vs. no-diapause comparisons, including cytoplasmic side of early endosome membrane, structural constituent of cuticle, and integral component of mitochondrial inner membrane, among others. These findings suggest that these processes may play important roles in metabolic regulation during diapause. Additionally, the diapause vs. pre-diapause comparison showed enrichment in endopeptidase activity and actin filament, while the diapause vs. no-diapause comparison showed enrichment in UDP-glycosyltransferase activity and chitin-based extracellular matrix, among others. These results suggest that important changes in metabolism and cellular function occur when insects enter or exit diapause. In the pre-diapause vs. no-diapause comparison, enrichment was found in several processes related to the extracellular matrix, such as chitin-based extracellular matrix

and basement membrane, as well as in integral component of plasma membrane and mitochondrial matrix. This suggests that changes in extracellular matrix composition and plasma membrane function may be important in the transition from pre-diapause to no-diapause state. Moreover, enrichment was found in endonuclease activity and axon, among others, indicating potential changes in DNA repair and nervous system function during this transition.

Similarly, KEGG enrichment analysis was conducted to further investigate the pathways associated with the different groups (Additional file 1: Fig. S7E). Both comparison between diapause vs. pre-diapause and diapause vs. no-diapause showed enrichment in many metabolic pathways, including purine metabolism, glycine, serine and threonine metabolism, and the biosynthesis of amino acids. This suggests that these pathways play important roles in metabolic regulation during diapause. Additionally, in comparison to the pre-diapause state, the diapause state also showed enrichment in pathways such as arachidonic acid metabolism and starch and sucrose metabolism, indicating that these pathways may undergo significant regulatory changes upon entering diapause. The comparison between pre-diapause and no-diapause showed significant enrichment in fatty acid metabolism pathways, such as biosynthesis of unsaturated fatty acids, fatty acid metabolism, and fatty acid elongation, as well as other metabolic pathways such as galactose metabolism and tryptophan metabolism. This suggests that significant changes in fatty acid metabolism occur upon entering the pre-diapause state, which may have important implications for insects entering diapause.

To further explore the potential role of lncRNAs and their target genes in the pre-diapause process, we constructed an interaction network between lncRNAs, target genes, and pathways (Additional file 1: Fig. S6F). Our analysis revealed that the target genes of Ofur-lnc-001066 and Ofur-lnc-005150, namely *LOC114362590*, *LOC114352303*, and *LOC114362590*, were enriched in three pathways associated with fatty acid metabolism: fatty acid elongation, biosynthesis of unsaturated fatty acids, and fatty acid metabolism. These pathways have been shown to play important roles in many biological processes, including energy metabolism, cell membrane composition, and signaling pathways, and are also important during insect diapause. Fatty acid elongation is a process that converts medium-chain fatty acids into long-chain fatty acids, which are important components of cell membranes and serve as an energy source during insect diapause. Biosynthesis of unsaturated fatty acids is also crucial for maintaining membrane fluidity and function. In addition, fatty acid metabolism plays a key role in regulating energy metabolism and maintaining

lipid metabolism balance. We also conducted interaction analysis of differentially enriched target genes and their interacting lncRNAs in the JH pathway in ACB diapause vs. pre-diapause. We found that out of the 10 interacting mRNAs, 5 showed significant differential expression, while 4 lncRNAs also exhibited significant differential expression. Among them, the change in Ofur-lnc-007063 and its host gene *LOC114360142* was remarkably significant, with both showing a significant decrease during diapause (adjusted p -value < 0.01, $|\log_2(\text{fold change})| > 1$) (Fig. 4I). This suggests the presence of a certain regulatory interaction between lncRNA and JH pathway-related mRNAs. We also selected 8 lncRNAs for differential expression gene validation in random (Fig. 4K–M). The qPCR results were consistent with the RNA-seq sequencing results, confirming the accuracy of our findings.

In this study, dsRNAi constructs targeting lncRNAs corresponding to five diapause-associated genes were created and used in feeding experiments to assess their impact on diapause rates (Fig. 4J). The dsRNAi for DHAMT was used as a positive control. The effectiveness of the dsRNAs was tested by RT-PCR (Additional file 1: Fig. S8A–E). The dsRNAi for Ofur-lnc-001217 and Ofur-lnc-005152 showed no significant effect, while Ofur-lnc-002794 exhibited a trend of increased diapause rates, although not statistically significant. RNA-seq data indicated a significant decrease in the expression of Ofur-lnc-002794 and its target genes during diapause initiation (Additional file 2: Table S1), suggesting their potential role in promoting diapause. The rise in diapause rates following dsRNAi interference with Ofur-lnc-002794 hints that it may normally inhibit diapause; thus, its disruption might relieve this inhibition, leading to an increased diapause rate. This observation suggests that Ofur-lnc-002794 and its target genes could be significant in diapause regulation. More extensive tests are needed in the future.

In conclusion, the enrichment of lncRNAs target genes in these pathways suggests their potential involvement in regulating critical metabolic processes during the pre-diapause stage. Further investigation of the intricate molecular mechanisms by which lncRNAs regulate these pathways may provide new insights into the physiological regulation of diapause.

Comparison of the DEGs profile of ACB diapause with other types of diapause revealed the molecular characteristics of larval diapause

In order to compare the regulatory pathways involved in different types of insect diapause, we conducted RNA-seq analysis on public data and performed KEGG enrichment analysis on the significantly DEGs. Through a comparative study of diapause in the ACB with other

types of diapause, we revealed the molecular characteristics of larval diapause. Our analysis primarily focused on conducting KEGG enrichment analysis across various diapause types, including egg diapause, larval diapause, and adult diapause (Fig. 5A). The results demonstrated that multiple diapause types exhibited enrichment in common metabolic pathways. Among the enriched pathways, the longevity regulating pathway—multiple species, amino sugar and nucleotide sugar metabolism, glutathione metabolism, carbon metabolism, pyruvate metabolism, drug metabolism—cytochrome P450, mTOR signaling pathway, peroxisome, and purine metabolism were consistently enriched in all three diapause types. This indicates the involvement of common regulatory mechanisms and metabolic pathways across diapause stages. Significantly, these enriched pathways in the ACB's diapause profile reveal the conservation of essential metabolic processes during larval diapause.

These pathways, such as autophagy, pentose and glucuronate interconversions, pyruvate metabolism, glycerolipid metabolism, sphingolipid metabolism, inositol phosphate metabolism, and galactose metabolism, are significantly enriched during the reproductive diapause of fruit flies, in contrast to ACB larvae. This exclusivity holds significant implications, indicating that these pathways may play distinct roles in nutrient utilization, energy regulation, and cellular adaptation during the reproductive diapause of fruit flies.

We also identified several pathways that were significantly enriched during diapause in ACB larvae. These pathways include insect hormone biosynthesis, DNA replication, alanine, aspartate, and glutamate metabolism, ECM-receptor interaction, Wnt signaling pathway, nucleotide metabolism and FoxO signaling pathway. These pathways exhibited prominent enrichment specifically during diapause in ACB larvae, highlighting their preferential involvement in this developmental stage.

During the diapause of ACB larvae, a pronounced enrichment of the α -linolenic acid metabolism, pyrimidine metabolism, caffeine metabolism, and sulfur metabolism pathways has been observed, potentially holding significant implications. The α -linolenic acid metabolism is linked to the synthesis of ω -3 fatty acids, which play a crucial role in various physiological processes such as cell membrane structure and signaling molecules. During diapause, larvae experience metabolic slowdown and reduced activity, and the alterations in fatty acid metabolism might reflect the demand for adapting to an energy-saving mode during diapause. The pyrimidine metabolism is associated with the synthesis of nucleotides for building DNA and RNA. Considering the metabolic deceleration during diapause, the enrichment of this pathway may be related to the maintenance

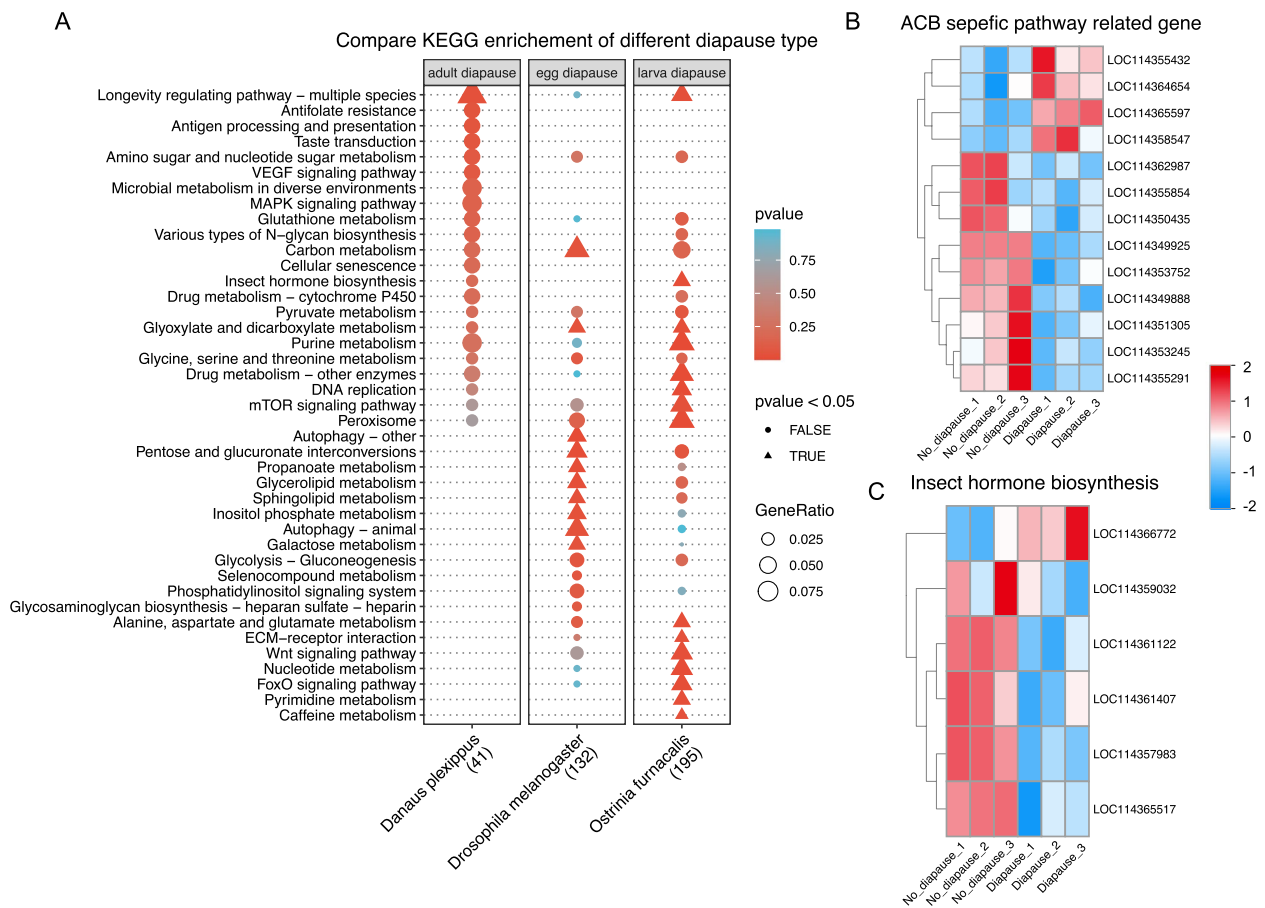


Fig. 5 Comparison of KEGG pathway enrichment analysis among different diapause types. **A** Dot plot of comparison of KEGG pathway. Each point or triangle represents a KEGG pathway, with the color intensity indicating the degree of enrichment, where redder the color, the more significant the pathway. The size of the points or triangle corresponds to the GeneRatio, showing a positive correlation. Notably enriched pathways are marked with triangular symbols (p -value < 0.05). **B** Heatmap of specific pathway genes of the ACB. **C** Heat map of insect hormone biosynthesis related gene

of fundamental cellular processes, such as DNA repair and maintenance, even under reduced metabolic activity. The enrichment of caffeine metabolism could indicate a response to environmental stress. Caffeine is believed to have a protective role against oxidative stress and participates in regulating various metabolic pathways. The accumulation of this pathway could signify an adaptive response to potential stress factors during diapause.

In these pathways, we also identified specific pathways and genes with significant changes unique to the diapause process of the ACB (Fig. 5B). The JH biosynthesis pathway is crucial in the insect diapause process. We observed a significant increase in the expression of the upstream gene, farnesyl diphosphate phosphatase (*FPPP*), promoting JH synthesis when compared to the no-diapause state. Simultaneously, we found a significant inhibition of *LOC114361407*, which encodes NADP⁺-dependent farnesol dehydrogenase, an enzyme in the corpora allata involved in JH synthesis. Additionally, the expression

of genes encoding aldehyde dehydrogenase (NAD⁺) (*LOC114357983*) and JH-III synthase (*LOC114365517*), involved in JH biosynthesis, was altered, leading to the induction of diapause in the ACB (Fig. 5C).

Exploring the influence of HSPs and the proteasome pathway on diapause in the ACB

We found that the HSP70 gene family showed a significant dynamic change during the diapause process in the ACB, and their expression levels were different at different diapause stages, both at the transcriptional and translational levels. At the transcriptional level, they generally showed a pattern of decreasing first and then increasing, indicating that the HSP70 family genes were suppressed and then facilitated the entry into diapause. At the translational level, this change was very significant ($p < 0.05$), for both hsp68-like and HSP70 genes (Fig. 6A). We also found that the *hsp70* (*LOC114359959*) gene had

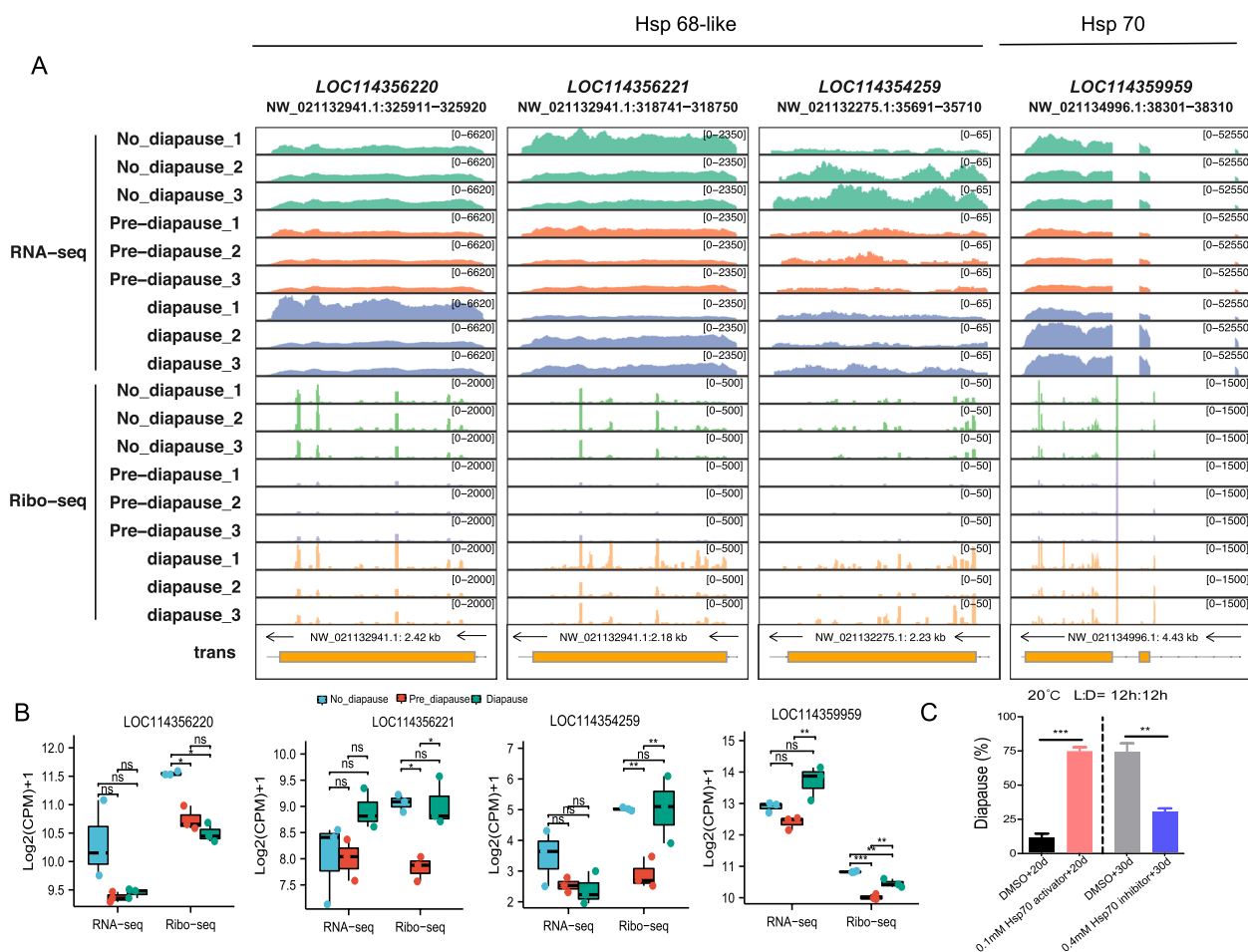


Fig. 6 Analysis of heat shock protein expression in relation to different diapause types. **A** Representative Integrative Genomics Viewer (IGV) plots for the transcription and translation profiles of four HSPs. **B** Comparative box plots of RNA-seq and Ribo-seq expression for the four corresponding HSPs. These expression patterns have been subjected to statistical analysis using the one-way analysis of variance (ANOVA). **C** Statistical representation of diapause rate under treatment with HSPs inhibitors and activators. **D** Comparison of diapause percentage of ACB under HSPs inhibitors and activators. * stands for $p < 0.05$, ** stands for $p < 0.01$, *** stands for $p < 0.001$, ns stands for not significant

a markedly increased transcription level and a markedly decreased translation level after entering diapause (Fig. 6B). When treated with the HSP activator, the diapause percentage significantly increased, while treating them with the HSP inhibitor significantly decreased the diapause percentage (Fig. 6C) [32, 33]. The effectiveness of the activator and inhibitors for HSP protein was verified by examining the expression of HSP proteins (Additional file 1: Fig. S8F–H). These results indicate that HSPs plays an important regulatory role in the diapause process of the ACB.

This study also revealed the involvement of proteasome pathway in diapause of ACB. In cluster 3, there is an initial and substantial decrease in gene expression levels, followed by a sustained, unchanging pattern.

Results shown above indicated that cluster 3 enriched a group of genes that are associated with cellular pathways related to proteasomes and ubiquitin-mediated degradation (Fig. 1G). In GSEA, to investigate RNA-seq data from different diapause stages, our analysis revealed the restriction and suppression of pathways related to proteasomes during the transition from the no-diapause state to diapause (Fig. 7A, B). To validate this process, treatment of ACB using MG132, an inhibitor of proteasome activity, led to an increase in percentage of diapause at a concentration of 100ug/ml (Fig. 7C). The effectiveness of the MG132 was verified by examining the expression of HSP proteins, which are subjected for degradation by proteasomes [34, 35] (Additional file 1: Fig. S8I–J). These results indicate that proteasome pathway plays an important regulatory role in the diapause process of the ACB.

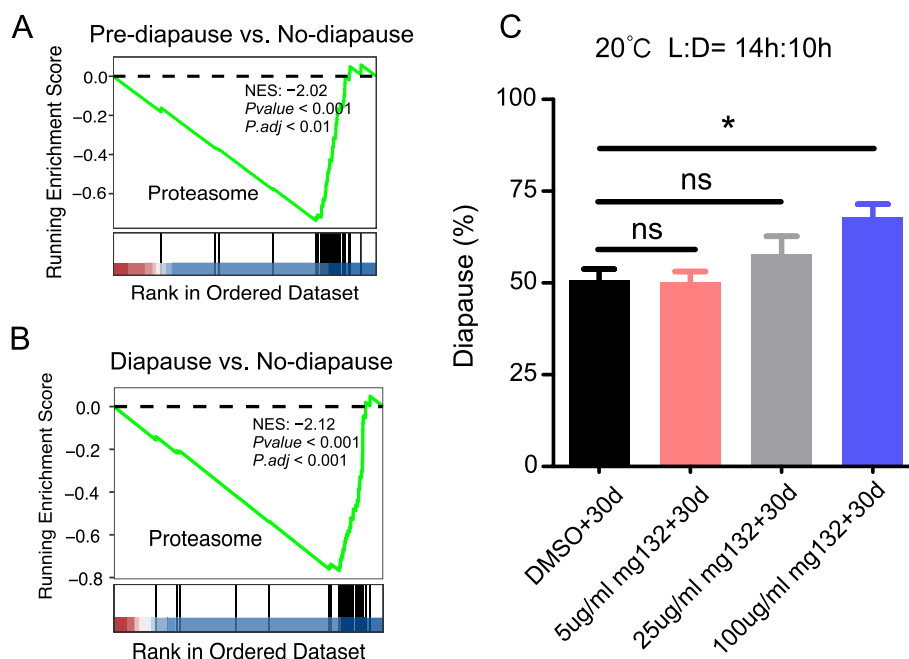


Fig. 7 Proteasome pathway inhibition promotes diapause in ACB. The GSEA analysis highlights gene expression variations in ACB, pre-diapause vs. no-diapause (A) and diapause vs. no-diapause (B). The running enrichment score reflects gene distribution in a set across the gene expression dataset, with peaks showing enrichment and valleys indicating suppression. This score reaches its maximum and minimum at these peaks and valleys, respectively. The Normalized Enrichment Score (NES), a standardized version of this score, allows comparison between gene sets. A positive NES means enrichment, while a negative NES suggests suppression. C Impact of MG132 treatment on the diapause rate of ACB. The experiment was conducted at 20 °C under a 14-h light/10-h dark photoperiod for 30 days. DMSO was used in the control group. The results show changes in the diapause percentage with MG132 treatments at 5 µg/ml, 25 µg/ml, and 100 µg/ml, compared to the control group. Statistical analysis of the differences among groups was performed using ANOVA. * stands for $p < 0.05$, ns stands for not significant

Discussion

Diapause is a vital adaptation that enables insects to endure harsh seasons, allowing them to strategically manage their life cycles in response to changing environmental conditions. This study extensively investigated the transcriptional and translational levels of the ACB during different diapause stages, revealing a series of significant findings. Firstly, ribosome dynamics analysis indicated substantial changes in ribosome conformation and translation activity across various diapause stages. Particularly, a broader distribution of RPFs was observed during the pre-diapause and diapause stages, suggesting structural or conformational differences in ribosomes at these stages. Ribosomal positioning analysis highlighted a strong preference for the P-site signal within the CDS, emphasizing ribosomes' preference for translating CDS. Simultaneously, lower P-site signals in the 5' UTR and 3' UTR suggested weaker translational activity in these regions. Furthermore, a combinational analysis of RNA-seq and Ribo-seq data revealed differences in translational efficiency between samples, providing insights into the regulation of protein synthesis during diapause.

In the study of insect diapause, some important genes and pathways have been discovered [3, 36, 37].

The regulation of diapause can be roughly divided into molecular regulation, hormone regulation [37–39], circadian clock [40, 41], photoperiod regulation [42], and energy metabolism regulation [43], etc. During diapause in the ACB, transcriptome analysis revealed differential expression of genes associated with crucial metabolic and regulatory pathways. These pathways include genes related to fatty acid metabolism, insect hormone synthesis, DNA replication, extracellular matrix-receptor interaction, Wnt signaling, nucleotide metabolism, and the FoxO signaling pathway. The JH plays a crucial role in the synthesis of insect hormones and is of great significance during the process of diapause. We observed a significant decrease in the expression of genes related to the synthesis of JH precursors in diapausing ACB larvae when compared to their no-diapause state. This reduction may contribute to the entry of the ACB into diapause.

Diapause manifests differently across various developmental stages in insects. By comparing different diapause types in insects, we have identified several conserved pathways related to diapause. Several studies have mentioned the longevity regulating pathway—multiple species in relation to diapause in insects. For example, one study indicated that the continuous upregulation of

relevant genes in *Streltzoviella insularis* during winter is associated with the longevity regulating pathway, which may help insects survive under low-temperature conditions [44]. Another study exploring the effects of high and low temperatures on *Spodoptera frugiperda* also found that the related genes play a role in preparing insects for diapause [45]. This pathway is significantly enriched in both ACB larval diapause and Monarch butterfly adult diapause. HSPs in both of them may play a key role in supporting stress resilience during diapause (Additional file 3, correspondence to Fig. 5A), enhancing adaptability and potentially contributing to lifespan extension. In mammals, mTOR regulates reproductive diapause by controlling energy metabolism and cellular growth, managing the pause and resumption of embryonic development [46]. mTOR influences the regulation of energy metabolism and hormone biosynthesis in *Bombus terrestris*, affecting growth and reproduction balance and inducing diapause [47]. Comparative enrichment analysis indicates the monarch butterfly appears to focus on controlling protein translation through the *EIF4EBP2* gene, an effective energy management strategy that conserves energy by reducing protein synthesis. In contrast, the fruit fly, like the ACB, not only involves the core mTOR signaling pathway but also utilizes genes from the Frizzled family to influence cellular behavior through the Wnt signaling pathway (Additional file 3, correspondence to Fig. 5). FoxO may thus be an important mediator of diapause in a range of species [48]. In a recent study, researchers discovered that the activation of the ubiquitin-proteasome system by *Helicoverpa armigera*'s FOXO protein leads to the translation of diapause-related responses [49]. The FoxO signaling pathway has been extensively studied in diapause research in various species, such as *Locusta migratoria* [50], *Galeruca daurica* [51], *C. pipiens* [52], and *Antheraea pernyi* [53]. A large body of research demonstrates that JH plays a critical role in inducing diapause across various insect species [54–56]. During fruit fly adult diapause, when JH production is blocked, it not only stops egg maturation but also triggers physiological changes to prepare for overwintering [57]. The JH may act as a regulatory switch in diapause control, such as the mosquito *Aedes* shown that diapause encompasses both the transcriptional down regulation of the pathway that synthesizes JH and the simultaneous enhancement of the production of JH esterase, an enzyme that degrades JH [58].

This study leads to the discovery of new pathways that potentially involve in diapause. The translation analysis uncovered DEGs linked to some known pathways, such as ribosome biogenesis [59, 60], cell cycle regulation [61, 62], and the PP2A signaling pathway [63], which potentially influencing physiological adaptations during

diapause. Significant changes in lncRNAs profiles were observed, indicating their potential roles in diapause regulation. GO and KEGG enrichment analysis of lncRNA targets demonstrated enrichment in various biological processes and metabolic pathways. Similarly, when comparing differential lncRNA target genes, we also found significant enrichment in pathways related to JH synthesis. This suggests that lncRNAs may also play a role in regulating the occurrence of diapause in the ACB. There are currently almost no similar reports available.

HSPs play a role in maintaining cellular stability, enhancing stress adaptation [64], and may be involved in the regulation of diapause processes in insects [9]. These functions contribute to the survival of insects in adverse environmental conditions and their successful entry into diapause. Like HSPs in other species, HSPs in the ACB are highly expressed before and during diapause. Their gene structures also have fewer or no introns, which facilitates gene expression, allowing for a rapid response to environmental changes [65]. In diapause, changes in the quantity of heat shock proteins may occur in anticipation of specific stressors, whether it involves an increase, decrease, or no change, rather than as a direct response to the stress itself [9, 66, 67]. This also corroborates our findings in the ACB. During the pre-diapause period leading up to entering diapause, a significant reduction in HSPs was observed. This rapid response of HSPs may serve as a crucial regulatory factor for the ACB's adaptation to adverse environmental conditions. Moreover, treatment with either an HSP activator or inhibitor significantly alters the diapause rate. These findings underscore the pivotal role of HSP in regulating diapause. The involvement of HSP in diapause regulation may be attributed to their function in mediating protein folding. Further exploration into the downstream molecular events of HSP during diapause could provide valuable insights into the underlying mechanism.

A comparative study with other types of insect diapause elucidated the molecular characteristics of larval diapause. Enriched metabolic pathways common to various diapause types included those related to longevity regulation, amino sugar and nucleotide sugar metabolism, glutathione metabolism, carbon metabolism, pyruvate metabolism, drug metabolism, mTOR signaling, peroxisomes, and purine metabolism. These shared pathways highlighted their critical roles in coordinating the physiological adaptations necessary for diapause survival across diverse diapause types. Meanwhile, exclusive enrichment in specific pathways and genes in ACB larval diapause suggested distinct metabolic and regulatory adaptations unique to this species.

In summary, our research extensively explored the molecular mechanisms, lncRNAs regulation, and

metabolic adjustments, and key genes associated with diapause in the ACB. It shed light on the broader field of insect diapause biology, offering insights into strategies for pest management and ecological research in the future.

Conclusions

By applying RNA-seq and Ribo-seq techniques on ACB samples at various stages of diapause, this study reveals DEGs at both the transcriptional and translational levels throughout the diapause process. It not only offers insights into dynamic expression changes during this process but also lays the foundation for further investigations into the regulation of diapause.

Methods

Rearing of ACB, diapause induction, and sample collection

We obtained a batch of laboratory-reared ACB from the Institute of Plant Protection, Chinese Academy of Agricultural Sciences, and also collected wild populations from the Beijing area to increase genetic diversity. We reared the insects in transparent and breathable transparent boxes, feeding them with artificial diet made of corn flour, soybean flour, yeast, sucrose, agar, etc. We placed the boxes in a constant temperature and light incubator and controlled the temperature at 27 °C, and the light-dark ratio was at 16:8, simulating the natural environment. To induce the ACB to enter diapause, we changed the light-dark ratio to 12:12 [68] and the temperature to 20 °C from the 3rd instar larvae for 40 days. We found that on the 20th day of induction, the larvae entered the pre-diapause state, that is, they began to slow down their development speed, but did not stop completely. On the 40th day of induction, the larvae entered diapause completely. To analyze the gene expression changes in the ACB, we extracted RNA from 5th instar larvae in different states and sequenced it. We selected no-diapause 5th instar larvae (no-diapause) under normal temperature and light, 5th instar larvae (pre-diapause) induced for 20 days, and 5th instar larvae (diapause) collected on the 10th day after entering diapause.

Sample preparation and library construction for Ribo-seq and RNA-seq

For ACB sample, we collected the final instar larvae of no-diapause and pre-diapause and 10 days diapause. For each state, we decapitated the larvae using a scalpel blade and immediately froze the heads in liquid nitrogen for RNA extraction. We collected 150 heads for each of the three states as one biological replicate and repeated this sample collection three times in total. Finally, RNA was extracted from the larvae of all three states, half of each sample for RNA-seq and the other for Ribo-seq. Each

sequencing analysis (RNA-seq and Ribo-seq) was performed with three biological replicates.

For Ribo-seq library, samples were treated with specific lysis buffer with cycloheximide (50 mg/mL) to acquire the lysate (Novogen, China). To digest the RNA other than RPFs, ACB tissue lysate was treated with unspecific endoribonuclease RNase I. Isolation of monosomes was performed by size-exclusion chromatography with MicroSpin S-400 HR columns. The RNA samples were then treated with rRNA depletion kit to deplete the samples of as much rRNA contamination as possible before PAGE purification of the relatively short (20~38 nt) RPFs. Following PAGE purification, the both ends of RPF were phosphorylated and ligated with 5' and 3' adapters respectively. Then the fragments were reversely transcribed to cDNAs and amplified by PCR, followed by quality testing and library construction. Ribo-seq libraries were sequenced using SE50 (Illumina NovaSeq 6000). For RNA-seq library, total RNA was extracted with a standard TRIzol RNA extraction protocol. The cDNA library building and sequencing using Illumina PE 150 (Illumina NovaSeq 6000) and RNA-seq analysis, which were performed by Novogen.

Ribo-seq and RNA-seq data were analyzed in the following way. Raw data (raw reads) in FASTQ format were firstly processed using Trimmomatic (v0.39) [69] to trim adapters and remove reads containing N bases and low-quality reads. Subsequently, the quality of the clean data was assessed using FastQC (v0.12.1), which calculated Q20, Q30 scores and GC content, providing a comprehensive quality control report.

RNA-seq analysis

For RNA-seq clean data, the index of the ACB reference genome was built using Hisat2 (v2.0.5) [70], and paired-end clean reads were aligned to the ACB reference genome (NCBI project: ASM419383v2) using Hisat2. The mapped reads of each sample were assembled by StringTie (v1.3.3b) [71] in a reference-based approach. FeatureCounts (v1.5.0) [72] was used to count the reads numbers mapped to each gene. Differential expression analysis was performed using the DESeq2 R package (v1.20.0) [73]. Genes with an adjusted p -value < 0.05 and $|\log_2$ (fold change) ≥ 1 were assigned as significant differentially expressed gene. GO enrichment, KEGG pathway enrichment, and GSEA analysis were done using clusterProfiler package (v4.5.2.002) [74].

In the process of comparing diapause types in different species, we applied the same analysis procedures to the raw data of the *Danaus plexippus* (NCBI project: PRJNA548105) and the *Drosophila melanogaster* (NCBI GEO: GSE28588). The gene lists used for analysis were

included in Additional file 4. The specific gene lists of the enrichment analysis were included in Additional file 3.

Ribo-seq analysis

At first, the clean reads were aligned to a FASTA file containing rRNA, tRNA, snoRNA, and snRNA from NCBI annotations using Bowtie (v1.2.2) [75] with a maximum of two mismatches (-v 2) by default, discarding the successful alignments and collecting all unaligned reads. Unaligned clean reads (ribosome-protected fragments, RPF) were then mapped to the ACB reference genome using STAR (v2.7.3a) [76] with default options. The unique genome-mapped reads were then mapped against protein coding transcripts using Bowtie (v1.2.2) [75] with parameters by default. The featureCounts program in Subread package (v1.6.3) was used to count the number of RPFs uniquely mapped to CDS regions based on ACB genome mapping file (-t CDS -g gene id), which were then normalized as RPF Per Kilobase per Million mapped RPFs (RPKM). The counts for all transcripts of a gene were aggregated to obtain the total expression level of that gene. This approach ensures that the gene expression levels reflect the cumulative abundance of all its transcripts. DESeq2 R package (v1.2.0) [73] was used for differential expression analysis. The threshold for significantly differential expression is $|\log_2(\text{fold change})| \geq 0.265$ and adjusted p -value < 0.05 . Differential TE analysis was performed using deltaTE [77]. TE is defined as the ratio of Ribo-seq RPKM to RNA-seq RPKM for each gene. Coding frame distribution and 3-nt periodicity analyses for Ribo-seq quality evaluation are performed based on Ribo-TISH (v0.2.6) [78]. Peptidyl site (P-site) offsets, region P-site percentages, length distributions, and meta-gene heatmaps were meticulously calculated and generated using riboWaltz (v1.0) [79]. Unless otherwise specified, the methods for GO and KEGG enrichment analysis, GSEA enrichment analysis, volcano plot, fuzzy clustering, and other graphical visualizations are the same as those mentioned in the RNA-seq section. The gene lists used for analysis were included in Additional file 4. The specific gene lists of the enrichment analysis were included in Additional file 4.

GSEA analysis

GSEA enrichment analysis includes GSEA GO and GSEA KEGG analyses. The gene sets for GO analysis were obtained from the GO annotation on the Bioconductor AnnotationHub website (<https://annotationhub.bioconductor.org/species/Ostrinia%20furnacalis>), and the KEGG annotations were sourced from the KEGG website (https://www.genome.jp/kegg-bin/show_organism) [80]. The significance threshold for the analysis was set to a p -value of 1 to enrich all results, and the

correction method used was Benjamini-Hochberg (BH). Enrichments were considered significantly activated if p -value < 0.05 , p .adjust < 0.25 , and NES > 1.5 , and significantly inhibited if p -value < 0.05 , p .adjust < 0.25 , and NES < -1.5 . Volcano plots and upset plots were plotted with the ggplot2 package. The fuzzy K-means cluster method from Mfuzz package (v2.62.0) were used to get appropriate cluster numbers. ClusterGVis was used for clustering and visualizing the time series gene expression (<https://github.com/junjunlab/ClusterGVis>) [81]. The BioSeqUtils package (<https://github.com/junjunlab/BioSeqUtils>) [82] was used to plot gene expression peak charts for RNA-seq and Ribo-seq.

Calculation of translational efficiency (TE) and classification of various regulatory categories

We classified the DEGs using the deltaTE method [77]. TE is defined as the ratio of RPFs to mRNA counts within a gene's CDS. The formula for calculating translational efficiency (TE) is as follows:

$TE = (\text{Ribo-seq RPKM}) / (\text{RNA-seq RPKM})$. A gene with concurrent changes in mRNA and RPFs is a differentially transcribed gene (DTG). A gene with RPF changes independent of mRNA alterations, leading to TE changes, is categorized as a differential translation efficiency genes (DTEG). The regulation group of each gene is determined by combining the changes in ribosome-protected fragments, mRNA counts, and translational efficiency (ΔTE). Forwarded genes are defined by significant alterations in both RNA levels and RPFs, but with no notable changes in TE, indicating that the increase or decrease in RNA directly influences RPF levels. Exclusive genes are those regulated solely at the translational level; these genes exhibit significant shifts in translational efficiency and RPFs without any corresponding changes in RNA levels, showcasing a regulation mechanism independent of RNA abundance. For genes that experience combined regulation, they fall into either intensified or buffered categories: Intensified genes see both RNA levels and translational efficiency change in the same direction, amplifying their expression, whereas buffered genes have RNA and translational efficiency changes that counteract each other. This buffering action can moderate or even negate the effects of RNA changes on protein production, with a special case where despite significant shifts in both RNA and translational efficiency, the RPFs remain unchanged.

LncRNA identifications

The lncRNA identification and filtering steps were as follows. First, a total of 110,752 transcripts were assembled for each sample, and then compared with the ACB genome to screen out 20,876 candidate

lncRNAs. Subsequently, a series of filtering strategies were employed to screen these candidate lncRNAs, including removing transcripts that overlap with coding genes, transcripts with length less than 200 bp or containing only one exon, low-expressed transcripts with $\text{FPKM} \leq 0.5$, and remaining transcripts with class code (i, j, o, u, x). The high-confidence ACB lncRNAs were selected through the comparison with the known protein structure domain database (Pfam) and the evaluation of protein-coding potential using three different lncRNA prediction software CPC2 [83], CNCI(v2.0) [84], and CPAT (v3.0.4) [85]. Finally, the Insectbase2 database [86] was used to screen known lncRNAs. Regarding the identification of target genes for lncRNAs, we performed the following procedures. For known lncRNAs, we extracted target genes from InsectBase2. For newly identified lncRNAs, we extracted the mRNA located within 100 kb upstream and downstream regions as their potential targets.

Drug and RNAi treatment

To explore the impact of the proteasome inhibitor MG132 (Biorigin, #BN26000) on diapause regulation, early 3rd instar larvae were subjected to MG132 treatment. MG132 was dissolved in DMSO and diluted with fresh artificial diet to final concentrations of 5 $\mu\text{g/ml}$, 25 $\mu\text{g/ml}$, or 100 $\mu\text{g/ml}$. Neither the solvent nor MG132 caused observable damage or stress responses. Larval feed was replenished every 48 h, with remaining feed removed and replaced. Incubation occurred in a controlled environment with a 14:10 h light-dark cycle at 20 °C and 60% humidity. After 30 days of treatment, diapause percentages under MG132 conditions were determined.

For HSP70 inhibitor and activator studies, concentrations of 0.4 mM HSP70 inhibitor VER-155008 (Target Mol, #T7010) and 0.1 mM HSP70 activator ML346 (Target Mol, #T3594) were separately dissolved into the artificial diet. Early 3rd instar larvae were fed with the respective drug-containing diet, with feeding regimen and environmental conditions similar to the MG132 experiment. After treatment, diapause percentages under different drug conditions were assessed.

For the plasmid construction and application of RNAi, the cDNA of the 5th instar larvae was used as the template for PCR amplification. Ofur-lnc-001217, Ofur-lnc-005152, and Ofur-lnc-002794 sequences were cloned into pET-30a to generate pET30-lncRNAs plasmids using standard ClonExpress technique (Vazyme, #C115). The sequence of primers is shown in Additional file 2: Table S2. pET28a-JHAMT was generated in a previous study [87]. To investigate the function of OfJHAMT and Ofur-lncRNAs using large amounts of dsRNA,

we constructed a pET30a (Tiandz, Mianyang, China) recombinant plasmid containing a hairpin structure of dsJHAMT and dsOfur-lncRNAs sequence. The constructed plasmid pET30a-JHAMT and pET30a-lncRNAs were transformed into BL21 (DE3) RNase III, a novel system that produces a large batch of dsRNA for insect gene silencing. The extraction of double strand RNA (dsRNA) followed the method previously described [88].

For large amounts of dsRNA induction and application. The concentration of dsRNA was determined by Thermo NanoDrope 2000, and the dsRNA was mixed with fresh artificial diet to a final concentration of 0.02% dsGFP, dsJHAMT, or dsLncRNAs. JHA (Alta Scientific, #1ST20290) serves as a positive control under feeding conditions. The dsRNA was separately mixed with an equal mass of SPc (Star Polycation, SPc). SPc is a gift from Dr. Jie Shen's laboratory at China Agricultural University at Beijing, China. The dsRNA was complexed with SPc to enhance stability and cellular uptake [89]. Subsequently, a final 1% surfactant, APG, was added to complete the complex solution. Early 3rd instar larvae were fed with the dsRNA-containing diet, with feeding and incubation conditions matching previous experiments. After 30 days, diapause percentages under different dsRNA treatments were evaluated.

Western blots

Protein extraction was conducted from 4th instar ACB larvae treated with HSP70 inhibitor and MG132. The heads of the larvae were homogenized and lysed in RIPA lysis buffer (CWbio, #CW2333S). The impact of MG132 was assessed by analyzing HSP70 protein levels [34, 35]. Antibodies from ABclonal, including anti-HSP70 (Cat# A12948, dilution rate: 1:2000) and anti-Actin (Cat# AC026, dilution rate: 1:3000), were utilized. Protein band intensities were quantified and analyzed using AZURE Biosystems (models C300-C600). Statistical significance was determined via unpaired Student's t-test, where "n.s." denotes no significant difference, and "*" represents $p < 0.05$.

Quantitative real-time PCR

In this study, we investigated the lncRNAs gene expression profiles of ACB larvae across no-diapause, pre-diapause, and diapause states using quantitative real-time PCR (qRT-PCR). The larvae used for qRT-PCR were part of the same batch collected for RNA-seq and Riboseq analyses. For RNA extraction, 10 heads from each biological replicate were utilized. The RNA was then converted to cDNA using the PrimeScript™ RT reagent Kit with gDNA Eraser (TakaRa, #RR047A), which also eliminates genomic DNA contamination. The cDNA amplification was performed using SuperReal PreMix Plus (SYBR

Green; TIANGEN, #FP205-02) on an Applied Biosystem StepOne Real-Time PCR system. *Actin* served as the normalization control. Statistical analysis of gene expression differences between the states was conducted using an unpaired two-tailed Student's *T*-test with R (v4.2.1). n.s. indicates no significant difference, * indicates $p < 0.05$, ** indicates $p < 0.01$, and *** indicates $p < 0.001$. All primers used are listed in Additional file 2: Table S3.

Abbreviations

ACB	Asian corn borer
RNA-seq	RNA sequencing
Ribo-seq	Ribosome sequencing
lncRNAs	Long non-coding RNAs
GO	Gene Ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
GSEA	Gene Set Enrichment Analysis
JH	Juvenile hormone
RPFs	Ribosome-protected fragments
P-site	Peptidyl site
TE	Translation efficiency
DTG	Differentially transcribed gene
DTEG	Differential translation efficiency gene
Spc	Star polycation
RNAi	RNA interference
dsRNA	Double strand RNA

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12915-024-02000-1>.

Additional file 1: Fig. S1. Sample collection and Quality control of the samples used for RNA-seq and Ribo-seq. (A) Schematic description of sample collection under different conditions for the Ribo-seq and RNA-seq experiments in ACB. (B, E) Meta analysis of gene expression data got from RNA-seq and Ribo-seq. (C, F) PCA of different samples for RNA-seq and Ribo-seq. (D, G) Correlation heatmap of normalized RNA and RPFs values between the three groups of samples. The color and size of the dots represented the Pearson correlation coefficient. Fig. S2. GO enrichment analysis for gene expression dynamics in RNA-seq data. (A-C) GO enrichment analysis for diapause vs. no-diapause (A), diapause vs. pre-diapause. (B) and pre-diapause vs. no-diapause (C) samples. Left panel: The bar plots represent the top 10 enriched GO terms in three categories: Biological Process (BP), Cellular Component (CC), and Molecular Function (MF). The X-axis indicates the number of genes associated with each GO term. The color of the bars represents the significance of the GO terms (P_{adj}), with darker colors indicating higher significance. Right Panel: The network plot displays the top 5 most significant GO terms and their corresponding genes. Different colors of the lines represent different GO terms. The size of each node indicates the number of genes associated with that term, while the color of each node represents the \log_2 (fold change) of the gene expression, with the color gradient showing the direction and magnitude of the change (red for up-regulated genes, blue for down-regulated genes). Fig. S3. KEGG enrichment analysis for gene expression dynamics in RNA-seq data. (A-C) KEGG enrichment analysis for diapause vs. no-diapause (A), diapause vs. pre-diapause. (B) and pre-diapause vs. no-diapause (C) samples. Left panel: The bar plots represent the top 20 enriched KEGG pathway. The X-axis indicates the number of genes associated with each pathway. The color of the bars represents the significance of the pathway, with darker colors indicating higher significance. Right Panel: The network plot displays the top 5 most significant pathway and their corresponding genes. Different colors of the lines represent different KEGG pathway. The size of each node indicates the number of genes associated with that term, while the color of each node represents the \log_2 (fold change) of the gene expression, with the color gradient showing the direction and

magnitude of the change (red for up-regulated genes, blue for down-regulated genes). Fig. S4. GSEA of RNA-seq data in ACB at different developmental stages. (A-R) GSEA for different comparisons of diapause stages. Panels A-C show GSEA for GO terms in the diapause vs. pre-diapause comparison, while panels D-F show GSEA for KEGG pathways in the same comparison. Panels G-H present GSEA for GO terms in the diapause vs. no-diapause comparison, and panels J-L show GSEA for KEGG pathways in this comparison. Panels M-O display GSEA for GO terms in the pre-diapause vs. no-diapause comparison, and panels P-R illustrate GSEA for KEGG pathways in this comparison. In all panels, the X-axis represents the comparison between the treatment group (left) and the control group (right). Running enrichment score is an integral enrichment score, reflecting the distribution of genes in the gene set across the entire gene expression dataset. Peaks indicate highly enriched gene sets, while valleys represent highly suppressed gene sets, corresponding to the maximum and minimum points of the running enrichment score. The NES is a standardized transformation of the running enrichment score for comparison between different gene sets. Positive NES indicates enrichment, while negative NES indicates suppression. Core contributing genes, which drive pathway activation or suppression, are annotated on the curves. Fig. S5. Enrichment analysis of differential expressed genes of lncRNAs target gene. (A-C) PCA plot showing the distribution of samples based on their expression profiles of lncRNAs transcripts in different developmental stages of the ACB. The plot illustrates the degree of similarity or dissimilarity among the samples, with samples that cluster together being more similar in their expression profiles. (D-E) Dot plot of differential expressed genes of lncRNAs target gene GO and KEGG enrichment. The circle color indicates the enrichment adjusted p -value, and the dot size indicates the number of DEGs in the functional class or pathway. (F) The relationship network among lncRNAs, target genes, and KEGG pathways of pre-diapause vs. no-diapause. Fig. S6. Exploring translational characteristics during diapause in Asian Corn Borer. (A-C) Quality check of Ribo-seq data using Ribo-TISH for the samples of no-diapause (A), pre-diapause, (B) and diapause (C). For each graph it includes the length distribution for the Ribo-seq reads as a histogram on the top. The 3-nucleotide periodicity of the RPFs mapped on all known protein-coding genes is shown for each read length. This is shown using a histogram of read coverage in the three frames, a bar plot of the number of RPFs in each position around the start codon and stop codon, and lastly a density plot for read coverage on the coding sequence across all genes. (D) Read length distribution of Ribo-seq data from different stages. Representative results are from three independent experiments. (E) Read length-based stratification of the percentage of P-sites in the three frames along the 5' UTR, CDS, and 3' UTR. (F) Ribo-seq data illustrating the percentage of P-sites in the 5' UTR, CDS, and 3' UTR of mRNAs (left). Percentage of mRNA sequence region lengths (right). (G) Schematic overview of ORF types detected by Ribo-seq. (H) Comparison between different sample regarding the statistics on active translated ORFs. (I) Meta-gene heatmaps for various read lengths show the signal at the 5' end (upper panel) and 3' end (lower panel) of reads aligned around the start and stop codons. (J) Bar plots showing the metagenome read distribution of the 5' end of reads from Ribo-seq data around the translation start codon and stop codon. Read position relative to frame (0, 1, 2) are shown in different colors. The schematic shows the ribosome binding around start/stop codon. Fig. S7. Enrichment analysis of differential expressed genes of lncRNAs target gene. (A-C) PCA plot showing the distribution of samples based on their expression profiles of lncRNAs transcripts in different developmental stages of the ACB. The plot illustrates the degree of similarity or dissimilarity among the samples, with samples that cluster together being more similar in their expression profiles. (D-E) Dot plot of differential expressed genes of lncRNAs target gene GO and KEGG enrichment. The circle color indicates the enrichment adjusted p -value, and the dot size indicates the number of DEGs in the functional class or pathway. (F) The relationship network among lncRNAs, target genes, and KEGG pathways of pre-diapause vs. no-diapause. Fig. S8. Verification of the effects of dsRNA, HSAP70 inhibitor, HSAP70 activator and MG132 treatment. (A) The schematic diagram illustrates the process of dsRNA production. Target genes were integrated into the

pET30a vector, and the resulting recombinant plasmid was then introduced into BL21(DE3) RNase III. The T7 promoter within the vector facilitated RNA transcription, induced by isopropyl-beta-D-thiogalactopyranoside (IPTG). (B-E) Expression levels of Ofur-JHMT, Ofur-Inc-001217, Ofur-Inc-005152, and Ofur-Inc-002794 in ACB heads treated with dsRNA were analyzed. (F) HSP70 protein levels were examined in the heads of ACB larvae treated with the HSP70 inhibitor VER15508. (G) Image J was utilized to measure the relative gray values of Western blot data from (F). (H) mRNA levels of HSP70 were assessed in the heads of ACB larvae treated with the HSP70 activator ML346. (I) HSP70 protein levels were determined in the heads of ACB larvae treated with MG132. (J) Image J was used to quantify the relative gray values of Western blot data from (I). Data are mean \pm SEM. *p*-values from unpaired two-tailed Student's *t*-test are indicated, n.s. represents no significant difference, * represents $p < 0.05$, ** represents $p < 0.01$, and *** represents $p < 0.001$.

Additional file 2: Table S1. Statistical analysis of JH synthesis-related IncRNAs and their corresponding target genes from diapause vs. pre-diapause. Table S2. dsRNA Primers used in this study. Table S3. qRT-PCR Primers used in this study.

Additional file 3: Table S4. All enrichment term/pathway gene annotation.

Additional file 4: Table S5. All gene lists corresponding to figures.

Additional file 5: Fig. S9. Uncropped blots for Fig. 8I of anti-Actin.

Additional file 6: Fig. S10. Uncropped blots for Fig. 8I of anti-Hsp70.

Additional file 7: Fig. S11. Uncropped blots for Fig. 8F of anti-Actin.

Additional file 8: Fig. S12. Uncropped blots for Fig. 8F of anti-Hsp70.

Acknowledgements

We would like to express our sincere gratitude to Professor Jie Shen from the College of Plant Protection at China Agricultural University for his invaluable assistance in providing vectors used in this study. We would also like to acknowledge the 2115 Talent Development Program of China Agricultural University.

Authors' contributions

XY designed and performed the experiments, analyzed the data, and wrote the manuscript. XZ designed and performed the experiments. ZZ analyzed the data. JD designed experiments, analyzed the data, and wrote and edited the manuscript. All authors have read and approved the final manuscript.

Funding

This research was supported by National Natural Science Foundation of China (Grant No. 32122017) to Juan Du.

Availability of data and materials

All data of RNA-seq and Ribo-seq have been deposited under NCBI and are available under PRJNA1026478 and PRJNA1035983. The main code for the analysis and visualization used in this study is publicly accessible on Zenodo (<https://zenodo.org/records/12789428>). This repository includes all scripts and tools necessary for replicating the analyses presented. For additional information or specific components not available in the repository, please contact the corresponding author. Supporting data for all figure include in Additional files 2, 3, and 4. All data generated or analyzed during this study are included in this published article, its supplementary information, files and publicly available repositories.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Consent for publication has been obtained from all relevant participants and organizations.

Competing interests

The authors declare that they have no competing interests.

Received: 14 December 2023 Accepted: 2 September 2024

Published online: 13 September 2024

References

- Horwath KL, Duman JG. Involvement of the circadian system in photoperiodic regulation of insect antifreeze proteins. *J Exp Zool.* 1982;219(2):267–70.
- Thompson AW, Ortí G. Annual killifish transcriptomics and candidate genes for metazoan diapause. *Mol Biol Evol.* 2016;33(9):2391–5.
- Saunders DS. Dormancy, diapause, and the role of the circadian system in insect photoperiodism. *Annu Rev Entomol.* 2020;65(1):373–89.
- Gao C, Li Q, Yu J, Li S, Cui Q, Hu X, Chen L, Zhang SO. Endocrine hormones couple fat rationing to dauer diapause through *HNF4a* nuclear receptors. *Sci China Life Sci.* 2021;64(12):2153–74.
- Reynolds JA, Bautista-Jimenez R, Denlinger DL. Changes in histone acetylation as potential mediators of pupal diapause in the flesh fly, *Sarcophaga bullata*. *Insect Mol Biol.* 2016;76:29–37.
- Burggren WW. Epigenetics in insects: mechanisms, phenotypes and ecological and evolutionary implications. In: Verlinden H, editor. *Insect epigenetics*. Iowa: Academic; 2017. p. 1–30.
- Reynolds JA, Peyton JT, Denlinger DL. Changes in microRNA abundance may regulate diapause in the flesh fly, *Sarcophaga bullata*. *Insect Biochem Mol Biol.* 2017;84:1–14.
- Zhang XC, Zabinsky R, Teng YD, Cui MX, Han M. microRNAs play critical roles in the survival and recovery of *Caenorhabditis elegans* from starvation-induced L1 diapause. *Proc Natl Acad Sci U S A.* 2011;108(44):17997–8002.
- King AM, MacRae TH. Insect heat shock proteins during stress and diapause. *Annu Rev Entomol.* 2015;60(1):59–75.
- Rinehart JP, Li A, Yocum GD, Robich RM, Hayward SAL, Denlinger DL. Up-regulation of heat shock proteins is essential for cold survival during insect diapause. *Proc Natl Acad Sci U S A.* 2007;104(27):11130–7.
- Zhang X, Du W, Zhang J, Zou Z, Ruan C. High-throughput profiling of diapause regulated genes from *Trichogramma dendrolimi* (Hymenoptera: Trichogrammatidae). *BMC Genomics.* 2020;21(1):864.
- Bogus MI, Scheller K. Allatotropin released by the brain controls larval molting in *Galleria mellonella* by affecting juvenile hormone synthesis. *Int J Dev Biol.* 1996;40(1):205–10.
- Denlinger DL, Yocum GD, Rinehart JP. Hormonal control of diapause. *Insect Endocrinol.* 2012;8:430–63.
- Cui J, Lin K, Xu L, Yue F, Yu L, Zhang Q. Transcriptome analysis of beet webworm shows that histone deacetylase may affect diapause by regulating juvenile hormone. *Insects.* 2022;13(9):835.
- Chippendale GM, Yin CM. Larval diapause of the European corn borer, *Ostrinia nubilalis*: further experiments examining its hormonal control. *J Insect Physiol.* 1979;25(1):53–8.
- Sieber R, Benz G. Juvenile hormone in larval diapause of the codling moth, *Laspeyresia pomonella* L. (Lepidoptera, Tortricidae). *Experientia.* 1977;33(12):1598–9.
- Eizaguirre M, Prats J, Abellana M, Carmen L, Llovera M, Canela R. Juvenile hormone and diapause in the mediterranean corn borer, *Sesamia nonagrioides*. *J Insect Physiol.* 1998;44(5–6):419–25.
- Eizaguirre M, Schafellner C, López C, Sehnael FE. Relationship between an increase of juvenile hormone titer in early instars and the induction of diapause in fully grown larvae of *Sesamia nonagrioides*. *J Insect Physiol.* 2005;51(10):1127–34.
- Schafellner C, Eizaguirre M, López C, Sehnael F. Juvenile hormone esterase activity in the pupating and diapausing larvae of *Sesamia nonagrioides*. *J Insect Physiol.* 2008;54(6):916–21.
- Munyiri FN, Ishikawa Y. Endocrine changes associated with metamorphosis and diapause induction in the yellow-spotted longicorn beetle, *Psacotheta hilaris*. *J Insect Physiol.* 2004;50(11):1075–81.
- Singtripop T, Manaboon M, Tatum N, Kaneko Y, Sakurai S. Hormonal mechanisms underlying termination of larval diapause by juvenile

- hormone in the bamboo borer, *Omphisa fuscidentalis*. *J Insect Physiol.* 2008;54(1):137–45.
22. Singtripop T, Wanichacheewa S, Sakurai S. Juvenile hormone-mediated termination of larval diapause in the bamboo borer, *Omphisa fuscidentalis*. *Insect Mol Biol.* 2000;30(8–9):847–54.
 23. Zhang G, Storey JM, Storey KB. Chaperone proteins and winter survival by a freeze tolerant insect. *J Insect Physiol.* 2011;57(8):1115–22.
 24. Wang HS, Wang XH, Zhou CS, Huang LH, Zhang SF, Guo W, Kang L. cDNA cloning of heat shock proteins and their expression in the two phases of the migratory locust. *Insect Mol Biol.* 2007;16(2):207–19.
 25. Wang H, Li K, Zhu JY, Fang Q, Ye GY, Wang H, Li K, Zhu JY. Cloning and expression pattern of heat shock protein genes from the endoparasitoid wasp, *Pteromalus puparum* in response to environmental stresses. *Arch Insect Biochem Physiol.* 2012;79(4–5):247–63.
 26. Michaud MR, Teets NM, Peyton JT, Blobner BM, Denlinger DL. Heat shock response to hypoxia and its attenuation during recovery in the flesh fly, *Sarcophaga crassipalpis*. *J Insect Physiol.* 2011;57(1):203–10.
 27. Liu G, Roy J, Johnson EA. Identification and function of hypoxia-response genes in *Drosophila melanogaster*. *Physiol Genomics.* 2006;25(1):134–41.
 28. Hwang JS, Go HJ, Goo TW, Yun EY, Choi KH, Seong SI, Lee SM, Lee BH, Kim I, Chun T, et al. The analysis of differentially expressed novel transcripts in diapausing and diapause-activated eggs of *Bombyx mori*. *Arch Insect Biochem Physiol.* 2005;59(4):197–201.
 29. Gkouvitsas T, Kontogiannatos D, Kourti A. Differential expression of two small Hsps during diapause in the corn stalk borer *Sesamia nonagrioides* (Lef.). *J Insect Physiol.* 2008;54(12):1503–10.
 30. Kankare M, Salminen T, Laiho A, Vesala L, Hoikkala A. Changes in gene expression linked with adult reproductive diapause in a northern malt fly species: a candidate gene microarray study. *BMC Ecol.* 2010;10(1):3.
 31. Lu MX, Cao SS, Du YZ, Liu ZX, Liu P, Li J. Diapause, signal and molecular characteristics of overwintering *Chilo suppressalis* (Insecta: Lepidoptera: Pyralidae). *Sci Rep.* 2013;3:211.
 32. Williamson DS, Borgognoni J, Clay A, Daniels Z, Dokurno P, Drysdale MJ, Foloppe N, Francis GL, Graham CJ, Howes R, et al. Novel adenosine-derived inhibitors of 70 kDa heat shock protein, discovered through structure-based design. *J Med Chem.* 2009;52(6):1510–3.
 33. Calamini B, Silva MC, Madoux F, Hutt DM, Khanna S, Chalfant MA, Saldanha SA, Hodder P, Tait BD, Garza D, et al. Small-molecule proteostasis regulators for protein conformational diseases. *Nat Chem Biol.* 2011;8(2):185–96.
 34. Awasthi N, Wagner BJ. Upregulation of heat shock protein expression by proteasome inhibition: an antiapoptotic mechanism in the lens. *Invest Ophthalmol Vis Sci.* 2005;46(6):2082–91.
 35. Kim HJ, Joo HJ, Kim YH, Ahn S, Chang J, Hwang KB, Lee DH, Lee KJ. Systemic analysis of heat shock response induced by heat shock and a proteasome inhibitor MG132. *PLoS One.* 2011;6(6):e20252.
 36. Teets NM, Marshall KE, Reynolds JA. Molecular mechanisms of winter survival. *Annu Rev Entomol.* 2023;68(1):319–39.
 37. Karp X. Hormonal regulation of diapause and development in nematodes, insects, and fishes. *Front Ecol Evol.* 2021;9:735924.
 38. Nouzova M, Rivera-Pérez C, Noriega FG. Omics approaches to study juvenile hormone synthesis. *Curr Opin Insect Sci.* 2018;29:49–55.
 39. Miki T, Shinohara T, Chafino S, Noji S, Tomioka K. Photoperiod and temperature separately regulate nymphal development through JH and insulin/TOR signaling pathways in an insect. *Proc Natl Acad Sci U S A.* 2020;117(10):5525–31.
 40. Numata H, Shintani Y. Diapause in univoltine and semivoltine life cycles. *Annu Rev Entomol.* 2023;68(1):257–76.
 41. Gotthard K, Wheat CW. Diapause: circadian clock genes are at it again. *Curr Biol.* 2019;29(23):R1245–6.
 42. Shaw PJ, Hasebe M, Shiga S. Clock gene-dependent glutamate dynamics in the bean bug brain regulate photoperiodic reproduction. *PLoS Biol.* 2022;20(9):e3001734.
 43. Olsen L, Thum E, Rohner N. Lipid metabolism in adaptation to extreme nutritional challenges. *Dev Cell.* 2021;56(10):1417–29.
 44. Pei J, Xu Y, Zong S, Ren L. Transcriptomic and metabolomic data reveal the key metabolic pathways affecting *Streltzoviella insularis* (Staudinger) (Lepidoptera: Cossidae) larvae during overwintering. *Front Physiol.* 2021;12:655059.
 45. Tao YD, Liu Y, Wan XS, Xu J, Fu DY, Zhang JZ. High and low temperatures differentially affect survival, reproduction, and gene transcription in male and female moths of *Spodoptera frugiperda*. *Insects.* 2023;14(12):958.
 46. Bulut-Karslioglu A, Biechele S, Jin H, Macrae TA, Hejna M, Gertsenstein M, Song JS, Ramalho-Santos M. Inhibition of mTOR induces a paused pluripotent state. *Nature.* 2016;540(7631):119–23.
 47. Chen H, Wu G, Zhou H, Dai X, Steeghs NWF, Dong X, Zheng L, Zhai Y. Hormonal regulation of reproductive diapause that occurs in the year-round mass rearing of *Bombus terrestris* queens. *J Proteome Res.* 2021;20(5):2240–50.
 48. Denlinger DL. Insect diapause: from a rich history to an exciting future. *J Exp Biol.* 2023;226(4):jeb245329.
 49. Li H, Lu Q, Li Y, Yan Y, Yin Z, Guo J, Xu W. Smurf participates in *Helicoverpa armigera* diapause by regulating the transforming growth factor- β signaling pathway. *Insect Sci.* 2022;29(5):1251–61.
 50. Hao K, Jarwar AR, Ullah H, Tu X, Nong X, Zhang Z. Transcriptome sequencing reveals potential mechanisms of the maternal effect on egg diapause induction of *Locusta migratoria*. *Int J Mol Sci.* 2019;20(8):1974.
 51. Duan TF, Li L, Wang HC, Pang BP. MicroRNA *miR-2765-3p* regulates reproductive diapause by targeting FoxO in *Galeruca daurica*. *Insect Sci.* 2022;30(2):279–92.
 52. Sim C, Kang DS, Kim S, Bai X, Denlinger DL. Identification of FOXO targets that generate diverse features of the diapause phenotype in the mosquito *Culex pipiens*. *Proc Natl Acad Sci U S A.* 2015;112(12):3811–6.
 53. Du J, Zhao P, Wang J, Ma S, Yao L, Zhu X, Yang X, Zhang X, Sun Z, Liang S, et al. Pupal diapause termination and transcriptional response of *Antheraea pernyi* (Lepidoptera: Saturniidae) triggered by 20-Hydroxyecdysone. *Front Physiol.* 2022;13:888643.
 54. Guo S, Tian Z, Wu QW, King-Jones K, Liu W, Zhu F, Wang XP. Steroid hormone ecdysone deficiency stimulates preparation for photoperiodic reproductive diapause. *PLoS Genet.* 2021;17(2):e1009352.
 55. Ma HY, Li YY, Li L, Tan Y, Pang BP. Regulation of juvenile hormone on summer diapause of *Galeruca daurica* and its pathway analysis. *Insects.* 2021;12(3):237.
 56. Smykal V, Bajgar A, Provaznik J, Fexova S, Buricova M, Takaki K, Hodkova M, Jindra M, Dolezel D. Juvenile hormone signaling during reproduction and development of the linden bug, *Pyrrhocoris apterus*. *Insect Mol Biol.* 2014;45:69–76.
 57. Baker DA, Russell S. Gene expression during *Drosophila melanogaster* egg development before and after reproductive diapause. *BMC Genomics.* 2009;10(1):242.
 58. Batz ZA, Brent CS, Marias MR, Sugijanto J, Armbruster PA. Juvenile hormone III but not 20-hydroxyecdysone regulates the embryonic diapause of *Aedes albopictus*. *Front Physiol.* 2019;10:1352.
 59. Gong J, Tian S, Zhou X, Yang H, Zhu Y, Hou Y. Transcriptional response of silkworm (*Bombyx mori*) eggs to O₂ or HCl treatment. *Int J Mol Sci.* 2016;17(12):1838.
 60. Liu Y, Wang R, Su L, Zhao S, Dai X, Chen H, Wu G, Zhou H, Zheng L, Zhai Y. Integrative proteomic and phosphoproteomic analyses revealed complex mechanisms underlying reproductive diapause in *Bombus terrestris* queens. *Insects.* 2022;13(10):862.
 61. Flannagan RD, Tammariello SP, Joplin KH, Cikra-Ireland RA, Yocum GD, Denlinger DL. Diapause-specific gene expression in pupae of the flesh fly *Sarcophaga crassipalpis*. *Proc Natl Acad Sci U S A.* 1998;95(10):5616–20.
 62. Košťál V, Šimůnková P, Kobelková A, Shimada K. Cell cycle arrest as a hallmark of insect diapause: changes in gene transcription during diapause induction in the drosophilid fly, *Chymomyces costata*. *Insect Mol Biol.* 2009;39(12):875–83.
 63. Li H-Y, Wang T, Yang Y-P, Geng S-L, Xu W-H. TGF- β signaling regulates p-Akt levels via *PP2A* during diapause entry in the cotton bollworm, *Helicoverpa armigera*. *Insect Mol Biol.* 2017;87:165–73.
 64. Jiang F, Chang G, Li Z, Abouzaid M, Du X, Hull JJ, Ma W, Lin Y. The HSP/co-chaperone network in environmental cold adaptation of *Chilo suppressalis*. *Int J Biol Macromol.* 2021;187:780–8.
 65. Sakano D, Li B, Xia Q, Yamamoto K, Fujii H, Aso Y. Genes encoding small heat shock proteins of the silkworm, *Bombyx mori*. *Biosci Biotechnol Biochem.* 2014;70(10):2443–50.
 66. Gkouvitsas T, Kontogiannatos D, Kourti A. Cognate *Hsp70* gene is induced during deep larval diapause in the moth *Sesamia nonagrioides*. *Insect Mol Biol.* 2009;18(2):253–64.

67. Gkouvtas T, Kontogiannatos D, Kourti A. Expression of the *Hsp83* gene in response to diapause and thermal stress in the moth *Sesamia nonagrioides*. *Insect Mol Biol*. 2009;18(6):759–68.
68. Guo J, Zhang H, Wang Z, He K. Effects of photoperiod and temperature on diapause induction in *Ostrinia furnacalis* (Lepidoptera: Crambidae). *Acta Entomol Sin*. 2013;56(9):996–1003.
69. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114–20.
70. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol*. 2019;37(8):907–15.
71. Pertea M, Pertea GM, Antonescu CM, Chang T-C, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol*. 2015;33(3):290–5.
72. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 2014;30(7):923–30.
73. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
74. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation*. 2021;2(3):100141.
75. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009;10(3):R25.
76. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15–21.
77. Chothani S, Adami E, Ouyang JF, Viswanathan S, Hubner N, Cook SA, Schafer S, Rackham OJL. deltaTE: detection of translationally regulated genes by integrative analysis of Ribo-seq and RNA-seq data. *Curr Protoc Mol Biol*. 2019;129(1):e108.
78. Zhang P, He D, Xu Y, Hou J, Pan B-F, Wang Y, Liu T, Davis CM, Ehli EA, Tan L, et al. Genome-wide identification and differential analysis of translational initiation. *Nat Commun*. 2017;8(1):1749.
79. Darling AE, Luria F, Tebaldi T, Bernabò P, Groen EJM, Gillingwater TH, Viero G. riboWaltz: optimization of ribosome P-site positioning in ribosome profiling data. *PLoS Comput Biol*. 2018;14(8):e1006169.
80. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27–30.
81. Zhang J. ClusterGVis: one-step to cluster and visualize gene expression matrix 2023. Accessed 3 Feb 2023.
82. Zhang J. BioSeqUtils: extract sequence from genome according to annotation file 2023. Accessed 3 Feb 2023.
83. Kang Y-J, Yang D-C, Kong L, Hou M, Meng Y-Q, Wei L, Gao G. CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. *Nucleic Acids Res*. 2017;45(W1):W12–6.
84. Sun L, Luo H, Bu D, Zhao G, Yu K, Zhang C, Liu Y, Chen R, Zhao Y. Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Res*. 2013;41(17):e166.
85. Wang L, Park HJ, Dasari S, Wang S, Kocher J-P, Li W. CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. *Nucleic Acids Res*. 2013;41(6):e74.
86. Mei Y, Jing D, Tang S, Chen X, Chen H, Duanmu H, Cong Y, Chen M, Ye X, Zhou H, et al. InsectBase 2.0: a comprehensive gene resource for insects. *Nucleic Acids Res*. 2022;50(D1):D1040–5.
87. Yu Z, Shi J, Jiang X, Song Y, Du J, Zhao Z. Neuropeptide F regulates feeding via the juvenile hormone pathway in *Ostrinia furnacalis* larvae. *Pest Manag Sci*. 2023;79(3):1193–203.
88. Ma ZZ, Zhou H, Wei YL, Yan S, Shen J. A novel plasmid-*Escherichia coli* system produces large batch dsRNAs for insect gene silencing. *Pest Manag Sci*. 2020;76(7):2505–12.
89. Ma Z, Zheng Y, Chao Z, Chen H, Zhang Y, Yin M, Shen J, Yan S. Visualization of the process of a nanocarrier-mediated gene delivery: stabilization, endocytosis and endosomal escape of genes for intracellular spreading. *J Nanobiotechnology*. 2022;20(1):124.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.