Proceedings of the 6th International Conference on Biological Engineering and Medical Science
DOI: 10.54254/2753-8818/2026.35945

Dynamic Mechanism of Hepatopancreatic Transcriptomic Response in Litopenaeus Vannamei under PirAB Toxin Challenge: A Time-Series Reanalysis Based on Public Databases

Yuetong Hao, Tianming Wang*

College of Fisheries and Life Science, Shanghai Ocean University, Shanghai, China
**Corresponding Author. Email: wtm4411@163.com*

Abstract. Acute Hepatopancreatic Necrosis Disease (AHPND) poses a major threat to the shrimp aquaculture industry, and its pathogenic agent is the PirAB toxin. Although it is known to target the hepatopancreas, the early events triggered by toxin attack and their dynamic mechanisms along disease progression remain to be elucidated. In this study, time-series transcriptomic data from public databases (control group and 0.5, 1, 2, 4, 6 hours after toxin treatment) were used to reveal a "double-hit" model of PirAB toxin pathogenesis through K-means clustering and key gene expression analysis. The first hit (0.5 h) was characterized by a sharp downregulation of transcription regulatory factors such as zinc finger proteins, leading to paralyzed transcriptional regulation. The second hit (6 h) manifested as sustained inhibition of metabolic genes and upregulation of apoptosis/degradation genes including matrix metalloproteinases, resulting in functional failure and tissue ulceration. This model identified molecular abnormalities at the very early stage, suggesting that zinc finger proteins, metallothioneins and other molecules can serve as biomarkers for ultra-early warning, providing new insights into the early diagnosis of AHPND.

Keywords: Litopenaeus vannamei, PirAB toxin, hepatopancreas, transcriptomic reanalysis, double-hit model, public database

1. Introduction

Litopenaeus vannamei is one of the most economically important shrimp species in global aquaculture [1]. However, Acute Hepatopancreatic Necrosis Disease (AHPND), caused by Vibrio parahaemolyticus carrying the pirA/pirB toxin genes, has severely impacted shrimp production [2]. AHPND is characterized by acute onset and high mortality, with rapid necrosis and functional failure of the hepatopancreatic tissue as the core pathological change [3].

As a multifunctional organ in crustaceans, the hepatopancreas undertakes multiple physiological functions including detoxification, metabolism, digestive enzyme secretion and immune defense [4]. Previous studies have shown that the PirAB toxin complex, composed of PirA and PirB proteins, is

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/2026.35945

the direct pathogenic agent of AHPND and can specifically target the shrimp hepatopancreas [5, 6]. Although PirAB toxin has been widely recognized as the core pathogenic factor of AHPND, there is still a lack of systematic dynamic analysis on how it induces the collapse of shrimp hepatopancreatic function step by step and across different time points at the molecular level. Most existing studies focus on terminal pathological observation [3] or transient expression of specific immune factors [7, 8], often ignoring the intrinsic regulatory logic from molecular warning at the initial stage of toxin attack to systematic pathological damage in the later stage.

In recent years, public databases such as NCBI GEO/SRA have accumulated a large amount of high-quality raw sequencing data, providing opportunities for multi-dimensional secondary development and in-depth mining [9]. Based on this, the present study used time-series transcriptomic data to redefine key stages of disease progression through the K-means clustering algorithm, and focused on exploring the "ultra-early" (0.5 h) regulatory signals neglected in original studies. The aim was to clarify the molecular cascade mechanism by which PirAB toxin gradually disrupts hepatopancreatic function and to investigate whether there are key early events that determine subsequent pathological progression. Based on the analytical results, we proposed a "double-hit" model, expecting to provide a theoretical basis for precise intervention in the pathogenesis of AHPND by constructing a dynamic molecular network.

2. Materials and methods

2.1. Data source and retrieval

All transcriptomic data in this study were obtained from the Gene Expression Omnibus (GEO) database of the National Center for Biotechnology Information (NCBI). The retrieval strategy was set as follows: in GEO DataSets, the search query (transcriptome AND "Litopenaeus vannamei") AND ("2020/01/01"[Update Date] : "3000"[Update Date]) was used for screening.

2.2. Computational environment and software platform

All bioinformatics analyses were performed using Python (v3.8.10) programming in the Jupyter Notebook (v6.4.5) environment. The main third-party libraries used included: Pandas (v1.5.3) and NumPy for data cleaning and matrix operations; Scikit-learn (v1.0.2) for machine learning and clustering analysis; and Matplotlib (v3.4.3) and Seaborn (v0.11.2) for data visualization.

2.3. Data preprocessing and quality control

The Series Matrix File and platform annotation file of GSE200137 [9] were downloaded from NCBI. Gene ID mapping was performed using Pandas. For cases where multiple probes corresponded to the same gene, the median expression value was adopted as the representative value for that gene.

To reduce interference from background expression noise, the raw data were normalized by calculating Counts Per Million (CPM) and performing $\log_2(\text{CPM}+1)$ transformation. Genes with expression levels below 10 in fewer than 3 samples were filtered out to ensure the reliability of subsequent clustering analysis.

DOI: 10.54254/2753-8818/2026.35945

2.4. Principal component analysis (PCA)

Principal component analysis was employed to evaluate the transcriptomic similarity among different groups of samples. Dimensionality reduction was conducted on the normalized expression matrix using Scikit-learn. The first two principal components (PC1 and PC2) were extracted as coordinate axes, and the proportion of explained variance for each principal component was calculated to quantify the evolutionary characteristics of the data along the time axis and the differences between groups.

2.5. K-means clustering of temporal expression patterns

To focus on the dynamic evolutionary trends of gene expression rather than absolute abundance, we first calculated the average expression difference of each gene at different time points (0.5, 1, 2, 4, 6 h) relative to the control group, followed by row-wise Z-score normalization. The K-means algorithm was used to cluster differentially expressed genes. The optimal number of clusters K was determined by the "elbow method" [10], i.e., by observing the slope of the within-cluster sum of squares (Inertia) as a function of K, and K=4 was finally selected. By plotting clustering trend graphs and heatmaps, the dynamic response patterns of different gene sets with the progression of toxin treatment time were systematically analyzed.

2.6. Gene annotation and heatmap construction of key functional genes

Based on the K-means clustering results, functional genes related to stress defense, transcriptional regulation, apoptosis, tissue remodeling, and metabolic detoxification were selectively screened. Temporal expression heatmaps were generated using Seaborn.clustermap with Z-score normalization, and columns were sorted by sampling time. By analyzing the expression kinetics of these key genes, their consistency with the K-means clustering patterns was verified, thereby constructing the molecular evidence chain for the "Double-hit" model.

3. Results and analysis

3.1. Dataset overview and global transcriptomic changes

This study carefully tested and analyzed the gene expression patterns of 30 samples in total. These samples include one control group and five separate groups treated with PirAB toxin at different time points, with n=5 for each group. After $\log_2(\text{CPM} + 1)$ transformation, the expression matrix showed good distribution features. Principal component analysis (PCA) results are shown in Figure 1. PC1 and PC2 together explain 36.0% of the total variance, accounting for 23.9% and 12.1% respectively. Samples from the 0.5 h control group cluster closely together, which shows the experiment has high repeatability. The PC1 axis (23.9%) shows a large numerical shift caused by one outlier sample in the control group. This deviation may come from physiological differences among individual shrimp or random technical noise during library construction. But the core infection trend is clearly reflected on the PC2 axis (12.1%). As treatment time increases, samples in each group show regular linear movement along the vertical axis. This fully proves the disease-causing process brought by PirAB toxin is step-by-step and stable. It also builds a reliable base for later screening of time-series genes with different expression levels.

As the length of toxin treatment time gets longer, samples show a clear time-related changing trend in the PCA space. In the early response period of 0.5 to 1 hour, treatment groups stay quite

DOI: 10.54254/2753-8818/2026.35945

close to the control group. This reflects a smooth transition at the very beginning of infection. As time goes on gradually, sample points show obvious linear separation along the PC2 axis. The overall deviation reaches its maximum level at the 6-hour time point. This whole change path shows a clear time-based rule for the hepatopancreas as it shifts from a stable normal state to a diseased state. The separation between the 6 h treatment group and the control group is the most obvious, which suggests that changes in gene expression patterns are closely linked to the length of treatment time.

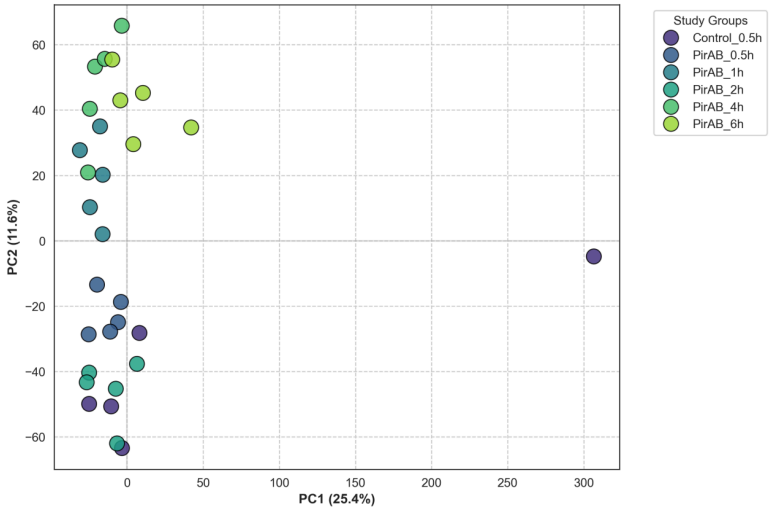

Figure 1. Principal component analysis (PCA) of the hepatopancreatic transcriptome in Litopenaeus vannamei. Different colors/shapes represent samples from the control group (Control_0.5h) and 0.5, 1, 2, 4, 6 h after PirAB toxin treatment, n=5 per group

3.2. Temporal distribution of differentially expressed genes (DEGs)

Gene screening was carried out with two set thresholds: $|\log_2 FC| > 1$ and $P < 0.05$. The results show that the number of DEGs caused by PirAB toxin presents a clear rising trend with infection time (Figures 2, 3):

At 0.5 hours after infection, 466 genes already showed clear responses to the toxin (247 upregulated, 219 downregulated). This shows PirAB toxin can quickly activate the host's signal sensing system. By the late 6-hour treatment stage, total DEG numbers rise to 1,105 (639 upregulated, 466 downregulated).

Volcano plot analysis shows that the degree of gene upregulation (Fold Change) rises obviously as time passes. Besides, many key functional genes that produce digestive enzymes and metabolism-related proteins show very obvious deep downregulation in the late stage. This change suggests gradual overall damage to hepatopancreatic tissue function.

DOI: 10.54254/2753-8818/2026.35945

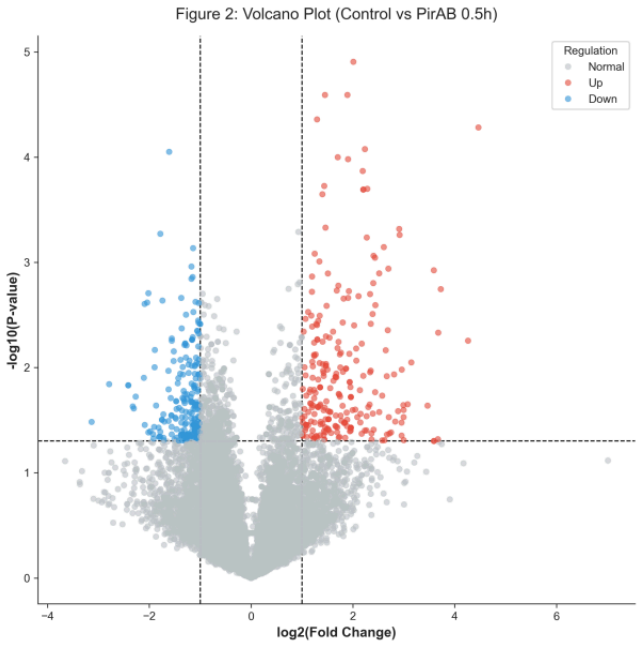

Figure 2. Volcano plot of differentially expressed genes: 0.5 h treatment group vs. control group

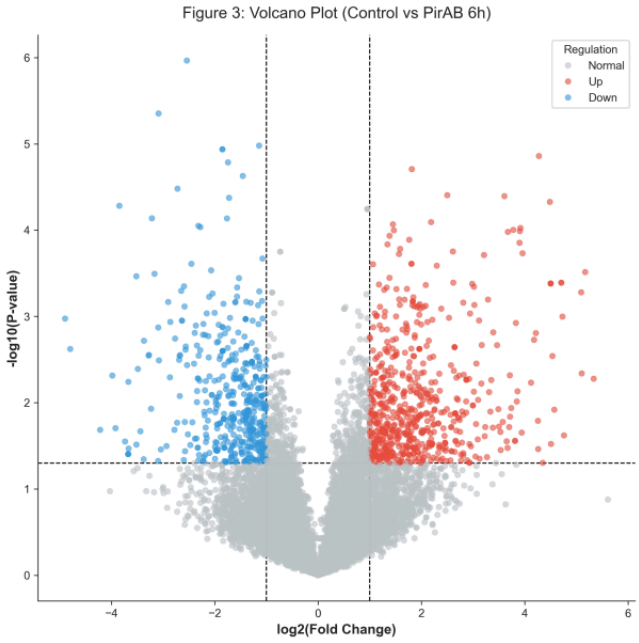

Figure 3. Volcano plot of differentially expressed genes: 6 h treatment group vs. control group

3.3. K-means clustering of temporal expression patterns of DEGs

To further reveal the dynamic regulatory logic of disease progression, this study used the K-means algorithm to divide DEGs into four transcription clusters (Cluster 1–4, Figure 4) with distinct characteristics, identifying four representative temporal expression patterns (Table 1). Elbow method analysis showed that when K=4, the within-cluster sum of squares (WCSS) tended to stabilize.

DOI: 10.54254/2753-8818/2026.35945

Table 1. Temporal expression characteristics and functional enrichment of genes in each cluster

Cluster	Number of Genes	Expression Pattern
Cluster 1	212	Rapidly downregulated immediately upon toxin exposure.
Cluster 2	562	Continuously upregulated, with a marked increase at the late stage.
Cluster 3	437	Continuously downregulated, exhibiting a sustained decline throughout the infection process.
Cluster 4	200	Rapidly upregulated at 0.5 h, indicating an acute response, followed by fluctuating expression.

Figure 4: K-means Clustering of Differential Expressed Genes

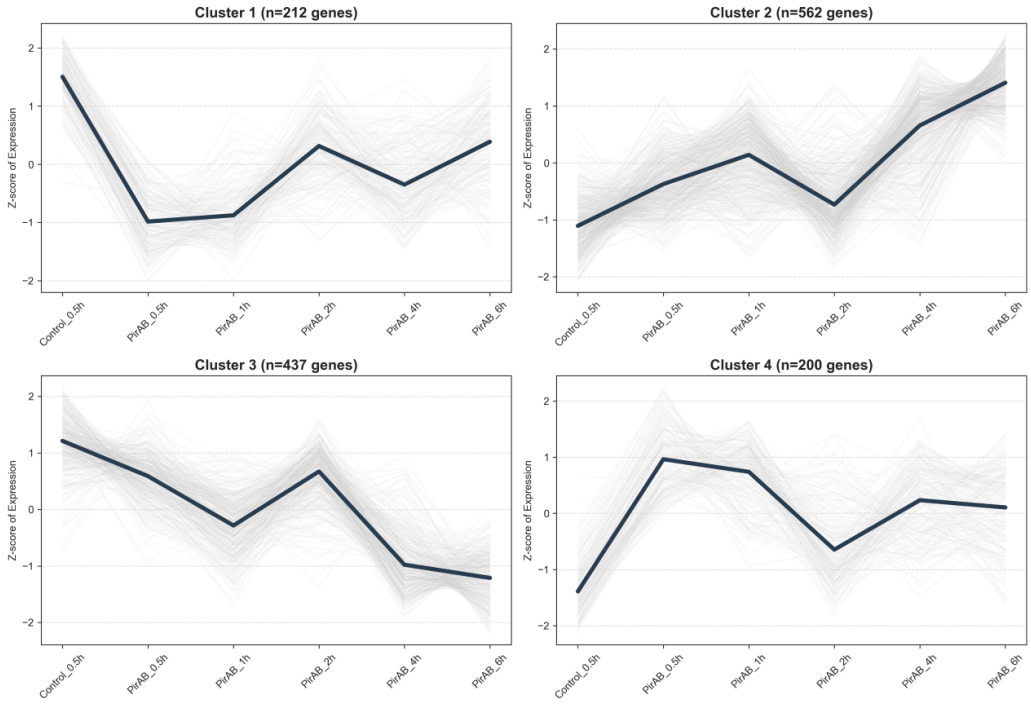

Figure 4. K-means clustering analysis of differentially expressed genes

From the analysis results above, comparison between Cluster 2 and Cluster 4 revealed that Cluster 4 showed acute elevation, representing the stress response after toxin invasion; Cluster 2 showed continuous elevation, representing systemic decline as toxin invasion deepened. Together, they constitute the "upregulation cascade" of toxin attack, corresponding to early stress response and late pathogenesis, respectively. Similarly, comparison between Cluster 1 and Cluster 3 showed that Cluster 1 exhibited a sharp decline, representing core metabolic switches instantly cut off by the toxin. Cluster 3 showed continuous decline, representing physiological functions gradually atrophying due to deteriorating health. The downregulation trajectories of these two clusters indicate the failure process of the metabolic network required for the hepatopancreas to maintain vital activities.

DOI: 10.54254/2753-8818/2026.35945

4. Temporal expression patterns of key functional genes

To further understand how the hepatopancreas shifts from active defense to functional damage under PirAB toxin attack, representative genes linked to defense, disease development and functional suppression were selected from k-means clusters. Their expression heatmaps were drawn (Figures 5–6), and a detailed study of their time-based expression changes was carried out.

The vertical axis stands for representative genes, and the horizontal axis stands for different time points. These include control_0.5h, PirAB_0.5h, PirAB_1h, PirAB_2h, PirAB_4h and PirAB_6h. The color shows the expression level after Z-score normalization.

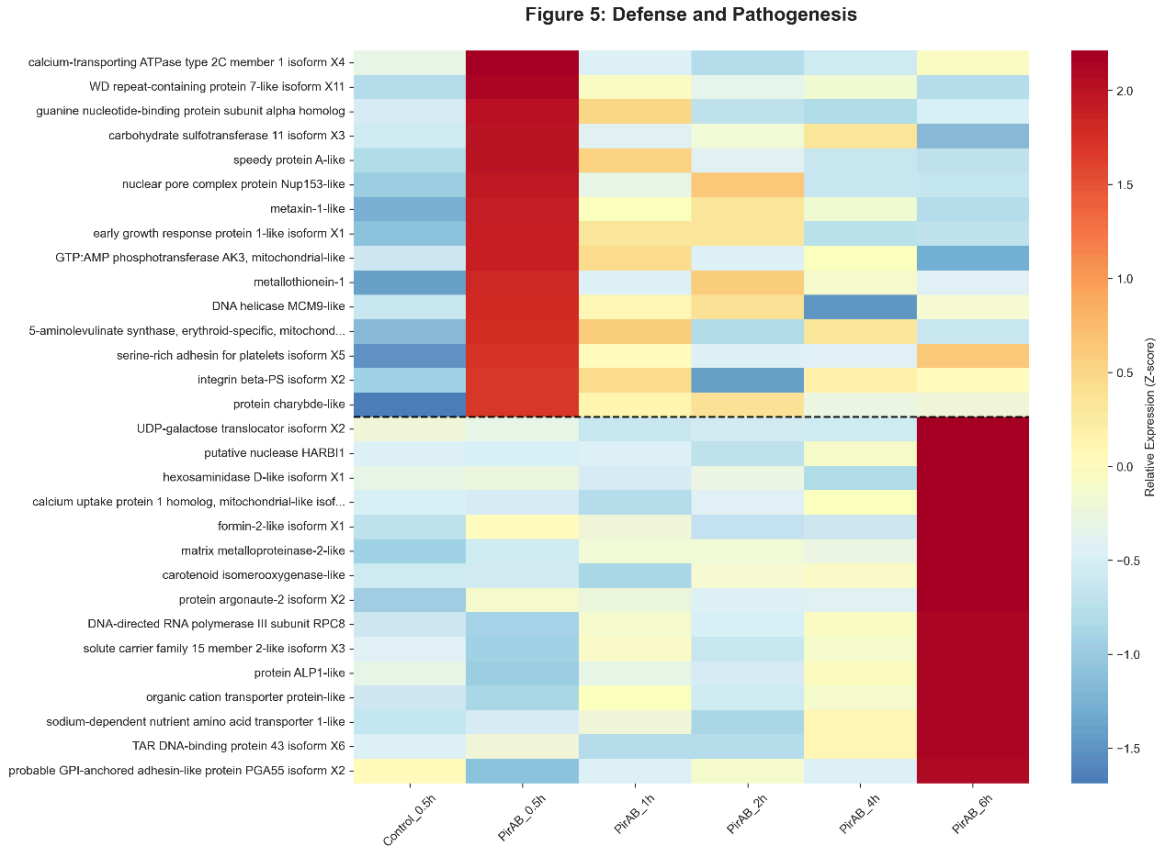

Figure 5. Temporal expression heatmap of genes related to defense and pathogenesis

DOI: 10.54254/2753-8818/2026.35945

Figure 6. Temporal expression heatmap of functional genes

4.1. First hit (0.5 h)

In the gene group linked to defense and disease development (Figure 5), genes above the dashed line showed clear upregulation at PirAB_0.5h. This change matches the rapid upregulation pattern of Cluster 4.

Environmental stress can cause abnormal metabolic changes in shrimp, which leads to the fast production of large amounts of reactive oxygen species (ROS) inside the body [11]. Metallothionein-1 (MT-1) is a key antioxidant protein, and its expression rose rapidly. It can bind to toxin-induced ROS and reduce the damage they cause [12], so it improves cell protection under oxidative stress [13]. At the same time, the expression of GTP:AMP Phosphotransferase AK3 (AK3) increased. It uses the TCA cycle in the mitochondrial matrix to replenish ADP and GDP, and helps keep energy metabolism in a stable state [14]. Besides, the upregulation of Calcium-transporting ATPase type 2C member 1 (ATP2C1) maintains calcium balance in the secretory pathway. It works by pumping Ca^{2+} into the Golgi apparatus, ensures normal protein modifications such as glycosylation, and supports the correct production of secretory proteins linked to detoxification and immune functions [15].

As shown in the upper part of Figure 6, the Cluster 1 pattern appeared together with the defense response. Many genes in the PirAB treatment group showed clear downregulation at 0.5 h. Among them, the expression levels of many members of the zinc finger protein family dropped quickly within 0.5 h. These include zinc finger SWIM domain-containing protein 8-like, zinc finger protein 90 homolog, zinc finger protein 501-like, zinc finger protein 570-like, and zinc finger MYM-type protein 2-like. Under environmental stress, zinc finger proteins act as key transcriptional regulatory

DOI: 10.54254/2753-8818/2026.35945

factors [16], and their expression levels fell obviously. This finding shows that these proteins may play a role in the early protection system of shrimp [17]. This result suggests that PirAB toxin may have very strong effects on suppressing transcription. It can quickly turn off the core genetic regulatory switches of the shrimp hepatopancreas at 0.5 h, and make the organism lose its ability to make up for functional damage.

4.2. Second hit (6 h)

The lower part of Figure 5 shows a group of genes with significantly increased expression levels at 6 h post-infection (Figure 5), corresponding to Cluster 1 (Figure 4). Among them, the expression level of matrix metalloproteinase-2 (MMP2) rose sharply, implying that extracellular matrix remodeling may occur in the late stage of infection [18]. Meanwhile, putative nuclease HARBI1 and proteins related to apoptosis and nucleic acid processing (such as TAR DNA-binding protein 43) were also synchronously upregulated. Given the widely conserved nature of TDP-43 [19], the upregulation at 6 h post-infection in the PirAB treatment group may indicate that the disease has entered the late stage, the organism can no longer maintain homeostasis, and apoptosis may have occurred [20].

As shown in the lower part of Figure 6, in the PirAB treatment group, some genes exhibited a long-term downregulation pattern over time, with expression levels reaching the lowest at 6 h, corresponding to the Cluster 3 pattern in Figure 4. Among them, cytochrome P450 18a1-like, an ecdysteroid-inactivating enzyme, has been mostly studied in insects and functions during developmental stages; its long-term downregulation may prevent the metabolism of ecdysteroids [21]. Probable citrate synthase 1, mitochondrial, a rate-limiting enzyme of the tricarboxylic acid cycle, when its expression is downregulated, the TCA cycle flux is significantly reduced, thereby inhibiting the amount of ATP produced from glucose via the glycolysis/TCA cycle pathway [22]. This explains the massive death of shrimp at 6 h, representing the comprehensive failure of hepatopancreatic physiological functions.

5. Discussion

The "two-hit" model proposed in this study carries clear biological significance. The first hit occurs within 0.5 h after toxin exposure, marked by the rapid downregulation of multiple zinc finger protein family members, which paralyzes the hepatopancreatic transcriptional regulatory network and causes the organism to lose its ability to coordinate defense responses at the very early stage. The second hit emerges at 6 h, manifesting as functional failure resulting from the sustained downregulation of metabolic enzymes (e.g., citrate synthase, CYP18A1) and tissue ulceration triggered by the upregulation of matrix metalloproteinase-2 (MMP2) and putative nuclease HARBI1, ultimately leading to irreversible collapse of the hepatopancreas. The scientific significance of this model lies in its first distinction, from a dynamic perspective, between two mechanistically independent stages—"transcriptional repression" and "tissue deconstruction"—and its advancement of the molecular warning window to 0.5 h. In terms of application value, early-sensitive genes that undergo significant changes as early as 0.5 h (e.g., zinc finger protein 570-like, metallothionein MT-1) can be targeted to develop ultra-early warning kits based on RT-LAMP or qRT-PCR, enabling rapid on-site detection of shrimp infection risk and buying precious time for aquaculture intervention. The main limitation of this study is that its conclusions are solely based on the transcriptomic level, and changes in mRNA may not fully reflect protein functions. Future studies are needed to further validate the molecular events and their causal relationships in this model by combining proteomic or metabolomic technologies.

DOI: 10.54254/2753-8818/2026.35945

References

[1] Tian, J., Wu, W., Li, J., et al. (2024). Development dilemma of Litopenaeus vannamei industry in China, current countermeasures taken and its implications for the world shrimp aquaculture industry.

[2] Lee, C.-T., Chen, I.-T., Yang, Y.-T., et al. (2015). The opportunistic marine pathogen Vibrio parahaemolyticus becomes virulent by acquiring a plasmid that expresses a deadly toxin. 112(34), 10798-803.

[3] Tran, L., Nunan, L., Redman, R. M., et al. (2013). Determination of the infectious nature of the agent of acute hepatopancreatic necrosis syndrome affecting penaeid shrimp. 105(1), 45-55.

[4] Vogt, G. J. D. O. A. O. (2020). Cytopathology and immune response in the hepatopancreas of decapod crustaceans. 138, 41-88.

[5] Kumar, V., Roy, S., Behera, B. K., et al. (2021). Acute hepatopancreatic necrosis disease (AHPND): Virulence, pathogenesis and mitigation strategies in shrimp aquaculture. 13(8), 524.

[6] Aldersey, J. E., Abernathy, J. W., Lange, M. D., et al. (2026). Single-cell transcriptomics of Pacific white shrimp (Litopenaeus vannamei) hepatopancreas reveal immune and metabolic responses to AHPND-causing Vibrio parahaemolyticus. Volume 17 - 2026.

[7] Soto-Rodriguez, S. A., Lozano-Olvera, R., Montfort, G. R.-C., et al. (2022). PirAB from Vibrio parahaemolyticus exhibits a binary toxin action with bactericidal activity and cytotoxicity in Penaeus vannamei.

[8] Mai, H. N., Caro, L. F. A., Cruz-Flores, R., et al. (2021). Differentially expressed genes in hepatopancreas of acute hepatopancreatic necrosis disease tolerant and susceptible shrimp (Penaeus vannamei). 12, 634152.

[9] Lange, M. D., Abernathy, J., Rawles, A. A., et al. (2023). Transcriptome analysis of Pacific white shrimp (Liptopenaeus vannamei) after exposure to recombinant Vibrio parahaemolyticus PirA and PirB proteins. 132, 108502.

[10] Syakur, M. A., Khotimah, B. K., Rochman, E. M. S., et al. (2018). Integration K-means clustering method and elbow method for identification of the best customer profile cluster. *IOP Conference Series: Materials Science and Engineering, 336*(1), 012017.

[11] Li, L., Wu, F., Liang, J., et al. (2021). Effects of cadmium (Cd) on reactive oxygen species (ROS) levels, hemocyte apoptosis, and expression of stress response genes in Litopenaeus vannamei. *Journal of Fishery Sciences of China, 28*(2), 167–175.

[12] Sutherland, D. E., Stillman, M. J. (2011). The "magic numbers" of metallothionein. *Metallomics, 3*(5), 444–463.

[13] Marek, D., & Marinescu, I. (2023). Role of metallothionein in regulating heavy metal detoxification and tissue-specific accumulation in the kidney, liver, and bone of small mammals. *International Journal of Sustainable Infrastructure for Cities and Societies, 8*(9), 36–44.

[14] Egan, B., Dowling, P., O'Connor, P. L., et al. (2011). 2-D DIGE analysis of the mitochondrial proteome from human skeletal muscle reveals time course-dependent remodelling in response to 14 consecutive days of endurance exercise training. *Proteomics, 11*(8), 1413–1428.

[15] Ramos-Castañeda, J., Park, Y.-N., Liu, M., et al. (2005). Deficiency of ATP2C1, a Golgi ion pump, induces secretory pathway defects in endoplasmic reticulum (ER)-associated degradation and sensitivity to ER stress. *Journal of Biological Chemistry, 280*(10), 9467–9473.

[16] Zhang, L., Sun, J., Ma, D., et al. (2010). Real-time quantitative analysis of three zinc finger protein gene families in salt-tolerant weedy rice. *Plant Physiology Communications, 46*(6), 529–536.

[17] Yuan, H., Jin, R., Wang, B., et al. (2025). Investigating the effect of microcystins on the transcriptomics of hepatopancreas in Penaeus vannamei using reverse-gavage. *Progresses in Fishery Sciences, 46*(3), 170–182.

[18] Wolosowicz, M., Prokopiuk, S., & Kaminski, T. W. (2024). The complex role of matrix metalloproteinase-2 (MMP-2) in health and disease. *International Journal of Molecular Sciences, 25*(24), 13691.

[19] Verde, E. M., Secco, V., Ghezzi, A., et al. (2025). Molecular mechanisms of protein aggregation in ALS-FTD: Focus on TDP-43 and cellular protective responses. *Cells, 14*(10), 680.

[20] Volik, P. I., Kopeina, G. S., Zhivotovsky, B., et al. (2025). p62-dependent caspase-2 activation governs TDP-43 clearance and neuronal fate in ALS. *Cell Death & Disease, 16*(1), 801.

[21] Guittard, E., Blais, C., Maria, A., et al. (2011). CYP18A1, a key enzyme of Drosophila steroid hormone inactivation, is essential for metamorphosis. *Developmental Biology, 349*(1), 35–45.

[22] Kang, P., Hang, Y., Chen, C., et al. (2023). Effects of replacing fishmeal with rapeseed meal and dietary condensed tannins on antioxidant capacity, immunity, and hepatic and intestinal health of largemouth bass (Micropterus salmoides). *Aquaculture Reports, 30*, 101548.