The Expression Dynamics of microRNAs in Mussel Under Cadmium Stress

A Study Case of miRNA Regulation of Stress Responses Based on MGI's DNBSEQ Sequencing Platform

Professor Peng Xiao's team at Shenzhen University has conducted extensive research for years on the complicated functions of microRNAs in aquatic species. Using the DNBSEQ sequencing platform from MGI, the team performed miRNA sequencing in different tissues (whole tissues, digestive glands, gills, and gonads) of mussels (*Mytilus galloprovincialis*) exposed to cadmium. They discovered significantly different expression of 107 known and 32 novel miRNAs.

The relevant findings were published in the journal "*Ecotoxicology and Environmental Safety*" in 2021, titled with "Stress responses in expressions of microRNAs in mussel *Mytilus galloprovincialis* exposed to cadmium"[1].

Recommended application: Marine Genomics

Recommended model: DNBSEQ-G400RS

- ### High sequencing accuracy

  The DNBSEQ sequencing platform employs advanced cPAS (combinatorial Probe Anchor Synthesis) and DNB technologies, which enhance signal intensity while effectively improving sequencing accuracy.

- ### A reliable platform for marine life research

  Based on the DNBSEQ sequencing platform, miRNA sequencing was performed in mussels exposed to cadmium. The findings reveal that 107 known and 32 novel miRNAs were significantly differentially expressed.

Background

Cadmium (Cd) pollution is widespread in water, sediments, and various marine species in the Bohai Sea[2]. Cd, as a typical heavy metal pollutant, can have various adverse effects on marine organisms like negative impacts on the immune system and the induction of oxidative stress[3]. Studies have shown that Cd can interfere with the energy metabolism of mussels (*Mytilus galloprovincialis*) and Manila clams (*Ruditapes philippinarum*). Given that Cd has emerged as one of the most severe metal pollutants in the Bohai Sea, it is essential to investigate its toxic effects on marine life.

miRNAs are pivotal in numerous marine species through various biological processes[4]. miRNAs, a class of non-coding RNA molecules with 18–24 nucleotides (nt), include a new class of small endogenous RNAs. They exert regulation over gene expression by guiding their target mRNAs toward degradation or translational repression[5]. In 2014, Chen et al. documented miRNA expression profiles in the scallop (Chlamys farreri) when exposed to acute viral necrosis virus infection. Additionally, miRNAs perform crucial functions in diverse tissues within the marine medaka (*Oryzias melastigma*), e.g. contributing to neural system development in the brain and regulating lipid metabolism in the liver[6]. Within Pacific oysters (Crassostrea gigas), cgi-miR-2d plays a role in regulating the cholinergic neuroendocrine system balance when facing *Vibrio splendidus* infection. These studies provide valuable insights into better understanding of the complex gene regulation mechanisms of marine species.

Marine organisms frequently encounter a range of environmental pollutants, prompting researchers to investigate how miRNAs respond to these environmental stresses. miRNAs may have pivotal roles in response of animals to stress by their gene expression fine-tuning capability. miRNAs have been reported to play a role in the response to environmental stress in corals, the nematode *Caenorhabditis elegans*, and numerous other species[1]. However, to date, there have been no studies on miRNAs in mussels under environmental stress[1].

Research Description

MicroRNAs (miRNAs) in aquatic organisms have complicated functions, yet their roles under environmental stress remain poorly understood in the scientific community. In this study, Professor Peng Xiao's team from Shenzhen University employed the DNBSEQ sequencing platform from MGI to conduct miRNA sequencing and subsequent bioinformatics analysis of different tissues (whole tissues, digestive glands, gills, and gonads) in mussels exposed to cadmium (Cd). The objective was to investigate the functions of differentially expressed miRNAs.

The results revealed that a total of 107 known mature miRNAs and 32 novel miRNAs in mussels were significantly ($p<0.05$) altered in response to Cd exposure. Functional analysis of the predicted target genes of differentially expressed miRNAs indicated that Cd induced immune toxicity and apoptosis in mussels, affecting their lipid metabolism and exogenous metabolism. These research findings provide valuable insights into further exploring the gene regulation mechanisms mediated by miRNAs in aquatic organisms under environmental stress.

Materials and Methods

Sample collection and RNA extraction

The research team collected 180 mussel samples, and randomly divided them into three groups, including a control group and two Cd-exposure groups with concentrations of 5 and 50 μg/L. After experimental treatment, they dissected whole tissues and three soft tissues (digestive gland, gills, and gonads) from each mussel and stored all samples at -80°C for heavy metal and miRNA analysis. Total RNA was extracted using TRIzol reagent (Invitrogen, Carlsbad, CA, USA). To determine the maximum number of miRNAs, total RNAs from all tissue samples were mixed in equal proportions.

MicroRNA Sequencing

The research team employed MGI's DNBSEQ sequencing platform for high-depth sequencing. They removed reads containing 5' adapter contaminants, reads without 3' adapters or insert tags, reads containing poly A, T, G, or C, low-quality reads, and reads shorter than 18 nt to obtain high-quality sequencing reads. Then, they map high-quality reads to a reference sequence without mismatch using Bowtie to analyze their expression and distribution thereon. To remove tags from protein-coding genes, repetitive sequences, rRNA, tRNA, snRNA, snoRNA, and sRNA tags were mapped to the RepeatMasker and Rfam databases.

The team compared high-quality reads to the miRNA precursor/mature miRNAs of all animals in the miRBase database to determine the sequence and quantity of miRNA families (non-species-specific) in the samples. New miRNAs were predicted based on the characteristics of miRNA precursor hairpin structures.

Sequencing Data Analysis

The miRNA data were mapped to the existing genome of mussels with close genetic relationships with mussels. Differentially expressed miRNAs in mussels between the control group and two doses of different Cd treatment groups

were determined by log_2-transforming the expression data and plotting them on a scatter plot. The process involved the following steps: (1) Normalizing the miRNA expression data from the three libraries to obtain the transcripts per million reads (TPM). The normalization formula was: Normalized expression = mapped reads/-total reads * 1×10^6. (2) Calculating the fold-change and *p-value* from the normalized expression, where the *p-value* was adjusted using the q-value. (3) Generating the log_2-ratio figure and scatter plot. In cases where the normalized expression of a miRNA between the three libraries was 0, its expression value was adjusted to 0.01 (since 0 cannot be plotted on a log plot). If the normalized expression of a specific miRNA in the three libraries was <1, further differential expression analysis did not need the miRNA. Consequently, the calculated gene expression values were suitable for directly comparing differences in gene expressions between various treatments.

The remaining reads were compared with the RepeatMasker, Rfam databases to remove potential mRNA, rRNA, tRNA, snRNA, snoRNA, and repetitive sequences. However, some sRNA tags might be mapped to multiple categories. To guarantee that each unique miRNA was only mapped to a single annotation, a priority rule was followed, which included: rRNA, etc. (GenBank > Rfam) > known miRNA > repeat > exon > intron. Then, the clean reads were all classified into categories such as exon_sense, exon_antisense, intron_sense, intron_antisense, miRNA, rRNA, repeat, scRNA, snRNA, snoRNA, srpRNA, tRNA, and unknown (sequences not mapped to any known reference database). Subsequent analyses included GO clustering and KEGG pathway analysis.

Sample collection	Library preparation and sequencing	Bioinformatics analysis	Result analysis
180 mussel samples, divided into control group, 5 μg/L and 50 μg/L cd-exposed groups. After experimental treatment, various tissues were collected from each mussel for further experiments.	 MGIEasy Small RNA Library Prep Kit  Genetic Sequencer DNBSEQ-G400	Bowtie miRBase RepeatMasker Rfam database Gene Ontology KEGG	miRNA sequencing analysis, different miRNA expression patterns analysis, GO enrichment and KEGG pathway.

Result

A summary of the detected miRNAs in miRNA sequencing data

A total of 12 miRNA libraries were constructed from mussel samples. The number of detected miRNAs in each mussel sample is shown in Table 1. Analysis of these 12 libraries confirmed the enrichment of miRNA sequences in these libraries. The length data of miRNAs showed that the size distribution of all libraries is similar, ranging between 17-32 nt, and they exhibit a bimodal distribution pattern. The most abundant class is the 22 nt miRNAs, followed by 29, 30, 21, and 28 nt.

Tissues	Treatments	Known miRNA count	Novel miRNA count
Whole tissue	Control	100	28
	5 µg/L Cd	103	29
	50 µg/L Cd	99	28
Digestive gland	Control	100	30
	5 µg/L Cd	96	29
	50 µg/L Cd	97	29
Gill	Control	99	30
	5 µg/L Cd	100	29
	50 µg/L Cd	104	29
Gonad	Control	96	24
	5 µg/L Cd	90	25

Table 1. A summary of detected miRNAs in mussel samples after 96 hours of exposure to cadmium

A summary of miRNA sequencing analysis mapping results

Table 2 provides a summary of the sequence information obtained from mussel soft tissues. After filtering out low-quality sequences, an average of 26,993,797 miRNAs was obtained from the 12 samples. Figure 1 illustrates the composition of RNA classes in each library. Within the clean reads, a substantial number of miRNAs were mapped to the mussel genome (Table 2).

Analysis of these 12 libraries confirmed the enrichment of miRNA sequences in these libraries. The length data of miRNAs showed that the size distribution of all libraries is similar, ranging between 17-32 nt, and they exhibit a bimodal distribution pattern. The most abundant class is the 22 nt miRNAs, followed by 29, 30, 21, and 28 nt.

Tissues	Treatments	Raw tag count	Total tag	Mapped tag	Percentage of Mapped tag (%)
Whole tissue	Control	29,405,689	28,050,597	4,057,381	14.46
	5 µg/L Cd	29,243,768	27,514,012	6,255,413	22.74
	50 µg/L Cd	26,857,869	25,324,868	3,438,338	13.58
Digestive gland	Control	28,068,451	26,789,951	11,677,717	43.59
	5 µg/L Cd	28,324,610	25,915,286	11,673,423	45.04
	50 µg/L Cd	29,936,911	27,854,771	11,799,772	42.36
Gill	Control	28,182,414	26,945,210	15,124,380	56.13
	5 µg/L Cd	27,975,463	25,967,304	21,144,264	81.43
	50 µg/L Cd	29,743,654	27,981,423	24,048,131	85.94
Gonad	Control	28,933,285	27,820,171	1,789,490	6.43
	5 µg/L Cd	29,385,455	27,584,619	217,1279	7.87

Table 2. A summary of sequencing data from mussel samples after 96 hours of exposure to cadmium, along with the mapping statistics of tags to the reference genome of cadmium-exposed mussels.

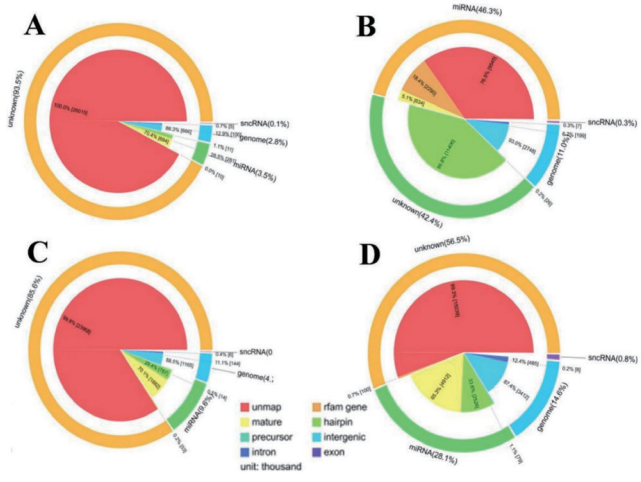

Figure 1. Proportions of different types of sRNAs in various tissues of mussels. To ensure that each sRNA corresponds to only one annotation, the following priority rule was followed: miRNA > piRNA > snoRNA > Rfam > other sRNA (A: Gonad, B: Gill, C: Whole Soft Tissue, D: Digestive Gland).

Different miRNA expression patterns

A total of 107 known miRNAs and 32 novel miRNAs were found in the mussel tissues (Table 1). Among them, 26 non-repetitive miRNAs were identified as differentially expressed ($p < 0.05$) under cadmium exposure. In general, 66 known miRNAs and 19 new miRNAs exhibited significant changes (Figure 2), with the majority of differentially expressed miRNAs upregulated following cadmium exposure.

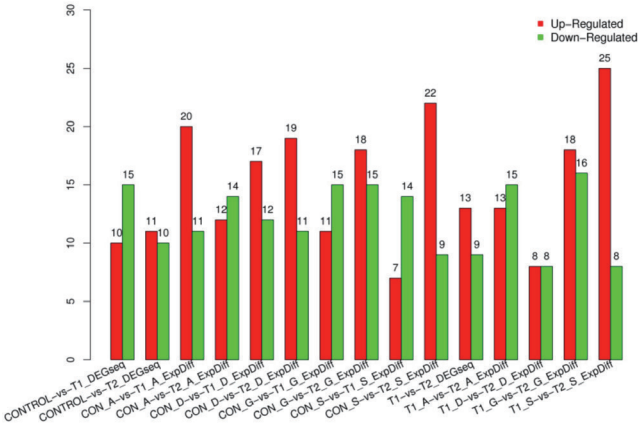

Figure 2. Differentially expressed miRNAs in mussels under cadmium exposure. The X-axis represents differentially expressed sRNAs, and the Y-axis represents the number of selected DESs. The green color indicates downregulation, while the red color represents upregulation. (CON: Control group, T1: 5μg/L cadmium-treated group, T2: 50μg/L cadmium-treated group; A stands for the whole soft tissue of blue mussels, G for the gill, D for the digestive gland, S for the gonad).

GO enrichment and KEGG pathway analyses

To further understand the biological functions of miRNAs, the research team predicted the target genes of differentially expressed miRNAs. Subsequently, enriched functional groups ($p<0.05$) were identified through GO analysis. These target genes were primarily involved in biological processes, cellular components, and molecular function categories. Moreover, enriched metabolic pathways and signal transduction pathways were also identified (Figure 3). 20 significantly enriched pathways for target genes ($p < 0.05$), primarily involving: spliceosome, regulation of actin cytoskeleton, acute myeloid leukemia, VEGF signaling pathway, lysosome, GnRH signaling pathway, endocytosis, and tuberculosis (Figure 3). In addition, according to KEGG analysis, proteoglycans in cancer (ko05205) and long-term depression (ko04730) were the most enriched pathways among the target genes associated with all the identified miRNAs under both 5 and 50 μg/L Cd exposure.

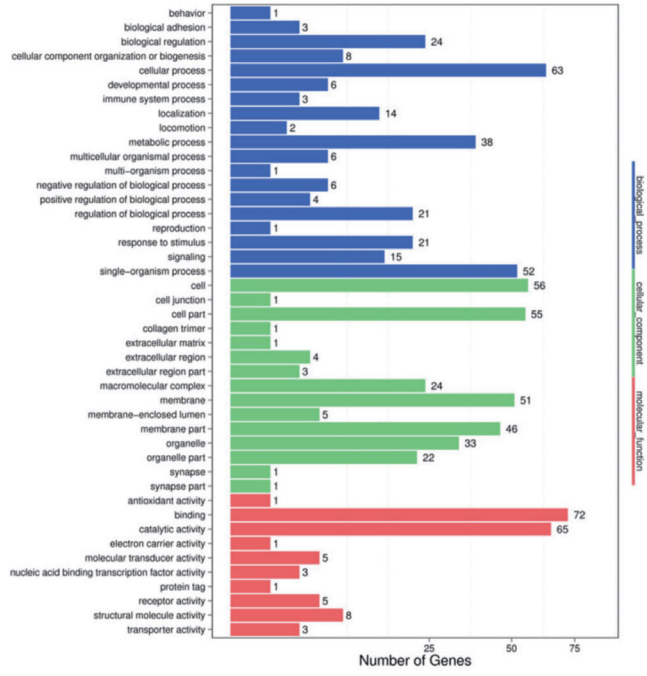

Figure 3. Classification of GO functions in mussels under cadmium exposure. The X-axis represents the number of differentially expressed genes (DEGs), and the Y-axis represents GO functions. All GO functions are divided into three categories: blue represents biological process, brown represents cellular components, and red represents molecular function.

Summary

The team led by Professor Peng Xiao at Shenzhen University employed high-depth small RNA sequencing technology and bioinformatics analysis based on the DNBSEQ sequencing platform to identify miRNAs in different tissues of blue mussels under cadmium stress. The results revealed a total of 107 known mature miRNAs and 32 new miRNAs. Among these, 26 non-repetitive miRNAs were identified as differentially expressed ($p < 0.05$) under cadmium exposure. Functional analysis of the predicted target genes of differentially expressed miRNAs indicated that Cd induced immune toxicity and apoptosis in mussels, affecting their lipid metabolism and exogenous metabolism. These research findings provide valuable insights into further exploring the gene regulation mechanisms mediated by miRNAs in aquatic organisms under environmental stress.

DNBSEQ-G400 Genetic Sequencer

References

1. Yu D, Peng Z, Wu H, Zhang X, Ji C, Peng X. Stress responses in expressions of microRNAs in mussel Mytilus galloprovincialis exposed to cadmium. *Ecotoxicol Environ Saf*. 2021 Apr 1; 212: 111927.
2. Gao, X., Zhou, F., Chen, C., 2014. Pollution status of the Bohai Sea: an overview of the environmental quality assessment related trace metals. *Environ. Int*. 62, 12-30.
3. Company, R., Serafim, A., Bebianno, M., Cosson, R., Shillito, B., Fiala-Medioni, A., 2004. Effect of cadmi-um, copper and mercury on antioxidant enzyme activities and lipid peroxidation on the gills of the hydrothermal vent mussel Bathymodiolus axoricus. Mar. *Environ*. Res. 58, 377-381.
4. Burgos-Aceves, M., Cohen, A., Paolella, G., Lepretti, M., Smith, Y., Faggio, C., Lionetti, L., 2018a. Modula-tion of mitochondrial functions by xenobiotic-in-duced microRNA: from environmental sentinel organisms to mammals. *Sci. Total Environ*. 645, 79-88.
5. Burgos-Aceves, M., Cohen, A., Smith, Y., Faggio, C., 2018b. A potential microRNA regulation of immune-related genes in invertebrate haemocytes. *Sci.Total Environ*. 621, 302-307.
6. Chen, G., Zhang, C., Jiang, F., Wang, Y., Xu, Z., Wang, C., 2014. Bioinformatics analysis of hemocyte miRNAs of scallop Chlamys farreri against acute viral necrobiotic virus(AVNV). *Fish. Shellfish Immu-nol*.37, 75-86.

Recommended Ordering Information

Category	Product	Cat. NO.
Instruments	Genetic Sequencer DNBSEQ-G400RS	900-000170-00
	MGISP-100RS Automated Sample Preparation System	900-000206-00
	MGISP-960RS Automated Sample Preparation System	900-000146-00
Software	MegaBOLT Bioinformatics analysis accelerator	900-000555-00
Library Prep	MGIEasy Small RNA Library Prep Kit（24 RXN)	940-000196-00
Sequencing Reagents	DNBSEQ-G400RS High-throughput Sequencing Set (FCL SE50)	1000016941

MGI Tech Co.,Ltd

Building 11, Beishan Industrial Zone, Yantian District, Shenzhen, CHINA, 518083

Version: November 2023

+86-4000-688-114
en.mgi-tech.com
MGI-service@mgi-tech.com

Authors: Chen Yiming, Zhou Hao
Editor-in-Charge: Wang Qiwei
Reviewer: Jiang Yao

1. For StandardMPS and CoolMPS: Unless otherwise informed, StandardMPS and CoolMPS sequencing reagents, and sequencers for use with such reagents are not available in Germany, Spain, UK, Sweden, Belgium, Italy, Finland, Czech Republic, Switzerland, Portugal, Austria and Romania. Unless otherwise informed, StandardMPS sequencing reagents, and sequencers for use with such reagents are not available in Hong Kong. No purchase orders for StandardMPS products will be accepted in the USA until after January 1, 2023.
2. For HotMPS sequencers: This sequencer is only available in selected countries, and its software has been specially configured to be used in conjunction with MGI's HotMPS sequencing reagents exclusively.
3. For HotMPS reagents: This sequencing reagent is only available in selected countries.