Full length article

The transcriptomic response to copper exposure by the gill tissue of Japanese scallops (*Mizuhopecten yessoensis*) using deep-sequencing technology

Xiaolin Meng\(^a\), Xue Tian\(^a\), Mei Liu\(^b\),*, Guoxing Nie\(^a\), Keyong Jiang\(^b\), Baojie Wang\(^b\), Lei Wang\(^b\),*

\(^a\)College of Fisheries, Henan Normal University, Xinxiang 453007, PR China
\(^b\)R&D Center of Marine, Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, PR China

**ABSTRACT**

The bivalve *Mizuhopecten yessoensis* has been greatly impacted by marine pollutants in northern China. To elucidate the toxicological mechanism of copper exposure on the immune system, we investigated differentially expressed genes (DEGs) and transcript abundance in *M. yessoensis* gill tissue using the deep-sequencing platform Illumina HiSeq™ 2000. In total, 1312 and 2237 genes were identified as significantly up- or down-regulated, respectively. In addition, significant enrichment analysis identified 9 GO terms and 38 pathways involved in the response to copper exposure. The analysis of immune-related transcripts revealed a complex repertoire of innate recognition receptors, including toll-like receptors, NOD-like receptors and RIG-like receptors. Downstream pathway effectors, such as apoptotic, lysosomal and C-type lectin transcripts, were also analyzed. These results will provide a resource for subsequent gene expression studies regarding heavy metal exposure and the identification of copper-sensitive biomarkers to monitor the aquaculture of *M. yessoensis*.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Many chemical contaminants, including organochlorine compounds, herbicides, domestic and municipal wastes, petroleum products and heavy metals, have recently been shown to induce serious adverse effects on ocean environments, even at low levels [1,2]. Heavy metals, which readily accumulate in ocean sediment, can reach significant accumulation levels in the tissues of aquatic organisms. These heavy metals are often concentrated and amplified in organisms at higher levels of the food chain, particularly in benthic animals. Copper (Cu), which is one of the most potentially harmful heavy metal pollutants, enters the environment in elevated concentrations from the use of copper sulfate (CuSO\(_4\)) in agricultural activity because it is widely used to control algal blooms and aquatic macrophyte infestations for aquaculture [3]. Although Cu is an essential element for most living organisms, it can be a toxic pollutant in high concentrations when present as the free metal ion.

The Japanese scallop (*Mizuhopecten yessoensis*) was introduced into China in 1982. However, the aquaculture of *M. yessoensis* has been devastated by large-scale death, which has resulted in serious economic losses. Environmental stress, such as excess Cu pollution in coastal waters, may be one of the main factors contributing to the high mortality rates of *M. yessoensis*. Previous studies regarding bivalves have focused on the antioxidant systems [4,5], lipid damage [6] and the mechanisms of Cu accumulation [7]; however, the molecular mechanism of Cu toxicity is still unclear.

One strategy to reduce the high mortality rates of *M. yessoensis* is to identify disease or stress-resistance genes in scallops and use them to genetically alter cultured stocks [8]. Physiological genomics, the global analysis of transcriptome responses to different conditions, offers an opportunity to achieve such a goal [9]. For instance, a genomic analysis using a microarray platform was recently used to investigate the responses of *Ruditapes philippinarum* in a highly polluted area [9] and the summer mortality rates of the Pacific oyster [10]. Nevertheless, this method has several limitations, such as reliance upon existing genome sequencing knowledge, high background levels and a limited dynamic range of detection [11]. For *M. yessoensis*, an overall understanding of the mechanism of action of Cu is extremely difficult due...
to a lack of genomic sequencing resources. With the advent of next generation sequencing technologies (NGS), a more comprehensive and accurate transcriptome analysis has become feasible and affordable. RNA-Seq (Quantification), a version of NGS, is used for transcriptome quantification and the gene expression analysis of biological specimens under specific conditions. In contrast to microarray technology, high-throughput RNA sequencing can provide a comprehensive assessment of RNA expression profiles with the advantages of high-throughput sequencing data acquisition, low background, high sensitivity and reproducibility [12].

Here, we used the RNA-seq (Quantification) approach to perform a deep transcriptome analysis of adult M. yessoensis exposed to Cu. The gill was chosen for these studies because of the key role it plays in filtration of suspended matter, its role as a defense barrier, and its high expression of putative immune-related contigs [13]. The present study aims to identify genes expressed in the gill that are sensitive to Cu exposure and to elucidate genes and pathways that can be used to examine the molecular and genetic basis of immunity and stress-resistance.

2. Materials and methods

2.1. Scallop acquisition

Adult Japanese scallops (M. yessoensis) of both sexes were obtained in 2012 from Dalian Zhangzidao (Dalian, Liaoning Province, China). Only healthy scallops with a homogeneous size (9.150 ± 0.162 cm (mean ± SD, n = 30), shell height) were tested. To minimize the effects of extraneous environmental factors other than Cu exposure, the scallops were cultured in running aerated sea water (salinity 30‰) at 16 ± 1 °C. Half of the seawater was changed daily, and the scallops were fed with the microalgae spirulina.

2.2. Copper exposure and RNA preparation

The scallops in the Cu-treated groups were exposed to CuCl2·2H2O (Sigma, USA), and the final Cu²⁺ concentration in the seawater was determined to be 0.1 mg/L, which is ten times the standard limit according to 'Water quality standard for fisheries of China' (Cu²⁺ ≤ 0.01 mg/L). All of the scallops were kept under the experimental conditions described above, and fresh seawater was replenished with a corresponding concentration of Cu²⁺. After 72 h of exposure, two gill samples from the control and treated groups were collected and used to construct libraries (control vs. treated). Six independent biological replicates for each sample were harvested and pooled to construct the libraries. Total RNA was extracted using TRIzol (Invitrogen, USA) according to the manufacturer's protocol.

2.3. Library preparation and Illumina sequencing

Using the reagents of the mRNA sequencing protocol provided by Illumina, poly(A)⁺ mRNA was obtained from total RNA samples using oligo(dt) magnetic beads. An RNA fragment kit (Ambion, Austin, TX) was used to cleave the mRNA into short fragments of approximately 200 bp. First strand cDNA was synthesized by random hexamer-priming, using the mRNA fragments as templates. Buffer, dNTPs, RNase H and DNA polymerase were added to synthesize the second strand. An 'end repair' reaction was then performed with Klenow polymerase, T4 DNA polymerase and dNTPs to blunt the ends of all mRNA fragments. The fragments were then phosphorylated by T4 polynucleotide kinase, followed by the addition of a single 'A' base to the 3'-end of the blunt-ended phosphorylated fragments. Finally, sequencing adaptors were ligated to the fragments. The required fragments were purified by agarose gel electrophoresis and enriched by PCR amplification, and the concentration was measured with a NanoDrop spectrophotometer. The library products were then sequenced and analyzed with an Illumina HiSeq™ 2000 instrument.

2.4. Data processing, read mapping and analysis

Raw sequencing reads averaging 49 bp in size were filtered through the Illumina pipeline. Image data acquired from the sequencing run were mirrored to an off-instrument computer using the Illumina platform, and all dirty raw reads (reads with adaptors, reads with unknown bases more than 10%, or low quality reads) were removed to generate a set of clean reads. The resultant reads were aligned to the M. yessoensis reference transcriptome, which was obtained via the Illumina platform (accession number: SRR653778) and the 454 sequencing platform (accession number: SRA027310) by SOAP2 [14]. Only 2 base mismatches were allowed. These high quality mapped reads were further subdivided into uniquely mapped reads and repetitive reads based on whether the best mapping for the read matched one location or more than 2 locations, respectively. Reads that mapped uniquely were used for digital gene expression counts and were normalized to reads per kilobase of transcript sequence per million mapped reads (RPKM) [15].

For differentially expressed gene analysis, a stringent algorithm p-value was used based on Audic’s method. The FDR (False Discovery Rate) was used to determine the threshold of the p-value in multiple tests [16]. We used a FDR ≤ 0.001 as the threshold to determine the significance of the differential gene expression data. More stringent criteria with smaller FDRs and larger fold-change values can be used to identify differentially expressed genes (DEGs).

2.5. Quantitative real-time PCR (qRT-PCR) verification

Seven genes were selected for confirmation of RNA-Seq (Quantification) data by qRT-PCR using a SYBR Premix Ex Taq kit (Takara, Dalian, China) according to the manufacturer’s instructions. The same RNA samples were used for both Illumina library synthesis and the qRT-PCR verification assay. The first strand cDNA was obtained from 2 μg of total RNA using a PrimeScript 1st strand cDNA synthesis kit (Takara, Japan). Melting curve analyses were performed following amplifications. The specific primers used for qRT-PCR are listed in Additional File 1, and β-actin was used as an endogenous control. The relative gene expression was analyzed using the comparative threshold (CT) cycle method established by Livak et al. [17].

2.6. Enrichment analysis of differentially expressed genes

Gene ontology (GO) enrichment analysis was performed by mapping each differentially expressed gene into the records of the GO database (http://www.geneontology.org/) [18]. The number of genes for each term was calculated, and GO terms significantly enriched for DEGs were identified by a hypergeometric test using the whole reference transcriptome as a measure of the background. The calculated p-value was subjected to Bonferroni Correction, and the corrected p-value < 0.05 was used as a threshold.

Pathway enrichment analysis was based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database [19]. Significantly enriched metabolic pathways or signal transduction pathways in DEGs were identified with respect to the whole transcriptome background, using the same formula as the GO enrichment analysis. FDR [16] was used to correct the p-value. Pathways with a FDR-value ≤ 0.05 were identified as significantly enriched.
3. Results and discussion

3.1. Sequencing and mapping

To characterize the response of *M. yessoensis* to Cu exposure, we performed a high throughput analysis based on the RNA-seq (Quantification) data. We sequenced two libraries: the control gill group (CG) and the treated gill group (TG). After removing the dirty reads, the number of clean reads ranged from 4.73 to 4.99 million. The data were deposited in the SRA database (accession number: SRR736121). Reads mapped to a unique sequence are the most critical, as they can explicitly identify a transcript. After mapping these clean reads to the reference database, 1.73 to 1.83 million sequences were identified as unique matches (Table 1). Next, the level of gene expression was analyzed by calculating the number of unambiguous reads for each gene and normalizing it to the number of RPKM [15]. As shown in Fig. 1(Additional File 2), the results indicated that the majority of uniquely expressed genes were represented in fewer than 10 copies and that a small portion of the genes were highly expressed. In additional File 3, we have listed the top 10 most abundantly expressed genes and annotated the differences between the two libraries. Both libraries contained alpha tubulin and ferritin CFC.

3.2. Validation of the RNA-seq data by qPCR

To further evaluate our library, the expression levels of 7 genes involved in stress response and immune system function were analyzed by qRT-PCR. The data exhibited expression tendencies similar to the DEG data (average spearman correlation = 0.964) (Fig. 2), demonstrating the reliability of the RNA-Seq (Quantification) results.

3.3. Changes in gene expression in gill after Cu exposure

A total of 3549 significantly expressed gene entities were detected between the TG and the CG, with 1312 up-regulated genes and 2237 down-regulated genes (Additional File 4), similar to data previously obtained after Cd exposure [20]. We also found that most of the highly differentially expressed genes in the gill were “orphan” sequences, which means that no homologs were found in the Blast Nr database until now. This is likely due to the absence of clear annotation information on mollusks, and the function of these unknown and predicted genes remains to be clarified in future studies.

3.4. Identification of GO enrichment analysis

Gene Ontology (GO) based on biological process (BP) enrichment analysis for sets of differentially expressed genes with significant cluster profiles was performed (Additional File 5). Only significant GO categories that had a p-value of <0.05 were chosen for analysis.

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
</tr>
<tr>
<td>Total clean reads</td>
</tr>
<tr>
<td>Total Basepairs (bp)</td>
</tr>
<tr>
<td>Total Mapped reads</td>
</tr>
<tr>
<td>Total Mapped reads (%)</td>
</tr>
<tr>
<td>Perfect match</td>
</tr>
<tr>
<td>Unique match reads</td>
</tr>
<tr>
<td>Unique match reads (%)</td>
</tr>
<tr>
<td>Total unmapped reads (%)</td>
</tr>
</tbody>
</table>

The predominant GO terms represented in a significant cluster profile in the gill after Cu exposure (TG vs. CG) were involved in the organic acid metabolic process, the carboxylic acid metabolic process, the cellular ketone metabolic process, the response to chemical stimulus, nitrogen compound transport, the ER-nucleus signaling pathway, amine transport, amino acid activation and tRNA aminoacylation. These results are the first report of the involvement of these pathways in response to Cu exposure, and their associated genes may play important roles in *M. yessoensis* following Cu exposure. The GO term ‘response to chemical stimulus’ (Go:0042221) is an important biological process after marine pollutant exposure. As shown in Fig. 3, dual oxidase (DUOX), asparagine synthase (AS), arylesterase, myeloid differentiation primary response gene88 (MyD88), tumor necrosis factor receptor superfamily (TNFRS) and aldehyde dehydrogenase (NAD+) were significantly induced in the Cu-treated gill. Conversely, caspase8, glutamate decarboxylase...
(GAD), indoleamine 2,3-dioxygenase (IDO), cholinergic receptor, alkaline phosphatase (ALP), proliferating cell nuclear antigen (PCNA) and insulin-like growth factor binding protein (IGFBP) were significantly suppressed upon Cu exposure. DUOX is a member of the NADPH oxidase family and was recently identified as a novel homolog of phagocytic oxidase, which could regulate reactive oxygen species (ROS) production by phagocytosis during the host defense response. Experiments in zebrafish [21] challenged with
bacteria and shrimp [22] with white spot syndrome virus (WSSV) support our conclusion. TNFRS mediates a wide spectrum of physiological and pathological events, such as cell activation, cell proliferation, inflammation, and cell death. However, there is only one EST of a TNFRS previously identified in a bivalve (hepatopancreas, member 25) in the NCBI database (1/26/2013). To our knowledge, this is the first example of TNFRS (member 11B) detection in *M. yessoensis*. Caspase8 is an initiator caspase and is an essential component of the extrinsic cell death pathway initiated by TNF family members. In response to the activation of the TNF family death receptors, caspase8 is recruited to the death-inducing signaling complex and results in caspase8 activation and cell death [23]. Genetic evidence from knockout mice has indicated that caspase8 is required for all known death receptor-mediated apoptotic pathways [24]. Furthermore, caspase8 was up-regulated in shrimp after infection with WSSV [25] and in rainbow trout hepatocytes exposed to cadmium (Cd) within 24 h [26]. Interestingly, the results we observed in Cu-treated scallops conflicted with those of previous studies. We note that when rainbow trout were exposed to Cd, 2,2,6,6-tetramethylpiperidinyl-1-oxyl (TEMPO) or N-acetylcysteine (NAC), caspase8 expression was significantly suppressed after 24h [26]. Therefore, time of exposure may be an important factor affecting caspase8 expression.

3.5. Identification of signaling pathway analysis and immune-related genes

To characterize the functional consequences of the gene expression changes associated with Cu exposure, we performed pathway analysis based on the KEGG database. Only significant pathway categories that had a FDR-value < 0.05 were chosen for the analysis. As shown in Fig. 4, the significantly enriched signaling pathways in the gill included the following: RIG-like receptor signaling; insulin signaling; lysosome and Toll-like receptor signaling; and the signaling involved in vascular smooth muscle contraction, measles, legionellosis, tuberculosis antigen processing and presentation, and apoptosis.

Abnormal vascular smooth muscle cell (VSMC) relaxation or contraction is known to induce hypertension in both humans and animal models. Heavy metals are one of the factors which contribute to the development of hypertension. For example, Cd causes the release of a variety of proinflammatory mediators, such as tumor necrosis factor alpha, from endothelial cells. The release of these factors stimulates the subsequent release of antithrombolytic agents, such as plasminogen activator inhibitor-1, and facilitates the adhesion of leukocytes and platelets to the endothelium. In addition, elevated intracellular smooth muscle Ca could lead to increased arterial tone, which would result in hypertension [27]. Our results show a significant enrichment of genes in this pathway in gill after exposure to Cu (Additional File 6). These genes included calmodulin (CaM), ROCK, and cytochrome P450 (CYP450), suggesting that the vascular smooth muscle of the gill was greatly affected by Cu exposure. However, the regulation mechanism of vascular smooth muscle in gill of bivalves upon exposure to marine pollutants has not been fully elucidated and further research is needed to clarify these mechanisms.

Previous studies in bivalves have focused on the immune system to understand specific acclimation processes resulting from changing environmental factors [13,28]. To identify the related genes involved in the immune defense of gill following Cu exposure, we analyzed immune related pathways.

3.5.1. Pattern recognition receptors (PRRs)

3.5.1.1. Toll-like receptors (TLRs). TLRs are an ancient family of pattern recognition receptors that play key roles in detecting non-self antigens and immune system activation [28]. TLRs are triggered

![Significant enrichment pathways of different expressed genes in the gill following Cu exposure](image)

*Fig. 4. Significant enrichment pathways of different expressed genes in the gill following Cu exposure. A P-value of < 0.05 and an FDR of < 0.05 were selected as the significant criteria. The vertical axis is the pathway category and the horizontal axis is the log 10 (p value) of these significant pathways.*
by ligand binding, which leads to the activation of a range of transcription factors, such as NF-κB and IFN regulatory factors (IRFs). This process eventually leads to the downstream activation of proinflammatory cytokines and receptors. Our results revealed that TLR-1,-2,-3,-4,-7/8 and their signaling intermediates (Rac1, MyD88, CASP8, IRAK4, TRAF3, TBK1 and RIP1) were dramatically altered after Cu exposure. These data provide evidence that TLRs (TLR1, TLR3) participate in M. yessoensis host defense after Cu exposure; however, novel TLRs (TLR13, TLR18, TLR21) were not detected in this pathway analysis.

3.5.1.2. NOD-like receptors (NLRs) and RIG-like receptors (RLRs). NLRs play essential roles in innate immunity by detecting intracellular pathogen-associated molecular patterns (PAMPs) and danger-associated molecular patterns (DAMPs). These molecules reveal the presence of pathogenic infection, abiotic stress, environmental insults, cellular damage, and cell death [29]. Among them, NOD1 and NOD2 are the best studied, both of which were down-regulated in the gill following Cu exposure. These data suggest that some changes occurred in the gill epithelium, such as necrosis and death. The RNA helicases retinoic acid inducible gene-I (RIG-I) and melanoma differentiation-associated gene 5 (MDA5) constitute a further PRR family called RIG-I-like receptors (RLRs). Similar to NOD-like receptors, RIG-I and MDA5 were expressed mainly in endothelial cells and were suppressed after Cu exposure. IPS-1 (interferon-β promoter stimulator 1), a central component in the RLR pathway, was also down-regulated, which may inhibit the expression of type I interferon. A similar pathway may be employed by NOD1/2. However, the detailed mechanism of PPR regulation in bivalves following heavy metal exposure is still unclear. Furthermore, exposure time and dose may have significant effects on the regulation of the innate immune system.

3.5.2. Apoptosis

Cell death is a fundamental response to immune stress, and three different forms of cellular destruction are known: apoptosis, autophagy and necrosis. Although the characteristics of apoptosis at the cellular level have been described in a variety of lower invertebrates, sequence information regarding the genes involved in these processes is still scarce. In the apoptosis pathway, a family of proteases known as caspases cleave target proteins at specific sites and are predominantly regulated by the extrinsic approach, which utilizes caspases 3, 6, 7, 8 and 10. Notably, the intrinsic pathways requiring mitochondrial permeability or caspase 3 activation were absent in gill cell apoptosis following Cu stress. Similar observations were found in molluscan hemocytes following Cd-induced apoptosis [30]. These results suggest that the caspase-independent apoptosis pathway in bivalves is still unclear.

3.5.3. Lysosome and C-type lectin

In addition to external barriers, such as shells, mucus and epithelia, internal humoral components play a key role in molluscan immune defense, including the involvement of lysosomal enzymes and lectins. In lysosomes, metals form a complex with polyvalent anions, such as sulfate or phosphate, generating an impermeable precipitate that sequesters the metal, leading to effective detoxification [31]. In this study, we detected abundant lysosomal enzymes such as cathepsins, lysosomal beta-galactosidase (GLB), alpha-N-acetylgalactosaminidase (NAGA), lysosomal alpha-mannosidase precursor (LAMAN), arylsulfatase (ARS) and lysosomal membrane proteins, such as tetraspavin (LIMP). However, most of these transcripts were down-regulated, which suggests that damage to the gill epithelial cell lysosome was occurring, leading to a lower detoxification efficiency in the Cu-exposed gill.

C-type lectin is a family of Ca2+ dependent carbohydrate-recognition proteins that play a crucial role in the innate immunity of invertebrates [32]. Although several C-type lectin genes have been identified and cloned in the scallop, to our knowledge, related information about M. yessoensis is scarce. Following Cu exposure, a total of 21 lectin family transcripts were found to be differentially regulated in the gill including C-type 1, 2, 4, 5, 7, 10 and B. Furthermore, 16 of those transcripts were significantly down-regulated. A clear, time-dependent pattern of CFlEc-1 expression has been observed in Zhikong scallop hemocytes after bacterial challenge and maximum expression was achieved after 2–8 h [32]. Together, these data provide evidence that C-type lectin expression exhibits a time-dependent relationship with Cu exposure.

Our data in this study will provide a rich source for gene expression studies following heavy metal pollution as well as for the identification of biomarkers for various applications such as the selection of sensitive genes to monitor the aquaculture of M. yessoensis.

Acknowledgments

Financial support for this study was provided by the Main Direction Program of Knowledge Innovation of Chinese Academy of Sciences (KZCX2-EW-Q215).

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jsi.2014.03.009.

References


