Latin American Journal of Aquatic Research, 50(2): 212-226, 2022
DOI: 10.3856/vol50-issue2-fulltext-2705

Research Article

De Novo assembly and annotation of the Pacific calico scallop (*Argopecten ventricosus*) transcriptome for immune-related gene discovery

Ruth Escamilla-Montes[1], Genaro Diarte-Plata[1], Gabriela B. Mendoza-Maldonado[1] Aarón Barraza[2], Carlos Angulo[3], Seyed Hossein Hoseinifar[4] Jesús A. Fierro-Coronado[1] & Antonio Luna-González[1]

[1]Instituto Politécnico Nacional, Centro Interdiciplinario de Investigación para el Desarrollo Integral Regional Guasave, Sinaloa, México

[2]CONACYT-Centro de Investigaciones Biológicas del Noroeste, Instituto Politécnico Nacional La Paz, Baja California Sur, México

[3]Immunology & Vaccinology Group, Centro de Investigaciones Biológicas del Noroeste La Paz, Baja California Sur, México

[4]Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran

Corresponding author: Antonio Luna-González (aluna@ipn.mx)

ABSTRACT. Invertebrates' immune defense mechanisms play a critical role in pathogen recognition and elimination. *De novo* assembly and annotation of the *Argopecten ventricosus* transcriptome were performed for the immune-related gene identification. Scallops (height: 4.4 cm) were challenged with inactivated *Vibrio parahaemolyticus* IPNGS16. The RNA from different tissues was pooled for a single cDNA library construction sequenced by NextSeq 500 platform 2×75 paired-end chemistry. Before *de novo* assembling with Trinity, reads were analyzed with FastQC, Trimmomatic, and Prinseq. Assembled sequences were analyzed by CD-HIT-EST and TransDecoder. The corresponding annotation was performed against NCBI-nr, RefSeq protein, and KAAS (KEGG) databases. The Trinity assembly yielded 107,516 contigs. TransDecoder yielded 25,285 sequences as CDSs of which, 16,123 were annotated against the NCBI-nr protein, most of them scored with *Crassostrea gigas* data. Gene ontology mapped sequences (15,262) were classified in molecular functions (~13,000), cellular components (~11,000), and biological processes (~13,000). The KAAS analysis showed biological categories for metabolism (13%), cellular processes (12%), genetic information processing (10%), organismal systems (19%), environmental information (13%), and human diseases (33 %). Within the organismal systems, 467 immune-related genes (KO) were identified. Sixty-four immune-related genes were annotated/blasted against the NCBI-nr and RefSeq protein databases. An RT-qPCR was performed to analyze the expression level of immune-related genes obtained in the transcriptome analysis in scallops (height 4.5 cm) treated with probiotic bacilli added to culture water. Bacilli significantly increased the expression of the HSP70 and PGRP genes. The gene transcripts analysis of *A. ventricosus* will better understand its immune response against pathogens in culture systems.

Keywords: *Argopecten ventricosus*; calico scallop; transcriptome; bivalve; immune genes

INTRODUCTION

Marine mollusks are essential invertebrates; many of them are used as food, craft making, and dye source (Flores-Garza et al. 2017). Among the bivalve mollusks, the Pacific calico scallop, *Argopecten ventricosus*, is distributed from Santa Barbara, California, USA, through the Gulf of California, Mexico to Paita in Peru (Coan & Valentich 2012). This scallop is a species of great economic value in Baja California Sur (Mexico) because they form vast banks in shallow waters, such as lagoons, bays, and estuaries (Maeda-Martínez et al. 1993, Ruiz-Verdugo et al. 2016). Scallop overfishing has led to studies focused on its cultivation, mainly for repopulation purposes in fishery areas (Medina et al. 2007, Ruiz-Verdugo et al. 2016). Spat production in

Corresponding editor: José Gallardo

hatcheries, nursery period, and grow-out until reaching commercial harvest size have been optimized (Mazón-Suástegui et al. 2010, Ruiz-Verdugo et al. 2016). However, as in other bivalves, the production has been negatively affected by infectious diseases, which have motivated research on bivalve immunology and biotechnological developments in hatcheries and nursery systems to strengthen the non-specific immune response against pathogens (Matozzo et al. 2016, Batista et al. 2019).

The scallop *A. ventricosus* is susceptible to pathogenic vibrios such as *Vibrio parahemolyticus* IPNGS16 (Mendoza-Maldonado et al. 2018). Understanding immune defense mechanisms in invertebrates are essential because they play a critical role in pathogen recognition and elimination. Immune defense mechanisms include cellular effector responses in which hemocytes are the central players. Hemocytes carry out their activity mainly through cytotoxic reactions, infiltration, aggregation, encapsulation, and phagocytosis of foreign particles; moreover, they are responsible for many molecular effectors, such as reactive oxygen intermediates, prophenoloxidase system, lectins, lysosomal hydrolytic enzymes, and antimicrobial peptides (Carballal et al. 1997, Hine 1999, Mitta et al. 2000, Donaghy et al. 2012). Immune-related gene expression can be modulated as it was found in *Penaeus monodon* fed dietary probiotic *Bacillus* sp. JL47 and challenged with pathogenic *Vibrio campbellii* LMG 21363 (Laranja et al. 2017).

Next-generation sequencing (NGS) was selected to compare studies in the research field of bivalve mollusk immunology, which has provided insights into the transcriptional activity of their whole genomes (Andrews 2010, Moreira et al. 2015, Batista et al. 2019). Therefore, in this work, the Pacific calico scallop (*A. ventricosus*), treated with inactivated *Vibrio* and live bacilli, was studied to perform a *de novo* assembly and annotation of its transcriptome to determine the transcriptional activity of immune-related genes.

MATERIALS AND METHODS

Experimental organisms and sample preparation

Adult *A. ventricosus* (sample size = 46, height 4.4 ± 0.51 cm) were collected by diving from El Caracol (Guasave, Sinaloa, Mexico, 108°74'W, 25°49'N) and transported in 250 L plastic tanks containing seawater and constant aeration to the laboratory facilities. The organisms were acclimated to new culture conditions in the laboratory by placing them in a 1000 L plastic tank with filtered seawater (20 μm), 30 of salinity and under constant aeration for three days. The scallops were fed with a Shellfish Diet 1800® microalgae concentrate (Reed Mariculture Inc., Campbell, CA, USA) composed of *Isochrysis* sp., *Pavlova* sp., *Tetraselmis* sp., and *Thalassiosira pseudonana*.

Samples for the transcriptomic analysis were taken from 10 healthy animals that showed food consumption, food-filled intestine, rapid valve closure, and sensitivity to light change. The animals were placed in a 130 L plastic tank with 120 L of filtered seawater (20 μm) and constant aeration. Scallops were fed as previously described. Animals were challenged with *Vibrio parahaemolyticus* IPNGS16 by injecting 1.5×10^6 CFU (colony forming units) in 100 μL saline solution (2.5% NaCl) into the adductor muscle as described (López-León et al. 2016). The bacteria were previously inactivated by heat at 60°C for 10 min. Tissue samples, including gill, adductor muscle, digestive gland, hemocytes, and gonad, were separately obtained from four scallops at 24 h post-inoculation. Samples of each tissue and hemocytes from each organism were placed in RNALater (Sigma-Aldrich, Saint Louis, MO, USA) for subsequent ribonucleic acid (RNA) extraction. This study complied with the Mexican Official Standard NOM-062-ZOO-1999 for the care and use of laboratory animals.

Ribonucleic acid extraction, library construction, and sequenciation

The RNA extraction, library construction, and sequencing were carried out at the Instituto de Biotecnología, at Unidad Universitaria de Secuenciación Masiva y Bioinformática core facility of the Universidad Nacional Autónoma de Mexico in Cuernavaca, Morelos, Mexico. The RNA extraction was performed with the Quick-RNA Miniprep Kit (Zymo Research, Irvine, CA, USA). Then, an RNA pool was generated from all extracted RNA tissues from four scallops. The RNA pool was used as a template for the library construction with Illumina's TruSeq Stranded mRNA Sample Prep kit and sequenced with NextSeq 500/550 v2 kit on a NextSeq 500 platform 2×75 paired-end chemistry (Illumina Inc., San Diego, CA, USA).

Bioinformatic analysis

First, the quality control of the raw reads yielded by the sequencer was carried out with FastQC version 0.11.8 (predetermined parameters) on the Galaxy platform (Andrews 2010). The obtained reads were analyzed with Trimmomatic software version 0.36 on the Galaxy platform (Bolger et al. 2014) to remove very short sequences, low-quality bases, and adapters. The threshold quality score was set at $Q \geq 20$. Next, the reads underwent another quality control analysis based

on sequence length, quality score, and the base content using Prinseq v.1.20.4 on the Galaxy platform (Schmieder & Edwards 2011). Sequences smaller than 60 bp were removed. Sequences with a mean quality score below 20 were removed. No treatment was applied based on a maximum mean quality score, and sequences with more than 2% ambiguities "N" bases were removed. The 3'-end trimming was made based on a quality score threshold set at $Q \geq 20$. The final reads were *de novo* assembled with Trinity v.2.2.0 on the Galaxy platform (Grabherr et al. 2011, Haas et al. 2013) with default parameters. The contigs were run against *Pecten maximus* transcriptome in rnaQAST in the Galaxy platform to assess the quality of the assembly (Bushmanova et al. 2016).

Assembled sequences (contigs) were then analyzed with CD-HIT-EST using default parameters to eliminate redundant sequences in the Galaxy platform (Li & Godzik 2006, Fu et al. 2012). The resulting non-redundant sequences were analyzed with TransDecoder with default parameters to identify candidate coding sequence (CDS) regions within the assembled sequences on the Galaxy platform (Hass et al. 2013). The resulting sequences from TransDecoder were annotated through the FunctionAnnnotator web server, applying a similarity search against the Center for Biotechnology Information non-redundant database (NCBI-nr protein database) (eukaryotes). Also, sequences were analyzed to assign functional annotations with gene ontology terms using Blast2GO (Chen et al. 2012). Immune-related genes found in several research works were selected from the annotated sequences (NCBI-nr protein database) with their respective Trinity ID. These immune-related genes were analyzed with BLASTP (protein *vs*. protein database) against RefSeq database with an E-value of 1×10^{-5} (http://www.genome.jp/tools/blast/), the analysis was performed with the deduced amino acid sequences of immune-related genes (obtained by TransDecoder). Furthermore, the amino acid sequences (20,398) obtained from the annotation in Function Annotation were annotated against KAAS (KEGG automatic annotation server) that provides functional annotations of genes in a given genome by amino acid sequence comparisons against a set of manually curated orthologous groups in KEGG (Kyoto encyclopedia of genes and genomes). The search program GHOSTX (amino acids) was used against sequences of *Crassostrea gigas* and *Mizuhopecten yessoensis* to obtain KEGG pathways (https://www.genome.jp/kaas-bin/kaas_main? mode=partial). The KEGG reconstruction was performed (https://www.genome.jp/kegg/tool/ map_pathway.html) using the Kegg orthology (KO) file for the active metabolic pathways (Kanehisa et al. 2010). The raw reads obtained from Illumina NextSeq 500 were deposited in the Sequence Read Archive (SRA) of NCBI under accession number PRJNA684757.

Effect of bacillus mixture (BM) on the immune system of *A. ventricosus*

The effect of a single addition of 1×10^6 CFU L^{-1} of the PM (2×10^5 CFU L^{-1} each strain) was evaluated. The bioassay lasted 48 h. Twenty organisms (height 4.0-5.0 cm) were placed in a 130 L plastic tank with 120 L filtered seawater (20 µL) and constant aeration. The BM consisted of *Bacillus horikoshii*, *B. aerius*, *B. licheniformis*, *B. subtilis* and *B. pumillis*. The scallops were fed daily with a microalgal concentrate (*Isochrysis*, *Pavlova*, *Tetraselmis*, *Thalassiosira pseudonana*). Gentle aeration was placed in the plastic bottle with microalgae, and the feeding was done by gravity drip. Six scallops were taken at time zero (before bacilli inoculation) and after 48 h (after bacilli inoculation). The hemolymph was extracted from the adductor muscle with an insulin syringe and placed in a 1.5 mL tube chilled on ice. The samples were centrifuged in a refrigerated centrifuge (Sigma-Aldrich, St. Louis, MO, USA) at 800x *g* for 10 min. Then, 400 µL of RNALater was added to the hemocyte pellet for the immune system-related genes transcriptional analysis through RT-qPCR. This study complied with the Mexican Official Standard NOM-062-ZOO-1999 for the care and use of laboratory animals.

The immune system-related genes selected for transcriptional expression analyses were the peptide-glycan recognition protein (PGRP, forward: 5'-AACCCAGCTCATAGAACGACA-3', reverse: 5'-AGTCTCCGTATGACACGTTCG-3') and the heat shock protein 70 (HSP70, forward: 5'-TCAATA CCGAGGGACAAAGG-3', reverse: 5'-ACGGTTCA GGAACTGTTTGG-3'). The Ubiquinone (forward: 5'-CCATTTTTGGGTTGTTGAGG-3', reverse: 5'-ACC AGCATCAAA GGAACCAG-3') and 60s ribosomal protein (forward: 5'-AGGTATCTGGTCA TGCA AACG-3', reverse: 5'-TCTTACTGTAGCGGCAGC ATT-3') were used as reference genes for expression level normalization.

Ribonucleic acid extraction and reverse transcription-polymerase chain reaction

Hemocyte RNA was extracted according to the manufacturer's protocol (Invitrogen, Waltham, MA, USA) and stored at -70ºC until it was used in reverse transcription. The concentration and purity of total RNA were determined by measuring absorbance at 260/280 nm in a nanophotometer (Implen®, Westlake Village, CA, USA). Subsequently, the RNA was treated with DNAse 1 (1 U μL^{-1}, Sigma-Aldrich, St. Louis,

MO, USA). Reverse transcriptase (Improm II, Promega®, Madison, WI, USA) was used to synthesize the first strand (cDNA) with oligo dT20 from 500 ng of total RNA at 42ºC for 60 min. The cDNA was resuspended in 80 µL of ultrapure water and stored at -70°C for use in real-time PCR reactions. The reverse transcription was performed in the Bio-Rad T100TM Thermal Cycler (Hercules, CA, USA) equipment under the following temperature conditions: 5 min at 25°C, 60 min at 42°C, and 15 min at 70°C. The amplifications were performed on the CFX96 Touch Real-Time System (Bio-Rad® Hercules, CA, USA), and the CFX Manager v.3.0 software (Bio-Rad® Hercules, CA, USA) was used to obtain the data.

Primer quantity was standardized at 10 pM for relative expression analysis. A pool was made, taking 3 µL cDNA from each sample. Subsequently, five serial dilutions were made (1:5, 100 µL cDNA pool/400 µL ultrapure water). The reaction efficiency was determined for each gene of interest using a standard curve from the threshold cycle values (Cq) of the serial dilutions (triplicate) as detailed below. A simple linear regression analysis followed correlation coefficient (R^2) determination. The reaction efficiency for each gene was calculated from the slope values. Efficiency values were up to 90% by the standardization procedure. A dissociation curve (Melting curve) analysis was performed by increasing from 65 to 95°C (0.5°C every 5 s) to verify the absence of primer dimers.

For gene expression analysis, amplifications were performed by triplicate in 96-well plates with a final reaction volume of 15 µL containing 7.5 µL of PCR Master Mix 2x (3 µL of 5x reaction buffer; 1.5 µL of 25 mM $MgCl_2$; 0.3 µL of 10 mM dNTPs; 0.75 µL of EvaGreen 20x (Biotium, Fremont, CA, USA; 0.15 µL of Go Taq (Promega®, Madison, WI, USA); 1.8 µL of ultrapure water), 0.7-1.0 µL of primers (forward and reverse, 10 pM per primer, Sigma-Aldrich, St. Louis, MO, USA), 1.5-1.8 µL of ultrapure water and 5 µL of quenching (cDNA). The thermal cycler conditions for qPCR were: denaturation at 95°C for 3 min, followed by 40 cycles at 95°C for 10 s, 60°C for 15 s, 72°C for 30 s, and 79°C for 5 s. A triplicate calibrator was included in each gene analysis plate as a correction factor for the Cq values due to pipetting and reading variation from one run to another. For the PGRP gene, we used the HSP70 as a calibrator and vice versa. For ubiquinone, we used 60s ribosomal protein as calibrator and vice versa. The copy number of the target gene and the Cq values were inversely related, so a sample containing a greater number of copies of the target gene had a lower Cq value than that of a sample with a lower number of copies of the same gen. The efficiency of the PCR reaction was determined by calculating a slope with five serial dilutions (dilution factor 5 or 10) from a representative pool of cDNA [E = 10 (-1 / slope) - 1]. The expression of the target genes was determined with reference genes. Therefore, a stability analysis was carried out to find the best genes with GeNorm (Vandesompele et al. 2002) and NormFinder (Andersen et al. 2004) algorithms, using the RefFinder application (http://www. ciidirsinaloa.com.mx /RefFinder-master/) (Xie et al. 2012). The expression of the target genes was normalized with ubiquinone and 60s ribosomal protein.

The Cq values were transformed to relative quantities using the equation RQij = E [(average Cq) - Cq (ij)] to calculate the expression of the genes of interest, where E is the specific efficiency of the interest gene and [(Cq mean - Cq (ij)] is the absolute difference between the Cq of each sample and the average Cq in the data set of each interest gene. The relative expression was calculated with the RQgen equation of interest/geometric mean of reference RQgenes (Vandesompele et al. 2002, Hellemans et al. 2007).

Statistical analyses

A one-way ANOVA ($P < 0.05$) was performed to analyze immune-related gene expression in hemocytes. If differences were observed in the ANOVA, a Tukey honest significant discriminant (HSD) test was performed to compare averages and determine between which treatments there are significant differences ($P < 0.05$) (Statistica Software, StatSoft).

Ethical approval

The authors followed the study regulations of the international and national guidelines for the care and use of animals.

Data availability

The data supporting this study's findings are available from the corresponding author upon reasonable request.

RESULTS

The RNA-seq sequencing runs generated 51,033,670 raw reads with an average length of 74.52 ± 1.06 bp (Table 1). The quality control processing with Trimmomatic and Prinseq yielded ~46 million clean reads (Table 1). The quality assessment for *de novo* transcriptome assembly against *Pecten maximus* showed 107,516 transcripts, 32,930 aligned contigs, and 74,586 unaligned contigs (Table 2). The CD-HIT-EST analysis yielded 82,793 representative sequences. After TransDecoder analysis, 25,285 coding sequences (ORFs) were obtained. The hits against the NCBI-nr

Table 1. *Argopecten ventricosus*, sequencing, and quality of reads.

Category	Value
Total raw reads	51,033,670
Read average length (bp)	74.52 ± 1.06
Quality reads (Trimmomatic, Prinseq)	46,432,832

protein database obtained were 16,123. The contig average length was 866.69 bp with a GC content of 45.56%. The contig size N50 was 1113 bp. Results from gene ontology analysis showed 15,263 sequence hits (Table 3, Fig. 1).

The mapping against the National Center for Biotechnology Information (NCBI) taxonomy database (phylogenetic distribution) resulted in the following annotations (16,123) against mollusks: *Crassostrea gigas* (12,442), *Lottia gigantea* (1175), *Lingula anatina* (816), *Aplysia californica* (727), *Octopus bimaculoides* (616), and *Biomphalaria glabrata* (347) (Fig. 2).

From the 16,123 hits against NCBI-nr protein database (FunctionAnnotator), 15,263 hits were mapped to gene ontology (FunctionAnnotator). Gene ontology (GO) is a functional classification system widely implemented; its terms were used to classify the annotated genes of *A. ventricosus* for its molecular functions, cellular components, and biological processes. At level 2 GO terms, 15,263 contigs were annotated by Blast2GO v.5.2.5 (Table 3). Approximately 13,000 contigs were classified into molecular functions (Fig. 3), ~11,000 into cellular components (Fig. 4), and ~13,000 into biological processes (Fig. 5). Most of the terms in the molecular functions category corresponded to binding genes and catalytic activity. The cellular components dominated cell parts, organelles, organelle parts, and macro-molecular complex. In biological processes, most of the terms corresponded to genes of cellular processes, metabolic processes, biological regulation, and response to a stimulus.

The amino acid sequences (16,123) obtained from FunctionAnnotator were annotated against KAAS (KEGG). The KEGG analysis identified biological categories for metabolism (13%), cellular processes (12%), processing of genetic information (10%), organismal systems (19%), environmental information (13%), and human diseases (33%) (Fig. 6). Within the organismal systems, 467 immune-related genes were found and distributed in 20 pathways, such as genes of the NOD-receptor signaling pathway, chemokine signaling pathway, Toll receptor signaling pathway, C-type lectin receptor, transendothelial leukocyte migration, Fc gamma-mediated phagocytosis, T-cell receptor signaling pathway, NK cell-mediated cytoto-xicity, among others (Fig. 7).

Table 2. *Argopecten ventricosus*, quality assessment for *de novo* transcriptome assembly against *Pecten maximus*.

Metrics/Transcripts	Assembly quality
Transcripts (assembled sequences)	107,516
Transcripts > 500 bp	44,047
Transcripts > 1000 bp	21,183
Aligned	32,939
Uniquely aligned	31632
Multiply aligned	434
Unaligned	74,586
Avg. aligned fraction	0.7
Avg. aligned length	686.84
Avg. mismatches per transcript	41.76
Misassemblies	864

The annotation with the NCBI-nr/RefSeq protein database identified 64 immune-related genes. Among these genes, molecular pattern recognition receptors (MPRRs), immune signaling pathways, and immune effectors were found (Tables 4-6).

The expression of the HSP70 and PGRP genes at 48 h as compared to 0 h (control) was significantly increased ($P < 0.05$) by the addition of BM in the culture water of *A. ventricosus* (Fig. 8).

DISCUSSION

The survival of invertebrates against pathogen attacks depends on their ability to recognize them and induce appropriate defense responses (Janeway & Medzhitov 2002). Invertebrates, such as *A. ventricosus*, do not have an adaptive immune system (they do not produce antibodies or memory cells) as vertebrates possess. Hence, their defense against an environment rich in potentially pathogenic microorganisms relies on the innate immune system composed of cellular (hemocytes) and humoral effectors (lysosomal hydrolytic enzymes, antimicrobial peptides, superoxide dismutase, catalase, lysozyme) (Zheng 2011, Gerdol 2017). A transcriptomic analysis (RNA-seq) of the sequences obtained with the next generation sequencing (NGS) approach was performed using the Illumina platform to identify immune system-related genes.

This study obtained 51,033,670 raw reads that were *de novo* assembled to yield 25,285 candidate coding sequence regions (contigs). In previous reports on other bivalve mollusks, such as *Mizuhopecten yessoensis*, the differences of generated reads and contigs assembly yields (112,265,296 raw reads, 194,839 contigs) (Meng

Table 3. Statistical summary of *Argopecten ventricosus* transcriptome analysis.

Category	Value
CD-HIT-EST representative sequences	82,793
TransDecoder CDS	25,285
Hits against NCBI nr protein database (FunctionAnnotator)	16,123
Contig average length (bp, FunctionAnnotator)	866.69
GC content (%, FunctionAnnotator)	45.56
Contig size N50 (bp, FunctionAnnotator)	1113
Mapping to Gene Ontology (FunctionAnnotator)	15,263

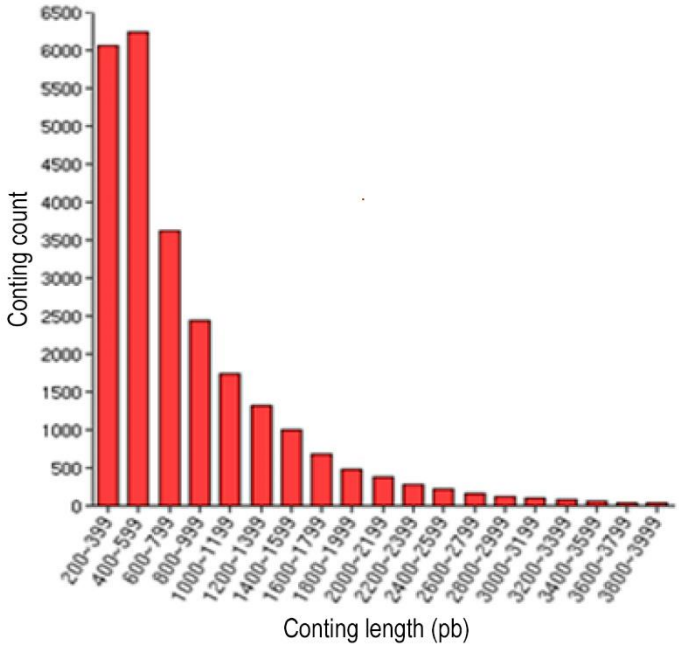

Figure 1. Contig total number and length obtained in the *Argopecten ventricosus* transcriptome assembly using FunctionAnnotator.

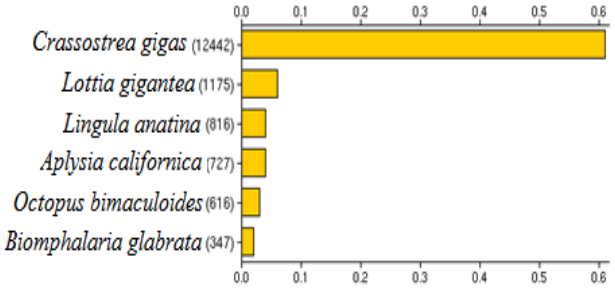

Figure 2. Taxonomic distribution of contigs annotated against the National Center for Biotechnology Information taxonomy database. Most of the sequences were scored with the Japanese oyster *Crassostrea gigas*. Data are in percentage.

et al. 2013), and *Crassostrea virginica* (52,857,842 raw reads, 61,147 contigs), may arise from the RNA treatment before NGS library preparation, raw reads quality, and even sequencing deep. Comparable to this study, Zhang et al. (2014) reported similar raw reads and contigs assembly yields for the bivalve mollusk *Mytilus galloprovincialis* (57,059,700 raw reads, 21,193 contigs). In addition, the average length contig in this study was 866.69 bp, similar to the average length contig reported previously for *C. virginica* (874 bp length), *M. galloprovincialis* (771 bp length), and *Cristaria plicata* (731 bp length) (Meng et al. 2013, Patnaik et al. 2016) but higher than *M. yessoensis* (461 bp length).

The mapping against the NCBI taxonomy database (phylogenetic distribution) showed that most of the sequences were annotated with *C. gigas* (12,442), which strongly suggested that this bivalve mollusk was a close phylogenetic or evolutionary relative of *A. ventricosus*. Within the annotated genes with GO terms, a wide variety of functional categories represented at

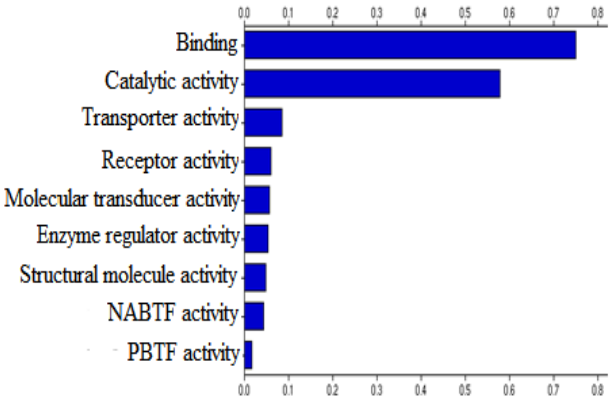

Figure 3. Gene ontology analysis of *Argopecten ventricosus* transcriptome at level 2 corresponding to molecular functions. The terms were obtained using Blast2GO. NABTF activity: nucleic acid binding transcription factor activity. PBTF activity: protein binding transcription factor activity. Data are in percentage.

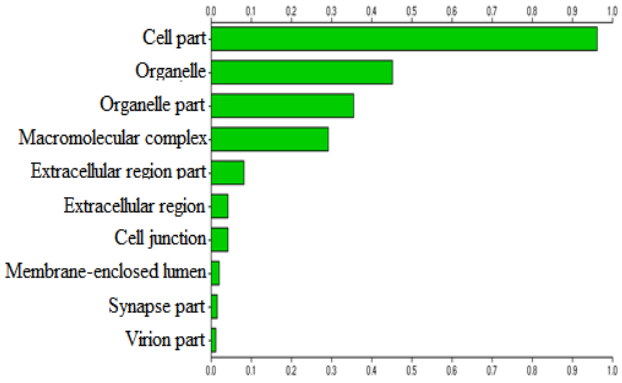

Figure 4. Gene ontology analysis of *Argopecten ventricosus* transcriptome at level 2 corresponding to cellular components. The terms were obtained using Blast2GO. Data are in percentage.

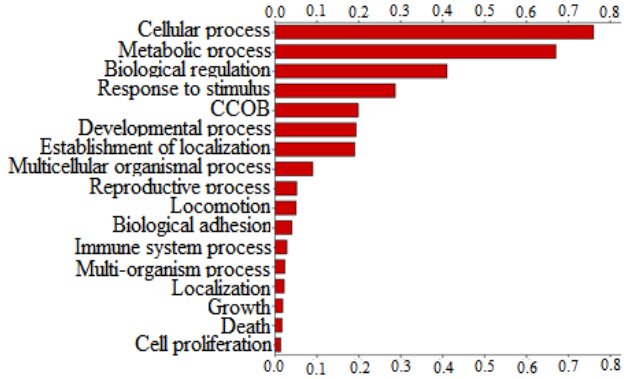

Figure 5. Analysis of *Argopecten ventricosus* transcriptome gene ontology at level 2 corresponding to biological processes. The terms were obtained using Blast2GO. CCOB: cellular component organization of biogenesis. Data are in percentage.

level 2 of the GO database corresponded to molecular functions (9 GO terms), cellular components (10 GO terms), and biological processes (17 GO terms). In the category of molecular functions, most of the terms corresponded to binding and catalytic activity, which was in accordance with the reported GO annotations for *M. yessoensis* (Meng et al. 2013). However, different results were reported in *Mytilus edulis* (Tanguy et al. 2013), where the terms of catalytic activity predominated over those of binding activity. In the cellular components, the terms cell part, organelles, organelle part, and macromolecular complexes predominated similarly to that found in *M. yessoensis* (Meng et al. 2013); however, no coincidence was found regarding the terms organelles and macromolecular complexes because the second and third most abundant term in *M. yessoensis* were intracellular and cytoplasm, respectively. In *M. edulis* predominated the terms cell part, intracellular, and membranes (Tanguy et al. 2013). In biological processes, most of the terms corresponded to genes of cellular processes, metabolic processes, and biological regulation, which coincide with those found in *M. yessoensis* (Meng et al. 2013); however, they differ from those found in *M. edulis* (Tanguy et al. 2013) in which the predominant terms were metabolic and cellular, respectively.

Interestingly, the GO annotation distribution of *A. ventricosus* was more similar for *M. yessoensis* exposed to heavy metal than *M. edulis* treated with an active *Vibrio splendidus*. Perhaps this bivalve mollusk makes no big difference between an inactivated pathogen (inactivated *V. parahaemolyticus* in this study) and an abiotic stressor, such as a high heavy metal concentration. Furthermore, the GO term annotation for *A. ventricosus* showed similarities with the GO term annotation of *Patinopecten yessoensis*, *Crassostrea hongkongensis*, and *C. plicata*, which strongly suggested that those bivalve mollusks and *A. ventricosus* were closely related with a basal and essential immune response to the environment, pathogens, growth, and development (Hou et al. 2011, Tong et al. 2015, Patnaik et al. 2016).

The information about the genes of mollusks directly involved in the innate immune response has increased but is still quite insufficient (Moreira et al. 2012). In this study, two different annotation analyses were carried out (NCBI-nr/RefSeq and KEGG KAAS), and both analyses identified several sequences that showed homology with pattern recognition receptors (PRRs), immune signaling pathways, and immune effectors following those found in the oyster *Saccostrea glomerata* (Ertl & O'Connor 2016). Most immune-related genes were obtained in the KO analysis (467

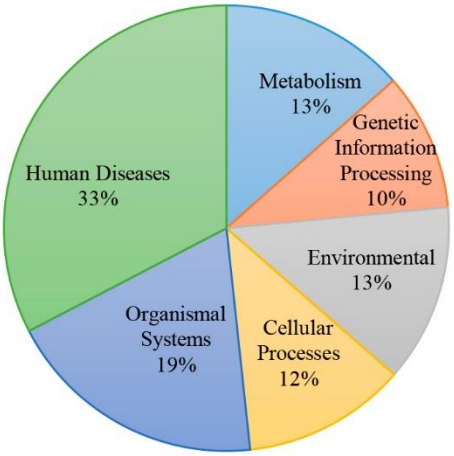

Figure 6. Annotation of *Argopecten ventricosus* unigenes with the KEGG database. Distribution of unigenes in biological categories of KEGG.

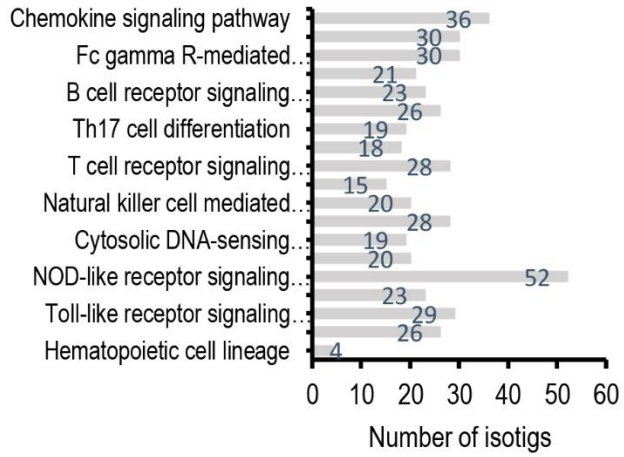

Figure 7. Unigenes classification of the KEGG analysis in the category "immune system". *Argopecten ventricosus* unigenes annotation with the KEGG database.

annotations) (Kanehisa et al. 2010) compared to those identified in NCBI-nr/RefSeq database (64 annotations). The above findings were consistent with *C. virginica* (Zhang et al. 2014), where 895 annotations were observed with KO and 62 annotations with NCBI-nr.

The innate defense system is found in plants, fungi, and all animals; based on this evolutionary wideness, its origin might be ancient (Macagno et al. 2010). Its efficacy, celerity, and robustness have remained the first defense line against pathogens (Macagno et al. 2010). According to Zhang et al. (2014), the invertebrate innate immune system is more sophisticated than that of vertebrates, which, in this case, allows coping with an aquatic environment rich in potential pathogens. Bivalve mollusks have an open circulatory system, and hemocytes are found in the hemolymph (circulating cells), as well as in tissues (infiltrating cells) (Song et al. 2011). Hemocytes have molecular pattern recognition receptors (MPRRs) (Zheng 2011). These receptors recognize the pathogen-associated molecular patterns (PAMPs) present in pathogens. PAMPs include a full variety of molecules, such as lipopolysaccharides from Gram-negative bacteria, peptidoglycans from Gram-positive and negative bacteria, β-1,3-glucans from fungi, bacterial DNA, and viral RNA. The PAMPs recognition triggers a varied immune response in invertebrates (Kawasaki et al. 2011, Tatematsu et al. 2018). In this study, several MPRRs were found in *A. ventricosus*, such as Toll-like receptors, peptidoglycan recognition protein, galectins, stabilin, and ficolins. Most of the found MPRRs has also been expressed in other mollusks, such as *Argopecten irradians* (Song et al. 2011), *M. edulis* (Philipp et al. 2012), *Pinctada margaritifera* (Teaniniuraitemoana et al. 2014), *C. virginica* (Zhang et al. 2014), *Venerupis philippinarum* (Mu et al. 2014), *C. gigas* (Huang et al. 2015), *Chlamys farreri* (Song et al. 2015) and *S. glomerata* (Ertl & O'Connor 2016).

In *A. ventricosus*, we were able to find coding sequences of proteins directly involved in regulating signaling pathways of Toll-like receptors (TOLLIP, IRAK1BP1, SOCS, NFKBIL1, and HSPs) similar to those found in *S. glomerata* (Ertl & O'Connor 2016). In bivalves, phagocytosis is an important defense mechanism where hemocytes (cellular effectors) are responsible (Donaghy et al. 2012). The internalization and phagocytosis of microorganisms are recognized by hemocyte receptors (TLRs, mannose receptor, and 2-stabilizer) (Malagoli et al. 2007, Underhill & Goodridge, 2012). Reactive oxygen species (ROS) production occurs during phagocytosis in the respiratory burst with increased intracellular oxygen consumption. ROS production is a mechanism dependent on NADPH oxidase. ROS are molecules that oxidize components of microbial cells (proteins, lipids, and nucleic acids). However, the intracellular production of antioxidant molecules is also activated to prevent hemocyte damage. These antioxidant molecules are carotenoids, glutathione S-transferase, superoxide dismutase, catalase, glutathione peroxidase, glutaredoxin, thioredoxin, and peroxiredoxin (Jenny et al. 2002). Similar to these results, transcripts of the antioxidant system have been reported in other bivalve mollusks, such as *C. virginica* (Jenny et al. 2002), *Ruditapes decussatus* (Leite et al. 2013), *Pinctada fucata* (Kuchel et al. 2012), *C. gigas* (Trevisan et al. 2012), and *S. glomerata* (Ertl & O'Connor 2016).

During the phagocytosis process, the phagosome is fused with the lysosome to form the phagolysosome for

Table 4. The transcriptome of *Argopecten ventricosus*. Sequences of immune and immune-related genes annotated/blasted against the National Center for Biotechnology Information NCBI-nr and RefSeq protein database. Molecular pattern recognition.

Molecular pattern recognition receptors (MPRRs)	Contig identification (Contig-ID)	E-value (nr-aa RefSeq)
Collectin-12	TRINITY_DN10857_c0_g1::TRINITY_DN10857_c0_g1_i1::g.2343::m.2343	4e-31 *Pecten maximus*
Peptidoglycan recognition protein	TRINITY_DN27560_c1_g1::TRINITY_DN27560_c1_g1_i1::g.10446::m.10446	1e-86 *Mizuhopecten yessoensis*
Lipopolysaccharide- and beta-1,3-glucan-binding protein	TRINITY_DN40092_c0_g1::TRINITY_DN40092_c0_g1_i1::g.25409::m.25409	2e-174 *P. maximus*
Lipopolysaccharide-binding protein	TRINITY_DN29543_c1_g1::TRINITY_DN29543_c1_g1_i2::g.12005::m.12005	5e-83 *M. yessoensis*
Ficolin-2-like	TRINITY_DN31029_c0_g1::TRINITY_DN31029_c0_g1_i1::g.13358::m.13358	7e-103 *Crassostrea virginica*
Scavenger receptor cysteine-rich protein	TRINITY_DN40442_c1_g1::TRINITY_DN40442_c1_g1_i3::g.26090::m.26090	3e-90 *P. maximus*
Fibrinogen-related protein	TRINITY_DN36559_c1_g1::TRINITY_DN36559_c1_g1_i1::g.19397::m.19397	8e-75 *Mytilus galloprovincialis*
Galectin	TRINITY_DN36670_c0_g1::TRINITY_DN36670_c0_g1_i1::g.19600::m.19600	3e-58 *P. maximus*
Galectin 4	TRINITY_DN30407_c0_g2::TRINITY_DN30407_c0_g2_i1::g.12795::m.12795	1e-25 *P. maximus*
C-type lectin 5	TRINITY_DN21812_c0_g1::TRINITY_DN21812_c0_g1_i1::g.6470::m.6470	1e-42 *P. maximus*
C-type lectin 2	TRINITY_DN25185_c0_g1::TRINITY_DN25185_c0_g1_i3::g.8679::m.8679	4e-78 *P. maximus*
C-type lectin 1	TRINITY_DN28185_c0_g1::TRINITY_DN28185_c0_g1_i2::g.10941::m.10941	6e-74 *P. maximus*
C-type lectin 7	TRINITY_DN29489_c0_g1::TRINITY_DN29489_c0_g1_i1::g.11981::m.11981	9e-108 *P. maximus*
Toll-like receptor 13	TRINITY_DN20515_c0_g1::TRINITY_DN20515_c0_g1_i1::g.5717::m.5717	4e-46 *P. maximus*
Toll-like receptor m	TRINITY_DN1375_c0_g1::TRINITY_DN1375_c0_g1_i1::g.314::m.314	1e-105 *P. maximus*
Stabilin-2	TRINITY_DN41727_c2_g1::TRINITY_DN41727_c2_g1_i2::g.29077::m.29077	9e-174 *Octopus vulgaris*

intracellular microbial degradation. The phagolysosome has an acidic pH (4.0-4.5) and contains cathepsins, superoxide anion, hydroxyl radicals, hydrogen peroxide, nitrogen radicals, lysosomal hydrolytic enzymes, proteins, antimicrobial peptides, and lysozymes (Yang et al. 2010, Ertl & O'Connor 2016). Hydrolytic enzymes have lysis properties because they can degrade peptidoglycans from the bacterial cell wall (Xue et al. 2010). Antimicrobial peptides, such as large defensins, are important immune effectors due to their bactericidal activity and immunomodulatory functions (Guaní-Guerra et al. 2010). Cathepsins are proteases that degrade proteins in the phagolysosome; in mollusks, the presence of cathepsins, cysteine-, aspartic- and serine-types have been reported (Conus & Simon 2008, Guha & Padh 2008, Turk et al. 2012). In addition to this study, gene coding sequences transcripts for proteins with antimicrobial activity and antimicrobial-associated peptides have been found in other bivalves, such as *Mactra veneriformis* (Fang et al. 2013), *Sinonovacula constricta* (Niu et al. 2014), and *S. glomerata* (Ertl & O'Connor 2016).

This study determined the expression of two genes related to the immune system (HSP70 and PGRP) in the Pacific calico scallop. The HSP proteins are expressed during heat and oxidative stress, malnutrition, viral and bacterial infections, among other stimulating conditions (Hickey & Singer 2004, Asea 2007, Multhoff 2007), and they have a role as chaperones in the immune response (Spagnolo et al. 2007). In the same species of mollusk studied in this work, López-Carballo

Table 5. The transcriptome of *Argopecten ventricosus*. Sequences of immune and immune-related genes annotated/blasted against the National Center for Biotechnology Information NCBI-nr and RefSeq protein database. Immune signaling pathways.

Immune signaling pathways	Contig identification (Contig-ID)	E-value (nr-aa RefSeq)
MyD88-1	TRINITY_DN42751_c2_g2::TRINITY_DN42751_c2_g2_i1::g.31980::m.31980	8e-81 *Crassostrea gigas*
Sterile alpha and TIR motif-containing protein 1	TRINITY_DN20462_c0_g1::TRINITY_DN20462_c0_g1_i1::g.5684::m.5684	3e-72 *Pecten maximus*
Interleukin-1 receptor-associated kinase 4-like	TRINITY_DN41086_c2_g1::TRINITY_DN41086_c2_g1_i2::g.27562::m.27562	5e-147 *C. virginica*
Tumor necrosis factor receptor-associated factor 3 TRAF	TRINITY_DN28258_c0_g1::TRINITY_DN28258_c0_g1_i1::g.11003::m.11003	4e-170 *Mizuhopecten yessoensis*
Inhibitor of nuclear factor kappa-B kinase subunit alpha	TRINITY_DN26033_c1_g1::TRINITY_DN26033_c1_g1_i1::g.9238::m.9238	5e-76 *P. maximus*
Nuclear factor related to kappa-B-binding protein	TRINITY_DN25793_c1_g1::TRINITY_DN25793_c1_g1_i1::g.9081::m.9081	8e-126 *P. maximus*
Inhibitor of nuclear factor-kappaB protein	TRINITY_DN41925_c0_g2::TRINITY_DN41925_c0_g2_i2::g.29645::m.29645	3e-158 *P. maximuns*
Interferon regulatory factor 2	TRINITY_DN30196_c1_g1::TRINITY_DN30196_c1_g1_i1::g.12590::m.12590	8e-146 *M. yessoensis*
Tyrosine-protein kinase JAK2-like	TRINITY_DN40830_c2_g3::TRINITY_DN40830_c2_g3_i1::g.26926::m.26926	5e-177 *C. gigas*
Mitogen-activated protein kinase 5	TRINITY_DN14762_c2_g1::TRINITY_DN14762_c2_g1_i1::g.3170::m.3170	8e-76 *P. maximus*
Protein pellino	TRINITY_DN15403_c1_g1::TRINITY_DN15403_c1_g1_i1::g.3399::m.3399	3e-147 *P. maximus*
Toll interacting protein (TOLLIP)	TRINITY_DN39263_c0_g1::TRINITY_DN39263_c0_g1_i2::g.23880::m.23880	2e-120 *C. gigas*
Suppressor of cytokine signaling 5	TRINITY_DN27846_c0_g1::TRINITY_DN27846_c0_g1_i1::g.10663::m.10663	3e-127 *C. gigas*
NF-kappa-B inhibitor-like protein 1	TRINITY_DN30626_c0_g1::TRINITY_DN30626_c0_g1_i1::g.13007::m.13007	9e-66 *Crassostrea virginica*
Interleukin-1 receptor-associated kinase 1-binding protein 1-like (IRAK1BP1)	TRINITY_DN29725_c0_g1::TRINITY_DN29725_c0_g1_i1::g.12152::m.12152	5e-175 *P. maximus*
Nuclear factor of activated T-cells	TRINITY_DN38502_c1_g2::TRINITY_DN38502_c1_g2_i1::g.22479::m.22479	6e-157 *C. gigas*
Endogenous ligands		
Fibrinogen-related protein	TRINITY_DN38885_c2_g1::TRINITY_DN38885_c2_g1_i1::g.23151::m.23151	3e-105 *P. maximus*
Heat shock protein 60	TRINITY_DN21539_c0_g1::TRINITY_DN21539_c0_g1_i1::g.6305::m.6305	5e-44 *C. virginica*
Heat shock protein 70 B2-like	TRINITY_DN27614_c2_g1::TRINITY_DN27614_c2_g1_i1::g.10508::m.10508	8e-60 *M. yessoensis*
Heat shock protein 70	TRINITY_DN39249_c0_g2::TRINITY_DN39249_c0_g2_i1::g.23723::m.23723	1e-256 *Crassostrea hongkongensis*
Heat shock protein 22	TRINITY_DN34749_c0_g1::TRINITY_DN34749_c0_g1_i1::g.17160::m.17160	4e-49 *C. gigas*
Stress-induced-phosphoprotein 1	TRINITY_DN40238_c2_g1::TRINITY_DN40238_c2_g1_i1::g.25691::m.25691	2e-179 *Lingula anatina*
Intraflagellar transport protein 25	TRINITY_DN37599_c0_g1::TRINITY_DN37599_c0_g1_i1::g.20854::m.20854	3e-94 *P. maximus*
Hypoxia-inducible factors		
Hypoxia-inducible factor 1-alpha inhibitor-like	TRINITY_DN39313_c3_g1::TRINITY_DN39313_c3_g1_i1::g.23978::m.23978	2e-173 *C. gigas*

Table 6. The transcriptome of *Argopecten ventricosus*. Sequences of immune and immune-related genes annotated/blasted against the National Center for Biotechnology Information NCBI-nr and RefSeq protein database. Immune effectors.

Molecular pattern recognition receptors (MPRRs)	Contig identification (Contig-ID)	E-value (nr-aa RefSeq)
Collectin-12	TRINITY_DN10857_c0_g1::TRINITY_DN10857_c0_g1_i1::g.2343::m.2343	4e-31 *Pecten maximus*
Peptidoglycan recognition protein	TRINITY_DN27560_c1_g1::TRINITY_DN27560_c1_g1_i1::g.10446::m.10446	1e-86 *Mizuhopecten yessoensis*
lipopolysaccharide- and beta-1,3-glucan-binding protein	TRINITY_DN40092_c0_g1::TRINITY_DN40092_c0_g1_i1::g.25409::m.25409	2e-174 *P. maximus*
lipopolysaccharide-binding protein	TRINITY_DN29543_c1_g1::TRINITY_DN29543_c1_g1_i2::g.12005::m.12005	5e-83 *M. yessoensis*
Ficolin-2-like	TRINITY_DN31029_c0_g1::TRINITY_DN31029_c0_g1_i1::g.13358::m.13358	7e-103 *Crassostrea virginica*
Scavenger receptor cysteine-rich protein	TRINITY_DN40442_c1_g1::TRINITY_DN40442_c1_g1_i3::g.26090::m.26090	3e-90 *P. maximus*
Fibrinogen-related protein	TRINITY_DN36559_c1_g1::TRINITY_DN36559_c1_g1_i1::g.19397::m.19397	8e-75 *Mytilus galloprovincialis*
Galectin	TRINITY_DN36670_c0_g1::TRINITY_DN36670_c0_g1_i1::g.19600::m.19600	3e-58 *P. maximus*
Galectin 4	TRINITY_DN30407_c0_g2::TRINITY_DN30407_c0_g2_i1::g.12795::m.12795	1e-25 *P. maximus*
C-type lectin 5	TRINITY_DN21812_c0_g1::TRINITY_DN21812_c0_g1_i1::g.6470::m.6470	1e-42 *P. maximus*
C-type lectin 2	TRINITY_DN25185_c0_g1::TRINITY_DN25185_c0_g1_i3::g.8679::m.8679	4e-78 *P. maximus*
C-type lectin 1	TRINITY_DN28185_c0_g1::TRINITY_DN28185_c0_g1_i2::g.10941::m.10941	6e-74 *P. maximus*
C-type lectin 7	TRINITY_DN29489_c0_g1::TRINITY_DN29489_c0_g1_i1::g.11981::m.11981	9e-108 *P. maximus*
Toll-like receptor 13	TRINITY_DN20515_c0_g1::TRINITY_DN20515_c0_g1_i1::g.5717::m.5717	4e-46 *P. maximus*
Toll-like receptor m	TRINITY_DN1375_c0_g1::TRINITY_DN1375_c0_g1_i1::g.314::m.314	1e-105 *P. maximus*
Stabilin-2	TRINITY_DN41727_c2_g1::TRINITY_DN41727_c2_g1_i2::g.29077::m.29077	9e-174 *Octopus vulgaris*

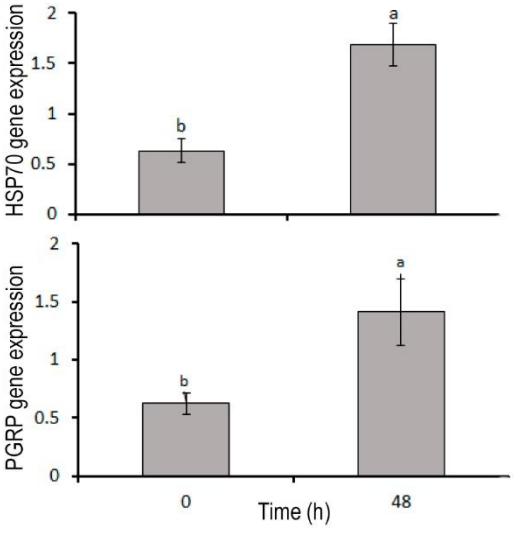

Figure 8. Relative expression of HSP70 and peptidoglycan recognition protein (PGRP) genes at 0 (before bacilli inoculation) and 48 h (after bacilli inoculation) in hemocytes of *Argopecten ventricosus* treated with bacilli (1×10^6 CFU L^{-1}). Columns and bars indicate the mean ± standard error. Different letters indicate significant differences ($P < 0.05$).

et al. (2020) observed a significant up-regulation of the HSP90 gene in the mantle of organisms treated with *Vibrio* lysates. In contrast to this study, Laranja et al. (2017) reported that HSP70 expression was unaffected in *Penaeus monodon* postlarvae fed *Artemia* enriched with *Bacillus* sp. JL47 and challenged with *Vibrio campbelli*.

On the other hand, the expression of the PGRP was significantly higher in scallops treated with bacilli compared to the control condition. The PGRP participates in the innate immune response by degrading the peptidoglycan wall in bacteria. For instance, in the fish *Oplegnathus fasciatus*, PGRP has a very important role in the agglutination of Gram (-) and Gram (+) bacteria and promotes phagocytosis in leukocytes (Choi et al. 2018).

CONCLUSION

In this is work, we found immune-related gene transcriptional activity in *A. ventricosus*, which highlights the complexity of its innate immune system. The expression of two genes was highly modulated upon a probiotic treatment. Thus far, NGS sequencing coupled with transcriptomic analysis (RNA-seq) identified several immune-related genes, which undoubtedly confirms that the Pacific calico scallop genome harbors genes related to the immune system, like MPRRs, immune signaling pathways, and immune effectors similar to that reported in other bivalve mollusks. This study opens the way for future research towards developing a sustainable *A. ventricosus* aquaculture.

ACKNOWLEDGMENT

This work was supported by Secretaría de Investigación y Posgrado del Instituto Politécnico Nacional (SIP-IPN, grant A160865). Gabriela Berenice Mendoza-Maldonado acknowledges support from the Consejo Nacional de Ciencia y Tecnología (CONACyT-Mexico, grant #416006) and SIP-IPN for grants to perform doctoral studies; to D. Fischer for editorial services.

REFERENCES

Andersen, C.L., Jensen, J.L. & Orntoft, T.F. 2004. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Research, 64: 5245-5250. doi: 10.1158/0008-5472.CAN-04-0496

Andrews, S. 2010. FastQC: a quality control tool for high throughput sequence data. [http://www.bioinformatics.babraham.ac.uk/projects/fastqc]. Reviewed: January 5, 2021.

Asea, A. 2007. Mechanisms of HSP72 release. Journal of Biosciences, 32: 579-584. doi: 10.1007/s12038-007-0057-5

Batista, F., Churcher, A.M., Manchado, M., Leitão, A. & Power, D.M. 2019. Uncovering the immunological repertoire of the carpet shell clam *Ruditapes decussatus* through a transcriptomic-based approach. Aquaculture and Fisheries, 4: 37-42. doi: 10.1016/j.aaf.2018.03.003

Bolger, A.M., Lohse, M. & Usadel, B. 20014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, 30: 2114-2120. doi: 10.1093/bioinformatics/btu170

Bushmanova, E., Antipov, D., Lapidus, A., Suvorov, V. & Prjibelski, A.D. 2016. rnaQUAST: a quality assessment tool for de novo transcriptome assemblies: Table 1. Bioinformatics, 32: 2210-2212. doi: 10.1093/bioinformatics/btw218

Carballal, M.J., López, C., Azevedo, C. & Villalba, A. 1997. Enzymes involved in defense functions of hemocytes of mussel *Mytilus galloprovincialis*. Journal of Invertebrate Pathology, 79: 96-105. doi: 10.1006/jipa.1997.4670

Chen, T.W., Ruei-Chi, R.G., Wu, T., Po-Jung, H., Cheng-Yang, L., Yi-Ywan, M.C., et al. 2012. FastAnnotator: an efficient transcript annotation web tool. BMC Genomics, 13: S9. doi: 10.1186/1471-2164-13-S7-S9

Choi, K.M., Joo, M.S., Cho, D.H., Bae, J.S., Jeong, J.M., Woo, W.S., et al. 2019. Molecular characterization, expression and functional analysis of peptidoglycan recognition protein-SC2 from rock bream, *Oplegnathus fasciatus*. Fish and Shellfish Immunology, 77: 286-293. doi: 10.1016/j.fsi.2018.04.006

Coan, E.V. & Valentich-Scott, P. 2012. Bivalve seashells of tropical West America. Marine bivalve mollusks from Baja California to Peru. Monographs, 6: 12-58.

Conus, S. & Simon, H.U. 2008. Cathepsins: key modulators of cell death and inflammatory responses. Biochemical Pharmacology, 76: 1374-1382. doi: 10.1016/j.bcp.2008.07.041

Donaghy, L., Kraffe, E., Le-Goïc, N., Lambert, C., Volety, A.K. & Soudant, P. 2012. Reactive oxygen species in unstimulated hemocytes of the Pacific oyster *Crassostrea gigas*: a mitochondrial involvement. Plos One, 7: e46594. doi: 10.1371/journal.pone.0046594

Ertl, N.G. & O'Connor, W.A., Papanicolaou, A., Wiegand, A.N. & Elizur, A. 2016. Transcriptome analysis of the Sydney rock oyster, *Saccostrea glomerata*: insights into molluscan immunity. Plos One, 11: e0156649. doi: 10.1371/journal.pone.0156649

Fang, Y., Yang, H., Liu, B. & Zhang, L. 2013. Transcriptional response of lysozyme, metallothionein, and superoxide dismutase to combined exposure to heavy metals and bacteria in *Mactra veneriformis*. Comparative Biochemistry and Physiology - Part C: Toxicology, 157: 54-62. doi: 10.1016/j.cbpc.2012.10.002

Flores-Garza, R., García-Ibáñez, S., Flores-Rodríguez, P., Torreblanca, C., Galeana, L., Valdés-González, A., et al. 2017. Commercially important marine mollusks for human consumption in Acapulco, Mexico. Natural Resources, 3: 11-17. doi: 10.4236/nr.2012.31003

Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics, 28: 3150-3152. doi: 10.1093/bioinformatics/bts565

Gerdol, M. 2017. Immune-related genes in gastropods and bivalves: a comparative overview. Invertebrate Survival Journal, 14: 95-111. doi: 10.25431/1824-307X/isj.v14i1.103-118

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I. & Regev, A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnology, 29: 644-652. doi: 10.1038/nbt.1883

Guaní-Guerra, E., Santos-Mendoza, T., Lugo-Reyes, S.O. & Teran, L.M. 2010. Antimicrobial peptides: general overview and clinical implications in human health and disease. Clinical Immunology, 135: 1-11. doi: 10.1016/j.clim.2009.12.004

Guha, S. & Padh, H. 2008. Cathepsins: fundamental effectors of endolysosomal proteolysis. Indian Journal of Biochemistry and Biophysics, 45: 75-90.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Philip, D., Bowden, J., et al. 2013. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nature Protocols, 8: 1494-1512. doi: 10.1038/nprot.2013.084

Hellemans, J., Mortier, G., De Paepe, A., Speleman, F. & Vandesompele, J. 2007. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biology, 8: R19. doi: 10.1186/gb-2007-8-2-r19

Hickey, D. & Single, G. 2004. Genomic and proteomic adaptations to growth at high temperature. Genome Biology, 5: 117. doi: 10.1186/gb-2004-5-10-117

Hine, P.M. 1999. The inter-relationships of bivalve haemocytes. Fish and Shellfish Immunology, 9: 367-385. doi: 10.1006/fsim.1998.0205

Hou, R., Bao, Z., Wang, S., Su, H., Li, Y. & Du, H. 2011. Transcriptome sequencing and de novo analysis for Yesso scallop (*Patinopecten yessoensis*) using 454 GS FLX. Plos One, 6: e21560. doi: 10.1371/journal.pone.0021560

Huang, B., Zhang, L., Li, L., Tang, X. & Zhang, G. 2015. Highly diverse fibrinogen-related proteins in the Pacific oyster *Crassostrea gigas*. Fish and Shellfish Immunology, 43: 485-490. doi: 10.1016/j.fsi.2015.01.021

Janeway, C.A. & Medzhitov, R. 2002. Innate immune recognition. Annual Review of Immunology, 20: 197-216. doi: 10.1146/annurev.immunol.20.083001.084359

Jenny, M.J., Ringwood, A.H., Lacy, E.R., Lewitus, A.J., Kempton, J.W. & Gross, P.S. 2002. Potential indicators of stress response identified by expressed sequence tag analysis of hemocytes and embryos from the American oyster, *Crassostrea virginica*. Marine Biotechnology, 4: 81-93. doi: 10.1007/s10126-001-0072-8

Kanehisa, M., Goto, S., Furumichi, M., Tanabe, M. & Hirakawa, M. 2010. KEGG for representation and analysis of molecular networks involving diseases and drugs. Nucleic Acids Research, 38: 355-360. doi: 10.1093/nar/gkp896

Kawasaki, T., Kawai, T. & Shizuo, A. 2011. Recognition of nucleic acids by pattern-recognition receptors and its relevance in autoimmunity. Immunological Reviews, 243: 61-73. doi: 10.1111/j.1600-065X.2011.01048.x

Kuchel, R.P., Nair, S. & Raftos, D.A. 2012. Changes in the transcriptional expression of oxidative stress response genes in Akoya pearl oysters (*Pinctada fucata*) exposed to air and mechanical agitation. Aquaculture, 362-363: 33-38. doi: 10.1016/j.aquaculture.2012.07.027

Laranja, J.L.Q., Amar, E.C., Ludevese-Pascual, G.L., Niu, Y., Geaga, M.J., De Schryver, P. & Bossier, P. 2017. A probiotic *Bacillus* strain containing amorphous poly-betahydroxybutyrate (PHB) stimulates the innate immune response of *Penaeus monodon* postlarvae. Fish and Shellfish Immunology, 68: 202-210. doi: 10.1016/j.fsi.2017.07.023

Leite, R.B., Milan, M., Coope, A., Bortoluzzi, S., Anjos, A. & Reinhardt, R. 2013. mRNASeq and microarray development fot the Grooved carpet shell clam, *Ruditaped decussatus*: a functional approach to unravel host-parasite interaction. BMC Genome, 14: 741. doi: 10.1186/1471-2164-14-741

Li, W. & Godzik, A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics, 22: 1658-1659. doi: 10.1093/bioinformatics/btl158

López-León, P., Luna-González, A., Escamilla-Montes, R., Flores-Miranda, M.C., Fierro-Coronado, J.A., Álvarez-Ruiz, P. & Diarte-Plata, G. 2016. Isolation and characterization of infectious *Vibrio parahaemolyticus*, the causative agent of AHPND, from the whiteleg shrimp (*Litopenaeus vannamei*). Latin American Journal of Aquatic Research, 44: 470-479. doi: 10.3856/vol44-issue3-fulltext-5

López-Carvallo, J.A., Mazón-Suástegui, J.M., Hernández-Oñate, M.A., Tovar-Ramírez, D., Abasolo-Pacheco, F. & Morelos-Castro, R.M. 2020. Transcriptome analysis of catarina scallop (*Argopecten ventricosus*) juveniles treated with highly-diluted immunomodulatory compounds reveals activation of non-self-recognition system. Plos One, 15: e0233064. doi: 10.1371/journal.pone.0233064

Macagno, E., Gaasterland, T., Edsall, L., Bafna, V., Soares, M., Scheetz, T., et al. 2010. Construction of a medicinal leech transcriptome database and its application to the identification of leech homologs of neural and innate immune genes. BMC Genome, 11: 407. doi: 10.1186/1471-2164-11-407

Maeda-Martínez, A.N., Reynoso-Granados, T., Solís-Marín, F., Leija-Tristán, A., Aurioles-Gamboa, D., Salinas-Zavala, C., et al. 1993. A model to explain the formation of catarina scallop, *Argopecten circularis* (Sowerby, 1835), beds, in Magdalena Bay, Mexico. Aquaculture and Fisheries Management, 24: 323-339. doi: 10.1111/j.1365-2109.1993.tb00554.x

Malagoli, D., Casarini, L., Sacchi, S. & Ottaviani, E. 2007. Stress and immune response in the mussel *Mytilus galloprovincialis*. Fish and Shellfish Immunology, 23: 171-177. doi: 10.1016/j.fsi.2006.10.004

Matozzo, V., Pagano, M., Spinelli, A., Caicci, F. & Faggio, C. 2016. *Pinna nobilis*: a big bivalve with big haemocytes. Fish and Shellfish Immunology, 55: 529-534. doi: 10.1016/j.fsi.2016.06.039

Mazón-Suástegui, J.M., Lodeiros, C., Avilés-Quevedo, A., Rodríguez-Jaramillo, C., Ortiz-Cornejo, N. & Abasolo-Pacheco, F. 2010. Cornstarch as a dietary supplement in conditioning broodstock and spat at a nursery of the Pacific calico scallop, *Argopecten ventricosus*. Latin American Journal of Aquatic Research, 45: 915-921. doi: 10.3856/vol45-issue5-fulltext-6

Medina, B.H., Guzmán, M. & Mair, J.M. 2007. Failed recovery of a collapsed scallop fishery in Las Perlas Archipelago, Panama. Journal of Shellfish Research, 26: 9-15. doi: 10.2983/0730-8000(2007)26[9:FROACS]2.0.CO;2

Mendoza-Maldonado, G.B., Diarte-Plata, G., Luna-González, A., Escamilla-Montes, R., Rodríguez-Jaramillo, C., Ruiz-Verdugo, C.A., et al. 2018. Susceptibility and immune response of Catarina scallop (*Argopecten ventricosus*) against *Vibrio parahaemolyticus*, the causative agent of AHPND in shrimp. Aquaculture, 482: 24-30. doi: 10.1016/j.aquaculture.2017.09.018

Meng, X., Liu, M., Jiang, K., Wang, B. & Tian, X. 2013. De novo characterization of Japanese scallop *Mizuhopecten yessoensis* transcriptome and analysis of its gene expression following cadmium exposure. Plos One, 8: e64485. doi: 10.1371/journal.pone.0064485

Mitta, G., Vandenbulcke, F., Hubert, F., Salzet, M. & Roch, P. 2000. Involvement of mytilins in mussel antimicrobial defense. Journal of Biological Chemistry, 275: 12954-12962. doi: 10.1074/jbc.275.17.12954

Moreira, R., Balseiro, P., Planas, J., Fuste, B., Beltrán, S. & Novoa, B. 2012. Transcriptomics of *in vitro* immune-stimulated hemocytes from the Manila clam *Ruditapes philippinarum* using high-throughput sequencing. Plos One, 7: e35009. doi: 10.1371/journal.pone.0035009

Moreira, R., Pereiro, P., Canchaya, C., Posada, D., Figueras, A. & Novoa, B. 2015. RNAseq in *Mytilus galloprovincialis*: comparative transcriptomics and expression profiles among different tissues. BMC Genome, 16: 728. doi: 10.1186/s12864-015-1817-5

Multhoff, G. 2007. Heat shock protein 70 (Hsp70): membrane location, export and immunological relevance. Methods, 43: 229-237. doi: 10.1016/j.ymeth.2007.06.006

Mu, C., Chen, L., Zhao, L. & Wang, C. 2014. Molecular cloning and expression of a C-type lectin gene from *Venerupis philippinarum*. Molecular Biology Reports, 41: 139-144. doi: 10.1007/s11033-013-2846-2

Niu, D., Xie, S., Bai, Z., Wang, L., Jin, K. & Li, J. 2014. Identification, expression, and responses to bacterial challenge of the cathepsin C gene from the razor clam *Sinonovacula constricta*. Developmental & Comparative Immunology, 46: 241-245. doi: 10.1016/j.dci.2014.04.012

Patnaik, B.B., Wang, T.H., Kang, S.W., Hwang, H.J., Park, S.Y. & Park, E.B. 2016. Sequencing, de novo assembly, and annotation of the transcriptome of the endangered freshwater pearl bivalve, *Cristaria plicata*, provides novel insights into functional genes and marker discovery. Plos One, 11: e0148622. doi: 10.1371/journal.pone.0148622

Philipp, E., Kraemer, L., Melzner, F., Poustka, A.J., Thieme, S. & Findeisen, U. 2012. Massively parallel RNA sequencing identifies a complex immune gene repertoire in the lophotrochozoan *Mytilus edulis*. Plos One, 7: e33091. doi: 10.1371/journal.pone.0033091

Ruiz-Verdugo, C.A., Koch, V., Félix-Pico, E., Beltrán-Lugo, A.I., Cáceres-Martínez, C., Mazón-Suástegui, J.M., et al. 2016. Scallop fisheries and aquaculture in

Mexico. Developments in Aquaculture and Fisheries Science, 40: 1111-1123.

Schmieder, R. & Edwards, R. 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics, 27: 863-864. doi: 10.1093/bioinformatics/btr026

Spagnolo, P., Sato, H., Marshall, S.E., Antoniou, K.M., Ahmad, T., Wells, A.U., et al. 2007. Association between heat shock protein 70/hom genetic polymerphisms and uveitis in patients with sarcoidosis. Investigative Ophthamology & Visual Sicence, 48: 3019-3025. doi: 10.1167/iovs.06-1485

Song, L., Wang, L., Zhang, H. & Wang, M. 2015. The immune system and its modulation mechanism in scallop. Fish and Shellfish Immunology, 46: 65-78. doi: 10.1016/j.fsi.2015.03.013

Song, X., Zhang, H., Wang, L., Zhao, J., Mu, C., Song, L., et al. 2011. A galectin with quadruple-domain from bay scallop *Argopecten irradians* is involved in innate immune response. Developmental & Comparative Immunology, 35: 592-602. doi: 10.1016/j.dci.2011.01.006

Tanguy, M., Mckenna, P., Gauthier-Clerc, S., Pellerin, J., Danger, J.M. & Siah, A. 2013. Sequence analysis of a normalized cDNA library of *Mytilus edulis* hemocytes exposed to *Vibrio splendidus* LGP32 strain. Research in Immunology, 3: 40-50. doi: 10.1016/j.rinim.2013.04.001

Tatematsu, M., Funami, K., Tsukasa, T. & Matsumoto, M. 2018. Extracellular RNA sensing by pattern recognition receptors. Journal of Innate Immunity, 10: 398-406. doi: 10.1159/000494034

Teaniniuraitemoana, V., Huvet, A., Levy, P., Klopp, C., Lhuillier, E., Gaertner-Mazouni, N., Gueguen, Y. & Le Moullac, G. 2014. Gonad transcriptome analysis of pearl oyster *Pinctada margaritifera*: identification of potential sex differentiation and sex determining genes. BMC Genome, 15: 491. doi: 10.1186/1471-2164-15-491

Trevisan, R., Arl, M., Sacchet, C.L., Engel, C.S., Danielli, N.M. & Mello, D.F. 2012. Antioxidant deficit in gills of Pacific oyster (*Crassostrea gigas*) exposed to chlorodinitrobenzene increases menadione toxicity. Aquatic Toxicology, 108: 85-93. doi: 10.1016/j.aquatox.2011.09.023

Received: March 10, 2021; Accepted: January 11, 2022

Tong, Y., Zhang, Y., Huang, J., Xiao, S., Zhang, Y. & Li, J. 2015. Transcriptomics analysis of *Crassostrea hongkongensis* for the discovery of reproduction-related genes. Plos One, 10: e0134280. doi: 10.1371/journal.pone.0134280

Turk, V., Stoka, V., Vasiljeva, O., Renko, M., Sun, T. & Turk, B. 2012. Cysteine cathepsins: from structure, function and regulation to new frontiers. Biochimica et Biophysica Acta, 1824: 68-88. doi: 10.1016/j.bbapap.2011.10.002

Underhill, D.M. & Goodridge, H.S. 2012. Information processing during phagocytosis. Nature Reviews Immunology, 12: 492-502. doi: 10.1038/nri3244

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B. & Van Roy, N. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biology, 3: 1-12. doi: 10.1186/gb-2002-3-7-research 0034

Xue, Q., Hellberg, M.E., Schey, K., Itoh, N., Eytan, R.I., Cooper, R.K. & La Peyre, J.F. 2010. A new lysozyme from the eastern oyster, *Crassostrea virginica*, and a possible evolutionary pathway for i-type lysozymes in bivalves from host defense to digestion. BMC Evolutionary Biology, 10: 213. doi: 10.1186/1471-2148-10-213

Yang. J., Wang, W., Wei, X., Qiu, L., Wang, L., Zhang, H. & Song, L. 2010. Peptidoglycan recognition protein of *Chlamys farreri* (*Cf*PGRP-S1) mediates immune defenses against bacterial infection. Developmental & Comparative Immunology, 34: 1300-1307. doi: 10.1016/j.dci.2010.08.006

Xie, F., Xiao, P., Chen, D., Xu, L. & Zhang, B. 2012. miRDeepFinder: a miRNA analysis tool for deep sequencing of plant small RNAs. Plant Molecular Biology, 80: 75-84. doi: 10.1007/s11103-012-9885-2

Zhang, L., Li, L., Zhu, Y., Zhang, G. & Guo, X. 2014. Transcriptome analysis reveals a rich gene set related to innate immunity in the Eastern oyster (*Crassostrea virginica*). Marine Biotechnology, 16: 17-33. doi: 10.1007/s10126-013-9526-z

Zheng, L.P. 2011. Expression pattern of a gram-negative bacteria-binding protein in early embryonic development of *Artemia sinica* and after bacterial challenge. Developmental & Comparative Immunology, 35: 35-43. doi: 10.1016/j.dci.2010.08.002