RNA-seq analysis of *Macrobrachium rosenbergii* hepatopancreas in response to *Vibrio parahaemolyticus* infection

RAMARAO SERIRAMALU AND SUBHA BHASSU
Animal Genomics and Evolutionary Biology, Genetics and Molecular Biology and Institute of Biological Sciences, Faculty of Science and CEBAR, University Malaya, 50603 Kuala Lumpur.

INTRODUCTION

The Malaysian giant freshwater prawn, *Macrobrachium rosenbergii*, an economically important crustacean worldwide are being affected by Vibriosis, a disease caused by *Vibrio* strains such as *Vibrio parahaemolyticus*. *M. rosenbergii* possesses an innate immune system which provides defence against pathogenic agents. The information regarding the regulation of innate immune system in this species is lacking, thus its necessary in providing solutions to control and minimize the loss of production due to this bacterial disease. In this study, we performed a transcriptome profiling of *M. rosenbergii* hepatopancreas infected with *V. parahaemolyticus* using the 'Next Generation' sequencing method (Illumina HiSeq^{TM}2000). A total of 54,295,342 and 54,708,014 high-quality reads obtained from *Vibrio*-infected and control *M. rosenbergii* cDNA libraries. The overall *de novo* assembly and clustering of both reads generated 64,411 unigenes, with an average length of 698 bp. Based on BLASTX search (E-value <10^{-5}) against NR, Swissprot, GO, COG and KEGG databases, 22,455 unigenes (34.86% of all unigenes) were annotated with gene descriptions, gene ontology terms, and metabolic pathways. The unigene differential expression analysis revealed 14,569 unigenes were differentially expressed in the infected shrimp compared to the controls. Several differentially expressed genes are involved in various animal immune functions. The large number of transcripts obtained in this study would provide valuable resources for further genomic research into freshwater prawns.

OBJECTIVES

- To perform a transcriptome profiling of the *M. rosenbergii* hepatopancreas infected with *V.parahaemolyticus*, using Illumina HiSeq^{TM} 2000 platform.
- To discover and determine the role of the immune genes involved in *V.parahaemolyticus* infection.

MATERIALS & METHODS

Wet lab
- Challenge test: One *V.parahaemolyticus* challenge group and one negative control group (10 prawns each group) prepared
- At 12 hours post-infection, the hepatopancreas tissues of prawns were dissected and immediately frozen in liquid nitrogen

NGS
- Total RNA extraction
- Transcriptome sequencing using the Illumina HiSeq^{TM} 2000, BGI Shenzhen, China

Dry Lab
- *De novo* assembly using the Trinity program
- Clustering of both unigenes done using TIGR Gene Indices clustering tools (TGICL)
- Annotation using BLASTX against NCBI Nr, Swissprot, COG and KEGG
- Gene Ontology (GO) assignment conducted using BLAST2GO software
- Differential gene expression analysis done using the FPKM approach

DISCUSSION/ CONCLUSION

- This study generates significant information on *M.rosenbergii* immune genes activity during *V.parahaemolyticus* infection.
- Several differentially expressed genes identified are involved in various animal immune functions, such as antimicrobial, proteases and protease inhibitors, pattern recognition proteins, heat shock proteins, cell death, oxidative stress, blood clotting system, and prophenoloxidase system.
- Results obtained provided a valuable insight into antibacterial mechanisms in *M.rosenbergii* and the role of the differentially expressed immune genes in response to *V. parahaemolyticus* infection.

RESULTS

	Control	*V.parahaemolyticus* Infected
Total number of reads	54,708,014	54,295,342
Total base pairs (bp)	4,923,721,260	4,886,580,780
Q20 value	97.73%	97.77%
Total number of contigs	95,645	123,141
Mean length of contigs	313	318
Total number of unigenes	59,050	73,946
Mean length of unigenes	479	532
NCBI Nr annotated	19,799	
Swiss-Prot annotated	16,832	
KEGG annotated	14,706	
COG annotated	7,856	
GO annotated	6,007	

Table 1. Summary of the control and infected transcriptome sequencing

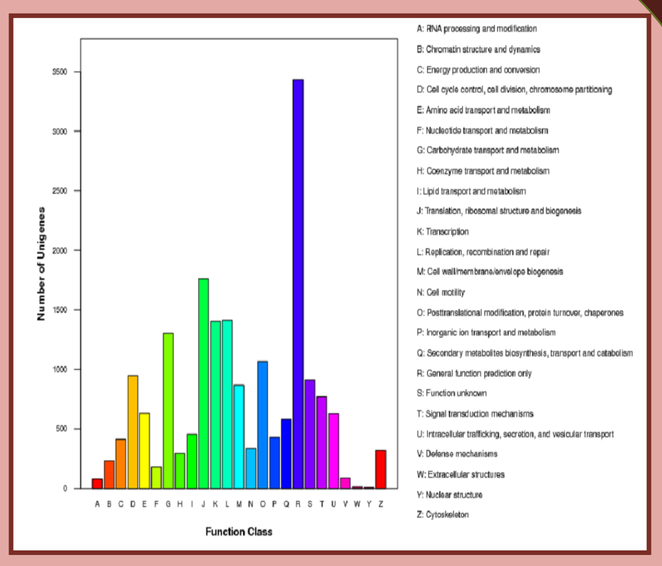

Figure 1. Histogram presentation of Cluster of Orthologus Groups (COG) classification of 7,856 known protein annotated unigenes. Each bar represents the number of unigenes classified into each of the 25 COG functional categories

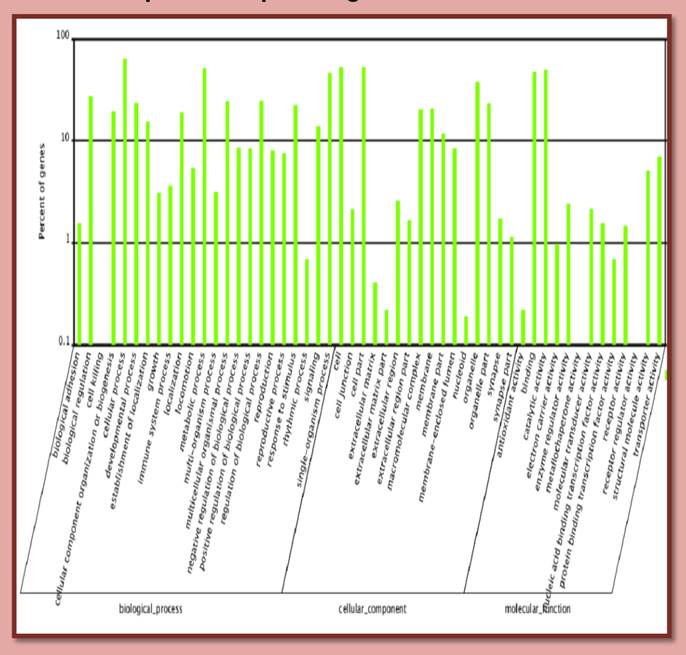

Figure 2. Gene ontology (GO) classification of the 6,007 protein annotated unigenes. Unigenes sequences were systematically classified into GO sub-categories under the Biological Process, Cellular Component and Molecular Function Gene Ontology Catalogue system. Each bar represents the relative abundance of unigenes classified under each sub-category

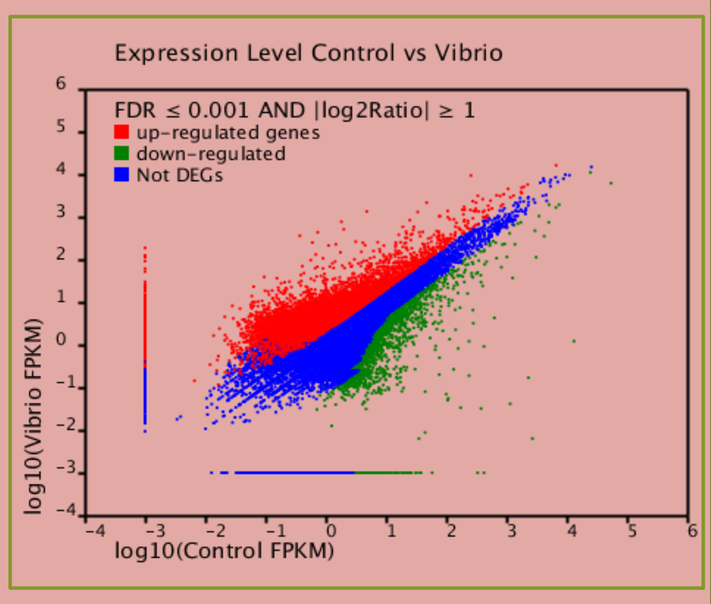

Figure 3. Digital gene expression between control group against *V.parahaemolyticus* infected group. Each point represents a unigene. The x- and y-axis are the log_{10} of the normalized expression level (FPKM) of unigene between the two groups. Red and green points indicate significant change at the absolute value of log_2 (FPKM ratio in two groups) ≥1 and fdr =0.001. Red points indicate up-regulated unigenes and green points indicate down-regulated unigenes in the two groups which its expression level is represented by the y-axis. Blue points indicate insignificant differentially expressed unigenes

Category or gene id	Homologues function	Species	FC*
Antimicrobial			
Unigene4120_All	Anti-lipopolysaccharide factor	*Macrobrachium rosenbergii*	4.13
Unigene37309_All	Crustin	*Macrobrachium rosenbergii*	5.13
Blood Clotting system			
Unigene13048_All	Clottable protein	*Marsupenaeus japonicus*	1.42
Unigene34308_All	Transglutaminase	*Macrobrachium rosenbergii*	11.34
PRPs			
Unigene10978_All	Lectin 1	*Macrobrachium rosenbergii*	4.4
Unigene23671_All	lipopolysaccharide and beta-1,3-glucan binding protein	*Macrobrachium rosenbergii*	1.71
Proteinases and Proteinases inhibitors			
Unigene12869_All	Alpha-2-macroglobulin	*Macrobrachium rosenbergii*	5.09
CL2365.Contig1_All	Caspase	*Marsupenaeus japonicus*	1.04
Unigene26757_All	Hemocyte kazal-type proteinase inhibitor	*Penaeus monodon*	3.87
CL1127.Contig2_All	Kazal-type serine proteinase inhibitor 4	*Procambarus clarkii*	-6.01
Heat shock Proteins			
Unigene3736_All	Heat shock protein 21	*Macrobrachium rosenbergii*	4.61
Unigene23034_All	Heat shock protein 40	*Frankliniella occidentalis*	1.8
Oxidative stress			
CL353.Contig2_All	Glutathione S transferase	*Procambarus clarkii*	1.01
CL2477.Contig1	Catalase	*Litopenaeus vannamei*	1.6
ProPO system			
Unigene12734_All	Prophenoloxidase	*Macrobrachium rosenbergii*	3.15
Unigene7353_All	Prophenoloxidase activating factor	*Fenneropenaeus chinensis*	3.79
Cell death			
Unigene2555_All	Beclin 1	*Megachile rotundata*	11.89
Unigene14045_All	Program cell death 5-like	*Penaeus monodon*	1.61

Table 2. Selected candidate genes involved in *M.rosenbergii* immune response. *Fold changes (Log_2 ratio) in gene expression, PRPs- pattern recognition proteins, ProPO- prophenoloxidase

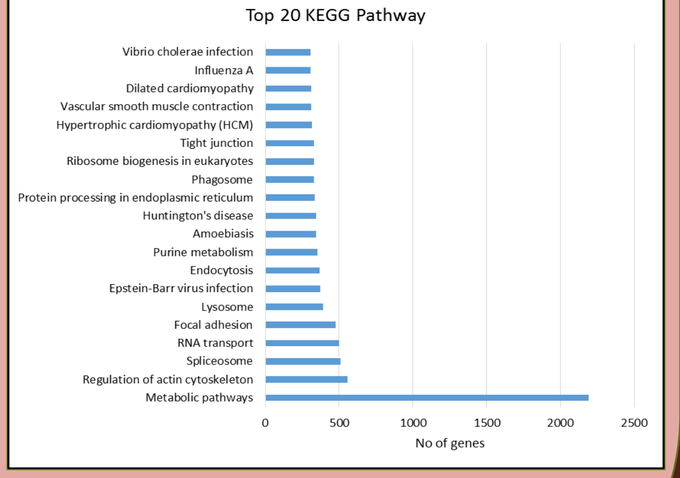

Figure 4. Top 20 KEGG biological pathway classification histograms for annotated unigenes

ACKNOWLEDGEMENT

•I thank my supervisor Assoc Prof Dr. Subha Bhassu and my fellow lab members for their guidance and motivation.
•This work was supported by a Flagship Grant No: FL002-201 granted by Cluster-Bio, University of Malaya and Postgraduate Research Grant (PPP) of the University of Malaya, Malaysia (PG088-2012B)

REFERENCES

- New MB: History and global status of freshwater prawn farming. Freshwater Prawns: Biology and Farming 2010:1-11.
- Tonguthai K: Diseases of the freshwater prawn, *Macrobrachium rosenbergii*, The Aquat. Anim Health Res Inst Newsletter 1995, 4:1-4
- Khuntia CP, Das BK, Samantaray BR, Samal SK, Mishra BK: Characterization and pathogenicity studies of *Vibrio parahaemolyticus* isolated from diseased freshwater prawn, *Macrobrachium rosenbergii* (de Man). Aquaculture research 2008, 39(3):301-310.
- Morozova O, Hirst M, Marra MA: Applications of new sequencing technologies for transcriptome analysis. Annual review of genomics and human genetics 2009, 10:135-151.