# IDENTIFICATION AND CHARACTERIZATION OF MICRO-RNA IN *MACROBRACHIUM ROSENBERGII*

TAN TIAN TIAN

DISSERTATION SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

INSTITUTE OF BIOLOGICAL SCIENCES
FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2012

**UNIVERSITI MALAYA**

**ORIGINAL LITERARY WORK DECLARATION**

Name of Candidate: Tan Tian Tian (I.C/Passport No: 850611-01-6994)

Registration/Matric No: SGR090134

Name of Degree: Master of Science

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):
IDENTIFICATION AND CHARACTERIZATION OF MICRO-RNA IN *MACROBRACHIUM ROSENBERGII*

Field of Study: Genetics and Molecular Biology

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any Copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date

Subscribed and solemnly declared before,

Witness's Signature Date

Name:
Designation:

## ABSTRACT

MicroRNAs are ~20-22nt non protein-coding RNA regulatory genes that post-transcriptionally regulate many protein–coding genes to influence critical biological and metabolic processes. While the number of known microRNA is increasing, there is no published data for microRNA in freshwater prawns (*Macrobrachium rosenbergii*) a commercially cultured and economically important food species. Through deep parallel sequencing and an *in silico* data analysis approach, 327 miRNA families have been identified from small RNA libraries with reference to both the *de novo* transcriptome of *M.rosenbergii* obtained from RNA-Seq and to miRBase (Release 18.0, November 2011). Based on the identified mature miRNA and recovered precursor sequences that form appropriate hairpin structures, three conserved miRNA (miR125, miR750, miR993) and 27 novel miRNA candidates encoding messenger-like non-coding RNA have been identified. miR-125, miR-750, G-m0002/H-m0009, G-m0005, G-m0008/H-m0016, G-m0011/H-m0027 and G-m0015 were selected for experimental validation with stem-loop quantitative RT-PCR and were found coherent with the expression profile of deep sequencing data as evaluated with Pearson's correlation coefficient ($r = 0.835178$ for miRNA in gill, $r = 0.724131$ for miRNA in hepatopancreas). Using a combinatorial approach of pathway enrichment analysis and inverse expression relationship of miRNA and mRNA, four co-expressed novel miRNA candidates (G-m0005, G-m0008/H-m0016, G-m0011/H-m0027, and G-m0015) were found to be associated with energy metabolism. In addition, the expression of the three novel miRNA candidates (G-m0005, G-m0008/H-m0016, and G-m0011/H-m0027) were also found to be significantly reduced at the 9 and 24 hours post infection in a controlled experiment of *M.rosenbergii* challenged with infectious hypodermal and haematopoietic necrosis virus. These findings provide a reference point to further improving the understanding of the repertoire of crustacean miRNAs.

# ABSTRAK

MikroRNAs adalah ~ 20-22 nt bukan protein-pengekodan RNA gen pengawalseliaan yang mengawal selia pengekodan gen protein pada peringkat post-transkripsi yang banyak mempengaruhi proses biologi dan metabolisme yang kritikal. Walaupun, bilangan mikroRNA dikenali semakin meningkat, tidak ada data yang diterbitkan bagi mikroRNA dari udang air tawar (*Macrobrachium rosenbergii*) yang merupakan makanan komersial yang penting dari segi ekonomi. Dengan mengunakan pendekatan jujukan dalam selari dan analisis data secara siliko, 327 keluarga miRNA telah dikenal pasti dari perpustakaan kecil RNA dengan merujuk kepada transkripsi data, *M.rosenbergii* yang dihimpun secara *de novo* yang diperolehi dari RNA-seq dan miRBase (Keluaran 18,0, November 2011). Berdasarkan miRNA matang yang dikenal pasti dan pelopor yang membentuk struktur jepit rambut yang sesuai, tiga miRNA konserve (miR125, miR750, miR993) dan 27 calon miRNA novel pengekodan mesej seperti bukan pengekodan RNA telah dikenal pasti. miR-125, miR-750, G-m0002/H-m0009, G-m0005, G-m0008/H-m0016, G-m0011/H-m0027 dan G-m0015 telah dipilih untuk pengesahan uji kaji dengan ranting-gelung kuantitatif RT -PCR dan ditemui koheren dengan profil ungkapan data penjujukan yang dalam seperti yang dinilai dengan pekali korelasi Pearson (r = 0,835178 untuk miRNA di insang, r = 0,724131 untuk miRNA dalam hepatopancreas). Menggunakan pendekatan kombinatorik analisis pengayaan laluan dan hubungan ungkapan songsang miRNA dan mRNA, empat calon miRNA novel (G-m0005, G-m0008/H-m0016, G-m0011/H-m0027, dan G-m0015) ditemui berkaitan dengan metabolisme tenaga. Di samping itu, ungkapan untuk tiga calon miRNA novel (G-m0005, G-m0008/H-m0016, dan G-m0011/H-m0027) juga didapati semakin berkurangan pada 9 dan 24 jam selepas jangkitan dalam eksperimen terkawal *M.rosenbergii* yang dicabar dengan virus nekrosis berjangkit hypodermal dan haematopoietic. Penemuan ini menyediakan titik rujukan untuk meningkatkan lagi pemahaman mengenai himpunan miRNAs krustacean.

## ACKNOWLEDGEMENTS

I am sincerely grateful to the people who have assisted me directly and indirectly, to complete my dissertation successfully.

I would like to express my deep appreciation to my supervisors, Professor Dr Jennifer Ann Harikrishna and Dr Subha Bhassu. Their constant encouragement, guidance, willingness to listen, discuss my results with me and provided helpful suggestion when I had problems or unexpected result as well as to share their times will always be greatly appreciated.

Beijing Genomic Institute (BGI), Shenzhen, China and Malaysia Palm Oil Board (MPOB) are acknowledged for granting me permission to use their respective bioinformatics facilities and dot-blot facilities. I would like to thanks Dr Zhang Guojie (BGI) for conceptual advice, Mr Chen Maoshan (BGI) for assistance in bioinformatics, Dr Rajinder Singh (MPOB) for permission to use chemical, consumable and equipment in his laboratory and Ms Leslie Ooi Cheng Li (MPOB) for technical assistance and advice in carrying out the miRNA dot-blot experiment.

This study was supported by grants from the Ministry of Science, Technology and Innovation, Malaysia (MOSTI Grant No: 53-02-03-1030, granted to Dr Subha Bhassu) and from University of Malaya (IPPP Grant No: PS260/2010A, granted to Ms Tan Tian Tian).

Thanks to my colleagues-laboratory mates for their generous assistance and patience throughout my research project and many thanks to my friends and family who always led me out of the darkness whenever I stumble as well as for their invaluable support and their belief in me throughout my life.

## TABLE OF CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

## LIST OF ABBREVIATIONS AND SYMBOLS

| | |
|---|---|
| **B** | |
| bp | Base pair |
| **C** | |
| cDNA | Complementary deoxyribonucleic acid |
| $^{o}$C | Degree Celsius |
| **D** | |
| DEPC | Diethyl pyrocarbonate |
| DNA | Deoxyribonucleic acid |
| DNase | Deoxyribonuclease |
| dNTP | Deoxynucleoside triphosphate |
| **E** | |
| e.g. | Example |
| **G** | |
| G | Gravity |
| H | |
| hpi | Hours post-injection |
| **I** | |
| i.e. | That is |
| **M** | |
| M | Molar |
| mM | Miliampere |
| mg | Miligram |

| | |
|---|---|
| ml | Mililiter |
| mM | Milimolar |
| **N** | |
| ng | Nanogram |
| **P** | |
| PCR | Polymerase chain reaction |
| Pmol | Picomole |
| % | Percentage |
| **R** | |
| RNA | Ribonucleic acid |
| RT-PCR | Reverse transcription polymerase chain reaction |
| **T** | |
| *Taq* | *Thermus aquaticus* |
| **U** | |
| µl | Microliter |
| USA | United States of America |
| **V** | |
| V | Voltage |

## LIST OF APPENDICES