UNIVERSITI PUTRA MALAYSIA

ISOLATION AND CHARACTERIZATION OF MICROSATELLITE LOCI FROM THE GIANT FRESHWATER PRAWN (*MACROBRACHIUM ROSENBERGII*)

SEE LENG MIN

FS 2008 44

ISOLATION AND CHARACTERIZATION OF MICROSATELLITE LOCI FROM THE GIANT FRESHWATER PRAWN (*MACROBRACHIUM ROSENBERGII*)

By

SEE LENG MIN

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirement for the Degree of Master of Science

August 2008

UPM

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Master of Science

ISOLATION AND CHARACTERIZATION OF MICROSATELLITE LOCI FROM THE GIANT FRESHWATER PRAWN (*MACROBRACHIUM ROSENBERGII*)

By

SEE LENG MIN

September 2008

Chairman: Professor Tan Soon Guan, PhD

Faculty: Science

The giant freshwater prawn, *Macrobrachium rosenbergii* , or locally know as udang galah has become the most popular freshwater prawn for commercial culture and a significant cash crop for many poor farmers throughout Asia and the Pacific region. Ninety microsatellite repeat sequences were successfully isolated from *M. rosenbergii* using the 5' anchored-PCR technique. BLAST analysis of the microsatellite marker flanking regions showed similarities towards expressed sequence tags (ESTs) in aquatic species. Sixty-two microsatellite primer pairs were designed with 29 perfect microsatellites, four were imperfect or interrupted microsatellites and the rest were compound microsatellites. Of these 62 single locus DNA microsatellite markers, 24 showed polymorphisms in the giant freshwater prawns of which four loci had dinucleotide, 16 trinucleotide, three tetranucleotide and one pentanucleotide core repeat units. Nine microsatellite primer pairs from the

green-lipped mussel (*Perna viridis*) were successful in cross-amplifying the giant freshwater prawn genome. However, only four of these cross-amplified microsatellite primer pairs were reliable and used in this study. Hence, the levels of genetic variability in 12 populations of wild *M. rosenbergii* and one cultured population in Malaysia were evaluated by using 28 microsatellite loci. The number of alleles per locus ranged from 2 to 26 and the total observed heterozygosity ranged from 0.2618 to 0.7265. A high level of polymorphism was also detected in each of the wild *M. rosenbergii* populations by using five RAPD and four LP-RAPD primers which generated 191 bands ranging in molecular weights from 150 bp to 2100 bp in 11 populations. The cross-amplifications of 32 of the 47 newly developed microsatellite primer pairs in nine other prawn species showed the presence of many highly conserved regions among the prawn species tested. However, some of the microsatellite motifs in the nine species tested differed slightly from the originally designed microsatellite loci for *M. rosenbergii*. These newly developed microsatellite loci were used to assess the genetic diversity and relationships of eleven wild *M. rosenbergii* stocks.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai keperluan ijazah Sarjana Sains

PEMENCILAN DAN PENCIRIAN LOKUS MICROSATELIT DARIPADA UDANG GALAH (*MACROBRACHIUM ROSENBERGII*)

Oleh

SEE LENG MIN

September 2008

Pengerusi: Professor Tan Soon Guan, PhD

Fakulti: Sains

Udang galah, *Macrobrachium rosenbergii* merupakan udang air tawar yang paling popular diternak secara komersial dan menjadi sumber pendapatan penting bagi kebanyakan penternak miskin. Sembilan puluh jujukan berulang mikrosatelit telah berjaya dipencilkan daripada spesies udang galah ini melalui teknik *5' anchored PCR*. Analisis *BLAST* penanda mikrosatelit menunjukkan kesamaan dengan tanda jujukan terungkap (*expressed sequence tags*) dalam spesies akuatik. Enam puluh dua pasang primer mikrosatelit telah direka, di mana 29 adalah mikrosatelit sempurna sementara empat darinya adalah mikrosatelit tidak sempurna atau terganggu dan yang selebihnya adalah mikrosatelit sebatian. Daripada 62 penanda mikrosatelit, 24 pasang primer menunjukkan polimorfik dalam mengamplifikasikan genom udang galah, dengan empat lokus adalah dinukleotid, l6 adalah trinukleotid, tiga adalah tetranukleotid dan satu adalah pentanukleotid. Sembilan pasang primer mikrosatelit daripada kepah (*Perna viridis*) telah berjaya mengamplifikasi-rentas genom udang

galah. Walau bagaimanapun, hanya empat pasang primer mikrosatelit yang digunakan dalam kajian udang galah ini. Oleh itu, variasi genetik dalam 12 populatsi spesies liar dan satu populasi ternakan *M. rosenbergii* telah dinilai dengan menggunakan 28 pasang primer mikrosatelit. Bilangan alel per lokus berjulat antara 2 hingga 26 dan jumlah heterozigositi cerapan berjulat antara 0.2618 hingga 0.7265. Polimorfik yang tinggi telah dikesan dalam setiap populasi liar *M. rosenbergii* dengan menggunakan lima primer RAPD dan empat primer LP-RAPD dan menghasilkan 191 jumlah jalur yang berat molekul berjulat antara 150 bp hingga 2100 bp dalam kesemua 11 populasi. Pengujian amplifikasi-rentas 32 pasang primer daripada 47 pasang primer yang baru direka dalam sembilan jenis spesies udang menunjukkan terdapat kehadiran kawasan terpulihara yang tinggi di kalangan spesies udang galah tersebut. Walau bagaimanapun, sebahagian motif mikrosatelilt dalam sembilan spesies udang adalah berbeza daripada lokus mikrosatelit yang direka untuk *M. rosenbergii*. Lokus mikrosatelit yang baru direka telah digunakan untuk mengkaji diversiti genetik dan perhubungan antara 11 populasi liar *M. rosenbergii*.

ACKNOWLEDGEMENTS

The core of research, besides having a good background of knowledge and a keen sense of curiosity, is about accepting successes and failures. The past two years of conducting this study have truly exposed me to obstacles that I believe, faced by every researcher who is out there. Living through these challenges, I now understand that no obstacle in life is hardy enough to resist a brave character and a valiant heart but most importantly, the strength of the human spirit. This dissertation is thus not just a product of my work alone, for many people have guided and supported me during the times when I needed them most.

My first appreciation definitely goes to my supervisor, Prof. Dr. Tan Soon Guan, Dr. Subha Bhassu and Assoc. Prof. Dr. Siti Shapor Siraj, who have always been there for me, guiding and supporting me all the way. I'm greatly indebted to each of them for their continuous support, cooperation and interest in my research and their comments and suggestions on many aspects of this work.

I would like to gratefully acknowledge the financial support from the Department of Fisheries, Malaysia project grant no. 22501-006 that allowed for the smooth running of the project. I am also deeply grateful to acknowledge the Yayasan Felda for providing me Felda Scholarship and The Southeast Asian Ministers of Education Organization (SEAMEO) for providing me with a Southeast Asian Regional Center for Graduate Study and Research in Agriculture (SEARCA) Scholarship.

Another word of thanks to Assoc. Prof. Dr. Siti Khalijah Daud who has been very kind in letting me using the instruments that I needed in the Tissue Culture Laboratory. My next appreciation goes to Mr. Azmi for helping me during my work in the Genetic Laboratory. Another word of thanks to Mrs. Mazlina and Mr. Rashidi who have been very patient in showing me the whereabouts of chemicals and apparatus that I needed in the Genetic Laboratory of Freshwater Fisheries Research Center, Jelebu.

My gratitude also goes out to Yanty, Nurul Izza, Ong Chin Chin, Yuzine bin Esa and Hoh Boon Peng who have revealed to me so much about the details of their research on population studies. I would like to thank my undergraduate students, Chew Mei Yee, Ng Lee Tin, Wan Ying Ying, Tan Shin Yee, Annie Wong, Foo Pei Boon, Lee Sook Jen, Liew Pui Ling, Mah Kah Phoon, Lee Ley Ching, Zheng Xiang and Kaviarasu, who have helped me with the giant freshwater prawn project and also made teaching an exciting experience for me.

I would not be here today without my parents, who have always given me full support in everything I do and still are. Thumbs up for my sisters, Leng Ling and Leng Li and my brother, Choon Wai for their never failing sense of optimism in me. Special thanks to my previous educators who have molded me into a better person. For my best friends, Kathy Ng Siang Chiing and Christine Cheryl Fernandez, thank you for believing in me and showing me that the sun will shine brighter after a heavy storm. My thanks also go to my coursemates, Tee Meng Han and Shafiq for their advices, friendship and the many interesting discussions we had on microsatellites

and data analysis. Without them, life in the laboratory would indeed be boring and uneventful.

TABLE OF CONTENTS

		Page
ABSTRACT		ii
ABSTRAK		iv
ACKNOWLEDGEMENT		vi
APPROVAL SHEETS		ix
DECLARATION FORM		xi
LIST OF TABLES		xv
LIST OF FIGURES		xvii
LIST OF ABBREVIATIONS		xx
CHAPTER		
1	**INTRODUCTION**	1
	1.1 Objective	3
2	**LITERATURE REVIEW**	4
	2.1 Taxonomic History of *Macrobrachium rosenbergii*	4
	2.2 The Giant Freshwater Prawns	6
	2.2.1 Nomenclature of *Macrobrachium rosenbergii*	8
	2.2.2 Subspecies of *Macrobrachium rosenbergii*	9
	2.3 *Macrobrachium rosenbergii* in Malaysia	9
	2.4 Shrimp species	11
	2.4.1 Suborder Dendrobranchiata	11
	2.4.2 Suborder Pleocyemata	12
	2.5 Microsatellites	13
	2.5.1 Application of Microsatellites	15
	2.5.2 Cross-amplification of Microsatellites	19
	2.6 Random Amplified Polymorphic DNA (RAPD)	19
	2.7 Long Primer RAPD (LP-RAPD)	21
	2.8 Random Amplified Microsatellites (RAMs)	22
3	**ISOLATION OF MICROSATELLITE LOCI**	24
	3.1 Introduction	24
	3.2 Methodology	25
	3.2.1 5' Anchored PCR Amplification	26
	3.2.2 Cloning	29
	3.2.3 Plasmid Extraction	30
	3.2.4 DNA Sequencing	31
	3.2.5 Submission of DNA Sequences to GenBank	31
	3.2.6 Designing Primers Flanking Microsatellite Regions	32
	3.2.7 Microsatellites Amplification	32
	3.2.8 BLAST Analyse of Microsatellite Marker	34
	3.3 Results	34
	3.3.1 Isolation of Microsatellites using 5' Anchored PCR	34
	3.3.2 Microsatellites Amplification	35
	3.3.3 BLAST Analyse of Microsatellite Marker Flanking Regions	47

	3.4	Discussion	49
	3.5	Conclusions	54
4	**CROSS-AMPLIFICATION OF *PERNA VIRIDIS* MICROSATELLITE IN *MACROBRACHIUM ROSENBERGII***		55
	4.1	Introduction	55
	4.2	Methodology	57
		4.2.1 Samples	58
		4.2.2 Screening for Cross-amplification	58
		4.2.3 Determination of Microsatellite Loci Amplification	59
	4.3	Results	61
		4.3.1 PCR Amplification	61
		4.3.2 Screening for Cross-amplification	61
	4.4	Discussion	67
	4.5	Conclusions	70
5	**GENETICS DIVERSITY STUDY OF *MACROBRACHIUM ROSENBERGII***		71
	5.1	Introduction	71
	5.2	Methodology	74
		5.2.1 Microsatellite Genotyping	74
		5.2.2 Cross-amplified Microsatellites Analysis	79
		5.2.3 RAPD and LP-RAPD Analysis	81
	5.3	Results	83
		5.3.1 Microsatellite Genotyping	84
		5.3.2 RAPD and LP-RAPD Analysis	113
	5.4	Discussion	125
	5.5	Conclusions	135
6	**CROSS-AMPLIFICATION OF *M. ROSENBERGII* MICROSATELLITES IN OTHER SPECIES**		137
	6.1	Introduction	137
	6.2	Methodology	139
		6.2.1 Samples	140
		6.2.2 Primers and Screening for Cross-Amplification	141
		6.2.3 PCR Amplification and Gel Electrophoresis	142
		6.2.4 Determination of Microsatellites Loci Amplification	142
	6.3	Results	142
		6.3.1 PCR Amplification	142
		6.3.2 Banding Pattern and Cross-Amplification	143
		6.3.3 Microsatellite Repeat Pattern	146
	6.4	Discussion	150
	6.5	Conclusions	152
7	**DISCUSSION**		153
8	**CONCLUSIONS**		157
REFERENCES			160

APPENDICES 182

BIODATA OF STUDENT 276

LIST OF TABLES

Table		Page
3.1	RAMs primer used for construction of genomic library enriched for microsatellites.	28
3.2	List of microsatellites isolated from *M. rosenbergii*.	37
3.3	Microsatellite loci in *M. rosenbergii*, primers sequences, GenBank accession number, PCR conditions and the expected PCR amplification product size.	40
3.4	Primer pair sequences and characteristics of polymorphic *Macrobrachium rosenbergii* microsatellite loci.	45
3.5	Comparison of nucleotide sequences of *M. rosenbergii* microsatellite flanking sequences with other species	48
4.1	Microsatellite repeat motifs flanked in *M. rosenbergii* by *Perna viridis* microsatellite primer pairs.	68
5.1	Locations of the sampling sites and samples size.	75
5.2	Locations of the sampling sites and samples size for cross-amplification microsatellite analysis.	80
5.3	List of microsatellite loci used for cross-amplification in *M. rosenbergii*.	80
5.4	RAPD and LP-RAPD primers used for PCR amplification.	82
5.5	The estimated null allele frequencies in 13 populations.	89
5.6	List of monomorphic microsatellite loci.	94
5.7	Microsatellite variations in 13 populations of *M. rosenbergii*.	95
5.8	Heterozygosity in 13 populations based on 28 microsatellite loci.	102
5.9	Estimation of exact *P*-value for Hardy-Weinberg equilibrium based on the Markov chain method for 13 populations.	106

Table		Page
5.10	Genetic distance of 13 populations based on 28 polymorphic microsatellite markers.	108
5.11	F_{ST} estimated based on 28 polymorphic microsatellite markers.	114
5.12	Analysis of Molecular Variance (AMOVA).	115
5.13	Overall polymorphism of bands produced by RAPD markers in the 11 population.	121
5.14	Overall polymorphism of bands produced by LP-RAPD markers in the 11 population.	122
5.15	Genetic distance of 11 populations derived from the similarity index of Nei and Li (1979) based on 191 dominant markers produced by five RAPDs.	123
5.16	Genetic distance of 11 populations derived from the similarity index of Nei and Li (1979) based on 191 dominant markers produced by four LP-RAPDs.	124
6.1	Habitat of the nine prawn species.	141
6.2	Cross-amplification of microsatellite markers developed for *M. rosenbergii* in 9 selected prawn species.	145
6.3	Microsatellite repeat pattern comparisons of *M. rosenbergii* with other species of prawns.	148

LIST OF FIGURES

Figure		Page
3.1	Agarose gel (2%) electrophoresis of PCR product obtained using degenerate primer, bp10.	36
3.2	Agarose gel (2%) electrophoresis of PCR product obtained from amplification of *Macrobrachium rosenbergii* using degenerate primers.	36
3.3	Plasmid DNA of 5' anchored PCR clones.	44
4.1	Microsatellite profiles of giant freshwater prawns using primer pair, OCC14.	62
4.2	Microsatellite profiles of giant freshwater prawns using primer pair, OCC18.	62
4.3	Microsatellite profiles of giant freshwater prawns using primer pair, OCC20.	62
4.4	Microsatellite profiles of giant freshwater prawns using primer pair, OCC26.	63
4.5	Microsatellite profiles of giant freshwater prawns using primer pair, OCC28.	63
4.6	Microsatellite profiles of giant freshwater prawns using primer pair, OCC34.	63
4.7	Microsatellite profiles of giant freshwater prawns using primer pair, OCC42.	64
4.8	Microsatellite profiles of giant freshwater prawns using primer pair, OCC43.	64
4.9	Microsatellite profiles of giant freshwater prawns using primer pair, OCC44.	64

Figure		Page
4.10	Microsatellite sequences amplified in *M. rosenbergii* obtained by using an automated DNA sequencer, ABI PRISM 377 for primer pair, OCC34	65
4.11	Microsatellite sequences amplified in *M. rosenbergii* obtained by using an automated DNA sequencer, ABI PRISM 377 for primer pairs, OCC26 and OCC28.	66
5.1	Location of the sampling sites.	76
5.2	Microsatellite profiles of SUGbp8-101b (expected size: 208 bp) showed additional locus.	86
5.3	Microsatellite profiles of SUGbp11-4a (expected size: 207 bp) showed additional locus.	86
5.4	Microsatellite profiles of wild *M. rosenbergii* samples from Sg. Pahang, Sg. Muda and Sg. Serian using primer pair SUGbp8-106c.	87
5.5	Microsatellite profiles of wild *M. rosenbergii* samples from Sg. Muar, Sg. Penarik and Sg. Pahang using primer pair SUGbp8-103a.	87
5.6	Microsatellite profiles of wild *M. rosenbergii* samples from Sg. Linggi and Sg. Endau using primer pair SUGbp8-109a.	88
5.7	Microsatellite profiles of wild *M. rosenbergii* sampoles from Sg. Penarik and Sg. Muda using primer pair SUGbp108-107a.	88
5.8	UPGMA dendrogram generated from Nei's (1978) unbiased genetic distance based on 28 polymorphic microsatellite loci.	110
5.9	Consensus tree generated out of 1000 trees using 28 polymorphic microsatellite loci.	111
5.10	RAPD banding profile of Sg. Pahang samples generated by primer OPA01.	115
5.11	RAPD banding profile of Sg. Linggi samples generated by primer OPA03.	116
5.12	RAPD banding profile of Sg. Kelantan samples generated by primer OPA04.	116

Figure		Page
5.13	RAPD banding profile of Sg. Penarik samples generated by primer OPA09.	117
5.14	RAPD banding profile of Sg. Terengganu samples generated by primer OPA10.	117
5.15	LP-RAPD banding profile of Sg. Perak samples generated by primer PUCM 13F.	118
5.16	LP-RAPD banding profile of Sg. Perak samples generated by primer BOXAIR.	118
5.17	LP-RAPD banding profile of Sg. Endau samples generated by primer PEH A3.	119
5.18	LP-RAPD banding profile of Sg. Endau samples generated by primer PEH A6.	119
5.19	UPGMA dendrogram generated based on 191 markers obtained using RAPD primers.	126
5.20	UPGMA dendrogram generated based on 191 markers obtained using LP-RAPD primers.	126
6.1	Three marine prawn species.	140
6.2	Microsatellite profile of 10 prawn species using primer pair, SUGbp8-109d.	144
6.3	Microsatellite profile of 10 prawn species using primer pair, SUGbp8-17a and SUGbp8-106c.	144
6.4	Microsatellite motifs amplified by SUGbp3-5a and SUGbp8-107b in *Macrobrachium lanchestri* and *Fenneropenaeus merguiensis*, respectively.	147

LIST OF ABBREVIATIONS

AFLP	amplified fragment length polymorphism
AMOVA	analysis of molecular variance
BLAST	basic local alignment search tool
bp	base pair
dATP	2'-deoxyadenosine 5'-triphosphate
DALP	direct amplification of length polymorphism
dCTP	2'-deoxycytidine 5'-triphosphate
dGTP	2'-deoxyguanosine 5'-triphosphate
D_N	genetic distance
dNTP	deoxyribonucleotide
dTTP	2'-deoxythymidine 5'-triphosphate
EST	expressed sequence tag
FAO	Food and Agriculture Organisation
HWE	Hardy-Weinberg equilibrium
*lac*Z	*lac* operon that encodes for β-galactosidase
LB	Luria-Bertani
LD	linkage disequilibrium
LP-RAPD	long primer random amplified polymorphic DNA
$MgCl_2$	magnesium chloride
NCBI	National Center for Biotechnology Information
PCR	polymerase chain reaction
PIC	polymorphism information content
QTL	quantitative trait loci
RAM	random amplified microsatellite
RAPD	random amplified polymorphic DNA
RFLP	restriction fragment length polymorphism
SNP	Single nucleotide polymorphism
SOC medium	super optimized culture medium
T_A	annealing temperature
TBE	tris borate ethylenediaminetetraacetic acid
T_M	melting temperature
UPGMA	unweighted pair group method with arithmetic averaging
VNTR	variable number of tandem repeats
X-gal	5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside

CHAPTER 1

INTRODUCTION

Freshwater prawn, *Macrobrachium rosenbergii*, or locally known as udang galah is hardy and fast growing, being able to grow in freshwater and low brackish water conditions. This species possesses many biological advantages for commercial culture including attaining maturation in capacity, a relatively large size, and rapid growth rate.

In the 1960's, fishing of giant freshwater prawn was lucrative and there was adequate supply in the country where they exist but at present, increased exploitation and better means of catching has greatly reduced natural stocks and eventually leads to extinction and loss of genetic diversity. Although *M. rosenbergii* has been considered commercially important, biological and genetic information essential to the sustainable management of the resource, such as knowledge of population structure, is lacking. This lack of information has lead to the loss of genetic variability or diversity in the prawn. Comparisons of genetic diversity levels within wild and cultured populations will enable understanding of the effects that intensive culture and small founder populations may have played on levels of genetic diversity in this species.

Assessing genetic variability within and among *M. rosenbergii* is considered important genetic information as it has direct benefit for conserving wild stocks, greater potential for improvement and serves as invaluable resource for different selection criteria, especially when planning breeding or crossbreeding programs. Efficient sampling and utilization of resources would facilitate the determination of genetic variability within and between accessions or varieties. However, the estimation of genetic variation is often limited by the availability of polymorphic genetic markers.

The development of molecular genetic markers has been an important approach toward studying population genetics and dynamics of economically important species. Dominant and codominant markers which have been developed in recent years to characterize population structure and genetic diversity include Random Amplified Polymorphic DNA (RAPD), Random Amplified Microsatellite (RAM) and Microsatellites. In spite of the many types of markers available, the most efficient and effective marker system is microsatellite with its ubiquitous in prokaryotes and eukaryotes genome, and high degree of polymorphism.

1.1 Objectives

The objectives of this study consisted of:

1. Identification and isolation of microsatellite loci in *M. rosenbergii.*
2. Characterization of the microsatellite markers.
3. Assessing the level of genetic variability in *M. rosenbergii* wild stocks.
4. Screening for the cross-amplification of microsatellite primers in nine closely related species.

1.1 Objectives

The objectives of this study consisted of:

1. Identification and isolation of microsatellite loci in *M. rosenbergii*.
2. Characterization of the microsatellite markers.
3. Assessing the level of genetic variability in *M. rosenbergii* wild stocks.
4. Screening for the cross-amplification of microsatellite primers in nine closely related species.