MINING AND VALIDATION OF EST-MICROSATELLITES IN FRESHWATER PRAWNS, *Macrobrachium rosenbergii*

SHAIRAH ABDUL RAZAK

DISSERTATION SUBMITTED IN FULFILMENT OF THE REQUIREMENT FOR THE DEGREE OF MASTER OF BIOTECHNOLOGY

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

2011

ABSTRACT

The current study conducted in purpose to illustrate the utility of EST-derived SSR in characterizing wild populations of *M. rosenbergii* in Malaysia's rivers. A novel set of EST-SSR generated from RNA *of M. rosenbergii* were validated in a full panel of 120 samples from four wild populations through Polymerase Chain Reaction (PCR). Seven EST-SSR loci were identified, characterized, and evaluated on 30 individuals each from the populations namely Negeri Sembilan, Kedah, Sarawak, and Terengganu. The average polymorphic informative content value (PIC) for these seven primers was found to be 0.6208 indicating considerable degree of polymorphism with number of alleles detected ranged from 4 to 10. The observed heterozygosity value count during multi-population analyses ranged from 0.5333 to 0.8333, whilst the expected ranged from 0.6288 to 0.7009. There was no linkage disequilibrium (LD) observed between all pairs of EST-SSRs loci. All loci, except for EST MR8 conformed to the Hardy-Weinberg equilibrium (HWE) suggesting factors violating the neutral expectation such as selection might have caused the locus to deviate from equilibrium. The F_{IS} index demonstrated no indication of inbreeding among individuals of each population. There was evidence that all populations assessed in this study are drawn from a single, large panmictic population; no genetic heterogeneity observed in population structure analysis, estimate of fixation index value in pairwise comparisons among the four localities revealed very low magnitude of differentiation (F_{ST} ranged between 0.00888 to the highest of 0.10644).

The results indicated that these polymorphic EST-SSR derived from *M. rosenbergii* would be useful for population genetic structure analysis and genetic diversity assessment in prawn populations as part of management policies of natural resources to ensure sustainability of wild broodstock for future development of prawn culture industries.

ABSTRAK

Kajian ini telah dijalankan dengan tujuan untuk memperlihatkan kegunaan EST-terbitan dari SSR dalam pencirian populasi liar *M. rosenbergii* di sungai-sungai di Malaysia. Set penanda EST-SSR baru telah dihasilkan dari RNA *M. rosenbergii*, dan5 penanda ini seterusnya ditentusahkan dengan 120 sampel daripada empat populasi semulajadi melalui Tindakbalas Berantai Polimerase (PCR). Tujuh lokus EST-SSR loci telah dikenalpasti, dikelaskan, serta diuji ke atas setiap 30 individu daripada populasi Negeri Sembilan, Kedah, Sarawak, dan Terengganu. Purata nilai Kandungan Polimorfik Informatif (PIC) bagi kesemua tujuh primer ialah 0.6208, menunjukkan tahap polimorfisme yang tinggi, dengan bilangan alel yang dikesan berjulat antara 4 hingga 10. Nilai heterozigositi yang diperoleh dalam analisis multi-populasi berjulat antara 0.5333 hingga 0.8333, sementara nilai dijangka berjulat antara 0.6288 hingga 0.7009. Tiada ketidakseimbangan rangkaian (LD) ditemui antara kesemua lokus EST-SSR. Semua lokus, kecuali EST MR8 memenuhi keseimbangan Hardy-Weinberg (HWE), dan ketidakseimbangan ini menunjukkan faktor melanggar jangkaan neutral seperti pemilihan mungkin mengakibatkan lokus tersebut melencong dari keseimbangan. Indeks F_{IS} menunjukkan tiada tanda berlakunya kacuk dalam antara individu bagi setiap populasi. Kesemua populasi yang telah dinilai dalam kajian ini berkemungkinan berasal daripada satu populasi besar yang membiak secara rawak; tiada keheterogenan dikesan dalam analisis struktur populasi, dan anggaran nilai indek penetapan dalam bandingan-secara-pasangan antara keempat-empat lokasi mencatatkan magnitud perbezaan yang sangat rendah (F_{ST} berjulat antara 0.00888 sehingga 0.10644 bagi yang tertinggi).

Hasil kajian ini menunjukkan EST-SSR yang polimorfik yang didapati dari *M. rosenbergii* ini berguna untuk analisis struktur genetik serta penilaian kepelbagaian genetik bagi populasi udang, sebagai sebahagian dari dasar pengurusan sumber semulajadi bagi memastikan kelestarian benih udang liar untuk pembangunan industri kultur udang pada masa hadapan.

ACKNOWLEDGEMENTS

I thank Allah for giving me the strength to accomplish this task and being my constant companion to went through these times.

Special gratitude to my supervisor, Dr Subha Bhassu, for giving me a chance to fly further & the opportunity to join her lab community also entrusted me to finish part of her project. I wish you the best in times to come.

To Ellie, you are my savior. To Izzah, you are like my guardian angel and to Yasmin, special thanks for being the best Mentor ever, and Jothi my comrade, we finally did it! Not to be forgotten, the whole lab members for being there to support me all this while.

I am also grateful and thankful to my very supportive husband, parents, family, classmates and colleagues for their endless support and their company through thick & thin during my journey. There's more to come in the future, but there will be none if it's not for what I learned here today.

Thank You All

Shairah Abdul Razak

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

A_e	effective number of alleles per locus
A_t	total number of alleles
BLAST	Basic local alignment search tool
bp(s)	Basepair(s)
cDNA	complementary DNA
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
dGTP	deoxyguanosine triphosphate
dNTP	deoxyribonucleotide triphosphate
dTTP	deoxythymidine triphosphate
ddH2O	Double distilled water
DNA	Deoxyribonucleic acid
DMSO	Dimethyl sulfoxide
EST	Expressed Sequence Tags
F_{IS}	Inbreeding Coefficient
F_{ST}	Fixation index
GFP	Giant Freshwater Prawn
H_E	Expected heterozygosity
H_O	Observed heterozygosity
HWE	Hardy Weinberg equilibrium
LD	Linkage disequilibrium
MAS	Marker-assisted selection
mRNA	Messenger RNA
N	effective sample size
NCBI	National Center for Biotechnology Information
OD	Optical density
PCR	Polymerase Chain Reaction

PIC	**Polymorphic information content**
QTL	**Quantitative trait Locus**
R_s	**allelic richness**
SSR	**Simple sequence repeat**
Std. Dev.	**Standard deviation**
T_A	**Annealing temperature**
UPGMA	**Unweighed Pair-Group Method of Arithmetic Averages**

LIST OF APPENDICES

APPENDICES